Meeting of the Working Group on Structural Equation Modeling

March 18-19, 2021

Digital

"Venue": Zoom meeting

Link: https://bokuvienna.zoom.us/j/94730221481

Online registration

In order to facilitate the planning of the event for us, we kindly ask you to register for the meeting **until March 16th** using the following Google Forms.

Link: https://forms.gle/8HCQvvMF2kYqyUuA8

Contact information local organizers

Local organizer:
Reinhard Hössinger
Institute for Transport Studies
University of Natural Resources and Life Sciences
r.hoessinger@boku.ac.at
+43 1 47654-85631

Co-organizers:
Juliane Stark, juliane.stark@boku.ac.at
Maria Juschten, maria.juschten@boku.ac.at

Secretariat:
Michaela Stern, michaela.stern@boku.ac.at
+43 1 47654-85619

Head of the Institute:
Astrid Gühnemann

Preliminary Timetable

Thursday, March 18th

Time	Title & Author(s)	
09.00 – 09.30	Welcome to Vienna!	
09.30 – 10.00	01	*David Duran-Rodas, Francisco Camara Pereira & Gebhard Wulfhorst* Exploring causality of built and social environment factors influencing the observed demand of bike-sharing systems
10.00. – 10.30	02	*Katharina Meitinger, Peter Schmidt & Michael Braun* Detecting and explaining measurement inequivalence: the case of patriotic feelings
10.30 – 11.00	03	*Suzanne Jak & Terrence Jorgensen* Modeling cluster-level constructs with individual-level measures
11.00 – 11.30	Viennese coffee break	
11.30 – 12.00	04	*Katharina Groskurth, Nivedita Bhaktha*, Matthias Bluemke*, Clemens M. Lechner*, & Thorsten Meiser* (*alphabetical order)* Moving Away From Fixed Cutoffs: What We Gain From ROC Analysis
12.00 – 12.30	05	*Steffen Grønneberg & Njål Foldnes* Nonparametric tetrachoric correlations
12.30 – 13.00	06	*Njål Foldnes & Steffen Grønneberg* The sensitivity of structural equation modeling with ordinal data to underlying non-normality and observed distributional forms
13.00 – 14.00	Lunch break	
14.00 – 14.30	07	*Maria Faust* SEM of Temporal Digital Change in Germany and China
14.30 – 15.00	08	*Daniel Seddig & Heinz Leitgöb* Decomposing Differences in Indicator Means across Groups. A Multi Group Confirmatory Factor Analysis Approach
15.00 – 15.30	09	*Henrik Andersen & Jochen Mayerl* Applying Panel Regression in the Structural Equation Modelling Framework to Assess Relationships between Environmental Values and Attitudes
15.30 – 15.45	Closing Day 1	

Friday, March 19th

Time	Title & Author(s)
09.00 – 09.30	10 *Yves Rosseel* The structural-after-measurement (SAM) approach for SEM
09.30 – 10.00	11 *Tamara Schamberger, Florian Schuberth & Jörg Henseler* Maximum Likelihood Estimator for Composite-Based Structural Equation Modeling
10.00 – 10.30	12 *Harry Garst* Growth curve models with estimated changepoints
10.30 – 11.00	Viennese coffee break
11.00 – 11.30	13 *Andrej Srakar & Tjaša Bartolj* Bayesian nonparametric estimation in longitudinal mediation: A Baron-Kenny based estimator for cross-lagged models
11.30 – 12.00	14 *Rolf Steyer* Why Interaction Matters
12.00 – 12.30	15 *Julien Irmer, Andreas G. Klein, Jana Gäde & Karin Schermelleh-Engel* When data are not perfect: Robustness of LMS compared to regression methods when categorical data are treated as continuous
12.30 – 13.30	Lunch break
13.30 – 14.00	16 *Roman Zachary, Holger Brandt* A Latent Auto-regressive Approach for Bayesian Structural Equation Modeling of Spatially or Socially Dependent Data
14.00 – 14.30	17 *Mariska Barendse & Yves Rosseel* Pairwise maximum likelihood for multilevel data
14.30 – 14.45	Closing Day 2

Abstracts of Presentations – Thursday, March 18th

01 - Exploring causality of built and social environment factors influencing the observed demand of bike-sharing systems

David Duran-Rodas [1], Francisco Camara Pereira [2], Gebhard Wulfhorst [1]

Chair of Urban Structure and Transport Planning, Department of Civil, Geo and Environmental Engineering, Technical University of Munich

Department of Management Engineering, Technical University of Denmark

Keywords: *bike-sharing, spatial analysis, built environment, lifestyle, causality*

Bike-sharing systems (BSSs) have been implemented in around 2000 cities worldwide and are still growing. For the optimal expansion of the systems, previous work has identified spatial factors associated with observed demand such as land use, transport infrastructure, points of interest and sociodemographic. However, some of these factors do not have a causal relationship with the observed demand. Therefore, the main objective of this research is to develop a theoretical framework to identify the causality of the most influencing factors on BSS using built and social environmental factors with structural equation models (SEM).

First, we developed our theoretical framework by merging the transport and land-use interaction model with the concept of urban mobility culture (see Figure 1). Then, we collected observed demand data from a BSS and built and social environment variables. We selected the most associated spatial factors to the observed demand based on linear and non-linear models. Finally, we built SEM based on the theoretical framework with the most influencing spatial factors. The method was applied to the hybrid BSS in Munich. Built environment variables included land use, transportation infrastructure, points of interest, and the social environment (sociodemographic, lifestyle milieus and social media usage). As an outcome, we expect to have a deeper understanding of the causality of the most influential spatial factors from the built and social environment on the observed demand in the BSS.

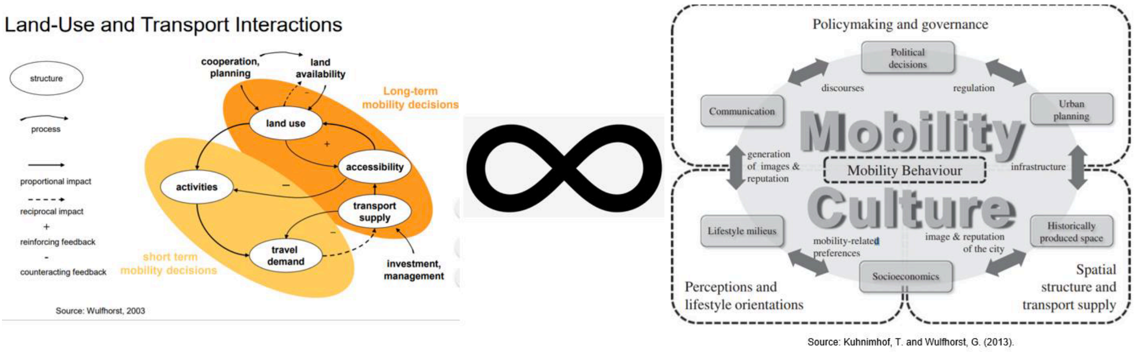

Figure 1: Theoretical framework

02 - Detecting and explaining measurement inequivalence: the case of patriotic feelings

Katharina Meitinger [1], Peter Schmidt [2], Michael Braun [3]

[1] Department of Methodology and Statistics, Utrecht University, Netherlands

[2] Institute of Political Science, University of Gießen, Germany

[3] GESIS – Leibniz Institute for the Social Sciences, Mannheim, Germany

Testing for measurement invariance is an important precondition to draw substantive conclusions from cross-national data. However, the traditional approach of multigroup confirmatory factor analysis (MGCFA) has been criticized as too strict and more liberal approaches have recently been proposed, such as alignment. Mixed-methods approaches combining quantitative measurement invariance tests and qualitative insights from web probing provide a powerful tool to address this issue. For this study, we selected the substantive example of the item battery on "Political Effects of Nationalism" from the 2013 ISSP Module on National Identity. With MGCFA, measurement invariance tests failed to show metric and scalar invariance indicating that structural coefficients and latent means should not be compared across countries. With alignment, scalar measurement invariance was confirmed. However, the web probing results point to several problematic issues that potentially question the comparability of results. The qualitative findings are mirrored in the MGCFA results but not in the alignment outcome. This study underlines the value of mixed methods approaches in the toolkit of cross-national researchers and generally those studying multiple groups since it provides the opportunity to detect and address issues of item and construct bias with qualitative insights.

