Dynamic Structural Equation Modeling of Intensive Longitudinal Data

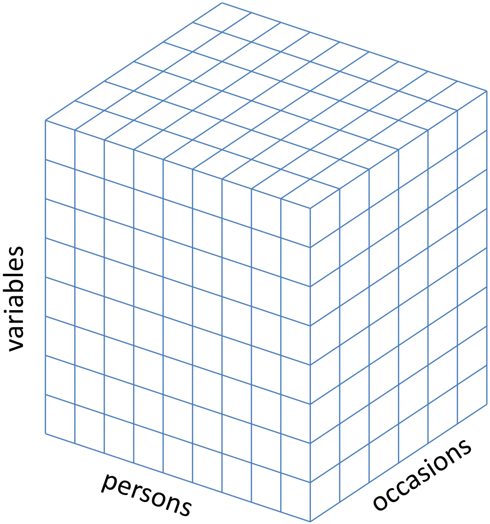
MuSt Zurich: Intro to Dynamic SEM
Day 2

Oisín Ryan

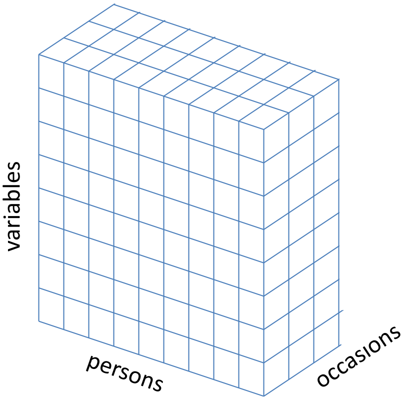
Utrecht University

March 2021

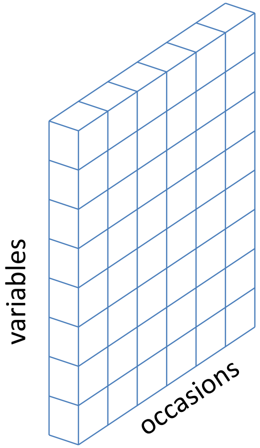
Cattell's data box



Panel Data: N = large and T is small (e.g. 2 -5 measurements)



Intensive Longitudinal Data: N=1 or many, T is large



Panel Data vs Intensive Longitudinal Data

Both consist of repeated measurements of the same set of variables over time

Panel Data:

- Many people, few measurements
- Time-points spaced far apart
- Longer Time-scale processes
- Suitable for describing global trajectories
- Captures between-person and average within-person dependencies
- “Wide” data format

Intensive Longitudinal Data:

- Single or multiple subjects. Many repeated measurements
- Time-points spaced closely together
- Attempts to directly capture short time-scale processes
- Focus primarily on individual within-person (idiographic) structure
- Sometimes called “time series” data
- “Long” format data

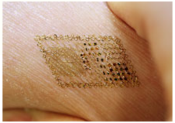
New technology

Smart phones

Smart glasses

Implants

Smart tattoo



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

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)

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 (in comparison to retrospective questionnaires)
- high ecological validity (in comparison to lab experiments)
- 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

Outline

- **Time series analysis**
  - Univariate Models
  - Multivariate Models
- Multilevel time series analysis
- DSEM application: Multilevel VAR(1) model
- Extensions and Issues
- 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.

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., **lagged** relationships)
- goal: forecasting

TSA in the social and medical sciences

In **sociology**:

- quarterly unemployment numbers
- effect of alcohol consumption per capita on criminal violence rates
- effect of suicide news on suicide rates

In **medical research**:

- effect of safety warnings on antidepressants use
- effects of pain control strategies
- effect of 9/11 attacks on weekly psychiatric patient admissions

In **psychology**:

- network of symptoms in depressive patient
- effect of feedback on academic performance
- effect of an intervention on the relationship between stress and affect

Similar to the models we considered yesterday, many time series models are based on *lagged regression*: Current values predicting future values.

Similar to the models we considered yesterday, many time series models are based on *lagged regression*: Current values predicting future values.

In TSA however, we typically pay more attention to the choice of time-series model

- Because we have more measurements over time, we have much more flexibility in the type of time-series model we consider.
- Should current Y be predicted by Y an hour ago, or Y two hours ago? How about Y 24 hours ago? Or a week ago?
- This is referred to the *order* or *lag* of a time-series model

Lags

Y
y_1
y_2
y_3
y_4
y_5
y_6
y_7
y_8
$\dots$
y_T

Lags

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

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

How to choose an appropriate Time-Series Model?

Box-Jenkins Method (1970). Aim: Model should explain all serial-dependencies present in the data – the **errors** of the regression model should contain no information about future observations (**white noise**)

Box-Jenkins Method (1970). Aim: Model should explain all serial-dependencies present in the data – the **errors** of the regression model should contain no information about future observations (**white noise**)

Model Components:

1. Remove **trends**
2. Remove **seasonal effects**
3. Choose order/lag of model

How to choose an appropriate Time-Series Model?

