Set-Theoretic Methods
PhD Advanced Methods Course, Winter 2016
2 CEU credits, 4 ETCS

November 16, 2016

Instructor
Carsten Q. Schneider
Professor, Head of Department
Department of Political Science
Central European University
Room: FT 903
E-mail: schneiderc@ceu.edu
Phone: +36 327-3086

Classes
Mondays, 13:30 - 15:10 and Wednesdays, 9:00 - 10:40
Room: TBA

Office Hours
Room: FT 903
NB: Sign up online 24 hours before at https://carstenqschneider.youcanbook.me

Course description
This is an advanced methodological course on set-theoretic methods for the social sciences. While the spectrum of a set-theoretic methods is broad, including techniques such as Mill’s methods or typological theory, this course primarily focuses on the crisp-set and fuzzy-set versions of Qualitative Comparative Analysis (QCA). Invented by Charles Ragin [1987], this technique has undergone various modifications, improvements, and ramifications [Ragin, 2000, 2008]. These methods are applied in fields as diverse as political science, public policy, international relations, sociology, business and management studies, or even musicology (see www.compasss.org). This course aims at enabling students to produce a publishable QCA of their own. In order to achieve this, this course provides both the formal set theoretical underpinnings of QCA and the technical and research practical skills necessary for performing a QCA.

The course is structured as follows. We start with some basics of formal logic and set theory. Then we introduce the notions of sets and how they are calibrated. After this, we move on to the concepts of causal complexity and of necessity and sufficiency, show how the latter denote subset relations, and then learn how such subset relations can be analyzed with so-called truth tables. All concepts and analytic steps are first introduced based on crisp sets and then it is shown how they apply to fuzzy sets. Once students master the current standard analysis practice, we discuss several extensions and possible improvements of QCA. Depending on the needs and interests of participants, we choose several topics from the following list: set-theoretic multi-method research, i.e. the combination of QCA with follow-up within-case analyses; the
integration of time into QCA; or theory-evaluation in set-theoretic methods. The course is split into three blocks of two weeks each. During these blocks, we will meet twice a week.

Since this is an advanced PhD course, students who plan to attend should first check for themselves and, in case of doubt, with me whether they fulfill the following requirements: Participants should have (a) some practical experience in empirical comparative social research; (b) undergone some thorough courses in basic research methodology; and (c) preferably some basic statistical training, or at least hands-on knowledge with some sort of spreadsheet programs (even if it is just Excel). The core reading of the course is Schneider and Wagemann [2012]. Students who wish to take the course and need more information as to what the course is about are invited to read the first chapters of the book.

From the beginning, we will use specialized software for performing the analytic steps learned in class. We will use R (and RStudio) and within it, the packages QCA [Dusa, 2007] and SetMethods [Medzihorsky et al., 2016]. A desired (and very likely) side effect of this course will be that we engage into discussions on more general methodological issues of good comparative research, such as principles and practices of case selection, concept formation, measurement validity, and forms of causal relations.

Course Requirements

Presence, Participation, and Homework

Students are expected to be actively present at all seminars. In case you are unable to attend, you need to inform me via email prior to the meeting you are going to miss. Unexcused missed classes count with 0 points for participation on that specific day. During the seminars you are expected to reflect critically on the mandatory readings and to engage in discussions with your fellow students and myself. In general, for the grade the quality of participation prevails over its quantity, but if quantity is zero, quality is zero, too. Students who are present but do not actively participate receive the lowest passing grade for participation. Feedback on the class performance (including grade) will be provided if and when students sign up for an appointment during the office hours.

At the end of the teaching blocks 1 and 2, students will receive homework assignments. They are meant to consolidate the conceptual knowledge and foster the practical skills in performing set-theoretic analyses with the help of the relevant computer software. How much of these homework is done is at the discretion of each students, but the performance on this task influences the participation component of the overall grade for the course.

Mid-term exam

The mid-term exam will take place at, well, mid-term. It consists of a closed-book, 100 minute long exam during which participants are asked to answer a set of questions. These questions relate to the material we will have covered in the course up until this point and test the participants knowledge of the conceptual and mathematical aspects of set-theoretic methods.

Final paper

In the final paper participants will be asked to (a) replicate a published QCA and to (b) go beyond that publication by adding meaningful set-theoretic analysis to the same data. Deadline for the final will be roughly two weeks after the course is over. The precise date will be communicated at teh beginning of the course.

The grading follows the standard scale adopted by the Department of Political Science:
A: 100-94; A-: 93-87; B+: 86-80; B: 79-73; B-: 72-66; C+: 65-59; F: 58-0
### Learning outcomes and their assessments

The overall grade will primarily indicate the ability of the students to comprehend two things: (a) to understand the distinct logic of social inquiry that one is buying into when applying set-theoretic methods such as QCA and (b) to master the practical tricks of the trade when performing a QCA. The learning outcomes of the doctoral program are supported and measured by the present course in the following ways: The ability to reflect on some of the major methodological schools in the discipline; to deploy effective oral presentation and discussion skills as measured primarily by the in-class participation. The skills to employ cutting-edge methods are reflected by the mid-term exam and the final paper.

### Course material

This course heavily draws on Schneider and Wagemann [2012]. Paper and electronic copies of the book are available at CEU’s (virtual) library. Data sets and other course material, such as the slides for each session will also be made available through a password protected e-learning site at [http://ceulearning.ceu.edu/](http://ceulearning.ceu.edu/). Students who are officially enrolled in the course will have access to the e-learning site.

### Technicalities and relevant software

Our seminar sessions will take place in a normal class room rather than in the computer lab. Students are therefore asked to bring their own laptops to class. Participants are asked to download and install the following free software packages on their laptops prior to the beginning of the course.

1. R, [https://cran.r-project.org/](https://cran.r-project.org/)
2. R packages QCA, QCAGUI, SetMethods, and their dependencies
3. RStudio, [https://www.rstudio.com/](https://www.rstudio.com/)

### Useful books and sources of information

The following books are particularly relevant for this course.


**Some journals publishing articles on and with QCA**

European Journal of Political Research, Field Methods, Political Analysis, Political Research Quarterly, Quality and Quantity, Sociological Methods and Research, Studies in Comparative International Development, Comparative Political Studies
Course outline

Part 1 – The Basics

Session 1: Causal Complexity, Basics of Formal Logic and Set Theory, Sets and Their Calibration

The aim of this week's session is to introduce a specific notion of causal complexity and to show how some basic notions of formal logic and set theory provide powerful tools for expressing such complexity. Before this, we discuss the difference between variables and sets and introduce the basics of set calibration.

Mandatory readings:

Recommended readings:
Gary Goertz and James Mahoney. A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences. Princeton University Press, Princeton, 2012, chapter 2

Session 2: The notions of necessity and sufficiency

We learn which empirical data patterns are postulated by claims about necessary and about sufficient conditions and why set-theoretic methods are better equipped than standard statistical techniques for unraveling such data pattern.

Mandatory readings:

Recommended readings:

Session 3: Basics of truth table analysis

In this session, we introduce the form and function of truth tables. We discuss how a data matrix can be represented with a truth table and how the information contained in this truth table is analyzed with the use of logical minimization.

Mandatory readings:

Recommended readings:
Part 2 - Empirical Data meets Formal Logic

Session 4: Parameters of fit

We discuss the problems that arise whenever the tools of formal logic are applied to empirical data, which, almost by default, are noisy and incomplete. We start with the problem of contradictory truth table rows. When discussing how to tackle them we will introduce the parameter of fit consistency and after that also coverage. Both parameters provide a numerical expression of how well the logical statement contained in a QCA solution term reflects the empirical evidence at hand. We also engage with the analytic pitfalls that can emerge when so-called skewed sets are part of the analysis. Along these lines, we introduce additional (rather than alternative) parameters of fit.

Mandatory readings:

Recommended readings:

Session 5: Limited diversity

The second problem arising when formal logical tools meet empirical data is that of limited diversity. We learn what it is; what it does to causal inference; and which strategies are available for mitigating its impact on the results produced. Here the notion of counterfactuals is crucial and we learn how to distinguish between more and less plausible counterfactuals.

Mandatory readings:

Recommended readings:

Session 6: The Truth Table Algorithm

In this session, we simply put together what we have learned up until this point and thus replicate the so-called Truth Table Algorithm, the modal algorithm for performing QCA.

Part 3 - Advanced Issues

Session 7: Enhanced Standard Analysis

From this session onwards, we are addressing hitherto under-researched or even unresolved issues in set-theoretic methods. We start by issues surrounding the current best practices in
handling logical remainders and learn about suggestions on how to improve these current practices.

**Mandatory readings:**


**Recommended readings:**


**Session 8: Mid-Term Exam**

**Session 9: Set-Theoretic Multi-Method Research**

In this session, we learn about the principles and practices of set-theoretic multi-method research. This term has come to mean the combination of a cross-case QCA with within-case process tracing. We define different types of cases (typical, deviant, individually irrelevant), discuss how in a given data set the best available cases are identified, and which pairs of cases are best for which inferential goal of comparative within-case analysis.

**Mandatory readings:**


**Recommended readings:**


**Session 10: Robustness tests**

We discuss how the results generated by set-theoretic methods can and should be subjected to tests of robustness. We ask questions, such as: Robustness against which changes to the analysis? What are robust results in set-theoretic methods? Are simulations an appropriate tool for robustness tests?
Mandatory readings:


Recommended readings:


Session 11+12: Notions of time; Theory evaluation

In this session, we address two separate issues. First, we engage with the intricacies that are caused when trying to integrate notions of time into an empirical analysis, especially at the cross-case level. We discuss partial solutions, such as temporal QCA (tQCA) and two-step QCA. Second, we explore the possibilities for evaluating theoretical hunches based on empirical results generated with set-theoretic methods and how this differs from hypothesis testing in statistical approaches.

Mandatory readings:


Recommended readings:


References