03 - Modeling cluster-level constructs with individual-level measures

Suzanne Jak & Terrence D. Jorgensen

[1] Methods and Statistics, Child Development and Education, University of Amsterdam

Researchers frequently use the responses of individuals in clusters to measure constructs at the cluster level. For example, student's evaluations may be used to measure the teaching quality of instructors, patient reports may be used to evaluate social skills of therapists, and residents ratings may be used to evaluate neighborhood safety. When multiple items are used to measure such cluster-level constructs, multilevel confirmatory factor models are useful. These models allow for the evaluation of the factor structure at the cluster level (modeling the (co)variances among item means across clusters), and at the individual level (modeling the (co)variances across individuals within clusters). If the cluster-level construct, for example teacher quality, would be perfectly measured using the responses of students, all students evaluating the same teacher would agree, and provide exactly the same item scores. In that case, there will not be any systematic variance in the item scores within clusters (only sampling error), so there will be nothing to model at the individual level.

In practice, the individuals do not all provide the same responses to the items, leading to systematic variance (and covariance) to be explained at the individual level. The question then arises how the variance within clusters should be modeled. Stapleton et al. (2016) proposed a model for cluster-level constructs with a saturated model at the individual level. This work was updated by proposing and evaluating several two-level factor models in Stapleton and Johnson (2019). In this article, we will provide a simulation study to evaluate under what scenario's the different models are able to provide sensible results, partly replicating the study by Stapleton and Johnson (2019).

04 - Moving Away From Fixed Cutoffs: What We Gain From ROC Analysis

Katharina Groskurth [1,2], Nivedita Bhaktha [1], Matthias Bluemke* [1], Clemens M. Lechner* [1] & Thorsten Meiser* [3] (*alphabetical order)*

[1] GESIS – Leibniz Institute for the Social Sciences, Mannheim, Germany

[2] Graduate School of Economic and Social Sciences, University of Mannheim, Germany

[3] School of Social Sciences, University of Mannheim, Germany

Keywords: *goodness-of-fit, fit index, cutoff, confirmatory factor analysis, structural equation modeling*

Evaluation of confirmatory factor analysis models is a day-to-day business of quantitative psychologists. Decisions on model acceptance or rejection are typically based on a fixed set of traditional cutoffs for fit indices (e.g., CFI, RMSEA, SRMR). The cutoffs are derived in methodological studies (e.g., Hu & Bentler, 1999) simulating a limited set of different model, analysis, and data settings. However, those cutoffs for fit indices are applied ubiquitously in diverse empirical settings. Problematically, fit indices do not only react to misspecification, as they are intended to, but also to model, analysis, and data characteristics in many complex ways. When empirical settings diverge from the initial simulated ones, the fixed set of cutoffs should not be applied, prohibiting valid model evaluation in the multiplicity of empirical settings. Although most researchers are aware of the erroneous conclusions fixed cutoffs might lead to, model evaluation via fixed cutoffs still remains the status quo. Easy-to-derive cutoffs tailored to the empirical model, analysis, and data at hand are in strong need. We provide a novel approach to generate tailored (i.e., flexible) cutoffs for fit indices suited for the model, analysis, and data of interest via a Receiver Operating Characteristic (ROC) analysis. We advance model evaluation by equipping researchers with an easy-to-use procedure to simulate their own data and to pick the most accurate cutoffs for their empirical problem. The approach can be implemented within a few steps only in standard statistical programs such as R.

05 - Nonparametric tetrachoric correlations

Steffen Grønneberg, Njål Foldnes

Department of Economics, BI Norwegian Business School

The tetrachoric correlation is a popular measure of association for binary data and estimates the correlation of an underlying normal latent vector. However, when the underlying vector is not normal, the tetrachoric correlation will be different from the underlying correlation. Since assuming underlying normality is often done on pragmatic and not substantial grounds, the estimated tetrachoric correlation may therefore be quite different from the true underlying correlation that is modeled in structural equation modeling. This motivates studying the range of latent correlations that are compatible with given binary data, when the distribution of the latent vector is partly or completely unknown. We show that nothing can be said about the latent correlations unless we know more than what can be derived from the data. We identify an interval constituting all latent correlations compatible with observed data when the marginals of the latent variables are known. Also, we quantify how partial knowledge of the dependence structure of the latent variables affect the range of compatible latent correlations. Implications for tests of underlying normality are briefly discussed.

