EBERHARD KARLS UNIVERSITÄT TÜBINGEN

METHODS CENTER

Separation of inter-individual differences, intra-individual changes, and time-specific effects in intensive longitudinal data using the NDLC-SEM framework

Augustin Kelava

International Statistical Conference in Memory of Professor Sik-Yum Lee

December 17 - 18, 2019

Outline

Some general thoughts

NDLC-SEM

Examples

Discussion

Some general thoughts

NDLC-SEM

Examples

Discussion

Some general thoughts

Large studies with longitudinal data

- inter-individual differences (e.g., some initial assessment, traits)
- intra-individual changes (e.g., development of competencies, states)
- unobserverd heterogeneity in trajectories (e.g., some subgroups)

Some general thoughts

Large studies with longitudinal data

- inter-individual differences (e.g., some initial assessment, traits)
- intra-individual changes (e.g., development of competencies, states)
- unobserverd heterogeneity in trajectories (e.g., some subgroups)

Some typical challenges

- poor data quality (e.g., initial assessement takes too much time)
- separation of different functional relationships between variables, data levels, heterogeneity etc.
- time-dependent variables/effects (interventions, specific events)
- lack of statistical procedures (and implementations)
- sparsity of models (many parameters, covariates etc.); regularization procedures

State of the art

Dynamic Latent Class Analysis (DLCA) framework

- Each individual is a member of a **latent class** at each time point with a specific probability (The latent class membership follows a *Hidden Markov Model process*.).
- The **individual-specific transition probabilities** are estimated as *(between-level) random effects* which are parameterized by a structural equation model or factor model. **The transition probabilities are not time-dependent!!**
- The **within-level** model is a *dynamic (time-series) model with autoregressive effects* of the latent variables.

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2017). Dynamic latent class analysis. *Structural Equation Modeling: A Multidisciplinary Journal, 24(2)*, 257-269. doi:10.1080/10705511.2016.1253479

State of the art

Dynamic Structural Equation Model (DSEM) framework

- The DSEM framework separates (a) **subject-specific** and (b) **time-specific random effects** (on the between-level).
- There is a **dynamic latent variable model**, which describes, for example, the *intra-individual (within-level) changes* using an autoregressive process of latent variables.
- Each **random within-level parameter** is explained by the subject-specific and time-specific random effects.

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 25(3)*, 359-388. doi:10.1080/10705511.2017.1406803

Some general thoughts

NDLC-SEM

Examples

Discussion

NDLC-SEM

based on ...

Kelava, A. & Brandt, H. (2019). Nonlinear Dynamic Latent Class Structural Equation Model. *Structural Equation Modeling, 26*, 509-528.
DOI: 10.1080/10705511.2018.1555692

NDLC-SEM – Components

Decomposition of observed variable Y_{it}

$$Y_{it} = Y_{1it} + Y_{2i} + Y_{3t} \tag{1}$$

Individual-specific component Y_{2i}

$$Y_{2i} = \nu_2 + \Lambda_2 \eta_{2i} + K_2 X_{2i} + \epsilon_{2i} \tag{2}$$

$$\eta_{2i} = \alpha_2 + B_2 \eta_{2i} + \Omega_2 h_2(\eta_{2i}) + \Gamma_2 X_{2i} + \zeta_{2i}. \tag{3}$$

Time-specific component Y_{3t}

$$Y_{3t} = \nu_3 + \Lambda_3 \eta_{3t} + K_3 X_{3t} + \epsilon_{3t} \tag{4}$$

$$\eta_{3t} = \alpha_3 + B_3 \eta_{3t} + \Omega_3 h_3(\eta_{3t}) + \Gamma_3 X_{3t} + \zeta_{3t}. \tag{5}$$

NDLC-SEM

Within model

$$[Y_{1it}|S_{it}=s] = \nu_{1s} + \Sigma_{l=0}^{L}\Lambda_{1ls}\eta_{1i(t-l)} + \Sigma_{l=0}^{L}R_{ls}Y_{1i(t-l)} + \Sigma_{l=0}^{L}K_{1ls}X_{1i(t-l)} + \epsilon_{1it} \quad (6)$$

$$[\eta_{1it}|S_{it}=s] = \alpha_{1s} + \Sigma_{l=0}^{L}B_{1ls}\eta_{1i(t-l)} + \Sigma_{l=0}^{L}\Sigma_{l'=0}^{L'}\Omega_{1ll's}h_{1ll'}(\eta_{1i(t-l)},\eta_{1i(t-l')}) + \Sigma_{l=0}^{L}Q_{ls}Y_{1i(t-l)} + \Sigma_{l=0}^{L}\Gamma_{1ls}X_{1i(t-l)} + \zeta_{1it} \quad (7)$$

