10th GESIS Summer School in Survey Methodology

[2nd Virtual GESIS Summer School]

28 July – 20 August 2021

Syllabus for Course 10:
Collecting and Analyzing Longitudinal Social Network Data

Instructors:	Dr. Lars Leszczensky	Dr. Sebastian Pink
E-mail:	lars.leszczensky@mzes.uni-mannheim.de	sebastian.pink@uni-mannheim.de
Homepage:	https://larsleszczensky.com	https://www.sowi.uni-mannheim.de

Date: 16-20 August 2021
Time: 09:00-12:30 + 13:30-17:00
Time zone: CEST/CEDT, course starts Monday at 09:00 am
Venue: Online via Zoom

About the Instructors:

Lars Leszczensky is a postdoctoral research fellow at the Mannheim Centre for European Social Research at the University of Mannheim. He is experienced in planning and conducting the collection of longitudinal social network data as well as analyzing these data. As a teacher, he regularly gives theoretical, methodological, and thematic seminars. Beyond social network analysis, he is interested in methods of social research and topics such as integration, intergroup relations, and social identity.

Sebastian Pink is a postdoctoral researcher at the Chair of General Sociology at the University of Mannheim. He is experienced in planning and conducting the collection of longitudinal social network data as well as analyzing these data. As a teacher, he regularly gives theoretical, methodological, and thematic seminars. Beyond social network analysis, he is interested in methods of social research and topics such as integration, fertility decision-making and social inequality.

Selected Publications:

- Leszczensky, L., & Pink, S. (2019): What Drives Ethnic Homophily? A Relational Approach on How Ethnic Identification Moderates Preferences for Same-Ethnic Friends. American Sociological Review, 84, 394-419.
- Leszczensky, L., & Pink, S. (2015): Ethnic Segregation of Friendship Networks in School: Testing a Rational-Choice Argument of Differences in Ethnic Homophily between Classroom- and Grade-Level Networks. Social Networks, 42, 18-26.
- Pink, S., Kretschmer, D., & Leszczensky, L. (2020): Choice Modelling in Social Networks Using Stochastic Actor-Oriented Models. Journal of Choice Modelling, 34, 100202.
- Leszczensky, L., & Pink, S. (2020): Are Birds of a Feather Praying Together? Assessing Friends' Influence on Muslim Youths' Religiosity in Germany. Social Psychology Quarterly, 83, 251-271.
- Jugert, P., Leszczensky, L. & Pink, S. (2020): Differential Influence of Same- and Cross-Ethnic Friends on Ethnic-Racial Identity Development in Early Adolescence. Child Development, 91, 949-963.

Short Course Description:

Social scientists often are interested in understanding how social networks emerge and/or how they shape individual behavior. These questions of network formation ("selection") and network effects ("influence") concern both human individuals and organizational units. Examples for selection are the emergence of friendship between people or cooperation between firms; examples for influence are adolescents start smoking because of their friends or firms copying other firms' strategies. Selection and influence are inherently dynamic processes, but few social scientists have been trained in collecting, processing, and analyzing longitudinal social network data.
This practical course guides participants who intend to collect and/or analyze longitudinal social network data. We start by conceptualizing and planning data collection, discussing both general challenges and, if applicable, participants' own projects. Thereafter, participants learn how to handle and manage network data in R by guided examples and exercises. The main part of the course focuses on specifying, estimating and interpreting stochastic actor-oriented models (SAOM) for network dynamics, again with a mix of guided examples and practical exercises using the R package RSiena. We consider selection and influence as well as how SAOM can help to empirically disentangle these competing processes.

Keywords:

social networks, data collection, data analysis, computational social science, R

Course Prerequisites:

- Basic knowledge in quantitative data analysis
- Prior knowledge of social network analysis and/or R is helpful but not necessarily required

Target Group:

Participants will find the course useful if:

- they (intend or consider to) collect longitudinal social network data
- they (intend or consider to) analyze longitudinal social network data to help them answering substantive research questions
- they already are analyzing social network data and want to discuss their work

Course and Learning Objectives:

By the end of the course participants will:

- know how to design and conduct a longitudinal social network study
- be able to manage and handle longitudinal network data
- know how to exploit the potential of stochastic actor-oriented models for their research aims
- understand how to specify and estimate stochastic actor-oriented models for network dynamics in R
- have learned how to interpret and communicate results of stochastic actor-oriented models

Organizational Structure of the Course:

This is a five-day course with a total amount of 30 hours of virtual class time. Participants can expect a mix of interactive lectures, hands-on exercises, and opportunities for group discussions and individual consultation. Guided exercises in R deepen the understanding of the course material. The lecturers will be available for individual consultations on participants' planned or current projects.

