1

Sensitivity Tests in Structural Equation Modeling

This document describes methods for conducting sensitivity tests similar to those used in OLS regression but doing so in an SEM context. I assume that you are familiar with Mplus programming and instrumental variables (see Chapter 6 and 16). I first illustrate a strategy I call the **fixed parameter approach** to sensitivity analysis. I then describe a second strategy I call the **estimation approach** to sensitivity analysis. I use a simplified example to convey the basic logic of the fixed parameter approach and then I use the social phobia example from Chapter 11 to illustrate the estimation approach. Finally, I discuss sensitivity tests for measurement error and covariate selection.

FIXED PARAMATER APPROACH TO SENSITIVITY ANALYSIS

The model I use to illustrate the fixed parameter approach is in Figure 1. An outcome, Y, is assumed to be influenced by a mediator, M. I identified a confounding covariate, C, that I measured and statistically controlled for. I assumed that this confound is the only additional source of correlation between M and Y other than the causal impact of M on Y. By controlling for the confound, I obtain a better estimate of the causal coefficient for M → Y. A critic might argue, however, that there are other confounds that I failed to measure that could be inflating the correlation between M and Y. These unmeasured confounds are residing in both disturbance terms d_1 and d_2 so the disturbance terms are correlated. I show this scenario in Figure 1b. I can address the critic by testing the model in Figure 1b in Mplus by adding a parameter for the correlated disturbances to the model. The problem with doing so is that the model in Figure 1b is statistically under-identified; there are more unknowns than knowns and the model cannot be estimated. Hence, I cannot estimate it.

The fixed parameter approach involves estimating the model in Figure 1b but instead of estimating the covariance/correlation between the two disturbances, I fix it at an *a priori* specified value so that the covariance does not have to be estimated. It turns out that by doing so, the model is no longer under-identified. Table 1 presents the Mplus syntax for evaluating the model but with a statement that fixes the covariance between the disturbances at zero (see the third line from the bottom of the syntax). The covariance between the disturbances is introduced using the `WITH` command and the @ sign fixes the covariance at the value to the right of the @ sign. In this case, the value it is fixed at is zero.

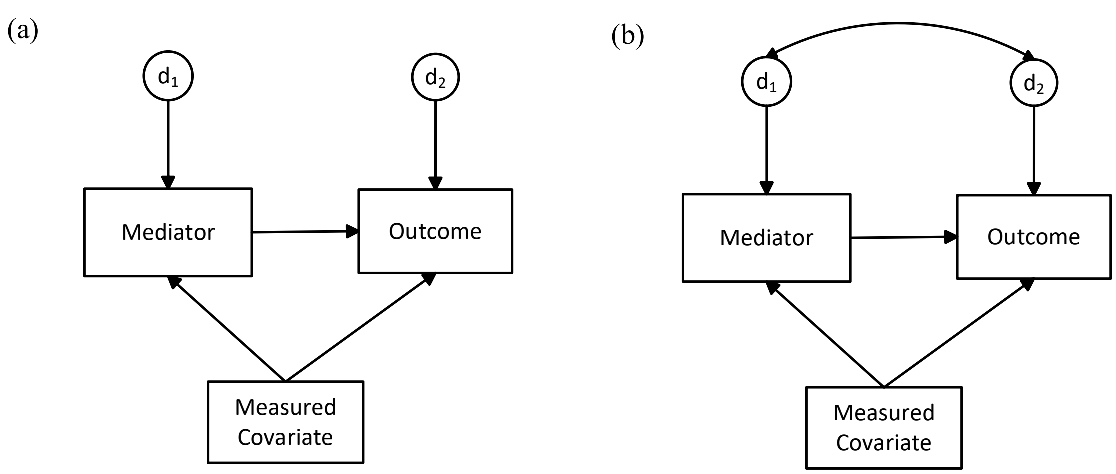

FIGURE 1. A model with and without correlated disturbances

Table 1: Mplus Syntax for Total Program Effect

```
TITLE: FIXED PARAMETER SENSITIVITY ANALYSIS ;
DATA: FILE IS c:\mplus\ret\sensitivity.dat ;
 VARIABLE:
    NAMES ARE m y c ;
    USEVARIABLES ARE m y c  ;
   MISSING ARE ALL (-9999) ;
 ANALYSIS:
   ESTIMATOR = MLR ;
 MODEL:
   y on m c ;
   m on c ;
   y with m@0 ; ! fix the disturbance covariances
 OUTPUT:
   SAMP STAND(STDYX) MOD(ALL 4) RESIDUAL CINTERVAL TECH4 ;
```

