Dynamic Structural Equation Modeling of Intensive Longitudinal Data

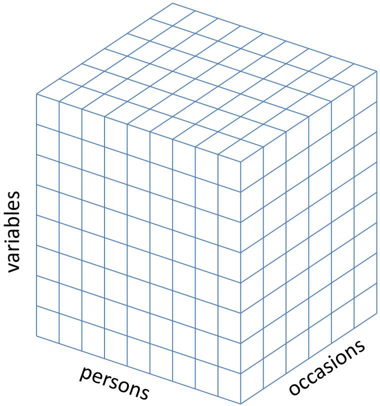
Part 1

Ellen L. Hamaker
Utrecht University
e.l.hamaker@uu.nl

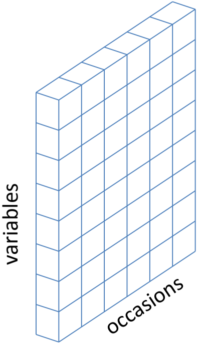
July, 2017

In collaboration with Bengt Muthén and Tihomir Asparouhov

Cattell's data box



Time series data: N=1 and T is large



Idiographic (N=1) research in psychology

N=1 research has included:

- Cattell's P-technique: factor analysis of N=1 data
- Dynamic factor analysis: considering lagged relationships
- Measurement burst design: multiple waves of intensive measurements
- Intervention research: ABAB design etc.

Critique of this kind of research:

- within-person fluctuations are just **noise**
- results are **not generalizable**
- no one has these data

New technology

Smart phones

Smart glasses

Secure continuous remote
alcohol monitor (SCRAM)

Smart watches

Activity trackers

Intensive longitudinal data

Different forms of intensive longitudinal data:

- daily diary (DD); self-report end-of-day
- experience sampling method (ESM); self-report of subjective experience
- ecological momentary assessment (EMA); healthcare related self-report
- ambulatory assessment (AA); physiological measurements
- event-based measurements; self-report after a particular event
- observational measurements; expert rater

For more info on **methodology**, check out:

- Seminar of Tamlin Conner and Joshua Smyth on YouTube (https://www.youtube.com/watch?v=nQBBVp9vBIQ)
- Society for Ambulatory Assessment (http://www.saa2009.org/)
- Life Data (https://www.lifedatacorp.com/)
- Quantified Self (http://quantifiedself.com/)

Characteristics of these kind of data

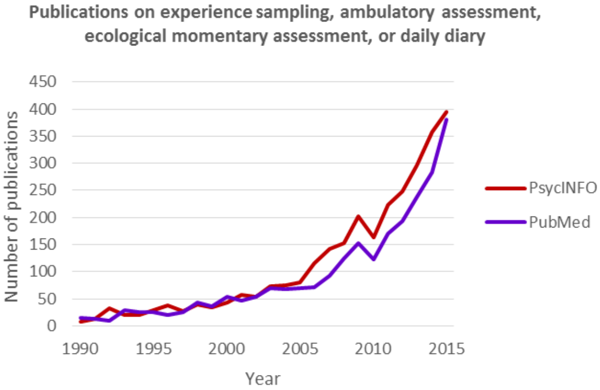
Data structure:

- one or more measurements per day
- typically for multiple days
- sometimes multiple waves (i.e., Nesselroade's measurement-burst design)

Advantages of ESM, EMA and AA

- no recall bias
- high ecological validity
- physiological measures over a large time span
- monitoring of symptoms and behavior, with new possibilities for feedback and intervention (e-Health and m-Health)
- window into the dynamics of processes

A paradigm shift



Taken from Hamaker and Wichers (2017)

Outline

- **Time series analysis**
- Multilevel time series analysis
- DSEM application 1: Multilevel VAR(1) model
- DSEM application 2: Mediation
- Discussion

What is time series analysis?

Time series analysis is a class of techniques that is used in econometrics, seismology, meteorology, control engineering, and signal processing.

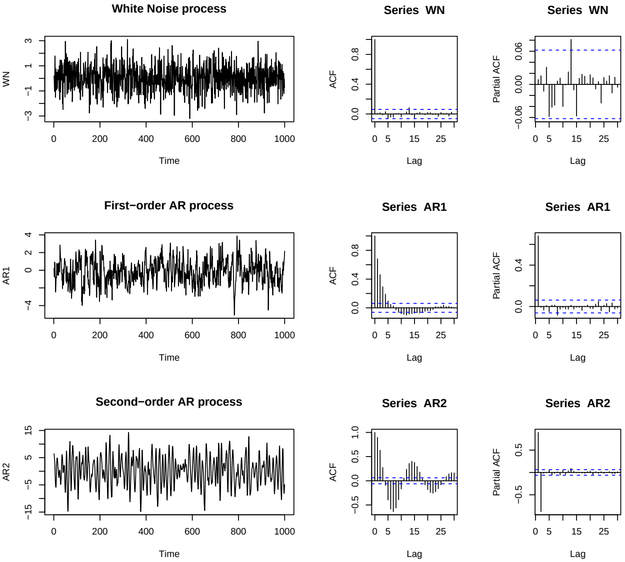
Main characteristics:

- N=1 technique
- T is large (say $>$50)
- concerned with *trends*, *cycles* and *autocorrelation structure* (i.e., serial dependency)
- goal: forecasting ($\neq$ prediction)

Lags

Y	Y at lag 1	Y at lag 2
y_1		
y_2	y_1	
y_3	y_2	y_1
y_4	y_3	y_2
y_5	y_4	y_3
y_6	y_5	y_4
y_7	y_6	y_5
y_8	y_7	y_6
$\dots$	$\dots$	$\dots$
y_T	y_{T-1}	y_{T-2}
	y_T	y_{T-1}
		y_T