Box-Jenkins Method (1970). Aim: Model should explain all serial-dependencies present in the data – the **errors** of the regression model should contain no information about future observations (**white noise**)

Model Components:

1. Remove **trends**
2. Remove **seasonal effects**
3. Choose order/lag of model

Diagnostic Tools:

- **Autocorrelation function (ACF).** Raw/marginal correlation between lagged versions of our variable y_t and y_{t-k}

How to choose an appropriate Time-Series Model?

Box-Jenkins Method (1970). Aim: Model should explain all serial-dependencies present in the data – the **errors** of the regression model should contain no information about future observations (**white noise**)

Model Components:

1. Remove **trends**
2. Remove **seasonal effects**
3. Choose order/lag of model

Diagnostic Tools:

- **Autocorrelation function (ACF).** Raw/marginal correlation between lagged versions of our variable y_t and y_{t-k}
- **Partial autocorrelation functions (PACF).** Partial correlation between y_t and y_{t-k} when controlling for the effect of intermediate observations y_{t-1} to y_{t-k+1}.

ACF and PACF explained

ACF Lag-1

$cor(Y_t, Y_{t-1})$

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

ACF and PACF explained

ACF Lag-2

$cor(Y_t, Y_{t-2})$

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

ACF and PACF explained

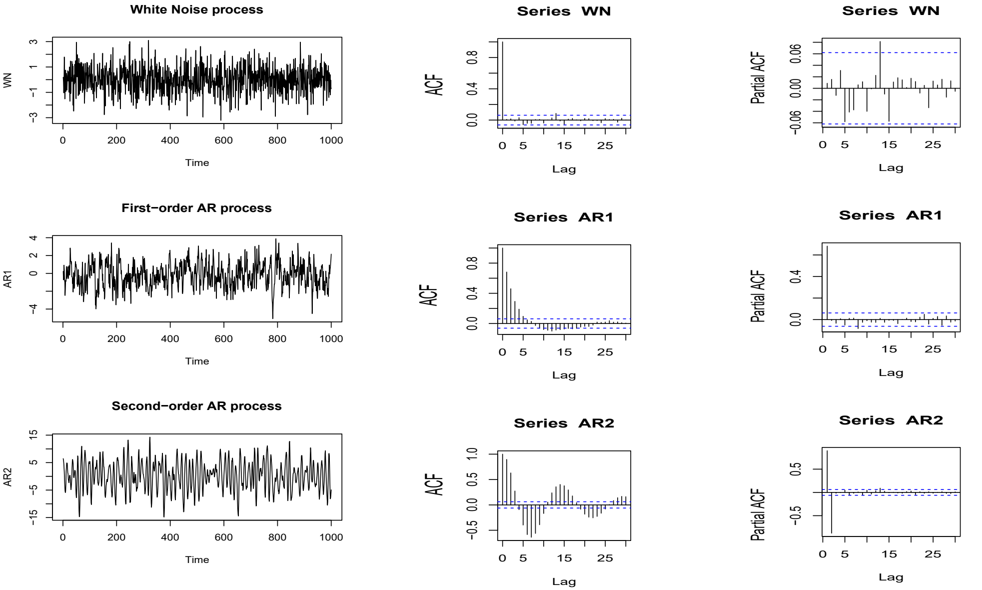
PACF Lag-2

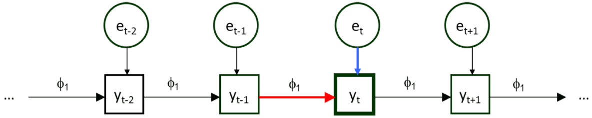
$cor(Y_t, Y_{t-2} \mid Y_{t-1})$

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
$\ldots$	$\ldots$	$\ldots$
y_T	y_{T-1}	y_{T-2}
	y_T	y_{T-1}
		y_T

