

Regression Moderation Examples

Does the effect of distance to work on satisfaction with housing location depend on the amount of traffic congestion? A hypothetical example.

SPSS Example using Simple1 Macro (download from <http://web.pdx.edu/~newsom/macros.htm>)¹

A macro is a syntax file that contains an elaborate set of syntax commands and is stored on your computer. The file can then be referenced with syntax by instructing SPSS to go find the macro and run the commands. As usual, you will need to access your data file first by opening the data window in SPSS or using a GET FILE command. Once the variables are renamed, the macro will center the variables, calculate the interaction variable, run the regression, and test and plot the simple slopes.

```
* I recommend getting your data file with a get file command in your syntax. Whether you take this
recommendation or not, you should make sure there are no data files open (except the original data file you are
analyzing if you do not use get file).
get file='c:\jason\spsswin\mvclass\congestion.sav'.
```

```
* predictor variable must be renamed x, the moderator must be renamed z, and the outcome must be renamed y.
Macro centers predictor variables automatically
```

```
compute x=distance.
compute z=congest.
compute y=satis.
```

```
cd 'C:\jason\temp'. *set default directory to location you have permission to save files in.
* on mac you can usually specify the directory location like this /Users/newsom/spssfiles/.
```

```
INCLUDE 'c:\jason\spsswin\macros\simple1.sps'. *location where you have stored the macro file.
SIMPLE1 VARS=. *covariates can be listed before the period if desired.
```

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients		t	Sig.
		B	Std. Error	Beta	Std. Error		
1	(Constant)	6.799	.104			65.623	.000
	x	-.497	.048	-.540	.053	-10.266	.000
	z	-.042	.004	-.506	.049	-10.392	.000
	xz	-.012	.002	-.343	.052	-6.527	.000

a. Dependent Variable: y

Correct Standardized Slope for XZ
 -.344

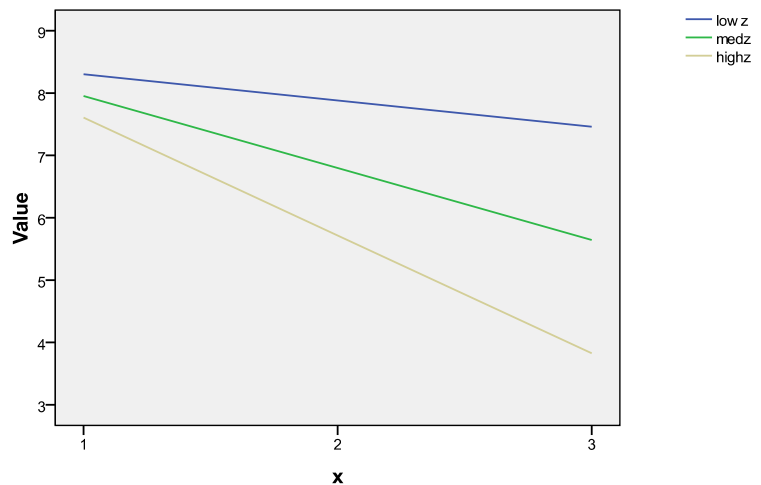
Unstandardized Simple Slopes for X
 b
 at low Z: -.18
 at mean Z: -.50
 at high Z: -.81

Standardized Simple Slopes for X
 beta
 at low Z: -.20
 at mean Z: -.54
 at high Z: -.88

Significance

	SE	t-value	p-value
at low Z:	.08	-2.25	.027
at mean Z:	.05	-10.27	.000
at high Z:	.05	-15.07	.000

Simple Slopes for Y on X at Values of Z



LOWZ = -1SD below mean, MEDZ=at mean, HIGHZ=+1SD above mean

¹ Errors may occur in using this macro (as in others too). These problems may be fixed by exiting SPSS and restarting. Otherwise, there are two potential reasons if the macro fails to work a) the cd statement does not contain a directory you have permission to write to on your computer, and b) it mostly runs but the last portion (text) of the output with the simple slope values and significance is not displayed. If the latter problem occurs, check to make sure that under **Edit -> Options -> Viewer** tab, under the **Item: log** make sure that **shown** is checked rather than hidden under **Contents are initially**. I have also experienced some issues related to Unicode mode that SPSS currently uses, so you might have to add `set Unicode=no.` to the first line of syntax. See the [trouble shooting page](#) for additional suggestions.