Sequence, ACF and PACF



Outline

- Time series analysis
- **Multilevel time series analysis**
- DSEM application 1: Multilevel VAR(1) model
- DSEM application 2: Mediation
- Discussion

Lagged relationships in multilevel data

If we have t**ime series data from multiple individuals**, we may want to study:

- individual differences in lagged relationships between a variable and itself: **autoregression**
- individual differences in lagged relationship between different variables: **cross-lagged relationships**

If we use multilevel modeling for this, we could refer to it as **multilevel time series analysis**, or **dynamic multilevel modeling.**

Creating lagged predictors

ID	y_{it}	y_{it-1}	x_{it-1}
1	y_{11}		
1	y_{12}	y_{11}	x_{11}
1	y_{13}	y_{12}	x_{12}
1	$\dots$	$\dots$	$\dots$
1	y_{1T}	y_{1T-1}	x_{1T-1}
2	y_{21}		
2	y_{22}	y_{21}	x_{21}
2	y_{23}	y_{22}	x_{22}
2	$\dots$	$\dots$	$\dots$
2	y_{2T}	y_{2T-1}	x_{2T-1}
$\dots$	$\dots$	$\dots$	$\dots$
N	y_{N1}		
N	y_{N2}	y_{N1}	x_{N1}
N	y_{N3}	y_{N2}	x_{N2}
N	$\dots$	$\dots$	$\dots$
N	y_{NT}	y_{NT-1}	x_{NT-1}

Inertia research based on multilevel AR(1) models

Level 1 model:

$$NA_{it} = c_i + \phi_i NA_{i,t-1} + \zeta_{it}$$

Level 2 model:

$$c_i = \gamma_{00} + u_{0i}$$
$$\phi_i = \gamma_{01} + u_{1i}$$

This research line was initiated by **Suls, Green and Hillis (1998)**, and continued by the group of **Kuppens.**

The focus is on individual differences in the **autoregressive parameter** ϕ_i (=inertia, carry-over, regulatory weakness), which is shown to be:

- positively related to current depression, neuroticism, and being female
- predictive of later depression (Kuppens and Koval)

Dynamic networks based on multilevel VAR(1) models

Level 1 model:

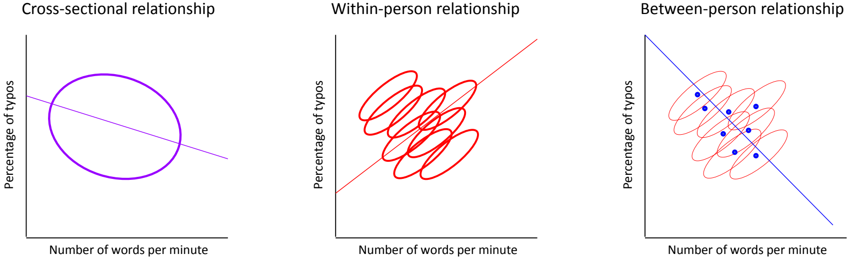
$$
\begin{aligned}
y_{1it} &= c_{1i} + \phi_{11i} y_{1it-1} + \cdots + \phi_{1ki} y_{kit-1} + \zeta_{1it} \\
y_{2it} &= c_{2i} + \phi_{21i} y_{1it-1} + \cdots + \phi_{2ki} y_{kit-1} + \zeta_{2it} \\
\ldots \\
y_{kit} &= c_{ki} + \phi_{k1i} y_{1it-1} + \cdots + \phi_{kki} y_{kit-1} + \zeta_{kit}
\end{aligned}
$$

Initiated by **Bringmann et al. (2013)**, and further popularized by the software from **Sacha Epskamp.**

The focus is on **cross-lagged parameters** between variables (=nodes; typically symptoms), and on measures based on these (e.g., centrality).

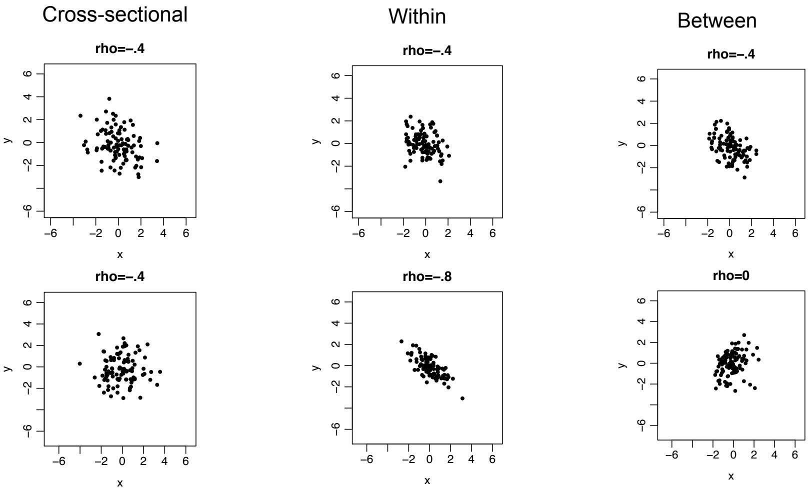
Main idea is that **stronger connections** lead to an **increased risk** of developing and maintaining psychopathology.