06 - The sensitivity of structural equation modeling with ordinal data to underlying non-normality and observed distributional forms

Njål Foldnes, Steffen Grønneberg

Department of Economics, BI Norwegian Business School

Structural equation modeling (SEM) of ordinal data is often performed using normal theory maximum likelihood estimation based on the Pearson correlation (cont-ML) or using least squares principles based on the polychoric correlation matrix (cat-LS). While cont-ML ignores the categorical nature of the data, cat-LS assumes underlying multivariate normality. Theoretical results are provided on the validity of treating ordinal data as continuous when the number of categories increases, leading to an adjustment to cont-ML (cont-ML-adj). Previous simulation studies have concluded that cat-LS outperforms cont-ML, and that it is quite robust to violations of underlying normality. However, this conclusion was based on a data simulation methodology equivalent to discretizing exactly normal data. The present study employs a new simulation method for ordinal data to re-investigate whether ordinal SEM is robust to underlying non-normality. In contrast to previous studies, we include a large set of ordinal distributions, and our results indicate that ordinal SEM estimation and inference is highly sensitive to the interaction between underlying non-normality and the ordinal observed distributions. Our results show that cont-ML-adj consistently outperforms cont-ML, and that cat-LS is less biased than cont-ML-adj. The sensitivity of cat-LS to violation of underlying normality necessitates the need for a test of underlying normality. A bootstrap test is found to reliably detect underlying non-normality.

07 - SEM of Temporal Digital Change in Germany and China

Maria Faust

Department for Empirical Communication and Media Research, University of Leipzig

Former research has shown that internet-mediated communication leads to a change in which we deal with time in everyday life and plan differently. This process is part of a cultural change in time. Such change is due to increasing social interaction on the internet and also because the routines of journalism have changed (Neuberger, 2010). However, this change was described on a theoretical level only (e.g. Castells, 2010; Eriksen, 2001; Hassan, 2003; Innis, 2004; Krotz, 2001; Neverla, 2010; Nowotny, 1995; Rosa, 2005etc.) There is a clear research desiderate in quantitative empirical analysis. This paper seeks to fill this gap and therefore suggests a structural equation model. The novelty of this approach lies in the first multivariate quantitative analysis of mediatized processes of temporal change on a societal level. Hereby a de-westernized (Gunaratne, 2010) most-different systems design is applied (Anckar, 2008) where German and Chinese Cultural Contexts are picked. Temporal understanding as the dependent variable is an eight-dimensional construct (Faust, 2016) and consists of Western and Chinese notions (Chinese Culture Connection, 1987). It integrates the anthropological constructs of past, polychronicity and monochronicity (Bluedorn, Kalliath, Strube, & Martin, 1999; Hall, 1984; Lindquist & Kaufman-Scarborough, 2007), fatalism, pace of life (Levine, 1998), temporal horizon (Klapproth, 2011) and a Chinese future sub-dimension (Chinese Culture Connection, 1987). All of these eight temporal sub-dimensions are subject to change with the overarching hypothesis:

> *Societal norms and values shift towards a more fatalistic, short-term, more multi-tasking oriented and less monochronic, yet accelerated lifestyle through internet-mediated communication.*

08 - Decomposing Differences in Indicator Means across Groups. A Multi Group Confirmatory Factor Analysis Approach

Daniel Seddig [1], Heinz Leitgöb [2]

[1] *University of Cologne, Germany)*

[2] *University of Eichstätt-Ingolstadt, Germany*

Keywords: *panel attrition, CFA panel model, measurement nonequivalence, cognitive model of survey response*

Drawing valid inferences about systematic differences in latent constructs between groups (including independently drawn random samples of some underlying population in the case of repeated cross-sectional survey designs) presupposes the measurement invariance (MI) assumption to be satisfied. If the assumption does not hold, group-specific differences in the understanding or meaning of the indicators applied to measure the latent construct of interest interfere with the between-group true differences in the respective latent variables (e.g., the differences in latent means). Strategies for detecting measurement non-invariance using multiple group confirmatory factor analysis (MGCFA) are well established. Thus, this contribution will focus on situations in which MI could not be achieved. We propose a decomposition approach that allows decomposing systematic differences in the group-specific indicator means into a component reflecting the true difference in the latent means and two components reflecting the differences in the underlying CFA model's parameters. Conceptually, the approach is based on the principles of the Blinder-Oaxaca decomposition (Blinder 1973; Oaxaca 1973) and has already been implemented for the CFA panel model (Oort 2005). In addition to extending the Oort's (2005) decomposition model to the multi-group-case, we contribute to the literature by (i) introducing a different strategy of scale identification, (ii) deriving a four-fold decomposition including an interaction term of differences in latent means and factor loadings, and (iii) providing standard errors for the respective decomposition components based on the delta method. We demonstrate the proposed decomposition approach with an artificial data example.

