EBERHARD KARLS UNIVERSITÄT TÜBINGEN

METHODS CENTER

Meeting of the SEM Working Group

28 February to 1 March 2019
Tübingen

1 Conference Venue

The meeting will be held at:

Festsaal Alte Aula
Münzgasse 30
72070, Tübingen.

2 Contact information

Local organizer:
Prof. Dr. Augustin Kelava
Methods Center
Hölderlinstr. 29
72074, Tübingen
Mail: augustin.kelava@uni-teuebingen.de
Phone: +49 7071 29 74933

Co-organizer:
Dr. Stefano Noventa,
Methods Center
Hölderlinstr. 29
72074, Tübingen
Mail:
stefano.noventa@uni-tuebingen.de
Phone: +49 7071 29 74928

Secretariat:
Martin Stegmaier
Methods Center
Hölderlinstr. 29
72074, Tübingen
Mail:
martin.stegmaier@uni-tuebingen.de
Phone: +49 7071 29 74932
fax: +49 7071 29 35264

3 Directions

3.1 Tübingen

Tübingen is located on the river Neckar in central Baden-Württemberg. The medieval old town, the striking market square and the picturesque Neckar front attract visitors form near and far. With over 87.000 inhabitants and 28.000 students the Swabian university town gives a modern and cosmopolitan atmosphere. The city centre is a car-free zone and can easily be explored by foot. Everything else can be reached by bus.

How to get to Tübingen:

- **By train**: Tübingen can be reached easily by train, usually via Stuttgart.
- **By plane**: Stuttgart Airport is situated 35 kilometres from Tübingen. By car or taxi Tübingen can be reached in approximately 20 minutes. We recommend taking the shuttle Bus 828 which is running between Tübingen and Stuttgart Airport several times a day.
- **By car**: As there is only a limited number of parking spaces in the city centre, it is not advisable to travel by car.

3.2 Locations

The coference venue is Alte Aula (see number 81 in the map below). Alte Aula is located in the city center, close to the Stiftskirche. The meeting room is located at the ground floor at the level of Münzgasse. The location is more easily accessed from the Münzgasse side, close to the church (see the black arrow). For more information see Alte Aula

The conference dinner will be held at 20:00 of Thursday the 28th at the Restaurant Pizzeria Alte Kunst in Marktgasse 8 (see the map below).

4 List of Hotels

Closest Hotels to the Meeting Venue

- Hotel Krone
- Hotel Domizil
- Ibis Styles Tübingen Hotel
- Hotel Metropol Garni
- Hotel Am Schloss
- Hotel Barbarina
- Hotel Meteora
- Hotel am Kupferhammer

Other Hotels can be found at the following link. Tübingen is a small city, where everything is within walking distance. There are also several bus connections. Scattered around the city there are also pensions, B&Bs and Air B&Bs.

5 Timetable

28 February	
Time	**Title and Presenter**
13:00-13:45	Registration
13:45-14:00	Opening address
14:00-14:30	*Augustin Kelava & Holger Brandt* A nonlinear dynamic latent class structural equation model
14:30-15:00	*Manuel Arnold* Detecting Heterogeneity in Dynamic Panel Models with Individual Parameter Contribution Regression
15:00-15:30	*Rebecca D. Büchner, Julien P. Irmer, & Andreas G. Klein* Model Evaluation of the Heterogeneous Growth Model
15:30-16:00	30 MIN BREAK
16:00-16:30	*Stefano Noventa, Andrea Spoto, Jürgen Heller, & Augustin Kelava* On a generalization of local independence in Item Response Theory
16:30-17:00	*Stella Bollmann & Irini Moustaki* Misspecification of Distribution in Multigroup Latent Variable Models
17:00-17:30	*Daniel W. Heck* Cognitive Psychometrics with Bayesian Hierarchical Multinomial Processing Tree Models
17:30-18:00	Meeting of the working group
20:00	Conference Dinner at Alte Kunst, Marktgasse 8.

