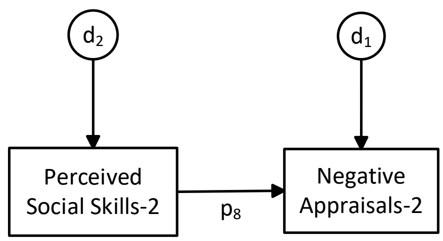
1

Dealing with Correlated Disturbances

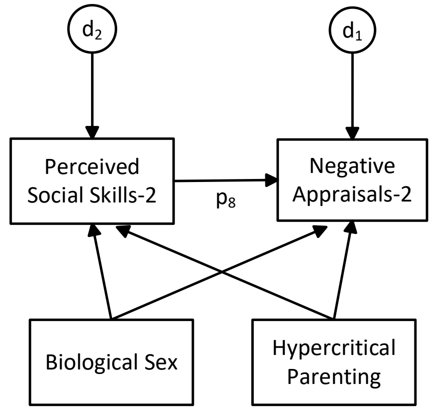
In the main text of Chapter 11, I described the importance of thinking about correlated disturbances with the idea that erroneously treating correlated disturbances as uncorrelated can introduce bias in path coefficients that are of substantive interest. This document elaborates this argument and describes correctives to help deal with unmeasured confounds. I use the main RET example in Chapter 11 to make my points.

Consider only the portion of the model that focuses on the causal effect of perceived social skills (PSS2) on negative causal appraisals:

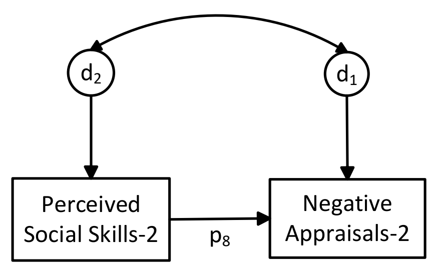


As noted in the main text, the strength of the causal effect of PSS2 on NCA2 typically is estimated using indices of the strength of the association between the two variables (usually the unstandardized or standardized regression coefficient). If, however, there are *unmeasured* confounds operating that impact both PSS2 and NCA2, then these confounds can inflate or deflate the association between the two variables and produce bias in the estimated causal effect coefficient, p_8.

One way of dealing with this problem is to measure the unmeasured confounds and to then statistically control for them, such as for biological sex and having hypercritical parents per the Chapter 11 example. Here is the portion of the influence diagram from Chapter 11 that illustrates the basic idea:

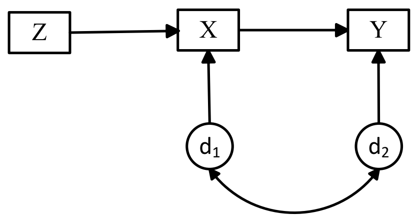


If I had not measured and include these two variables in the model, they would reside, unidentified, in d_1 and d_2, causing a correlation between d_1 and d_2. This is diagrammed as follows:



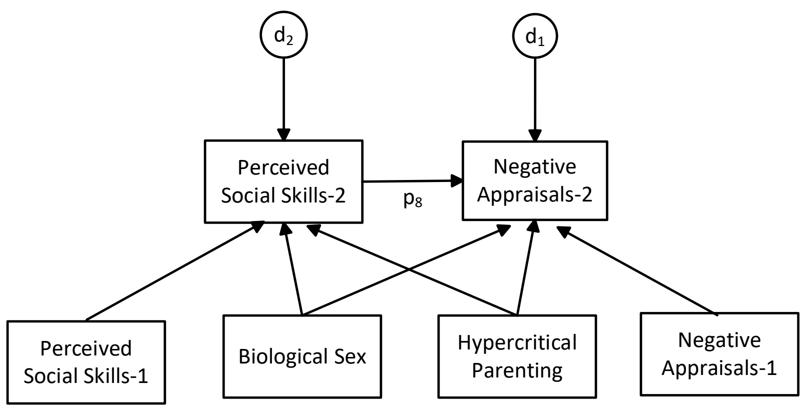
Suppose that I did not measure biological sex or parental hypercriticism in my study so that the option of statistically controlling for them is not possible. One strategy for dealing with this scenario is to add a correlation parameter between the two disturbances in Mplus that captures the degree of correlation between PSS and NCA2 that results from the omitted unmeasured confounds. This strategy explicitly models the above dynamic and takes into account the correlated disturbances when estimating p_8 in the model. A problem with this strategy, however, is that the above model portion is underidentified; I have two parameters to estimate (p_8 and the correlation between the two disturbances) but only one empirical known, namely an index of the association between PSS and NCA2.