Similar to a regression coefficient – controlling for Y_{t-1}

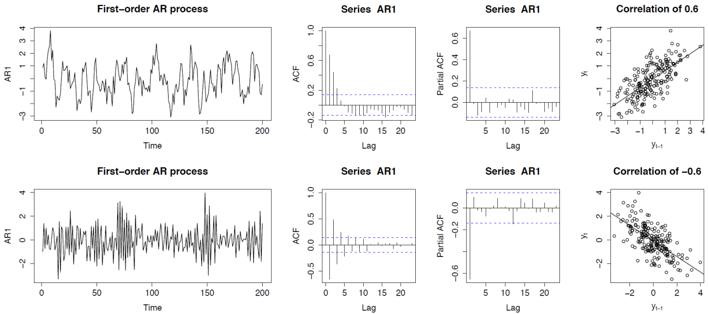
Sequence, ACF and PACF



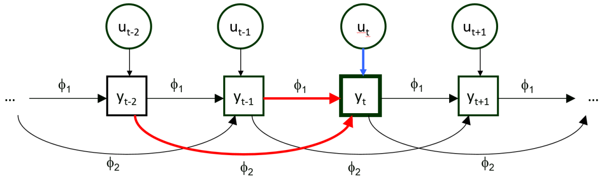


$$y_t = \phi_1 y_{t-1} + e_t$$

$$\phi_1 = \{0.6; -.6\}$$

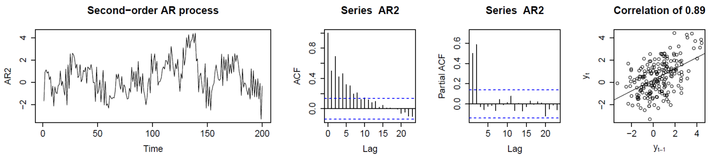


AR(2) Model



$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + e_t$$

$$\phi_1 = 0.2;\ \phi_2 = 0.6$$



Substantive Interpretation of AR processes

Granger and Morris (1976): An AR process is a **momentum effect in a random variable that varies smoothly over time**

Substantive Interpretation of AR processes

Granger and Morris (1976): An AR process is a **momentum effect in a random variable that varies smoothly over time**

We think about a variable with a stable **equilibrium position.** The errors represent "shocks" that push the system away from equilibrium. The AR parameter describes the "carry-over" or "regulation" of those shocks from one moment to the next.

Substantive Interpretation of AR processes

Granger and Morris (1976): An AR process is a **momentum effect in a random variable that varies smoothly over time**

We think about a variable with a stable **equilibrium position.** The errors represent "shocks" that push the system away from equilibrium. The AR parameter describes the "carry-over" or "regulation" of those shocks from one moment to the next.

AR(1) model: The **auto-regressive parameter** pulls the system **back** to equilibrium. The bigger the **absolute value**, the slower the system is to return to baseline

Substantive Interpretation of AR processes

Granger and Morris (1976): An AR process is a **momentum effect in a random variable that varies smoothly over time**

We think about a variable with a stable **equilibrium position.** The errors represent "shocks" that push the system away from equilibrium. The AR parameter describes the "carry-over" or "regulation" of those shocks from one moment to the next.

AR(1) model: The **auto-regressive parameter** pulls the system **back** to equilibrium. The bigger the **absolute value**, the slower the system is to return to baseline

Note: We always need to assume **stationarity**: the system returns to baseline after a shock. This only happens when the model parameters take on a certain value.

- For an AR(1), we need $-1 < \phi < 1$

Multivariate Time Series Models

Typically we have ILD that consist of **more than one** variable:

- mother's depression and child's disruptive behaviour
- stress and anxiety
- physical activity and happiness
- sleep quality and mood
- ...

In that situation, we might be interested in what time-series researchers call "**lead-lag relations**" between these variables

A simple multivariate time-series model

To analyze these lead-lag relations, we can make use of a **vector autoregressive (VAR) model**

A VAR(1) model is a multivariate version of the popular AR(1) model. We can write it as a set of equations:

$$\begin{aligned} A_t &= c_A + \phi_{11} A_{t-1} + \phi_{12} S_{t-1} + e_{A_t} \\ S_t &= c_S + \phi_{12} A_{t-1} + \phi_{22} S_{t-1} + e_{S_t} \end{aligned}$$

A simple multivariate time-series model

To analyze these lead-lag relations, we can make use of a **vector autoregressive (VAR) model**

A VAR(1) model is a multivariate version of the popular AR(1) model. We can write it as a set of equations:

$$A_t = c_A + \phi_{11} A_{t-1} + \phi_{12} S_{t-1} + e_{A_t}$$
$$S_t = c_S + \phi_{12} A_{t-1} + \phi_{22} S_{t-1} + e_{S_t}$$

where:
ϕ_{11} and ϕ_{22} are the **autoregressive coefficients**
ϕ_{12} and ϕ_{21} are the **cross-lagged coefficients**

A simple multivariate time-series model

To analyze these lead-lag relations, we can make use of a **vector autoregressive (VAR) model**

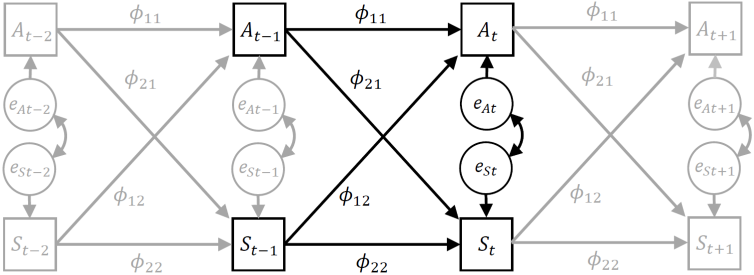
A VAR(1) model is a multivariate version of the popular AR(1) model. We can write it as a set of equations:

$$A_t = c_A + \phi_{11} A_{t-1} + \phi_{12} S_{t-1} + e_{A_t}$$
$$S_t = c_S + \phi_{12} A_{t-1} + \phi_{22} S_{t-1} + e_{S_t}$$

where:
ϕ_{11} and ϕ_{22} are the **autoregressive coefficients**
ϕ_{12} and ϕ_{21} are the **cross-lagged coefficients**

or in **matrix form**

$$\begin{bmatrix} A_t \\ S_t \end{bmatrix} = \begin{bmatrix} c_A \\ c_S \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{bmatrix} \begin{bmatrix} A_{t-1} \\ S_{t-1} \end{bmatrix} + \begin{bmatrix} e_{A_t} \\ e_{S_t} \end{bmatrix}$$



