Online Appendix to accompany
Sterba, S.K., Preacher, K.J., Forehand, R., Hardcastle, E.J., Cole, D.J. & Compas, B.E. (2014). "Structural Equation Modeling Approaches for Analyzing Partially Nested Data." ***Multivariate Behavioral Research***

Contents:

Section A: SSEM-PN & MSEM-PN Equations for Empirical Example

This section of the online appendix presents SSEM-PN and MSEM-PN equations for the empirical example depression intervention analysis. Notation corresponds to Figures 6 and 7. Empirical example variables corresponding with y_{ij}, x_{ij}, and m_j were mentioned in the text. Parameters pertaining to the *y*-outcome equation are superscripted *y* (similarly for *m* superscripts). *t* or *c* superscripts denote treatment or control arm.

MSEM-PN **<u>Empirical example</u>** **equations:**

Treatment arm:

$$y_{ij} = \tilde{y}_j^t + \tilde{y}_{ij}^t$$
$$x_{ij} = \tilde{x}_j^t + \tilde{x}_{ij}^t$$
$$m_j = \tilde{m}_j^t$$

Between:

$$\tilde{y}_j^t = \mu^{yt} + b_{1B}^{yt}\tilde{m}_j^t + b_{2B}^{yt}\tilde{x}_j^t + \zeta_j^{yt}$$
$$\tilde{m}_j^t = \mu^{mt} + b_{1B}^{mt}\tilde{x}_j^t + \zeta_j^{mt}$$
$$\tilde{x}_j^t = \mu^{xt} + \zeta_j^{xt}$$

Within:

$$\tilde{y}_{ij}^t = 0 + b_{3W}^{yt}\tilde{x}_{ij}^t + \varepsilon_{ij}^{yt}$$
$$\tilde{x}_{ij}^t = 0 + \varepsilon_{ij}^{xt}$$

Control arm:

$$y_{ij} = \tilde{y}_j^c + \tilde{y}_{ij}^c$$
$$x_{ij} = \tilde{x}_j^c + \tilde{x}_{ij}^c$$

Between:

$$\tilde{y}_j^c = \mu^{yc} + b^{yc}\tilde{x}_j^c + \zeta_j^{yc}$$
$$\tilde{x}_j^c = \mu^{xc} + \zeta_j^{xc}$$

Within:

$$\tilde{y}_{ij}^c = 0$$
$$\tilde{x}_{ij}^c = 0$$

Where

$$\begin{bmatrix}\zeta_j^{yt}\\ \zeta_j^{xt}\\ \zeta_j^{mt}\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}\psi^{yt} & & \\ 0 & \psi^{xt} & \\ 0 & 0 & \psi^{mt}\end{bmatrix}\right) \quad \text{and} \quad \begin{bmatrix}\zeta_j^{yc}\\ \zeta_j^{xc}\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}\psi^{yc} & \\ 0 & \psi^{xc}\end{bmatrix}\right)$$

$$\begin{bmatrix}\varepsilon_{ij}^{yt}\\ \varepsilon_{ij}^{xt}\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}\theta_\varepsilon^{yt} & \\ 0 & \theta_\varepsilon^{xt}\end{bmatrix}\right)$$

SSEM-PN <u>Empirical example</u> equations:

$$\begin{bmatrix} y_{1j}^{c} \\ \mathbf{y}_j^{t} \\ x_{1j}^{c} \\ \mathbf{x}_j^{t} \\ m_j^{t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \mathbf{0}_{1\times4} & \mathbf{0}_{1\times4} \\ \mathbf{0}_{4\times1} & \mathbf{1}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{I}_{4\times4} & \mathbf{0}_{4\times4} \\ 0 & 0 & 1 & 0 & 0 & \mathbf{0}_{1\times4} & \mathbf{0}_{1\times4} \\ \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{1}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times4} & \mathbf{I}_{4\times4} \\ 0 & 0 & 0 & 0 & 1 & \mathbf{0}_{1\times4} & \mathbf{0}_{1\times4} \end{bmatrix} \begin{bmatrix} \eta_j^{yc} \\ \eta_j^{yt} \\ \eta_j^{xc} \\ \eta_j^{xt} \\ \eta_j^{mt} \\ \boldsymbol{\eta}_j^{yt} \\ \boldsymbol{\eta}_j^{xt} \end{bmatrix}$$

$$\begin{bmatrix} \eta_j^{yc} \\ \eta_j^{yt} \\ \eta_j^{xc} \\ \eta_j^{xt} \\ \eta_j^{mt} \\ \boldsymbol{\eta}_j^{yt} \\ \boldsymbol{\eta}_j^{xt} \end{bmatrix} = \begin{bmatrix} \mu^{yc} \\ \mu^{yt} \\ \mu^{xc} \\ \mu^{xt} \\ \mu^{mt} \\ \mathbf{0}_{4\times1} \\ \mathbf{0}_{4\times1} \end{bmatrix} + \begin{bmatrix} 0 & 0 & b^{yc} & 0 & 0 & \mathbf{0}_{1\times4} & \mathbf{0}_{1\times4} \\ 0 & 0 & 0 & b_{2B}^{yt} & b_{1B}^{yt} & \mathbf{0}_{1\times4} & \mathbf{0}_{1\times4} \\ 0 & 0 & 0 & 0 & 0 & \mathbf{0}_{1\times4} & \mathbf{0}_{1\times4} \\ 0 & 0 & 0 & 0 & 0 & \mathbf{0}_{1\times4} & \mathbf{0}_{1\times4} \\ 0 & 0 & 0 & b_{1B}^{mt} & 0 & \mathbf{0}_{1\times4} & \mathbf{0}_{1\times4} \\ \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times4} & b_{3W}^{yt}\mathbf{I}_{4\times4} \\ \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times4} & \mathbf{0}_{4\times4} \end{bmatrix} \begin{bmatrix} \eta_j^{yc} \\ \eta_j^{yt} \\ \eta_j^{xc} \\ \eta_j^{xt} \\ \eta_j^{mt} \\ \boldsymbol{\eta}_j^{yt} \\ \boldsymbol{\eta}_j^{xt} \end{bmatrix} + \begin{bmatrix} \zeta_j^{yc} \\ \zeta_j^{yt} \\ \zeta_j^{xc} \\ \zeta_j^{xt} \\ \zeta_j^{mt} \\ \boldsymbol{\varepsilon}_{j_{4\times1}}^{yt} \\ \boldsymbol{\varepsilon}_{j_{4\times1}}^{xt} \end{bmatrix}$$

where $\mathbf{y}_j^t$ and $\mathbf{x}_j^t$ are 4×1 in this example and where:

$$\begin{bmatrix} \zeta_j^{yc} \\ \zeta_j^{yt} \\ \zeta_j^{xc} \\ \zeta_j^{xt} \\ \zeta_j^{mt} \\ \boldsymbol{\varepsilon}_{j_{4\times1}}^{yt} \\ \boldsymbol{\varepsilon}_{j_{4\times1}}^{xt} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \mathbf{0}_{4\times1} \\ \mathbf{0}_{4\times1} \end{bmatrix}, \begin{bmatrix} \psi^{yc} & & & & & & \\ 0 & \psi^{yt} & & & & & \\ 0 & 0 & \psi^{xc} & & & & \\ 0 & 0 & 0 & \psi^{xt} & & & \\ 0 & 0 & 0 & 0 & \psi^{mt} & & \\ \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \theta_\varepsilon^{yt}\mathbf{I}_{4\times4} & \\ \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times4} & \theta_\varepsilon^{xt}\mathbf{I}_{4\times4} \end{bmatrix} \right).$$

Section B: MLM-PN Matrix expressions

In this section MLM-PN specifications that were given in the manuscript in scalar expressions are presented in corresponding matrix expressions.

MLM-PN, using the Laird & Ware (1982) MLM matrix expression

$$\mathbf{y}_j = \mathbf{X}_j\boldsymbol{\gamma} + \mathbf{Z}_j\mathbf{u}_j + \mathbf{r}_j$$

Below is each MLM-PN specification from the manuscript, *with matrix elements shown for a prototypic treatment arm cluster, and a prototypic control arm cluster*.