A fundamental problem in a nutshell



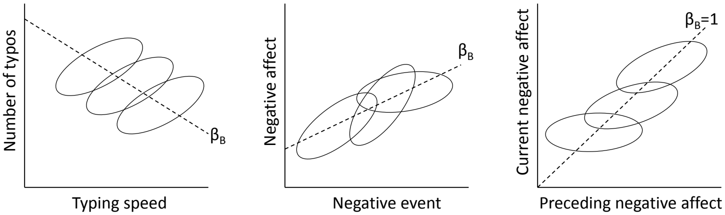
Taken from Hamaker (2012).

Three perspectives on data



Taken from Hamaker (2012).

Between-person differences in within-person slopes



Taken from Hamaker and Grasman (2014).

In conclusion: To study within-person processes we need

- (intensive) **longitudinal** data
- to **decompose** observed variance into within and between
- to consider **individual differences** in within-person dynamics

Estimating the multilevel AR(1) model

When **estimating** the multilevel AR(1) model, we can decide to:

- **not center** the lagged predictor (NC)
- center with the **sample mean** $\bar{y}_{\cdot i}$
- center with the **estimated mean from an empty multilevel model** $\hat{\mu}_i$
- center with the **true mean** $\hat{\mu}_i$ (in case of simulations)

Table 4 | Bias and coverage rates for fixed autoregressive parameter ϕ in multilevel autoregressive model under diverse scenarios.

AR parameter	Sample size		Bias				$CR_{0.95}$			
	N	T	NC	$C(\bar{y}_{\cdot i})$	$C(\hat{\mu}_i)$	$C(\mu_i)$	NC	$C(\bar{y}_{\cdot i})$	$C(\hat{\mu}_i)$	$C(\mu_i)$
$\phi_i \sim N(0.3, 0.1)$	20	20	0.002	−0.072	−0.069	−0.068	0.928	0.762	0.785	0.787
		50	0.000	−0.027	−0.027	−0.026	0.940	0.900	0.901	0.898
		100	0.000	−0.013	−0.013	−0.013	0.932	0.932	0.932	0.932
	50	20	0.005	−0.071	−0.069	−0.067	0.893	0.480	0.512	0.518
		50	0.001	−0.027	−0.026	−0.026	0.936	0.800	0.804	0.805
		100	0.000	−0.013	−0.013	−0.013	0.946	0.902	0.902	0.903
	100	20	0.006	−0.070	−0.068	−0.066	0.892	0.196	0.227	0.242
		50	0.001	−0.027	−0.027	−0.027	0.930	0.623	0.630	0.637
		100	0.000	−0.013	−0.013	−0.013	0.930	0.851	0.854	0.851

Disadvantages of using regular multilevel software

If we are interested in **dynamic multilevel modeling**, we may run into the following problems/limitation when using **standard multilevel software**:

- **negative bias in autoregression** when centering the lagged predictor (Nickell's bias)
- only **one outcome variable** (thus, separate models for multivariate outcomes)
- only **observed variables** (no measurement error, moving average terms, factor models)
- **missing data** result in many missing cases
- **unequally spaced** observations

Dynamic structural equation modeling (DSEM) in Mplus tackles all these problems.

Outline

- Time series analysis
- Multilevel time series analysis
- **DSEM application 1: Multilevel VAR(1) model**
- DSEM application 2: Mediation
- Discussion

Data: Daily measurements affect

Data come from the **COGITO study** of the MPI in Berlin; goal is to study aging using a younger and older sample.

Analyses here are based on Hamaker et al. (in preparation).

Characteristics of the **younger** and **older sample**:

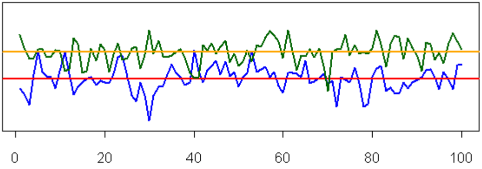
- aged 20-31; aged 65-80
- 101 individuals; 103 individuals
- about 100 daily measurements of positive affect (PA) and negative affect (NA)

Decomposition into a between part and a within part

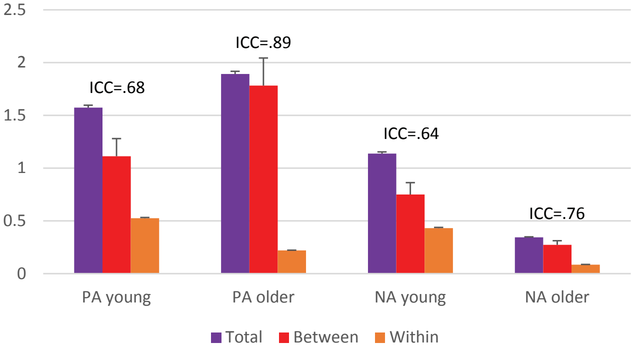
$$PA_{it} = \mu_{PA,i} + PA^*_{it}$$
$$NA_{it} = \mu_{NA,i} + NA^*_{it}$$

where

- $\mu_{PA,i}$ and $\mu_{NA,i}$ are the individual's **means** on PA and NA (i.e., baseline, trait, or equilibrium scores) $\Rightarrow$ between-person part
- PA^*_{it} and NA^*_{it} are the **within-person centered** (cluster-mean centered) scores $\Rightarrow$ within-person part



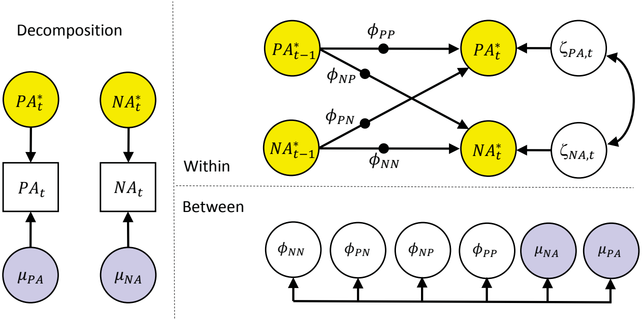
Total, between-, and within-person variance



Intraclass correlation:

$$\frac{\sigma^2_{between}}{\sigma^2_{between}+\sigma^2_{within}} = \frac{\sigma^2_{between}}{\sigma^2_{total}}$$

Bivariate model: Multilevel vector AR(1) model



Within-person level model

Lagged within-person model:

$$PA^*_{it} = \phi_{PP,i} PA^*_{i,t-1} + \phi_{PN,i} NA^*_{i,t-1} + \zeta_{PA,it}$$
$$NA^*_{it} = \phi_{NN,i} NA^*_{i,t-1} + \phi_{NP,i} PA^*_{i,t-1} + \zeta_{NA,it}$$

where

- $\phi_{PP,i}$ is the **autoregressive parameter** for PA (i.e., inertia, carry-over)
- $\phi_{NN,i}$ is the **autoregressive parameter** for NA (i.e., inertia, carry-over)
- $\phi_{PN,i}$ is the **cross-lagged parameter** for NA to PA (i.e., spill-over)
- $\phi_{NP,i}$ is the **cross-lagged parameter** for PA to NA (i.e., spill-over)
- $\zeta_{PA,it}$ is the **innovation** for PA (residual, disturbance, dynamic error)
- $\zeta_{NA,it}$ is the **innovation** for NA (residual, disturbance, dynamic error)

Parameters estimated at this level are the residual variances and covariance:

$$\begin{bmatrix} \zeta_{PA,it} \\ \zeta_{NA,it} \end{bmatrix} \sim MN \left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \theta_{11} & \\ \theta_{21} & \theta_{22} \end{bmatrix} \right]$$

Between-person level model

Between level: fixed and random effects

$$\begin{bmatrix} \mu_{PA,i} \\ \mu_{NA,i} \\ \phi_{PP,i} \\ \phi_{PN,i} \\ \phi_{NP,i} \\ \phi_{NN,i} \end{bmatrix} = \begin{bmatrix} \gamma_P \\ \gamma_N \\ \gamma_{PP} \\ \gamma_{PN} \\ \gamma_{NP} \\ \gamma_{NN} \end{bmatrix} + \begin{bmatrix} u_{P,i} \\ u_{N,i} \\ u_{PP,i} \\ u_{PN,i} \\ u_{NP,i} \\ u_{NN,i} \end{bmatrix} \quad \boldsymbol{u}_i \sim MN(0, \boldsymbol{\Psi})$$

Where:

- γ_P to γ_{NN} $\Rightarrow$ fixed effects
- $u_{P,i}$ to $u_{NN,i}$ $\Rightarrow$ random effects

Parameters estimated at this level are:

- 6 fixed effects (i.e., γ's)
- 6 variances for random effects (i.e., diagonal elements of $\boldsymbol{\Psi}$)
- 15 covariances between the random effects (i.e., off-diagonal elements in $\boldsymbol{\Psi}$)

Bivariate model: Mplus code

```
VARIABLE:   NAMES ARE id sessdate
            na1 na2 na3 na4 na5 na6 na7 na8 na9 na10
            pa1 pa2 pa3 pa4 pa5 pa6 pa7 pa8 pa9 pa10
            sessionNr age_pre sex CESDpre CESDpost dayNA dayPA older;

            CLUSTER = id; ! Specify the person id variable
            USEVAR = dayPA dayNA; ! Specify which variables are used in the model
            MISSING = ALL(-999);

            LAGGED = dayPA(1) dayNA(1); ! This creates lagged variables
            TINTERVAL = sessdate(1); ! This is to account for unequal intervals

ANALYSIS:   TYPE IS TWOLEVEL RANDOM; ! This allows for random slopes
            ESTIMATOR = BAYES; ! DSEM requires Bayesian estimation
            PROC = 2; ! Using 2 processors makes it faster
            BITER = (5000); ! This implies at least 5000 iterations are used
            THIN = 10; ! Thinning helps with getting more stable results
```

Bivariate model: Mplus code

```
MODEL:      %WITHIN% ! Specify the random lagged relationships
            p_pp | dayPA ON dayPA&1;
            p_pn | dayPA ON dayNA&1;
            p_np | dayNA ON dayPA&1;
            p_nn | dayNA ON dayNA&1;

            %BETWEEN% ! Allow all 6 random effects to be correlated
            p_pp WITH p_pn-p_nn dayPA dayNA;
            p_pn WITH p_np-p_nn dayPA dayNA;
            p_np WITH p_nn dayPA dayNA;
            p_nn WITH dayPA dayNA;
            dayPA WITH dayNA;

OUTPUT:     TECH1 TECH8 STDYX;

PLOT:       TYPE = PLOT3;
            FACTORS =ALL;
```

Mplus results: Within-person (younger sample)

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

Within Level

 DAYNA     WITH
    DAYPA            -0.069       0.004      0.000      -0.076      -0.061      *

 Residual Variances
    DAYPA             0.414       0.006      0.000       0.403       0.426      *
    DAYNA             0.302       0.004      0.000       0.294       0.311      *
```

Mplus results: Between-person (younger sample)

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	Upper 2.5%	Significance
[...]						
Between Level						
[...]						
Means						
DAYPA	3.090	0.110	0.000	2.875	3.308	*
DAYNA	0.977	0.077	0.000	0.826	1.128	*
P_PP	0.334	0.026	0.000	0.283	0.387	*
P_PN	0.050	0.022	0.016	0.006	0.093	*
P_NP	0.038	0.015	0.006	0.008	0.068	*
P_NN	0.370	0.027	0.000	0.315	0.423	*
Variances						
DAYPA	1.178	0.189	0.000	0.886	1.618	*
DAYNA	0.595	0.101	0.000	0.443	0.832	*
P_PP	0.055	0.010	0.000	0.039	0.079	*
P_PN	0.024	0.006	0.000	0.014	0.039	*
P_NP	0.013	0.003	0.000	0.008	0.021	*
P_NN	0.062	0.012	0.000	0.044	0.089	*

Comparing cross-lagged parameters

Standardization in multilevel models is a **tricky issue.**

Schuurman, Ferrer, Boer-Sonnenschein and Hamaker (2016) discuss four forms of **standardization in multilevel models**, using:

- total variance (i.e., grand standardization)
- between-person variance (i.e., between standardization)
- average within-person variance
- within-person variance (i.e., within standardization)

Conclusion: last form is most meaningful, as it **parallels standardizing when N=1.**

Standardized fixed effect should be the **average standardized within-person effect.**

Mplus standardized results (younger sample)

```
STDYX Standardization

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

Within-Level Standardized Estimates Averaged Over Clusters

 P_PP | DAYPA ON
    DAYPA&1            0.335       0.011      0.000       0.312       0.358      *

 P_PN | DAYPA ON
    DAYNA&1            0.034       0.013      0.006       0.008       0.059      *

 P_NP | DAYNA ON
    DAYPA&1            0.038       0.011      0.000       0.017       0.059      *

 P_NN | DAYNA ON
    DAYNA&1            0.370       0.012      0.000       0.347       0.394      *

 DAYNA    WITH
    DAYPA             -0.194       0.010      0.000      -0.213      -0.175      *

 Residual Variances
    DAYPA              0.816       0.008      0.000       0.799       0.832      *
    DAYNA              0.792       0.008      0.000       0.775       0.808      *
```

Mplus standardized results (younger sample)

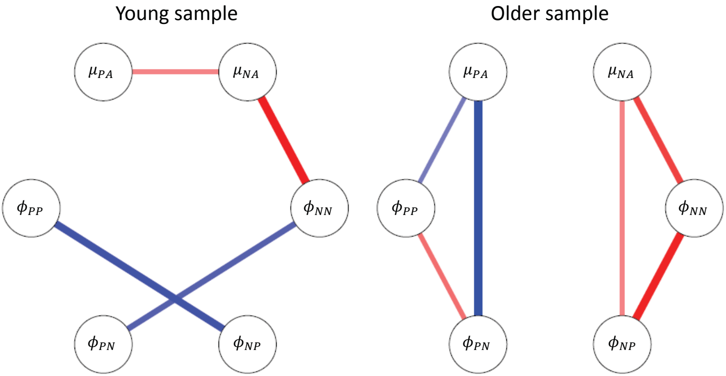
```
R-SQUARE

Within-Level R-Square Averaged Across Clusters

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

    DAYPA              0.184      0.008      0.000       0.168       0.201
    DAYNA              0.208      0.008      0.000       0.192       0.225
```

To **represent the correlation matrices** of the 6 random effects in each group, we can use the network representation (with qgraph from Sacha Epskamp in R):



Outline

- Time series analysis
- Multilevel time series analysis
- DSEM application 1: Multilevel VAR(1) model
- **DSEM application 2: Mediation**
- Discussion

Including level 2 predictor and outcome

Depression was measured prior to the ILD phase and afterwards, using the CESD; we include these measures at the between-person level as a **predictor** and an **outcome.**

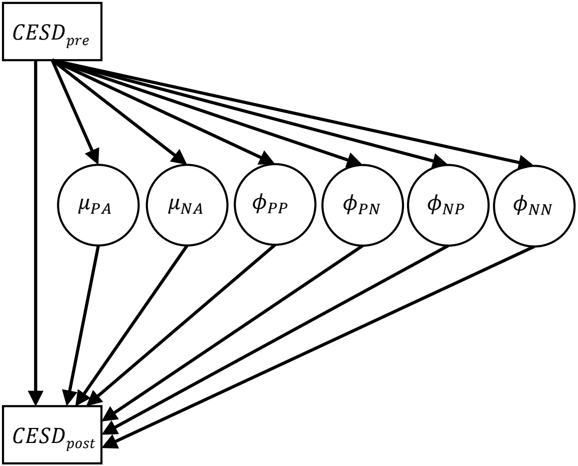
Between level: Including a level 2 predictor

$$\mu_{PA,i} = \gamma_{00} + \gamma_{01} CESDpre_i + u_{0i}$$
$$\mu_{NA,i} = \gamma_{10} + \gamma_{11} CESDpre_i + u_{1i}$$
$$\phi_{PP,i} = \gamma_{20} + \gamma_{21} CESDpre_i + u_{2i}$$
$$\phi_{PN,i} = \gamma_{30} + \gamma_{31} CESDpre_i + u_{3i}$$
$$\phi_{NN,i} = \gamma_{40} + \gamma_{41} CESDpre_i + u_{4i}$$
$$\phi_{NP,i} = \gamma_{50} + \gamma_{51} CESDpre_i + u_{5i}$$

Between level: Including a level 2 outcome

$$CESDpost_i = \gamma_{60} + \gamma_{61} CESDpre_i + \gamma_{62}\mu_{PA,i} + \gamma_{63}\mu_{NA,i} + \gamma_{64}\phi_{PP,i} + \gamma_{65}\phi_{PN,i} + \gamma_{66}\phi_{NN,i} + \gamma_{67}\phi_{NP,i} + u_{6i}$$

Dynamic mediation model



Mplus input mediation model

