TEXAS TECH UNIVERSITY

STATS CAMP

statscamp.org

Lecture 11: Inference in CPA

Kyle M. Lang

Institute for Measurement, Methodology, Analysis & Policy
Texas Tech University
Lubbock, TX

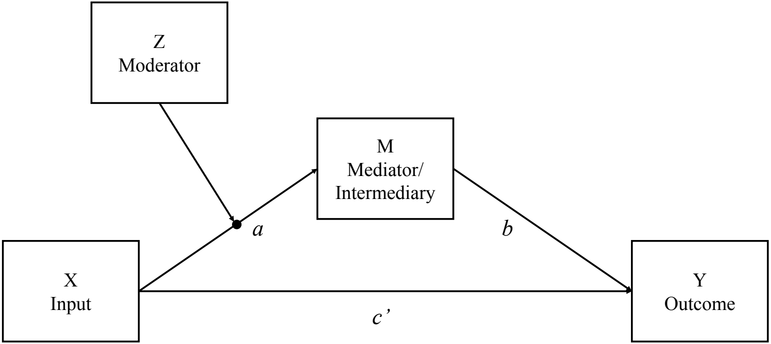
2016 Stats Camp

Outline

- Index of moderated mediation
- Examples of inference in conditional process models
- Inference when multiple paths are moderated
- General steps of inference in conditional process analysis

From Last Time

Last time, we saw how to calculate the conditional indirect effects in a model such as the following:



From Last Time

The preceding diagram corresponds to the following equations:

$$Y = i_1 + bM + c'X + e_Y \tag{1}$$

$$M = i_2 + a_1X + a_2Z + a_3XZ + e_M \tag{2}$$

The indirect effect must be interpreted as conditional on Z due to the a path being moderated by Z.

We rearrange Equation 2 to get:

$$M = i_2 + a_2Z + (a_1 + a_3Z)\,X + e_M,$$

and the conditional indirect effect is defined by:

$$IE = (a_1 + a_3Z)\,b$$

Index of Moderated Mediation

Probing conditional indirect effects is all well and good, but we'd like a single index to test the overall hypothesis of moderated mediation.

- Enter the *Index of Moderated Mediation* (IMM) introduced by Hayes (2015).
- The IMM quantifies the linear effect of the moderator on the indirect effect.
- When IMM is different from zero, we know that the indirect effect is moderated, generally.

Index of Moderated Mediation

Applying some basic algebra to the preceding conditional indirect effect formula produces:

$$(a_1 + a_3 Z)\, b = a_1 b + a_3 b Z,$$

which is a linear function describing the effect of Z on the indirect effect

- $a_1 b$ is the intercept term
- $a_3 b$ is the slope linking Z to the indirect effect

The $a_3 b$ term is the IMM.

- We test $a_3 b \neq 0$ to infer moderated mediation.
- Normal theory tests are possible, but we want to use bootstrapping.

Example

```
## Prep stuff:
library(lavaan)
nBoot <- 1000
dat1 <- readRDS("../data/lecture11Data.rds")
##
## Create product term:
dat1$xz <- dat1$x * dat1$z
##
## Specify model:
mod1 <- "
y ~ cp*x + b*m1
m1 ~ a1*x + a2*z + a3*xz

imm := a3*b
"
##
## Fit model:
out1 <- sem(mod1, data = dat1, se = "boot", boot = nBoot)
summary(out1)
```

Example

```
lavaan (0.5-20) converged normally after  16 iterations

  Number of observations                           100

  Estimator                                         ML
  Minimum Function Test Statistic                2.380
  Degrees of freedom                                 2
  P-value (Chi-square)                           0.304

Parameter Estimates:

  Information                                 Observed
  Standard Errors                            Bootstrap
  Number of requested bootstrap draws             1000
  Number of successful bootstrap draws            1000

Regressions:
                   Estimate  Std.Err  Z-value  P(>|z|)
  y ~
    x         (cp)    0.082    0.259    0.318    0.751
    m1         (b)    1.390    0.215    6.465    0.000
  m1 ~
```

