6th GESIS Summer School in Survey Methodology
Cologne, August 2017

Syllabus for Course 6: “Structural Equation Modeling with Panel Data”

Instructors: Dr. Daniel Seddig Georg Kessler
Email: dseddig@uni-koeln.de georg.kessler@univie.ac.at

Date: August 14-18 August, 2017
Time: 11:00-13:00, 14:00-18:00
Course starts Monday morning at 11:00

About the Instructors:
Daniel Seddig is a Research Assistant at the Institute of Sociology and Social Psychology, University of Cologne, statistical consultant at the Institute of Psychology, University of Zurich, and chair of the “European Working Group on Quantitative Methods in Criminology (EQMC)” of the “European Society of Criminology (ESC)”. His research interests in the area of structural equation modeling include measurement invariance, count data analysis, panel data analysis, and Bayesian SEM. Applications include values, attitudes, hedonism, school adaptation problems, school bonding and adolescent deviant behavior. Recent publications appeared in Sociological Methods and Research, Crime and Delinquency, European Journal of Criminology, and Advances in Statistical Analysis.

Georg Kessler teaches methodology of the social sciences at the University of Würzburg. He graduated in Sociology from the University of Vienna, where he specialized in methodology, and wrote his master thesis on a mixed methods design, applying SEM and Cognitive Interviews on the Schwartz Value Scale used in the ESS. He received his SEM-training from Prof. Reinecke and Prof. Schmidt from the University of Gießen. He worked as a lecturer at the University of Vienna in the department of Sociology and as a freelancer at a consulting and human development firm in Vienna.

Selected Publications:

Short Course Description:
The course will expand on the course “Introduction to Structural Equation Modeling: Confirmatory Factor Analysis with Mplus” and show how to apply the SEM approach to longitudinal data using the Mplus computer program. In the first part of the course, we will introduce the autoregressive model (ARM) and the cross-lagged panel model (CLPM) to study the stability of a single variable over time as well as reciprocal effects and ‘causal’ predominance between two variables over time. Each type of model will be discussed as a single-indicator as well as a multiple-indicator model. All models will be applied to data from a longitudinal study on authoritarianism and anomia in Germany. In the second part of the course we will focus on modeling development and change over time with the latent growth model (LGM) applying the same dataset. We will begin with the univariate case to analyze growth of a single variable and extend the basic model to the multivariate case to analyze parallel
processes of two variables over time. Finally, two ‘hybrid-models’ will be discussed: the autoregressive latent trajectory model (ALT) and the more recently developed random-intercept cross-lagged panel model (RI-CLPM). Topics in both parts include parameterization of autocorrelations, Socratic effects, latent means, and MIMIC models. Furthermore, we will extend our discussion to multiple group comparisons as well as the issue of (longitudinal) measurement invariance as a prerequisite for comparisons across groups and over time.

Keywords:
structural equation modeling, longitudinal analysis, panel data, stability and reciprocity, change

Course Prerequisites:
Basic knowledge of and basic experience with confirmatory factor analysis and structural equation modeling. This could be acquired in the course “Introduction to Structural Equation Modeling: Confirmatory Factor Analysis with Mplus” in week 1.

We will use the software package Mplus. We will briefly introduce Mplus during the first exercise. Mplus will also be used in the course “Introduction to Structural Equation Modeling” in week 1. The short course “Introduction to Data Analysis Using Mplus” in week 0 will explicitly focus on how to use Mplus effectively.

Target Group:
Participants will find the course useful if:
- they are interested in assessing and explaining change over time
- they are interested in the relationship between processes of change in different variables
- they are interested in assessing the degree of reciprocity in the relationship between variables over time
- they are interested in causal predominance of one variable over another
- they are interested in conducting meaningful comparisons of (latent) variables over time (or across groups)

Course and Learning Objectives:
By the end of the course participants will:
- know how to specify autoregressive and cross-lagged structural equation models to test for stability and reciprocity in manifest and latent variables over time as predicted by a social scientific theory
- know how to test for measurement invariance of latent variables across time (and groups)
- know how to examine various forms of change with different specifications of the latent growth model
- be able to specify all models with the software package Mplus;
- (if time allows), learn to separate interpersonal from intrapersonal variance in the variables of interest and run a random-intercept cross-lagged panel model test causal predominance as predicted by a social scientific theory

Organizational Structure of the Course:
First four days:
- about 180 minutes – lecture
- about 180 minutes – exercises on data prepared by the instructors

Note: In some of the days we may divide the theory and exercises time differently, depending on our progress.

Last day:
- expansion of topics and exercises due to participants’ preferences
- it will be possible to discuss participants’ projects in small groups focusing on conceptual and analytical problems; groups will rotate so that each group spends time on each project; possible solutions and difficulties will be discussed in class
- general summary and discussion; open questions

Free study time and what we expect from participants:
- Participants are encouraged to discuss the topics of the course every day with each other.
- Participants are expected to repeat the exercises conducted in class on their own.
- Participants should read the course literature.
- Participants are encouraged to work on their own projects and analyze their own data individually or in groups.