Researchers often interpret the cross-lagged parameters in terms of **causal dominance** or **Granger causal** relationships. Note: the **standardized cross-lagged regression coefficients** should be used!

VAR(1) vs CLPM

This should be starting to look pretty familiar to you - the path model looks essentially the same as the CLPM!

The VAR(1) is a time-series model for long single-subject data and the CLPM is a model for (short, multi-subject) panel data.

But there are some other **key differences** to keep in mind. Notably, time-series models like the VAR(1) models are only able to model *stationary processes*

Stationarity

Stationary processes are those for which the mean, variance, and auto-correlation structure **stay the same over our window of observation**

- Lagged regression parameters stay fixed across waves
  - Always necessary in some form - otherwise impossible to estimate!
- Only certain AR(1) and VAR(1) parameters are allowed
  - For AR(1) $-1 < \phi < 1$
  - For VAR(1) the $-1 <$ eigenvalues of $\mathbf{\Phi} < 1$
- Mean is fixed over time (no trends or trends removed)
  - Not so different to CLPM approaches - there either we are modeling the trend only, or removing the trend first!

This concludes our brief tour of time-series analysis

We only focused on a limited range of time-series models based on auto-regressive (AR) models. Actually, there is a much broader suite of time-series models, sometimes known as ARIMA models. See **Hamaker & Dolan (2009)**

AR(1) and VAR(1) models are some of the most popular time-series models. Largely because they are the easiest to interpret.

For that reason, they have been the most popular choices for *multilevel* time series analysis

Outline

- Time series analysis
- **Multilevel time series analysis**
- DSEM application: Multilevel VAR(1) model
- Extensions and Issues
- Discussion

Lagged relationships in multilevel data

Single-subject time series data allows us to study (for a *stationary process*):

- Lagged relationships between a variable and itself: **autoregression**
- Lagged relationships between different variables: **cross-lagged relationships**

Lagged relationships in multilevel data

Single-subject time series data allows us to study (for a *stationary process*):

- Lagged relationships between a variable and itself: **autoregression**
- Lagged relationships between different variables: **cross-lagged relationships**

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

- individual differences in **autoregression**
- individual differences in **cross-lagged relationships**

Lagged relationships in multilevel data

Single-subject time series data allows us to study (for a *stationary process*):

- Lagged relationships between a variable and itself: **autoregression**
- Lagged relationships between different variables: **cross-lagged relationships**

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

- individual differences in **autoregression**
- individual differences in **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
1
1
1
1
1
2
2
2
2
2
$\dots$
N
N
N

Creating lagged predictors

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

Creating lagged predictors

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

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}

Inertia research based on multilevel AR(1) models

Level 1 model:

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

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:

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

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:

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

Dynamic networks based on multilevel VAR(1) models

Level 1 model:

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

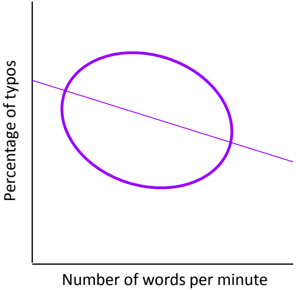
Initiated by **Bringmann et al. (2013)**, and further popularized by the software from **Sacha Epskamp**.

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

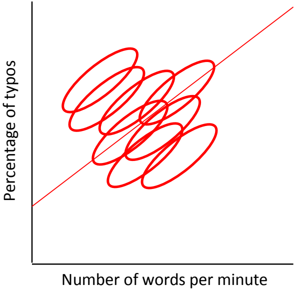
Main idea is that **stronger connections** lead to an **increased risk** of developing and maintaining psychopathology.

