1

Traditional Omnibus Mediation Tests for Communication Example

This document presents Mplus output and approaches to evaluating omnibus mediation effects through mediational chains when the primary outcome variable is binary. I use the parent communication example from Chapter 12 and I assume you are familiar with it. I also assume you have read the document on omnibus effects for Chapter 11 on the resources tab (titled "Traditional Omnibus Mediation Tests for Social Phobia Example"). By an omnibus effect, I mean estimating the effect of the treatment on the outcome through a specific mediator. For example, what is the effect of the treatment condition on parental communication through the mediational chain of perceived advantages.

Most approaches to omnibus testing seek to decompose the overall total effect of the program on the outcome into that which is due to each separate mediational chain. Often the intent is to use this information to identify the relative importance or relative contribution of the different mediators as influencers of the outcome. I argue in Chapter 17 that this is not optimal practice for purposes of program evaluation and I provide what I think are better approaches. In this primer, I first describe approaches to omnibus mediation tests with binary outcomes using LISEM and then I describe approaches using FISEM.

APPROACHES BASED ON LISEM

Consider the simplified model in Figure 1 where my focus is on estimating the omnibus mediating effect for perceived advantages of communication (`PA2`) relative to the effect of the treatment (`TREAT`) on communication (`COM3`). In the Figure, I have ignored the other formal mediators in the model, perceived knowledge and perceived embarrassment (`PK2` and `PE2`). Path *c* is the direct effect of `TREAT` on `COM3` independent of the other mediators in the model but in Figure 1, there is only one mediator. This means that path *c* includes the effect variance it shares with all omitted mediators considered as a collective, including `PK2` and `PE2`. Perceived advantages is not reflected in path *c* because it is statistically held constant via path *b*.

The methods I describe estimate the effect of `TREAT` on `COM3` through paths *a* and *b*, holding constant variables captured by path c, which includes the omitted mediators. The methods do so for each formal mediator in the model considered separately. The methods only provide unbiased estimates of such indirect effects if there are no correlated disturbances and no causal relationships among the mediators in the larger model. These conditions are satisfied for the communication example, but they are restrictive more generally. If they are violated, then one must expand the LISEM model beyond the two equations I discuss so that the omitted dynamics can be modeled. If the violations are minor in magnitude, one might safely ignore them if you judge the bias they create to be inconsequential.

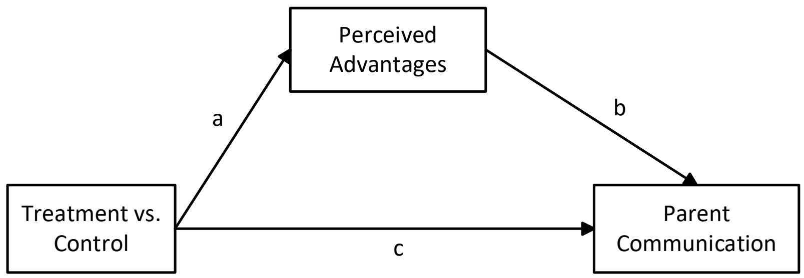

FIGURE 1. Simplified mediation model

The Modified Linear Probability Model using LISEM

The approach uses Mplus with MLR to estimate two equations, hence it is not single equation LISEM approach. It might be used if the modeling context does not permit the application of FISEM due to model complexity relative to sample size or some other reason. The two equations I estimate are Equations 12.5 and 12.8 from the full model:

$$PA2 = a_1 + p_1\ TREAT + b_1\ BS1 + b_2\ CQ1 + b_3\ PA1 + d_1 \quad [12.5]$$

$$COM3 = a_4 + p_4\ PA2 + p_5\ PK2 + p_6\ PE2 + p_7\ TREAT + b_{10}\ BS1 + b_{11}\ CQ1 \quad [12.8]$$

The Mplus syntax is in Table 1:

Table 1: Syntax for indirect effects using the MLPM with LISEM

```
1. TITLE: MLPM Analysis of communication  ;
2. DATA: FILE IS c:\mplus\communication.dat ;
3. DEFINE:
4.   CENTER PA1 CQ1 BS1 (GRANDMEAN) ;
5. VARIABLE:
6.   NAMES ARE ID COM3 PA2 PK2 PE2 CQ1 PA1 PK1 PE1 TREAT BS1 ;
7.   USEVARIABLES ARE COM3 PA2 CQ1 PA1 TREAT BS1 ;
```

```
  MISSING ARE ALL (-9999) ;
ANALYSIS:
  ESTIMATOR=MLR ;
MODEL:
  PA2 ON BS1 CQ1 TREAT PA1 ;
  COM3 ON PA2 TREAT BS1 CQ1 ;
MODEL INDIRECT:
  COM3 IND TREAT ;
OUTPUT: SAMP STANDARDIZED(STDYX) MOD(ALL 4) RESIDUAL
CINTERVAL TECH4 ;
```