SPSS Process Macro

Alternatively, you can use the Process Macro, which was developed by Andrew Hayes (see Hayes, 2018)² and is very convenient for conducting a number of different types of regression analyses that involve moderation and mediation (discussed soon). For simple moderation models (*/model=1* is the simplest form), the process macro will automatically center the variables (if you use the */center=1* subcommand; note that if you do not use */moment=1* it will use the 16th, 50th, and 84th percentiles by default), compute the interaction term, run the regression model with the interaction term, and then test the simple slopes. It does not provide standardized coefficients for this type of model³ and it does not plot the simple slopes. It will provide plot points, which can be used for creating a scatterplot in SPSS of the simple slope groups. Download the macro from <https://haskayne.ucalgary.ca/CCRAM/resource-hub> and unzip the folder, and then save the SPSS syntax file somewhere on your computer in a location that is easy to refer to.

Then you can use the following syntax if you modify my location to match the location on your computer (for Mac, use just the forward slash then location, such as */Users/blah/blah/blah*). Make sure that the CD (short for change directory) command, which sets the working directory, specifies a location that you have access to for saving files. I used version 4.3 of the macro for this analysis.

```
*p r o c e s s macro.  
output close *.  
CD 'C:\jason\temp'.  
insert file='C:\Jason\SPSSWIN\macros\process.sps'.  
execute.
```

*note: order subcommand seems to matter, when using the *plot=1* subcommand for instance, it must go before the *moment=1* subcommand.

```
process y=satis  
  /x=distance  
  /w=congest  
  /model=1  
  /center=1  
  /plot=1  
  /moments=1.
```

Run MATRIX procedure:

```
***** PROCESS Procedure for SPSS Version 4.3.1 *****  
          Written by Andrew F. Hayes, Ph.D.      www.afhayes.com  
          Documentation available in Hayes (2022). www.guilford.com/p/hayes3
```

```
*****  
Model   : 1  
Y       : satis  
X       : distance  
W       : congest  
Sample Size: 100  
*****
```

OUTCOME VARIABLE:

satis

Model Summary

	R	R-sq	MSE	F	df1	df2	p
	.8796	.7737	1.0672	109.4174	3.0000	96.0000	.0000

² See <https://processmacro.org/index.html>. Documentation for the macro is only available in the Hayes book, Hayes, A.F. (2022) *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach, Third Edition*. Guilford Press. Note that I only illustrate the syntax approach to running the process macro, but you can also install the macro into SPSS and run it through windows. See <https://processmacro.org/uploads/3/4/8/5/34858384/dialoginstall.pdf> for more information on that approach.

³ Although Process does provide standardized values for mediation models starting with version 3.2. For moderation models, you can standardize the variable first and then run the macro, but a) do not use the significance tests from that output, b) be careful that the same N is used when standardizing as in the model, and c) the interaction value will be incorrect unless the interaction variable is formed from the standardized variables, although usually only slightly off.

Model	coeff	se	t	p	LLCI	ULCI
constant	6.7987	.1036	65.6231	.0000	6.5931	7.0044
distance	-.4970	.0484	-10.2665	.0000	-.5930	-.4009
congest	-.0421	.0041	-10.3918	.0000	-.0502	-.0341
Int_1	-.0123	.0019	-6.5268	.0000	-.0160	-.0086

Product terms key:

Int_1 : distance x congest

Test(s) of highest order unconditional interaction(s):

	R2-chng	F	df1	df2	p
X*W	.1004	42.5989	1.0000	96.0000	.0000

Focal predict: distance (X)
 Mod var: congest (W)

Conditional effects of the focal predictor at values of the moderator(s):

congest	Effect	se	t	p	LLCI	ULCI
-25.7188	-.1806	.0805	-2.2452	.0271	-.3403	-.0209
.0000	-.4970	.0484	-10.2665	.0000	-.5930	-.4009
25.7188	-.8133	.0540	-15.0733	.0000	-.9204	-.7062

Data for visualizing the conditional effect of the focal predictor:
 Paste text below into a SPSS syntax window and execute to produce plot.

```
DATA LIST FREE/
  distance  congest  satis  .
BEGIN DATA.
  -2.3241  -25.7188   8.3016
   .0000  -25.7188   7.8817
   2.3241  -25.7188   7.4619
  -2.3241   .0000   7.9537
   .0000   .0000   6.7987
   2.3241   .0000   5.6437
  -2.3241  25.7188   7.6059
   .0000  25.7188   5.7157
   2.3241  25.7188   3.8255
END DATA.
GRAPH/SCATTERPLOT=
  distance WITH  satis  BY  congest  .
```

***** ANALYSIS NOTES AND ERRORS *****

Level of confidence for all confidence intervals in output:
 95.0000

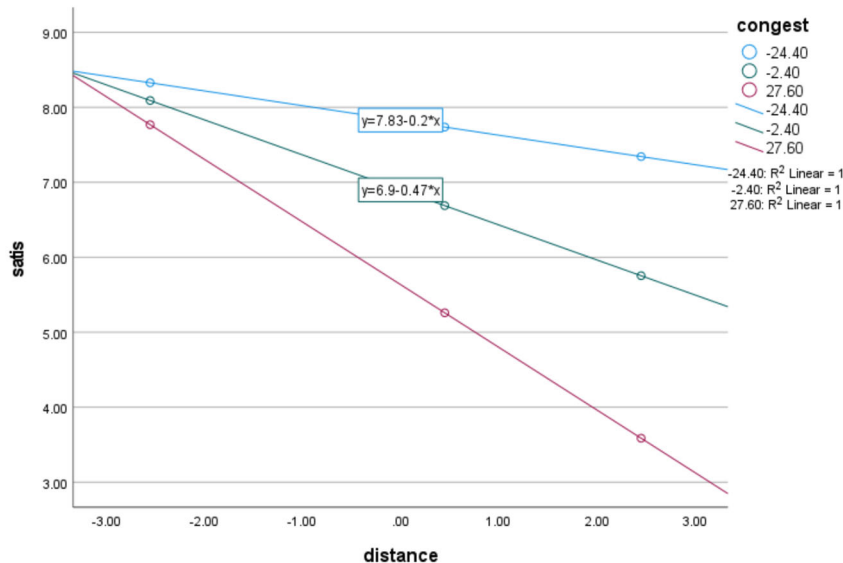
W values in conditional tables are the mean and +/- SD from the mean.

NOTE: The following variables were mean centered prior to analysis:

congest distance

----- END MATRIX -----

To obtain the following plot, I copied and pasted the plot syntax above, starting with the `data list` command through the end of the `graph` command into a new syntax file. Double click on graph in the output, choose **Options**, choose **Elements** and **Fit line at subgroups line from equation**. Click apply, exit from the **Elements** dialog.



R Example using *pequod* Package

The *pequod* package is simple to use, provides standardized estimates, simple slope tests, and a plot. It was recently removed from CRAN, so extra code is needed to install it from github using the *remotes* package.

```
> #clear active frame from previous analyses
> rm(d)
>
> library(haven)
> d = read_sav("c:/jason/spsswin/da2/congestion.sav")

> #pequod has been removed from CRAN - get it from github
> #install.packages("remotes")
> #library(remotes)
> #install_github("cran/pequod")
> #install.packages("pequod") #run on first use, packages car and ggplot2 are required also
> library(pequod)
>
> model<-lmres(satis~distance+congest+distance*congest, centered=c("distance", "congest"), data=d)
> summary(model)
Formula:
satis ~ distance + congest + distance.XX.congest
<environment: 0x0000000018b195b0>

Models
  R      RA2   Adj. RA2    F      df1 df2      p.value
Model 0.880 0.774  0.767 109.417  3.000  96 <0.0000000000000002

Residuals
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-2.5260 -0.6227  0.0620  0.0000  0.8918  1.5190

Coefficients
              Estimate StdErr  t.value   beta      p.value
(Intercept)   6.79871  0.10360  65.62307  < 0.00000000000000022
distance      -0.49696  0.04841 -10.26649 -0.5401 < 0.00000000000000022
congest       -0.04211  0.00405 -10.39180 -0.5064 < 0.00000000000000022
distance.XX.congest -0.01230  0.00188 -6.52678 -0.3426 < 0.00000000000000022

Collinearity
              VIF Tolerance
distance      1.1741  0.8517
congest       1.0076  0.9924
distance.XX.congest 1.1688  0.8556

>
> simples<-simpleSlope(model, pred="distance", mod1="congest")
> summary(simples)

** Estimated points of satis **

              Low distance (-1 SD) High distance (+1 SD)
Low congest (-1 SD)      8.3016      7.4619
High congest (+1 SD)     7.6059      3.8255

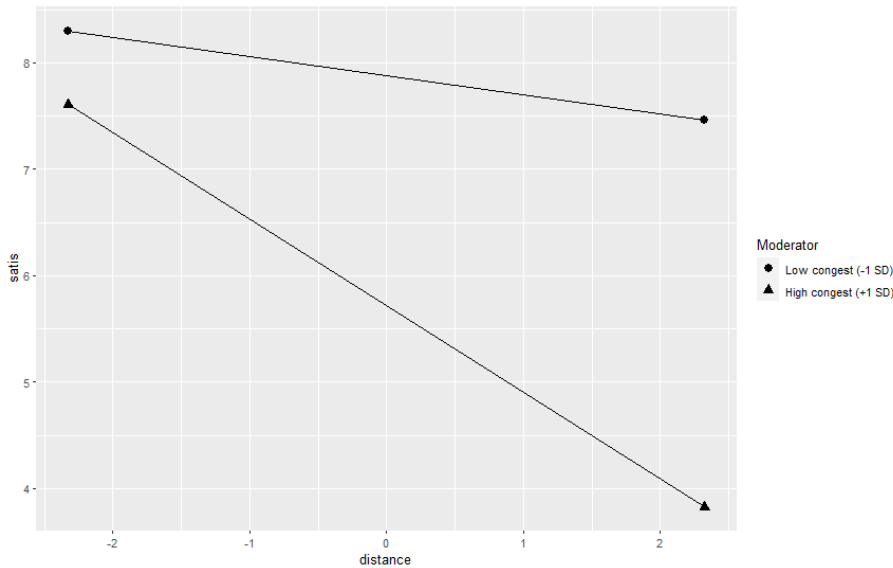
** Simple slopes analysis ( df= 96 ) **

              simple slope standard error t-value      p.value
Low congest (-1 SD)      -0.1806      0.0805   -2.25      0.027
High congest (+1 SD)     -0.8133      0.0540  -15.07 <0.0000000000000002
```

** Bauer & Curran 95% CI **

```

lower CI upper CI
congest -63.543 -27.21
>
> plot<-PlotsSlope(simples)
> print(plot)
    
```



R using the Process Macro⁴

The PROCESS macro is now available for R. Download from <https://haskayne.ucalgary.ca/CCRAM/resource-hub> (appears in the zip folder with SPSS and SAS versions of the macro).

After downloading, open the `process.R` macro code and run it in RStudio. This takes several minutes to run. Go ahead and close the tab for the macro.

Then in a new or existing script read in your data as usual, and, below that, include the process code. Here is my code for the above example

```

> rm(d) #clear active frame from previous analyses
> if(!is.null(dev.list())) dev.off(dev.list()[ "RStudiogD"]) #clear plots
> library(haven)
> d = read_sav("c:/jason/spsswin/da2/congestion.sav")
>
> #use process macro
> #make sure you have opened process macro and have run it in Rstudio first
> process(data=d,y="satis",x="distance",w="congest",model=1)
    
```

Note no plots or standardized coefficients are provided with this method. For standardized coefficients, they can be computed on a calculator or programming in R with the usual formula.

$$\beta_{yx} = B_{yx} \left(\frac{sd_x}{sd_y} \right)$$

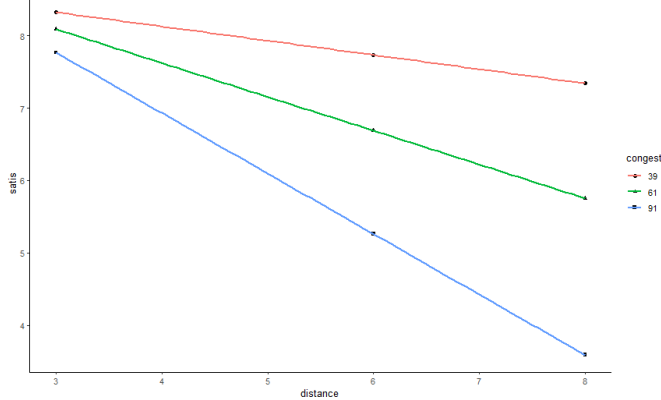
For plotting, perhaps the simplest method is to paste the plot information into a new script file. Then save this as text and read the text file into a new script file for plotting. I provide my code for reading in the data and plotting separate lines.

```

> d <- read.table(file=paste('C:/Jason/R/mvclass/congestion data.txt',sep="/"), header=TRUE)
> #plotting required moderator data to be factor
> d$congest <- as.factor(d$congest)
> ggplt <- ggplot(d,aes(x=distance,y=satis,shape=congest))+
+   geom_point()+
    
```

⁴ See also this video for a clear explanation of how to use the process macro in R (but with a mediation example) <https://www.youtube.com/watch?v=ld6zjPZoKvo> from Mike Crowson.

```
+ theme_classic()
>
> ggplot+geom_smooth(method=lm, se=FALSE, fullrange=TRUE,
+ aes(color=congest))
```



Write-up

A multiple regression model was tested to investigate whether the association between distance to work and satisfaction with housing location depends on the amount of traffic congestion in the area. After centering distance to work and traffic congestion and computing the distance-by-congestion interaction term (Aiken & West, 1991), the two predictors and the interaction were entered into a simultaneous regression model. Results indicated that greater traffic congestion ($B = -.042$, $SE = .004$, $\beta = -.506$, $p < .001$) and farther distance to work ($B = -.497$, $SE = .048$, $\beta = -.540$, $p < .001$) were both associated with lower satisfaction with housing location. The interaction between distance to work and traffic congestion was also significant ($B = -.012$, $SE = .002$, $\beta = -.344$, $p < .001$), suggesting that the effect of distance on satisfaction depended on the level of traffic congestion. Together, the variables accounted for approximately 77% of the variance in satisfaction, $R^2 = .77$, $F(3,96) = 109.42$, $p < .001$. Simple slopes for the association between distance and satisfaction were tested for low (-1 SD below the mean), moderate (mean), and high (+1 SD above the mean) levels of traffic congestion. Each of the simple slope tests revealed a significant negative association between distance to work and satisfaction with housing location, but the distance to work was more strongly related to satisfaction for high levels of traffic congestion ($B = -.81$, $SE = .05$, $\beta = -.88$, $p < .001$) than for moderate ($B = -.50$, $SE = .05$, $\beta = -.54$, $p < .001$) or lower levels ($B = -.18$, $SE = .08$, $\beta = -.20$, $p = .03$) of traffic congestion. Figure 1 plots the simple slopes for the interaction.