Journal of Asia Pacific Studies (2018) Volume 5 Issue 1, 1-61

A Basic Understanding to Mediation Analysis and Statistical Procedures in Management Research

***Dr. Meena Madhavan,** Lecturer-Faculty of Business Administration, St.Theresa International College, Thailand
Email: meena2priya@gmail.com

Abstract

The purpose of this study is to facilitate beginners to understand mediation analysis and its' statistical procedures in the field of management. This article explained the evolution of mediation analysis, which provides a basic understanding towards the idea of mediation. And further, the general assumptions to carry out mediation analysis were highlighted. Next, the study covered the different approaches towards mediation; and statistical procedures were outlined to ease the understanding of practicing researchers in the field of management. And it was noted that, bootstrapping is the best method to conduct mediation analysis. One of the major limitations of this study is, the technical part of the statistical theory about variance, estimates, confidence intervals, effect size, etc., were not explained in detail. Also, few modern methods were not included as the motive is to ease the understanding of practicing researchers in the field.

Keywords: *mediation analysis; traditional method; modern methods; hypothesis testing*

Introduction

The purpose of this study is to understand the mediation analysis and statistical tests used to investigate the mediation effects. Scholars in the field of Psychology and Social science widely use mediation analysis to test the causality. There is a high demand among researchers to test the causal effects of the intermediate variable, which exert influence on dependent variable. The mediation analysis is often conducted to examine the type of relationship or effects between independent variable and dependent variable by using the proposed mediating variable. The causal mechanisms assess the indirect effects produced by predictor variable. This causal chain is referred as mediation analysis. Baron and Kenny (1986) developed the causal chain model to test the mediation effects. This model is widely used in social science research. Baron and Kenny's model was popularized in the field of social science research with 75008 citations (according to Google Scholar). Zhao, Lynch and Chen (2010) pointed out that many research projects was revoked at the early stages or staggered at the finishing stage as it was not conformed to Baron and Kenny's condition. The authors presented the nontechnical flaws in the Baron and Kenny's logic and also provided the alternative decision-tree & step-by-step framework for mediation tests. Many researchers have criticized Baron and Kenny's approach with valid logic; but still it is popular in the field. One of the possible reasons may be that the practicing researchers develop their basic understanding to mediation tests by following the traditional Baron and Kenny's approach. Nevertheless, there are modern technological advancements with advanced statistical tools dismiss the use of older mediation tests. The researchers are curious to apply the statistical advancements and consider it as meaningless to use the obsolete tests. This study would ponder out the different approaches i.e. from traditional Baron and Kenny to modern SEM model to ease the understanding of practicing researchers/scholars in the field of Management, which will also cover the application of mediation tests and its statistical procedures.

Mediation Analysis

This part aims to answer the following questions: 1. What is mediation analysis? 2. Why mediation analysis? 3. What is the difference between mediating and moderating variables?

1. What is mediation analysis?

Mediation analysis is more prominent in psychological research (Mackinnon, Fairchild, and Fritz, 2007). The use of mediation analysis in the field of psychology prompted from stimulus-response formula, in which it is stated as *"Mediating refers to the possibility that the process may act as a link between a sensory input and a response not directly connected with it- one main function of such a process".* Also in other words *"mediating processes are fundamentally a means of modifying the way in which sensory control acts, not an absence of it".* This was supported with various examples and experiments in the text book of Psychology by Hebb (1958). Hence, the mediation process or analysis rooted from the field of psychology. Kenny (2018 & n.d.) stated that, the history of mediation tests emerged from the researchers Wright (1934), Fisher (1935), and Hyman (1955). In general mediation analysis is referred as the mechanism to study the cause and effect relationship between the predictor variable (X) and dependent variable (Y), where a mediating variable is hypothesized to intermediate the relationship between X and Y. In other words, the effect of X on Y is intervened or mediated by the mediating variable M and still the causal variable X affects the dependent variable Y (Kenny, 2018). The most common approaches widely used by researchers for testing the meditational hypothesis are Sobel test (Sobel, 1982) and Baron & Kenny's 4 step process (Baron and Kenny, 1986). Later, these tests were considered as obsolete by other researchers in the field due to the raise of modern approaches. And few researchers have compared the models and reported that traditional mediation approaches have low statistical power when compared to the modern approaches (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002; Biesanz, Falk, & Savalei, 2010). The general understanding to mediation process is depicted in the following figure (Kenny, 2018):

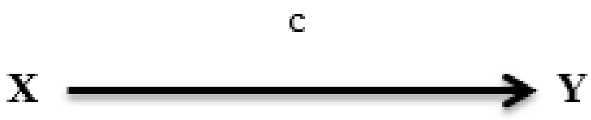


Figure 1

From the Figure 1 it can be understood that X is the causal variable and that causes Y the outcome. This is unmediated model and path c in the above model is called the total effect (Kenny 2018).

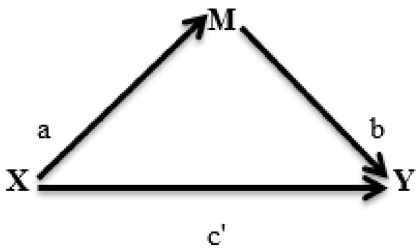


Figure 2

Source: Kenny (2018)

Figure 2 is the mediated model, the effect of X on Y is mediated by a mediating or intervening variable M and still X may cause Y. The path *c'* in the figure 2 is called the direct effect. To probe the evidence of mediation it is important to demonstrate that the effect of the treatment on the outcome variable is zero after the mediator is controlled (Judd and Kenny, 1981). Complete mediation would happen when X no longer affects Y after mediating variable M has been controlled and making path c' zero. Partial mediation would occur when the mediator is introduced the effect of X to Y is reduced or different from zero. Mediation model is the causal model and it is presumed to cause the outcome Y (Kenny, 2018).

When Mediation Analysis should be used?

Major assumptions for mediation analysis as follows:

1. The researchers' research scope is about testing the relationship between three variables, among which one is intervening or mediating variable. Judd and Kenny (1981) stated that it is necessary to follow the process analysis to specify the causal chain that is responsible for treatment effects. This analysis has value in evaluation research for three reasons: firstly, it examines and specifies the causal mechanisms that produce outcomes; secondly, once the theoretical model has been framed for the outcome behavior it is easy to generalize the results in other research settings; thirdly, the researcher knows the process and variables that have direct impact. Also the authors mentioned, in order to claim the mediation there must be three conclusions: 1. the predictor variable causes outcome variable, without this there is no mediation, 2. the predictor variable causes potential mediator, and 3. the mediator must cause the outcome variable controlling for the predictor variable, unless it directly affects the outcome variable it can't be claimed as mediator. It is considered to have mediation effects if there is evidence for the above three conclusions.
2. The mediating hypothesis should be framed accordingly with the proper theoretical support stating the relationships with X or Y or based on its practical applicability. If there is no theoretical support, the final results would be opposite as 'no mediation' and this will not meet the stated research scope/objectives/hypothesis.

3. The researchers' should not confuse the mediation and moderation analysis. The mediation analysis is the analysis which explains the relationship between X and Y, whereas the moderation analysis influences the relationship between X and Y. Baron and Kenny (1986) stated that a moderator is a qualitative (e.g., sex, race, class) or quantitative (e.g., level of reward) variable that affects the direction and/or strength of the relation between an independent or predictor variable and a dependent or criterion variable. And also a moderator variable is introduced when there is a weak or inconsistent relationship between predictor and criterion variable, whereas mediator is introduced when there is a strong relation between the predictor and criterion variable. The understanding towards mediating variable is given in figure 2.
4. Judd and Kenny (2010) stated that it is necessary to have the valid causal assumptions for the mediation to be valid.
5. It is important to consider the standard assumptions for the general linear model such as linearity, normality, homogeneity of error variance, and independence of errors (Kenny, 2018).

Full Mediation and Partial Mediation

The term full mediation is also referred as complete mediation. Baron and Kenny (1986) stated this as perfect mediation, the situation where the independent variable has no effect when the mediator is controlled. Hair, Black, Babin, Anderson, and Tatham (2006) stated that full mediation is where the relationship between predictor variable and outcome variable becomes insignificant after the inclusion of mediating variable. Partial mediation is where the effect of relationship between predictor and outcome variable is reduced and still it is significant after inclusion of the mediating variable. Kenny (2018) stated the difference between complete and partial mediation as follows: *Complete mediation is the case in which variable X no longer affects Y after M has been controlled, making path c' zero. Partial mediation is the case in which the path from X to Y is reduced in absolute size but is still different from zero when the mediator is introduced.*

Statistical analysis for testing mediation effects

Regression

It was suggested by the researchers Judd and Kenny (1981), James and Brett (1984), Baron and Kenny (1986) to conduct the series of regression equations to identify the mediation effects with four step approaches. The equations could be framed accordingly (Testing Mediation with Regression Analysis, 2017).

Step1

$X \xrightarrow{c} Y$ (c path) $Y = B_0 + B_1X + e$ ………………… *Equation (1)*

Step 2

$X \xrightarrow{a} M$ (a path) $M = B_0 + B_1X + e$ ………………… *Equation (2)*

Step 3

$M \xrightarrow{b} Y$ (test path b) $Y = B_0 + B_1M + e$ ………………..... *Equation (3)*

Step 4

$X \xrightarrow{c'}$ / $M \longrightarrow Y$ (test path c', controlling M)

$Y = B_0 + B_1X + B_2M + e$ ………………………………………………………. *Equation (4)*

Difference in these approaches has been discussed by Kenny (2018) that James and Brett (1984) specified that it should not be controlled by the X (causal variable) and assumed the full mediation implicitly, whereas Judd and Kenny (1981) & Baron and Kenny (1986) control for X in step 3. In Baron and Kenny's approach step 4 is not required and Judd and Kenny's approach include all four steps (Kenny, n.d).

For the first 3 steps simple linear regression should be conducted and for the step 4 multiple regression should be conducted (Testing Mediation with Regression Analysis, 2017). If all the four steps are met, it indicates M completely mediates the relationship of X and Y and the first three steps are met and step 4 is not met then it indicates partial mediation (Kenny, 2018).

Example:

If the researcher would like to test the hypothesis using below concept model about the supply chain management practices of the manufacturing industry and its impact on the business performance. Here, the mediating variable is introduced i.e. competitive advantage, because in the recent times the link between supply chain management and strategic management has been addressed by many studies. Testing the model for reliability or validity is not conducted here, as the research area is already existing in pace and the main objective of the study is to discuss about the statistical procedures of mediation analysis. Here X is Supply Chain Management (Predictor/Independent Variable), Y is Business Performance (Criterion/Dependent Variable), and M is Competitive Advantage (Mediating Variable). In this example, there are 4 observed variables for X (Supply Chain Management), 5 observed variables for M (Competitive Advantage), and 4 observed variables for Y (Business Performance).

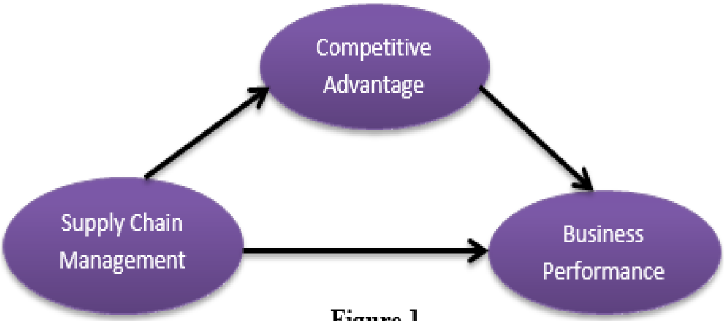


Figure 1

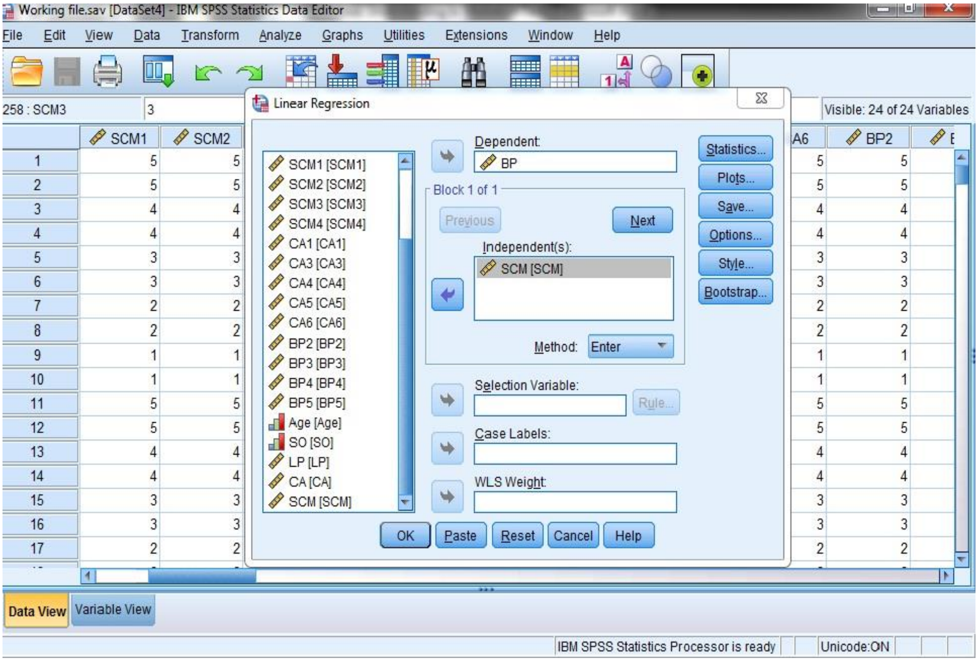
Let's assume that the sample respondents for the study were around 286, which is sufficient to run the advanced statistical models. The dummy dataset were prepared for the study. Also, the following hypotheses were constructed appropriately for the study.