Categorical variables

$$[Y_{1jit}=k|S_{it}=s] \Leftrightarrow \tau_{j(k-1)s} \leq [Y^{*}_{1jit}|S_{it}=s] < \tau_{jks} \quad (8)$$

with $\tau_{j0s} = -\infty$ and $\tau_{j(m_j)s} = \infty$ for all latent states $s = 1, ..., K$.

NDLC-SEM

Within model

$$[Y_{1it}|S_{it}=s] = \nu_{1s} + \Sigma_{l=0}^{L}\Lambda_{1ls}\eta_{1i(t-l)} + \Sigma_{l=0}^{L}R_{ls}Y_{1i(t-l)} + \Sigma_{l=0}^{L}K_{1ls}X_{1i(t-l)} + \epsilon_{1it} \quad (6)$$

$$[\eta_{1it}|S_{it}=s] = \alpha_{1s} + \Sigma_{l=0}^{L}B_{1ls}\eta_{1i(t-l)} + \Sigma_{l=0}^{L}\Sigma_{l'=0}^{L'}\Omega_{1ll's}h_{1ll'}(\eta_{1i(t-l)},\eta_{1i(t-l')}) + \Sigma_{l=0}^{L}Q_{ls}Y_{1i(t-l)} + \Sigma_{l=0}^{L}\Gamma_{1ls}X_{1i(t-l)} + \zeta_{1it} \quad (7)$$

Categorical variables

$$[Y_{1jit}=k|S_{it}=s] \Leftrightarrow \tau_{j(k-1)s} \leq [Y^{*}_{1jit}|S_{it}=s] < \tau_{jks} \quad (8)$$

with $\tau_{j0s}=-\infty$ and $\tau_{j(m_j)s}=\infty$ for all latent states $s=1,...,K$.

The Markov switching model

The latent state variable S_{it} follows a Markov switching model with **person- and time-specific transition probability**:

$$P(S_{it}=d|S_{i(t-1)}=c) = \frac{\exp(\alpha_{itdc})}{\Sigma_{k=1}^{K}\exp(\alpha_{itkc})} \quad (9)$$

α_{itdc} are person- and time-specific **random effects** with $\alpha_{itKc}=0$.

NDLC-SEM

Random effects

Any random within-level parameter p_{it} (e.g., elements from ν_{1s}, Λ_{1ls} etc.) can be decomposed as

$$p_{it} = p_{2i} + p_{3t} \tag{10}$$

- p_{2i} **is a subject-specific random effect which is an element of vector** η_{2i} in the between-level model.
- p_{3t} **is a time-specific random effect which is an element of vector** η_{3t}.

Generalized measurement and structural models

$$[Y_{1it}|S_{it} = s] = \nu_{1s} + \Sigma_{l=0}^{L}\Lambda_{1ilts}\eta_{1i(t-l)} + \Sigma_{l=0}^{L}R_{ilts}Y_{1i(t-l)} + \Sigma_{l=0}^{L}K_{1ilts}X_{1i(t-l)} + \epsilon_{1it} \tag{11}$$

$$\begin{aligned}[\eta_{it}|S_{it} = s] = \alpha_{1s} &+ \Sigma_{l=0}^{L}B_{1ilts}\eta_{1i(t-l)} + \Sigma_{l=0}^{L}\Sigma_{l'=0}^{L'}\Omega_{1ill'ts}h_{1ll'}(\eta_{1i(t-l)}, \eta_{1i(t-l')}) \\ &+ \Sigma_{l=0}^{L}Q_{ilts}Y_{1i(t-l)} + \Sigma_{l=0}^{L}\Gamma_{1ilts}X_{1i(t-l)} + \zeta_{1it}\end{aligned} \tag{12}$$

NDLC-SEM – properties

The NDLC-SEM framework is a **comprehensive approach** which is capable of

a) **intra-individual changes** (as a dynamic structural equation model),
b) **inter-individual differences**, which have an effect on the individual trajectories (e.g., within-level random parameters)
c) **time-specific effects** (as random effects),
d) **dynamic latent class memberships**, which capture *heterogeneity of the trajectories* or which can reflect *nominal latent variables* (such as knowledge mastery), and
e) **flexible nonlinear effects** (e.g., splines or interactions) in models, in order to account for (multiple) complex relationships between the latent variables.

