Qualitative Comparative Analysis in Practice: Why, When, and How?

Claude Rubinson
University of Houston—Downtown
Houston, TX

rubinsonc@uhd.edu
http://gator.uhd.edu/~rubinsonc/
http://grundrisse.org/qca/

University of Colorado
School of Medicine
Adult and Child Consortium for Health Outcomes Research and Delivery Science
April 26, 2018

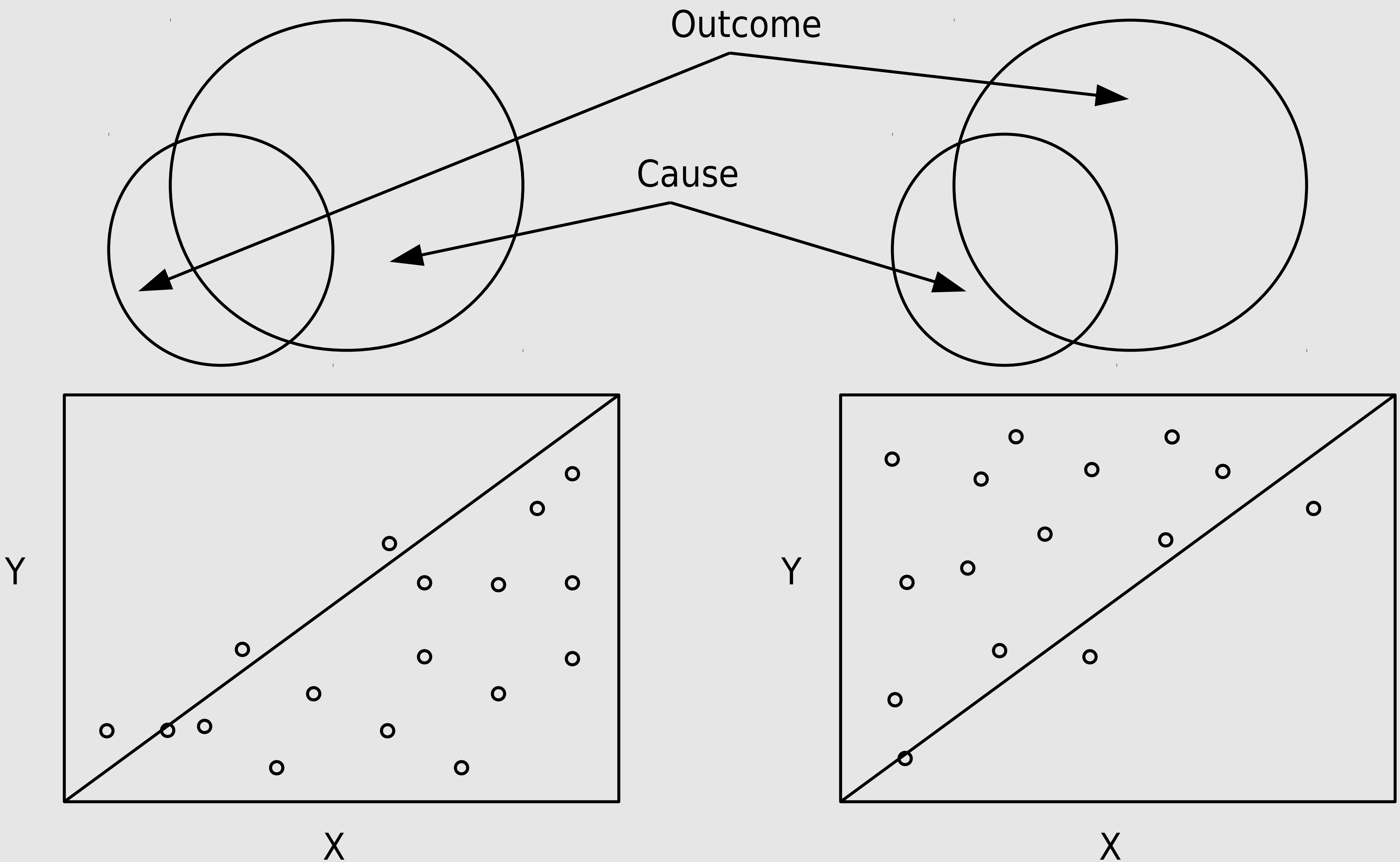
Outline

- Day 1: Introduction and Overview of QCA
    - Review of QCA resources, publications, and software
    - QCA as an investigation of invariance
    - Three analytic components of QCA:
      dataset calibration, necessity analysis, and sufficiency analysis
    - Three types of QCA projects:
      identifying causal recipes, uncovering taxonomies, understanding context
- Day 2: The Details of Why, When, and How
    - Review
    - Dataset calibration
    - Necessity analysis
        - Consistency and coverage measures for necessity
        - Testing for necessary conditions
    - Sufficiency analysis
        - Consistency and coverage measures for sufficiency
        - Constructing and reducing truth tables
        - Interrogating the analysis and deriving solutions