MODEL: Basic model with treatment as only predictor

Treatment cluster (trt_{ij} =1)	$\begin{bmatrix} y_{1j} \\ y_{2j} \\ y_{3j} \\ y_{4j} \\ y_{5j} \end{bmatrix} = \begin{bmatrix} 1 & trt_{1j} \\ 1 & trt_{2j} \\ 1 & trt_{3j} \\ 1 & trt_{4j} \\ 1 & trt_{5j} \end{bmatrix} \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \end{bmatrix} + \begin{bmatrix} trt_{1j} \\ trt_{2j} \\ trt_{3j} \\ trt_{4j} \\ trt_{5j} \end{bmatrix} \begin{bmatrix} u_{1j} \end{bmatrix} + \begin{bmatrix} r_{1j} \\ r_{2j} \\ r_{3j} \\ r_{4j} \\ r_{5j} \end{bmatrix}$
Control cluster (trt_{1j} =0)	$\begin{bmatrix} y_{1j} \end{bmatrix} = \begin{bmatrix} 1 & trt_{1j} \end{bmatrix} \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \end{bmatrix} + \begin{bmatrix} trt_{1j} \end{bmatrix} \begin{bmatrix} u_{1j} \end{bmatrix} + \begin{bmatrix} r_{1j} \end{bmatrix}$

Note: Residual and random effect variances were given in the manuscript.

MODEL: Adding exogenous covariates

Treatment cluster (trt_{ij} =1)	$\begin{bmatrix} y_{1j} \\ y_{2j} \\ y_{3j} \\ y_{4j} \\ y_{5j} \end{bmatrix} = \begin{bmatrix} 1 & trt_{1j} & x_{1j} & w_j trt_{1j} \\ 1 & trt_{2j} & x_{2j} & w_j trt_{2j} \\ 1 & trt_{3j} & x_{3j} & w_j trt_{3j} \\ 1 & trt_{4j} & x_{4j} & w_j trt_{4j} \\ 1 & trt_{5j} & x_{5j} & w_j trt_{5j} \end{bmatrix} \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \\ \gamma_{20} \\ \gamma_{11} \end{bmatrix} + \begin{bmatrix} trt_{1j} \\ trt_{2j} \\ trt_{3j} \\ trt_{4j} \\ trt_{5j} \end{bmatrix} \begin{bmatrix} u_{1j} \end{bmatrix} + \begin{bmatrix} r_{1j} \\ r_{2j} \\ r_{3j} \\ r_{4j} \\ r_{5j} \end{bmatrix}$
Control cluster (trt_{1j} =0)	$\begin{bmatrix} y_{1j} \end{bmatrix} = \begin{bmatrix} 1 & trt_{1j} & x_{1j} & w_j trt_{1j} \end{bmatrix} \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \\ \gamma_{20} \\ \gamma_{11} \end{bmatrix} + \begin{bmatrix} trt_{1j} \end{bmatrix} \begin{bmatrix} u_{1j} \end{bmatrix} + \begin{bmatrix} r_{1j} \end{bmatrix}$

Notes: Missing data handling procedures when w_j is missing in the control arm are discussed in the manuscript. Residual and random effect variances were given in the manuscript.

MODEL: Adding treatment arms

Trt1 cluster $\begin{pmatrix} trt1_{ij}=1 \\ trt2_{ij}=0 \end{pmatrix}$	$\begin{bmatrix} y_{1j} \\ y_{2j} \\ y_{3j} \\ y_{4j} \\ y_{5j} \end{bmatrix} = \begin{bmatrix} 1 & trt1_{1j} & x_{1j} & w_j trt1_{1j} & trt2_{1j} & w_j trt2_{1j} \\ 1 & trt1_{2j} & x_{2j} & w_j trt1_{2j} & trt2_{2j} & w_j trt2_{2j} \\ 1 & trt1_{3j} & x_{3j} & w_j trt1_{3j} & trt2_{3j} & w_j trt2_{3j} \\ 1 & trt1_{4j} & x_{4j} & w_j trt1_{4j} & trt2_{4j} & w_j trt2_{4j} \\ 1 & trt1_{5j} & x_{5j} & w_j trt1_{5j} & trt2_{5j} & w_j trt2_{5j} \end{bmatrix} \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \\ \gamma_{20} \\ \gamma_{11} \\ \gamma_{30} \\ \gamma_{31} \end{bmatrix} + \begin{bmatrix} trt1_{1j} & trt2_{1j} \\ trt1_{2j} & trt2_{2j} \\ trt1_{3j} & trt2_{3j} \\ trt1_{4j} & trt2_{4j} \\ trt1_{5j} & trt2_{5j} \end{bmatrix} \begin{bmatrix} u_{1j} \\ u_{3j} \end{bmatrix} + \begin{bmatrix} r_{1j} \\ r_{2j} \\ r_{3j} \\ r_{4j} \\ r_{5j} \end{bmatrix}$
Trt2 cluster $\begin{pmatrix} trt1_{ij}=0 \\ trt2_{ij}=1 \end{pmatrix}$	$\begin{bmatrix} y_{1j} \\ y_{2j} \end{bmatrix} = \begin{bmatrix} 1 & trt1_{1j} & x_{1j} & w_j trt1_{1j} & trt2_{1j} & w_j trt2_{1j} \\ 1 & trt1_{2j} & x_{2j} & w_j trt1_{2j} & trt2_{2j} & w_j trt2_{2j} \end{bmatrix} \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \\ \gamma_{20} \\ \gamma_{11} \\ \gamma_{30} \\ \gamma_{31} \end{bmatrix} + \begin{bmatrix} trt1_{1j} & trt2_{1j} \\ trt1_{2j} & trt2_{2j} \end{bmatrix} \begin{bmatrix} u_{1j} \\ u_{3j} \end{bmatrix} + \begin{bmatrix} r_{1j} \end{bmatrix}$
Control cluster $\begin{pmatrix} trt1_{ij}=0 \\ trt2_{ij}=0 \end{pmatrix}$	$\begin{bmatrix} y_{1j} \end{bmatrix} = \begin{bmatrix} 1 & trt1_{1j} & x_{1j} & w_j trt1_{1j} & trt2_{1j} & w_j trt2_{1j} \end{bmatrix} \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \\ \gamma_{20} \\ \gamma_{11} \\ \gamma_{30} \\ \gamma_{31} \end{bmatrix} + \begin{bmatrix} trt1_{1j} & trt2_{1j} \end{bmatrix} \begin{bmatrix} u_{1j} \\ u_{3j} \end{bmatrix} + \begin{bmatrix} r_{1j} \end{bmatrix}$

Notes: Missing data handling procedures when w_j is missing in the control arm are discussed in the manuscript. Residual and random effect variances were given in the manuscript.

MODEL: Strategy A for disaggregating effects, with observed cluster means

Trt cluster ($trt_{ij}=1$)	$\begin{bmatrix} y_{1j} \\ y_{2j} \\ y_{3j} \\ y_{4j} \\ y_{5j} \end{bmatrix} = \begin{bmatrix} 1 & trt_{1j} & \dot{x}_{1j} & w_j trt_{1j} & x_{1j}(1-trt_{1j}) & \bar{x}_{.j} trt_{1j} \\ 1 & trt_{2j} & \dot{x}_{2j} & w_j trt_{2j} & x_{2j}(1-trt_{2j}) & \bar{x}_{.j} trt_{2j} \\ 1 & trt_{3j} & \dot{x}_{3j} & w_j trt_{3j} & x_{3j}(1-trt_{3j}) & \bar{x}_{.j} trt_{3j} \\ 1 & trt_{4j} & \dot{x}_{4j} & w_j trt_{4j} & x_{4j}(1-trt_{4j}) & \bar{x}_{.j} trt_{4j} \\ 1 & trt_{5j} & \dot{x}_{5j} & w_j trt_{5j} & x_{5j}(1-trt_{5j}) & \bar{x}_{.j} trt_{5j} \end{bmatrix} \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \\ \gamma_{20} \\ \gamma_{11} \\ \gamma_{30} \\ \gamma_{12} \end{bmatrix} + \begin{bmatrix} trt_{1j} \\ trt_{2j} \\ trt_{3j} \\ trt_{4j} \\ trt_{5j} \end{bmatrix} \begin{bmatrix} u_{1j} \end{bmatrix} + \begin{bmatrix} r_{1j} \\ r_{2j} \\ r_{3j} \\ r_{4j} \\ r_{5j} \end{bmatrix}$
Control cluster ($trt_{ij}=0$)	$\begin{bmatrix} y_{1j} \end{bmatrix} = \begin{bmatrix} 1 & trt_{1j} & \dot{x}_{1j} & w_j trt_{1j} & x_{1j}(1-trt_{1j}) & \bar{x}_{.j} trt_{1j} \end{bmatrix} \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \\ \gamma_{20} \\ \gamma_{11} \\ \gamma_{30} \\ \gamma_{12} \end{bmatrix} + \begin{bmatrix} trt_{1j} \end{bmatrix} \begin{bmatrix} u_{1j} \end{bmatrix} + \begin{bmatrix} r_{1j} \end{bmatrix}$

Notes: Missing data handling procedures when w_j is missing in the control arm are discussed in the manuscript. Residual and random effect variances were given in the manuscript.