Example

```
    x          (a1)    0.729    0.091    8.014    0.000
    z          (a2)    0.641    0.089    7.183    0.000
    xz         (a3)    0.451    0.097    4.643    0.000

Variances:
                   Estimate  Std.Err  Z-value  P(>|z|)
    y              5.102    0.676    7.550    0.000
    m1             0.669    0.089    7.501    0.000

Defined Parameters:
                   Estimate  Std.Err  Z-value  P(>|z|)
    imm            0.628    0.169    3.706    0.000
```

Example

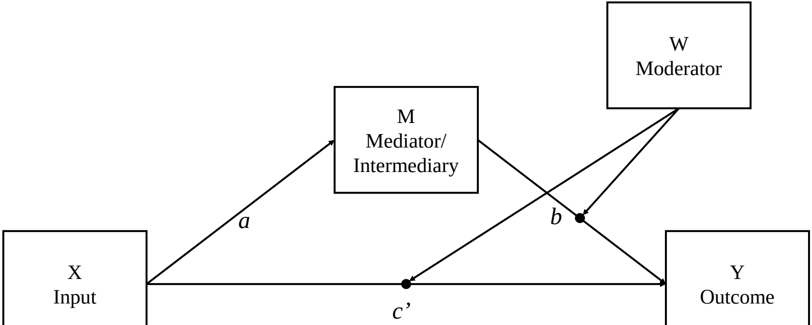
```
parameterEstimates(out1,
                    boot = "bca.simple")[-c(6 : 13), -c(1 :
                        3)]
```

```
   label   est    se     z pvalue ci.lower ci.upper
1     cp 0.082 0.259 0.318  0.751   -0.408    0.641
2      b 1.390 0.215 6.465  0.000    0.939    1.786
3     a1 0.729 0.091 8.014  0.000    0.539    0.908
4     a2 0.641 0.089 7.183  0.000    0.469    0.818
5     a3 0.451 0.097 4.643  0.000    0.223    0.629
14   imm 0.628 0.169 3.706  0.000    0.296    0.980
```

A Little Different

Moderation of the direct effect doesn't change the calculation of the IMM.

Consider the following model:



A Little Different

This analytic diagram implies the following equations:

$$Y = i_1 + c_1' X + b_1 M + b_2 W + c_2' XW + b_3 MW + e_Y \tag{3}$$

$$M = i_2 + aX + e_M \tag{4}$$

The direct effect is conditional:

$$DE = c_1' + c_2' W$$

The conditional indirect effect is defined by:

$$IE = a\,(b_1 + b_3 W) = ab_1 + ab_3 W,$$

which implies $IMM = ab_3$

Example

```
##
## Create product term:
dat1$m1w <- dat1$m1 * dat1$w
dat1$xw <- dat1$x * dat1$w
##
## Specify model:
mod2 <- "
y ~ cp1*x + cp2*xw + b1*m1 + b2*w + b3*m1w
m1 ~ a*x

imm := a*b3
"
##
## Fit model:
out2 <- sem(mod2, data = dat1, se = "boot", boot = nBoot)
summary(out2)
```

Example

```
lavaan (0.5-20) converged normally after  23 iterations

  Number of observations                           100

  Estimator                                         ML
  Minimum Function Test Statistic                5.078
  Degrees of freedom                                 3
  P-value (Chi-square)                           0.166

Parameter Estimates:

  Information                                 Observed
  Standard Errors                            Bootstrap
  Number of requested bootstrap draws             1000
  Number of successful bootstrap draws            1000

Regressions:
                   Estimate  Std.Err  Z-value  P(>|z|)
  y ~
    x         (cp1)  -0.103    0.190   -0.543    0.587
    xw        (cp2)   1.099    0.204    5.380    0.000
    m1         (b1)   1.615    0.167    9.649    0.000
```

Example

```
    w           (b2)    0.381    0.173    2.206    0.027
    m1w         (b3)    0.571    0.173    3.297    0.001
  m1 ~
    x            (a)    0.741    0.131    5.638    0.000

Variances:
                     Estimate  Std.Err  Z-value  P(>|z|)
    y                   2.950    0.344    8.575    0.000
    m1                  1.182    0.145    8.161    0.000

Defined Parameters:
                     Estimate  Std.Err  Z-value  P(>|z|)
    imm                 0.424    0.153    2.773    0.006
```