Hypothesis 1: Supply chain management practices of the firms' impact on business performance.
Hypothesis 2: Supply chain management practices of the firms' impact on competitive advantage.
Hypothesis 3: Competitive advantage of the firms' impact on business performance.

The following are the procedures based on Baron and Kenny; the simple linear regression was performed using SPSS 25 Trial version. The results are presented below for all three hypotheses:

For the **Hypothesis 1:** Supply chain management practices of the firms' impact on business performance. The effect of X (Supply Chain Management) on Y (Business Performance) is assessed in the step 1. If the results are not significant, there may be no possibilities for mediation.
After calculating the mean for all 4 observed variables in Supply Chain Management (X) and 4 observed variables in Business Performance (Y), the

variable Supply Chain Management (X) and Business Performance (Y) has been entered in the appropriate boxes below.



Now, after clicking the Ok button the following results were displayed in the SPSS output window.

Table 1.1

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.213[a]	.045	.042	1.21775
a. Predictors: (Constant), SCM				

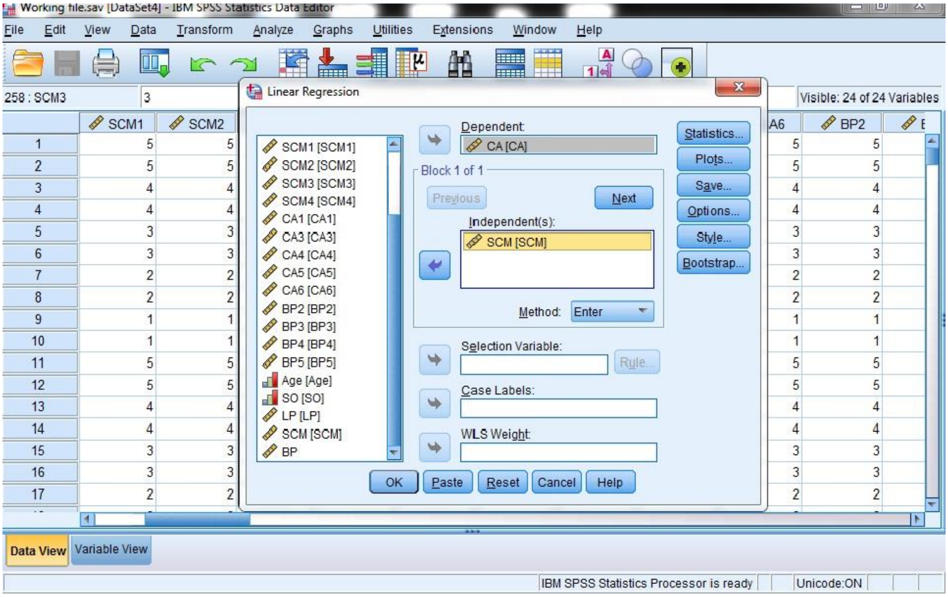
Table 1.2

ANOVA[a]						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	20.029	1	20.029	13.506	.000[b]
	Residual	421.150	284	1.483		
	Total	441.178	285			
a. Dependent Variable: BP						
b. Predictors: (Constant), SCM						

Table 1.3

Coefficients[a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.492	.212		11.768	.000
	SCM	.229	.062	.213	3.675	.000
a. Dependent Variable: BP						

For the **Hypothesis 2:** Supply chain management practices of the firms' impact on competitive advantage. In this step competitive advantage M is regressed against supply chain management X. So while writing it we state competitive advantage as Y, because it is the dependent variable here. The effect of X (Supply Chain Management) on Y (Competitive Advantage) is assessed in the step 2. After calculating the mean for all 4 observed variables in Supply Chain Management (X) and 5 observed variables in Competitive Advantage (M), the variable Supply Chain Management (X) and Competitive Advantage (M) has been entered in the appropriate boxes below.



Now, after clicking the Ok button following results were displayed in the SPSS output window.

Table 1.4

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.252[a]	.063	.060	1.13479
a. Predictors: (Constant), SCM				

Table 1.5

ANOVA[a]						
Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	24.728	1	24.728	19.202	.000[b]
	Residual	365.720	284	1.288		
	Total	390.447	285			

a. Dependent Variable: CA
b. Predictors: (Constant), SCM

Table 1.6

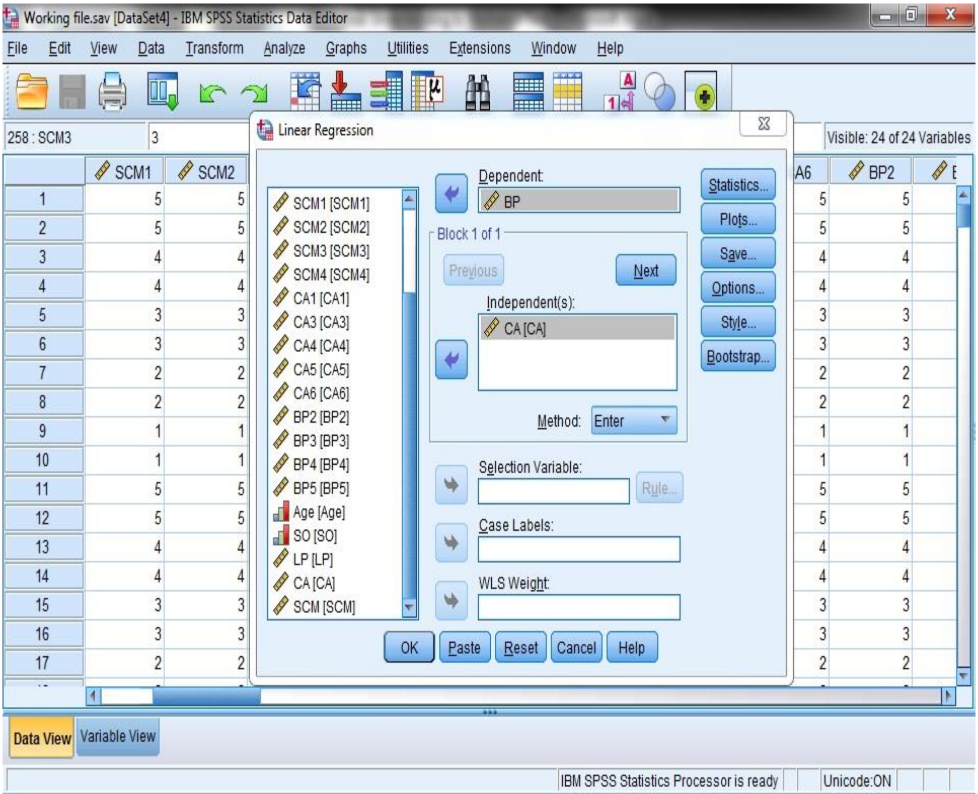
Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.424	.197		12.284	.000
	SCM	.254	.058	.252	4.382	.000

a. Dependent Variable: CA

For the **Hypothesis 3:** Competitive advantage of the firms' impact on business performance. In this step business performance Y is regressed against competitive advantage M. So while writing it we state competitive advantage as X, because it is the independent variable here. The effect of X (Competitive Advantage) on Y (business performance) is assessed in the step 3.

In the previous two hypotheses the mean for all 5 observed variables in Competitive Advantage (M) and 4 observed variables in Business Performance (Y) has been calculated. So, now just the researcher should enter the variables in appropriate boxes as shown below.



Now, after clicking the Ok button following results were displayed in the SPSS output window. This is the simple regression and it is not required to click any other options like statistics, options, etc. The researchers' can go with the default selection in the statistics tab in the right hand side (in first above the plots option).

Table 1.7

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.472[a]	.223	.220	1.09867
a. Predictors: (Constant), CA				

Table 1.8

ANOVA[a]						
Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	98.368	1	98.368	81.492	.000[b]
	Residual	342.811	284	1.207		
	Total	441.178	285			
a. Dependent Variable: BP						
b. Predictors: (Constant), CA						

Table 1.9

Coefficients[a]						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.599	.191		8.356	.000
	CA	.502	.056	.472	9.027	.000
a. Dependent Variable: BP						

The below is the summarized table (Step 1, Step 2, and Step 3) of all three simple linear regressions for the stated hypotheses. This format may be followed when the researchers are reporting it into the dissertation.

Summarized Table 1.10

Hypothesis	Predictor/ Independent Variable X	Dependent Variable Y	a intercept value	B	Standard error b	β	t-value	p-value	Hypothesis Support
Hypothesis 1	Supply Chain Management[a]	Business Performance	2.492	.229	.062	.213	3.675	.000	Yes*
Hypothesis 2	Supply Chain Management[b]	Competitive advantage	2.424	.254	.058	.252	4.382	.000	Yes*
Hypothesis 3	Competitive Advantage[c]	Business Performance	1.599	.502	.056	.472	9.027	.000	Yes*

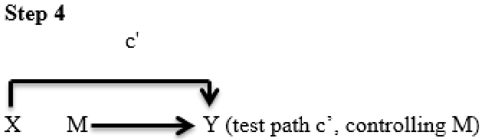
${}^{a}R^2 = .045$, ${}^{b}R^2 = .063$, ${}^{c}R^2 = .223$, $P \leq 0.01$, $n = 286$*

From the above simple linear regression and summary tables, it is observed that all the three stated hypotheses were supported. Using Baron and Kenny’s approach is to find out at the initial stage whether the independent variable is the significant predictor of dependent variable. It is evident from the p-value, which is statistically significant at p<0.01. Following the three stage approach, at the first step business performance was regressed against supply chain management (Hypothesis 1); it is observed that supply chain management has significant and positive impact (b = .213, p<0.01) on business performance F (1,284) = 13.506, p<0.01. In the second step, competitive advantage was regressed against supply chain management (Hypothesis 2); it is observed that supply chain management has significant and positive impact (b = .252, p<0.01) on competitive advantage F (1,284) = 19.202, p<0.001. In the third step, business performance was regressed against competitive advantage (Hypothesis 3); it is observed that competitive advantage has significant and positive impact ($b = .472$, $p<0.01$) on business performance $F (1,284) = 81.492$, $p<0.01$. And also the R square values indicate that the independent variable supply chain management accounted for 4% of the variance in business performance and supply chain management accounted for 6% of the variance in competitive advantage, whereas the variable competitive advantage accounted for 22% of the variance in business performance. It is evident that the Mediator M is having considerable amount of impact when it is regressed against business performance compared to supply chain management. At this stage, the researcher can assess whether the mediator is statistically significant to run the advanced model, although it was claimed by many researchers that Baron and Kenny's approach is obsolete. Hence, this approach gives clear idea to the researchers' interested in doing mediation analysis.

Multiple Regression

Using multiple regression analysis researchers' are interested in identifying the best predictors. And also there is a need to identify those predictors that are supportive of theory. The two approaches which could deduct the best predictors are stepwise regression and hierarchical regression. Hierarchical regression would analyze the effect of predictor variables after controlling for other variables (Lewis, 2007). The researchers can choose the entry of variables in each step based on the theory, the variable which they would like to control. In the first step, the variable which is controlled could be entered. And in the second step, other variables could be entered. Depending on the experiment design of the researcher, the variables could be either entered separately in the each step or based on the hierarchy it could be entered with supporting theory. Also, the researchers' may opt for stepwise methods if they are interested in identifying the predictors that are most effective instead of enter method, although it depends on the researchers' experiment design. Wampold and Freund (1987) stated that hierarchical regression is specifically used to test the theory based hypothesis.