MODEL: Strategy B for disaggregating effects, with observed cluster means

Trt cluster ($trt_{ij} = 1$)	$\begin{bmatrix} y_{1j} \\ y_{2j} \\ y_{3j} \\ y_{4j} \\ y_{5j} \end{bmatrix} = \begin{bmatrix} 1 & trt_{1j} & x_{1j}trt_{1j} & w_j trt_{1j} & x_{1j}(1-trt_{1j}) & \bar{x}_{.j}trt_{1j} \\ 1 & trt_{2j} & x_{2j}trt_{2j} & w_j trt_{2j} & x_{2j}(1-trt_{2j}) & \bar{x}_{.j}trt_{2j} \\ 1 & trt_{3j} & x_{3j}trt_{3j} & w_j trt_{3j} & x_{3j}(1-trt_{3j}) & \bar{x}_{.j}trt_{3j} \\ 1 & trt_{4j} & x_{4j}trt_{4j} & w_j trt_{4j} & x_{4j}(1-trt_{4j}) & \bar{x}_{.j}trt_{4j} \\ 1 & trt_{5j} & x_{5j}trt_{5j} & w_j trt_{5j} & x_{5j}(1-trt_{5j}) & \bar{x}_{.j}trt_{5j} \end{bmatrix} \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \\ \gamma_{20} \\ \gamma_{11} \\ \gamma_{30} \\ \gamma_{12} \end{bmatrix} + \begin{bmatrix} trt_{1j} \\ trt_{2j} \\ trt_{3j} \\ trt_{4j} \\ trt_{5j} \end{bmatrix} \begin{bmatrix} u_{1j} \end{bmatrix} + \begin{bmatrix} r_{1j} \\ r_{2j} \\ r_{3j} \\ r_{4j} \\ r_{5j} \end{bmatrix}$
Control cluster ($trt_{ij} = 0$)	$\begin{bmatrix} y_{1j} \end{bmatrix} = \begin{bmatrix} 1 & trt_{1j} & x_{1j}trt_{1j} & w_j trt_{1j} & x_{1j}(1-trt_{1j}) & \bar{x}_{.j}trt_{1j} \end{bmatrix} \begin{bmatrix} \gamma_{00} \\ \gamma_{10} \\ \gamma_{20} \\ \gamma_{11} \\ \gamma_{30} \\ \gamma_{12} \end{bmatrix} + \begin{bmatrix} trt_{1j} \end{bmatrix} \begin{bmatrix} u_{1j} \end{bmatrix} + \begin{bmatrix} r_{1j} \end{bmatrix}$

Notes: Missing data handling procedures when w_j is missing in the control arm are discussed in the manuscript. Residual and random effect variances were given in the manuscript.

Section C: MSEM-PN Matrix expressions

Below is a matrix MSEM-PN expression for each scalar MSEM-PN specification described in the manuscript, using an adaptation of Muthén and Asparouhov's (2009) MSEM matrix formulation.

Each expression consists of 3 matrix equations: the first equation is a measurement model relating observed variables to latent within and between components; the second equation is a structural model for latent components; and the third equation is a structural model for random effects. (Scalar equations for the level 1 measurement model, and scalar reduced form equations for the level 1 and 2 structural models were given in the manuscript).

Importantly, note that, following Muthén and Asparouhov (2009), the below MSEM-PN matrix representations are *for a generic level 1 unit i,* nested in generic cluster *j,* which contrasts with the Laird-Ware MLM representation above, which was depicted for a generic cluster *j.* The matrix expressions below are shown in a conventional "multiple-group" matrix notation (here, multiple *arm,* since study arm is the grouping variable, designated *g*).

Matrices **in grey** are not used in a given model specification but are part of the general MSEM expression and may be used in a later specification.

Note that because these matrix expressions are stated generically for the g^{th} arm, rather than restated separately for the treatment arm vs. the control arm, these expressions do not convey the arm-specific constraints imposed to accommodate the partially nested design (e.g., if a residual was constrained to 0 in the control arm only). See the manuscript for details on these constraints.

MODEL: Basic model with treatment as only predictor.

Superscript *g* stands for study arm, such that *g*=t or c.

$$\mathbf{y}_{ij}^{(g)} = \boldsymbol{\Lambda}^{(g)}\tilde{\boldsymbol{\eta}}_{ij}^{(g)}$$

$$\left[y_{ij}^{(g)} \right] = [1\ 1]\begin{bmatrix} \tilde{y}_{ij}^{(g)} \\ \tilde{y}_{j}^{(g)} \end{bmatrix}$$

$$\tilde{\boldsymbol{\eta}}_{ij}^{(g)} = \boldsymbol{\alpha}_{j}^{(g)} + \mathbf{K}_{j}^{(g)}\mathbf{x}_{ij}^{(g)} + \mathbf{B}_{W}^{(g)}\tilde{\boldsymbol{\eta}}_{ij}^{(g)} + \boldsymbol{\varepsilon}_{ij}^{(g)}$$

$$\begin{bmatrix} \tilde{y}_{ij}^{(g)} \\ \tilde{y}_{j}^{(g)} \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha_{j}^{(g)} \end{bmatrix} + \begin{bmatrix} \varepsilon_{ij}^{(g)} \\ 0 \end{bmatrix}$$

$$\tilde{\boldsymbol{\eta}}_{j}^{(g)} = \boldsymbol{\mu}^{(g)} + \boldsymbol{\Omega}^{(g)}\mathbf{w}_{j}^{(g)} + \mathbf{B}_{B}^{(g)}\tilde{\boldsymbol{\eta}}_{j}^{(g)} + \boldsymbol{\zeta}_{j}^{(g)}$$

$$\left[\alpha_{j}^{(g)} \right] = \left[\mu^{(g)} \right] + \left[\zeta_{j}^{(g)} \right]$$

Notes: Residual and random effect variances were given in the manuscript.

MODEL: Adding exogenous covariates

$$\mathbf{y}_{ij}^{(g)} = \mathbf{\Lambda}^{(g)}\tilde{\boldsymbol{\eta}}_{ij}^{(g)}$$
$$\begin{bmatrix} y_{ij}^{(g)} \end{bmatrix} = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \tilde{y}_{ij}^{(g)} \\ \tilde{y}_{j}^{(g)} \end{bmatrix}$$

$$\tilde{\boldsymbol{\eta}}_{ij}^{(g)} = \boldsymbol{\alpha}_{j}^{(g)} + \mathbf{K}_{j}^{(g)}\mathbf{x}_{ij}^{(g)} + \mathbf{B}_{W}^{(g)}\tilde{\boldsymbol{\eta}}_{ij}^{(g)} + \boldsymbol{\varepsilon}_{ij}^{(g)}$$
$$\begin{bmatrix} \tilde{y}_{ij}^{(g)} \\ \tilde{y}_{j}^{(g)} \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha_{j}^{(g)} \end{bmatrix} + \begin{bmatrix} \kappa^{(g)} \\ 0 \end{bmatrix} \begin{bmatrix} x_{ij}^{(g)} \end{bmatrix} + \begin{bmatrix} \varepsilon_{ij}^{(g)} \\ 0 \end{bmatrix}$$

$$\tilde{\boldsymbol{\eta}}_{j}^{(g)} = \boldsymbol{\mu}^{(g)} + \mathbf{\Omega}^{(g)}\mathbf{w}_{j}^{(g)} + \mathbf{B}_{B}^{(g)}\tilde{\boldsymbol{\eta}}_{j}^{(g)} + \boldsymbol{\zeta}_{j}^{(g)}$$
$$\begin{bmatrix} \alpha_{j}^{(g)} \end{bmatrix} = \begin{bmatrix} \mu^{(g)} \end{bmatrix} + \begin{bmatrix} \omega_{1}^{(g)} \end{bmatrix} \begin{bmatrix} w_{j}^{(g)} \end{bmatrix} + \begin{bmatrix} \zeta_{j}^{(g)} \end{bmatrix}$$

Notes: Missing data handling procedures when w_j is missing in the control arm are discussed in the manuscript. Residual and random effect variances were given in the manuscript.