Software and Hardware Requirements:
SPSS/Stata and Mplus Version 7.4. Course participants will not need to bring a laptop computer for this course. This course will take place in a computer lab.

Long Course Description:
Social scientists are often interested in explaining stability and change in variables such as values, attitudes, beliefs, intentions or behaviors. Furthermore, researchers want to explore in a theory-driven way which variables are causally predominant over others and how processes of change are interrelated over time. Many of these issues can be addressed with survey-based non-experimental panel data and longitudinal structural equation modeling (SEM). Panel data contain measurements of the same dimensions for the same respondents over multiple time-points. The first part of our lectures (day 1-2) will begin with a discussion of the concept of stability, i.e. the assessment of the same variable at (at least) two subsequent time-points (e.g., t-1 and t). In the autoregressive model (ARM), stability refers to the degree of maintenance of the response pattern over time and stable differences in the response pattern reflect stable between-person characteristics. Further, the concept of correlated errors over time (autocorrelations) and the implementation in autoregressive models will be explained. We will then show how to interpret the estimates obtained from the ARM. Next, we will describe how to extend the ARM to address the issues of reciprocity and causality between two variables by means of the classic cross-lagged panel model (CLPM). Further, we will explain the basic ideas of testing (longitudinal) measurement invariance, which is a prerequisite to compare estimates obtained from variables measured at different time-points or in different contexts. We will also discuss the ARM and CLPM for latent variables with multiple indicators. At this point, we will also look at how to draw meaningful comparisons of estimates obtained from the ARM or CLPM across multiple groups (e.g., countries, cultures, regions) and at MIMIC models. The second part of our lectures (day 3-4) will focus on the latent growth model (LGM) to analyze development and change. We will begin with a discussion of the basic univariate LGM and show why conclusions about change cannot be easily drawn from the ARM. We will talk about the parameterization of change in terms of latent intercept and slope growth factors and show how these factors are different from the understanding of latent variables in CFA. Next, we will introduce different forms of change (linear, nonlinear, piecewise) as well as different strategies for the coding of time. Further, the univariate LGM will be extended from a single-indicator model to a multiple-indicator model. Again, correlated errors and measurement invariance will be of crucial importance.

We will also discuss the explanation of change by exogenous variables as well as the explanation of exogenous variables by the developmental process (MIMIC models). Further, we will extend the univariate LGM to a multivariate LGM and show how to test correlations between two growth processes, or in other words, if change in one variable is related to change in another variable (factor-of-curves LGM, parallel process LGM). Finally, we want to present two types of hybrid-models. The first is the autoregressive latent trajectory model (ALT), which combines features of the autoregressive cross-lagged and latent growth models to study reciprocity and change simultaneously. The second is the recently developed random-intercept cross-lagged panel model (RI-CLPM), which enables to control for unobserved stable variables and to separate stable between-person differences from within-person processes by. Due to this extension, any cross-lagged effect can be allocated at the within-person level and strengthen the causal implications of cross-lagged SEM.
We encourage participants to bring their own data and apply the methods we discuss and study on their data during the course. On the last day of the course, participants will have the opportunity to present their findings and discuss with the class problems they encountered.

We will focus our lectures on the analysis of data on a continuous scale. However, all models discussed in the course can be estimated with categorical as well.

**Day-to-day Schedule and Literature:**

<table>
<thead>
<tr>
<th>Day</th>
<th>Topic(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction; autoregressive (AR) and cross-lagged panel models (CLPM): single indicator models; autocorrelations; interpretation of parameters; examples</td>
</tr>
</tbody>
</table>

**Compulsory reading:**

**Suggested reading:**

<table>
<thead>
<tr>
<th>Day</th>
<th>Topic(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Autoregressive (AR) and cross-lagged panel models (CLPM): multiple-indicator models; measurement invariance over time; multiple-group analysis; MIMIC models; treatment of missing values; examples</td>
</tr>
</tbody>
</table>

**Compulsory reading:**

**Suggested reading:**
- Vandenberg, R. J. & Lance, C. E. (2000). A review and synthesis of the measurement invari-
3

- **Latent growth model (LGM):** single-indicator LGM; MIMIC LGM; types of change; coding of time; interpretation of parameters

**Compulsory reading:**

**Suggested reading:**

4

- **Multivariate representations of growth; hybrid models: autoregressive latent trajectory model (ALT) and random-intercept cross-lagged panel model (RI-CLPM); examples**

**Compulsory reading:**

**Suggested reading:**

5

- **Expansion of topics and exercises; discussion of participants’ work; general summary; open questions; filling in the online evaluation questionnaire followed by a debriefing/feedback round**

**Preparatory Reading:**
Additional Recommended Literature: