Pak. J. Statist.
2020 Vol. 36(4), 363-378

MODIFIED DOMINANCE ANALYSIS

El-Houssainy Abd El-Bar Rady[1], **Ahmed Amin ELSheikh**[1]
and **Mohamed Rashed Ezzeldin**[2§]

[1] Department of Applied Statistics and Econometric,
Faculty of Graduate Studies for Statistical Research,
Cairo University, Egypt.

[2] Department of Statistics, Deanship of Educational Services,
Qassim University, Saudi Arabia.

[§] Corresponding author Email: m.ezzeldin@qu.edu.sa

ABSTRACT

Relative importance analysis is a very useful supplement to regression analysis. The purpose of determining predictor importance is not model selection but rather uncovering the individual contributions of predictors relative to each other within a selected model.

Dominance analysis considers one of the most powerful relative important techniques which provides a framework for assessing the relative importance of predictors in univariate and multivariate multiple regression models.

This method is based on pairwise comparisons of the contribution of predictors in all possible subset models. One of the great weaknesses of Dominance analysis (DA) that is, it will become more computationally difficult due to the exponentially growing number of sub-models involved that is due to the fact for predictors, there are $(2^P - 1)$ sub-models.

Here we try to extend the current studies practice by developing a technique to overcomes this disadvantage. We determined that the new technique gives the same result for classical general dominance and furthermore, we get a huge reduction in the sub-models, for the result to be robust a simulation study has been conducted.

KEYWORDS

Dominance Analysis, Multiple regression, Relative importance, Statistical significance, Simulation, Predictor Importance.

1. INTRODUCTION

As mentioned in many articles, regression analysis is arguably one of the most essential tools used by researchers and practitioners in figuring out the relationship among a response variable and its' predictor variables. Frequently times but there may be a preference to further examine the model created to decide which variable or variables are considered the most vital or make a contribution the most variance explanation of the overall model. Evaluation of this type is known as Relative importance analysis and serves as an essential supplement to traditional regression models. Relative importance analysis aims to partition defined variance among more than one predictors with a purpose to better understand the role played by each predictor in a regression model.

The concept of predictor" dominance" proposed by Budescu (1993) is a new way to compare predictors in a multiple regression context. Dominance Analysis (Budescu 1993) approaches the problem of relative importance by examining the change in R^2 resulting from adding a predictor to all possible subset regression models. One predictor is said to completely dominate another if its additional contribution in R^2 is greater than that of the other predictor under each of the subset models; if the average additional contribution in R^2 within each model size is greater for one predictor than the other, then that predictor is said to conditionally dominate the other. By averaging the conditional dominance weights, one obtains a predictor's general dominance weight. The three levels of dominance are related in a hierarchical fashion: complete dominance implies conditional dominance, which, in turn, implies general dominance. However, for p > 3, the converse does not necessarily hold; that is, general dominance does not necessarily imply conditional dominance, and conditional dominance does not necessarily imply complete dominance. The general dominance weight can be interpreted as the partial effect of the variable because it addresses a variable's contribution in combination with the other predictors while overcoming the problems associated with correlated predictors. It represents the proportion of variance of the variable explained by all other variables in the model (Azen and Budescu 2003) and is very similar to other measures of relative importance.

However, the general dominance weight, because it is based on the averaging of conditional dominance weights, introduces certain degrees of bias. Note that the additional contribution of R^2 of a predictor is more substantial for small models; as the number of predictors increases, the additional contribution of any extra predictors will get lesser. Using an equal-weighted average of conditional dominance weights across different model sizes puts more credence on the small model. Hence, the order of dominance tends to be influenced more by the conditional dominance weights of the models with a small number of predictors.

2. METHODOLOGY

Dominance analysis is accomplished by performing pairwise comparisons of each predicting variable to a competing predicting variable on all of the $k = 0$ to $(p - 1)$ subset levels. The two 'competing' predictors are compared in the context of all $2^{(\mathrm{p}-2)}$ models that contain some subset of the other predictors. The table below found in Budescu (1993) outlines how to calculate the additional contribution for each predicting variable. To determine the dominance hierarchy of the model in question several steps must be completed. To begin, the researcher must determine what value or statistic they will use to measure the 'model adequacy'.

Upon deciding on an adequacy measure the next step is to determine all the possible subset bases models containing $k = 0$ to $k = p - 1$ variables. Once these subset models have been determined, the adequacy measure is calculated and used as a base case scenario. Following this, each of the two predicting variables that are pitted against each other is added to the base case scenario. The adequacy value of the model is then calculated for the model with the recently added predicting variables, X_i and X_j, assuming that X_i and X_j are not included in the base case scenario. This process is referred to as computing the additional contribution for each predicting variables. The two additional contribution values are then compared to each other. If the additional contribution is greater after X_i than it is for X_j, then X_i is said to dominate X_j for that particular subset model.