If I run this syntax, I will obtain the exact same results if I were to analyze in Mplus the model in Figure 1a because, after all, the two models are identical. The coefficient from M to Y in the model is 0.523, with a critical ratio of 14.75, $p < 0.05$.

Suppose to address the critic's concerns about unmeasured confounds I change the fixed value for the disturbance covariances from 0 to a value that reflects a correlation of 0.20 between them. The formula for a covariance, using sample notation, is

$$cov_{MY} = r_{MY}\, s_M\, s_Y$$

where cov_{MY} is the covariance between M and Y, r_{MY} is the correlation between M and Y, s_M is the standard deviation of M and s_Y is the standard deviation of Y. In the data for this example, s_M was 1.03 and s_Y was 1.01. The covariance associated with a correlation of 0.20 coupled with these standard deviations is

$cov_{MY} = (0.20)(1.03)(1.01) = 0.21$

I therefore change the relevant command in Table 1 from

```
y with m@0 ;
```

to

```
y with m@0.21 ;
```

and I re-estimate the model. The path coefficient for M influencing Y is estimated but now after adjusting for the unmeasured confounds on the assumption that the disturbances are correlated 0.20. When I did so, the result for the M→Y coefficient was 0.223 with a critical ratio of 5.77, $p < 0.05$. The coefficient is indeed reduced, as expected, but it is still non-zero and it remains statistically significant. I might argue back to the critic that even when I allowed for correlated disturbances, the effect of M on Y remained intact, albeit somewhat weakened. The critic might retort that a correlation of 0.20 between the disturbances is too weak and ask that I explore the case where the correlation is 0.30. I might counterargue that a correlation of 0.30 is too high and challenge the critic to name the unmeasured variables that would create such a large correlation between the disturbances over and above the measured confound I already controlled for in the model.

Using the fixed parameter strategy, you can evaluate different disturbance correlation scenarios that you think are plausible.

Mplus offers a tool that plots direct and indirect effects in a mediation model as a function of the correlation between disturbances. Here is sample code for a model with one distal variable (x), one mediator (m), and one outcome (y):

```
TITLE: PLOT FOR CORRELATED DISTURBANCES ;
DATA: FILE IS c:\mplus\ret\sensitivity.dat ;
 VARIABLE:
   NAMES ARE m y x ;
   MISSING ARE ALL (-9999) ;
 ANALYSIS: ESTIMATOR = MLR ;
 MODEL:
   y on m x  ;
   m on x ;
 MODEL INDIRECT:
```

```
  y IND m x  ;
OUTPUT: SAMP STAND(STDYX) MOD(ALL 4) RESIDUAL CINTERVAL TECH4 ;
PLOT: TYPE = SENSITIVITY PLOT3 ;
```

The line called `PLOT` after the `OUTPUT` line requests the sensitivity plot. I also need to specify both the mediator and the distal determinant to the right of the `IND` keyword in the `MODEL INDIRECT` command. After executing the program, I click on the PLOT menu item from the Mplus interface and then choose the option *Sensitivity Plots*. The tool is somewhat limited for the analysis of RETs because it can only deal with a single as opposed to multiple mediator models and it focuses only on the direct effect of X on Y or the full indirect effect through the mediator rather than the component parts of the indirect effect.