Software and Hardware Requirements:

The practical examples and exercises will be done in R, so participants should have a recent version installed on their local computer. For working with R in general, we recommend using RStudio.

Prior to the course, participants should install the following R-packages from CRAN, with dependencies: igraph, ggraph, ggplot2, tidygraph, tidyverse, reshape2, and gridExtra
Participants also should install the newest version of the R-package "RSiena" from R-Forge, with dependencies. (The version on CRAN tends to be outdated.) The command is: install.packages("RSiena", repos="http://R-Forge.R-project.org")

Long Course Description:

Like many other social scientific terms, the term "social networks" is used in different ways. In this course, we use "social network" in reference to a finite set of actors and one or more relationships between them. Examples are friendship relations among classmates, working relations among firms in a city, or diplomatic relations among countries. Since actors within social networks are by definition related to each other, conventional methods of statistical analysis that assume independence cannot be applied. This necessitates specific methods of social network analysis that, however, are not included in the standard curricula of social science education.

Researchers can use social network analysis to answer two broad kinds of questions. First, social networks are dynamic by nature, with classmates, firms, or countries both creating new friendship, working, or diplomatic relations and terminating existing ones. Understanding how and why actors *select* their network partners and, therefore, why the structure of social network changes over time, is the first kind of question that social network analysis can help to answer. Second, the behaviour of actors such as classmates, firms, or countries often is affected by their network partners; for example, classmates might start to smoke because their friends do or firms might copy procedures of successful partners. Understanding how and why social networks *influence* the behaviour of individual actors is the second kind of question that social network analysis can help to answer.

These two types of research questions are labelled as "selection" and "influence". Selection and influence are inherently dynamic processes, but few social scientists are trained in collecting, processing, and analyzing longitudinal social network data. Separating selection and influence processes further is challenging because they often lead to the same outcome. For example, if we observe a friendship group of smokers, did they influence each other's smoking behaviour or did they become friends because they already were smokers? And do two successful firms cooperate with each other because they are both successful, or did they become successful because they cooperated? Answering these types of questions by disentangling competing mechanisms requires *longitudinal* social network data and methods of longitudinal social network analysis.

This course guides participants who intend to collect and/or analyze longitudinal social network data. For this purpose, we rely on a mix of interactive lectures, guided examples, and practical exercises. We use R for all guided examples and exercises, so participants should install a recent version on their local computer beforehand; they also should install the R-packages listed in the software requirements. We will use and provide exemplary school-based friendship network data that we collected for a project called "Friendship and Identity in School".

On the first day, we start by introducing basic concepts, typical research questions, and longitudinal social network data. We will provide numerous examples, but also enable participants to bring forward their own research aims. In the morning of the second day, we discuss the design of longitudinal social network studies and the collection of longitudinal social network data, discussing both general challenges and, if applicable, participants' own data collection projects. In the afternoon, participants learn how to handle and manage network data in R by guided examples and exercises. This includes the visualization of networks and the calculation of descriptive network measures to get a grip on how networks change over time.

On the third and fourth day, we will address how to analyze selection (Day 3) and influence (Day 4) with stochastic actor-oriented models (SAOM). In the morning of both days, we will introduce the respective and show and practice how to specify and estimate it using R. In the afternoon of both days, we will practice how to interpret the results of selection models and how to graphically communicate findings. On both days, we thus proceed in a similar fashion, with the overarching aim of enabling participants to address their own research question by showing them how to set up and estimate models and interpret and convey their results.

On the final day, we will address various more advanced topics with which participants likely will be confronted when working with SAOM for their own projects, such as criteria for convergence in parameter estimates, goodness of fit, and different means of analyzing multiple networks. We close by giving participants group-based and individual feedback on their own projects.