This procedure is carried out for each of the determined subset models. Once all the predicting variables have been accounted for, the next step is to determine the average additional contribution for each predicting variable for each of the $k = 0$ to $k = p-1$ levels. Additionally, the overall contribution for each of the predicting variables is calculated by calculating the simple average of each different 3 levels. See Table 1 for an example of a theoretical model of $p = 4$ predicting variables, using the squared partial correlation as the measure of model adequacy.

In the final step of the process, the different classifications of dominance are determined. The diagrams below (Figure 1) outline the process of determining which category of dominance is prevalent.

- **Complete**
  For complete dominance to be achieved, one of the two competing predicting variables must have a greater additional contribution to the subset model than the other.
- **Conditional**
  For a predicting variable to conditionally dominate another, it must exhibit a greater additional contribution for each of the $k = 0$ to $k = p-1$ level averages.
- **General**
  For a variable to generally dominate another, the average of all the different k level additional contributions, in this case, $k = 0$ to 3 must be greater for that variable than the variable it is competing against.

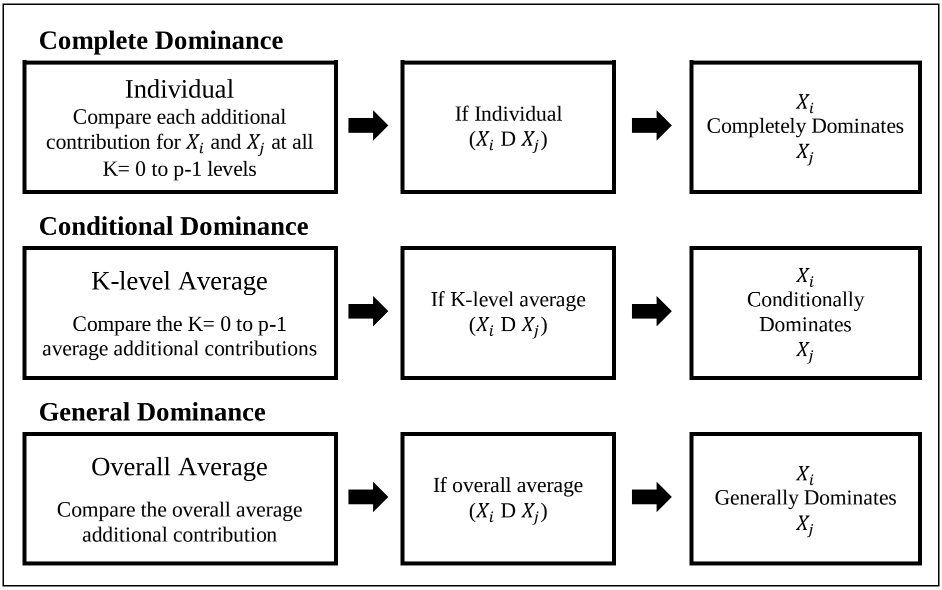

Figure 1: The Different Classifications of Dominance Hierarchies

It can be shown that if one variable completely dominates another, then it will also conditionally and generally dominate that variable as well. That is to say that a complete dominance hierarchy implies the lesser forms of dominance also. Similarly, if a dominance hierarchy holds in the conditional form, it will also hold at the general level. As shown in the example though, the reverse is not necessarily true. It cannot be asserted that if a hierarchy is established at the general or conditional form, that it will hold in the conditional and complete form, or complete form, respectively.

3. DOMINANCE ANALYSIS STRENGTHS AND WEAKNESSES

One of this technique's strengths is that Dominance Analysis can come across and identify suppressor variables due to the fact the Dominance Analysis table will display a negative dominance index, rather than it being masked as in other techniques. The extra power of this method is the capability to carry out a constrained Dominance Analysis. This is, one predictor variable may be constrained as being essential to a model, to locate which predictors are the maximum essential and best supplement the variable required within the model. This could be an advantage to a practitioner remodeling a selection system that has been given direction that a selected tool(s) need to stay within the system. A final strength of this technique is that importance estimates may be divided by using the R^2 to calculate the share of explainable variance in the overall R^2 via a given predictor variable.

A weakness of this approach is that as the number of analyzed predictors will increase, Dominance Analysis turns into extra computationally hard because of the exponentially growing number of sub-models involved. That is due to the fact for predictors, there are $(2^P - 1)$ sub-models. As an instance, with 3 predictors a Dominance Analysis summary table will have seven possible models (i.e. (1) X_1, (2) X_2, (3) X_3, (4) X_1, X_2, (5) X_2, X_3, (6) X_1, X_3, (7) X_1, X_2, X_3). Likewise, 10 predictors will yield 1023 possible sub-models.

4. MODIFIED DOMINANCE ANALYSIS

As mentioned before one of the weaknesses of the Dominance Analysis approach is that, as the number of analyzed predictors will increase, Dominance Analysis turns into extra computationally hard because of the exponentially growing number of sub-models involved.