```
VARIABLE:    NAMES ARE id sessdate
             na1 na2 na3 na4 na5 na6 na7 na8 na9 na10
             pa1 pa2 pa3 pa4 pa5 pa6 pa7 pa8 pa9 pa10
             sessionNr age_pre sex CESDpre CESDpost dayNA dayPA older;
             CLUSTER = id;
             USEVAR = dayPA dayNA CESDpre CESDpost; ! Plus level 2 variables
             BETWEEN = CESDpre CESDpost; ! Specify these as level 2 variables
             LAGGED = dayPA(1) dayNA(1);
             TINTERVAL = sessdate(1);
             MISSING = ALL(-999);

DEFINE:      CENTER CESDpre CESDpost (GRANDMEAN);! Grand mean centering

ANALYSIS:    TYPE IS TWOLEVEL RANDOM;
             ESTIMATOR = BAYES;
             PROCESSORS = 2;
             BITER = (5000);
             THIN = 10;
```

Bivariate model: Mplus code

```
MODEL:              %WITHIN% ! Same as before
                    p_pp | dayPA ON dayPA&1;
                    p_pn | dayPA ON dayNA&1;
                    p_np | dayNA ON dayPA&1;
                    p_nn | dayNA ON dayNA&1;

                    %BETWEEN% ! Mediation model with parameter names
                    p_pp-p_nn dayPA dayNA ON CESDpre (a1-a6);
                    CESDpost ON p_pp-p_nn dayPA dayNA CESDpre (b1-b7);

MODEL CONSTRAINT:   ! Compute the indirect effects
                    new (ab_p_pp); ab_p_pp=a1*b1;
                    new (ab_p_pn); ab_p_pn=a2*b2;
                    new (ab_p_np); ab_p_np=a3*b3;
                    new (ab_p_nn); ab_p_nn=a4*b4;
                    new (ab_dayPA); ab_dayPA=a5*b5;
                    new (ab_dayNA); ab_dayNA=a6*b6;

OUTPUT:             TECH1 TECH8 STDYX;

PLOT:               TYPE = PLOT3;
                    FACTOR =ALL;
```

Mplus output mediation model (younger sample)

```
                                Posterior  One-Tailed         95% C.I.
                    Estimate       S.D.      P-Value   Lower 2.5%  Upper 2.5%  Significance
[...]
Between Level
[...]

 Intercepts
    CESDPOST           0.104       0.136      0.223      -0.167       0.365
    DAYPA              3.088       0.103      0.000       2.888       3.293      *
    DAYNA              0.989       0.076      0.000       0.844       1.146      *
    P_PP               0.338       0.024      0.000       0.289       0.386      *
    P_PN               0.031       0.020      0.057      -0.008       0.071
    P_NP               0.035       0.014      0.006       0.007       0.062      *
    P_NN               0.376       0.024      0.000       0.329       0.423      *

 Residual Variances
    CESDPOST           0.067       0.012      0.000       0.048       0.095      *
    DAYPA              1.049       0.158      0.000       0.798       1.416      *
    DAYNA              0.517       0.091      0.000       0.377       0.729      *
    P_PP               0.045       0.008      0.000       0.032       0.064      *
    P_PN               0.019       0.005      0.000       0.011       0.030      *
    P_NP               0.010       0.003      0.000       0.005       0.016      *
    P_NN               0.043       0.008      0.000       0.031       0.062      *