Continuing with the Example and Step 4, Hierarchical multiple regression is conducted. The procedures of SPSS are given below:



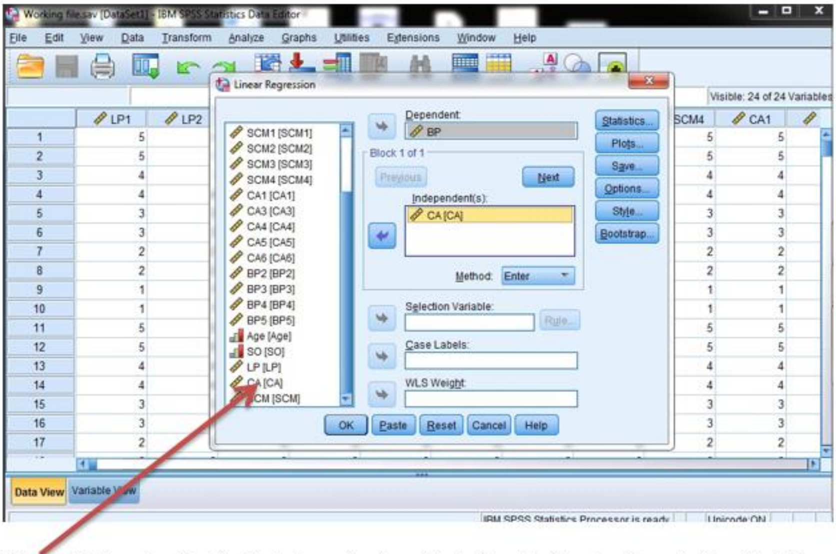
$$Y = B_0 + B_1X + B_2M + e \text{ .... } \textit{Equation (4)}$$

The major purpose for performing hierarchical regression in mediation test is to assess is there any multicollinearity effects and to analyze the effect of predictor variable after controlling for M. Researchers also use hierarchical regression basing Baron and Kenny's approach for testing mediation by analyzing the amount of variance in R^2 after introducing the mediating variables. But here we assess the multicollinearity effects between the two variables (predictor and mediator) in determining the level of dependent variable.

Multicollinearity Effects

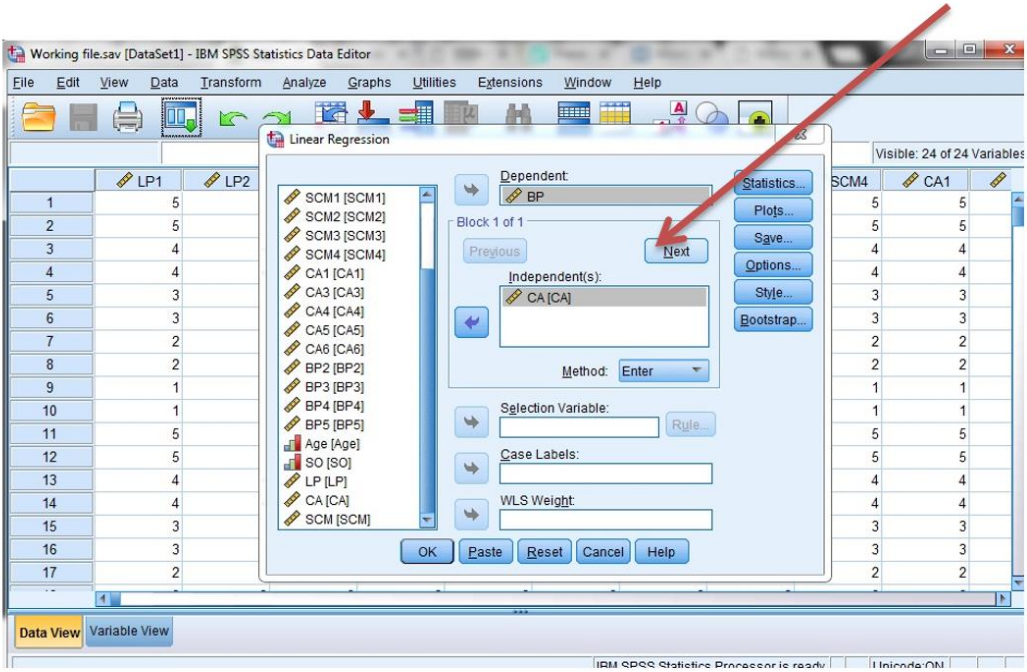
Multicollinearity refers to the situation where two or more explanatory variables are highly linearly related (Hawking, 1983). Perfect multicollinearity is where the relationship between two independent variables is equal to +1 or -1. This occurs in rare datasets due to the redundancy of information. Once the items with redundancies are removed it will be free from multicollinearity effects. In other words, it can be stated as the correlation between two or more explanatory variables is larger than the correlation between the predictor and criterion variables, the perfect multicollinearity exists (Klein, 1962). Basically, multicollinearity problem can be detected under following cases: a. large change in the regression coefficients when the new variables is added or removed, b. insignificant regression coefficients for a variable and rejecting the joint hypothesis those coefficients are statistically significant, c. Hair et. al (2006) mentioned that when pearson correlation coefficient between two independent variables are above 0.8, d. O'Brien (2007) & Hair et al (2006) mentioned about the threshold levels of value of tolerance and variance inflation factor, which value of tolerance less than 0.20 or 0.10, and whereas variance inflation factor above 5.00 or 10.00 indicates a serious problem of multicollinearity. And also the condition index value above 30 indicates the problem of multicollinearity.

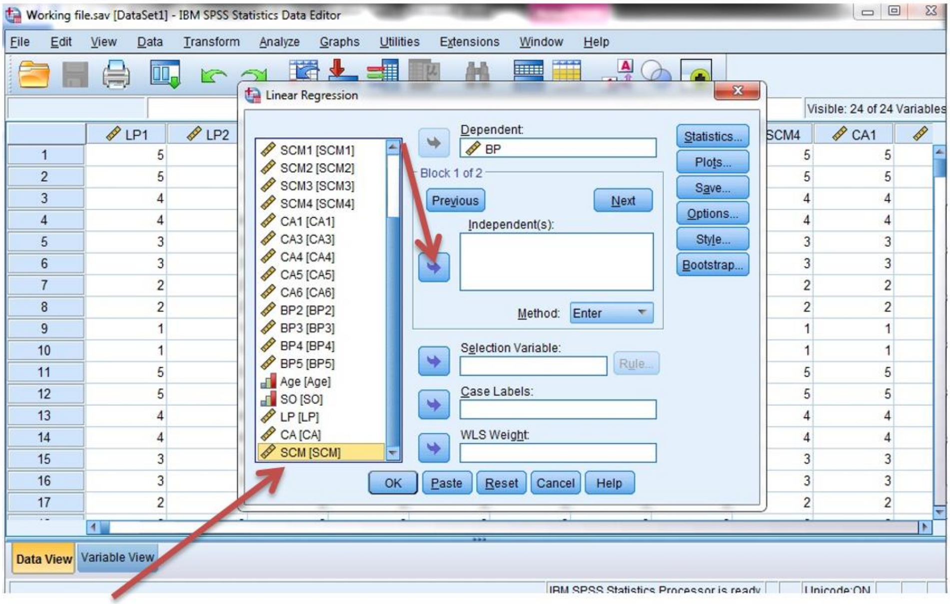
Procedure:



Where, CA is entered in the first step as basic predictor (as stated in step 4, controlling for M).

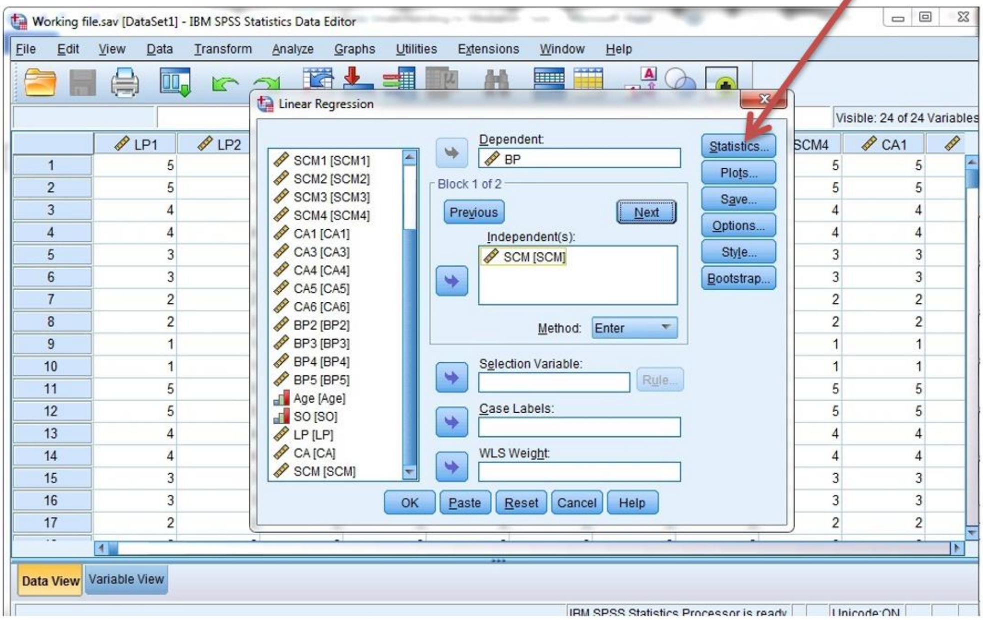
Click Next

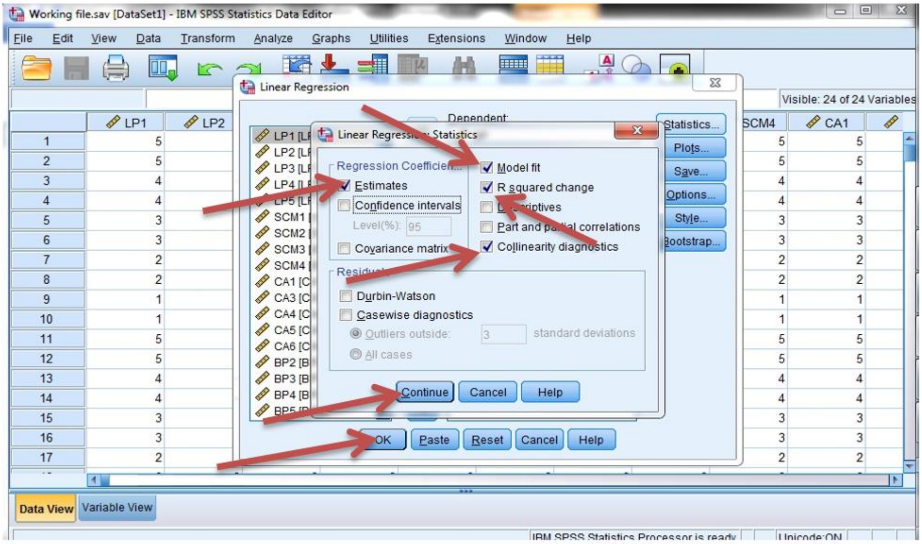

Working file.sav [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Linear Regression
Dependent:
BP
Block 1 of 1
Previous
Next
Independent(s):
CA [CA]
Method: Enter
Selection Variable:
Rule...
Case Labels:
WLS Weight:
SCM1 [SCM1]
SCM2 [SCM2]
SCM3 [SCM3]
SCM4 [SCM4]
CA1 [CA1]
CA3 [CA3]
CA4 [CA4]
CA5 [CA5]
CA6 [CA6]
BP2 [BP2]
BP3 [BP3]
BP4 [BP4]
BP5 [BP5]
Age [Age]
SO [SO]
LP [LP]
CA [CA]
SCM [SCM]
Statistics...
Plots...
Save...
Options...
Style...
Bootstrap...
OK Paste Reset Cancel Help
Visible: 24 of 24 Variables
Data View
Variable View



In the second step, independent variable SCM is entered as an exploratory predictor.

Click Statistics





Select the above items in statistics click continue and then click ok.

The results of multiple hierarchical regression are displayed below.