1 March

Time	**Title and Presenter**
9:00-9:30	*Tim Schaffland, Stefano Noventa, & Augustin Kelava* Factor Score Estimation: A Comparison of traditional methods with a Non-Parametric Approach
9:30-10:00	*Mariska T. Barendse & Yves Rosseel* Pairwise maximum likelihood for generalized linear mixed models
10:00-10:30	*Andrej Srakar & Miroslav Verbič* Endogeneity corrections in MIMIC models: instrumental variable approaches
10:30-11:00	30 MIN BREAK
11:00-11:30	*Boris Sokolov* How Much Should We Trust Conventional SEM Goodness-of-FitMeasures in Large Cross-National Samples?
11:30-12:00	*Miriam Reußner* A Monte Carlo Simulation Study to Simplify the Examination ofthe Goodness of a Structural Equation Model
12:00-12:30	*Julien P. Irmer, Rebecca D. Büchner, & Andreas G. Klein* Global Model Fit Test for Nonlinear SEM Using the QuasiMaximum Likelihood Method
12:30-14:00	LUNCH
14:00-14:30	*Erik-Jan van Kesteren & Daniel L. Oberski* Extending SEM with insights from deep learning
14:30-15:00	*Manuel Rademaker* cSEM: composite-based Structural Equation Modeling
15:00-15:30	30 MIN BREAK
15:30-16:00	*Christoph Kiefer & Axel Mayer* Analyzing Effects in Regressions with Logarithmic Link Function and Latent Covariates
16:00-16:30	*Tamara Schamberger, Florian Schuberth, Jörg Henseler & Theo K. Dijkstra* Robust Partial Least Squares Path Modeling
16:30	Closing address

6 Abstracts

6.1 Thursday 28

A Nonlinear Dynamic Latent Class Structural Equation Model

Augustin Kelava, University of Tübingen
Holger Brandt, The University of Kansas

In this talk, we propose a nonlinear dynamic latent class structural equation model (NDLC-SEM). It can be used to examine intra-individual processes of observed or latent variables. These processes are decomposed into parts which include individual- and time-specific components. Unobserved heterogeneity of the intra-individual processes are modeled via a latent Markov process that can be predicted by individual-specific and time-specific variables as random effects. We discuss examples of sub-models which are special cases of the more general NDLC-SEM framework. Furthermore, we provide empirical examples and illustrate how to estimate this model in a Bayesian framework. Finally, we discuss essential properties of the proposed framework, give recommendations for applications, and highlight some general problems in the estimation of parameters in comprehensive frameworks for intensive longitudinal data.

Detecting Heterogeneity in Dynamic Panel Models with Individual Parameter Contribution Regression

Manuel Arnold, Humboldt-University of Berlin

Dynamic panel models are among the most popular statistical techniques for modeling the interrelationships of one or multiple repeatedly measured variables. In psychological research, dynamic panel models are often specified and estimated within the structural equation modeling (SEM) framework. An endemic problem that threatens the validity of inferences from dynamic panel models is unmodelled heterogeneity due to stable differences between individuals or groups. Recently, individual parameter contribution (IPC) regression was proposed as a flexible method to study heterogeneity in SEM parameters as a function of observed covariates. In this talk, we show how IPCs can be derived for general maximum likelihood estimates (including SEM as a special case) and evaluate the performance of IPC regression to estimate group differences in bivariate dynamic panel models in discrete and continuous time. We demonstrate that IPC regression can be slightly biased in samples with large group differences and, as a solution to this problem, present a correction procedure. In a simulation study, IPC regression showed promising results for discrete-time dynamic panel models but exhibited small power to detect heterogeneity in continuous-time models.

Model Evaluation of the Heterogeneous Growth Model

Rebecca D. Büchner, Goethe University of Frankfurt Julien P. Irmer, Goethe University of Frankfurt Andreas G. Klein, Goethe University of Frankfurt

Klein and Muthén (2006) proposed an extension of conventional growth curve models for the modeling of heterogeneous growth rates (heterogeneous growth curve model, HGM). The HGM allows the modeling of heterogeneous growth rates as a continuous function of latent initial status and time-invariant covariates. A quasi maximum-likelihood approach is used to estimate the HGM. Standard tests for the assessment of model fit in ordinary latent growth models, for instance the χ^2 test, are not appropriate for the HGM. Therefore, we propose a new quasi-likelihood ratio test (Q-LRT) to assess the global model fit of the HGM. Results of a simulation study evaluating the performance of the test show that the Q-LRT performs well with satisfactory Type I error and desirable power rates. A comparison with the χ^2 test shows that - in contrast to the Q-LRT - the χ^2 test ails to detect important misspecifications.

