



Advanced Topics in Empirical Social Research (ATESR)

Seminar, 2nd term 2019-2020

Organised by Fabrizio Bernardi and Elias Dinas

Please register [online](#)

Contact: Jennifer Dari: jennifer.dari@eui.eu

Course Outline

The course touches on a series of cutting-edge and cross-cutting topics in empirical social research. The course runs in parallel to the course Causal Inference Analysis and, in certain occasions, complements it. The first part of course covers issues related to directed acyclic graphs and problems related to colliders and causal mediation analysis. The central sessions cover issues on colliders, post-treatment bias, mechanisms, mediation analysis, G-methods, parametric and non-parametric inference and bounding analysis. The last two ad hoc sessions are on the misuse of statistical significance and implication analyses.

Classes and topics:

- o Week 1: Introduction to cause seeking designs & practical arrangements [14 January]
- o Week 2: Potential Outcomes framework and Directed Acyclic Graphs [21 January]
- o Week 3: Endogenous selection bias or the problem of colliders [28 January]
- o Week 4: G-Computation: An introduction – Guest speaker: Juho Harkonen [4 February]
- o Week 5: Causal Mediation Analysis [11 February]
- o Week 6: Bounds [18 February]
- o Week 7: Inference I: parametric and non-parametric inference [3 March]
- o Week 8: Inference II: clustering (when and how to cluster) [6 March]
- o Week 9: Debates on statistical significance [10 March]
- o Week 10: Implication analysis [17 March]

Participation

Given the heterodox nature of the seminars, researchers are also allowed and encouraged to attend ad hoc-sessions of the course. In that case we expect that **they make the required assignments.**

Requirements

In order to obtain the 20 credits participants, in addition to attending regularly, are required to make the assignments (about 1 page) for all weeks, following the indications on the syllabus. Your assignments should be posted on the course folder by the Sunday preceding the class. A selection of these response papers will be discussed in class.

Schedule

The seminar takes place on **Tuesdays at 11:00-13:00**, in Seminar Room 2 in the Badia, and will run **from Tuesday 14 January to Tuesday 17 March**. Please note there will be **no lecture on the week between 24 and 28 February**. Use this week as your reading week.

The full list of the seminar sessions is the following:

Tuesday	14 January	11:00-13:00	Seminar Room 2
Tuesday	21 January	11:00-13:00	Seminar Room 2
Tuesday	28 January	11:00-13:00	Seminar Room 2
Tuesday	4 February	11:00-13:00	Seminar Room 2
Tuesday	11 February	11:00-13:00	Seminar Room 2
Tuesday	18 February	11:00-13:00	Seminar Room 2
Tuesday	3 March	11:00-13:00	Seminar Room 2
Friday	6 March	15:00-17:00	Seminar Room 2
Tuesday	10 March	11:00-13:00	Seminar Room 2
Tuesday	17 March	11:00-13:00	Seminar Room 2

SYLLABUS (please note that the reading list may still be subject to change)

Week 1. Introduction to cause seeking designs & practical arrangements (Tuesday, January 14, 2020)

Compulsory readings

Joshua D. Angrist & Jörn-Steffen Pischke 2013. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Chapter 1, which can be found here: <http://press.princeton.edu/chapters/s8769.pdf>

Freedman, D.A. 1991. Statistical models and Shoe Leather. *Sociological Methodology* 21: 291-313.

Elif Batuman. 2013. "Poisoned Land. On the trail of a mystery disease in the Balkans". *New Yorker*, August 12, 2013

Other readings

James Lind (1716-1794), a pioneer: http://en.wikipedia.org/wiki/James_Lind

Freedman, D.A. 2010. "On types of scientific enquiry: Nine success stories in medical research." in *The Oxford Handbook of Political Methodology* pp. 300–18. Janet M. Box-Steffensmeier, Henry E. Brady and David Collier, editors.