Table 2.1

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.472[a]	.223	.220	1.09867	.223	81.492	1	284	.000
2	.482[b]	.232	.227	1.09388	.009	3.496	1	283	.063

a. Predictors: (Constant), CA

b. Predictors: (Constant), CA, SCM

Table 2.2

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	98.368	1	98.368	81.492	.000[b]
	Residual	342.811	284	1.207		
	Total	441.178	285			
2	Regression	102.551	2	51.275	42.852	.000[c]
	Residual	338.628	283	1.197		
	Total	441.178	285			

a. Dependent Variable: BP

b. Predictors: (Constant), CA

c. Predictors: (Constant), CA, SCM

Table 2.3

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	1.599	.191		8.356	.000		
	CA	.502	.056	.472	9.027	.000	1.000	1.000
2	(Constant)	1.341	.235		5.695	.000		
	CA	.475	.057	.447	8.305	.000	.937	1.068
	SCM	.108	.058	.101	1.870	.063	.937	1.068

a. Dependent Variable: BP

Table 2.4

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics		
						Tolerance	VIF	Minimum Tolerance
1	SCM	.101[b]	1.870	.063	.110	.937	1.068	.937
a. Dependent Variable: BP								
b. Predictors in the Model: (Constant), CA								

Table 2.5

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions		
				(Constant)	CA	SCM
1	1	1.941	1.000	.03	.03	
	2	.059	5.716	.97	.97	
2	1	2.863	1.000	.01	.01	.01
	2	.086	5.757	.00	.62	.63
	3	.050	7.532	.99	.36	.36
a. Dependent Variable: BP						

The below is the summarized table for multiple hierarchical regression, which the researchers normally use for dissertation writing.

Table No. 2.6 (Summarized Table) Hierarchical Regression results of Business Performance against Competitive Advantage and Supply Chain Management

Variables		**B**	**Std. Error**	**β**	**Tolerance**	**VIF**
Step 1	(Constant)	1.599	.191			
	Competitive Advantage	.502	.056	.472**	1.000	1.000
Step 2	(Constant)	1.341	.235			
	Competitive Advantage	.475	.057	.447**	.937	1.068
	Supply Chain Management	.108	.058	.101*	.937	1.068

*R^2 = .223 for Step 1: ΔR^2 = .009 for Step 2. *P ≥ 0.05, **P ≤ 0.01, n=286.*

The following is the general APA write-up for hierarchical regression. The hierarchical regression was performed to predict the business performance. The regression results revealed that competitive advantage has significant and positive impact (b=.472, p<0.01) on business performance F (1, 284) = 81.492, p<0.01. The multiple correlation coefficient R was at .472 and R square value indicated that competitive advantage accounted for 22% variance in business performance in the first step. And in the second step it is observed that competitive advantage has significant and positive impact (b=.447, p<0.01) on business performance F (2, 283) = 42.852, p<0.01. The multiple correlation coefficient R was at .482 and R square value indicated that competitive advantage accounted for 23% of variance in business performance in the second step. (This is the write up from tables 2.1, 2.2 and 2.3; F value and degrees of freedom could be found in ANOVA table 2.2, b (beta) values could be found in coefficients table 2.3, R and R square values could be found in Model summary table 2.1.)

The multicollinearity effects can be analyzed using the collinearity statistics given in table 2.3. It shows the value of tolerance between competitive advantage and supply chain management is .937, which is ≥0.20, indicates that there is no problem of multicollinearity. And it is further substantiated with Variance of Inflation (VIF) value. The VIF value between competitive advantage and supply chain management is 1.068, which is ≤5.0 indicates that there is no problem of multicollinearity. Therefore, both the mediating and predictor variable, competitive advantage and supply chain management doesn't influence each other in determining the level of business performance.

For General Understanding

Table 2.4 shows the excluded variables from the model. Here the variable supply chain management is excluded in model 1 and competitive advantage alone entered.

Table 2.5 shows the collinearity diagnostics, which indicates relationship between variables and how they vary each other. The values above 1 in condition index indicate the correlation between two or more predictor variables. The values greater than 15 indicate a problem, whereas values of 1 are independent (Stepwise Linear Regression, n.d.).

Indirect Effects

The amount of variation is called as indirect effect. It measures the indirect effect or ab. The total effect can be written as, Total Effect = Direct Effect + Indirect Effect; it can also be denoted in symbols as c = c' + ab. It also equals the reduction of the effect of the causal variable on the outcome or ab = c - c' (Kenny, 2018), which is indirect effect = total effect – direct effect.

Difference of Coefficients and Products of Coefficients Approach

Baron and Kenny’ approach should be supplemented by the difference of coefficients or products of coefficients approach; because most of the researchers fail to calculate the indirect effects. Also, Baron and Kenny’s approach failed to analyze the true mediation effects, which causes type II errors (Mackinnon, Fairchild and Fritz, 2007). To estimate the indirect effects the difference of coefficients and products of coefficients method could be used (Testing mediation with regression analysis, n.d.). The Difference of coefficients approach is proposed by Judd and Kenny (1981). Judd and Kenny recommended finding the difference between of regression coefficients. Referring to the equation 1 and equation 4 in this paper, it is subtracting the partial regression coefficients B_1 (equation 4) obtained through multiple regression from the coefficient B obtained from simple linear regression (equation). This can be written as $B_{indirect} = B - B_1$ (Testing mediation with regression analysis, n.d.).

The products of coefficients approach was proposed by Sobel (1982). It is the product of regression coefficients obtained from two regression models i.e. Equation 4 and Equation 2 in this paper. This could be written as $B_{indirect} = (B_2)(B)$. This approach is about X and M relationship, which is different from the difference approach. The partial regression coefficient for M predicting Y is referred as B_2, whereas the coefficient from simple linear regression X predicting M is referred as B. The products of two regression coefficients reveal the indirect effects (Testing mediation with regression analysis, n.d.).

Both Judd and Kenny’s difference of coefficients approach and Sobel’s products of coefficients approach produce identical values (Mackinnon, Warsi, & Dwyer, 1995). The difference of coefficients approach is about X and Y relationship and the products of coefficients approach is about X and M relationship.

Procedure

The unstandardized regression coefficients from the linear and multiple regression should be included for the analysis.

Calculation from the example, refer to the linear and multiple regression in this paper.

Judd and Kenny's Difference of coefficients approach

$B_{indirect} = B - B_1$ → 0.229 (from table 1.3) – 0.108 (from table 2.3) = **0.121**

Sobel's Products of coefficients approach

$B_{indirect} = (B_2)(B)$ → 0.475 (from table 2.3) * 0.254 (from table 1.6) = **0.121**

Sobel Test

Sobel test was proposed by Sobel (1982), and is often referred by researchers as Delta Method. Sobel test is more conservative and has low power (Mackinnon, Warsi, & Dwyer, 1995). This is because the sampling distribution of ab is highly skewed Kenny (2018). Sobel test is computed from the regression coefficients and standard errors. In order to determine the statistical significance of the indirect effect, a statistic based on the indirect effect must be compared to its null sampling distribution. The Sobel test uses the magnitude of the indirect effect compared to its estimated standard error of measurement to derive a t statistic (Sobel, 1982). Alternatively z or t distributions could be used to determine the significance (Mackinnon, Lockwood, Hoffman, West, and Sheets, 2002).

Procedure

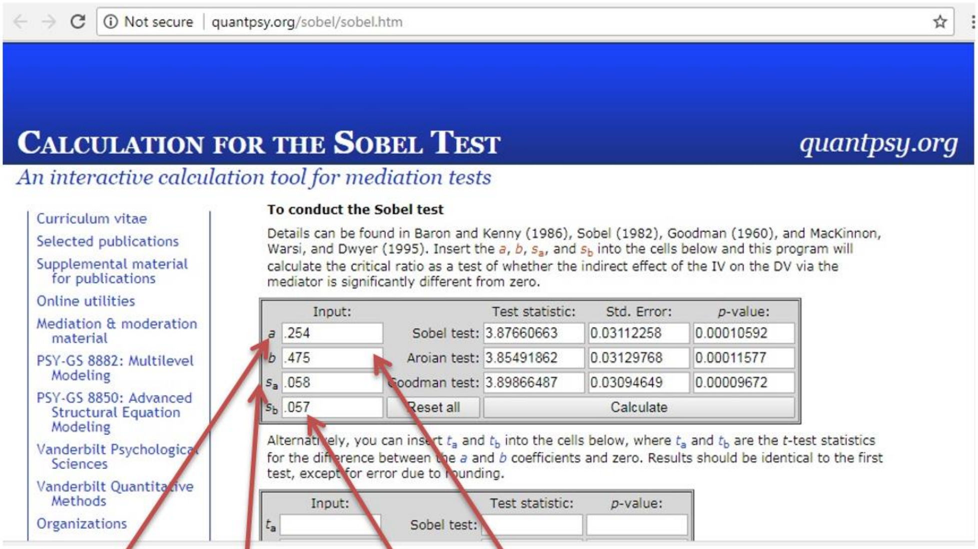
The researchers' can refer to the following link for calculating Sobel test http://quantpsy.org/sobel/sobel.htm (Preacher and Leonardelli, 2001). The webpage serves as an interactive calculation tool for mediation tests developed by Kristopher J. Preacher and Geoffrey J. Leonardelli.

The formulae for all three versions of mediation test were given below by referring to the above link and Mackinnon, Warsi and Dwyer (1995). All the three versions are performed using the above link. The usual Sobel test (Sobel, 1982) omits the third denominator, next the test that one adds the third denominator was popularized by Baron and Kenny (1986) is Aroian Test (1944/1947), whereas the Goodman tests Goodman (1960) subtracts it, and all three tests were tested in this example.

Sobel Test z value = $ab/\sqrt{(b^2 S_a^2 + a^2 S_b^2\ )}$ ……..………… Equation (5)
Aroian Test z value = $ab/\sqrt{(b^2 S_a^2 + a^2 S_b^2 + S_a^2 S_b^2)}$ ……..... Equation (6)
Goodman Test z value = $ab/\sqrt{(b^2 S_a^2 + a^2 S_b^2 - S_a^2 S_b^2)}$ …….. Equation (7)

Mackinnon, Warsi and Dwyer (1995) stated that Sobel and Aroian tests perform well.

Researchers' may enter the values of a and b from respective regression coefficient table to measure the indirect effect or mediation is statistically significant or not. Considering the sample example in this paper and following the above given link, Sobel test and other tests has been calculated.



Value a from table 1.6 **Value b from table 2.3**

Value S_a from table 1.6 **Value S_b from table 2.3**

Source: http://quantpsy.org/sobel/sobel.htm

The results can be written as:
Table 3.1

Input	Test	Test Statistic	Standard Error	P-Value
a = .254	Sobel Test	3.87660663	0.03112258	0.00010592
b = .475	Aroian Test	3.85491862	0.03129768	0.00011577
S_a = .058	Goodman test	3.89866487	0.03094649	0.00009672
S_b = .057				

a = raw regression coefficient (unstandardized) for the association between independent variable and mediator

S_a = standard error of a
b = raw regression coefficient (unstandardized) for the association between mediator and predictor variable, when independent variable is also a predictor of dependent variable.
S_b = standard error of b

Interpretation

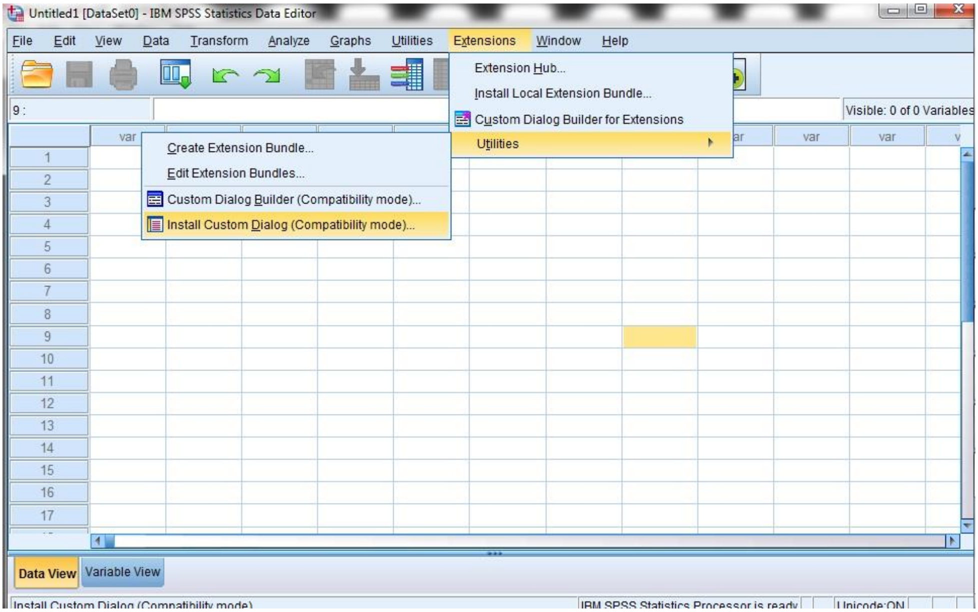
The raw regression coefficient for the supply chain management (SCM) and competitive advantage (CA) is .254 (a) with a standard error of .058 S_a. The raw regression coefficient for the competitive advantage (CA) and business performance (BP) is .475 (b) with a standard error of .057 S_b, when independent variable supply chain management is also a predictor of business performance. The test statistic for Sobel test is 3.876, Aroian test is 3.854, and Goodman test is 3.898 with an associated p-value of 0.00. The standard errors of Sobel, Aroian, and Goodman test are 0.031, mostly identical for all the three tests. The z test value is >1.96 with an associated p-value of <0.05 indicates that the relationship between supply chain management and business performance is mediated by competitive advantage. Hence, there is evidence of complete mediation or indirect effects. The indirect effects are statistically significant.

