gesis Leibniz Institute for the Social Sciences

# GESIS Summer School in Survey Methodology 2022

## Syllabus for course:
## "Using Directed Acyclic Graphs for Causal & Statistical Inference"


| | |
|---|---|
| Lecturer: | Julian Schuessler |
| E-mail: | julians@ps.au.dk |
| Homepage: | www.julianschuessler.net |


Date: 03-05 August 2022
Time: 09:00-11:00 | 13:00-15:00
Venue: Online via Zoom

## About the Lecturer:
*Julian Schuessler* is a post-doc at the Centre for the Experimental-Philosophical Study of Discrimination, Aarhus University, Denmark. He defended his PhD on the use of causal graphs for causal and statistical inference problems in the social sciences at the University of Konstanz in 2020. In 2019, he received the American Statistical Association's "Causality in Statistics Education Award". Beyond causal inference and statistical methods, his research interests are in the empirical study of discrimination, political behaviour, and public opinion, using observational and experimental data.

## Selected Publications:
- Schuessler, Julian, and Peter Selb. 2019. "Graphical Causal Models for Survey Inference." SocArXiv.doi:10.31235/osf.io/hbg3m.

## Course Description:
This course uses causal graphs (or "directed acyclic graphs", DAGs) as a remarkably simple, yet general and powerful framework to describe and discuss a large set of problems that empirical social scientists need to tackle. Is my question of interest descriptive or causal? How can I communicate my assumptions effectively to others, and can I test them? How can I tell correlation from causation? How do I choose control variables for my regression models? After discussing how DAGs can be used to answer these foundational questions, the course also covers basics of causal interaction and effect heterogeneity, causal mediation, nonresponse/selection bias (and adjustments for it) and, if time permits, instrumental variables and panel data analysis from a DAG perspective.

## Keywords:
causal graphs; causal inference; mediation; instrumental variables; nonresponse

## Course Prerequisites:
Participants should be willing to learn and use formal reasoning and must have at least Bachelor-level knowledge of statistics. Basic knowledge of R is helpful.

## Target Group:
Participants will find the course useful if:
- They are interested in causal questions and want to understand the assumptions associated with regression control, mediation analysis and instrumental variables better

- They are interested in non-causal question, want to use data suffering from nonresponse, and want to understand how to use causal assumptions in this case

## Course and Learning Objectives:
By the end of the course participants will:
- Know how to use causal graphs to visualize causal assumptions, define quantities of interest, and to determine testability of assumptions via d-separation
- Know how to graphically determine identification of causal and descriptive quantities like average causal effects, causal interaction, effect heterogeneity, natural direct and indirect effects, and population distributions from data with nonresponse
- Know under what graphical assumptions instrumental variable and panel data analysis typically operate
- Will have some basic knowledge about how all of this relates to implementation in standard statistical software

## Organizational Structure of the Course:
This short course throughout will change between short lecture-style inputs and individual or small-group hands-on exercises supervised by the lecturer and a teaching assistant (4hrs/day). Participants are encouraged to bring their own research ideas to develop them further using the material from the class. The lecturer will also be available for individual consultation in the afternoons (1hr/day).

## Software and Hardware Requirements:
We will use briefly discuss some elements of the R packages “dagitty”, “sensemakr”, “mediation”, “AER”, “estimatr”. Most of the course will not depend on using R. For those who have never used R, installation instructions and a short introductory video are linked to below under “Preparatory Reading”.

## Day-to-day Schedule and Literature:

<table>
<tr><td>Day</td><td>Topic(s)</td></tr>
<tr><td rowspan="2">1</td><td>Introductions, interests and preferences of participants, self-assessments<br>Statistical control and Simpson’s paradox<br>Descriptive and causal questions<br>Graph basics and d-separation<br>Interventions and causality<br>Back-door criterion and unobserved confounding<br>Post-treatment bias</td></tr>
<tr><td><u>Compulsory reading (have to be read before class):</u><br>▪ Felix Elwert. “Graphical causal models". In: Handbook of causal analysis for social research. Springer, 2013, pp. 245—273.<br>Only pp. 245—252 needed!<br><br>Play around with DAGitty: http://dagitty.net/dags.html<br><br><u>Suggested reading (suggested, yet do not have to be read before class):</u><br>▪ Luke Keele, Randolph T Stevenson, and Felix Elwert. “The causal interpretation of estimated associations in regression models". In: Political Science Research and Methods 8.1 (2020), pp. 1—13.</td></tr>
<tr><td>2</td><td>Sensitivity analysis for unobserved confounding<br>Causal interaction and effect heterogeneity<br>Causal mediation: Controlled and natural effects, post-treatment confounding, sensitivity analyses</td></tr>
</table>