09 - Applying Panel Regression in the Structural Equation Modelling Framework to Assess Relationships between Environmental Attitudes, Behavioural Intentions and Behaviours

Henrik Andersen & Jochen Mayerl

Faculty of Behavioural and Social Sciences, Institute of Sociology, Chemnitz University of Technology

Keywords: *Environmental attitudes and behaviour, reciprocal effects, unobserved heterogeneity, panel data analysis, cross-lagged panel models*

The question as to whether pro-environmental behaviour can be encouraged by creating more positive attitudes towards the environment has been investigated for many decades. Namely, there are a number of theoretical and empirical hurdles to assessing environmental attitude-behaviour consistency. For one, effects tend to depend on the specific behaviour in question. Secondly, the sheer number of suggested mechanisms is staggering, ranging from sociodemographic, to attitudinal to situational. Third, it is not even clear whether the assumed causal direction is correct. Due to cognitive dissonance or the so-called behavioural snowball effect, pro-environmental behaviour may cause positive attitudes towards the environment. And finally, the constructs of interest are notoriously difficult to measure. Panel regression in the structural equation modeling (SEM) framework provides a good basis for tackling many of these issues. Latent variable measurement models can deal with measurement error in the observed variables. Fixed effects (FE) panel regression

allows us to control for the majority of the often-cited confounders. Panel regression allows us to investigate reverse causality as well as measurement invariance testing to ensure the meaning of the constructs of interest has not changed over time. Using six waves of the GESIS Panel survey, we implement FE panel regression in SEM to investigate the relationship between several different specific types of environmental attitudes and behaviours.

Abstracts of Presentations – Friday, March 19th

10 - The structural-after-measurement (SAM) approach for SEM

Yves Rosseel

Department of Data Analysis, Ghent University, Belgium

Keywords: *two-step approaches, estimation, small-samples*

In the structural-after-measurement (SAM) approach, estimation proceeds in several steps. In a first step, only parameters related to the measurement part of the model are estimated. In a second step, parameters related to the structural part (only) are estimated. Several implementations of this old idea will be presented. A distinction will be made between local and global SAM, and it will be suggested that various alternative estimators (including non-iterative estimators) could be used for the different model parts. It turns out that this approach is not only effective in small samples, but it is also robust against many types of model misspecification. Many existing alternatives (factor score regression with Croon corrections, sum scores with fixed reliabilities, model-implied instrumental variables estimation, Fuller's method, …) turn out to be special cases of this general framework.

11 - Maximum Likelihood Estimator for Composite-Based Structural Equation Modeling

Tamara Schamberger [1,3], Florian Schuberth [1] & Jörg Henseler [1, 2]

[1]Department of Design, Production and Management, Faculty of Engineering Technology, University of Twente

[2] NOVA Information Management School, Universidade Nova de Lisboa

[3] Faculty of Business Management and Economics, University of Würzburg

Keywords: *composites, structural equation modeling, maximum likelihood estimation*

Over the last two decades, composite-based structural equation modeling (SEM), i.e., structural equation modeling with concepts modeled as composites, gained more and more attention. However, the choice of estimators for estimating composite-based structural equation models is currently rather limited. Somewhat surprisingly, the estimation by typical estimators applied in the context of SEM such as the maximum likelihood estimator is currently limited to models containing composites in an exogenous position in the structural model. To overcome this issue, we introduce a maximum likelihood estimator for composite-based structural equation models.

To investigate the nite behavior of our proposed ML estimator and to compare its results to existing estimators such as partial least squares path modeling (PLS-PM) and generalized structured component analysis (GSCA), we conduct a Monte-Carlo simulation. The results

showed that ML performs similar to PLS-PM and GSCA.