Bootstrapping Method

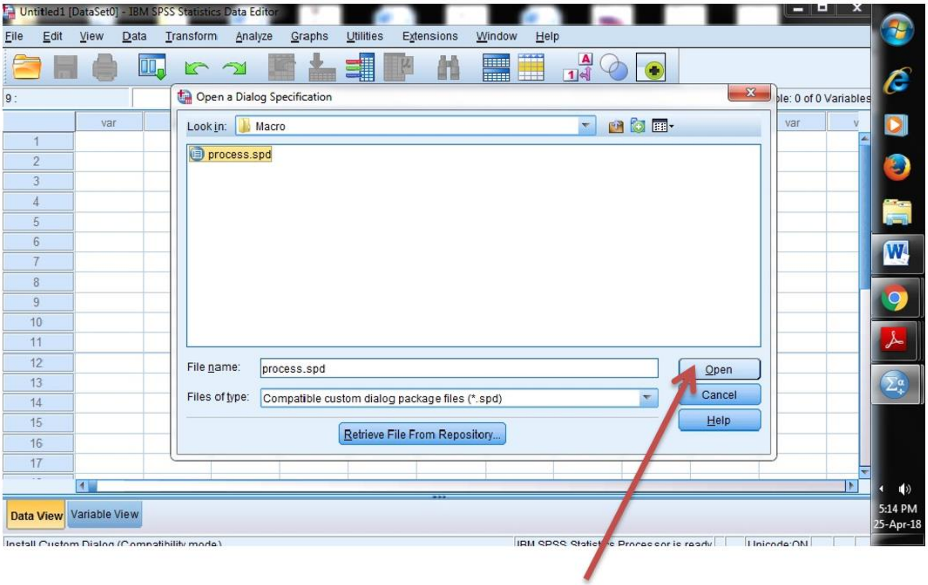
In order to further corroborate the mediation test, bootstrapping method could be considered to measure the direct and indirect effects effectively. Bootstrapping method by professor Preacher & Hayes (2004) performs well, when compared to Baron and Kenny's approach and Sobel's test; because Baron and Kenny's approach and Sobel's test has several criticisms regarding the sample size and type II error. Bootstrapping is a resampling method which is used to estimate confidence interval for indirect effects (Preacher and Hayes, 2004). It can be performed using SPSS Macros, Mplus, R package Lavaan, and Amos. The bootstrapping method can be simply used in SPSS by using process macro 2.16 version written by Hayes (2013).

Bootstrapping method is popular and used by many researchers for measuring indirect effects (Bollen and Stine, 1990; Shrout and Bolger, 2002). Bootstrapping method is a non-parametric method by resampling with replacement eg. 5000 times. From each of these samples the indirect effect is computed and a sampling distribution can be empirically generated. Because the mean of the bootstrapped distribution will not exactly equal the indirect effect a correction for bias can be made. With the distribution, a confidence interval, a *p* value, or a standard error can be determined (Kenny, 2018). If power of the indirect effect is the major concern bias-corrected bootstrap should be used, and if type I error is the major concern percentile bootstrap method is suggested (Hayes and Scharkow, 2013). Basically, Model 4 could be used in Process macro by bootstrapping of 5000 samples in the given example. Model 4 allows up to 10 mediators operating in parallel. Process macro could be downloaded from the following link: http://www.processmacro.org/download.html. Templates of conceptual model are available in the download folder of process macro. The models could be chosen according to the study objective. In this paper, the example has single mediator, hence model 4 has been chosen and demonstrated with steps.

It is easy to install the PROCESS MACRO after download, click utilities under extensions tab, then select Install Custom Dialog (Compatibility mode).

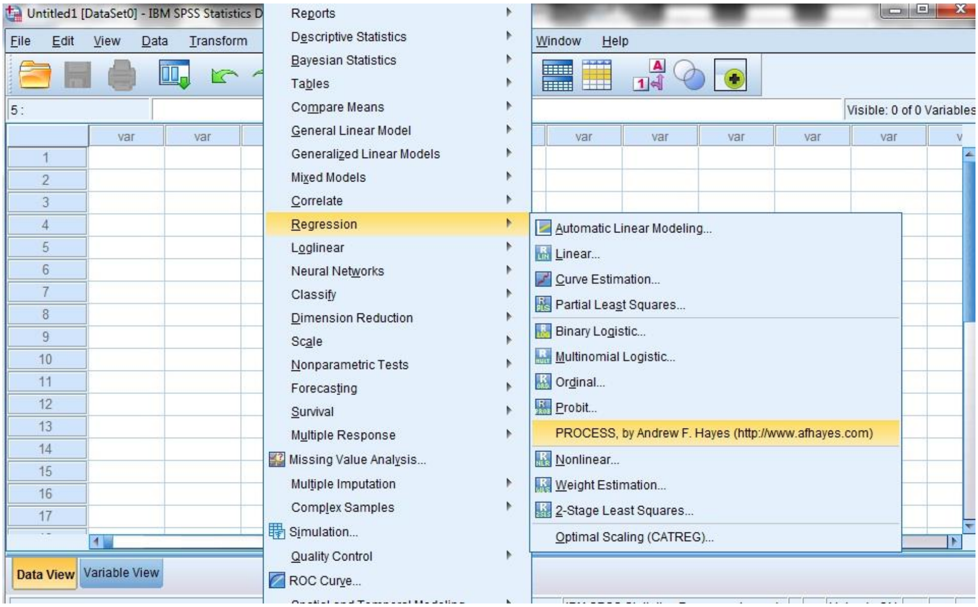


After selecting the Install Custom Dialog (Compatibility mode), the following window will appear.

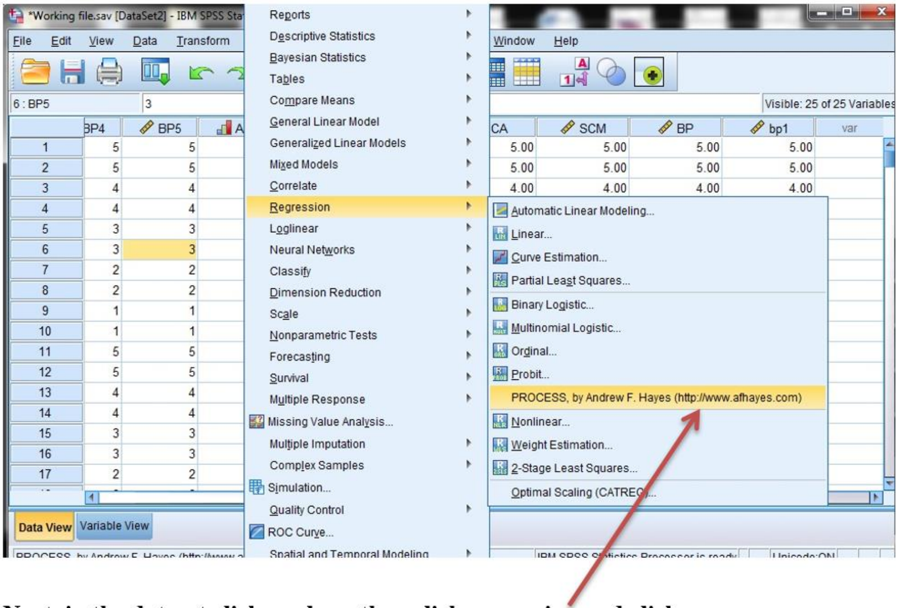


Click Open to Install the Macro

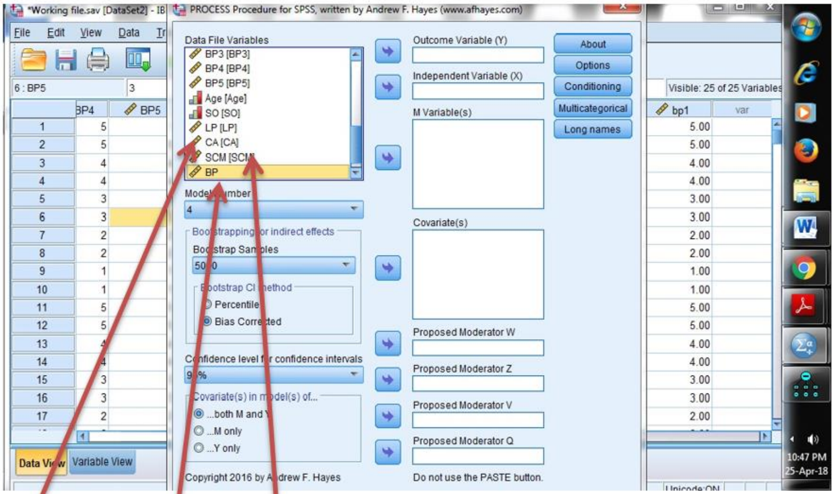
After installation, the macro could be found in SPSS under the following tab.


Untitled1 [DataSet0] - IBM SPSS Statistics D
File
Edit
View
Data
Transform
Window
Help
Visible: 0 of 0 Variables
Reports
Descriptive Statistics
Bayesian Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
ROC Curve...
Automatic Linear Modeling...
Linear...
Curve Estimation...
Partial Least Squares...
Binary Logistic...
Multinomial Logistic...
Ordinal...
Probit...
PROCESS, by Andrew F. Hayes (http://www.afhayes.com)
Nonlinear...
Weight Estimation...
2-Stage Least Squares...
Optimal Scaling (CATREG)...
Data View
Variable View

Next, mean should be calculated for independent variable, mediator and dependent variable.



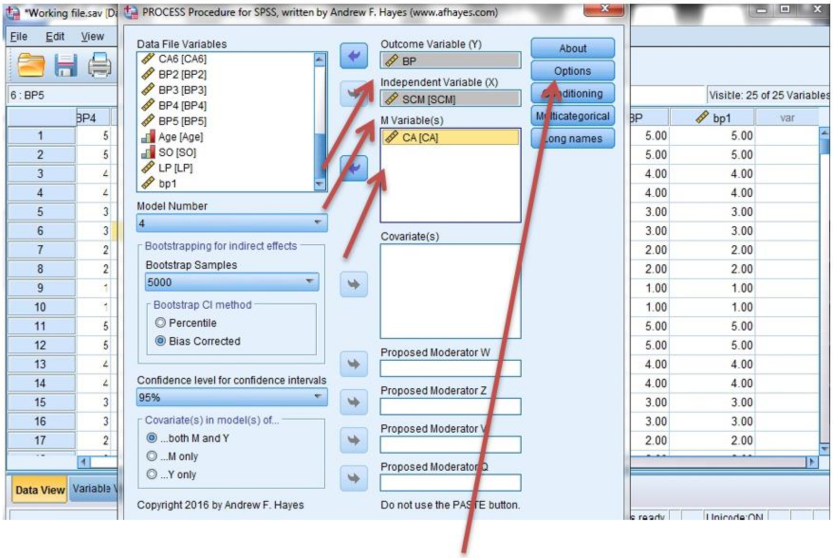
Next, in the dataset click analyze, then click regression and click process.