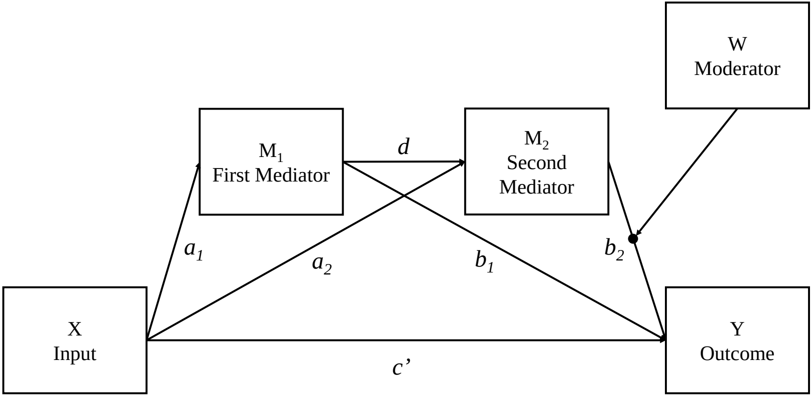
Example

```
parameterEstimates(out2,
                   boot = "bca.simple")[-c(7 : 18), -c(1 :
                      3)]
```

```
   label    est    se      z pvalue ci.lower ci.upper
1    cp1 -0.103 0.190 -0.543  0.587   -0.473    0.259
2    cp2  1.099 0.204  5.380  0.000    0.639    1.472
3     b1  1.615 0.167  9.649  0.000    1.270    1.912
4     b2  0.381 0.173  2.206  0.027    0.034    0.720
5     b3  0.571 0.173  3.297  0.001    0.220    0.921
6      a  0.741 0.131  5.638  0.000    0.471    0.984
19   imm  0.424 0.153  2.773  0.006    0.174    0.781
```

There's no reason that our baseline mediation model needs to be a simple, three-variable model.

Consider the following model:



Getting More Complicated

The preceding implies the following equations:

$$Y = i_1 + c'X + b_1 M_1 + b_2 M_2 + b_3 W + b_4 M_2 W + e_Y \tag{5}$$

$$M_2 = i_2 + a_2 X + dM_1 + e_{M2} \tag{6}$$

$$M_1 = i_3 + a_1 X + e_{M1} \tag{7}$$

We now have several specific indirect effects:

$$\begin{aligned} IE_1 &= a_1 b_1 \\ IE_2 &= a_2 (b_2 + b_4 W) = a_2 b_2 + a_2 b_4 W \\ IE_3 &= a_1 d (b_2 + b_4 W) = a_1 db_1 + a_1 db_4 W, \end{aligned}$$

which imply $IMM_2 = a_2 b_4$ and $IMM_3 = a_1 db_4$.

Example

```
## Create product term:
dat1$m2w <- dat1$m2 * dat1$w
##
## Specify model:
mod4 <- "
y ~ cp*x + b1*m1 + b2*m2 + b3*w + b4*m2w
m2 ~ a2*x + d*m1
m1 ~ a1*x

ab1 := a1*b1
imm2 := a2*b4
fullImm := a1*d*b4
"
##
## Fit model:
out4 <- sem(mod4, data = dat1, se = "boot", boot = nBoot)
summary(out4)
```

Example

```
lavaan (0.5-20) converged normally after  24 iterations

  Number of observations                           100

  Estimator                                         ML
  Minimum Function Test Statistic                1.655
  Degrees of freedom                                 4
  P-value (Chi-square)                           0.799

Parameter Estimates:

  Information                                 Observed
  Standard Errors                            Bootstrap
  Number of requested bootstrap draws             1000
  Number of successful bootstrap draws            1000

Regressions:
                   Estimate  Std.Err  Z-value  P(>|z|)
  y ~
    x         (cp)   -0.082    0.215   -0.383    0.701
    m1        (b1)    0.462    0.218    2.114    0.035
    m2        (b2)    0.822    0.165    4.998    0.000
```

Example

```
    w          (b3)    0.656    0.196    3.355    0.001
    m2w        (b4)    0.627    0.113    5.545    0.000
  m2 ~
    x          (a2)    0.035    0.130    0.271    0.787
    m1          (d)    1.262    0.097   13.079    0.000
  m1 ~
    x          (a1)    0.741    0.135    5.503    0.000

Variances:
                    Estimate  Std.Err  Z-value  P(>|z|)
    y                  2.784    0.373    7.462    0.000
    m2                 1.051    0.141    7.466    0.000
    m1                 1.182    0.142    8.345    0.000

Defined Parameters:
                    Estimate  Std.Err  Z-value  P(>|z|)
    ab1                0.342    0.178    1.924    0.054
    imm2               0.022    0.083    0.266    0.790
    fullImm            0.587    0.170    3.449    0.001
```

Example

```
parameterEstimates(out4,
                   boot = "bca.simple")[-c(9 : 17), -c(1 :
                     3)]
```

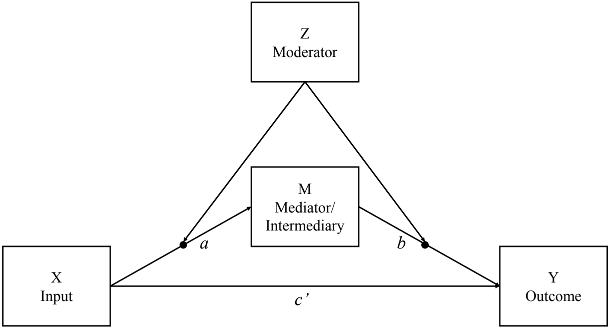
```
     label    est    se      z pvalue ci.lower ci.upper
1       cp -0.082 0.215 -0.383  0.701   -0.515    0.327
2       b1  0.462 0.218  2.114  0.035    0.034    0.931
3       b2  0.822 0.165  4.998  0.000    0.514    1.158
4       b3  0.656 0.196  3.355  0.001    0.222    1.032
5       b4  0.627 0.113  5.545  0.000    0.419    0.862
6       a2  0.035 0.130  0.271  0.787   -0.217    0.296
7        d  1.262 0.097 13.079  0.000    1.071    1.450
8       a1  0.741 0.135  5.503  0.000    0.483    0.995
18     ab1  0.342 0.178  1.924  0.054    0.040    0.745
19    imm2  0.022 0.083  0.266  0.790   -0.131    0.189
20 fullImm  0.587 0.170  3.449  0.001    0.332    1.012
```

Limitations of IMM

The IMM is only well-defined for *linear* relations between the moderator and the indirect effect.

- Indirect effects with multiple constituent paths moderated imply non-linear relations between the moderators and the indirect effect.

Consider this model:



Limitations of IMM

The preceding model implies the following equations:

$$Y = i_1 + c'X + b_1 M + b_2 Z + b_2 MZ + e_Y \tag{8}$$
$$M = i_2 + a_1 X + a_2 Z + a_3 XZ + e_{M1}, \tag{9}$$

so we have the following conditional indirect effect:

$$\begin{aligned} IE &= (a_1 + a_3 Z)(b_1 + b_3 Z) \\ &= a_1 b_1 + (a_1 b_3 + a_3 b_1) Z + a_3 b_3 Z^2 \end{aligned}$$

which represents a quadratic linkage between the moderator and indirect effect.

With no single term describing the association of the moderator and the indirect effect, the Hayes (2015) IMM isn't directly applicable.

Alternative Strategy

When we have multiple moderated paths, we need to employ an alternative method discussed by Edwards and Lambert (2007) and Wang and Preacher (2015), among others:

1. Fit our moderated mediation model, as usual
2. Compute the conditional indirect effects at two values of the moderator
3. Test for significant differences between these two conditional indirect effects.

Finding significant differences by this method suggests overall moderation.

- The converse does not hold
- A lack of significance does not imply no moderation
- Pairwise comparisons of conditional indirect effects are dependent on the values chosen for the moderator values

Example

```
## Compute product term:
dat2$mz <- dat2$m * dat2$z
##
## Specify model:
mod5 <- "
y ~ cp*x + b1*m + b2*z + b3*mz
m ~ a1*x + a2*z + a3*xz

fullIE1 := (a1 + a3 * (-1.244962)) *
           (b1 + b3 * (-1.244962)) * b2
fullIE2 := (a1 + a3 * 1.369550) *
           (b1 + b3 * 1.369550) * b2
mmTest := fullIE1 - fullIE2
"
##
## Fit model:
out5 <- sem(mod5, data = dat2, se = "boot", boot = nBoot)
summary(out5)
```

Example

```
lavaan (0.5-20) converged normally after  18 iterations

  Number of observations                           100

  Estimator                                         ML
  Minimum Function Test Statistic                1.230
  Degrees of freedom                                 2
  P-value (Chi-square)                           0.541

Parameter Estimates:

  Information                                 Observed
  Standard Errors                            Bootstrap
  Number of requested bootstrap draws             1000
  Number of successful bootstrap draws            1000

Regressions:
                   Estimate  Std.Err  Z-value  P(>|z|)
  y ~
    x         (cp)    0.258    0.114    2.263    0.024
    m         (b1)    0.792    0.093    8.549    0.000
    z         (b2)    0.784    0.110    7.111    0.000
```

Example

```
    mz          (b3)     0.413     0.063     6.550     0.000
  m ~
    x           (a1)     0.729     0.090     8.115     0.000
    z           (a2)     0.641     0.084     7.631     0.000
    xz          (a3)     0.451     0.098     4.626     0.000

Variances:
                       Estimate   Std.Err   Z-value   P(>|z|)
    y                    0.593     0.080     7.385     0.000
    m                    0.669     0.086     7.805     0.000

Defined Parameters:
                       Estimate   Std.Err   Z-value   P(>|z|)
    fullIE1              0.036     0.044     0.819     0.413
    fullIE2              1.434     0.208     6.880     0.000
    mmTest              -1.398     0.225    -6.218     0.000
```

Example

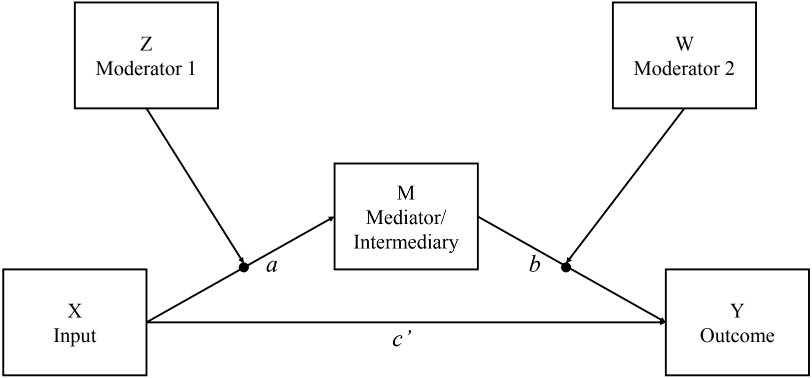
```
parameterEstimates(out5,
                   boot = "bca.simple")[-c(8 : 19), -c(1 :
                       3)]
```

```
      label    est    se      z pvalue ci.lower ci.upper
1        cp  0.258 0.114  2.263  0.024    0.006    0.461
2        b1  0.792 0.093  8.549  0.000    0.613    0.991
3        b2  0.784 0.110  7.111  0.000    0.570    1.019
4        b3  0.413 0.063  6.550  0.000    0.302    0.550
5        a1  0.729 0.090  8.115  0.000    0.552    0.902
6        a2  0.641 0.084  7.631  0.000    0.476    0.809
7        a3  0.451 0.098  4.626  0.000    0.246    0.632
20 fullIE1  0.036 0.044  0.819  0.413   -0.026    0.162
21 fullIE2  1.434 0.208  6.880  0.000    1.051    1.879
22  mmTest -1.398 0.225 -6.218  0.000   -1.866   -1.004
```

Limitations of IMM

The same issue arises when the indirect effects has multiple paths moderated by different variables.