MODEL: Adding treatment arms

The matrix expression from immediately above applies, only now g=t_1, t_2, or c.

MODEL: Adding a level 2 outcome

$$\mathbf{y}_{ij}^{(g)} = \mathbf{\Lambda}^{(g)}\tilde{\boldsymbol{\eta}}_{ij}^{(g)}$$
$$\begin{bmatrix} y_{ij}^{(g)} \\ z_{j}^{(g)} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{y}_{ij}^{(g)} \\ \tilde{y}_{j}^{(g)} \\ \tilde{z}_{j}^{(g)} \end{bmatrix}$$

$$\tilde{\boldsymbol{\eta}}_{ij}^{(g)} = \boldsymbol{\alpha}_{j}^{(g)} + \mathbf{K}_{j}^{(g)}\mathbf{x}_{ij}^{(g)} + \mathbf{B}_{W}^{(g)}\tilde{\boldsymbol{\eta}}_{ij}^{(g)} + \boldsymbol{\varepsilon}_{ij}^{(g)}$$
$$\begin{bmatrix} \tilde{y}_{ij}^{(g)} \\ \tilde{y}_{j}^{(g)} \\ \tilde{z}_{j}^{(g)} \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha_{j}^{y(g)} \\ \alpha_{j}^{z(g)} \end{bmatrix} + \begin{bmatrix} \kappa^{y(g)} \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} x_{ij}^{(g)} \end{bmatrix} + \begin{bmatrix} \varepsilon_{ij}^{y(g)} \\ 0 \\ 0 \end{bmatrix}$$

$$\tilde{\boldsymbol{\eta}}_{j}^{(g)} = \boldsymbol{\mu}^{(g)} + \mathbf{\Omega}^{(g)}\mathbf{w}_{j}^{(g)} + \mathbf{B}_{B}^{(g)}\tilde{\boldsymbol{\eta}}_{j}^{(g)} + \boldsymbol{\zeta}_{j}^{(g)}$$
$$\begin{bmatrix} \alpha_{j}^{y(g)} \\ \alpha_{j}^{z(g)} \end{bmatrix} = \begin{bmatrix} \mu^{y(g)} \\ \mu^{z(g)} \end{bmatrix} + \begin{bmatrix} \omega_{1}^{y(g)} \\ \omega_{1}^{z(g)} \end{bmatrix} \begin{bmatrix} w_{j}^{(g)} \end{bmatrix} + \begin{bmatrix} \zeta_{j}^{y(g)} \\ \zeta_{j}^{z(g)} \end{bmatrix}$$

Notes: Missing data handling procedures when w_j is missing in the control arm are discussed in the manuscript. Residual and random effect variances were given in the manuscript.

MODEL: Strategy A for disaggregating effects, with observed cluster means

$$\mathbf{y}_{ij}^{(g)} = \mathbf{\Lambda}^{(g)}\tilde{\boldsymbol{\eta}}_{ij}^{(g)}$$

$$\left[y_{ij}^{(g)} \right] = [1\ 1]\begin{bmatrix} \tilde{y}_{ij}^{(g)} \\ \tilde{y}_{j}^{(g)} \end{bmatrix}$$

$$\tilde{\boldsymbol{\eta}}_{ij}^{(g)} = \boldsymbol{\alpha}_{j}^{(g)} + \mathbf{K}_{j}^{(g)}\mathbf{x}_{ij}^{(g)} + \mathbf{B}_{W}^{(g)}\tilde{\boldsymbol{\eta}}_{ij}^{(g)} + \boldsymbol{\varepsilon}_{ij}^{(g)}$$

$$\begin{bmatrix} \tilde{y}_{ij}^{(g)} \\ \tilde{y}_{j}^{(g)} \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha_{j}^{(g)} \end{bmatrix} + \begin{bmatrix} \kappa_{W}^{(g)} \\ 0 \end{bmatrix}\left[\dot{x}_{ij}^{(g)} \right] + \begin{bmatrix} \varepsilon_{ij}^{(g)} \\ 0 \end{bmatrix}$$

--where, for the control arm, substitute the raw x_{ij} instead of the cluster mean centered $\dot{x}_{ij}$

$$\tilde{\boldsymbol{\eta}}_{j}^{(g)} = \boldsymbol{\mu}^{(g)} + \mathbf{\Omega}^{(g)}\mathbf{w}_{j}^{(g)} + \mathbf{B}_{B}^{(g)}\tilde{\boldsymbol{\eta}}_{j}^{(g)} + \boldsymbol{\zeta}_{j}^{(g)}$$

$$\left[\alpha_{j}^{(g)} \right] = \left[\mu^{(g)} \right] + \left[\omega_{1}^{(g)}\ \ \omega_{2B}^{(g)} \right]\begin{bmatrix} w_{j}^{(g)} \\ \bar{x}_{.j}^{(g)} \end{bmatrix} + \left[\zeta_{j}^{(g)} \right]$$

--note that the *W* and *B* subscripts only pertain to the treatment arm effects, and are deleted in the control arm's scalar expressions in the article

Notes: Missing data handling procedures when w_j is missing in the control arm are discussed in the manuscript. Residual and random effect variances were given in the manuscript.

MODEL: Disaggregating effects, with latent cluster means

$$\mathbf{y}_{ij}^{(g)} = \mathbf{\Lambda}^{(g)}\tilde{\boldsymbol{\eta}}_{ij}^{(g)}$$

$$\begin{bmatrix} y_{ij}^{(g)} \\ x_{ij}^{(g)} \\ w_{j}^{(g)} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \tilde{y}_{ij}^{(g)} \\ \tilde{y}_{j}^{(g)} \\ \tilde{x}_{ij}^{(g)} \\ \tilde{x}_{j}^{(g)} \\ \tilde{w}_{j}^{(g)} \end{bmatrix}$$

$$\tilde{\boldsymbol{\eta}}_{ij}^{(g)} = \boldsymbol{\alpha}_{j}^{(g)} + \mathbf{K}_{j}^{(g)}\mathbf{x}_{ij}^{(g)} + \mathbf{B}_{W}^{(g)}\tilde{\boldsymbol{\eta}}_{ij}^{(g)} + \boldsymbol{\varepsilon}_{ij}^{(g)}$$

$$\begin{bmatrix} \tilde{y}_{ij}^{(g)} \\ \tilde{y}_{j}^{(g)} \\ \tilde{x}_{ij}^{(g)} \\ \tilde{x}_{j}^{(g)} \\ \tilde{w}_{j}^{(g)} \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha_{j}^{y(g)} \\ 0 \\ \alpha_{j}^{x(g)} \\ \alpha_{j}^{w(g)} \end{bmatrix} + \begin{bmatrix} 0 & 0 & b_{3W}^{y(g)} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \tilde{y}_{ij}^{(g)} \\ \tilde{y}_{j}^{(g)} \\ \tilde{x}_{ij}^{(g)} \\ \tilde{x}_{j}^{(g)} \\ \tilde{w}_{j}^{(g)} \end{bmatrix} + \begin{bmatrix} \varepsilon_{ij}^{y(g)} \\ 0 \\ \varepsilon_{ij}^{x(g)} \\ 0 \\ 0 \end{bmatrix}$$

$$\tilde{\boldsymbol{\eta}}_{j}^{(g)} = \boldsymbol{\mu}^{(g)} + \mathbf{\Omega}^{(g)}\mathbf{w}_{j}^{(g)} + \mathbf{B}_{B}^{(g)}\tilde{\boldsymbol{\eta}}_{j}^{(g)} + \boldsymbol{\zeta}_{j}^{(g)}$$

$$\begin{bmatrix} \alpha_{j}^{y(g)} \\ \alpha_{j}^{x(g)} \\ \alpha_{j}^{w(g)} \end{bmatrix} = \begin{bmatrix} \mu^{y(g)} \\ \mu^{x(g)} \\ \mu^{w(g)} \end{bmatrix} + \begin{bmatrix} 0 & b_{2B}^{y(g)} & b_{1B}^{y(g)} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \alpha_{j}^{y(g)} \\ \alpha_{j}^{x(g)} \\ \alpha_{j}^{w(g)} \end{bmatrix} + \begin{bmatrix} \zeta_{j}^{y(g)} \\ \zeta_{j}^{x(g)} \\ \zeta_{j}^{w(g)} \end{bmatrix}$$

--note that the *W* and *B* subscripts only pertain to the treatment arm effects, and are deleted in the control arm's scalar expressions in the article

Notes: Residual and random effect variances were given in the manuscript.