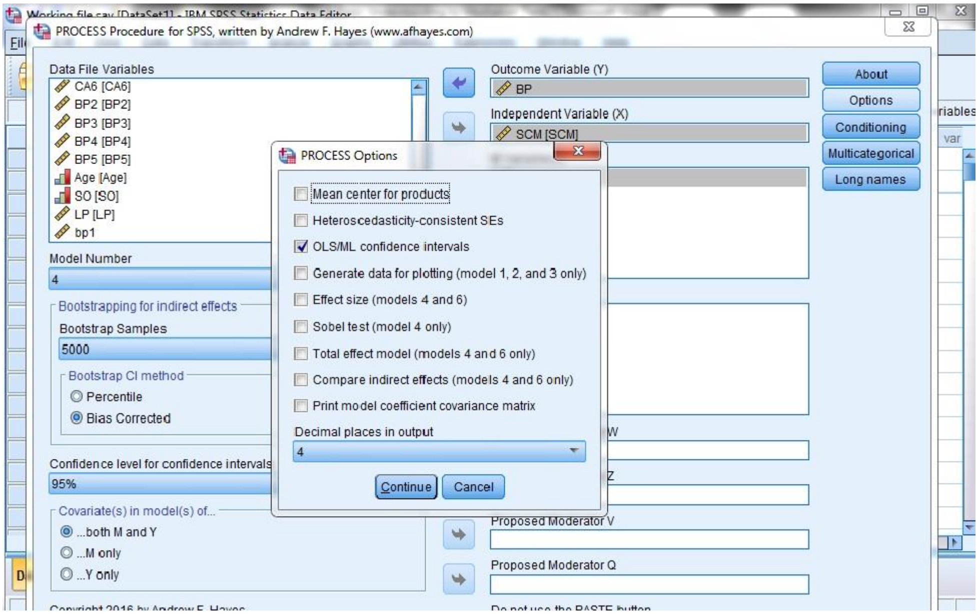
Enter CA in M variables box   Enter SCM in Independent Variable X

Enter BP in Outcome Variable Y

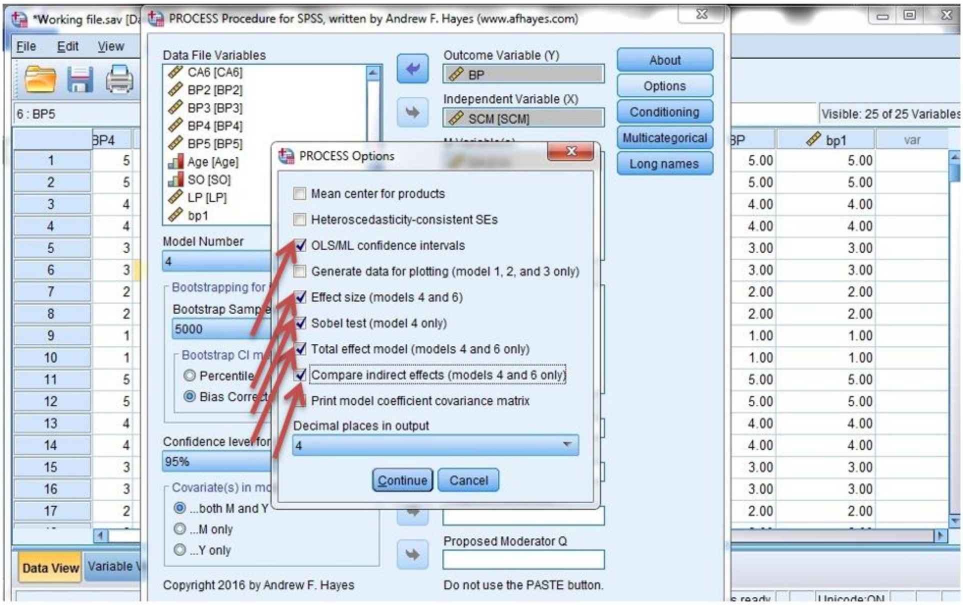
Referring to the example in this paper, the variables should be entered into the respective boxes, SCM (Supply Chain Management) should be entered in the independent variable X box, BP (Business Performance) should be entered in outcome variable Y box, and CA (Competitive Advantage) in M variable box. Then, Model number, bootstrap samples and confident intervals are selected in default. So, based on the example the default selection is more appropriate to use.



After entering variables into the respective boxes, click options.



After clicking options, the box PROCESS Options will be appeared, then the respective boxes should be checked to measure the total effects, indirect effects, and direct effects (Shown below).



Matrix

```
Run MATRIX procedure:

***** PROCESS Procedure for SPSS Release 2.16.1 ****

          Written by Andrew F. Hayes, Ph.D.
www.afhayes.com
    Documentation available in Hayes (2013).
www.guilford.com/p/hayes3

**************************************************************
Model = 4
    Y = BP
    X = SCM
    M = CA
```

```
Sample size
        286

*************************************************************
**************
Outcome: CA

Model Summary
          R       R-sq       MSE          F        df1
df2           p
      .2517      .0633     1.2877    19.2022     1.0000
284.0000      .0000

Model
              coeff          se           t          p
LLCI       ULCI
constant     2.4239       .1973     12.2836      .0000
2.0355     2.8123
SCM           .2542       .0580      4.3820      .0000
.1400       .3683

*************************************************************
Outcome: BP

Model Summary
          R       R-sq       MSE          F        df1
df2           p
      .4821      .2324     1.1966    42.8521     2.0000
283.0000      .0000

Model
              coeff          se           t          p
LLCI       ULCI
constant     1.3405       .2354      5.6952      .0000
.8772     1.8039
CA            .4750       .0572      8.3046      .0000
.3624       .5876
SCM           .1080       .0578      1.8697      .0626      -
.0057       .2217

******* TOTAL EFFECT MODEL *************
```

```
Outcome: BP

Model Summary
          R       R-sq        MSE          F        df1
df2          p
      .2131      .0454     1.4829    13.5062     1.0000
284.0000      .0003

Model
              coeff         se          t          p
LLCI        ULCI
constant      2.4919      .2118    11.7681      .0000
2.0751      2.9087
SCM            .2287      .0622     3.6751      .0003
.1062       .3513

******** TOTAL, DIRECT, AND INDIRECT EFFECTS ******

Total effect of X on Y
     Effect         SE          t          p       LLCI
ULCI
      .2287      .0622     3.6751      .0003      .1062
.3513

Direct effect of X on Y
     Effect         SE          t          p       LLCI
ULCI
      .1080      .0578     1.8697      .0626     -.0057
.2217

Indirect effect of X on Y
       Effect    Boot SE   BootLLCI   BootULCI
CA      .1207      .0360      .0576      .2009

Partially standardized indirect effect of X on Y
       Effect    Boot SE   BootLLCI   BootULCI
CA      .0970      .0282      .0468      .1595

Completely standardized indirect effect of X on Y
       Effect    Boot SE   BootLLCI   BootULCI
```

```
CA       .1125      .0333      .0537      .1854

Ratio of indirect to total effect of X on Y
       Effect    Boot SE   BootLLCI   BootULCI
CA       .5278      .4733      .2790     1.0193

Ratio of indirect to direct effect of X on Y
       Effect    Boot SE   BootLLCI   BootULCI
CA      1.1178    30.7850      .3022     9.1097

R-squared mediation effect size (R-sq_med)
       Effect    Boot SE   BootLLCI   BootULCI
CA       .0359      .0194      .0090      .0879

Normal theory tests for indirect effect
     Effect         se          Z          p
      .1207      .0313     3.8538      .0001

******* ANALYSIS NOTES AND WARNINGS ******

Number of bootstrap samples for bias corrected bootstrap
confidence intervals:
     5000

Level of confidence for all confidence intervals in output:
    95.00

NOTE: Kappa-squared is disabled from output as of version
2.16.

------ END MATRIX -----
```

The above displayed is the results from process macro by bootstrapping of 5000 samples. Now, the researcher should be able to interpret the above results about the indirect effects and direct effects.

Interpretation

Considering the example discussed in this paper, the following interpretation has been made for the understanding about mediation effects.

Table 4.1 Direct, Indirect and Total Effects

Interactions	**Path**	**Coefficient**	**Standard Error**	**t-Value**	**P-Value**
b(YX)	*c*	.2287	.0622	3.6751	.0003
b(MX)	*a*	.2542	.0580	4.3820	.0000
b(YMX)	*b*	.4750	.0572	8.3046	.0000
b(YXM)	*c'*	.1080	.0578	1.8697	.0626

The results are summarized in the above table according to their paths. The coefficient values, standard error, t-value, and p-value were taken from the matrix table respectively based on their paths. The variables were regressed accordingly and the results could be interpreted as the following. The mediation effects was measured using process macro at 95% confidence interval with 5000 bootstrap resamples (Preacher and Hayes, 2008). It is observed from the results that supply chain management (SCM) practices was positively associated with business performance (BP) (b=.2287, t (284) = 3.6751, p<0.01). Also, it is observed that supply chain management (SCM) practices was positively associated with competitive advantage (CA) (b=.2542, t (284) = 4.3820, p<0.01). It is revealed from the results that the mediator competitive advantage (CA) is positively associated with business performance (BP) (b=.4750, t (283) = 8.3046, p<0.01). It is noted from the results that the paths a and b are statistically significant, hence the basic criteria has been satisfied. The bias-confidence interval estimates has been used as default (Preacher and Hayes, 2004; Mackinnon, Lockwood, and Williams, 2004). It is evident from the results that competitive advantage (CA) mediates the relationship between supply chain management (SCM) and business performance (BP) i.e. (b = c – c’ is .1207, Standard Error = .0360, and Confidence Interval CI = .0576 to .2009 (CI should be different from zero)), which indicates that the indirect effects are statistically significant. Although, the results revealed that the direct effects of supply chain management (SCM) practices on business performance (BP) is statistically non-significant when controlling the mediator competitive advantage (CA) (b=.1080, t (283) = 1.8697, p = 0.0626, which is greater than p value 0.05), means supply chain management (SCM) practices no longer predicts Y or is lessened predicting Y i.e. path c'. The Sobel test statistic was reported at z = .1207, with significant p value at p<0.01. Hence, it can be concluded that there is an evidence of complete mediation. And this is depicted in the following diagram.

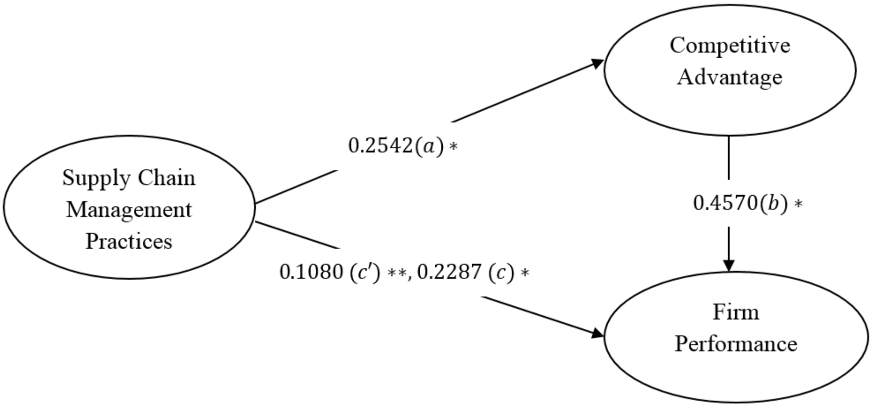


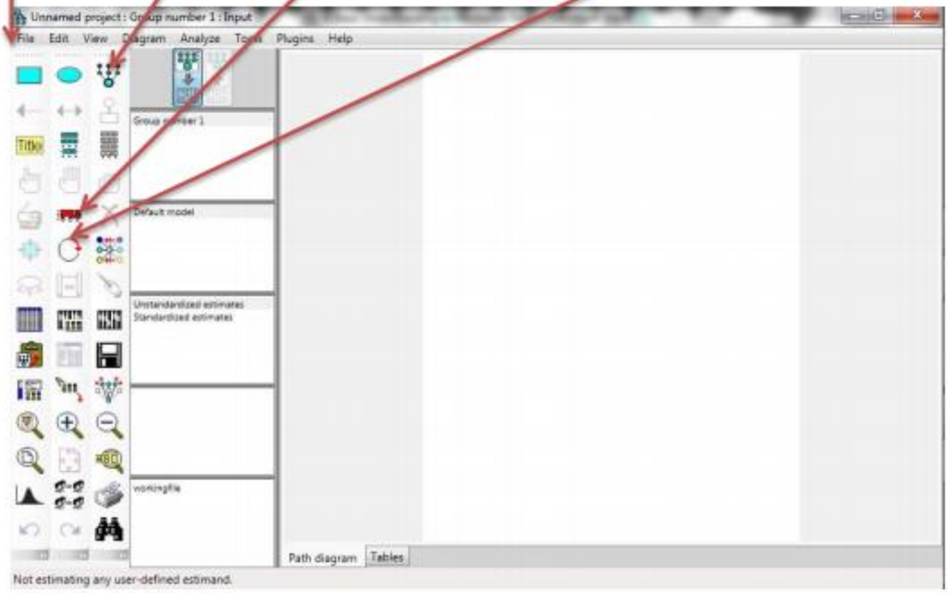
Figure 4.1 Mediating role of Competitive Advantage between Supply Chain Management Practices and Business Performance