Luke Keele, The Statistics of Causal Inference: A View from Political Methodology. *Political Analysis* (2015) pp. 1–23 [esp introductory section and section 3]

Assignment

We expect you to do the compulsory readings in the syllabus for class 1 during the Christmas break, and to post on the course foulder by the previous Sunday night a page or so, *the research question of your dissertation* (or one sub-question derived from it, or one new question you are interested in) and then following Angrist & Pischke fundamental prescription that you find in their chapter 1: imagining you are in a world with no constraints – ethical, financial, temporal etc, -- *can you conceive an ideal experiment that could answer that question?*"

Week 2. Directed acyclic graphs: an introduction (Tuesday, January 21, 2020)

Compulsory reading

Morgan, Stephen L. and C. Winship. 2007. Counterfactuals and Causal Inference. Methods and principles for Social Research. Cambridge: Cambridge University Press. *Chapter 1 and 2*

Rohrer, Julia M. 2018. "Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data." *Advances in Methods and Practices in Psychological Science* 1(1):27-42. doi: 10.1177/2515245917745629.

Other readings (Optional)

Antonakis, J, Bendahan, S., Jacquart, P. and Lalive, R. . 2010. On making causal claims: A review and recommendations. *The Leadership Quarterly* 21:1086-1120.

Gangl, M. 2010. Causal Inference in Sociological Research. *Annual Review of Sociology* 36: 21-47.

Luke Keele, The Statistics of Causal Inference: A View from Political Methodology. *Political Analysis* (2015) pp. 1–23 [esp introductory section and section 3]

Ni Bhrolchain, M., and Dyson, T. 2007. On Causation in Demography: Issues and Illustrations. *Population and Development Review* 33 (1):1–36.

Assignment

After reading the compulsory readings, present the research question that you have discussed in the first assignment of last week using a DAG, including in the DAG possible confounders that challenge your causal inference.

Week 3. Endogenous selection bias or the problem of colliders (Tuesday, January 28, 2020)

Elwert, Felix and Christopher Winship. 2014. "Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable." *Annual Review of Sociology* 40(1):31-53.

Morgan, Stephen L. and C. Winship. 2007. *Counterfactuals and Causal Inference. Methods and Principles for Social Research*. Cambridge: Cambridge University Press. Chapter 3 pp. 61-74

Other readings

Breen, Richard. 2018. "Some Methodological Problems in the Study of Multigenerational Mobility." *European Sociological Review* 34(6):603-11. doi: 10.1093/esr/jcy037.

Stephen R Cole, Robert W Platt, Enrique F Schisterman, Haitao Chu, Daniel Westreich, David Richardson, Charles Poole, Illustrating bias due to conditioning on a collider, *International Journal of Epidemiology*, Volume 39, Issue 2, April 2010, Pages 417–420, <https://doi.org/10.1093/ije/dyp334>

Acharya, Avidit, Matthew Blackwell and Maya Sen. 2016. "Explaining Causal Findings without Bias: Detecting and Assessing Direct Effects." *American Political Science Review* 110(3):512-29.

Assignment

Read the two compulsory reading and one among the other readings of your choice. Look for an example of an article in your research area that controls for a collider (“who is without sin, cast the first stone”..). Write a short comment (1 page) and represent the causal structure of the problem of the example that you have found, using a DAG.

Week 4. G-computation: An Introduction -- Guest speaker: Juho Harkonen (Tuesday, February 4, 2020)

Compulsory Readings

Hernán MA, Robins JM (2020). Causal Inference: What If. Boca Raton: Chapman & Hall/CRC.”, Ch. 12-14 & Ch. 19-21.

Naimi, Ashley I., Stephen R. Cole, and Edward H. Kennedy. "An introduction to g methods." *International journal of epidemiology* 46, no. 2 (2017): 756-762.

Other Readings

Lee, Dohoon, and Sara McLanahan. "Family structure transitions and child development: Instability, selection, and population heterogeneity." *American sociological review* 80.4 (2015): 738-763.

Witteaman, Jacqueline CM, Ralph B. D'Agostino, Theo Stijnen, William B. Kannel, Janet C. Cobb, Maria AJ de Ridder, Albert Hofman, and James M. Robins. "G-estimation of causal effects: isolated systolic hypertension and cardiovascular death in the Framingham Heart Study." *American Journal of Epidemiology* 148, no. 4 (1998): 390-401.

