PhD Course Description: Field Experiments in Child Research and Social Science

Teachers: Morten Hjortskov+* (coordinator), Simon Calmar Andersen+* and Michael Rosholm+†
+ TrygFonden's Centre for Child Research
* Department of Political Science, Aarhus University
† Department of Economics and Business Economics, Aarhus University

Time: Week 36, 2019, Monday through Friday, 9-16 – 24 lectures

ECTS: 5

Course description:
Field experiments are becoming a very important tool in the social science researcher's toolbox. Combining randomization and its advantages over many of our usual designs towards being able to claim causality with the natural contexts in the field holds an opportunity to maximize both internal and external validity. However, as with every other research design, it takes knowledge and skill to be able to design and carry out field experiments. And threats to causal inference still lure in the shadows. Even randomization sometimes must give in to the realities of the field!

TrygFondens's Center for Child Research at Aarhus University was established in 2013 as an interdisciplinary research center and has since worked with experimental and econometric methods in order to deliver state of the art research on child development, learning and education. One of the stated goals of the center is to disseminate knowledge on how to design, carry out and analyze large-scale field experiments within education – a method that was considered impossible to use in an educational setting only a few years ago.

The present course will take as its starting point the need for careful causal studies in education and social sciences. Policy and decision makers need research and evidence that addresses causal questions, and we as researchers need it to advance science. Examples of bad evidence that has turned in to actual policy is abound.

We start out by introducing the notation of potential outcomes as a common language well fitted to communicate about causal inference (Rubin 2005). Then we move on to the basics of experimental design and analysis: randomization, balance, hypothesis tests and stratification. Moreover, this part of the course will also focus on different design possibilities, such as clustered random assignment, factorial designs and power analysis. After this introduction, we start focusing on the potential threats to causal inference that may also exist in experimental research, not least when doing field experiments in real world settings: noncompliance, attrition and interference. The students will learn how to identify these threats and how to counter them both in the design and analytical phase. Lastly, we will look at how experiments can be used further down the road in downstream experimentation and structural models.

Another important part of the course will be to introduce the students to the many practical circumstances that play a substantial role when doing field experiments. The child research center now has comprehensible experience in running field experiments, and our secretariat will also visit the course to give the participants a feeling for the practicalities of it all. Aside from the workload in this part of running field experiments, the students will also get an opportunity to discuss when such circumstances result in some of the theoretical threats to causal inference which the course also covers.

Finally, students will have an opportunity to develop their own research design or project during the course. Each afternoon there will be exercises aiming at qualifying the students' projects in relation to the themes of the day. During these hours, the teachers will also be available for discussing specific designs and research questions regarding the students' projects.

Learning objectives:

At the end of the course the participants are expected to be able to:

- Describe and explain the pros and cons in different experimental designs
- Asses strengths and weaknesses in applied experimental research
- Use different experimental designs to address causal questions
- Analyze experimental data with relevant statistical methods
- Independently develop a research question and an experimental design during the course that can be used in the exam

Exam: Students hand in a written assignment after the course has ended. It will have to contain a motivated research question, an experimental design that is able to answer the research question and an analysis plan. One of the three teachers marks the assignment (failed/passed).

Course fee: The course is free of charge for PhD students in Social Sciences from Danish PhD programs. PhD students in Social Sciences from other universities need pay a tuition of 150 euros. The tuition covers refreshments and lunch on all days as well as a workshop dinner on September 5th , 2019. Participants are expected to cover their own transportation and accommodation costs (if needed).

Target audience and prerequisites: The course is targeted PhDs in social science programs such as economics, political science, psychology and sociology and related fields. However, any research project that plan to use field experiments involving people in social contexts could benefit from the course. The prerequisites for joining are a basic understanding of statistics like the standard error, t-tests, OLS regression, and logistic regression comparable to a master's degree in political science or psychology.