All of it should be familiar to you. The key output is from the `TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS`:

```
                                                       Two-Tailed
                    Estimate       S.E.    Est./S.E.    P-Value

Effects from TREAT to COM3

  Total                0.190      0.025      7.591      0.000
  Total indirect       0.097      0.025      3.828      0.000

  Specific indirect 1
    COM3
    PA2
    TREAT              0.097      0.025      3.828      0.000

  Direct
    COM3
    TREAT              0.093      0.036      2.587      0.010
```

The primary effect of interest is listed under `Specific indirect 1`. The variable in the last row underneath this heading is assumed to influence the variable in the next to last row which, in turn, influences the variable in the top row. So, this chain refers to `TREAT` → `PA2` → `LSP3`. We work our way from top to bottom to isolate the mediational chain. The omnibus coefficient *for this chain* is 0.097 ±0.05 ($z = 3.83$, $p < 0.05$). Because `TREAT` is a dummy variable with dummy coding, the coefficient is the mean social phobia difference between the intervention condition minus the control condition through the mediational chain focused on perceived advantages. The value and significance test maps closely onto the result found in the FISEM analysis using the MLPM.

The `Direct` effect in the output is the combined mediation effects of all omitted mediators. It was 0.093 ±0.07 ($z = 2.59$, $p < 0.05$). It includes `PK2` and `PE2` within it.

Probit Modeling and the Causal Mediation Approach using LISEM

As noted in previous chapters, the causal mediation framework (CMF) is challenged by the presence of multiple mediators, so the approach outlined above that uses the logic model for Figure 1 lends itself well to CMF application. However, it brings with it the restrictive assumptions noted above and a few more besides those, as I elaborate shortly. Mplus offers the CMF as an option for single mediator models with binary/continuous mediators and/or binary outcomes. Table 2 presents the syntax for a probit-based analysis using the CMF as focused on `PA2`:

Table 2: Syntax for causal mediation analysis using probit-based LISEM

```
1. TITLE: Probit Analysis of communication  ;
2. DATA: FILE IS c:\mplus\communication.dat ;
3. DEFINE:
4.   CENTER PA1 CQ1 BS1 (GRANDMEAN) ;
5. VARIABLE:
6.   NAMES ARE ID COM3 PA2 PK2 PE2 CQ1 PA1 PK1 PE1 TREAT BS1 ;
7.   USEVARIABLES ARE COM3 PA2 CQ1 PA1 TREAT BS1 ;
8.   MISSING ARE ALL (-9999) ;
9.   CATEGORICAL COM3 ;
10. ANALYSIS:
11.  ESTIMATOR=ML ; LINK=PROBIT ;
12. MODEL:
13.  PA2 ON BS1 CQ1 TREAT PA1 ;
14.  COM3 ON PA2 TREAT BS1 CQ1 ;
15. MODEL INDIRECT:
16.  COM3 IND PA2 TREAT ;
17. OUTPUT: SAMP STANDARDIZED(STDYX) RESIDUAL
18.  CINTERVAL TECH4 ;
```

The primary differences from Table 1 are (a) specifying `COM3` as categorical on Line 9, (b) changing the estimator on Line 11, (c) adding the mediator `PA2` on Line 16, which activates a CMF analysis when the outcome is categorical/ordinal binary, and (d) removing the modification indices request from the output line (Mplus does not allow it with the CMF).

When applying the CMF strategy inMplus, the default is to hold the covariates constant at a value of 0. If a value of zero is meaningless or impossible for the metrics, then the covariate needs to be re-scaled so that a score of 0 is meaningful. For example, one might mean center the covariates; or, if a covariate is, say income, and one wants to hold it constant at a value of \$30,000, one would subtract 30,000 from each person's raw income score before analyzing the data so that a score of 0 on the transformed variable represents a score of 30,000 on the original metric.

Here is edited output for the estimated pure indirect effect and the total natural indirect effect for `PA2` where I mean center the covariates:

```
                                                          Two-Tailed
                    Estimate       S.E.   Est./S.E.    P-Value

  Tot natural IE       0.101      0.026       3.848      0.000
  Pure natural IE      0.099      0.026       3.768      0.000
```