Assignment

Come up with an example from a published paper in which G-estimation could help to either: a) produce unbiased estimates of causal effects or b) help to shed light on mechanisms. Write a short comment (1 page) to explain the design, the problem and the G-based solution.

Week 5. Causal Mediation (Tuesday, February 11, 2020)

Imai K, Keele L, Tingley D, Yamamoto T. Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review*. 2011 Nov;105(4):765-89.

Acharya A, Blackwell M, Sen M. Explaining causal findings without bias: Detecting and assessing direct effects. *American Political Science Review*. 2016 Aug;110(3):512-29.

Acharya A, Blackwell M, Sen M. The political legacy of American slavery. *The Journal of Politics*. 2016 Jul 1;78(3):621-41.

Assignment

Causal mediation is everywhere in social sciences, in the sense that whenever we find a treatment effect we want to know the mechanism through which this effects manifests itself. Thinking of your own research, come up with an example of how you could use mediation analysis to shed light on the underlying mechanisms. Work alone or in a group to implement the actual analysis. Describe the approach and the results in a one-page response paper. Some of you will be asked to present their analysis in the lecture.

Week 6: Bounds (Tuesday, February 18, 2020)

Gerber, Alan S., and Donald P. Green. *Field experiments: Design, analysis, and interpretation*. WW Norton, 2012, Chapter 7.

Lee, David S. "Training, wages, and sample selection: Estimating sharp bounds on treatment effects." *The Review of Economic Studies* 76.3 (2009): 1071-1102.

Hirano, Shigeo, and James M. Snyder, Jr. "Using multimember district elections to estimate the sources of the incumbency advantage." *American Journal of Political Science* 53.2 (2009): 292-306.

Other Readings

Oster, Emily. "Unobservable selection and coefficient stability: Theory and evidence." *Journal of Business & Economic Statistics* 37.2 (2019): 187-204.

Assignment

Either: Use bounds to address the problem of non-response bias either in an experimental or in an observational study. Use either an example from your own research or let us know so we provide you with data and a setup to work on. Describe the approach and the results in a one-page response paper. Some of you will be asked to present their analysis in the lecture.

Or: Read the Oster paper and apply the bounding estimation to examine the sensitivity of your key estimate of interest (the regression coefficient of the X that is of interest to you on Y) across different control variables.

Week 7. Inference I: Parametric and non-parametric inference (Tuesday, March 3, 2020)

Gerber, Alan S., and Donald P. Green. *Field experiments: Design, analysis, and interpretation*. WW Norton, 2012, Chapter 3.

Efron, Bradley. "RA Fisher in the 21st century." *Statistical Science* (1998): 95-114.

Yates, Frank. "Sir Ronald Fisher and the design of experiments." *Biometrics* 20.2 (1964): 307-321.

Other Readings

Ho, Daniel E., and Kosuke Imai. "Randomization inference with natural experiments: An analysis of ballot effects in the 2003 California recall election." *Journal of the American Statistical Association* 101.475 (2006): 888-900.

Glynn, Adam N., and Nahomi Ichino. "Using qualitative information to improve causal inference." *American Journal of Political Science* 59.4 (2015): 1055-1071.

Assignment

From your own research, implement some analysis that entails both point and interval estimation and use at least two out of three following methods for the estimation of accompanying uncertainty of the point estimates: parametric, bootstrapped and randomization inference. If you do not have data to implement the analysis, do let us know so that we provide you with data and an example to work on.

Week 8: Inference II: clustering—when to cluster, how to cluster and what to do when having few clusters (Friday, March 6, 2020 at 15:00-17:00)

Abadie, Alberto, et al. *When should you adjust standard errors for clustering?*. No. w24003. National Bureau of Economic Research, 2017.

Canay, Ivan A., Andres Santos, and Azeem Shaikh. "The Wild Bootstrap with a Small Number of Large Clusters." *University of Chicago, Becker Friedman Institute for Economics Working Paper 2019-17* (2018).