ESTIMATION APPROACH TO SENSITIVITY ANALYSIS

Another approach to sensitivity analysis is to add a parameter to estimate the covariance between the targeted disturbances and examine the impact that doing so has on the model parameters. For the model in Figure 1b, this was not possible because adding the parameter created an under-identified model. However, if your model has an appropriate instrumental variable in it that allows the parameter to be added without creating under-identification (see Chapter 6), then you might want to compare the values of your substantively important parameters both with and without the added correlated disturbances parameter. I now illustrate this approach using the social phobia example from Chapter 11. I repeat here for reference the influence diagram and primary syntax code for the social phobia example.

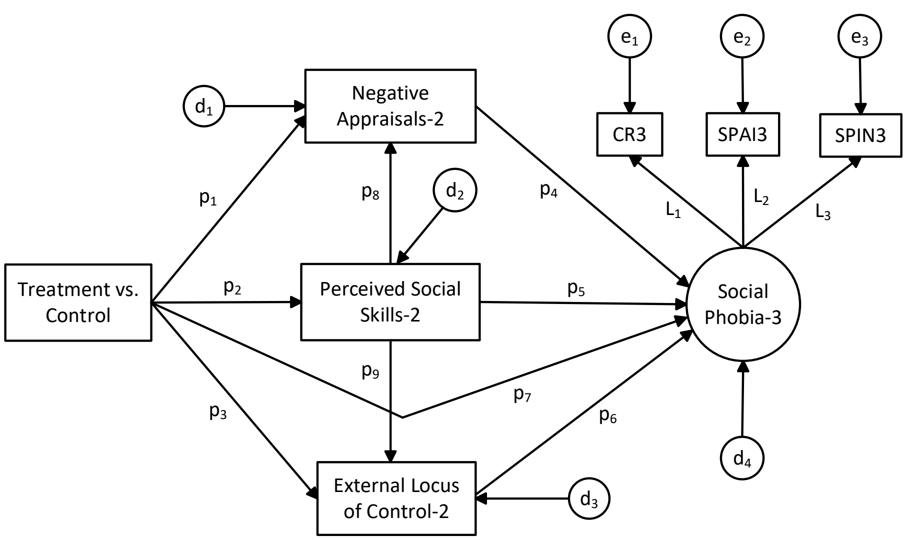

FIGURE 11.2. Social phobia example

Table 2: Mplus Syntax for Social Phobia Example

```
TITLE: EXAMPLE CHAPTER 11 ;
DATA: FILE IS c:\mplus\ret\chap11M.txt ;
VARIABLE:
NAMES ARE ID CR1 SPAI1 SPIN1 CR3 SPAI3 SPIN3
NEGAPP2 PSKILLS2 EXTERN2 NEGAPP1 PSKILLS1 EXTERN1
HYPER SEX TREAT ;
USEVARIABLES ARE CR1 SPAI1 SPIN1 CR3 SPAI3 SPIN3
NEGAPP2 PSKILLS2 EXTERN2 NEGAPP1 PSKILLS1 EXTERN1
HYPER SEX TREAT ;
MISSING ARE ALL (-9999) ;
ANALYSIS:
ESTIMATOR = MLR ;  !Robust maximum likelihood
MODEL:
!Specify latent variables
   LSP1 BY CR1 SPAI1 SPIN1 ;
   LSP3 BY CR3 SPAI3 SPIN3 ;
[CR1@0] ; [CR3@0] ; [LSP1] (mean1) ; [LSP3] (int1) ;
!Specify equations
LSP3 ON LSP1 NEGAPP2 PSKILLS2 EXTERN2 TREAT SEX (b10 p4-p7 b11) ;
LSP3 ON HYPER (b12) ;
NEGAPP2 ON TREAT HYPER SEX NEGAPP1 PSKILLS2 (p1 b1-b3 p8) ;
PSKILLS2 ON TREAT HYPER SEX PSKILLS1 (p2 b4-b6) ;
EXTERN2 ON TREAT HYPER SEX EXTERN1 PSKILLS2 (p3 b7-b9 p9) ;
!Specify correlations of latent variable with exogenous variables
LSP1 WITH NEGAPP1 PSKILLS1 EXTERN1 TREAT SEX HYPER ;
MODEL INDIRECT:
LSP3 IND TREAT ;
LSP3 IND PSKILLS2 ;
NEGAPP2 IND TREAT ;
EXTERN2 IND TREAT ;
OUTPUT:
SAMP STANDARDIZED(STDYX) MOD(ALL 4) RESIDUAL CINTERVAL TECH4 ;
```