Note: *p<0.01 and **p>0.05

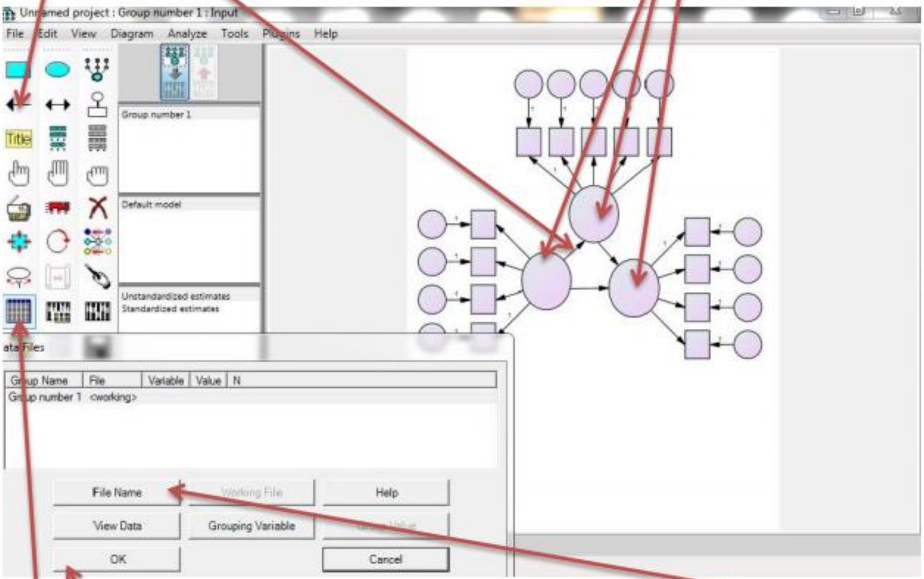
Structural Equation Model

The further substantiation could be made if the researchers' are interested in using structural equation model to probe the indirect effects are statistically significant (Shrout and Bolger, 2002). The programs AMOS, R Lavaan, LISREL can perform structural equation models, and the more programs with newer versions are introduced with more compatibility. Here considering the same example, it is not necessary to develop two models with mediator and without mediator, according to Kenny (2018); because the program AMOS is more compatible to give the results of Direct and Indirect Effects. Also, the total effects c could be seen by calculating the value for c' + ab. For the example discussed in this paper structural equation model was performed by bootstrapping 2000 samples using biased-confidence interval method. It would be good to use structural equation modeling for causal research as other models like Sobel test, and Process macro uses unstandardized coefficients of regression. Also, it could be noted that there is no standard suggestions for bootstrap sample numbers for obtaining accurate results with standard errors (Nevitt and Hancock, 2001); but if the number of bootstrap samples are high, there is possibility for holding good statistical power (Davidson and Mackinnon, 2000). Usually, the bootstrap works well with n = 300 sample size (Ishikawa and Konishi, 1995), the present study has samples closer to 300. The bootstrap of 2000 samples was considered and used here based on the assumption that non-normality may exist at different conditions.

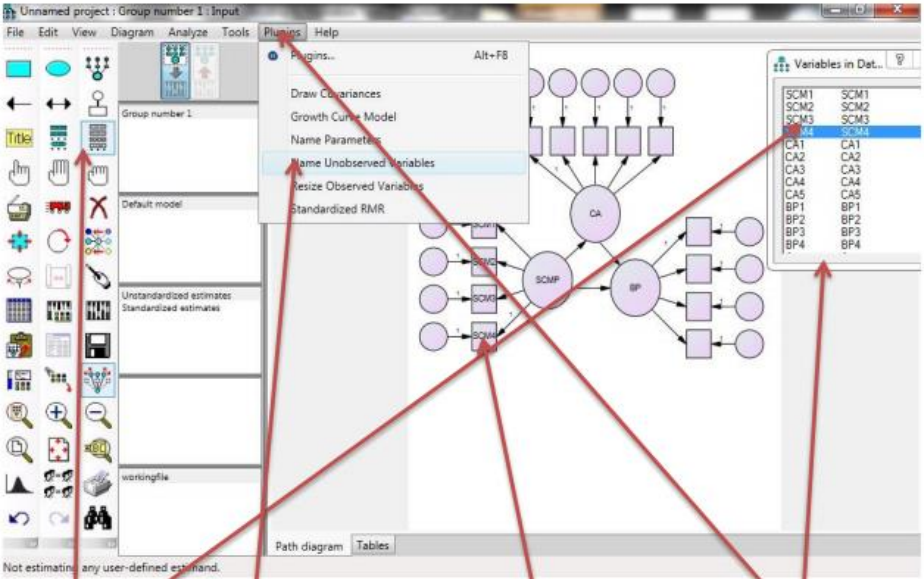
The sample data was ran in AMOS 21 trial version. Open Amos, then click File to open to New file, then Click the symbol below to draw latent variables and observed variables, after clicking the symbol, bring the cursor to the page and click one time to draw the latent variable and 4 times (SCM has 4 observed variables in this examples) to draw observed variables. Similarly repeat this procedure to draw the latent and observed variables. Next, click this symbol to rotate the variables, and then click this symbol to move the variables, so that it can very well fit in the page.



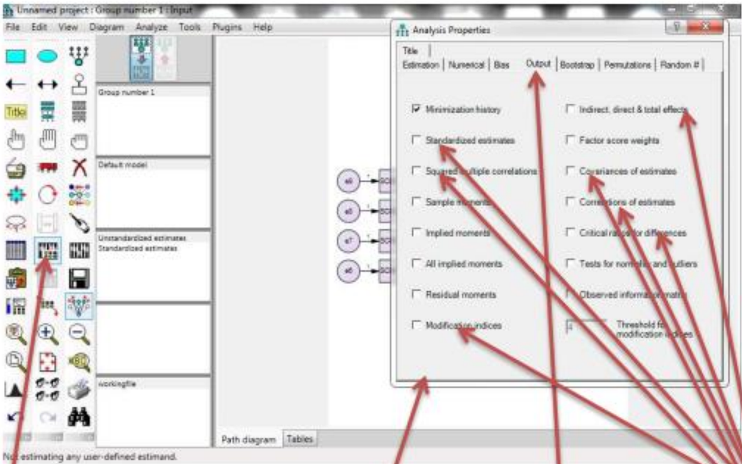
After completing the procedures in the previous step, the figure may look like this. In this step, the symbol below can be used to draw the arrow between the independent variable, dependent variable, and mediating variables, which specify the paths. Next, click each latent variable to name it.



Click the above object to select data files, then browse the data files by clicking File Name and then click ok.

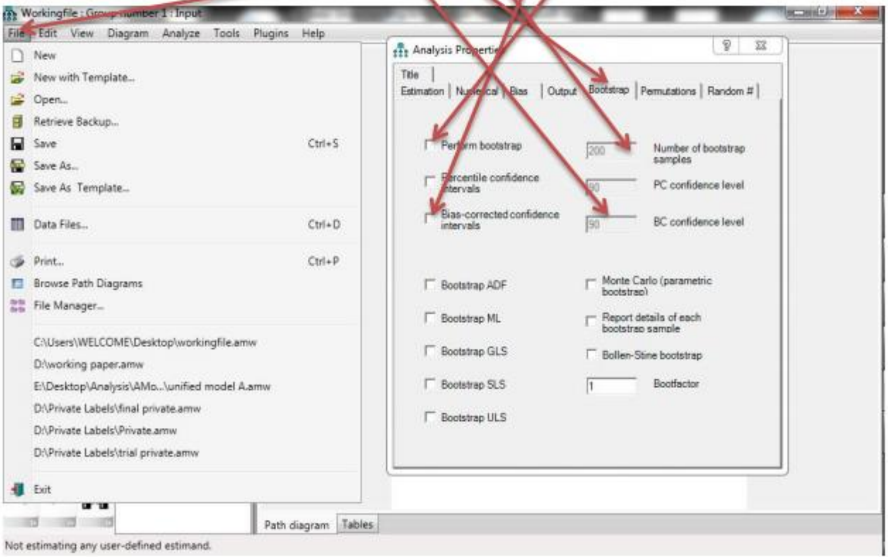


Next, Click the above object to list to variables in the data set, then the box will appear, now we may select the variables and enter into the respective observed variable boxes. Next, click the plugins tab and select name unobserved variables to name the unobserved variables.

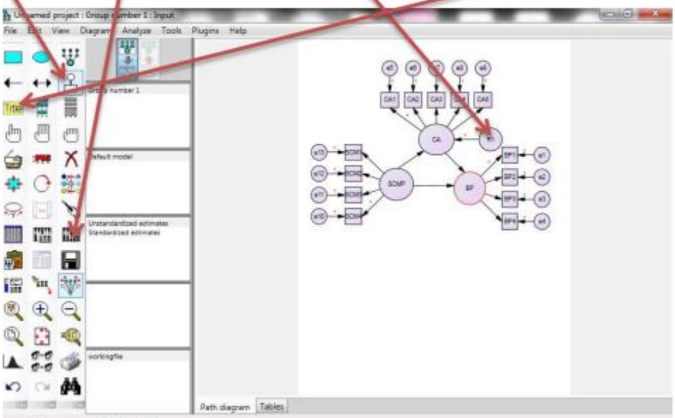


Next, click Analysis Properties, a box will appear then select output section and select the relevant analysis required to measure the direct and indirect effects.

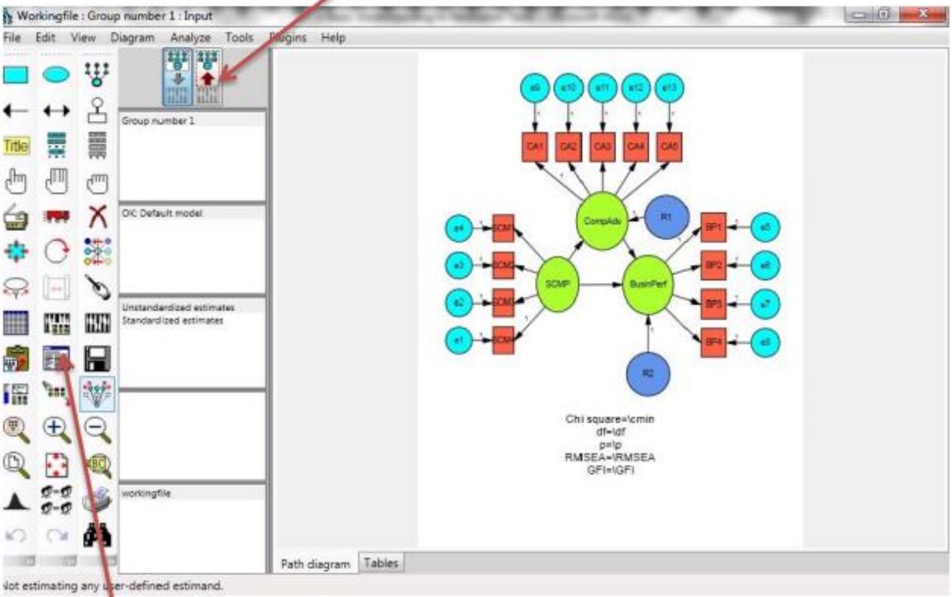
Next, in the analysis properties click the bootstrap section and select perform bootstrap and change the number of bootstrap samples as 2000 and select bias-corrected confidence interval and change the BC confidence level as 95. And then click the file and select Save As and save the file.



Next, click unique variable symbol to create residual for latent variable and it can be rotated clockwise, then click calculate estimates to calculate the results. Click title & add cmin, GFI, etc.



Once we run the program, AMOS may give error sometimes as Eg: 'the observed variable SCMP is denoted by the ellipse in the path diagram'. For this, we need to rename the variable. Ignore the error and run again. For coloring each variable, double click the variable box and then select colors. After running the model,Click this to view results. if the model is fit, this object will turn bright.



Next, the output can be viewed by clicking this. In the results section, we have to check the estimates, standardized regression estimates to see the paths and p-values. To check the bootstrapping results click estimates, then click matrices and then finally click indirect effects. Now we can see the estimates of indirect effects, click Bootstrap confidence at the bottom to check the lower bound and upper bound bootstrap bias corrected confidence level to understand whether the indirect effects are statistically significant. The confidence level should be different from zero and the p-value should be less than 0.05.