MODEL: Empirical example model with level 2 outcome *m* and disaggregated effect of x, using latent cluster mean

$$\mathbf{y}_{ij}^{(g)} = \mathbf{\Lambda}^{(g)}\tilde{\boldsymbol{\eta}}_{ij}^{(g)}$$

$$\begin{bmatrix} y_{ij}^{(g)} \\ x_{ij}^{(g)} \\ m_j^{(g)} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{y}_{ij}^{(g)} \\ \tilde{y}_j^{(g)} \\ \tilde{x}_{ij}^{(g)} \\ \tilde{x}_j^{(g)} \\ \tilde{m}_j^{(g)} \end{bmatrix}$$

$$\tilde{\boldsymbol{\eta}}_{ij}^{(g)} = \boldsymbol{\alpha}_j^{(g)} + \mathbf{K}_j^{(g)}\mathbf{x}_{ij}^{(g)} + \mathbf{B}_W^{(g)}\tilde{\boldsymbol{\eta}}_{ij}^{(g)} + \boldsymbol{\varepsilon}_{ij}^{(g)}$$

$$\begin{bmatrix} \tilde{y}_{ij}^{(g)} \\ \tilde{y}_j^{(g)} \\ \tilde{x}_{ij}^{(g)} \\ \tilde{x}_j^{(g)} \\ \tilde{m}_j^{(g)} \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha_j^{y(g)} \\ 0 \\ \alpha_j^{x(g)} \\ \alpha_j^{m(g)} \end{bmatrix} + \begin{bmatrix} 0 & 0 & b_{3W}^{y(g)} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{y}_{ij}^{(g)} \\ \tilde{y}_j^{(g)} \\ \tilde{x}_{ij}^{(g)} \\ \tilde{x}_j^{(g)} \\ \tilde{m}_j^{(g)} \end{bmatrix} + \begin{bmatrix} \varepsilon_{ij}^{y(g)} \\ 0 \\ \varepsilon_{ij}^{x(g)} \\ 0 \\ 0 \end{bmatrix}$$

--note that the *W* and *B* subscripts only pertain to the treatment arm effects, and are deleted in the control arm's scalar expressions given at the beginning of this online appendix

$$\tilde{\boldsymbol{\eta}}_j^{(g)} = \boldsymbol{\mu}^{(g)} + \mathbf{\Omega}^{(g)}\mathbf{w}_j^{(g)} + \mathbf{B}_B^{(g)}\tilde{\boldsymbol{\eta}}_j^{(g)} + \boldsymbol{\zeta}_j^{(g)}$$

$$\begin{bmatrix} \alpha_j^{y(g)} \\ \alpha_j^{x(g)} \\ \alpha_j^{m(g)} \end{bmatrix} = \begin{bmatrix} \mu^{y(g)} \\ \mu^{x(g)} \\ \mu^{m(g)} \end{bmatrix} + \begin{bmatrix} 0 & b_{2B}^{y(g)} & b_{1B}^{y(g)} \\ 0 & 0 & 0 \\ 0 & b_{1B}^{m(g)} & 0 \end{bmatrix} \begin{bmatrix} \alpha_j^{y(g)} \\ \alpha_j^{x(g)} \\ \alpha_j^{m(g)} \end{bmatrix} + \begin{bmatrix} \zeta_j^{y(g)} \\ \zeta_j^{x(g)} \\ \zeta_j^{m(g)} \end{bmatrix}$$

Notes: Residual and random effect variances for this model were given in the first section of this online appendix.

Section D: Software Syntax

The next section of this online appendix contains M*plus* 7 (Muthén & Muthén, 1998-2012) code for simulated and empirical examples of multivariate single-level structural equation models for partial nesting (SSEM-PN) and multiple-arm multilevel structural equation models for partial nesting (MSEM-PN) in the accompanying manuscript. LISREL 8.8 (Jöreskog & Sörbom, 2006) code is also provided for example SSEM-PN models.

For SSEM-PN and MSEM-PN, the dataset structure was described in the manuscript text and footnotes, as was the coding of missing-by-design covariates for conditional likelihood specifications with exogenous predictors.

Note that *output=nochisquare* command is requested in the below M*plus* code because if a saturated model likelihood is provided by default, and used to compute absolute fit indices, these are untrustworthy. This is because they do not employ the design based constraints described in the manuscript (see also Bauer, 2003). If researchers desire to interpret absolute model fit for non-saturated SSEM-PN or MSEM-PNs, they need to separately specify a saturated model that incorporates these design-based constraints (two examples are given in the manuscript). Then they need to obtain the saturated model likelihood, along with the likelihood of the hypothesized model, to calculate absolute fit indices by hand. Note that if the *output=nochisquare* command is deleted in MSEM-PN, M*plus* may list error messages pertaining to its attempt to fit the incorrect saturated model; these can be ignored.

LISREL does not have a command analogous to *output=nochisquare*, to our knowledge, that suppresses display of a default saturated likelihood and absolute fit indices; hence those quantities outputted by default should be ignored. Interpretable versions computed using a correct saturated model can be computed by hand as described above and in the manuscript.

I. SSEM-PN Syntax in M*plus.*

M*plus* syntax for 2-arm unconditional SSEM-PN with homoscedastic residual variances across arm (Table 1)

```
DATA: file is genexwideA.dat; variances=nocheck;
VARIABLE: Names are clusterID arm y1c yt1_1-yt1_5 ;
Usevariables are y1c yt1_1-yt1_5;
missing=.;
ANALYSIS: estimator=ml;
 MODEL:
 fact1 by yt1_1-yt1_5@1;
 fact1; [fact1] (mut1);
 yt1_1-yt1_5 (1);
 [yt1_1-yt1_5@0];
 facc by y1c@1;
 facc@0; [facc] (muc);
 y1c (1);
 [y1c@0];
 facc with fact1@0;
 model constraint:
 new tx1eff ;
 tx1eff=mut1-muc;
 output: nochisquare;
```

M*plus* syntax for 2-arm unconditional SSEM-PN with heteroscedastic residual variances across arm (Table 1)

```
DATA: file is genexwideB.dat; variances=nocheck;
VARIABLE: Names are clusterID arm y1c yt1_1-yt1_5 ;
Usevariables are y1c yt1_1-yt1_5;
missing=.;
ANALYSIS: estimator=ml;
  MODEL:
  fact1 by yt1_1-yt1_5@1;
  fact1; [fact1] (mut1);
  yt1_1-yt1_5 (2);
  [yt1_1-yt1_5@0];
  facc by y1c@1;
  facc@0; [facc] (muc);
  y1c (1); [y1c@0];
  facc with fact1@0;
  model constraint:
  new tx1eff ;
  tx1eff=mut1-muc;
  output: nochisquare;
```

M*plus* syntax for 2-arm conditional SSEM-PN, homoscedastic residual variances across arm (Table 2). Shown here, as in Table 2, for effect of *x* held equal across arms (i.e. $\kappa^t = \kappa^c$).

```
DATA: file is genexwideC.dat; variances=nocheck;
VARIABLE:
Names are clusterID arm y1c yt1_1-yt1_5 x1-x5 w;
Usevariables are y1c yt1_1-yt1_5 x1-x5 w;
missing=.;
 ANALYSIS: estimator=ml;
  MODEL:
  fact1 by yt1_1-yt1_5@1;
  fact1; [fact1] (mut1);
  yt1_1-yt1_5 (1);
  [yt1_1-yt1_5@0];
  yt1_1-yt1_5 pon x1-x5 (2);
  fact1 on w;
  facc by y1c@1;
  facc@0; [facc] (muc);
  y1c (1); [y1c@0];
  facc with fact1@0;
  y1c on x1 (2);
  model constraint:
  new tx1eff;
  tx1eff=mut1-muc;
  output: nochisquare;
```

