1

Random Coefficients in Dynamic Structural Equation Models

This document illustrates how to analyze DSEM data with random coefficients. I assume have read and mastered the relevant material in Chapter 16. We will analyze the sleep quality example but I will use a different set of data that I simulated to conform to a random coefficient model. So the results will differ because it is a different data set. I first repeat details of the study here as a reminder.

The RET has a binary treatment variable (scored 0 = control, 1 = intervention), two continuous mediators, M1 and M2, and a continuous outcome, Y. Y is a measure of sleep quality on a given night that is obtained on each day for 28 consecutive days, post-intervention. M1 is a measure of psychological detachment from work (PDW) on a given day, reflecting the extent to which the person was able to leave the stresses and demands of work at the office. M2 is a measure of the extent to which the individual engages in appropriate wind-down behaviors relative to sleep on a given day. These include such sleep hygienic behaviors (SHBs) as turning off all electronic devices (phones, laptops, TVs) at least 30 to 60 minutes before going to bed, going to bed about the same time each night, engaging in short low-stimulation activities before going to bed, such as reading a book or stretching, creating a dark, quiet, and cool environment for sleeping, and avoiding late-day stimulants, such as coffee.

The measures of M1, M2 and Y were obtained on each of the 28 days. More stable multi-item baseline measures of M1 and M2 were obtained to use as time invariant covariates for the respective M1, M2 as well as a key time invariant covariate for all three variables that I call COV1. The metrics of M1, M2 and Y all were such that their approximate standard deviations were 1.5. The intervention sought to teach participants techniques for separating work life from home life (thereby addressing M1) and helping them understand and see the importance of engaging in sleep hygienic behaviors (thereby addressing M2). Control individuals received materials on an unrelated topic.

Figure 1 presents the two model forms tested using DSEM, (1) a between-person model on person-level mean scores calculated across the 28 days, and (2) a within-person model on the causal impact of M1 and M2 on Y. I omit covariates and disturbance terms

from the Figure to avoid clutter but are included in the Mplus syntax.

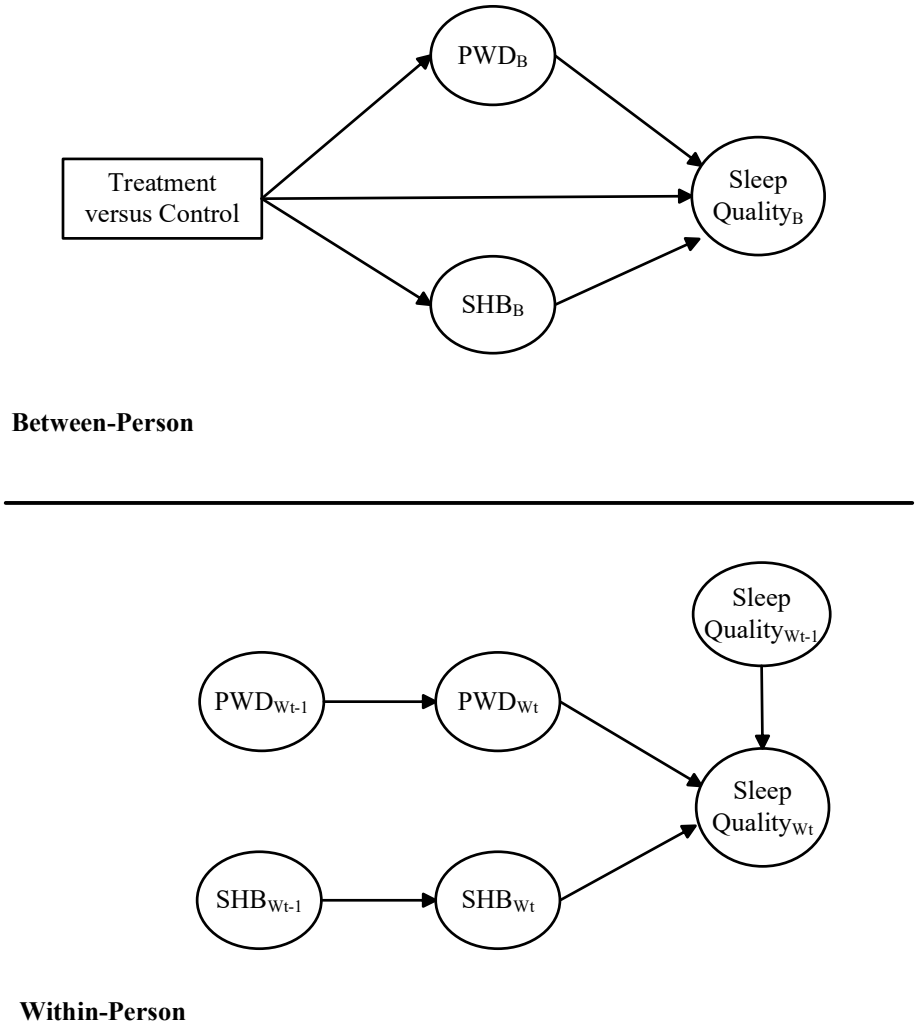

FIGURE 1 RET design for continuous time survival analysis

Consider first the within-person model. This model evaluates the estimated effect of the two mediators on the outcome on a within-individual basis. A positive path coefficient *p* for M1 means on days when a person's M1 is higher than their usual average across days (holding M2 constant), their Y also is higher. Stated somewhat differently, a one unit increase in M1 from one day to the next is associated with a *p* unit increase in Y, holding M2 constant. Note that because these coefficients are estimated within-individuals, all time invariant confounds, measured or unmeasured, are controlled for.