Consider this model:



Limitations of IMM

The preceding model implies the following equations:

$$Y = i_1 + c'X + b_1 M + b_2 W + b_2 M_2 W + e_Y \tag{10}$$

$$M = i_2 + a_1 X + a_2 Z + a_3 XZ + e_M, \tag{11}$$

so we have the following conditional indirect effect:

$$\begin{aligned} IE &= (a_1 + a_3 Z)(b_1 + b_3 W) \\ &= a_1 b_1 + a_1 b_3 W + a_3 b_1 Z + a_3 b_3 ZW \end{aligned}$$

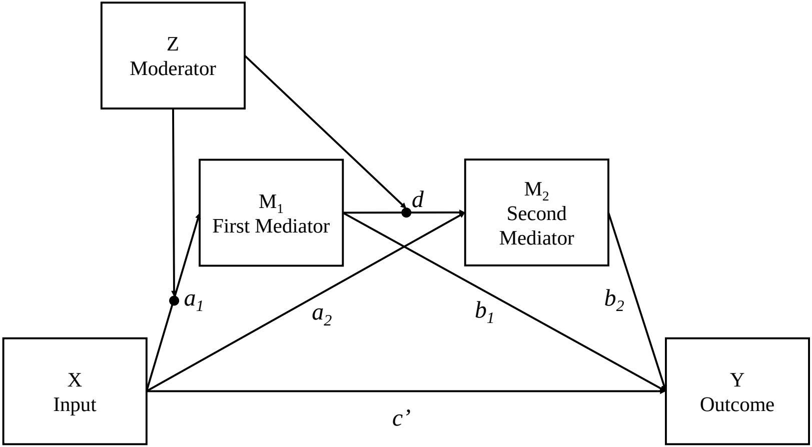
which represents an interactive linkage between the moderator and indirect effect.

We still have no single term describing the association of the moderator and the indirect effect, so the Hayes (2015) IMM still isn't directly applicable.

More Complicated Model

Okay, let's put this all together into a pretty complicated conditional process model:

Consider this model:



More Complicated Model

The preceding model implies the following equations:

$$Y = i_1 + c'X + b_1 M + b_2 M_2 + e_Y \tag{12}$$

$$M_2 = i_2 + a_2 X + d_1 M_1 + d_2 Z + d_3 M_1 Z + e_{M2} \tag{13}$$

$$M_1 = i_3 + a_1 X + a_2 Z + a_3 XZ, \tag{14}$$

so we have the following conditional indirect effects:

$$IE_1 = (a_1 + a_3 Z)\, b_1$$

$$IE_2 = a_2 b_2$$

$$IE_3 = (a_1 + a_3 Z)(d_1 + d_3 Z)\, b_2$$

We can employ multiple inferential strategies:

1. IE_2 is not moderated, so we can make direct inferences
2. IE_1 only contains one moderated path, so we can make inferences via $IMM_1 = a_3 b1$
3. IE_3 contains two moderated paths, so we need to use the pairwise comparison approach.

Example

```r
## Prep stuff:
dat1$m1z <- dat1$m1 * dat1$z
quantile(dat1$z, c(0.05, 0.95))
```

```
       5%        95%
-1.244962  1.369550
```

```r
##
## Specify the model:
mod6 <- "
y ~ cp*x + b1*m1 + b2*m2
m2 ~ a2*x + d1*m1 + d2*z + d3*m1z
m1 ~ a1*x + a2*z + a3*xz

imm1 := a3*b1
ab2 := a2*b2
fullIE1 := (a1 + a3 * (-1.244962)) *
           (d1 + d3 * (-1.244962)) * b2
fullIE2 := (a1 + a3 * 1.369550) *
           (d1 + d3 * 1.369550) * b2
mmTest := fullIE1 - fullIE2
"
```

Example

```r
## Fit the model:
out6 ← sem(mod6, data = dat1, se = "boot", boot = nBoot)
summary(out6)
```

```
lavaan (0.5-20) converged normally after  23 iterations

  Number of observations                           100

  Estimator                                         ML
  Minimum Function Test Statistic                8.849
  Degrees of freedom                                 6
  P-value (Chi-square)                           0.182

Parameter Estimates:

  Information                                 Observed
  Standard Errors                            Bootstrap
  Number of requested bootstrap draws             1000
  Number of successful bootstrap draws            1000

Regressions:
                   Estimate  Std.Err  Z-value  P(>|z|)
```