The Δ[1]

- The DLCA framework is capable of a). b) are used to predict d). Again, there is no inclusion of time-dependent information in d).
- The DSEM framework is capable of a). b) and c) are used for random effects on the within-level.

[1] Delta is by the way a great album of the British band Mumford & Sons.

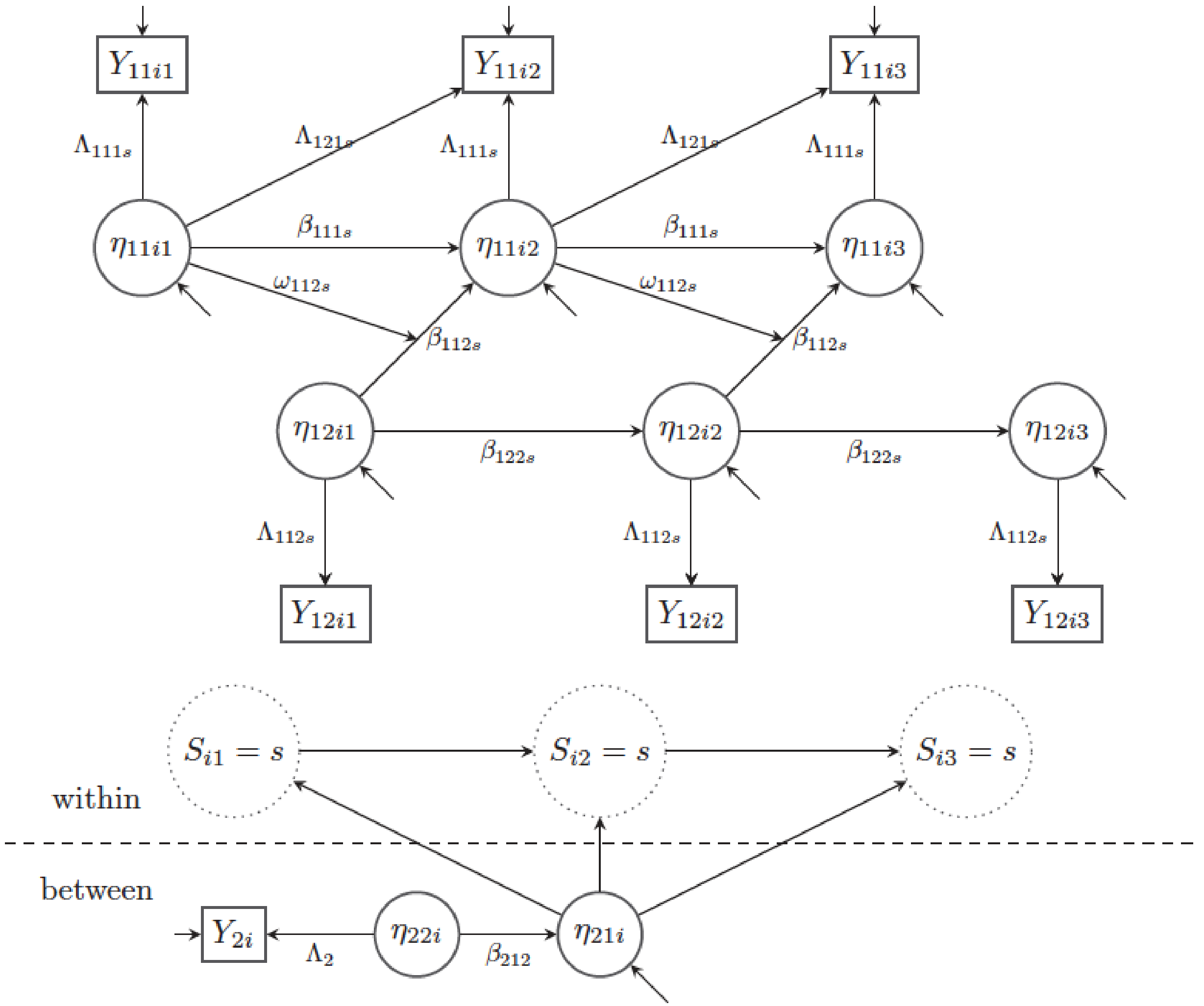
Some general thoughts

NDLC-SEM

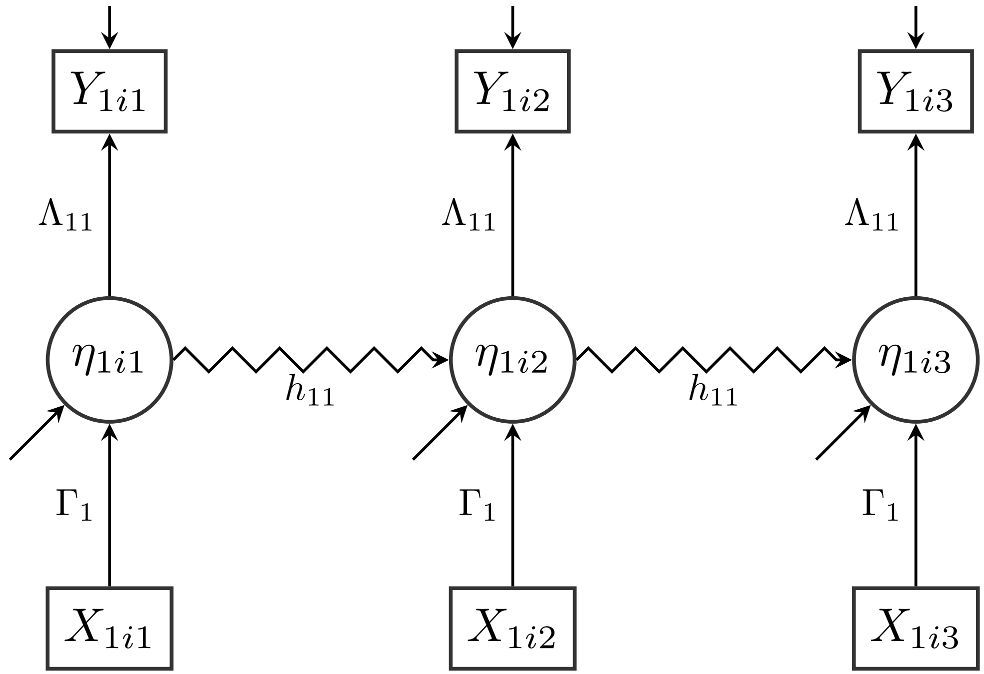
Examples

Discussion

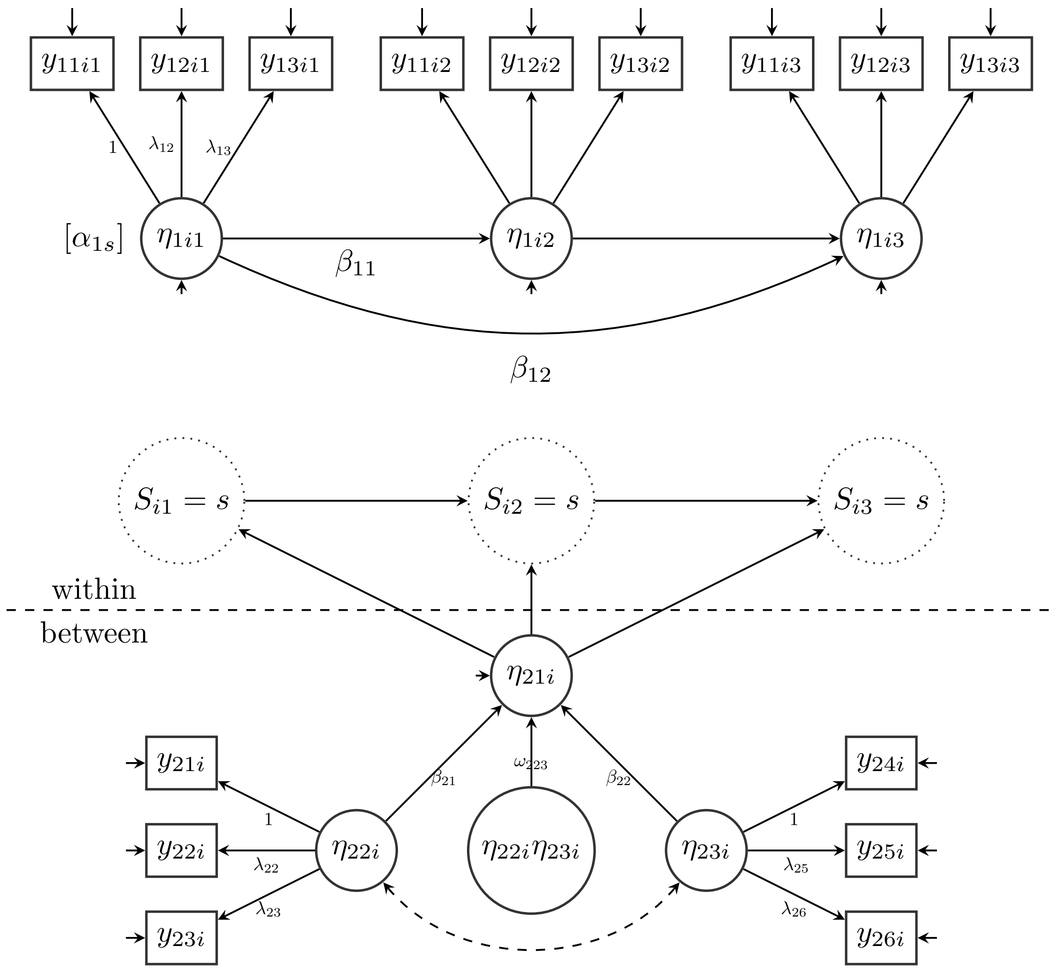
Example – Achievement and effort



Example – Single-level semiparametric dynamic structural equation model



Example – College drop-out study



Example – Development of Math Skills

Sample

- Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999 (ECLS-K; Tourangeau, Nord, Lê, Pollack, & Atkins-Burnett, 2009)
- 7 measurement occasions from kindergarten to grade 8
- random sample of 500 students

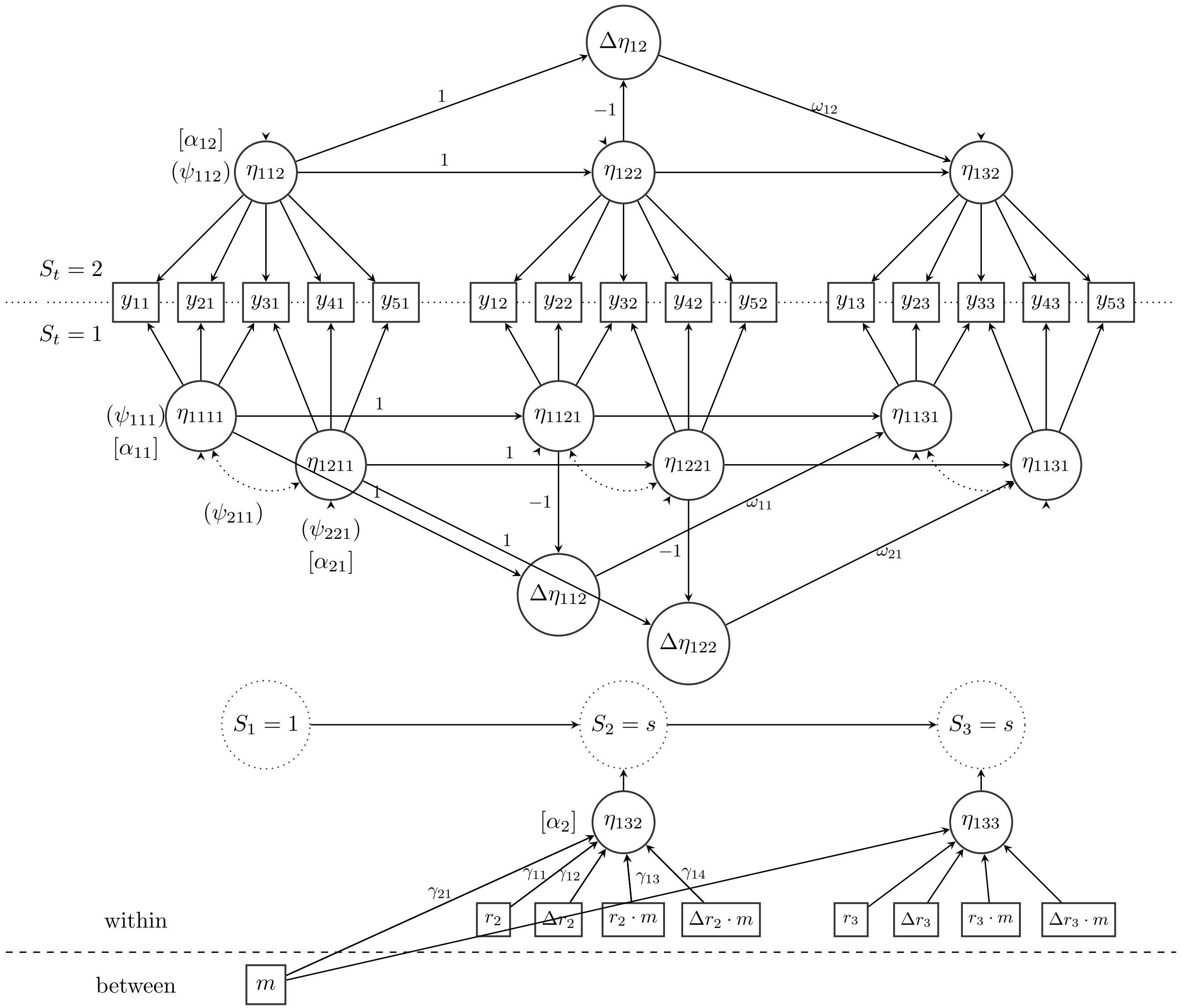
Variables

- math skills: 5 out of 9 scales (ordinality/sequence, add/subtract, multiply/divide, place value, rate & measurement)
- time-specific reading skills were used as observed covariates
- fine motor skills were used as covariates (initial measurement occasion)

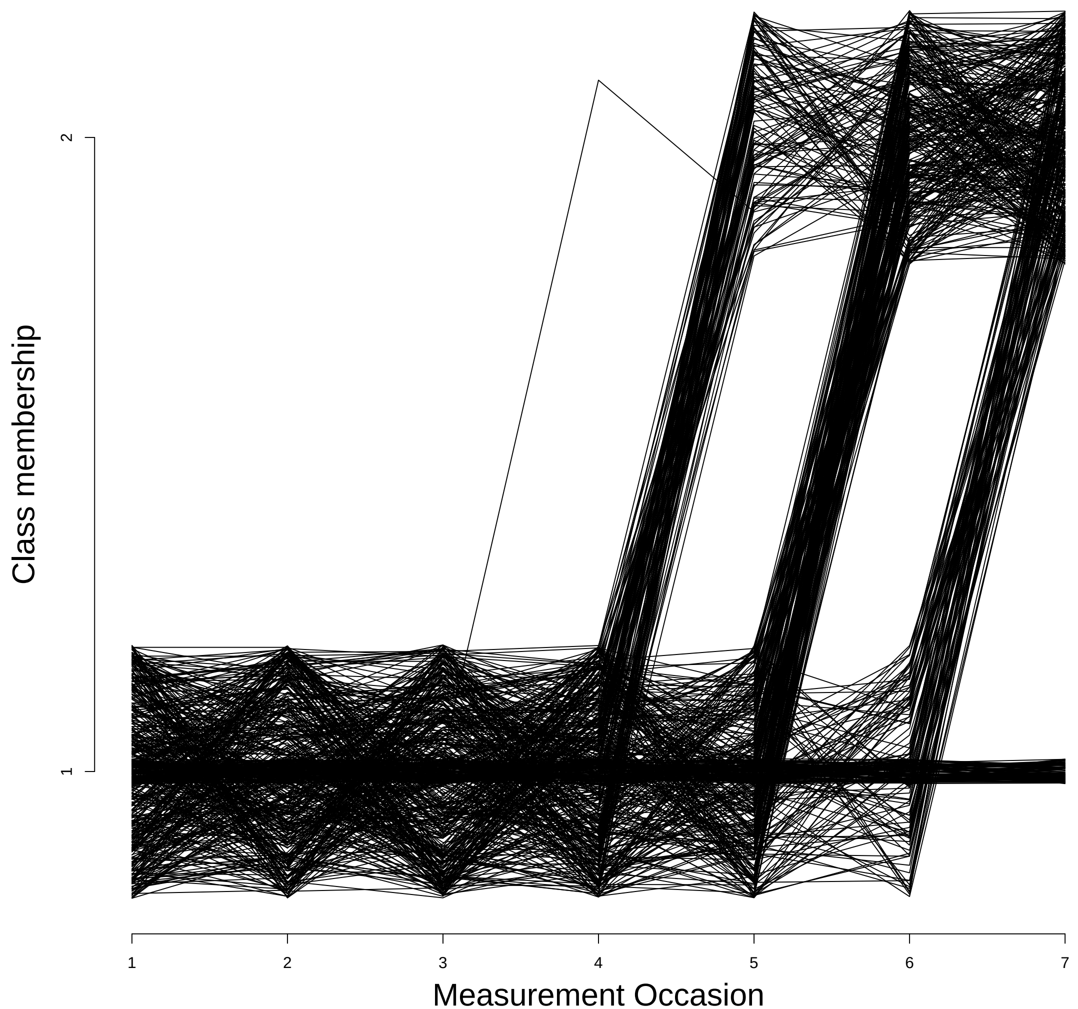
Hypotheses

- There is **nominal change in the math skill constructs** over time (state S_{it}). Whereas at the beginning two constructs (concrete and abstract math) are necessary, only one construct for math skills will be sufficient at a later time point.
- Students who have a strong increase from one time point to the next in **reading skills**, will also be more likely to master math (switch from state 1 to state 2).
- We assume that increasing **reading skills can compensate low fine motor skills**, which implies a nonlinear interaction effect between these variables.

Example – Development of Math Skills



Example – Development of Math Skills



Some general thoughts

NDLC-SEM

Examples

Discussion

Discussion

Longitudinal studies

- Integration and separation of
  a) intra-individual changes,
  b) inter-individual differences,
  c) time-specific effects,
  d) unobserved heterogeneity, and
  e) flexible relationships

  is **important** to both modern psychometric modeling of longitudinal data AND **to comprehensive substantial theories** (e.g., college student drop out theories).
- Large (longitudinal) studies include **many scales and covariates/variables**.
- However, the **complexity of typical models** is increasing dramatically.

So what?

Usefulness

- **Separation of different types of concepts/measures** (e.g., traits, states)
- **Simultaneous inclusion of both time-specific information AND inter-individual differences** when explaining (unobserved or observed) hetereogenity in trajectories (e.g., unobserved (!) decision to quit college might depend on both vulnerabilty factors and specific life events)
- **Structural changes** of concepts or behavior (e.g., changes of dimensionality of measures)
- **Interactions and flexible semiparametric effects** are available on all data levels.

Some general problems

Bayesian estimation

- **Regularization techniques** are important fields of future methods development, especially for complex longitudinal models.
- **Suitable variable and parameter selection techniques** are strongly required (not just shrinkage; but selection!).
- **Need for simulation studies** (e.g., sample size requirements, prior distributions)

Some general problems

Bayesian estimation

- **Regularization techniques** are important fields of future methods development, especially for complex longitudinal models.
- **Suitable variable and parameter selection techniques** are strongly required (not just shrinkage; but selection!).
- **Need for simulation studies** (e.g., sample size requirements, prior distributions)

Frequentist estimation

- Features like multilevel data, mixtures, and nonlinearity **challenge optimization procedures** (e.g., dimensionality of integrals) and **theoretical method developments** (e.g., correct likelihoods, quadratures).
- For example, **expectation maximization (EM) algorithm extensions** are needed.
- What is the role of **factor scores** or **2-step procedures**?

Transfer to continuous time dynamic models

Thanks!

Eberhard Karls Universität Tübingen
Wirtschafts- und Sozialwissenschaftliche Fakultät
Methodenzentrum
Hölderlinstr. 29
72074 Tübingen
Germany

Augustin Kelava, Prof. Dr.
augustin.kelava@uni-tuebingen.de

Model setup and priors for the empirical example

Distribution of variables

All observed ($math_{1jit}$) and latent variables η were assumed to follow a normal distribution with the respective mean structure and variance:

$$[math_{1jit}|S_{it} = s] \sim N(\mu_{1jits}, \sigma^2_{\epsilon_{1j}}) \quad (13)$$

$$[\eta_{1it}|S_{it} = 1] \sim MVN(\mu_{\eta,it1}, \Phi_{\zeta_{11}}) \quad (14)$$

$$[\eta_{1it}|S_{it} = 2] \sim N(\mu_{\eta,it2}, \sigma^2_{\zeta_{12}}) \quad (15)$$

where $N(\mu, \sigma^2)$ was the normal distribution with mean μ and variance σ^2.