One assumption I made in my original social phobia analysis was the absence of correlated disturbances between the negative cognitive appraisals disturbance term (d_1) and the latent social phobia disturbance term (d_4). I assumed the other covariates I included in the model surrounding these variables were adequate for controlling confounds and thereby removed any non-trivial correlation between d_1 and d_4. If my assumption is incorrect and there are non-trivial unmeasured confounds at work, then this can bias the causal coefficient between negative cognitive appraisals and social phobia. It turns out I can add a parameter reflecting the covariance between d_1 and d_4 without introducing under-identification if negative cognitive appraisals is non-trivially impacted by an instrumental variable but the latent social phobia variable is not impacted by that same variable. The baseline measure of negative cognitive appraisals plays such a role (see Figure 2).[1] I can therefore add the correlated disturbance and determine how key parameter estimates are affected. To do so, I add the following command to the Mplus syntax in Table 2 after Line 23:

```
NEGAPP2 WITH LSP3 ;
```

Here is the output that estimates the impact of `NEGAPP2` on `LSP3` in the original analysis:

```
                                                 Two-Tailed
                     Estimate       S.E.  Est./S.E.    P-Value
LSP3      ON
    NEGAPP2             0.390      0.095      4.100      0.000
    PSKILLS2           -0.707      0.099     -7.109      0.000
    EXTERN2            -0.002      0.091     -0.017      0.986
    TREAT              -0.488      0.136     -3.581      0.000
    SEX                -0.002      0.088     -0.026      0.979
    HYPER              -0.186      0.103     -1.803      0.071
    LSP1                0.347      0.072      4.835      0.000
```

and here is the corresponding output with the estimation of the correlation between d_1 and d_4 added:

```
                                                  Two-Tailed
                     Estimate       S.E.  Est./S.E.     P-Value
LSP3      ON
    NEGAPP2             0.463      0.224      2.069      0.039
    PSKILLS2           -0.674      0.132     -5.108      0.000
    EXTERN2            -0.002      0.091     -0.017      0.986
    TREAT              -0.444      0.179     -2.475      0.013
    SEX                 0.001      0.087      0.010      0.992
    HYPER              -0.196      0.106     -1.854      0.064
    LSP1                0.343      0.074      4.645      0.000
```

The pattern of statistical significance is unchanged for all of the coefficients and the coefficients are comparable in magnitude in both analyses. Also, the estimated correlation between the disturbances for `LSP3` and `NEGAPP2` was only -0.05 ($z = 0.36$, $p < 0.72$). If a critic argues that I should not have omitted the correlation between d_1 and d_4, I can reply that inclusion of it is moot.

If I believe that the correlation between d_1 and d_4 is trivial when I first specify my model, I may want to be somewhat assertive about excluding the parameter from the model. This is because the inclusion of instrumental variables coupled with the correlated disturbances can inflate standard errors and weaken statistical power. Do not move into such

[1] Some methodologists argue against using the baseline counterpart of a mediator or an outcome as an instrument, but I set that aside here in order to show you the general logic of sensitivity analysis via estimation strategies.

analyses lightly. However, it also is important to be sensitive to the possibility of biased estimates due to unmeasured confounders.

ADDITIONAL SENSITIVITY ANALYSES

There are other forms of sensitivity analysis you can undertake. For constructs measured by single indicators, one can test for result sensitivity to measurement error by using the methods described in the document associated with the link on my Resources tab called *adjust single indicators for measurement error* for Chapter 11. For result sensitivity to the choice of covariates, you can use the methods for covariate choices described in the link *preliminary analyses for the social phobia example* for Chapter 11. I urge you to check this latter document because it includes additional tests under the rubric of preliminary analyses that also can be conceptualized as sensitivity tests.