M*plus* syntax for 3-arm conditional SSEM-PN, homoscedastic residual variances across arm (Table 2). Shown here, as in Table 2, for effect of *x* held equal across arms (i.e. $\kappa^{c} = \kappa^{t_1} = \kappa^{t_2}$) and where there is clustering in both treatment arms.

```
DATA: file is genexwideD.dat; variances=nocheck;
VARIABLE:
Names are clusterID arm txt1 txt2 y1c yt1_1-yt1_5 yt2_1-yt2_2 x1-x5 w;
Usevariables are y1c yt1_1-yt1_5 yt2_1-yt2_2 x1-x5 w;
missing=.;
ANALYSIS: estimator=ml;
  MODEL:
  fact1 by yt1_1-yt1_5@1;
  fact1; [fact1] (mut1);
  yt1_1-yt1_5 (1);
  [yt1_1-yt1_5@0];
  yt1_1-yt1_5 pon x1-x5 (2);
  fact1 on w;
  fact2 by yt2_1@1 yt2_2@1;
  fact2; [fact2] (mut2);
  yt2_1-yt2_2 (1);
  [yt2_1-yt2_2@0];
  yt2_1-yt2_2 pon x1-x2 (2);
  fact2 on w;
  facc by y1c@1;
  facc@0; [facc] (muc);
  y1c (1); [y1c@0];
  facc with fact1@0;
  facc with fact2@0;
  fact1 with fact2@0;
  y1c on x1 (2);
  model constraint:
  new tx1eff tx2eff;
  tx1eff=mut1-muc;
  tx2eff=mut2-muc;
  output: nochisquare;
```

M*plus* syntax for 2-arm SSEM-PN disaggregating within vs. between effects of person-level predictor in treatment arm with Strategy B: Using exogenous observed mean of *x* (Table 3)

```
DATA: file is genexwideE.dat; variances=nocheck;
VARIABLE:
Names are clusterID arm y1c yt1_1-yt1_5 x1-x5 w x_mean;
Usevariables are y1c yt1_1-yt1_5 x1-x5 w x_mean;
missing=.; ANALYSIS: estimator=ml;
  MODEL:
  fact1 by yt1_1-yt1_5@1;
  fact1; [fact1] (mut);
  yt1_1-yt1_5 (1);
  [yt1_1-yt1_5@0];
  yt1_1-yt1_5 pon x1-x5 (withineffx);
  fact1 on w (btweffw);
```

```
  fact1 on x_mean (contexteffx);
  facc by y1c@1;
  facc@0; [facc] (muc);
  y1c (1); [y1c@0];
 facc with fact1@0;
y1c on x1 (effx);
 model constraint:
 new txeff betweeneffx;
 txeff=mut-muc;
 betweeneffx=contexteffx+withineffx;
 output: nochisquare;
```

M*plus* syntax for 2-arm SSEM-PN disaggregating within vs. between effects of person-level predictor in treatment arm: Using endogenous latent cluster mean of *x* and endogenous *w* (Table 4)

```
DATA: file is genexwideF.dat; variances=nocheck;
VARIABLE:
Names are clusterID arm y1c yt1_1-yt1_5 w xc1 xt1-xt5;
Usevariables are y1c yt1_1-yt1_5 w xc1 xt1-xt5;
 missing=.; ANALYSIS: estimator=ml;
  MODEL:
  fact1 by yt1_1-yt1_5@1;
  fact1; [fact1] (mut);
  yt1_1-yt1_5@0; [yt1_1-yt1_5@0];
  facyt1_1-facyt1_5 PON facxt1-facxt5 (withineffx);
  fact1 on facw (btweffw);
  fact1 on facxt (btwneffx);
  facxt by xt1-xt5@1;
  facxt;  [facxt];
  xt1-xt5@0; [xt1-xt5@0];
  facyt1_1 by yt1_1@1; facyt1_2 by yt1_2@1; facyt1_3 by yt1_3@1;
  facyt1_4 by yt1_4@1;  facyt1_5 by yt1_5@1;
  facyt1_1-facyt1_5 (1); [facyt1_1-facyt1_5@0];
  facxt1 by xt1@1; facxt2 by xt2@1; facxt3 by xt3@1;
  facxt4 by xt4@1; facxt5 by xt5@1;
  facxt1-facxt5 (3); [facxt1-facxt5@0];
  facyt1_1-facyt1_5 facc fact1 facw with facyt1_1-facyt1_5@0;
  facxt1-facxt5 facxt facxc facw with facxt1-facxt5@0;
  facw by w@1;
  facw; w@0; [facw]; [w@0];
  facxc by xc1@1;
  facxc; xc1@0; [facxc]; [xc1@0];
  facc by y1c@1;
  facc (1); [facc] (muc);
  y1c@0; [y1c@0];
  facc on facxc (effx);
  facc with fact1@0;
  facxc with facxt@0;
  facxc with facw@0;
  facxt with facw;
```

```
  facxc with facw@0;
  model constraint:
  new txeff;
  txeff=mut-muc;
  output: nochisquare;
```

M*plus* syntax for SSEM-PN Empirical Example model. (Empirical example combines cluster-level outcome and disaggregated between/within effects of person-level predictors in the treatment arm, using endogenous, latent cluster mean of *x*. Empirical example involves 2 arms, and heteroscedastic residual variances across arm.) (Table 5)

```
DATA: file is exwide_empirical.dat; variances=nocheck;
 VARIABLE:
 NAMES ARE clusterID arm y1c yt1-yt4 m xc1 xt1-xt4;
USEVARIABLES ARE y1c yt1-yt4 m xc1 xt1-xt4;
 MISSING ARE .;
 ANALYSIS: estimator=ml;
  MODEL:
  facyt BY yt1-yt4@1;
  facyt (varzyt); [facyt];
  facyt ON facxt (btwnx_y);
  facyt ON facm (btwnm_y);
  yt1-yt4@0; [yt1-yt4@0];
  xt1-xt4@0; [xt1-xt4@0];
  facyt1 by yt1@1; facyt2 by yt2@1; facyt3 by yt3@1; facyt4 by yt4@1;
  facyt1-facyt4 (vy); [facyt1-facyt4@0];
  facxt1 by xt1@1; facxt2 by xt2@1;  facxt3 by xt3@1; facxt4 by xt4@1;
  facxt1-facxt4 (vxt); [facxt1-facxt4@0];
  facyt1-facyt4 PON facxt1-facxt4 (withx_y);
  facyt1-facyt4 facyc facyt with facyt1-facyt4@0;
  facxt1-facxt4 facxt facxc with facxt1-facxt4@0;
  facxt BY xt1-xt4@1; facxt; [facxt];
  facm BY m@1; facm (varzm); [facm];
  facm ON facxt; m@0; [m@0];
   facyc BY yc1@1;
  facyc (vareyc); [facyc](muc_y);
   yc1@0; [yc1@0];
  facyc ON facxc (effx_yc);
   facxc BY xc1@1;
  facxc (varexc); [facxc](muc_x);
   xc1@0; [xc1@0];
  facyt WITH facyc@0;
  facyt WITH facxc@0;
  facxc WITH facxt@0;
   output: nochisquare;
```

II. SSEM-PN Syntax in LISREL

LISREL syntax for 2-arm unconditional SSEM-PN with homoscedastic residual variances across arm

```
DA NI=6 MI=-999
RA FI=genexwide1_lisrel.dat
MO NY=6 NE=2 LY=FI TY=ZE AL=FR PS=DI,FI TE=DI,FR
LE
FACC FACT1
LA
y1c yt11 yt12 yt13 yt14 yt15
VA 1 LY 1 1 LY 2 2 LY 3 2 LY 4 2 LY 5 2 LY 6 2
EQ TE 1 1 - TE 6 6
FR PS 2 2
PD
OU ND=4 ME=ML
```

LISREL syntax for 2-arm unconditional SSEM-PN with heteroscedastic residual variances across arm

```
DA NI=6 MI=-999
RA FI=genexwide2_lisrel.dat
MO NY=6 NE=2 LY=FI TY=ZE AL=FR PS=DI,FI TE=DI,FR
LE
FACC FACT1
LA
y1c yt11 yt12 yt13 yt14 yt15
VA 1 LY 1 1 LY 2 2 LY 3 2 LY 4 2 LY 5 2 LY 6 2
EQ TE 2 2 - TE 6 6
FR PS 2 2
PD
OU ND=4 ME=ML
```