On a Generalization of Local Independence in Item Response Theory

Stefano Noventa, University of Tübingen
Andrea Spoto, University of Padova
Jürgen Heller, University of Tübingen
Augustin Kelava, University of Tübingen

Knowledge Space Theory (KST) structures can be introduced within Item Response Theory (IRT) as a possible way to model local dependence (LD) between items. This allows to generalize the usual characterization of local independence without introducing new parameters and to merge the information provided by the IRT and KST perspectives. In detail, connections are established between the KST Simple Learning Model (SLM) and the IRT General Graded Response Model (GRM), and between the KST Basic Local Independence model (BLIM) and IRT models in general. As a consequence, IRT models become a subset of KST models and both local independence assumption and IRT likelihood functions can be generalized to account for the existence of prerequisite relations between the items. Considerations are drawn for the modeling of both dichotomous and polytomous items, and for their interpretation (e.g., relevance and meaning of the parameters, definition of polytomous items as knowledge structures of dichotomous ones, interpretation of Rasch model as a probabilistic version of Guttman's scale).

Misspecification of Distribution in Multigroup Latent Variable Models

Stella Bollmann, UZH Zurich.
Irini Moustaki, LSE London

Marginal maximum likelihood estimation for 2PL Item Response Theory (IRT) models assumes normally distributed latent variables. Previous research has shown that when this assumption is violated, estimation of item parameters is biased (e.g. Stone, 1992) and estimation of person parameters (i.e. of the latent variable) is also error-prone (Sass, Schmitt, Walker, 2008). In this study, we want to investigate how mis-specification of the latent variable distribution can affect the estimation and fit of multigroup latent variable models and in particular measurement equivalence. Multigroup latent variable models are designed for the use in multigroup analyses as e.g. cross-national surveys where countries are compared on constructs such as behavior, attitudes, ability etc. A grouping variable for different populations is included that allows for different parameter estimates in each group (e.g. nation). In multigroup analysis, it is important to establish measurement equivalence to allow for comparative analysis. Nonequivalence is operationalized as an association between the group and an item (Kuha & Moustaki, 2015). The research question of this work is, whether nonequivalence might be erroneously detected in data sets that are fully equivalent but exhibit a violation of normality in the latent variable. Non-normality will be generated using the sn-package by Azzalini (2014).

References

Azzalini, A. (2014). *The R 'sn'package: The skew-normal and skew-t distributions (version 1.0-0).* Università di Padova, Italia.

Kuha, J., & Moustaki, I. (2015). Nonequivalence of measurement in latent variable modeling of multigroup data: A sensitivity analysis. *Psychological methods*, 20 (4), 523.

Sass, D. A., Schmitt, T. A., & Walker, C. M. (2008). Estimating non-normal latent trait distributions within item response theory using true and estimated item parameters. *Applied Measurement in Education*, 21 (1), 65-88.

Stone, C. A. (1992). Recovery of marginal maximum likelihood estimates in the two-parameter logistic response model: An evaluation of MULTILOG. *Applied Psychological Measurement*, 16 (1), 1-16.

Cognitive Psychometrics with Bayesian Hierarchical Multinomial Processing Tree Models

Daniel W. Heck, University of Mannheim

Many psychological theories assume that different cognitive processes can result in identical observable responses. A typical example is the recognition-memory paradigm, in which correct responses may be either due to the retrieval of the test item or due to lucky guessing. Multinomial processing tree (MPT) models allow disentangling such latent cognitive processes based on observed response frequencies. Recently, MPT models have been extended to explicitly account for participant and item heterogeneity by assuming specific group-level distributions of the person parameters. These hierarchical Bayesian MPT models provide the opportunity to connect traditionally isolated areas of psychology, since they combine cognitively meaningful parameters with concepts from psychometrics and personality research. Whereas cognitive psychology has often focused on the experimental validation of MPT models on the group level, item-response theory provides the necessary tools and concepts for measuring MPT parameters on the item or person level. Moreover, MPT parameters can be regressed on covariates to predict latent cognitive processes using personality traits or other person characteristics. Given that user-friendly software for the estimation of hierarchical MPT models has recently become available, this model class can serve as a general conceptual framework for cognitive psychometrics that includes both MPT models and item response theory as special cases.

6.2 Friday 1

Factor Score Estimation: A Comparison of traditional methods with a non-parametric approach

Tim Schaffland, University of Tübingen
Stefano Noventa, University of Tübingen
Augustin Kelava, University of Tübingen

Estimation of factor scores in latent variable models has repeatedly attracted the interest of researchers for decades. Already in 1935 Thurstone proposed the regression method, and in 1937 Bartlett suggested his well-known approach. Still today, factor score estimation and their properties, for example the bias of their moments, raise debate and interest (see, e.g., Grice, 2001; Hoshino & Bentler, 2013). In this talk we will compare the Bartlett estimator, the regression method, the least square estimation, and one new approach (Kelava, Kohler, Krzyzak, & Schaffland 2017) which makes no distributional assumptions on the latent variables. Factor scores are estimated by combining the empirical CDF and the independence assumption between the measurement errors and the latent factors. This results in factor score estimates that in theory could consistently replicate the true joint distribution of the latent variables and the measurement error. In a simulation study we vary the (multivariate) distribution of the underlying factors and examine the performance of the different approaches in recovering the first four moments of the joint distribution of the latent variables. This talk concludes with the implications and recommendations for factor score estimation in an applied context.

Pairwise Maximum Likelihood for Generalized Linear Mixed Models

Mariska T. Barendse, University of Ghent
Rosseel Yves, University of Ghent

Generalized linear mixed models (GLMM) are often used with hierarchical data or multilevel data, where units are nested in clusters, which in turn may be nested in even larger clusters. Inherent in GLMMs and discrete data is that the marginal likelihood is obtained after integrating over the random effects. This approach is very intensive as it involves intractable integrals. An alternative approach is the pairwise maximum likelihood (PML) estimation method, where only the sum of all bivariate likelihoods is maximized. For this reason the PML estimation method is a compromise between computational burden and loss of efficiency. In the context of multilevel models with a probit link and binary responses, Renard, Molenberghs, and Geys (2004) studied the PML estimation method with a random slope and a random intercept. Little is yet known about how the PML estimation method performs with more extended GLMM models

and/or different kind of response options. The current study will further investigate the possibilities of the PML estimation method with GLMM models and compare it to the often used marginal maximum likelihood. In a small simulation study, we vary the type of response options and the number of random/fixed effects to determine the accuracy and the efficiency of the estimated model parameters.

Endogeneity Corrections in MIMIC Models: Instrumental Variable approaches

Andrej Srakar, University of Ljubljana
Miroslav Verbič, University of Ljubljana

Endogeneity is known to arise in the situations that feature: (1) omitted relevant variables, (2) measurement error in the explanatory variables, (3) self-selection, (4) simultaneity, and (5) serially correlated errors injected by lagged explanatory variables (Ruud, 2000; Ebbes, 2004; Hueter, 2016). In MIMIC models, endogeneity has so far been explored for measurement error models (Tekwe et al., 2014). We provide a new estimator appropriate for simultaneity endogeneity problems, based on basic two-stage and 2SLS approaches. We derive the asymptotic theory for the estimator in light of the original contribution of Jöreskog and Goldberger (1975) and show the estimator is consistent and superefficient (Le Cam, 1953). We discuss nonparametric and Bayesian extensions. The performance of the estimator is compared to "regular" MIMIC estimator in a simulation study and applied to examples of composite indicators for precarious work (cross-sectional application) and deaccessioning (panel application).

How Much Should We Trust Conventional SEM Goodness-of-Fit Measures in Large Cross-National Samples?

Boris Sokolov, Higher School of Economics, St. Petersburg

Using Monte Carlo simulation experiments, this project examines how well popular SEM goodness-of-fit measures (namely, CFI, TLI, RMSEA, and SRMR) perform in the context of measurement invariance testing in large samples. Its contribution to the existing methodological literature on cross-national survey research is three-fold. First, it explores how sensitive are the aforementioned fit measures to various amounts of measurement non-invariance in large samples (10-30-50 groups) under various conditions imitating typical features of such type of survey data. Second, it tests how other model misspecifications affect model fit in the multi-group setting, thus disentangling the impact of different fit-worsening factors (non-invariance vs. other misspecifications). The results suggest that CFI and SRMR are superior to RMSEA and TLI as measures of model misfit due to non-invariance, but the existing cut-off values for all these measures are too strict and must be somewhat softened. Finally, it examines how critical are different levels of non-invariance in terms of bias in the latent means hierarchy. The results show that the danger of measurement invariance might be somewhat exaggerated since even in conditions with the highest levels of metric and scalar non-invariance the estimated latent means do not deviate strongly from the corresponding true population values.

A Monte Carlo Simulation Study to Simplify the Examination of the Goodness of a Structural Equation Model

Miriam Reußner, University of Bremen

The rating of a linear structural equation model (SEM) is an essential part of the estimation process. One way of assessing the fit of a SEM is to use the χ^2 test. Due to this test's sensitivity to the sample size, fit indices have become a popular alternative for evaluating a SEM. However, based on established cut-off values of the fit indices, the evaluation of a SEM via multiple fit indices may produce contradictory results – with one index rejecting the model that is accepted by another index. My simulation study evaluates the effects of different data and model structures - including variations of the sample size, the distributions of the indicators, model complexity, and model specification - on four of the most established fit indices, namely: TLI, CFI, RMSEA and SRMR. Contrary to large parts of the literature, the results reveal sensitivities of all fit indices to skewed indicators. Additionally, using the TLI and the RMSEA results in more strict and thereby correct ratings of misspecified models than solely relying on the CFI and the SRMR. Surprisingly, different indicator distributions affect the TLI and CFI and therefore need to be further studied.

Global Model Fit Test for Nonlinear SEM Using the Quasi Maximum Likelihood Method

Julien P. Irmer, Goethe University of Frankfurt
Rebecca D. Büchner, Goethe University of Frankfurt
Andreas G. Klein, Goethe University of Frankfurt

Nonlinear structural equation modeling (SEM) has received much attention in recent years, enabling a more detailed specification of a model by extending the structural equation. A problem is still the evaluation of model fit. Goodness of fit of conventional SEM is usually evaluated by performing a likelihood ratio test (the χ^2 test), which compares the target model (the model derived by theory) to a saturated model (a well fitting more general model). For quadratic SEM, the χ^2 test cannot be used because an adequate saturated model does not exist. Within a master's thesis we, therefore, propose a new quasi-likelihood ratio test (Q-LRT) for a special case of nonlinear SEM equivalent to the likelihood ratio test of linear SEM based on the quasi-maximum likelihood method (QML, Klein & Muthén, 2007). The new test, a revised version of the Q-LRT derived from the simplified QML method developed by Büchner and Klein (2019), includes a proper saturated model especially tailored for quadratic SEM. First results from a Monte Carlo study conducted for this master's thesis show that the fit measure Q-LRT performs well with regard to Type I error rates and power rates, when sample size is sufficiently large.

Extending SEM with insights from deep learning

Erik-Jan van Kesteren, Utrecht University
Daniel L. Oberski, Utrecht University

How can we perform SEM on data with more variables than cases, such as with MRI data or genetic data? What do we do when SEM models such as MTMM models are identified but empirically underdetermined, leading to nonconvergence in practice? How can our SEM model be updated in an online fashion if sensor data keeps streaming in but we never have the full data vector? Such problems have equivalents in the field of deep learning, where solutions such as dropout, norm regularisation, and stochastic gradient descent are readily available. In this talk, I will show how SEM can be represented as a TensorFlow program. TensorFlow is an efficient optimisation library targeted at estimating neural networks. This representation thus enables SEM researchers to leverage the computational solutions developed by the deep learning community, paving the way for extending SEM in new directions. At the core of this library is a computation graph which represents the operations leading from parameters to a loss value (see Figure 1). This computation graph allows for automatic gradient computation. I will illustrate how two extensions to SEM - Regularised SEM and Least Absolute Deviation Estimation - can be conveniently implemented within this optimisation framework.