Recall from Chapter 9 that the total natural indirect effect is

$$\text{TNIE} = P(\text{COM}_{\text{TREAT}}|\text{M}=m_{TREAT}) - P(\text{COM}_{\text{TREAT}}|\text{M}=m_{CTRL})$$

and the pure natural indirect effect is

$$\text{PNIE} = P(\text{COM}_{\text{CTRL}}|\text{M}=m_{TREAT}) - P(\text{COM}_{\text{CTRL}}|\text{M}=m_{CTRL})$$

where P(COM) is the probability of communication, M = *m* is the mediator mean, and the subscript TREAT or CTRL indicates if the parameter refers to the value for the treatment group or the control group. Roughly speaking, the TNIE is the indirect effect if all participants were in the treatment group, and the PNIE is the indirect effect if all participants were in the control group, with a cross-world counterfactual operative in both expressions. As discussed in Chapter 9, the two effects can be thought of as reflecting the generalizability of the indirect effect across different instantiations of counterfactual profiles, one that determines how the outcome varies as a function of change brought about in the mediator by the program and one that determines how the outcome varies as the same amount of change occurs in the natural, non-intervention world. In the current case, the two indirect effects are quite similar and they map closely onto the result for the MLPM.

When I applied the approach to the analysis of `PK2`, the TNIE was 0.100 ±0.05 ($z = 3.78$, $p < 0.05$) and the PNIE was 0.098 ±0.052 ($z = 3.71$, $p < 0.75$). For `PE2`, the TNIE was -0.001 ±0.01 ($z = 0.33$, $p < 0.75$) and the PNIE was -0.001 ±0.01 ($z = 0.33$, $p < 0.75$). These results also map reasonably onto those I found with the MLPM.

Some researchers who work in the causal mediation tradition like to sum the indirect effects to get an index of the total mediational effect (also called the total indirect effects). In this case, it equals 0.101 + 0.100 + -0.001 = 0.200 for the TNIEs and 0.099 + 0.098 + -.001 = 0.196 for the PNIEs. `PA2` and `PK2` each account for about half of the total indirect effects. A strict application of this approach is problematic in the current case because some of the indirect effects are positive and others are negative (see Chapter 10).

APPROACHES BASED ON FISEM

The Modified Linear Probability Model

For the FISEM MLPM approach using the syntax in Chapter 12, in the output section `TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS`, Mplus prints out the indirect effect analysis for each of the three mediators, as follows:

```
                                                      Two-Tailed
                   Estimate       S.E.     Est./S.E.    P-Value

Effects from TREAT to COM3

  Specific indirect 1
    COM3
    PA2
    TREAT             0.097      0.025      3.863      0.000

  Specific indirect 2
    COM3
    PK2
    TREAT             0.092      0.025      3.749      0.000

  Specific indirect 3
    COM3
    PE2
    TREAT            -0.001      0.003     -0.332      0.740
```

These represent omnibus probability/proportion differences between the treatment and control conditions for each of the mediational chains considered separately.

Probit Modeling

For probit-based FISEM that uses the Mplus syntax in Table 12.13 of Chapter 12, the indirect effects are also reported in the in the output section `TOTAL, TOTAL INDIRECT, SPECIFIC INDIRECT, AND DIRECT EFFECTS` as follows:

```
                                                   Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

  Total                0.505      0.068      7.465      0.000
  Total indirect       0.506      0.095      5.325      0.000
```

```
  Specific indirect 1
    COM3
    PA2
    TREAT                0.257      0.068      3.770      0.000

  Specific indirect 2
    COM3
    PK2
    TREAT                0.251      0.068      3.681      0.000

  Specific indirect 3
    COM3
    PE2
    TREAT               -0.002      0.008     -0.331      0.741

  Direct
    COM3
    TREAT               -0.001      0.113     -0.012      0.991
```

These effects do not take the form of probabilities but rather they are mean differences on the latent propensity y* between the treatment and control groups through each designated mediational chain. As I noted in the main text, many researchers prefer to use a partially standardized version of the analysis based on the `STANDARDIZED (STDY)` option in Mplus, which translates these mean differences into an analog of Cohen's *d*. Here is the relevant output:

```
STDY Standardization
                                                        Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

Effects from TREAT to COM3

  Total                  0.471      0.060      7.911      0.000
  Total indirect         0.472      0.086      5.490      0.000

  Specific indirect 1
    COM3
    PA2
    TREAT                0.240      0.063      3.826      0.000

  Specific indirect 2
    COM3
    PK2
    TREAT                0.234      0.063      3.734      0.000
```

```
  Specific indirect 3
    COM3
    PE2
    TREAT              -0.002      0.007     -0.331      0.741

  Direct
    COM3
    TREAT              -0.001      0.105     -0.012      0.991
```

For example, the standardized mean difference on the underlying y* propensity between the treatment and control conditions through PA2 was 0.24 ±0.13 ($z = 3.83$, $p < 0.05$).

Bayesian modeling uses the same probit-based approach but from a Bayesian perspective and a reliance on credibility intervals.

You can't use the CMF in these models because there are multiple mediators specified. For an approach that uses a "single-mediator-at-a-time", see the main text.