To overcome the problem, we introduced a new technique, which is, involved in the following main stages:

First: All predictor variables should be examined for bivariate correlations. We calculate the predictor's direct effect,

Second: When we obtain the predictors' direct effect from the first step we arrange predictors in an ascending or descending order according to their direct effect.

Third: After arranging the predictors, we split the predictors into two (or more) groups concerning the given order. This step is aiming to reduce the number of sub-models obtained which is as mentioned earlier increase exponentially as the number of predictors increase.

Forth: We run a separate Dominance Analysis for each group, as shown in Table 1.

Fifth: We get the dominance weight for the whole variables.

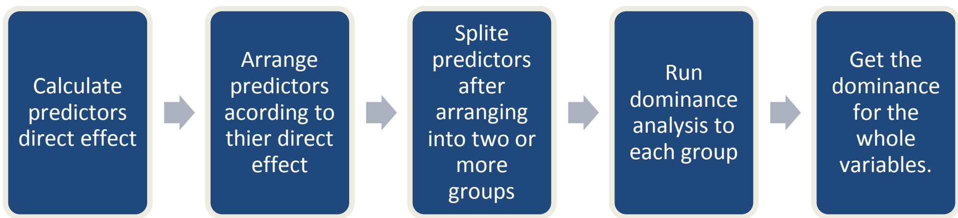

Figure 2: Steps of the Modified Dominance Analysis

Table 1
Dominance Analysis Formulas for Three Predictors Case

Subset Model (X)	$\rho^2{}_{Y.X}$	**Additional Contribution**		
		X_1	X_2	X_3
Null and k=0 average (CD)		$\rho^2{}_{Y.X_1}$	$\rho^2{}_{Y.X_2}$	$\rho^2{}_{Y.X_3}$
X_1	$\rho^2{}_{Y.X_1}$	-	$\rho^2{}_{Y.X_1X_2} - \rho^2{}_{Y.X_1}$	$\rho^2{}_{Y.X_1X_3} - \rho^2{}_{Y.X_1}$
X_2	$\rho^2{}_{Y.X_2}$	$\rho^2{}_{Y.X_1X_2} - \rho^2{}_{Y.X_2}$	-	$\rho^2{}_{Y.X_2X_3} - \rho^2{}_{Y.X_2}$
X_3	$\rho^2{}_{Y.X_3}$	$\rho^2{}_{Y.X_1X_3} - \rho^2{}_{Y.X_3}$	$\rho^2{}_{Y.X_2X_3} - \rho^2{}_{Y.X_3}$	-
$k = 1$ average (CD)		$\frac{(\rho^2{}_{Y.X_1X_2} - \rho^2{}_{Y.X_2}) + (\rho^2{}_{Y.X_1X_3} - \rho^2{}_{Y.X_3})}{2}$	$\frac{(\rho^2{}_{Y.X_1X_2} - \rho^2{}_{Y.X_1}) + (\rho^2{}_{Y.X_2X_3} - \rho^2{}_{Y.X_3})}{2}$	$\frac{(\rho^2{}_{Y.X_1X_3} - \rho^2{}_{Y.X_1}) + (\rho^2{}_{Y.X_2X_3} - \rho^2{}_{Y.X_2})}{2}$
X_1X_2	$\rho^2{}_{Y.X_1X_2}$	-	-	$\rho^2{}_{Y.X_1X_2X_3} - \rho^2{}_{Y.X_1X_2}$
X_1X_3	$\rho^2{}_{Y.X_1X_3}$	-	$\rho^2{}_{Y.X_1X_2X_3} - \rho^2{}_{Y.X_1X_3}$	-
X_2X_3	$\rho^2{}_{Y.X_2X_3}$	$\rho^2{}_{Y.X_1X_2X_3} - \rho^2{}_{Y.X_2X_3}$	-	-
$k = 2$ average (CD)		$\rho^2{}_{Y.X_1X_2X_3} - \rho^2{}_{Y.X_2X_3}$	$\rho^2{}_{Y.X_1X_2X_3} - \rho^2{}_{Y.X_1X_3}$	$\rho^2{}_{Y.X_1X_2X_3} - \rho^2{}_{Y.X_1X_2}$
$X_1X_2X_3$	$\rho^2{}_{Y.X_1X_2X_3}$	-	-	-
Overall Average (GD)		$\frac{\sum_{K=0}^{2} CDX_1}{3}$	$\frac{\sum_{K=0}^{2} CDX_2}{3}$	$\frac{\sum_{K=0}^{2} CDX_3}{3}$

Note: (CD) Conditional Dominance(GD) Complete Dominance (-) cell are not applicable

5. SIMULATION STUDY

To make the presentation more concrete, we conduct a simulation study to illustrate the robustness of the modified Dominance Analysis procedure. We used a parent sample with a defined correlation matrix (Table 2), the sample followed the normal distribution and its size $n = 1000$. The pattern for correlations with the criterion is increasingly correlated with two equally high correlation coefficients.