Example

```
  y ~
    x          (cp)    0.055    0.232    0.238    0.811
    m1         (b1)    0.427    0.252    1.693    0.090
    m2         (b2)    0.763    0.204    3.735    0.000
  m2 ~
    x          (a2)    0.473    0.073    6.480    0.000
    m1         (d1)    0.689    0.090    7.652    0.000
    z          (d2)    0.863    0.110    7.843    0.000
    m1z        (d3)    0.400    0.065    6.139    0.000
  m1 ~
    x          (a1)    0.720    0.089    8.110    0.000
    z          (a2)    0.473    0.073    6.480    0.000
    xz         (a3)    0.476    0.098    4.871    0.000

Variances:
                    Estimate  Std.Err  Z-value  P(>|z|)
    y                  4.490    0.704    6.375    0.000
    m2                 0.617    0.088    7.011    0.000
    m1                 0.691    0.095    7.302    0.000

Defined Parameters:
                    Estimate  Std.Err  Z-value  P(>|z|)
    imm1               0.203    0.128    1.580    0.114
```

Example

ab2	0.361	0.108	3.342	0.001
fullIE1	0.019	0.034	0.551	0.582
fullIE2	1.295	0.387	3.344	0.001
mmTest	-1.276	0.388	-3.291	0.001

Example

```
parameterEstimates(out6,
                   boot = "bca.simple")[-c(11 : 23), -c(1 :
                         3)]
```

	label	est	se	z	pvalue	ci.lower	ci.upper
1	cp	0.055	0.232	0.238	0.811	-0.431	0.474
2	b1	0.427	0.252	1.693	0.090	-0.068	0.929
3	b2	0.763	0.204	3.735	0.000	0.334	1.151
4	a2	0.473	0.073	6.480	0.000	0.329	0.625
5	d1	0.689	0.090	7.652	0.000	0.509	0.870
6	d2	0.863	0.110	7.843	0.000	0.670	1.116
7	d3	0.400	0.065	6.139	0.000	0.257	0.528
8	a1	0.720	0.089	8.110	0.000	0.521	0.883
9	a2	0.473	0.073	6.480	0.000	0.329	0.625
10	a3	0.476	0.098	4.871	0.000	0.296	0.674
24	imm1	0.203	0.128	1.580	0.114	-0.013	0.518
25	ab2	0.361	0.108	3.342	0.001	0.164	0.596
26	fullIE1	0.019	0.034	0.551	0.582	-0.020	0.134
27	fullIE2	1.295	0.387	3.344	0.001	0.562	2.102
28	mmTest	-1.276	0.388	-3.291	0.001	-2.082	-0.547

General Steps for Conditional Process Analysis

At this point, we can lay out a few general steps that should help structure any conditional process analysis:

1. Draw a conceptual diagram representing your hypothesized process
2. Translate that conceptual diagram into an analytic diagram
3. Translate the analytic diagram into equations
4. Solve for all of your conditional indirect effects
5. Group your indirect effects into three categories:
   - Not Moderated
   - Linearly Moderated
   - Non-Linearly Moderated

6. For linearly moderated conditional indirect effects, derive the appropriate IMMs

7. For non-linearly moderated conditional indirect effects, choose conditional values of the moderator(s) at which to test for differences in the conditional direct effects

8. Use the bootstrapping capabilities of path modeling software to fit the model implied by the equations from Step 3 and test the hypotheses implied by Steps 4 – 7.

References

Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: a general analytical framework using moderated path analysis. *Psychological methods*, *12*(1), 1–22.

Hayes, A. F. (2015). An index and test of linear moderated mediation. *Multivariate Behavioral Research*, *50*(1), 1–22.

Wang, L., & Preacher, K. J. (2015). Moderated mediation analysis using bayesian methods. *Structural Equation Modeling: A Multidisciplinary Journal*, *22*(2), 249–263.