Assignment

Continuing from last week's assignment, consider whether you should cluster your errors and if so within which set should the errors be clustered. Try different clustered-error estimators and examine how much the results differ. Summarize your choices and results in a one-page response paper.

Week 9. The insignificance of statistical significance (Tuesday, March 10, 2020)

Compulsory readings

Fabrizio Bernardi, Lela Chakhaia, Liliya Leopold, 'Sing Me a Song with Social Significance': The (Mis)Use of Statistical Significance Testing in European Sociological Research, *European Sociological Review*, Volume 33, Issue 1, February 2017, Pages 1–15, <https://doi.org/10.1093/esr/jcw047>

Other readings (for comments)

Gelman, A., & Carlin, J. (2014). Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspectives on Psychological Science*, 9(6), 641–651. <https://doi.org/10.1177/1745691614551642>

Benjamin, D. J., Berger, J., Johannesson, M., Nosek, B. A., Wagenmakers, E., Berk, R., ... Johnson, V. (2017, July 22). Redefine statistical significance. <https://doi.org/10.31234/osf.io/mky9j>

Brodeur, A., M. Lé, M. Sangnier, and Y. Zylberberg. 2016. "Star Wars: The Empirics Strike Back." *American Economic Journal: Applied Economics* 8(1):1-32.

Assignment

Read Bernardi et al. 2017 "Sing me Song..." .Consider the last number of your favorite journal in your research area and the the 5 item questionnaire used in Bernardi et al. 2017 to each article. In your written assignment:

- provide an example of misuse of statistical significance (for instance interpreting statistical significance as an indicator of substantive significance, or equating a statistical insignificant effect to a 0 effect. The template text could be: "in the article .., the authors estimates a model to study the relationship between. Intepreting the findings they comment HERE the quote"
- make a table with the results of the 5 item questionnaire to the articles you have examined.

Week 10. Implication analysis (Tuesday, March 17, 2020)

Compulsory readings

S. Lieberson 2002. Barking up the wrong tree. *Scientific Alternatives to the Current Model of Sociological Research. Annual. Review of Sociology* 28:1–19.

Lieberson, S. and Horwich, J. 2008. Implications analysis: a pragmatic proposal for linking theory and data in the social sciences. *Sociological Methodology* 38(1): 1-50.

[the article is discussed by several authors here:

<http://onlinelibrary.wiley.com/doi/10.1111/some.2008.38.issue-1/issuetoc>]

Tilly, C. (2008). Comment: Disturbing Implications. *Sociological Methodology*, 38(1), 85–89. <https://doi.org/10.1111/j.1467-9531.2008.00210.x>

Firebaugh, Glenn. 2008. "Comment: Implications of Implication Analysis." *Sociological Methodology* 38:51-58.

Other readings

Durkheim E. 1897. *Suicide*. [The whole book is a great read. It relies on implications which Durkheim derives from various conjectures on what might be causes of suicide; read especially chapter 4, on Imitation] [In the original French the book can be found here

<http://www.gutenberg.org/ebooks/40489>; in English here

<http://archive.org/details/suicidestudyinso00durk>]

Goldthorpe, John H. 2001. "Causation, Statistics, and Sociology." *European Sociological Review* 17(1):1-20.

S. Lieberson 2002. Barking up the wrong tree. *Scientific Alternatives to the Current Model of Sociological Research. Annual. Review of Sociology* 28:1–19.

<http://themonkeycage.org/2013/06/14/is-theory-getting-lost-in-the-identification-revolution>

<https://tompepinsky.com/2011/12/21/omfg-exogenous-variation-or-can-you-find-good-nails-when-you-find-an-indonesian-politics-hammer/>

Assignment

Read Lieberson, S. and Horwich, J. 2008 (disclaimer, the article is rather thick; suggestion: focus on the examples that are presented in the article and read the articles by Tilly and Firebaugh; plus other article in the same special issue). Browse the last number of your favorite journal in your research area. Does an article include something close to the implication analyses proposed by Lieberson/Horwich? Can you find an example of an article in your research area that that include an implication analysis? Write a short comment summarizing schematically such implication analysis (main theory/hypothesis and how it was tested drawing different implications)