A solution to this dilemma is to introduce into the model what is known as an **instrumental variable** (sometimes referred to as an **instrument** for short) which I introduced and discussed in Chapter 6. An instrumental variable is a variable that may or may not be a part of your core theoretical narrative but it serves a function, namely to allow me to estimate the two parameters in an otherwise underidentified model. Diagrammed using general notation, an instrumental variable, Z, has a direct impact on the cause in the targeted causal relationship but not on the outcome, like this:



The key property of the instrument in the Chapter 11 example is that it should influence PSS2 but there should be no causal path that goes directly from Z to NCA2. To be sure, Z can influence NCA2 (and, hence, be correlated with it) but it must do so exclusively through PSS2 and not directly. Z also is assumed to be uncorrelated with the two disturbances. It can be shown mathematically that if you have an instrumental variable in your model targeting the underidentified causal relationship of interest, then you can estimate the causal coefficient p_8 after adjusting for the correlation between the two disturbance terms as well as obtain an estimate of the correlation between the two disturbances.

For the RET example in Chapter 11, I included four covariates in the model in accord with the following influence diagram (note: I omit exogenous correlations to reduce clutter):



Note that neither biological sex nor hypercriticism qualifies as being an instrument because each variable directly influences both PSS2 and NCA2. However, the baseline measure of perceived social skills, PSS1, satisfies the criteria for an instrument for PSS2. This means that if I truly think there are unmeasured confounds operating in d_1 and d_2, then I can control for them by parameterizing the correlation between them in my model. By including the instrument, I circumvent the underidentification.

Here is the amended Mplus code from the main chapter (the original Table 11.1) that accomplishes this, with the key additional syntax highlighted in red:

Table 1: 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 ;
!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 ;
PSKILLS2 WITH NEGAPP2 ;
MODEL INDIRECT:
LSP3 IND TREAT ;
LSP3 IND PSKILLS2 ;
NEGAPP2 IND TREAT ;
EXTERN2 IND TREAT ;
OUTPUT:
SAMP STANDARDIZED(STDYX) MOD(ALL 4) RESIDUAL CINTERVAL TECH4 ;
```

When I executed this syntax, I found that the estimated correlation between the disturbances was relatively low, 0.085 ($p > 0.05$) and that the path coefficient p_8 between PSS2 and NCA2 was relatively unchanged from the model where I did not estimate the correlated disturbances. It seems there were not egregious unmeasured confounds in the model. The convenient presence of the instrumental variable in the model as originally formulated allowed me to address the issue of bias due to unmeasured confounds for p_8. Note that if there had not been an instrumental variable in the model to begin with, than I could still address the issue by trying to identify a viable instrument in my data set and then bringing it into the model for purposes of sensitivity analyses. Failing to identify such a variable, there are other forms of sensitivity analyses that I could conduct (see the document for how to conduct sensitivity tests in full information SEM on the Resources tab of my web page in Chapter 11).

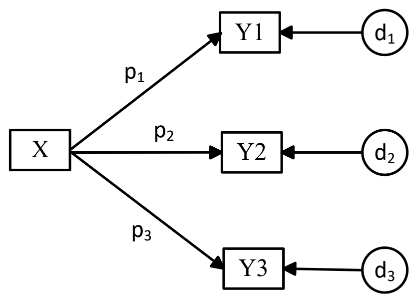
If you decide to correlate disturbances in your model, you will want to make sure you can justify doing so by explicitly specifying what the unmeasured confounds are that are problematic. Introducing correlated disturbances is not done lightly because it often comes at a cost, namely the lowering of statistical power and inflating standard errors. Correlated disturbances should be theoretically justified. Also, the instrumental variable must not be what is known as a weak instrument. Modeling with instrumental variables works best when the instrumental variables have strong relationships with the causal variable they are assumed to influence directly. If the relationships are weak, then the instrumental variables are said to be **weak instruments** and their use can actually make estimation worse. Given this, several formal tests or diagnostics have been proposed to identify weak instruments. If the instruments are weak, then one should abandon the instrument. One somewhat crude diagnostic is if the path coefficient linking the instrumental variable to the variable it directly influences is statistically significant. If it is not, then it is treated as a weak instrument. Angrist and Pischke (2009) suggest that the critical ratio associated with the significance test of the coefficient linking the instrument to the target causal variable should be 3.0 or larger. Other tests include the **Wu-Hausman test** and the **Sargan test** (see Woolridge, 2010 for details).

There is controversy in the literature about using a baseline measure of the target cause as an instrument, as I did earlier (Wang & Bellemare, 2020). The argument is that the baseline measure (in this case, PSS1) might be correlated with the disturbance term for the cause (PSS2) which then violates the criteria for qualifying as an instrument. If the violation is weak, it will not matter much. But if the violation is strong enough, then including the baseline instrument can make estimation of p_8 even more biased. If you believe that this is the case, then you should seek to find an instrument other than the baseline counterpart of the target cause to allow you to estimate the correlated disturbances. Or, at the least, note the assumption as a limitation in your report. If the baseline measure is indeed relatively independent of the disturbance term or if it can be assumed to be independent of it, then the baseline can be effectively used as an instrument as long as it is not a weak instrument (Wang & Bellemare, 2020).

The mathematics of instrumental variable analysis are rather technical so I do not delve into them here. A good discussion is in Woolridge (2010). There exist other methods for dealing with unmeasured confounds that do not rely on instrumental variables or covariate inclusion (see Chapter 6), but in my opinion these require further development.

CORRELATED DISTURBANCES FOR MULTIPLE OUTCOMES

In addition to coordinated correlated disturbances with a causal effect, another scenario where correlated disturbances are likely can occur when one or more predictors are linked to multiple endogenous variables like this:



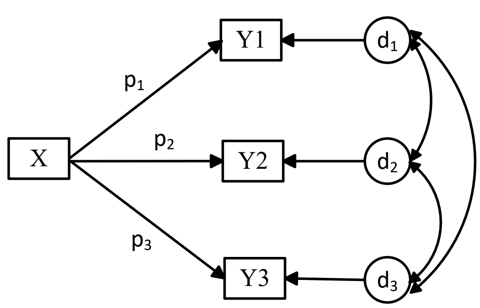
This (sub)model is defined by three linear regression equations, one for each endogenous variable:

$Y1 = a_1 + p_1 X + d_1$

$Y2 = a_2 + p_2 X + d_2$

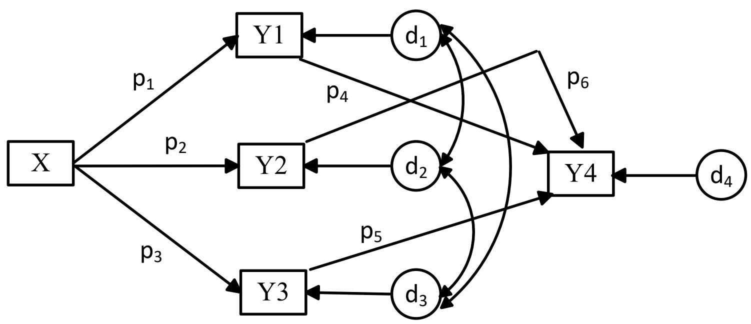
$Y3 = a_3 + p_3 X + d_3$

According to the model, Y1, Y2 and Y3 should be correlated with one another because they have a common cause, X. What some theorists do not realize, however, is that this model holds that the *only* reason the three variables are correlated is because they share X as a common cause. This assumption often is unreasonable. Rather, there are likely other unmeasured variables that simultaneously influence Y1, Y2, and Y3 (e.g., age, ethnicity) and that each reside in the disturbance terms of the various Y, i.e., there are unmeasured common causes of Y1, Y2 and Y3. To accommodate these other variables, we need to add correlations between the disturbance terms, like this:



When we add such correlations to the disturbances, we obtain what economists call a **seemingly unrelated regressions model**. These correlations often have little impact on the values of p1, p2, and p3 but they can, in some cases, affect their standard errors. They also can affect the global model fit indices of the overall SEM model.

The need to accurately reproduce the correlations between the Y variables takes on greater importance if they also serve as predictors of another endogenous variable in the model, like this:



This model adds a fourth equation to the model that takes the form of a three predictor regression equation, namely

$$Y4 = a_4 + p_4\, Y1 + p_5\, Y2 + p_6\, Y3 + d_4$$

When estimating the coefficients in a linear equation with multiple predictors, the quality of your estimates will be impacted by how well you have captured the intercorrelations among the predictors. In this case, adding the correlation parameters between the disturbances for Y1, Y2 and Y3 ensures that you accurately capture the correlations among the predictors. Without them, you may underestimate the predictor intercorrelations which can then bias your estimates of p_4, p_5 and p_6.

As before, we need to be careful about adding correlated disturbances liberally and carelessly. Such additions need to be theoretically justifiable and well articulated. Adding correlated disturbance contributes to model complexity and can pose both estimation and interpretational challenges. Having said that, you should not shy away from including them if they are theoretically appropriate and key to formulating a well-behaved model.

REFERENCES

Angrist, J. & Pischke, J. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Wang, Y & Bellemare, M. (2020). Lagged variables as instruments. Downloaded from https://marcfbellemare.com/wordpress/wp-content/uploads/2020/09/WangBellemare LagIVsJuly2020.pdf

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT Press.