New/Additional Parameters
    AB_P_PP            0.010       0.025      0.266      -0.028       0.076
    AB_P_PN           -0.002       0.032      0.439      -0.074       0.062
    AB_P_NP           -0.004       0.037      0.401      -0.089       0.067
    AB_P_NN            0.195       0.070      0.000       0.081       0.359      *
    AB_DAYPA           0.049       0.035      0.029      -0.001       0.135
    AB_DAYNA           0.028       0.043      0.234      -0.052       0.119
```

Mplus output mediation model (older sample)

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%	Significance
[...]						
Between Level						
[...]						
Intercepts						
CESDPOST	0.015	0.113	0.448	-0.210	0.236	
DAYPA	4.566	0.120	0.000	4.336	4.796	*
DAYNA	0.313	0.052	0.000	0.210	0.417	*
P_PP	0.421	0.026	0.000	0.370	0.472	*
P_PN	0.133	0.039	0.000	0.057	0.212	*
P_NP	0.016	0.017	0.167	-0.018	0.051	
P_NN	0.239	0.027	0.000	0.185	0.291	*
Residual Variances						
CESDPOST	0.039	0.006	0.000	0.029	0.053	*
DAYPA	1.416	0.221	0.000	1.079	1.918	*
DAYNA	0.269	0.041	0.000	0.203	0.365	*
P_PP	0.056	0.010	0.000	0.039	0.079	*
P_PN	0.083	0.021	0.000	0.051	0.131	*
P_NP	0.024	0.004	0.000	0.018	0.035	*
P_NN	0.051	0.009	0.000	0.037	0.072	*
New/Additional Parameters						
AB_P_PP	0.005	0.016	0.302	-0.018	0.049	
AB_P_PN	-0.004	0.025	0.396	-0.061	0.045	
AB_P_NP	0.012	0.027	0.268	-0.035	0.076	
AB_P_NN	-0.036	0.038	0.112	-0.130	0.025	
AB_DAYPA	0.028	0.038	0.209	-0.042	0.110	
AB_DAYNA	0.027	0.036	0.194	-0.040	0.108	

Outline

- Time series analysis
- Multilevel time series analysis
- DSEM application 1: Multilevel VAR(1) model
- DSEM application 2: Mediation
- **Discussion**

Advantages of using DSEM in Mplus (thus far)

Compared to standard multilevel software:

- **Multiple outcome variables**: this allows for correlated residuals and correlated random effects
- **Unequal time interval**: can be handled by choosing a grid for inserting missings
- **Outcomes** at between-person level
- **Person-mean centering** integral part of model estimation (solves Nickell's bias)

Compared to other Bayesian software (e.g., WinBUGS, jags, Stan):

- **Easy to use** due to tailor-made code
- **Default uninformative priors** for parameters (even for small variances)
- **Fast** (which makes a difference in case of Bayes)

Other recent developments: mlVAR, ctsem and open Mx (in R); Bayesian Ornstein-Uhlenbeck Model (BOUM); GIMME.

More advantages of using DSEM in Mplus

Other options offered by DSEM in Mplus version 8:

- **Diverse plotting options**: allows for inspection of data and results
- **Latent variables**: allows for measurement error to be split off and for moving average terms
- **Cross-classified models**: allows for random effects of time
- **Random variance**: allows for individual difference in variability

Future options Mplus will offer:

- **Regime-switching models**: allows for a process to switch between distinct states
- **Residual dynamic modeling**: allows for easy combination of time trends and residual lagged relationships

Random innovation variance (univariately)

Within level: AR(1) with random ϕ_i

$$NA^*_{it} = \phi_i NA^*_{i,t-1} + \zeta_{it} \qquad \zeta_{it} \sim N(0, \sigma_i^2)$$

Between level: fixed and random effects

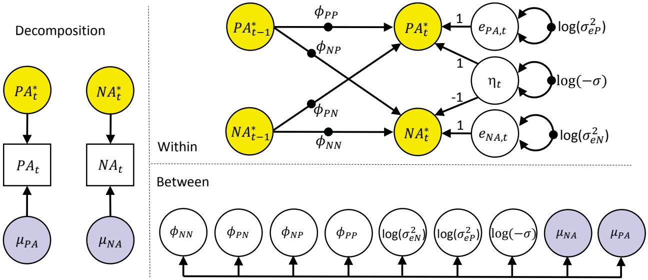
$$\begin{aligned} \mu_i &= \gamma_\mu + u_{0i} \\ \phi_i &= \gamma_\phi + u_{1i} \\ \log(\sigma_i^2) &= \gamma_{\log(\sigma^2)} + u_{2i} \end{aligned} \qquad \begin{bmatrix} u_{0i} \\ u_{1i} \\ u_{2i} \end{bmatrix} \sim MN \left[\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \psi_{11} & & \\ \psi_{21} & \psi_{22} & \\ \psi_{31} & \psi_{32} & \psi_{33} \end{bmatrix} \right]$$

Reasons to assume **individual differences** for σ^2:

- individuals may differ with respect to the **variability in exposure** to external factors
- individuals may differ with respect to their **reactivity** to external influences (see reward experience and stress sensitivity research)

In the bivariate case, we want **random innovation variances** AND **random innovation covariance**.

The latter is modeled with an additional factor η_t:



Where:

- -η_t is the shared part (we assume a negative covariance)
- $e_{PA,t}$ and $e_{NA,t}$ are the unique parts

Mplus code: Within model

```
MODEL:     %WITHIN%
           p_pp | dayPA ON dayPA&1;
           p_pn | dayPA ON dayNA&1;
           p_np | dayNA ON dayPA&1;
           p_nn | dayNA ON dayNA&1;

           ! Create latent variable that represents negative covariance
           Cov BY dayPA1 dayNA-1;

           ! Create random (log) variances
           logvarPA | dayPA;
           logvarNA | dayNA;
           logCov | Cov;

           %BETWEEN%
           p_pp-p_nn WITH p_pn-p_nn logvarPA logvarNA logCov dayPA dayNA;
           logvarPA WITH logvarNA logCov dayPA dayNA;
           logvarNA WITH logCov dayPA dayNA;
           logCov WITH dayPA dayNA;
           dayPA WITH dayNA;