Day-to-day Schedule and Literature:

Day	Topics
1	**Introduction to Longitudinal Network Analysis** ▪ Introduction, Objectives, Schedule, and Organization ▪ Social networks, research questions and potential of (longitudinal) social network analysis ▪ (Longitudinal) Social network data: specifics and challenges ▪ Discussion of participants' goals and plans for their projects
	Compulsory reading: ▪ Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network Analysis in the Social Sciences. Science, 323, 892-895. ▪ Wasserman, S. and Faust, K. (1994). Social Network Analysis: Methods and Applications. Chapters 1-3. Cambridge: Cambridge University Press
2	**Data Collection and Descriptive Analysis** ▪ Data requirements for longitudinal social network analysis and how to design a longitudinal network study ▪ How to collect longitudinal network data ▪ Data processing and network visualization ▪ Calculation of statistical network measures
	Compulsory reading: ▪ Jackson, M. O. (2008). Social and Economic Networks. Chapter 2. Princeton: Princeton University Press. ▪ Leszczensky, L., Beier, H., Kruse, H., & Pink, S. (2016). Collecting Network Panel Data in Schools: Practical Guidance Based on the Experiences of Three German Research Projects. Mannheim Centre for European Social Research, Working Paper 166.
3	**Stochastic Actor-Oriented Models for Network Dynamics I: Selection** ▪ Logic of network tie selection in stochastic actor-oriented models ▪ Specification and estimation of selection part of the model ▪ Interpretation of results (based on linear combinations) ▪ (Graphical) Communication of findings
	Compulsory reading: ▪ Snijders, T. A., Van de Bunt, G. G., & Steglich, C. E. (2010). Introduction to Stochastic Actor-Based Models for Network Dynamics. Social Networks, 32, 44-60. ▪ Pink, S., Kretschmer, D., & Leszczensky, L. (2020): Choice Modelling in Social Networks Using Stochastic Actor-Oriented Models. Journal of Choice Modelling, 34. ▪ Ripley, R., Snijders, T. A. B., Boda, Z., Vörös, A., & Preciado, P. (2021). Manual for RSiena. University of Oxford and University of Groningen. Chapter 13.3.
4	**Stochastic Actor-Oriented Models for Network Dynamics II: Influence** • Logic of influence from network ties in stochastic actor-oriented models • Specification and estimation of influence part of the model • Interpretation of results (based on linear combinations) • (Graphical) Communication of findings
	Compulsory reading: ▪ Steglich, C., Snijders, T. A., & Pearson, M. (2010). Dynamic networks and behavior: Separating selection from influence. Sociological Methodology, 40, 329-393. ▪ Ripley, R., Snijders, T. A. B., Boda, Z., Vörös, A., & Preciado, P. (2021). Manual for RSiena. University of Oxford and University of Groningen. Chapter 13.4.
5	**Stochastic Actor-Oriented Models for Network Dynamics III and Project Outlook** ▪ Convergence and goodness of fit ▪ Combining multiple networks: Meta-analysis, multi-group option, and SienaBayes ▪ Missing Data ▪ Group-based and individual feedback on participants' projects
	Compulsory reading: ▪ Lospinoso, J., & Snijders, T. A. B. (2019). Goodness of fit for stochastic actor-oriented models. Methodological Innovations, 12, 1-18. (Example section) ▪ Ripley, R., Snijders, T. A. B., Boda, Z., Vörös, A., & Preciado, P. (2021). Manual for RSiena. University of Oxford and University of Groningen. Chapter 11.

Preparatory Reading:

None.

Additional Recommended Literature:

In this course, we treat the basic version of stochastic actor-oriented models. For further information, including on various extensions and related topics, we refer participants to the Siena webpage, which is maintained by Tom Snijders and provides tons of useful materials. This includes RSiena scripts on numerous topics including but vastly extending the ones discussed in this course. The website also lists literature and teaching materials on various topics, including introductions to SIENA in various languages.
The url is: https://www.stats.ox.ac.uk/~snijders/siena/siena.html
If participants are interested in literature on specific topics, we are also happy to provide more information.