Within- and Between- Variance

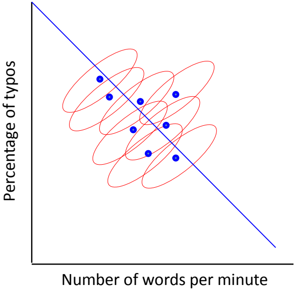
Cross-sectional relationship



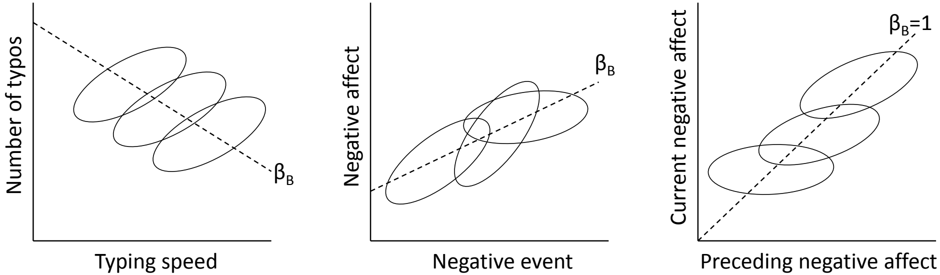
Within-person relationship



Between-person relationship

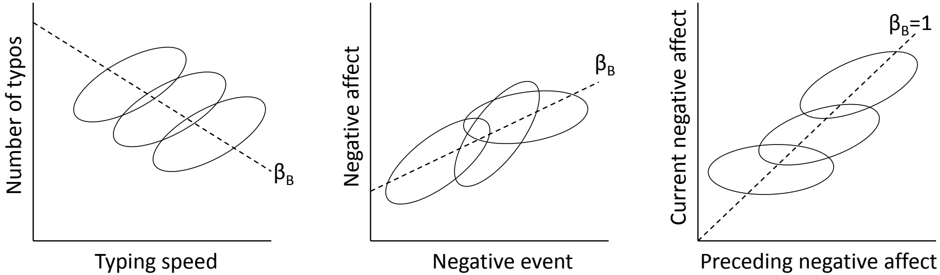


Between-person differences in within-person slopes



Taken from Hamaker and Grasman (2014).

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

Multilevel vs Single-Subject Time Series

Advantages of Multi-Level Time Series Models:

- Allow to model within-person and between-person structure simultaneously
- Regularizes estimates of within-person parameters
    - Person-specific parameters get “pulled” towards the mean. Estimates have lower variance across samples
- In general, higher power than single-subject models - the more similar people are, and the more people we have, the less observations we need per person

Multilevel vs Single-Subject Time Series

Advantages of Multi-Level Time Series Models:

- Allow to model within-person and between-person structure simultaneously
- Regularizes estimates of within-person parameters
    - Person-specific parameters get “pulled” towards the mean. Estimates have lower variance across samples
- In general, higher power than single-subject models - the more similar people are, and the more people we have, the less observations we need per person

Disadvantages of Multi-Level Time Series Models:

- Need to assume the same type of time series model (or a very general model) for everyone
- Regularizes estimates of within-person parameters
    - If we have enough time points, and we really care about the individual-specific parameters, we don’t want this to happen
    - If people are not that similar, we might get a very strange/uninformative “average” model

Why use DSEM in Mplus?

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

Why use DSEM in Mplus?

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

Why use DSEM in Mplus?

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:

- A specific implementation of multilevel time-series models within an SEM framework
- Uses Bayesian estimation techniques
- Tackles the above problems

Outline

- Time series analysis
- Multilevel time series analysis
- **DSEM application: Multilevel VAR(1) model**
- Extensions and Issues
- 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. (2018).

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. (2018).

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

Decomposition into a between part and a within part

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

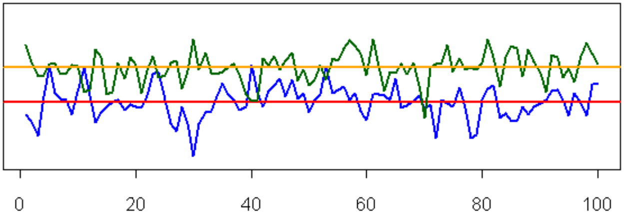
Decomposition

Decomposition into a between part and a within part

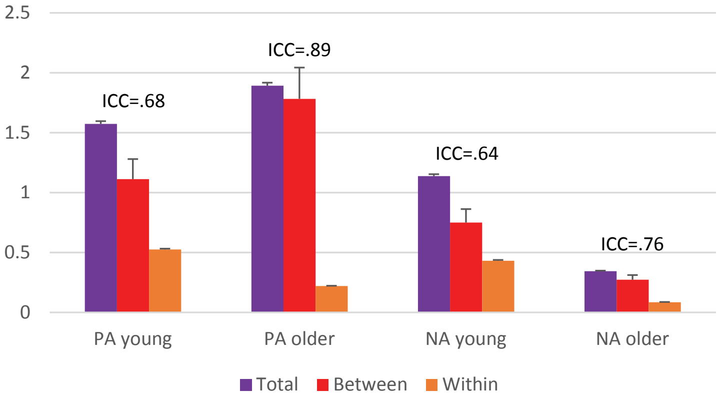
$$PA_{it} = \mu_{PA,i} + PA^*_{it}$$
$$NA_{it} = \mu_{NA,i} + NA^*_{it}$$