OUTPUT:    TECH1 TECH8 STDYX FSCOMPARISON;
```

What about many variables?

Emilio Ferrer obtained data from **193 dyads** for **52-108 days** on **8 variables** (i.e., general and relationship specific PA and NA).

Within level: Vector autoregressive model

$$\begin{bmatrix} GPAM^*_{it} \\ GNAM^*_{it} \\ RSPAM^*_{it} \\ RSNAM^*_{it} \\ GPAF^*_{it} \\ GNAF^*_{it} \\ RSPAF^*_{it} \\ RSNAF^*_{it} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} & \phi_{14} & \phi_{15} & \phi_{16} & \phi_{17} & \phi_{18} \\ \phi_{21} & \phi_{22} & \phi_{23} & \phi_{24} & \phi_{25} & \phi_{26} & \phi_{27} & \phi_{28} \\ \phi_{31} & \phi_{32} & \phi_{33} & \phi_{34} & \phi_{35} & \phi_{36} & \phi_{37} & \phi_{38} \\ \phi_{41} & \phi_{42} & \phi_{43} & \phi_{44} & \phi_{45} & \phi_{46} & \phi_{47} & \phi_{48} \\ \phi_{51} & \phi_{52} & \phi_{53} & \phi_{54} & \phi_{55} & \phi_{56} & \phi_{57} & \phi_{58} \\ \phi_{61} & \phi_{62} & \phi_{63} & \phi_{64} & \phi_{65} & \phi_{66} & \phi_{67} & \phi_{68} \\ \phi_{71} & \phi_{72} & \phi_{73} & \phi_{74} & \phi_{75} & \phi_{76} & \phi_{77} & \phi_{78} \\ \phi_{81} & \phi_{82} & \phi_{73} & \phi_{84} & \phi_{85} & \phi_{86} & \phi_{87} & \phi_{88} \end{bmatrix} \begin{bmatrix} GPAM^*_{it-1} \\ GNAM^*_{it-1} \\ RSPAM^*_{it-1} \\ RSNAM^*_{it-1} \\ GPAF^*_{it-1} \\ GNAF^*_{it-1} \\ RSPAF^*_{it-1} \\ RSNAF^*_{it-1} \end{bmatrix} + \begin{bmatrix} \zeta_{1it} \\ \zeta_{2it} \\ \zeta_{3it} \\ \zeta_{4it} \\ \zeta_{5it} \\ \zeta_{6it} \\ \zeta_{7it} \\ \zeta_{8it} \end{bmatrix}$$

which gives:

$$GPAM^*_{it} = \phi_{11} GPAM^*_{it-1} + \phi_{12} GNAM^*_{it-1} + \cdots + \phi_{18} RSNAF^*_{it-1} + \zeta_{1it}$$

$\dots$

$$RSNAF^*_{it} = \phi_{81} GPAM^*_{it-1} + \phi_{82} GNAM^*_{it-1} + \cdots + \phi_{88} RSNAF^*_{it-1} + \zeta_{8it}$$

Multilevel VAR(1)

Within level: Residual covariance matrix

$$\begin{bmatrix} \zeta_{1it} \\ \zeta_{2it} \\ \dots \\ \zeta_{8it} \end{bmatrix} \sim MN(0, \mathbf{\Theta}^*)$$

Hence, we estimate $8 \times 8 = 64$ lagged parameters, and $8 \times 9/2 = 36$ variances and covariances at the within-person level.

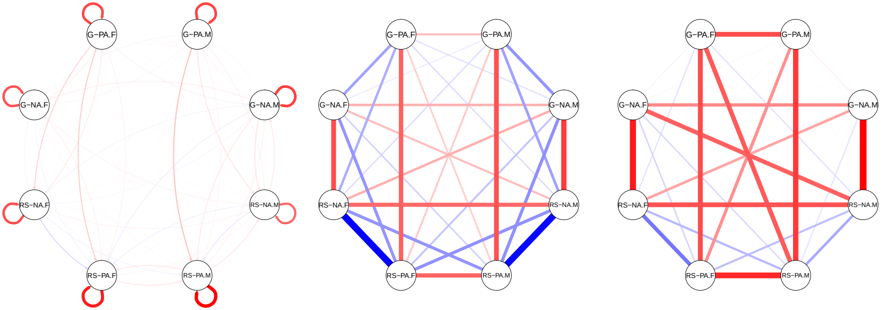
Between level: Fixed and random effects

$$\begin{bmatrix} \mu_{1i} \\ \mu_{2i} \\ \dots \\ \mu_{8i} \end{bmatrix} \sim MN(\boldsymbol{\gamma}, \mathbf{\Psi})$$

Hence, we estimate 8 grand means, and $8 \times 9/2 = 36$ variances and covariances at the between-person level. In total: 144 parameters.

Three networks

Lagged, within-person (residual), and between-person:



Note:

- lagged network = within-person standardized lagged relationships
- within-person residual network = correlations of within-person residuals
- between-person network = correlations of within-person means

References

- Bringmann, Vissers, Wichers, Geschwind, Kuppens, Peeters, Borsboom & Tuerlinckx (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS ONE, 8*, e60188, 1-13.
- Hamaker (2012). Why researchers should think “within-person”: A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.). *Handbook of Research Methods for Studying Daily Life*, 43-61. New York, NY: Guilford Publications.
- Hamaker, Asparouhov, Brose, Schmiedek & Muthén (submitted). At the frontiers of modeling intensive longitudinal data: Dynamic structural equation models for the affective measurements from the COGITO study. *Multivariate Behavioral Research*.
- Hamaker & Grasman (2015). To center or not to center? Investigating inertia with a multilevel autoregressive model. *Frontiers in Psychology, 5*, 1492.
- Hamaker & Wichers (2017). No time like the present: Discovering the hidden dynamics in intensive longitudinal data. *Current Directions in Psychological Science, 26*, 10-15.
- Jongerling, Laurenceau & Hamaker (2015). A multilevel AR(1) model: Allowing for inter-individual differences in trait-scores, inertia, and innovation variance. *Multivariate Behavioral Research, 50*, 334-349.

References

- Koval, Kuppens, Allen & Sheeber (2012). Getting stuck in depression: The roles of rumination and emotional inertia. *Cognition & Emotion, 26*, 1412-1427.
- Kuppens, Allen & Sheeber (2010). Emotional inertia and psychological maladjustment. *Psychological Science, 21*, 984-991.
- Kuppens, Sheeber, Yap, Whittle, Simmons & Allen (2012). Emotional inertia prospectively predicts the onset of depressive Multilevel AR(1) model 33 disorder in adolescence. *Emotion, 12*, 283-289.
- Schuurman, Ferrer, de Boer-Sonnenschein & Hamaker (2016). How to compare cross-lagged associations in a multilevel autoregressive model. *Psychological Methods, 21*, 206-221.
- Suls, Green & Hillis (1998). Emotional reactivity to everyday problems, affective inertia, and neuroticism. *Personality and Social Psychology Bulletin, 24*, 127-136.