| | |
|---|---|
| | Compulsory reading:<br>None.<br><br>Suggested reading:<br>▪ Cinelli, Carlos, and Chad Hazlett. "Making sense of sensitivity: Extending omitted variable bias." Journal of the Royal Statistical Society: Series B (Statistical Methodology) 82.1 (2020): 39-67.<br>▪ Luke Keele and Randolph T Stevenson. "Causal interaction and effect modification: same model, different concepts". In: Political Science Research and Methods 9.3 (2021), pp. 641—649.<br>▪ Imai, Kosuke, and Teppei Yamamoto. "Identification and sensitivity analysis for multiple causal mechanisms: Revisiting evidence from framing experiments." Political Analysis (2013): 141-171. |
| 3 | Non-reponse/missing data and its consequences<br>Adjustment for non-response from a graphical perspective<br>Instrumental variables in linear & nonparametric models<br>Choosing control variables in IV models<br>Panel data and fixed effects<br>Wrap-up |
| | Compulsory reading:<br>▪ Schuessler, Julian, and Peter Selb. "Graphical causal models for survey inference." Working Paper (2021).<br>Suggested reading:<br>▪ Knox, Dean, Will Lowe, and Jonathan Mummolo. "Administrative records mask racially biased policing." American Political Science Review 114.3 (2020): 619-637.<br>▪ Felix Elwert and Christopher Winship. "Endogenous selection bias: The problem of conditioning on a collider variable". In: Annual review of sociology 40 (2014), pp. 31--53.<br>▪ Imai, Kosuke, and In Song Kim. "When should we use unit fixed effects regression models for causal inference with longitudinal data?." American Journal of Political Science 63.2 (2019): 467-490.<br>▪ Schuessler, Julian, Glynn, Adam N., and Rueda, Miguel. R. "Post-Instrument Bias". Working paper. (2021) |

## Preparatory Reading:

- Downloading R and RStudio: https://rstudio-education.github.io/hopr/starting.html
- For those who have never used R before (18 mins): https://www.youtube.com/watch?v=DuQSQQa6Ssw

## Additional Recommended Literature:

### Textbooks:

*(Aronow/Miller and Imai do not discuss graphs)*

- Aronow, Peter M., and Benjamin T. Miller. Foundations of agnostic statistics. Cambridge University Press, 2019.
- Hernán MA, Robins JM. Causal Inference: What If. Boca Raton: Chapman & Hall/CRC, 2020. https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/
- Imai, Kosuke. Quantitative social science: An introduction. Princeton University Press, 2018.
- Pearl, Judea, Madelyn Glymour, and Nicholas P. Jewell. Causal inference in statistics: A primer. John Wiley & Sons, 2016.
- Peters, Jonas, Dominik Janzing, and Bernhard Schölkopf. Elements of causal inference: foundations and learning algorithms. The MIT Press, 2017.

- Shalizi, Cosma. "Advanced data analysis from an elementary point of view." (2021) http://www.stat.cmu.edu/~cshalizi/ADAfaEPoV/

## Other books and papers:

- Chen, Bryant, and Judea Pearl. "Regression and causation: a critical examination of six econometrics textbooks." Real-World Economics Review, Issue 65 (2013): 2-20.
- Knox, Dean, Will Lowe, and Jonathan Mummolo. "Administrative records mask racially biased policing." American Political Science Review 114.3 (2020): 619-637.
- Ogburn, Elizabeth L., and Tyler J. VanderWeele. "Causal diagrams for interference." Statistical science 29.4 (2014): 559-578.
- Pearl, Judea. Causality. Cambridge university press, 2009.
- Pearl, Judea. "Linear models: A useful “microscope” for causal analysis." Journal of Causal Inference 1.1 (2013): 155-170.
- Pearl, Judea. "Causes of effects and effects of causes." Sociological Methods & Research 44.1 (2015): 149-164.
- Pearl, Judea. "Conditioning on post-treatment variables." Journal of Causal Inference 3.1 (2015): 131-137.
- Pearl, Judea. "Generalizing experimental findings." Journal of causal inference 3.2 (2015): 259-266.
- Pearl, Judea. "Indirect Confounding and Causal Calculus (On three papers by Cox and Wermuth)." (2015).
- Pearl, Judea. "Theoretical impediments to machine learning with seven sparks from the causal revolution." arXiv preprint arXiv:1801.04016 (2018).