Table 2
Population Correlation Matrix Used in the Simulation Study

	ρ_{XiY}	ρ_{XiXj}					
	Y	$X1$	$X2$	$X3$	$X4$	$X5$	$X6$
Y	1						
$X1$	0.11	1					
$X2$	0.13	0.10	1				
$X3$	0.22	0.16	0.08	1			
$X4$	0.25	0.17	0.11	0.11	1		
$X5$	0.32	0.28	0.17	0.18	0.11	1	
$X6$	0.32	0.30	0.20	0.15	0.11	0.10	1

Note: ρ_{XiY} represents the correlation coefficient between predictor X_i and criterion Y; ρ_{XiXj} represents the correlation coefficient between predictor X_i and another predictor X_j;

First, we will calculate the classical Dominance Analysis of the parent sample generated. The result of the classical dominance analysis shown in Table 3. The column labeled ${\rho^2}_{Y.X}$ represents the variance in Y explained by the model appearing in the corresponding row. Columns labeled X_i contain the additional contributions to the explained variance gained by adding the column variable (X_i) to the row model. Blank cells indicate that data are not applicable.

Table 3
Dominance Analysis for the Hypothetical Example with Six Predictors

Classical								
Subset Model (X)	Level	$\rho^2 Y.X$	**Additional Contribution**					
			X_1	X_2	X_3	X_4	X_5	X_6
	0	0	0.012	0.018	0.05	0.063	0.107	0.107
X_1	1	0.012		0.015	0.043	0.055	0.096	0.095
X_2	1	0.018	0.009		0.046	0.056	0.096	0.093
X_3	1	0.05	0.005	0.014		0.052	0.085	0.088
X_4	1	0.063	0.004	0.011	0.039		0.091	0.091
X_5	1	0.107	0	0.006	0.028	0.046		0.088
X_6	1	0.107	0	0.005	0.031	0.047	0.089	
Average level 1	1		0.004	0.01	0.037	0.051	0.091	0.091
X_1X_2	2	0.027			0.041	0.051	0.087	0.084
X_1X_3	2	0.055		0.012		0.048	0.079	0.082
X_1X_4	2	0.067		0.01	0.036		0.086	0.087
X_1X_5	2	0.107		0.006	0.027	0.046		0.092
X_1X_6	2	0.107		0.005	0.031	0.047	0.093	
X_2X_3	2	0.063	0.004			0.047	0.077	0.078
X_2X_4	2	0.074	0.003		0.036		0.083	0.082
X_2X_5	2	0.114	0		0.027	0.043		0.082
X_2X_6	2	0.111	0		0.03	0.045	0.085	
X_3X_4	2	0.102	0.001	0.009			0.073	0.076
X_3X_5	2	0.135	0	0.005		0.04		0.077
X_3X_6	2	0.138	0	0.004		0.04	0.074	
X_4X_5	2	0.153	0	0.004	0.022			0.077
X_4X_6	2	0.153	0	0.002	0.025		0.077	
X_5X_6	2	0.195	0.005	0.001	0.016	0.035		
Average level 2	2		0.001	0.006	0.029	0.044	0.081	0.082
$X_1X_2X_3$	3	0.067				0.044	0.073	0.074
$X_1X_2X_4$	3	0.077			0.034		0.08	0.079
$X_1X_2X_5$	3	0.114			0.026	0.043		0.087
X_1X_2X6	3	0.111			0.03	0.045	0.089	
$X_1\ X_3X_4$	3	0.103		0.008			0.073	0.076
$X_1X_3X_5$	3	0.135		0.005		0.041		0.083
$X_1X_3X_6$	3	0.138		0.004		0.042	0.08	
$X_1X_4X_5$	3	0.153		0.004	0.022			0.085
$X_1X_4X_6$	3	0.154		0.002	0.026		0.085	
$X_1X_5X_6$	3	0.2		0.001	0.018	0.039		
$X_2X_3X_4$	3	0.11	0.001				0.068	0.069
$X_2X_3X_5$	3	0.14	0			0.038		0.072