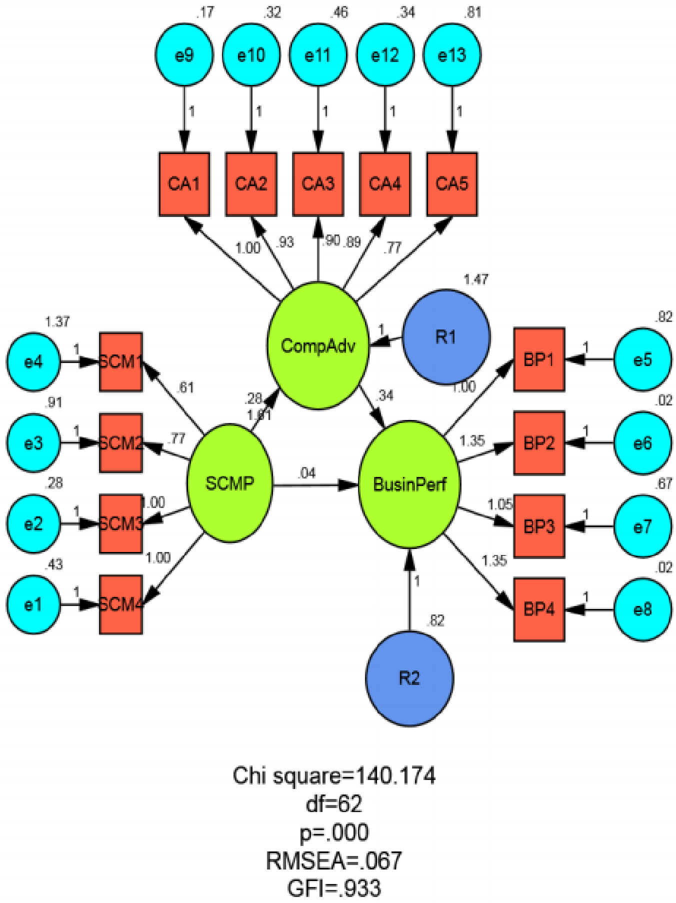


Figure 5.1 SEM Model

Interpretation (For the example discussed in this paper)

The estimates revealed that the effect of supply chain management on business performance is non-significant with a p-value > 0.05 (Refer to table 5.1). The model converged with a chi-square $\chi_{(62)}$ value of 140.174 at $p \leq 0.05$ (Refer to Figure 1). To analyze the existence of mediation, bootstrapping results were considered. The indirect effect of supply chain management practices on business performance revealed that b = .093, Standard error = .030, confidence interval values of lower bound and upper bound is .045 and .164, which is different from zero and the indirect effects are statistically significant at $p \leq 0.01$ (Refer to table 5.2). Hence, it proves the evidence of mediation and the indirect effect of supply chain management practices on business performance is statistically significant. In order to assess the type of mediation existing direct effects were considered. The direct effect of supply chain management practices on business performance revealed that b = .041, Standard error = .049, confidence interval values of lower bound and upper bound is -.047 and .143, which is zero or negative and the direct effects are statistically non-significant at $p \geq 0.05$ (Refer to table 5.3). The direct effect of supply chain management practices on firm performance is non-significant, and this confirms the evidence of complete mediation. Hence, the relationship between supply chain management practices and business performance advantage is mediated by competitive advantage. The model is fit at all respect according to the threshold levels suggested by Hu and Bentler (1999) for GFI (Figure 5.1), AGFI, RMSEA, RMR, and SRMR.

Table 5.1 Estimates

Paths			Estimate	S.E.	C.R.	P	Label
CompAdv	<---	SCMP	.275	.062	4.467	***	par_11
BusinPerf	<---	CompAdv	.340	.050	6.852	***	par_12
BusinPerf	<---	SCMP	.041	.047	.882	.378	par_13

Table 5.2 Indirect Effects

Estimates			
	SCMP	**CompAdv**	**BusinPerf**
CompAdv	.000	.000	.000
BusinPerf	.093	.000	.000
Bootstrap Standard Errors			
	SCMP	**CompAdv**	**BusinPerf**
CompAdv	.000	.000	.000
BusinPerf	.030	.000	.000
Bias corrected Confidence Interval Lower Bound			
	SCMP	CompAdv	BusinPerf
CompAdv	.000	.000	.000
BusinPerf	.045	.000	.000
Bias corrected Confidence Interval Upper Bound			
	SCMP	**CompAdv**	**BusinPerf**
CompAdv	.000	.000	.000
BusinPerf	.164	.000	.000
Bias Corrected Two Tailed Significance			
	SCMP	**CompAdv**	**BusinPerf**
CompAdv	...	...	...
BusinPerf	.001	...	...

Table 5.3 Direct Effects

Estimates			
	SCMP	**CompAdv**	**BusinPerf**
CompAdv	.275	.000	.000
BusinPerf	.041	.340	.000
Bootstrap Standard Errors			
	SCMP	**CompAdv**	**BusinPerf**
CompAdv	.068	.000	.000
BusinPerf	.049	.057	.000
Bias corrected Confidence Interval Lower Bound			
	SCMP	CompAdv	BusinPerf
CompAdv	.138	.000	.000
BusinPerf	-.047	.234	.000
Bias corrected Confidence Interval Upper Bound			
	SCMP	**CompAdv**	**BusinPerf**
CompAdv	.398	.000	.000
BusinPerf	.143	.455	.000
Bias Corrected Two Tailed Significance			
	SCMP	**CompAdv**	**BusinPerf**
CompAdv	.001	...	...
BusinPerf	.331	.001	...

The researchers' may write in detail about model fit indices. Here it is not given in detail as the objective is to explain the indirect effects.

Discussion

For performing the mediation tests it is important for the researchers' to frame the hypothesis based on the theory and practical applicability, otherwise the researchers' may get negative results as 'no mediation'. In this article, an example has been used with dummy dataset to explain the statistical procedures with interpretation. In this article, Firstly, Baron and Kenny's approach was demonstrated through series of regression analysis i.e. simple linear regression. Secondly, hierarchical regression was performed to test the multicollinearity effects, and reported that there is no multicollinearity effects. Because, for analyzing the mediation effect it is important to check the multicollinearity issues. Thirdly, indirect effects and the calculation of indirect effects using difference of coefficients and products of coefficients approach were discussed. Also, Sobel test and other versions of the related test were performed to analyze the mediation and its statistical significance. Fourthly, Bootstrapping method was utilized. Bootstrapping is a popular method in testing the mediation (Shrout and Bolger, 2002). Bootstrap method was initially published by Efron (1979), later the statisticians in the field developed the extensions like improved estimates of the variance, Bayesian approaches, Bias-Corrected, and Bias-corrected and accelerated bootstrap, etc. It is resampling the sample data to control the stability of the results. It was performed using process macro by professor (Hayes, 2013). The bootstrapping method helps to estimate the confidence interval for indirect effect. That's the reason bootstrapping method is strongly recommended by Hayes (2013) for mediation analysis. Based on the example dummy dataset used in this article, the results of direct, indirect effects and Sobel test were reported appropriately. Fifthly, Structural Equation Modeling was performed to analyze the Direct, Indirect and Total Effects using Bootstrapping method and the results of the sample data were reported appropriately. Hence, five approaches (Baron and Kenny, Difference of coefficients, Products of coefficients, Sobel Test, and Bootstrapping) were used to estimate the indirect or mediation effects of the variables used in this example. Although, each approach has its own criticisms, this article with step-by-step procedures helps the beginners to develop the understanding towards mediation effects. Kenny, Kashy, and Bolger (1998) restated that four step procedures should be undertaken for testing the mediation. Researcher's like Collins, Graham and Flaherty (1998) questioned the Baron and Kenny's first step i.e. testing the relationship between X and Y when the researchers' are supposed to test the mediation. Also, it is possible to find the total effect by calculating c' + ab. Still, logically the question raised by Collins, Graham and Flaherty (1998) seems to be fine; however the researchers might be interested in framing and testing the hypothesis for X to Y without the mediator to probe the theory. Shrout and Bolger (2002) stated that first step of Baron and Kenny's approach should be dropped unless the effect of X on Y is large or medium based on the theory; also Baron and Kenny's approach has low statistical power (Mackinnon, Lockwood, Hoffman, West, and Sheets, 2002). In overall bootstrapping procedure seems to perform well when compared to Sobel's test (Kenny, 2018). Shrout and Bolger (2002) further stated that bootstrapping method is more appropriate to test the mediation, where the mediator and outcome variable is not normally distributed. Bootstrapping method could also be used in structural equation modeling, where the option is available in statistical software AMOS (Arbuckle, 1999). Finally, it is observed that structural equation models using bootstrap method perform better than regressions and Sobel test. The researchers' may read more from Kenny (2018 & n.d.).

Conclusion

This article is an attempt to provide basic understanding towards mediation tests for the beginners. The dummy dataset was used to run the mediation models using the traditional approach to modern approach; this would develop an understanding towards mediation analysis and the statistical procedures. This article is also with certain limitations that it doesn't focus on few modern methods like Monte Carlo method, Hierarchical Bayesian method, and Likelihood-based confidence interval. And it considered an example with single mediator, where multiple mediator models, mediated moderation, moderated mediation, and multilevel mediation was not demonstrated and discussed. Also, the technical part regarding the statistical theory about variance, estimates, confidence intervals, effect size, etc., were not explained in detail. However, this article would serve the beginners as a guideline to conduct mediation tests in the field of management.

References

Arbuckle, J. L. (1997). *AMOS users" guide*, Version 3.6. Chicago: SmallWaters.

Aroian, L. A. (1944/1947). The probability function of the product of two normally distributed variables. *Annals of Mathematical Statistics, 18*, 265-271.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, *51*(6), 1173.

Biesanz, J. C., Falk, C. F., & Savalei, V. (2010). Assessing mediational models: Testing and interval estimation for indirect effects. *Multivariate Behavioral Research*, *45*(4), 661-701.

Bollen, K. A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological methodology*, 115-140.

Collins, L. M., Graham, J. J., & Flaherty, B. P. (1998). An alternative framework for defining mediation. *Multivariate Behavioral Research*, *33*(2), 295-312.

Davidson, R., & MacKinnon, J. G. (2000). Bootstrap tests: How many bootstraps?. *Econometric Reviews*, *19*(1), 55-68.

Efron, B. (1979). Computers and the theory of statistics: thinking the unthinkable. *SIAM review*, *21*(4), 460-480.

Fisher, R.A. (1935). *The design of experiments.* Oxford: Oliver & Boyd.

Goodman, L. A. (1960). On the exact variance of products. *Journal of the American Statistical Association*, *55*, 708-713.

Hair, J. F., B. Black, B. Babin, R. E. Anderson, and R. L. Tatham. (2006). *Multivariate Data Analysis (6th ed.).* Upper Saddle River, New Jersey: Pearson Education, Inc.

Hawking, R. R. and Pendleton, O. J. (1983). The regression dilemma. *Commun. Stat.- Theo. Meth, 12*, 497-527.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* Guilford Press.

Hayes, A. F., & Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter?. *Psychological science*, *24*(10), 1918-1927.

Hebb, D.O. (1958). *A textbook of Psychology*, W.B. Saunders Company: London

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55.

Hyman, H. H. (1955). *Survey design and analysis*. New York: Glencoe, IL: The Free Press.

Ichikawa, M., & Konishi, S. (1995). Application of the bootstrap methods in factor analysis. *Psychometrika*, *60*, 77–93.

James, L. R., & Brett, J. M. (1984). Mediators, moderators, and tests for mediation. *Journal of Applied Psychology*, *69*(2), 307.

Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation review*, *5*(5), 602-619.

Judd, C. M., & Kenny, D. A. (2010). *Data analysis in social psychology: Recent and recurring issues*. John Wiley & Sons, Inc..

Kenny, D.A. (n.d.). Mediation: Background and Basics. Retrieved from: http://www.da.ugent.be/cvs/pages/en/Presentations/Presentation%20David%20Kenny.pdf

Kenny, D.A. (2018). *Mediation.* Retrieved from: http://davidakenny.net/cm/mediate.htm

Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology (In D. Gilbert, S. Fiske, & G. Lindzey (Eds.). The handbook of social psychology (Vol. 1, pp. 233–265).

Klein, L. R. (1962). *An Introduction to Econometrics*, Prentice-Hall.

Lewis, M. (2007). Stepwise versus Hierarchical Regression: Pros and Cons. *Online Submission*.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annu. Rev. Psychol.*, *58*, 593-614.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological methods*, *7*(1), 83.

MacKinnon, D. P., Warsi, G., & Dwyer, J. H. (1995). A simulation study of mediated effect measures. *Multivariate behavioral research*, *30*(1), 41-62.

Nevitt, J., & Hancock, G. R. (2001). Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. *Structural equation modeling*, *8*(3), 353-377.

O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*(5), 673-690.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*, 717-731.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods*, *40*(3), 879-891.

Preacher, K. J., & Leonardelli, G. J. (2001). *Calculation for the Sobel test: An interactive calculation tool for mediation tests.* Retrieved from: http://quantpsy.org/sobel/sobel.htm

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*, 422-445.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological methodology*, *13*, 290-312.

Stepwise Linear Regression, (n.d.). http://www.geog.leeds.ac.uk/courses/other/statistics/spss/stepwise/

Testing Mediation with Regression Analysis. (2017). Retrieved from: http://web.pdx.edu/~newsomj/semclass/ho_mediation.pdf

Wampold, B. E., & Freund, R. D. (1987). Use of multiple regression in counseling psychology research: A flexible data-analytic strategy. *Journal of Counseling Psychology*, *34*(4), 372.

Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics, 5,* 161-215.

Zhao, X., Lynch Jr, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of consumer research*, *37*(2), 197-206.