12 - Growth curve models with estimated changepoints

Harry G.J.A. Garst

Faculty of Social and Behavioural Sciences, University of Amsterdam

The use of intensive longitudinal datasets requires specific analysis techniques. In the SEM framework, growth curve models may not be appropriate for long series of longitudinal data. For linear growth curve models the limits may have already been reached by data sets extending more than five measurement occasions. Nonlinear growth curve models will clearly extend the range of datasets that can be described reasonably well even in studies with extended timeframes and consisting of many measurement occasions. However, in case there is one or more theoretical changepoints to be expected in the trajectories there is need for models that include change parameters. Using traditional SEM models changepoints can be estimated in piecewise growth curve models, but only for fixed changepoints. This a severe limitation because the timing of changepoints varies in most applications between individuals.

Another complication is that as the number of measurement occasions grows, it is often difficult to have multiple indicators for the same construct at each measurement occasion and therefore models for growth curves for latent variables are out of reach. The alternative to go outside the SEM framework and use estimation techniques for the parameters for each trajectory individually may become an attractive alternative. In this presentation piecewise growth models with unknown changepoint models in the SEM framework will be compared with ALS estimators (alternating least squares) aimed at estimating changepoints in each trajectory separately.

13 - Bayesian nonparametric estimation in longitudinal mediation: A Baron-Kenny based estimator for cross-lagged models

Andrej Srakar, Tjaša Bartolj

Institute for Economic Research (IER) and School of Economic and Business, University of Ljubljana, Slovenia

Keywords: *longitudinal mediation analysis, cross-lagged panel models, Bayesian nonparametrics, centered Dirichlet process mixtures*

Mediation analysis has its roots in linear and nonlinear structural equation modelling. Limitations of cross-sectional models to analyze mediation can be overcome by longitudinal modelling which is particular reason for studying longitudinal mediation (LMA), being an uncommon and underresearched methodology. Existing models for LMA are estimated under strong parametric assumptions. We derive both nonparametric and Bayesian nonparametric (BNP) estimators for cross-lagged LMA models (based on standard Baron-Kenny approach to mediation). As LMA for cross-lagged models (CLM) demands a dynamic panel modelling, we extend Su and Lu (2013) iterative local kernel-based approach with sieves as initial estimators to derive a novel nonparametric estimator for CLM LMA. To derive a Bayesian nonparametric estimator we use dynamic Bayesian approach of Kim et al. (2019) extending their usage of Dirichlet process mixtures to centrered DPM (Yang, Dunson and Baird, 2010). We show the constructed BNP estimator attains optimal information rate (Alaa and van der Schaar, 2018). We explore the properties of the estimators in a Monte Carlo simulation study, comparing performance to parametric estimators for cross-sectional and longitudinal mediation. We discuss possibilities of extension to the framework of multiple mediators as well as multiple outcomes. In an application, we estimate the effects of long term care provision on reduction of costs in health care using dataset of Survey of Health, Ageing and Retirement in Europe (SHARE).

14 - Why Interaction Matters

Rolf Steyer

Institute of Psychology, Department of Methodology and Evaluation Research, Friedrich-Schiller-University Jena

Keywords: *Analysis of pre-post designs, ANCOVA, conditional effects, average effects, multi-group SEM analysis, analysis of change scores*

I present a simulated data example in which there are three relevant variables, a continuous pretest Z, a dichotomous treatment (or intervention) variable X, and a continuous outcome variable Y. The outcome variable Y depends on the pretest Z and on the treatment variable X, but the treatment effect of X on Y linearly depends on the pretest Z, that is, there is an interaction between X and Z. Most often data from such a pretest-posttest intervention, often called a nonequivalent control group design, is analyzed by an analysis of covariance (ANCOVA), that is, a linear regression of Y on X and Z, ignoring the interaction term X Z. In the example presented, the 'treatment effect' estimated via ANCOVA is negative although the true average treatment effect is positive. This reversal of effects also occurs in this example if we analyze the change scores. In contrast, if the interaction term is included in the list of regressors – which is the standard procedure, for example in EffectLiteR – then conditional and average treatment effects are estimated correctly. Hence, this example shows that considering the interaction is not only important in order to learn about differential (conditional) treatment effects but also about the average treatment effect.