Classical								
Subset Model (X)	Level	$\rho^2 Y.X$	**Additional Contribution**					
			X_1	X_2	X_3	X_4	X_5	X_6
$X_2X_3X_6$	3	0.141	0			0.038	0.071	
$X_2X_4X_5$	3	0.157	0		0.021			0.073
X_2X_4X6	3	0.156	0		0.024		0.074	
$X_2X_5X_6$	3	0.196	0.005		0.016	0.034		
$X_3X_4X_5$	3	0.175	0.001	0.003				0.068
$X_3X_4X_6$	3	0.178	0.001	0.002			0.065	
$X_3X_5X_6$	3	0.211	0.006	0		0.031		
$X_4X_5X_6$	3	0.23	0.008	0	0.013			
Average level 3	3		0.002	0.003	0.023	0.04	0.076	0.076
$X_1X_2X_3X_4$	4	0.111					0.068	0.07
$X_1X_2X_3X_5$	4	0.14				0.039		0.078
$X_1X_2X_3X_6$	4	0.141				0.04	0.077	
$X_1X_2X_4X_5$	4	0.157			0.022			0.081
$X_1X_2X_4X_6$	4	0.156			0.025		0.082	
$X_1X_2X_5X_6$	4	0.2			0.018	0.038		
$X_1X_3X_4X_5$	4	0.176		0.003				0.077
$X_1X_3\ X_4X_6$	4	0.179		0.002			0.074	
$X_1\ X_3\ X_5\ X_6$	4	0.218		0		0.035		
$X_1\ X_4\ X_5X_6$	4	0.238		0	0.014			
$X_2\ X_3\ X_4\ X_5$	4	0.178	0.001					0.065
$X_2\ X_3\ X_4X_6$	4	0.18	0.002				0.063	
$X_2\ X_3\ X_5X_6$	4	0.212	0.006			0.031		
$X_2\ X_4\ X_5X_6$	4	0.23	0.008		0.013			
$X_3\ X_4\ X_5X_6$	4	0.243	0.01	0				
Average level 4	4		0.005	0.001	0.018	0.037	0.073	0.074
$X_1\ X_2\ X_3\ X_4X_5$	5	0.179						0.074
$X_1\ X_2\ X_3\ X_4X_6$	5	0.181					0.072	
$X_1\ X_2\ X_3\ X_5X_6$	5	0.218				0.035		
$X_1\ X_2\ X_4\ X_5X_6$	5	0.238			0.014			
$X_1\ X_3\ X_4\ X_5X_6$	5	0.253		0				
$X_2\ X_3\ X_4\ X_5X_6$	5	0.243	0.01					
Average level 5	5		0.01	0	0.014	0.035	0.072	0.074
$X_1X_2X_3X_4X_5X_6$	6	0.253						
Overall Average		0.144	0.003	0.005	0.027	0.043	0.080	0.081

Note: The column labeled ${\rho^2}_{Y.X}$ represents the variance in Y explained by the model appearing in the corresponding row. Columns labeled X_i contain the additional contributions to the explained variance gained by adding the column variable (X_i) to the row model. Blank cells indicate that data are not applicable.

To begin, the researcher must determine all the possible subset bases models containing $k = 0$ to $k = p - 1$ variable. Once these subset models have been determined, the adequacy measure is calculated and used as a base case scenario. Following this, each of the two predicting variables that are pitted against each other is added to the base case scenario. The adequacy value of the model is then calculated for the model with the recently added predicting variables, X_i and X_j, assuming that X_i and X_j are not included in the base case scenario. This process is referred to as computing the additional contribution for each predicting variables. The two additional contribution values are then compared to each other. If the additional contribution is greater after X_i then it is for X_j, then X_i is said to dominate X_j for that particular subset model. This procedure is carried out for each of the determined subset models. Once all the predicting variables have been accounted for, the next step is to determine the average additional contribution for each predicting variable for each of the $k = 0$ to $k = p - 1$ level. Additionally, the overall contribution for each of the predicting variables is calculated by calculating the simple average of each different 5 levels. See Table 3 for our example of $p = 6$ predicting variables.

For a variable to generally dominate another, the average of all the different k level additional contributions, in this case, $k = 0$ to 5 must be greater for that variable than the variable it is competing against. The overall averages for this example can be found in Table 3 in the last row of the table. As shown in Table 3 the additional average contributions for predictors form X_6 to X_1 in are 0.084, 0.083, 0.045, 0.029, 0.006 and 0.006 respectively. These values give rise to the general dominance hierarchy of $X_6 > X_5 > X_4 > X_3 > X_2 > X_1$. The detailed structure is given in the following table.

Table 4
Complete, Conditional and General dominance

	Complete	Conditional	General
X_1			
X_2			X_1
X_3	X_1X_2	X_1X_2	X_1X_2
X_4	$X_1X_2X_3$	$X_1X_2X_3$	$X_1X_2X_3$
X_5	$X_1X_2X_3X_4$	$X_1X_2X_3X_4$	$X_1X_2X_3X_4$
X_6	$X_1X_2X_3X_4$	$X_1X_2X_3X_4$	$X_1X_2X_3X_4X_5$

As we can see from the table that X_6 generally dominates $X_1X_2X_3X_4X_5$, X_5 generally dominates $X_1X_2X_3X_4$, X_4 generally dominates $X_1X_2X_3$, X_3 generally dominates X_1X_2, and X_2 generally dominates X_1.

5.1 Modified Dominance Analysis

Now we will use the same sample, split the predictors into two groups each group with three variables according to the direct effect, and run a Dominance Analysis for each group.