LISREL syntax for 2-arm conditional SSEM-PN, homoscedastic residual variances across arm Shown here for effect of *x* held equal across arms (i.e. $\kappa^t = \kappa^c$). *Note: The SSEM-PN specification in the text for this model corresponds with the Mplus syntax, which uses a conditional likelihood and exogenous predictors. The below LISREL syntax uses a joint likelihood and endogenous predictors, yielding matching parameter estimates and standard errors but a different likelihood. Missing data on endogenous covariates do not need to be recoded as in Footnote 8.*

```
DA NI=13 MI=-999
RA FI=genexwide3_lisrel.dat
MO NY=13 NE=9 LY=FU,FI TY=ZE AL=FR PS=SY,FI TE=DI,FI BE=FU,FI
LE
FACC FACT1 FX1T FX2T FX3T FX4T FX5T FW FXC
LA
y1c yt11 yt12 yt13 yt14 yt15 x1c x1t x2t x3t x4t x5t w
VA 1 LY 1 1 LY 2 2 LY 3 2 LY 4 2 LY 5 2 LY 6 2 LY 7 9
FR TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6
EQ TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6
```

```
FR PS 2 2 PS 3 3 PS 4 4 PS 5 5 PS 6 6 PS 7 7 PS 8 8 PS 9 9
FR PS 4 3 PS 5 3 PS 5 4 PS 6 3 PS 6 4 PS 6 5 PS 7 3 PS 7 4
FR PS 7 5 PS 7 6 PS 8 3 PS 8 4 PS 8 5 PS 8 6 PS 8 7
VA 1 LY 8 3 LY 9 4 LY 10 5 LY 11 6 LY 12 7 LY 13 8
FR BE 2 8
FR LY 2 3 LY 3 4 LY 4 5 LY 5 6 LY 6 7 LY 1 9
EQ LY 2 3 LY 3 4 LY 4 5 LY 5 6 LY 6 7 LY 1 9
PD
OU ND=4 ME=ML
```

LISREL syntax for 3-arm conditional SSEM-PN, homoscedastic residual variances across arm Shown here for effect of *x* held equal across arms (i.e. $\kappa^{c} = \kappa^{t_1} = \kappa^{t_2}$) and where there is clustering in both treatment arms. *Note: The SSEM-PN specification in the text for this model corresponds with the Mplus syntax, which uses a conditional likelihood and exogenous predictors. The below LISREL syntax uses a joint likelihood and endogenous predictors, yielding matching parameter estimates and standard errors but a different likelihood. Missing data on endogenous covariates do not need to be recoded as in Footnote 8.*

```
DA NI=18 MI=-999
RA FI=genexwide4_lisrel.dat
MO NY=18 NE=13 LY=FU,FI TY=ZE AL=FR PS=SY,FI TE=DI,FI BE=FU,FI
LE
FACC FACT1 FACT2 FXC FX1T1 FX2T1 FX3T1 FX4T1 FX5T1 FX1T2 FX2T2 FWT1 FWT2
LA
y1c yt11 yt12 yt13 yt14 yt15 yt21 yt22 x1c x1t1 x2t1 x3t1 x4t1 x5t1 x1t2 x2t2 wt1 wt2
FR BE 2 12 BE 3 13
VA 1 LY 1 1 LY 2 2 LY 3 2 LY 4 2 LY 5 2 LY 6 2 LY 7 3 LY 8 3
VA 1 LY 9 4 LY 10 5 LY 11 6 LY 12 7 LY 13 8 LY 14 9 LY 15 10 LY 16 11 LY 17 12 LY 18 13
FR LY 1 4 LY 2 5 LY 3 6 LY 4 7 LY 5 8 LY 6 9 LY 7 10 LY 8 11
EQ LY 1 4 LY 2 5 LY 3 6 LY 4 7 LY 5 8 LY 6 9 LY 7 10 LY 8 11
FR PS 2 2 PS 3 3 PS 4 4 PS 5 5 PS 6 6 PS 7 7 PS 8 8 PS 9 9 PS 10 10 PS 11 11 PS 12 12 PS
13 13
FR TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6 TE 7 7 TE 8 8
EQ TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6 TE 7 7 TE 8 8 TE 1 1
FR PS 6 5 PS 7 5 PS 8 5 PS 9 5 PS 7 6 PS 8 6 PS 9 6 PS 8 7 PS 9 7 PS 9 8
FR PS 12 5 PS 12 6 PS 12 7 PS 12 8 PS 12 9
FR PS 11 10 PS 13 10 PS 13 11
PD
OU ND=4 ME=ML
```

LISREL syntax for 2-arm SSEM-PN disaggregating within vs. between effects of person-level predictor in treatment arm: Using endogenous latent cluster mean of *x* and endogenous *w*.

```
DA NI=13 MI=-999
RA FI=genexwide5_lisrel.dat
MO NY=13 NE=10 LY=FU,FI TY=ZE AL=FI PS=SY,FI TE=DI,FI BE=FU,FI
LE
FACC FACT1 FXC FX1T FX2T FX3T FX4T FX5T FW FXT
LA
y1c y1t1 y1t2 y1t3 y1t4 y1t5 x1c x1t1 x2t1 x3t1 x4t1 x5t1 w
FR LY 2 4 LY 3 5 LY 4 6 LY 5 7 LY 6 8
```

```
VA 1 LY 1 1 LY 2 2 LY 3 2 LY 4 2 LY 5 2 LY 6 2 LY 7 3
VA 1 LY 8 4 LY 9 5 LY 10 6 LY 11 7 LY 12 8 LY 13 9
EQ LY 2 4 LY 3 5 LY 4 6 LY 5 7 LY 6 8
FR PS 1 1 PS 2 2 PS 3 3 PS 4 4 PS 5 5
FR PS 6 6 PS 8 8 PS 9 9 PS 10 10 PS 10 9
EQ PS 4 4 PS 5 5 PS 6 6 PS 7 7 PS 8 8
FR TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6
EQ TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6 PS 1 1
FR BE 1 3 BE 2 9 BE 2 10
VA 1 BE 4 10 BE 5 10 BE 6 10 BE 7 10 BE 8 10
FR AL 1 AL 2 AL 3 AL 9 AL 10
ST 1 ALL
PD
OU ND=4 ME=ML
```

LISREL syntax for SSEM-PN Empirical Example model. (Empirical example combines cluster-level outcome and disaggregated between/within effects of person-level predictors in treatment arm, using endogenous latent cluster mean of *x*. Empirical example involves 2 arms, and heteroscedastic residual variances across arm.)

```
DA NI=11 MI=-999
RA FI=exwide_empirical_lisrel.dat
MO NY=11 NE=9 LY=FU,FI TY=ZE AL=FI PS=SY,FI TE=DI,FI BE=FU,FI
LE
FXT FXT1 FXT2 FXT3 FXT4 FACT1 FM FXC FACC
LA
xt1 xt2 xt3 xt4 yt1 yt2 yt3 yt4 m xc1 yc1
FR LY 5 2 LY 6 3 LY 7 4 LY 8 5
VA 1 LY 1 2 LY 2 3 LY 3 4 LY 4 5 LY 5 6 LY 6 6 LY 7 6 LY 8 6 LY 9 7 LY 10 8 LY 11 9
EQ LY 5 2 LY 6 3 LY 7 4 LY 8 5
FR BE 6 1 BE 7 1 BE 6 7 BE 9 8
VA 1 BE 2 1 BE 3 1 BE 4 1 BE 5 1
FR AL 1 AL 6 AL 7 AL 8 AL 9
FR PS 1 1 PS 2 2 PS 3 3 PS 4 4 PS 5 5 PS 6 6 PS 7 7 PS 8 8 PS 9 9
ST 1 ALL
EQ PS 2 2 PS 3 3 PS 4 4 PS 5 5
FR TE 5 5 TE 6 6 TE 7 7 TE 8 8
EQ TE 5 5 TE 6 6 TE 7 7 TE 8 8
PD
OU ND=4 ME=ML
```

III. MSEM-PN Syntax in M*plus.*

Mplus syntax for 2-arm unconditional MSEM-PN with homoscedastic residual variances across arm (Table 1)

```
DATA: file is genexlongA.dat; variances=nocheck;
VARIABLE: Names are clusterID y arm;
Usevariables are y arm;
cluster is clusterID;
grouping is arm (0=cont 1=txt1);
 missing=.;
 ANALYSIS: type=twolevel; estimator=ml;
  MODEL:
  %WITHIN%
  y (1);
  %BETWEEN%
  y;
  [y] (mut1);
  MODEL cont:
  %WITHIN%
  y (1);
   %BETWEEN%
   y@0;
  [y] (muc);
  model constraint:
  new tx1eff ;
  tx1eff=mut1-muc;
  output: nochisquare;
```