Maximum number of participants: 30

Course outline and readings

The course consists of 24 lectures in 5 days. This amounts to 5 lectures each of the 4 first days and 4 the last day. On top of the formalized lectures (including discussions, group work and statistical assignments in Stata), there will be a set of slots where the students are expected to work on a specific project – either their own project or an example project provided by the organizers. During these slots teachers will also be available for questions.

Day 1

Time: Monday September 2nd, 9-16

The first day of the course concentrates on introducing the fundamental building blocks of experiments and a common notational language (potential outcomes). We discuss the reasons for randomizing, the fundamental assumptions of experiments (excludability and non-interference), how we can test hypotheses in experimental designs and take a look at the advantages of blocking.

Lecture 1: Introduction to the course, teachers and field experiments

Lecture 2: Potential outcomes

Lecture 3: Why randomize? Randomization, causal inference and fundamental assumptions in experiments

Lecture 4: Sampling distributions and hypothesis tests

Lecture 5: Stratification/block randomization

Lecture 6: Instructive examples of field experimental designs

Project time: Students analyze simple and block randomized experimental data and work with their own project designs

Readings:

(Gerber and Green 2012, chapter 1-3 and 12, except pp. 80-86) (106 pages)

(Glennerster and Takavarasha 2013, chapter 1-2) (65 pages)

(LaLonde 1986) (17 pages)

(Czibor *et al.* 2019) (56 pages)

Pages: 244

Day 2

Day 2 starts out by looking more thoroughly at the different possibilities of utilizing covariates in experimental design and analysis. This includes more on the practice of blocking, an introduction to balance checks, interactions and how to enhance power by controlling for covariates in the analysis. Furthermore, an important design choice when doing field experiments in the education sector is whether to randomize at the individual (or unit) level, or randomize at some sort of cluster level. The cluster randomized experiments has many practical and design advantages, but also some pitfalls that should be taken seriously. Next, we will treat how to design appropriately powered experiments that are actually capable of finding the treatment effect (if it is out there). Day 2 will end with a set of inspiring examples of experimental design that might be used in the work with projects afterwards.

Lecture 7: Using covariates in experimental design and analysis

Lecture 8: Cluster randomized experiments

Lecture 9: Power analysis I

Lecture 10: Power analysis II

Project time: Students analyze experimental data with and without covariates and with and without clusters and work with their own project designs (calculate power and consider different designs)

Readings:

(Gerber and Green 2012, chapter 4 and pp. 80-86) (36 pages)

(Glennerster and Takavarasha 2013, chapter 6 and pp. 161-179) (58 pages)

(Raudenbush 1997) (13 pages)

(Bruhn and McKenzie 2009) (33 pages)

(Lortie-Forgues and Inglis 2019) (52 pages)

Pages: 211

Day 3

On day 3 we will start looking at the possible threats to causal inference that may also occur when doing field experiments. Specifically, we will look at compromised randomization, noncompliance and attrition. Moreover, we will dive a little more into the practicalities of running field experiments. The center's project managers will do a lecture on their work with pilot/feasibility studies, data collection, measurement and

how to keep the wheels of experimentation running. All of this is work that the students will have to take care of one way or another while running field experiments. When working in the field there are many practical issues to be taken care of, and in the worst case scenario, these will affect the experiment negatively.

Lecture 11: One-sided noncompliance

Lecture 12: Two-sided noncompliance

Lecture 13: Attrition

Lecture 14: Practicalities of running field experiments (project managers)

Lecture 15: Practicalities of running field experiments (project managers)

Project time: Students analyze experiments with noncompliance and assess the threat of noncompliance and attrition in their own projects.

Readings:

(Gerber and Green 2012, chapter 5-7) (103 pages)

(Glennerster 2017) (69 pages)

Pages: 172

<u>Day 4</u>

On day 4 we will take a deep dive into understanding the econometrics of instrumental variables that we use to solve the noncompliance challenge. Second, an often overlooked but severe challenge of field experiments is the non-interference assumption. We do not want spillover between experimental units, because it jeopardizes internal validity. We will look at how we can avoid interference, but also how we can seek to model it in multilevel experiments and how we can use non-experimental units to estimate spillover.