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



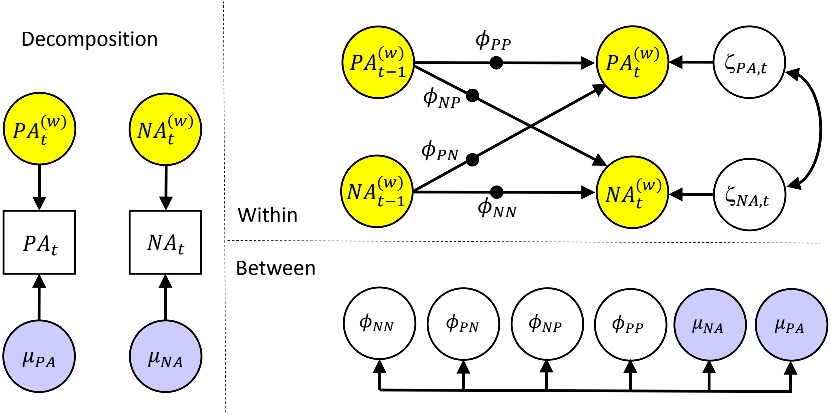
Total, between-, and within-person variance



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)

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)

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)

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

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

Where:

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

Between-person level model

Between level: fixed and random effects

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

Where:

- γ_P to $\gamma_{NN} \Rightarrow$ 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)

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

```
                                Posterior  One-Tailed         95% C.I.
                    Estimate       S.D.      P-Value   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.**

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)

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**.

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

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

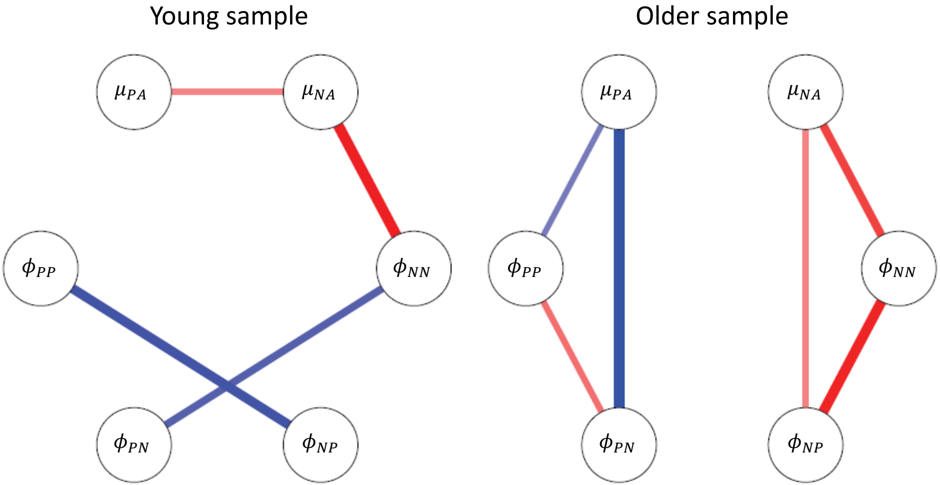
R-SQUARE

Within-Level R-Square Averaged Across Clusters

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

Between-person level: Correlated random effects

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: Multilevel VAR(1) model
- **Extensions and Issues**
    - More SEM
    - Time-Intervals
    - Stationarity
- Discussion

One of the benefits of fitting Multilevel time series model in a SEM framework is that we can build up models in much the same way as we build up SEM models

For instance, we can include a measurement model and fit a multilevel-VAR(1) on latent variables

View of random effects as *latent variables* allows us to do other creative things

Disclaimer: In DSEM, testing general fit of models is not so straightforward, due to the Bayesian estimation techniques used. **DIC** is the Bayesian information criteria, but model comaprison can be tricky due to irregularities in how model complexity is counted

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

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

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.**

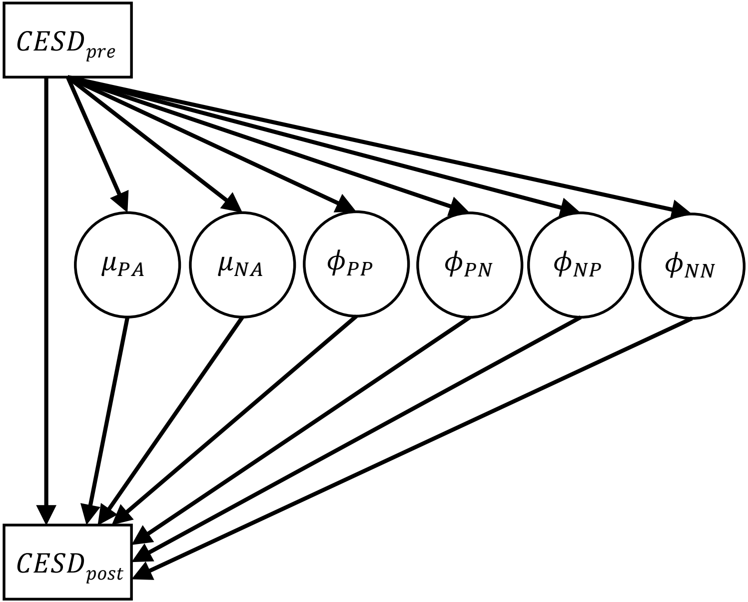
Between level: Including a level 2 predictor

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

Between level: Including a level 2 outcome

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

Dynamic mediation model



Mplus input level-2 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;
```