M*plus* syntax for 2-arm unconditional MSEM-PN with heteroscedastic residual variances across arm (Table 1)

```
DATA: file is genexlongB.dat; variances=nocheck;
 VARIABLE: Names are clusterID y arm;
 Usevariables are y arm;
 cluster is clusterID;
 grouping is arm (0=cont 1=txt1);
 missing=.;
 ANALYSIS: type=twolevel; estimator=ml;
  MODEL:
  %WITHIN%
  y (2);
  %BETWEEN%
  y;
  [y] (mut1);
  MODEL cont:
  %WITHIN%
  y (1);
   %BETWEEN%
   y@0;
  [y] (muc);
```

```
 model constraint:
 new tx1eff ;
 tx1eff=mut1-muc;
 output: nochisquare;
```

M*plus* syntax for 2-arm conditional MSEM-PN, homoscedastic residual variances across arm (Table 2). Shown here, as in Table 2, for effect of *x* held equal across arms (i.e. $\kappa^t = \kappa^c$).

```
DATA: file is genexlongC.dat; variances=nocheck;
VARIABLE:
 Names are clusterID y x w arm;
 Usevariables are y arm x w;
 cluster is clusterID;
 grouping is arm (0=cont 1=txt1);
 missing=.;
 WITHIN is x;
 BETWEEN IS w;
 ANALYSIS: type=twolevel; estimator=ml;
  MODEL:
  %WITHIN%
  y (1); y on x (2);
  %BETWEEN%
  y;  y on w;
  [y] (mut1);
  MODEL cont:
  %WITHIN%
  y (1); y on x (2);
   %BETWEEN%
   y@0;
  y on w@0;
  [y] (muc);
  model constraint:
  new tx1eff ;
  tx1eff=mut1-muc;
  output: nochisquare;
```

M*plus* syntax for 3-arm conditional MSEM-PN, homoscedastic residual variances across arm (Table 2). Shown here, as in Table 2, for effect of *x* held equal across arms (i.e. $\kappa^c = \kappa^{t_1} = \kappa^{t_2}$) and where there is clustering in both treatment arms.

```
DATA: file is genexlongD.dat; variances=nocheck;
 VARIABLE:
 Names are clusterID y x w arm;
 Usevariables are y arm x w;
 cluster is clusterID;
 grouping is arm (0=cont 1=txt1 2=txt2 );
 missing=.;
 WITHIN is x;
 BETWEEN IS w;
  ANALYSIS: type=twolevel; estimator=ml;
  MODEL:
```

```
 %WITHIN%
 y (1); y on x (2);
 %BETWEEN%
 y on w; y;
 [y] (mut2);
 MODEL cont:
 %WITHIN%
 y (1); y on x (2);
  %BETWEEN%
  y@0; y on w@0;
 [y] (muc);
 model txt1:
 %WITHIN%
  y (1); y on x (2);
 %BETWEEN%
 y; y on w (3);
 [y] (mut1);
 model constraint:
 new tx1eff tx2eff;
 tx1eff=mut1-muc;
 tx2eff=mut2-muc;
 output: nochisquare;
```

M*plus* syntax for 2-arm MSEM-PN disaggregating within vs. between effects of person-level predictor in treatment arm with Strategy A: Using exogenous observed mean of *x* (Table 3).

```
DATA: file is genexlongE.dat; variances=nocheck;
VARIABLE:
Names are clusterID y x x_grpmc x_mean w arm;
Usevariables are y arm w x x_mean x_grpmc;
 grouping is arm (0=cont 1=txt);
 cluster is clusterID;
 missing=.;
 WITHIN is x x_grpmc;
 BETWEEN IS w x_mean;
 ANALYSIS: type=twolevel; estimator=ml;
 MODEL:
 %WITHIN%
 y (1); y on x_grpmc (within);
y on x@0;
  %BETWEEN%
  y;  y on w;
  y on x_mean (between);
  [y] (mut1);
  MODEL cont:
  %WITHIN%
  y (1); y on x (2);
  y on x_grpmc@0;
  %BETWEEN%
  y@0; y on w@0;
  y on x_mean@0;
```

```
  [y] (muc);
  output: nochisquare;
  model constraint:
  new txeff context;
  txeff=mut1-muc;
  context=between-within;
```

Note: The below alternative syntax provides identical results to the specification above, but is slightly simpler.

```
DATA: file is genexlongE.dat; variances=nocheck;
VARIABLE:
Names are clusterID y x_grpmc x_mean w arm;
Usevariables are y arm w x_mean x_grpmc;
 grouping is arm (0=cont 1=txt);
 cluster is clusterID;
 missing=.;
 WITHIN is x_grpmc;
 BETWEEN IS w x_mean;
 ANALYSIS: type=twolevel; estimator=ml;
 MODEL:
 %WITHIN%
 y (1); y on x_grpmc (within);
  %BETWEEN%
  y;  y on w;
  y on x_mean (between);
  [y] (mut1);
  MODEL cont:
  %WITHIN%
  y (1); y on x_grpmc (2);
  %BETWEEN%
  y@0; y on w@0;
  y on x_mean (2);
  [y] (muc);
  output: nochisquare;
  model constraint:
  new txeff context;
  txeff=mut1-muc;
  context=between-within;
```

M*plus* syntax for 2-arm MSEM-PN disaggregating within vs. between effects of person-level predictor in treatment arm: Using endogenous latent cluster mean of *x* and endogenous *w* (Table 4).

Note: For MSEM-PN using a latent cluster mean of x, x is no longer declared a WITHIN variable using WITHIN=x. When x is not mentioned as a WITHIN variable, Mplus anticipates having both within and between variability in x in both arms, though in the control arm, there is no between-variability in x. To address this, a constraint can be placed on the effect of x in the control arm to declare the within and between effects equal (a.k.a. just a simple/total effect of x for controls, as shown below) or the effect of x can be set to 0 at one level in the control arm.

```
DATA: file is genexlongF.dat; variances=nocheck;
VARIABLE: Names are clusterID y x w arm;
Usevariables are y x w ;
cluster is clusterID;
grouping is arm (0=cont 1=txt1);
missing=.;
BETWEEN IS w;
ANALYSIS: type=twolevel; estimator=ml;
  MODEL:
  %WITHIN%
  y (1);  y on x (withineffx); x;
  %BETWEEN%
  y;   y on w;
  y on x (betweeneffx);
  [y] (mut1);  [ w x];  w x;  w with x;
  MODEL cont:
  %WITHIN%
  y@0; y on x (effx); x@0;
   %BETWEEN%
   y (1); y on w@0; x;
  y on x (effx);
  [y] (muc); [w@0  x]; w@0; x with w@0;
  output: nochisquare;
```

M*plus* syntax for MSEM-PN Empirical Example model. (Empirical example combines cluster-level outcome and disaggregated between/within effects of person-level predictors in treatment arm, using endogenous latent cluster mean of *x*. Empirical example involves 2 arms, and heteroscedastic residual variances across arm) (Table 5). *Note: For MSEM-PN using a latent cluster mean of x, x is no longer declared a WITHIN variable using WITHIN=x. When x is not mentioned as a WITHIN variable, Mplus anticipates having both within and between variability in x in both arms, though in the control arm, there is no between variability in x. To address this, a constraint can be placed on the effect of x in the control arm to declare the within and between effects equal (a.k.a. just a simple/total effect of x for controls), or the effect of x can be set to 0 at one level in the control arm (as shown below).*

```
DATA: FILE IS exlong_empirical.dat; variances=nocheck;
VARIABLE: NAMES ARE clusterID kidID arm y x m;
USEVARIABLES ARE y x m; CLUSTER IS clusterID;
GROUPING IS arm (0=cont 1=txt);
MISSING ARE .;
BETWEEN IS m;
ANALYSIS: type=twolevel; estimator=ml;
MODEL:
%WITHIN%
y (vareyt); x;
y ON x (withx_y);
```

```
%BETWEEN%
y (varzyt); m (varzm); x;
y ON m (btwm_y);
m ON x;
y ON x (btwx_y);
[y x m];
MODEL cont:
%WITHIN%
y@0; x@0; y ON x@0;
%BETWEEN%
y ON m@0; m ON x@0;
y; x; m@0;
y ON x (effx_yc);
[y x m@0];
output: nochisquare;
```