BSSn4495: Qualitative research in security studies

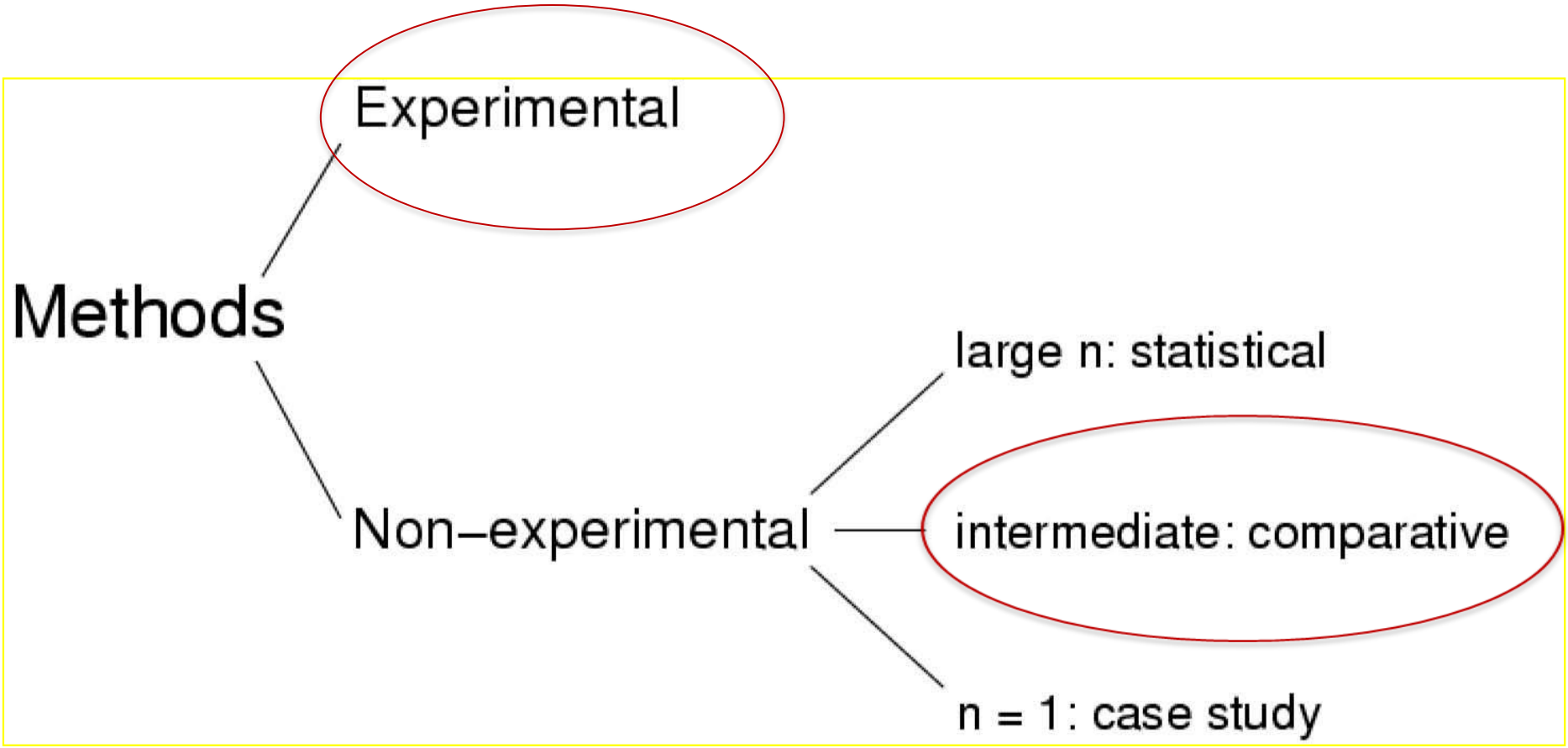
Strategies for causal identification: experiments and QCA

March 21, 2024

Miriam Matejova, PhD

Agenda

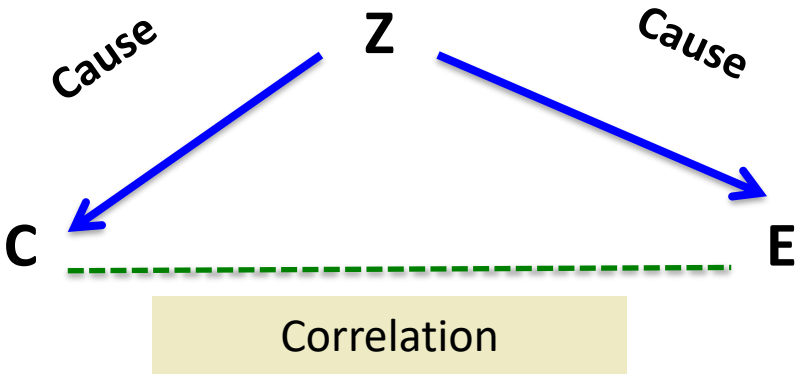
- Why experiments/QCA?
- When should/can we use experiments/QCA?
- What are the advantages/disadvantages of the experimental/QCA method?


Experimental
Methods
large n: statistical
Non–experimental
intermediate: comparative
n = 1: case study

Spurious correlation

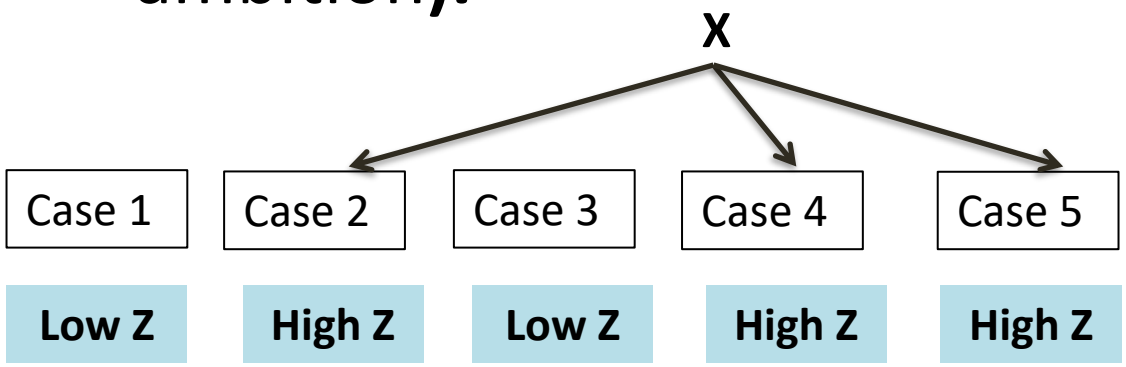
We may observe a **covariation (correlation)** between **C** and **E.**

BUT, this may be because **C** is NOT a cause of **E,** but because **Z** is a cause of BOTH **C** and **E**.



"Assignment" of Causes

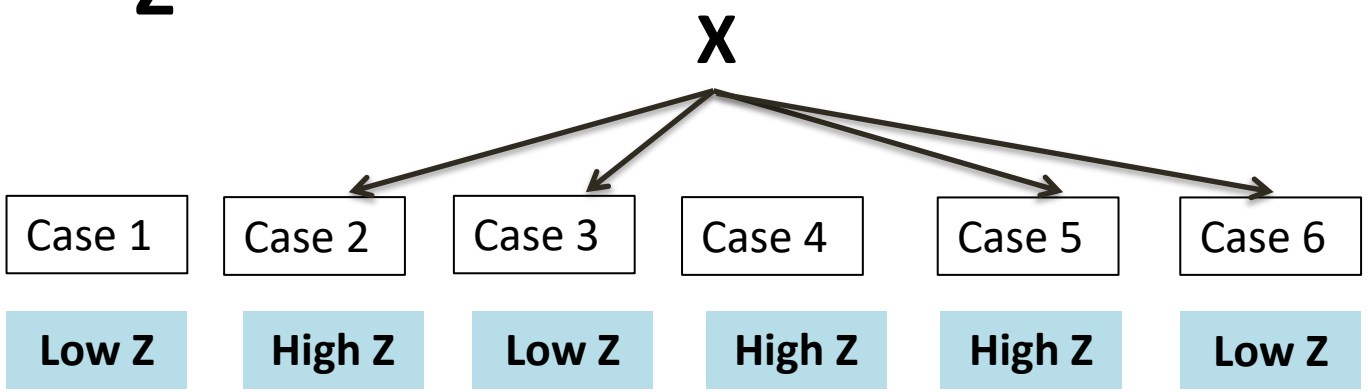
- Causal claim: **Attending University (X)** Causes **Higher Future Earnings (Y).**
- Each case represents an individual (a potential student)
- How is **Attending University (X)** **"assigned"** across cases in the real world? **X** is typically chosen by individuals on the basis of some **Z (e.g.** ambition**).**



Here, only cases with High Z get X=> Spurious Correlation between X and Y

“Assignment” of Causes

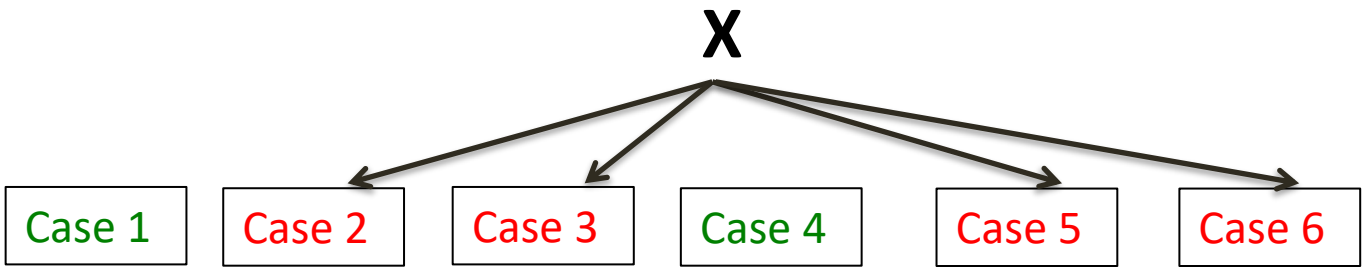
- But what if we let **researchers** assign **X** across cases in such a way that it does not depend on **Z**



Here Cases with High and Low Z are equally likely to be assigned **X => No Spurious Correlation between X and Y**

"Random Assignment"

"**Random assignment**" is a procedure for assigning **X** to cases that ensures that the difference in the value of the **Zs** between the cases that are assigned **X** and the cases that are not assigned **X** disappears as the number of cases gets large (law of large numbers)



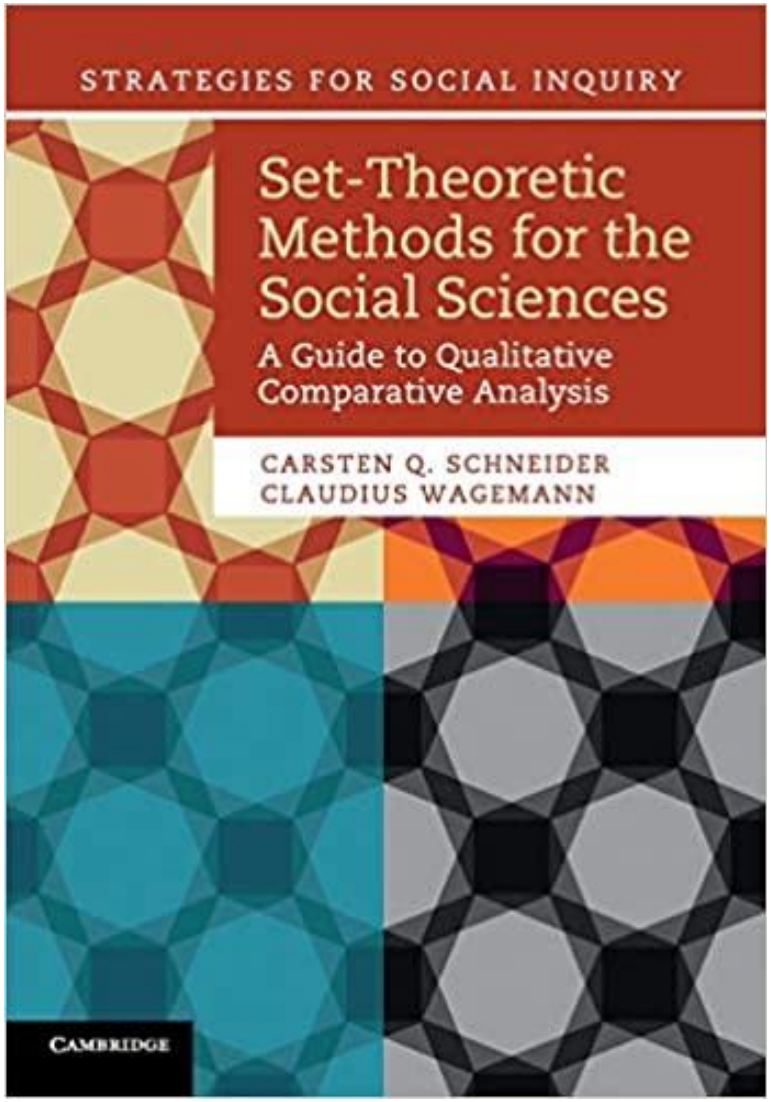
This procedure works even if the researcher does not know what the Z variables are or cannot measure them

Limitations of “random assignment” in social sciences

- Cost and ethics
- Artificial intervention by the researcher vs. real world applicability
  - The problem of generalization
- Cannot study the effects of things that have already happened
- Can get biased result if inappropriately designed

Qualitative comparative analysis (QCA)

- A set-theoretic method
- QCA as an approach and a data analysis technique



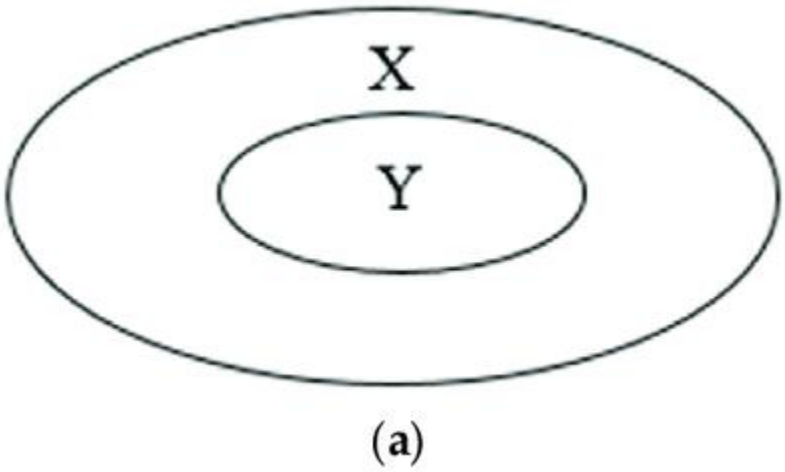
Set-theoretic methods

- The data consist of set membership scores
  - crisp, fuzzy, multi-value
- Relations between social phenomena modeled in terms of set relations
  - necessity, sufficiency, etc.
- The focus is on causal complexity
  - equifinality, conjunctural causation, etc.

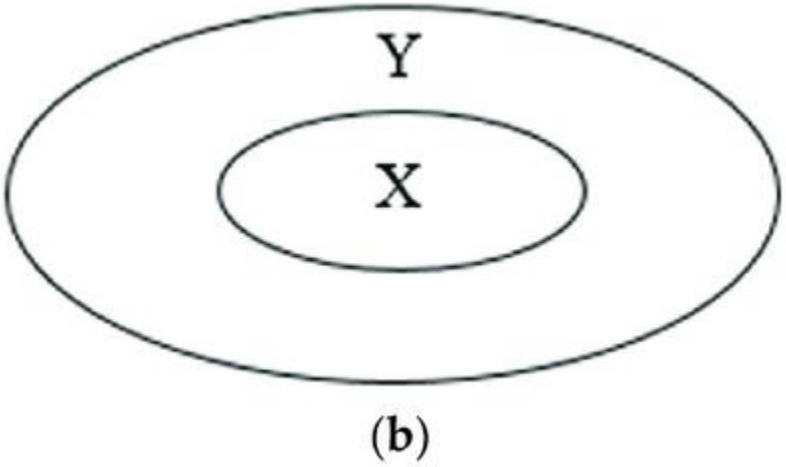
When do we use QCA?

- Causal complexity
  - *Multifinality*: same factor, different outcomes
  - *Equifinality*: different factors, same outcome
  - *Asymmetric causality*:
    - presence and absence of outcome have different explanations
      - economic growth → democratization
      - clientelism → non-democratization
    - Presence and absence of condition produce different outcomes
- Mid-sized N

Sets: necessary and sufficient conditions



(a)



(b)

What are sets?

- Establish qualitative, not quantitative, differences between cases
  - height ← not a set
  - tall person ← set

Sets vs. variables

	Sets	**Variables**
Labeling	Noun (object) and adjective (property of object): 'tall man'	Noun: 'height'
Data	Set membership scores (between 0-1)	Numbers (preferably unbound)
Information	Difference in type (qualitative differences)	Difference in degree
Data generation	Calibration	Measurement
Operations	Formal logical rules	Standard math

Schneider 2017

Types of sets: crisp set

- Dichotomous sets
- Full member (1) vs. full non-member (0)
  - Establishes qualitative, not quantitative, differences between cases
  - E.g., set of big countries
    - China, Russia (1) vs. Hungary, Lichtenstein (0)

Types of sets: fuzzy sets

- Allow for degree of membership in set
- Partial membership in sets
  - Any value between 0 and 1
  - Three qualitative anchors (0, 0.5, 1)
  - Qualitative and quantitative differences
- NOT probabilities

CRISP VERSUS FUZZY SETS

Crisp set	Three-value fuzzy set	Four-value fuzzy set	Six-value fuzzy set	"Continuous" fuzzy set
1 = fully in	1 = fully in	1 = fully in	1 = fully in	1 = fully in
			.8 = mostly but not fully in	Degree of membership is more "in" than "out": $.5 < x_i < 1$
		.75 = more in than out	.6 = more or less in	
	.5 = neither fully in nor fully out			.5 = cross-over: neither in nor out
		.25 = more out than in	.4 = more or less out	Degree of membership is more "out" than "in": $0 < x_i < .5$
			.2 = mostly but not fully out	
0 = fully out	0 = fully out	0 = fully out	0 = fully out	0 = fully out

Schneider 2017

QCA: Steps

1) Assemble the universe of cases
2) Collect raw data
3) Calibrate conditions sets and outcome sets
4) Search for necessary conditions
5) Represent empirical evidence in a truth table
6) Identify sufficient conditions by logically minimizing the truth table
7) Do within-case analyses in typical and deviant cases

QCA challenges

- Location of qualitative anchors
- Sometimes false impression of precision
- Resources, time, data availability