For the development of each factor of math skills η_{1kit}, we assume an ARIMA(1,1,0) model:

$$[\eta_{1ki1}|S_{i1} = s] = \alpha_{1ks} + \zeta_{1ki1}$$

$$[\eta_{1ki2}|S_{i2} = s] = \alpha_{1ks} + \eta_{1ki1} + \zeta_{1ki2}$$

$$[\eta_{1kit}|S_{it} = s] = \alpha_{1ks} + \eta_{1ki(t-1)} + \omega_{ks} \underbrace{(\eta_{1ki(t-1)} - \eta_{1ki(t-2)})}_{\Delta\eta_{1ki(t-1)s}} + \zeta_{1kit} \text{ for } t > 2 \quad (16)$$

with $\Psi_{1kts} = \Psi_{1ks}$ for all t. Measurement models were assumed to be time invariant.

Markov Switching Model

The probabilities for state membership were modeled using a time- and person-specific latent variable α_{itcd} for $t > 1$ (all persons were assumed to be in state $S_{i1} = 1$ at $t = 1$).

$$P(S_{it} = 1|S_{i(t-1)} = 1) = \frac{\exp(\alpha_{it11})}{\sum_{k=1}^{2} \exp(\alpha_{itk1})} \tag{17}$$

$$P(S_{it} = 2|S_{i(t-1)} = 1) = 1 - P(S_{it} = 1|S_{i(t-1)} = 1) \tag{18}$$

$$P(S_{it} = 1|S_{i(t-1)} = 2) = 0.01 \tag{19}$$

$$P(S_{it} = 2|S_{i(t-1)} = 2) = 1 - P(S_{it} = 1|S_{i(t-1)} = 2) \tag{20}$$

where we chose a very small probability for those students that mastered math to switch back to a non-mastery state of $\pi = 0.01$. The latent variable α_{it11} was specified as:

$$\begin{aligned}\alpha_{it11} = \alpha_{11} + \beta_{11} \cdot read_{i,t-1} + \omega_{13} \cdot (read_{it} - read_{i(t-1)}) + \beta_{21} \cdot motor_i \\ + \omega_{21} \cdot motor_i \cdot read_{i(t-1)} + \omega_{22} \cdot motor_i \cdot (read_{it} - read_{i(t-1)})\end{aligned} \tag{21}$$

Prior distributions

Priors were chosen as weakly informative priors throughout the model. For the **measurement model** on the within level, factor loading and intercept priors were specified as

$$\lambda_{1j} \sim N(1,1), \text{ for } j = 1 \ldots 5 \quad (22)$$

$$\tau_{1j} \sim N(0,2), \text{ for } j = 1 \ldots 4. \quad (23)$$

For the **structural models** coefficients on the within and between levels, again weakly informative priors were chosen:

$$\beta_{11} \sim N(0,1) \quad (24)$$

$$\beta_{21} \sim N(0,2) \quad (25)$$

$$\omega_{13} \sim N(0,1) \quad (26)$$

$$\omega_{2p} \sim N(0,2) \text{ for } p = 1,2 \quad (27)$$

$$\alpha_{11} \sim N(0,2) \quad (28)$$

where the constraint $\alpha_{11} = 0$ was necessary for model identification. Note that this constraint always holds in this model if data are rescaled by $Y^c_{1jit} = Y_{1jit} - \bar{Y}_{111}$ because $Y_{11i1} = \eta_{1i1}$ and all persons are in state $S_{i1} = 1$ at the first measurement occasion.

Standard priors were chosen for the **precisions** as

$$\sigma_{\epsilon_{1j}}^{-2} \sim Gamma(9,4), \text{ for } j = 1 \ldots 5 \quad (29)$$

$$\Phi_{\zeta 11}^{-1} \sim Wishart(\Phi_0^{-1}, 4) \quad (30)$$

$$\sigma_{\zeta_{12}}^{-2} \sim Gamma(9,4) \quad (31)$$

where Φ_0 was a 2×2 identity matrix.