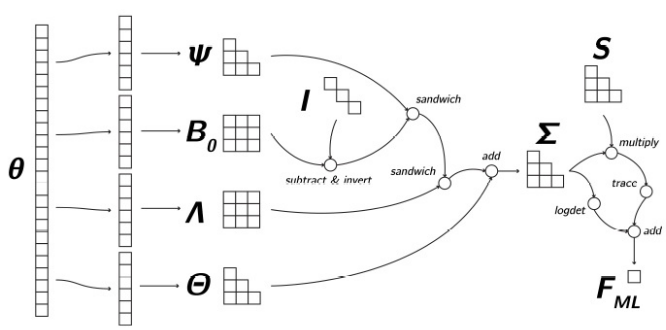

Figure 1: Computation graph for maximum likelihood-based SEM in Tensorflow

`cSEM`: composite-based Structural Equation Modeling

Manuel Rademaker, University of Würzburg
Florian Schuberth, University of Twente

The goal of the `cSEM` package is to provide the user with a unified interface to (eventually) all algorithms, approaches, methods, estimators, as well as typical postestimation procedures (e.g. assess, test, verify/check and summarize the model) that qualify as composite-based. Currently, this includes PLS and its descendants (PLSc or ordPLSc), GSCA (and descendants such as GSCAm), all of Kettenring approaches and simpler schemes such as SumScores. In its current state, the package automatically distinguishes constructs modeled as common factors and constructs modeled as composites and is able to handle linear models, non-linear models, models containing second-order constructs, and models/data with a group structures. Apart from its content that is only party found in other packages, the package' design philosophy is **strictly user centered** in a sense that the often observed hesitation related to using a software like `R` is minimized by designing the package API to be as independent as possible from the type of data, the procedure, and the type of model used: users provide a model (in `lavaan syntax`) and a dataset to cSEM's main function `csem()`. Independent of the model or dataset provided, a concise set of postestimation verbs (`assess()`, `test()`, `verify()`, `summarize()`) can then be applied on the resulting object to carry out typical postestimation tasks.

Analyzing Effects in Regressions with Logarithmic Link Function and Latent Covariates

Christoph Kiefer, RWTH Aachen University
Axel Mayer, RWTH Aachen University

The effectiveness of a treatment on a count outcome is often assessed using a regression with logarithmic link function (e.g., negative binomial regression). A treatment effect (a.k.a. marginal effect) is commonly defined as the difference between the expected outcome under treatment and under control. In a regression with logarithmic link function, such treatment effects can to date only be estimated if covariates are manifest (observed). Nevertheless, in some scenarios it might be beneficial to control for latent covariates, for example, in order to avoid attenuation bias and to get unbiased treatment effect estimates. We propose a new approach to compute average and conditional treatment effects in regressions with logarithmic link function involving latent covariates. We provide an illustrative example to explain the application and estimation in Mplus.

Robust Partial Least Squares Path Modeling

Tamara Schamberger, University of Twente and University of Würzburg
Florian Schuberth, University of Twente
Jörg Henseler, University of Twente
Theo K. Dijkstra, University of Groningen

Outliers can seriously distort the results of statistical analyses and threaten the validity of structural equation models. As a remedy, this article introduces a robust variant of Partial Least Squares Path Modeling and its consistent version called robust PLS/PLSc, which is robust against distortion caused by outliers. Robust PLS/PLSc allows to consistently estimate structural models with constructs modeled as composites and common factors even if empirical data are contaminated with outliers. A Monte Carlo simulation with various population models, sample sizes and extents of outliers shows that robust PLS/PLSc can deal with outlier shares of up to 50% without distorting the estimates. The simulation also shows that robust PLS/PLSc should always be preferred over its traditional counterparts if the data contain outliers.