Lecture 16: Instrumental variables and the LATE estimator I

Lecture 17: Instrumental variables and the LATE estimator II

Lecture 18: Interference between experimental units and spatial spillover I

Lecture 19: Interference between experimental units and spatial spillover II

Lecture 20: Heterogeneous treatment effects and interactions

Project time: Students work with data examples on instrumental variable estimation and continue the conversation over dinner in the evening.

Readings:

(Wooldridge 2010, pp.937–954) (18 pages)

(Sovey and Green 2011) (13 pages)

(Sinclair *et al.* 2012) (15 pages)

(Gerber and Green 2012, chapter 8 and 9) (55 pages)

Pages: 101

Day 5

The last day of the course will first focus on the difficulties in measurement of mechanisms in experiments. Second, the integration of research findings in broader research agendas, replication and treatments that vary in intensity will be introduced. Lastly, Mette Trier Damgaard will introduce structural behavioral economics as an example of how to gain more insight into mechanisms of experiments.

Lecture 21: Mechanisms and integration of research findings

Lecture 22: Structural behavioral economics (Mette Trier Damgaard) I

Lecture 23: Structural behavioral economics (Mette Trier Damgaard) II

Lecture 24: Final remarks

Project time: Students work with their project proposals

Readings:

(Gerber and Green 2012, chapter 10 and 11) (52 pages)

Pages: 52

Pages in total: 780

Literature:

Bruhn, M., McKenzie, D. (2009) 'In pursuit of balance: Randomization in practice in development field experiments', *American economic journal: applied economics*, 1(4), 200–232.

Czibor, E., Jimenez-Gomez, D., List, J.A. (2019) *The Dozen Things Experimental Economists Should Do (More Of)*, Working Paper 25451, National Bureau of Economic Research, available: http://www.nber.org/papers/w25451 [accessed 16 May 2019].

Gerber, A.S., Green, D.P. (2012) *Field Experiments: Design, Analysis, and Interpretation*, W. W. Norton: New York.

Glennerster, R. (2017) 'The Practicalities of Running Randomized Evaluations: Partnerships, Measurement, Ethics, and Transparency', in *Handbook of Field Experiments*, North-Holland: Amsterdam, Netherlands, 175–243.

Glennerster, R., Takavarasha, K. (2013) *Running Randomized Evaluations - A Practical Guide*, Princeton University Press: Princeton and Oxford.

LaLonde, R.J. (1986) 'Evaluating the Econometric Evaluations of Training Programs with Experimental Data', *The American Economic Review*, 76(4), 604–620.

Lortie-Forgues, H., Inglis, M. (2019) 'Rigorous Large-Scale Educational RCTs are Often Uninformative: Should We Be Concerned? (in press)', *Educational Researcher*, 1–52.

Raudenbush, S.W. (1997) 'Statistical analysis and optimal design for cluster randomized trials.', *Psychological Methods*, 2(2), 173–185.

Rubin, D.B. (2005) 'Causal inference using potential outcomes', *Journal of the American Statistical Association*, 100(469), available: http://amstat.tandfonline.com/doi/full/10.1198/016214504000001880 [accessed 28 Oct 2013].

Sinclair, B., McConnell, M., Green, D.P. (2012) 'Detecting Spillover Effects: Design and Analysis of Multilevel Experiments', *American Journal of Political Science*, 56(4), 1055–1069.

Sovey, A.J., Green, D.P. (2011) 'Instrumental variables estimation in political science: A readers' guide', *American Journal of Political Science*, 55(1), 188–200.

Wooldridge, J.M. (2010) *Econometric Analysis of Cross Section and Panel Data*, 2nd Edition. ed, MIT Press: Cambridge, MA.