As the univariate correlation between predictor and criterion are predefined, so after arranging the predictors in ascending order $X_1 X_2 X_3 X_4 X_5, X_6$. We will split the predictors into two groups in such that the first group contains predictors $X_1 X_2 X_3$, and the second group contains predictors $X_4 X_5, X_6$. We run a separate dominance analysis for each group. The results appeared below in Table 5.

Table 5
Dominance Analysis for the First Group with Three Predictors

		Additional Contribution		
Subset Model (X)	$\boldsymbol{\rho^2} Y.X$	$\boldsymbol{X_1}$	$\boldsymbol{X_2}$	$\boldsymbol{X_3}$
Null and k=0		0.012	0.018	0.05
$\mathbf{X_1}$	0.012		0.015	0.043
$\mathbf{X_2}$	0.018	0.009		0.046
$\mathbf{X_3}$	0.05	0.005	0.014	
$k = 1$		0.007	0.014	0.045
$\mathbf{X_1 X_2}$	0.027			0.041
$\mathbf{X_1 X_3}$	0.055		0.012	
$\mathbf{X_3\ X_3}$	0.063	0.004		
$k = 2$		0.004	0.012	0.041
$\mathbf{X_1 X_2\ X_3}$	0.067	-	-	-
Overall Average		0.008	0.015	0.045

Note: The column labeled ${\rho^2}_{Y.X}$ represents the variance in Y explained by the model appearing in the corresponding row. Columns labeled X_i contain the additional contributions to the explained variance gained by adding the column variable (X_i) to the row model. Blank cells indicate that data are not applicable.

We perform Dominance Analysis for the first group which now contains the three variables X_1, X_2 and X_3. Table 5 shows the average additional contributions for X_1, X_2 and X_3 are 0.008, 0.015 and 0.045 respectively. These values give rise to the general dominance hierarchy of $X_3 > X_2 > X_1$. Given this structure, we can say that X_3 generally dominates X_2, X_2 generally dominates X_1. Because of the transitivity property of dominance, we can say that X_3 will dominate X_2 and X_1. X_2 will dominate X_1. X_1, however, will not generally dominate any of the other variables. This hierarchy is the same as classical dominance analysis.

The following step is to perform Dominance Analysis for the second group which now contains the three variables $X_4, X_5,$ and X_6. And the result of Dominance Analysis appears in Table 6.

Table 6
Dominance Analysis for the Second Group with Three Predictors

		Additional Contribution		
Subset Model (X)	$\rho^2 Y.X$	X_4	X_5	X_6
Null and k=0		0.063	0.107	0.107
$\mathbf{X_4}$	0.063		0.091	0.091
$\mathbf{X_5}$	0.107	0.046		0.088
$\mathbf{X_6}$	0.107	0.047	0.089	
$k = 1$		0.046	0.09	0.089
$\mathbf{X_4X_5}$	0.153			0.077
$\mathbf{X_4X_6}$	0.153		0.077	
$\mathbf{X_5X_6}$	0.195	0.035		
$k = 2$		0.035	0.077	0.077
$\mathbf{X_4X_5X_6}$	0.23	-	-	-
Overall Average		0.048	0.091	0.091

Note: The column labeled ${\rho^2}_{Y.X}$ represents the variance in Y explained by the model appearing in the corresponding row. Columns labeled X_i contain the additional contributions to the explained variance gained by adding the column variable (X_i) to the row model. Blank cells indicate that data are not applicable.

As shown in Table 6 the additional contributions for $X_4, X_5,$ and X_6 are 0.048, 0.091 and 0.091. These values give rise to the general dominance hierarchy of $X_6 > X_5 > X_4$. Given this structure, we can say that X_6 generally dominates X_5, X_5 generally dominates X_4. Because of the transitivity property of dominance, we can say that X_6 will dominate X_5 and X_4. X_5 will dominate X_4.

If we combine the results for the previous two tables and as we will get a general dominance hierarchy of $X_6 > X_5 > X_4 > X_3 > X_2 > X_1$, which is identical to the hierarchy of the full model.

5.2 Bootstrapping Procedure

To evaluate the robustness of our results, we can use bootstrap analysis (Azen and Budescu, 2006). We applied a bootstrap analysis method with R^2 as a fit index and for precise results, we run 1000 replications. We get the following result for the two groups:

Table 7
The First Group Bootstrapping Result

Dominance	i	k	D_{ij}	mD_{ij}	$SE.D_{ij}$	P_{ij}	P_{ji}	P_{noij}	Rep
Complete	$X1$	$X2$	0	0.216	0.3748	0.158	0.726	0.116	0.726
Complete	$X1$	$X3$	0	0.003	0.0499	0.002	0.996	0.002	0.996
Complete	$X2$	$X3$	0	0.021	0.1326	0.015	0.973	0.012	0.973
Conditional	$X1$	$X2$	0	0.216	0.3748	0.158	0.726	0.116	0.726
Conditional	$X1$	$X3$	0	0.003	0.0499	0.002	0.996	0.002	0.996
Conditional	X2	$X3$	0	0.021	0.1326	0.015	0.973	0.012	0.973
General	X1	$X2$	0	0.215	0.411	0.215	0.785	0	0.785
General	X1	$X3$	0	0.003	0.0547	0.003	0.997	0	0.997
General	X2	X3	0	0.018	0.133	0.018	0.982	0	0.982

Note: Fit index: R^2

Table 7 presents the results for the bootstrap analysis for the first group. D_{ij} shows the original result, and mD_{ij}, the mean for D_{ij} on bootstrap samples and SE. D_{ij} its standard error.

P_{ij} is the proportion of bootstrap samples where i dominates j, P_{ji} is the proportion of bootstrap samples where j dominates i and P_{noij} is the proportion of samples where no dominance can be asserted. *Rep* is the proportion of samples where original dominance is replicated. We used reproducibility to show the stability and robustness of dominance patterns. (Azen & Budescu, 2003; Azen & Traxel, 2009) suggested that reproducibility values under 70% might indicate that the relationship observed in the sample may not hold in the population.

We can see that the value of complete, conditional and general dominance for $X_1X_2X_3s$ are strongly robust overall variables, as *Rep* is grater than 0.7.

Table 8
The Second Group Bootstrapping Result

Dominance	i	k	D_{ij}	mD_{ij}	$SE.D_{ij}$	P_{ij}	P_{ji}	P_{noij}	Rep
Complete	$X4$	$X5$	0	0.013	0.094	0.005	0.979	0.016	0.979
Complete	$X4$	$X6$	0	0.0345	0.1691	0.025	0.956	0.019	0.956
Complete	$X5$	$X6$	0.5	0.506	0.4819	0.47	0.458	0.072	0.072
Conditional	$X4$	$X5$	0	0.013	0.094	0.005	0.979	0.016	0.979
Conditional	$X4$	X6	0	0.0345	0.1691	0.025	0.956	0.019	0.956
Conditional	$X5$	X6	0.5	0.506	0.4819	0.47	0.458	0.072	0.072
General	$X4$	$X5$	0	0.01	0.0995	0.01	0.99	0	0.99
General	$X4$	$X6$	0	0.035	0.1839	0.035	0.965	0	0.965
General	$X5$	$X6$	1	0.509	0.5002	0.509	0.491	0	0.509

Table 8 presents the results for the bootstrap analysis for the first group. D_{ij} shows the original result, and mD_{ij}, the mean for D_{ij} on bootstrap samples and SE.D_{ij} its standard error.

P_{ij} is the proportion of bootstrap samples where i dominates j, P_{ji} is the proportion of bootstrap samples where j dominates i and P_{noij} is the proportion of samples where no dominance can be asserted. *Rep* is the proportion of samples where original dominance is replicated.

We can see that the value of complete, conditional and general dominance for $X_4X_5X_6$ are strongly robust as *Rep* is grater than 0.7, except X_5X_6 as they are with equal correlation with the creterion, so the dominance cannot be determine.

6. CONCLUSION

Overall, results from this study indicate that dominance analysis can be modified to save time and effort, and also eliminate many redundant data when performing it and with the same efficiency as a tool for researchers wishing to determine the relative contribution of predictors. As shown the great weakness of Dominance Analysis is that as the number of analyzed predictors will increase, it turns into extra computationally hard because of the exponentially growing number of sub-models involved. Therefore, we try to overcome this weakness by the following steps: First: We examine the predictor's direct effect. Second: when we obtain the predictors' direct effect from the first step we arrange predictors in an ascending or descending order according to their direct effect. Third: After arranging the predictors, we split the predictors into two or more groups according to the predictor's direct effect. This step is aiming to reduce the number of sun-models obtained which is as mentioned earlier increase exponentially as the number of predictors increase. Forth: we run a separate Dominance Analysis for each group. Fifth: We get the dominance weight for the whole variables. A simulation study has been conducted, we use the bootstrapping technique and used reproducibility to show the stability and robustness of dominance patterns. The result for the new technique shows an identical result for classical general dominance and we get 14 sub-model instead of 63 sub-model. This is a huge reduction in sub-models, which will lead to saving effort and eliminate many redundant data. With the modified technique for a 10 predictor's model, we will get a 62 sub-model instead of 1024 sub-model. This technique will lead to saving effort and eliminate many redundant data.

REFERENCES

1. Azen, R., Budescu, D.V. and Reiser, B. (2001). Criticality of predictors in multiple regression. *British Journal of Mathematical and Statistical Psychology*, 54, 201-225.
2. Azen, R. and Budescu, D.V. (2003). The Dominance Analysis Approach for Comparing Predictors in Multiple Regression. *Psychological Methods*, 8, 129-148.
3. Azen, R., and Traxel, N.M. (2009). Using dominance Analysis to determine predictor importance in logistic regression. *Journal of Educational and Behavioral Statistics*, 34, 319-347.
4. Barni, D. (2015). The relative importance analysis for the study of the family: Accepting the challenge of correlated predictors. *TPM - Testing Psychometrics Methodology in Applied Psychology*, 22, 235-250.

5. Braun, M.T., Converse, P.D. and Oswald, F.L. (2019). The accuracy of dominance analysis as a metric to assess relative importance: The joint impact of sampling error variance and measurement unreliability. *Journal of Applied Psychology*, 104(4), 593-602.
6. Budescu, D.V. (1993). Dominance analysis: a new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin*, 114(3), 542-551.
7. Bradley, R.A. and Terry, M.E. (1952). Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons. *Biometrika*, 39(3/4), 324-334.
8. Courville, T. and Thompson, B. (2001). Use of structure coefficients in published multiple regression articles: b is not enough. *Educational and Psychological Measurement*, 61, 229-248.
9. Darlington, R.B. (1968). Multiple regression in psychological research and practice. *Psychological Bulletin*, 69, 161-182.
10. Miranda-Villatoro, F., Forni, F. and Sepulchre, R. (2018). Dominance analysis of linear complementarity systems, Proceedings *23rd International Symposium on Mathematical Theory of Networks and Systems* (Hong-Kong).
11. Fabbris, L. (1980). Measures of predictor variable importance in multiple regression: An additional suggestion. *Quality and Quantity*, 4, 787-792.
12. Forni, F. and Sepulchre, R. (2019). Differential Dissipativity Theory for Dominance Analysis. In *IEEE Transactions on Automatic Control*, 64(6), 2340-2351.
13. Johnson, J. (2000). A Heuristic Method for Estimating the Relative Weight of Predictor Variables in Multiple Regression. *Multivariate Behavioral Research*, 35, 1-19.
14. Johnson, J.W. and LeBreton, J.M. (2004). History and use of relative importance indices in organizational research. *Organizational Research Methods*, 7(3), 238-257.
15. Joseph N. Luchman, (2019). DOMME: Stata module to conduct dominance analysis with multiple equation models. *Statistical Software Components S458673*, Boston College Department of Economics.
16. Kurskal, W. and Majors, R. (1989). Concepts of Relative Importance in Recent Scientific Literature. *The American Statistician*, 43(1), 2-6.
17. Lebreton, J.M., Ployhart, R.E. and Ladd, R.T. (2004). A Monte Carlo comparison of relative importance methodologies. *Organizational Research Methods*, 7(3), 258-282.
18. Luchman, J.N. (2014). Relative Importance Analysis with Multicategory Dependent Variables: An Extension and Review of Best Practices. *Organizational Research Methods*, 17(4), 452-471.
19. Luo, W. and Azen, R. (2013). Determining Predictor Importance in Hierarchical Linear Models Using Dominance Analysis. *Journal of Educational and Behavioral Statistics*, 38, 3-31.
20. Brusco, M.J., Cradit, J.D. and Brudvig, S. (2019). An integrated dominance analysis and dynamic programing approach for measuring predictor importance for customer satisfaction. *Communications in Statistics-Theory and Methods*, 48(21), 5290-5307.
21. Pratt, J.W. (1987). Dividing the Indivisible: Using Simple Symmetry to Partition Variance Explained. Paper presented at the Proceedings of the *2nd International Conference in Statistics*, Tampere, Finland.

22. Shou, Y. and Smithson, M. (2015). Evaluating Predictors of Dispersion: A Comparison of Dominance Analysis and Bayesian Model Averaging, *Psychometrika*, 80(1), 236-256.
23. Tang, S. (2014). *Inferential procedures for dominance analysis measures in multiple regression* (Doctoral dissertation). Retrieved from UWM Digital Commons (dc.uwm.edu/etd/645/
24. Thompson, D. (2009). Ranking Predictors in Logistic Regression, Correlation, and Dependence. [online] Scribd. Available at: https://www.scribd.com/document/286483947/Ranking-Predictors-in-Logistic-Regression.
25. Tonidandel, S., LeBreton, J. and Johnson, J. (2009). Determining the statistical significance of relative weights. *Psychological Methods*, 14(4), 387-399.
26. Tonidandel, S. and LeBreton, J.M. (2010). Determining the relative importance of predictors in logistic regression: An extension of relative weight analysis. *Organizational Research Methods*, 13(4), 767-781.
27. Tonidandel, S. and LeBreton J.M. (2011). Relative Importance Analysis: A Useful Supplement to Regression Analysis. *Journal of Business and Psychology*, 26(1), 1-9.
28. Traxel, N.M. and Azen, R. (2006). Predictor importance in logistic regression: An extension of dominance analysis. Poster presented at the *114th Annual Convention of the American Psychological Association*, New Orleans, LA.
29. Wang, X., Duverger, P. and Bansal, H.S. (2013). Bayesian Inference of Predictors Relative Importance in Linear Regression Model Using Dominance Hierarchies. *International Journal of Pure and Applied Mathematics*, 88(3), 321-339.