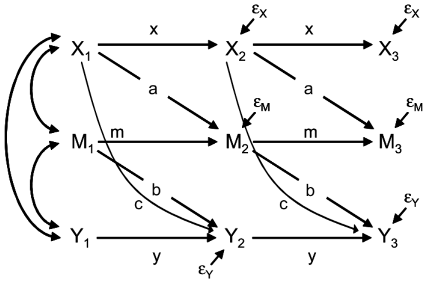
Mplus output level-2 mediation model (younger sample)

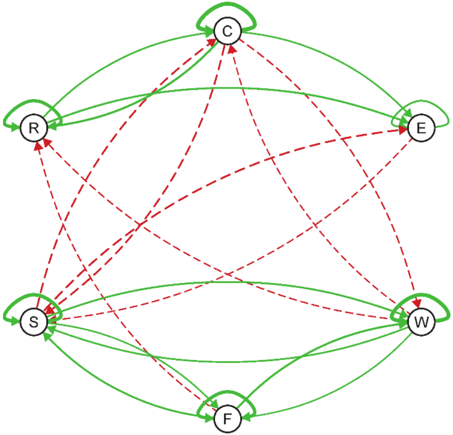
```
                                Posterior  One-Tailed         95% C.I.
                    Estimate       S.D.      P-Value   Lower 2.5%  Upper 2.5%  Significance
[...]
Between Level
[...]

 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 level-2 mediation model (older sample)

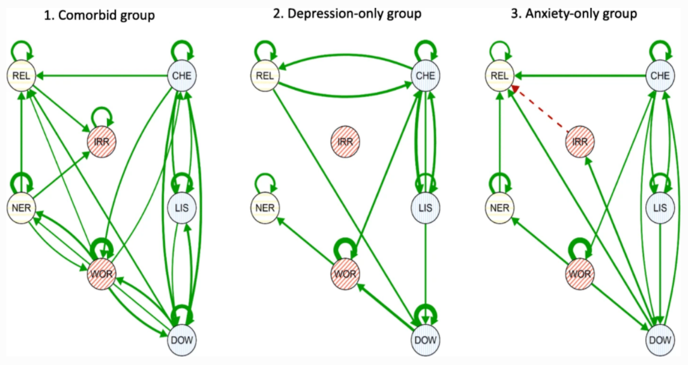
	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%	Significance
[...]						
Between Level						
[...]						
Intercepts						
CESDPOST	0.015	0.113	0.448	-0.210	0.236	
DAYPA	4.566	0.120	0.000	4.336	4.796	*
DAYNA	0.313	0.052	0.000	0.210	0.417	*
P_PP	0.421	0.026	0.000	0.370	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	

We may also be interested in *direct*, *indirect* and *total effects* **across lags.** Often the interest in *psychological network analysis*





Groen, Ryan, Wigman et al (2020): Average within-person VAR(1) network structure of different “symptom-state” items in ESM data, across three groups: Comorbid, Anxiety-only and Depression-only.



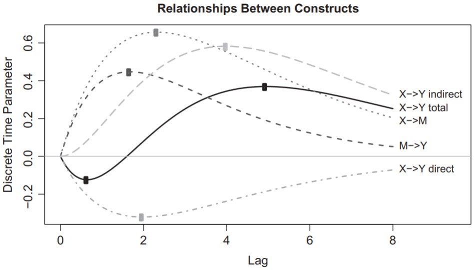
No evidence that the comorbid group displayed higher indirect effects through “bridge symptoms” than the other groups. All analyses in DSEM

The Time-Interval Problem

VAR(1) models are **discrete time**: Do not account in any way for time-interval information

As we discussed yesterday, all lagged regression models (potentially) suffer from the time-interval problem.

Effects can change sign, size and relative ordering depending on how measurements are spaced in time (Kuiper & Ryan, 2018)



The Time-Interval Problem

Problem 1: If you sample at a fixed interval, Δt, relationships may not generalize to **other** intervals

Problem 2: If you sample at a mix of intervals $\Delta t_1, \Delta t_2 \ldots$ parameters may be a *mix* of effects at all different intervals, not correctly characterizing relationships at **any** interval

Problem 1: If you sample at a fixed interval, Δt, relationships may not generalize to **other** intervals

Problem 2: If you sample at a mix of intervals $\Delta t_1, \Delta t_2 \ldots$ parameters may be a *mix* of effects at all different intervals, not correctly characterizing relationships at **any** interval

DSEM cannot *solve* these problems, but does try to *mitigate* them

- Choose a *target interval* and try to estimate effects at that interval
- Do this by inserting missing values in a grid such that rows are approximately "equally spaced" with Δt
- 'TINTERVAL' command

Alternative Approach: Continuous-Time Modeling

Continuous-Time models **explicitly** model time-interval dependencies

Based on **differential equation** models (tradition of physics, ecology, systems biology):

- DE models describe the **truly** moment-to-moment dependencies
- Relationships at longer lags a result of “adding up” or “path-tracing” moment-to-moment effects
- Can still be estimated from ESM type data

Relating continuous and discrete time solutions

$$\boldsymbol{\Phi}(\Delta t_i) = e^{\boldsymbol{A} \times \Delta t_i}$$

Alternative Approach: Continuous-Time Modeling

Continuous-Time models for panel data:

- Voelkle, Oud, Davidov & Schmidt (2012)
- Deboeck & Preacher (2018) Mediation

Continuous-Time models for ILD/ESM data in psychology:

- Ryan, Kuiper, Hamaker (2018)
- Ryan & Hamaker (under review) - Continuous Time Network Analysis

Estimation:

- *ctsem* (Driver, Oud, Voelkle, 2017). Bayesian, multilevel, CT analysis.
- *dynr* (Ou, Hunter, Chow, 2017).
- **Disclaimer:** Less user-friendly than Mplus!

Beyond Stationarity

Time-Varying VAR(1) models:

- Haslbeck, Bringmann & Waldorp (2020)
- Assumes "local stationarity" - all parameters can be smooth functions of time
- Mostly single-subject, needs even more time-points

Regime-Switching Models:

- Time-series models where the system "switches" between different states with different parameters
- See Hamaker, Grasman, Kamphuis (2010); Haslbeck* & Ryan* (2021); dynr (Ou, Hunter, Chow, 2017)

In Mplus:

- **Residual dynamic modeling**: Allows for de-trending in the model
- **Cross-classified models**: Allows for random effects of time

Outline

- Time series analysis
- Multilevel time series analysis
- DSEM application: Multilevel VAR(1) model
- Extensions and Issues
- **Discussion**

Summary

Time-series analysis aims to capture the structure of within-person variation over time when we have many repeated measures

When we have ILD from multiple individuals, **multi-level time series models** can be used to try and explore how within-person structure varies across people

The VAR(1) is just one type of time-series model which happens to be very popular - simple, lag-1 linear relationships.

The VAR(1) is related to the cross-lagged panel model - both use auto-regressive and cross-lagged relationships to model change over time

RI-CLPM for Panel Data:

- “Wide” data format: Many subjects, few measurements per subject
- Random-Intercepts capture between-person variation in **means**
- No random slopes. Logic: Not enough measurements within-person
- Relationships can be different at every wave
- Means can change at each point in time

Multilevel VAR(1) for ILD

- “Long” data format: Many repeated measurements, 1 or more subjects
- Random-Intercepts capture between-person variation in **means**
- Random slopes capture between-person variation in within-person lagged parameters
- Relationships must be the same at every wave (*stationarity*)
- Trends/changes in mean must be removed, but can be done in the model (residual DSEM)

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)

DSEM in Mplus

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 possibilities: mlVAR, ctsem, dynr, GIMME

Discussion

Of course, big issue is: which time-series models are substantively interesting?

- Which captures/reproduces patterns that are interesting?
- How can this be driven by theory?

Approach direct (and causal) interpretation with care

Tip: Pick a model and simulate time-series data from it (e.g. in R) - get a feel for what types of patterns it produces

These models are just tools. Proper use must be informed by theory (Haslbeck*, Ryan*, Robinaugh* et al, in press)

Thank you!

o.ryan@uu.nl

Thank you!

o.ryan@uu.nl

For more info, some relevant summer schools at UU:

- Introduction to SEM using Mplus
- Advanced course on Mplus
- Modeling the Dynamics of Intensive Longitudinal Data

References

- Box, George; Jenkins, Gwilym (1970). Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day.
- 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.
- Haslbeck, J. M., Bringmann, L. F., & Waldorp, L. J. (2020). A tutorial on estimating time-varying vector autoregressive models. Multivariate behavioral research, 1-30.
- Haslbeck*, J., Ryan*, O., Robinaugh*, D., Waldorp, L., Borsboom, D. (in press). Modeling psychopathology: From data models to formal theories. *Psychological Methods*. Pre-print: psyarxiv.com/jgm7f/
- 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.
- Ryan, O., Kuiper, R. M., & Hamaker, E. L. (2018). A continuous time approach to intensive longitudinal data: What, Why and How? In K. v. Montfort, J. H. L. Oud, & M. C. Voelkle (Eds.), *Continuous time modeling in the behavioral and related sciences* (pp. 27-54). New York, NY: Springer. Pre-print: ryanoisin.github.io
- Ryan, O. & Hamaker, E.L. (under review) Time to Intervene: A Continuous-Time Approach to Network Analysis and Centrality. Pre-print: /psyarxiv.com/2ambn
- 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.
- Voelkle, M. C., Oud, J. H., Davidov, E., & Schmidt, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. Psychological methods, 17(2), 176.