The model includes an autoregressive effect with lag 1 for each of the variables. Doing so controls for lingering or carry over effects of a variable from the prior day to the current day. For the logic underlying inclusion of these effects, see the main chapter text. The between-person model works with the average values of M1, M2 and Y for each person across the 28 days. This portion of the model includes the measured time invariant variables, most notably the treatment condition and the baseline measures because they are, by definition, between-person in character; they do not vary across time. The path coefficient for M1 regressed onto the binary treatment condition estimates the effect of treatment on the person's average level of M1 aggregated across the 28 days. This also applies to M2. A positive coefficient for when Y is regressed onto M1 indicates that individuals who have a higher average level of M1 across time also tend to have a higher average level of Y compared to individuals with lower average levels of M1 across time, holding M2 and the treatment condition constant. Similar interpretations apply to M2 and to the direct effect of the treatment condition on Y independent of M1 and M2.

The model in Chapter 16 assumed that the within-person coefficients were fixed. In this document, I assume all of them are random; that is they vary across individuals in conformity with a normal distribution. Table 1 presents the Mplus syntax for the model. I highlight in red the syntax differences between the random coefficient model and the fixed coefficient model reported in the main chapter text.

Table 1. Mplus Syntax for DSEM Example

```
1.  DSEM Random Coefficient Analysis;
2.  DATA:  FILE = random.dat;
3.  DEFINE:
4.    CENTER cov1 m1b m2b (GRANDMEAN) ;
5.  VARIABLE:
6.    NAMES = treat m1b m2b cov1 y m1 m2 id time  lm1 lm2 ly ;
7.    USEVARIABLES = y m1 m2 m1b m2b cov1 treat ;
8.    CLUSTER = id ;
9.    TINTERVAL = TIME(1);
10.   LAGGED = y(1) m1(1) m2(1);
11.   BETWEEN = treat m1b m2b cov1 ;
12.   MISSING ALL (-9999) ;
13. ANALYSIS:
14.   TYPE = TWOLEVEL RANDOM ;
15.   ESTIMATOR = Bayes;
16.   BITERATIONS = 10000 (100000); BCONVERGENCE =.05 ; CHAINS = 2 ;
17. MODEL:
18.   %WITHIN%
19.    SYM1 | y ON m1;              !Define random coefficients
20.    SYM2 | y ON m2;
21.    SARY | y ON y&1;
```

```
22.    SARM1 | m1 ON m1&1;
23.    SARM2 | m2 ON m2&1;
24.    y; m1; m2;           !estimate variances
25.    m1 WITH m2;        !allow correlated disturbances
26.  %BETWEEN%
27.    m1 ON treat m1b cov1 (p1 b1 b2) ;
28.    m2 ON treat m2b cov1 (p2 b3 b4) ;
29.    y ON m1 m2 treat cov1 (p3 p4 p5 b5);
30.    m1 WITH m2 ;
31.    y; m1; m2;
32.    [y] (ay) ; [m1] (am1) ; [m2] (am2) ;
33.   [SYM1] (asm1); [SYM2] (asm2);
34.   [SARM1 SARM2 SARY];   !means of random coeffs
35.   SARY SYM1 SYM2 SARM1 SARM2;                    !var of random coeffs
36.   SYM1 WITH SYM2 SARM1 SARM2 SARY;               !cor random coeffs
37.   SYM2 WITH SARM1 SARM2 SARY;
38.   SARM1 WITH SARM2 SARY;
39.   SARM2 WITH SARY;
40. MODEL CONSTRAINT:
41.   NEW (indm1 indm2 toteff) ;
42.   indm1 = p1 * p3;
43.   ndm2 = p2 * p4;
44.   toteff = indm1 + indm2 + p5 ;
45.   NEW (m1meanc m1meant m2meanc m2meant ymeanc ymeant) ;
46.   m1meanc = am1 + p1*0 ;
47.   m1meant = am1 + p1*1 ;
48.   m2meanc = am2 + p2*0 ;
49.   m2meant = am2 + p2*1 ;
50.   ymeanc = ay + p3*m1meanc + p4*m2meanc  + p5*0 ;
51.   ymeant = ay + p3*m1meant + p4*m2meant + p5*1 ;
52. PLOT: TYPE = PLOT3; FACTORS = ALL (500);
53. SAVEDATA:  FILE = datawithlags.dat;
54. OUTPUT: STDYX RESIDUAL CINTERVAL(HPD) TECH4 TECH8;
```