15 - When data are not perfect: Robustness of LMS compared to regression methods when categorical data are treated as continuous

Julien P. Irmer, Andreas G. Klein, Jana C. Gäde, Karin Schermelleh-Engel

Goethe University, Institute of Psychology, Frankfurt, Germany

Keywords: *Nonlinear SEM, categorical variables, LMS, factor score regression, scale regression, skewed data, model misspecification*

One of the most often used methods to analyze nonlinear structural equation models (SEM) is latent moderated structural equations (LMS; Klein & Moosbrugger, 2000). This method assumes that variables are continuous, that predictor and error variables are normally distributed, and that the model at hand generated the data. Although in empirical research indicator variables are generally item responses with ordered categorical data that are often also asymmetric, the variables are often treated as if they were continuous. For linear models, Rhemtulla, Brosseau-Liard, and Savalei (2012) already suggested that at least five categories are needed in order to treat indicators as continuous. For nonlinear models, this suggestion might not be valid as higher moment information is necessary.

Using a Monte Carlo study, we investigated the performance of LMS for the analysis of nonlinear effects compared to factor score regression and multiple regression using categorical data that were treated as if they were continuous. Data were generated with varying numbers of categories, patterns of category thresholds, skewness and kurtosis, and model misspecification. For different population and analysis models, bias of parameter estimates, power rates to detect existing nonlinear effects and Type I error rates to detect spurious nonlinear effects were assessed.

16 - A Latent Auto-regressive Approach for Bayesian Structural Equation Modeling of Spatially or Socially Dependent Data

Roman Zachary, Holger Brandt

Department of Psychology, University of Zürich

Keywords: *Measurement invariance, Differential item functioning, Shrinkage priors*

Spatial analytic approaches are classic models in econometric literature (LeSage & Pace, 2009), but relatively new in social sciences. Spatial analysis models are synonymous with social network autoregressive models which are also gaining popularity in the behavioral sciences. These models have two major benefits. First, dependent data, either socially or spatially, must be accounted for to acquire unbiased results. Second, analysis of the dependence provides rich additional information such as spillover effects (Valente et al.,1997; LeSage & Pace, 2009). Structural Equation Models (SEM) are widely used in psychological research for measuring and testing multi-faceted constructs (Bollen, 1989). While SEM models are widely used, limitations remain, in particular latent interaction/polynomial effects are troublesome (Brandt et al., 2014). Recent work has produced methods to account for these issues (Brandt et al., 2018) as well as integrating spatial and network effects in SEM (Oud & Folmer, 2008) to a limited extent. However, a cohesive framework which can simultaneously estimate latent interaction/polynomial effects and account for spatial effects with both exogenous and endogenous latent variables has not been established. To accommodate this, I provide a novel model, the Bayesian Auto-Regressive Dependence Structural Equation Model (BARDSEM). First I briefly outline classic spatial auto-regressive models. Next I present the BARDSEM and introduce simulation results to exemplify its performance. Finally, I provide an empirical example using the spatially dependent extended US southern homicide data (Messner et al., 1999; Land et al., 1990) to show the rich interpretations made possible by the BARDSEM. Finally, I discuss results, implications, limitations, recommendations, and future research.

17 - Pairwise maximum likelihood for multilevel data

Mariska T. Barendse [1], Yves Rosseel [2]

[1] Department of Psychology, Education & Child Studies, Erasmus University Rotterdam, Netherlands

[2] Department of Data Analysis, Ghent University, Belgium

Keywords: *structural equation modeling, multilevel, discrete data, random slopes*

The pairwise maximum likelihood (PML) estimation method seems very promising to estimate complex multilevel structural equation models (SEM) with discrete data. With multilevel models we take into account that observations within a cluster tend to be more alike than observations from different clusters. The pairwise likelihood with multilevel models is obtained as the product of bivariate likelihoods for within-cluster pairs of observations (see Renard et al., 2004; Bellio and Varin, 2005; Cho and RabeHesketh, 2011). As casewise bivariate likelihoods can be calculated, the PML estimation method is able to estimate models with random slopes. In this presentation, we will discuss the possibilities of the PML estimation method with complex multilevel SEM models. We also investigate the PML estimation method in a simulation study, where we vary the type of response scale (binary, four response options) and the number of random slopes (one, two).