With a few exceptions, all of the syntax should be familiar to you. On Lines 19 to 23, I define five coefficients as being random instead of fixed. On Line 19, the entry to the left of the pipe | is a label I assign to the coefficient for the regression expression to the right of the pipe, using standard 8 character Mplus conventions. The label SYM1 is an acronym for the slope of Y on m1, per the right hand side of the pipe. Lines 20 to 23 have a similar format. Although the two mediators and three lagged effects (m1, m2, y&1, y&2, y&3) are listed on separate lines, Mplus treats them multivariately and regresses the outcome onto all of the predictors simultaneously. On Lines 33 and 34, I tell Mplus to estimate the mean coefficient across individuals (hence the request appears in the %BETWEEN% portion of the model) for each of the random coefficients. On Line 35, I tell Mplus to report the variance of the random coefficients across individuals. Lines 36 to 39 allow the random

coefficients to be correlated with one another.

To save space, I do not review evaluation of model fit. The indices I examine are the same as what I used for the fixed coefficient model in Mplus. The model fit was satisfactory. The estimates of the total effect of treat→Y and treat→M occur in the %BETWEEN% portion of the output and follow closely the output for the fixed coefficient model in the main text. Where the two forms of analysis diverge is with respect to the Mplus output for the analysis of the effect of the mediators on the outcome. I concentrate on it.

The first output I examine pertains to the average within-person coefficient for the estimated effect of M1 and M2 on Y. Here is the relevant output:

```
Between Level

                          Posterior  One-Tailed         95% C.I.
                Estimate       S.D.      P-Value   Lower 2.5%  Upper 2.5%  Sig
Means
    SYM1           0.514      0.011      0.000       0.491       0.540      *
    SYM2           0.494      0.010      0.000       0.474       0.516      *
    SARY           0.311      0.007      0.000       0.296       0.325      *
    SARM1          0.294      0.014      0.000       0.264       0.320      *
    SARM2          0.291      0.012      0.000       0.264       0.310      *
```

The average coefficient across individuals for the effect of M1 on the outcome was 0.51 with a 95% credible interval of 0.49 to 0.54. This result is statistically significant from zero ($p < 0.05$) as reflected by the fact that the credible interval does not contain the value of zero. The average coefficient across individuals for the effect of M2 on the outcome was 0.49 with a 95% credible interval of 0.47 to 0.52. This result also is statistically significant from zero ($p < 0.05$) given the credible interval does not contain the value of zero.

The section of the output on variances gives me a sense of how much across individual variability there is for each of the above coefficients. Here are the results:

```
Between Level

                          Posterior  One-Tailed         95% C.I.
                Estimate       S.D.      P-Value   Lower 2.5%  Upper 2.5%  Sig

Variances
         SYM1      0.014      0.003      0.000       0.007       0.020      *
         SYM2      0.014      0.004      0.000       0.008       0.020      *
         SARY      0.009      0.002      0.000       0.006       0.012      *
         SARM1     0.015      0.004      0.000       0.007       0.021      *
         SARM2     0.016      0.003      0.000       0.011       0.022      *
```

` The variance of M1 coefficients across individuals was 0.014 and if I calculate the square root of this I obtain the standard deviation, which is 0.12. The coefficients for M1 differed on average from the mean coefficient by 0.12 units. This variability is statistically significant ($p < 0.05$) because the value of 0 is not contained in the credible interval. The task for the researchers is to make a judgment about whether this variability is sufficiently large to be treated as meaningful. Parenthetically, the square root of the limits of the credible interval yields the credible interval for the standard deviation. Comparable results were found for M2.

The links between M1, M2 and the direct effect of the treatment condition on Y independent of M1 and M2 are obtained from the results from the syntax line `y ON m1 m2 treat cov1 (p3 p4 p5 b5);` (Line 29 in Table 1). Here is the output:

```
Between Level
```

		Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	Upper 2.5%	Sig
Y	ON						
	M1	0.362	0.048	0.000	0.269	0.447	*
	M2	0.520	0.055	0.000	0.409	0.624	*
	TREAT	0.013	0.121	0.455	-0.224	0.239	
	COV1	0.172	0.052	0.000	0.059	0.271	*

At the between-person level, for every one unit that M1 increases, Y is predicted to increase by 0.36 units with a 95% credible interval of 0.27 to 0.45. This result is statistically significant, $p < 0.05$. For every one unit that M2 increases, Y is predicted to increase by 0.52 units with a 95% credible interval of 0.41 to 0.62. This result also is statistically significant, $p < 0.05$. The direct effect of the treatment condition on Y independent of the two mediators was trivial in magnitude (coefficient = 0.013 with a 95% credible interval of -0.22 to 0.0.24) and statistically non-significant.

One final point. For the omnibus mediation effect for M1 and M2, you can evaluate it using the joint significance test or you can use the you use the same syntax as that for the fixed coefficient model in the main text (see Lines 42 and 43 in Table 1). An exception would be if the two links in the mediational chain are *both* random coefficients. In such cases you need to add to the product term the covariance of the two random coefficients in the `MODEL CONSTRAINT` command. In the current case it is not necessary because the coefficients involved are not both random coefficients.