QCA as the Study of Invariance

- QCA is a case-oriented, set-theoretic technique for identifying invariant relationships by analyzing the strength of superset/subset relationships
  - Tenured faculty tend to have many publications
  - Religious fundamentalists tend to be politically conservative
  - HIV causes AIDS (i.e., the set of people with AIDS is a subset of the set of people exposed to HIV)
  - Pregnancy termination may occur due to miscarriage or elective abortion

QCA as the Study of Invariance



Subset relationship consistent with *necessity* (X≥Y)

Subset relationship consistent with *sufficiency* (Y≥X)

Boolean Algebra

- UPPERCASE for the presence of a condition
- lowercase or ~ for the absence of a condition
- Negation

```
~A = 1 – A
 a = 1 – A
```

- Logical and (Boolean multiplication/Set intersection)

```
A•b = Ab = min(A,b)
```

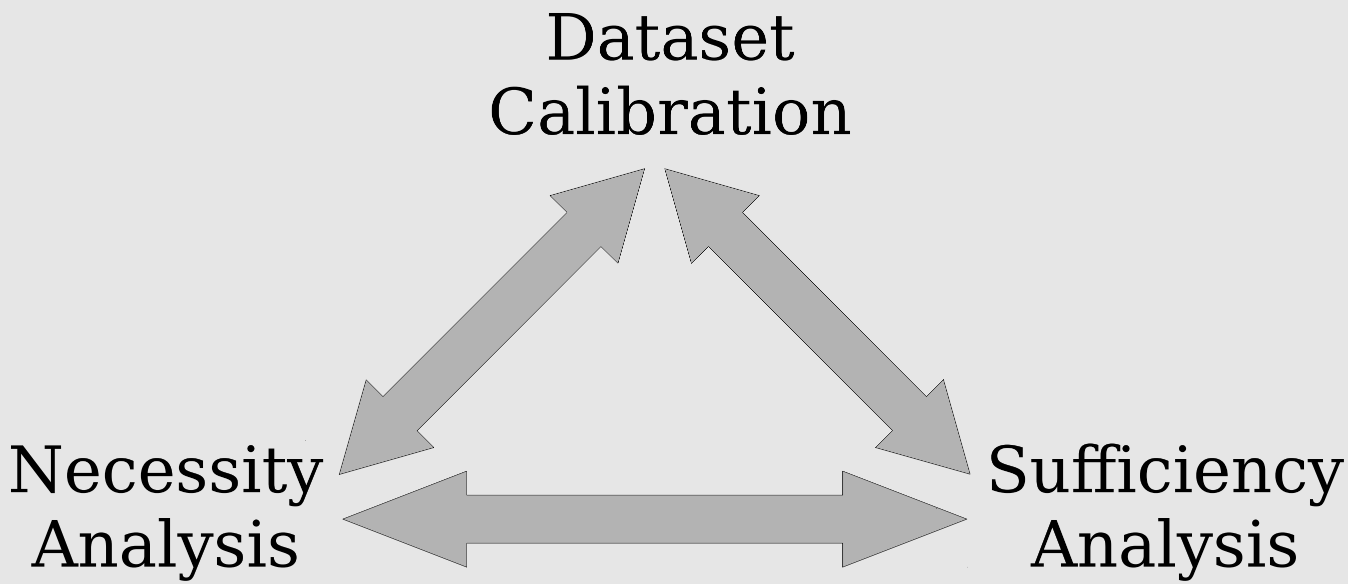
- Logical or (Boolean addition/Set union)

```
A+b = max(A,b)
```

Distinguishing Features of QCA

- Assumption of invariance
- Assumption of causal complexity
  - Identification of necessary and sufficient conditions
  - There can be multiple paths to the same outcome
- No degrees-of-freedom restrictions
  - Appropriate for small-, medium-, and large-N analysis
- Encourages retroductive analysis (moving back and forth between theory and data)
  - Uses a malleable analytic frame
  - Must identify, measure, and scale (calibrate) your causal conditions and outcome
  - Dataset must include both positive and negative outcomes
  - Identifying and resolving contradictions is key

Three Analytic Components of QCA



Calibrating Datasets

Data Set Calibration

- Instrument calibration is routine in the natural sciences; largely absent in the social sciences.
- Scientific data collection and measurement typically emphasizes relative effects: Paul is poorer than Peter; the United States' infant mortality rate is greater than that of Japan.
- Calibration allows us to state that an individual is poor or that a country's infant mortality rate is high.
- Calibration requires application of theoretical and substantive knowledge: A T-cell count of below 200 µL of blood is sufficient to diagnose AIDS

Calibrating 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
	0.67 = more in than out	0.67 = more in than out	0.8 = mostly but not fully in 0.6 = more or less in	Degree of membership is more "in" than "out" $0.5 < X < 1$
-------------------------------- 0.5 = Crossover Point ---------------------------------				
		0.33 = more out than in	0.4 = more or less out 0.2 = mostly but not fully out	Degree of membership is more "out" than "in" $0.0 < X < 0.5$
0 = fully out	0 = fully out	0 = fully out	0 = fully out	0 = fully out

Calibrating Fuzzy Sets

- Methods of calibration:
  - Manually
  - “Direct” Method
  - “Indirect” Method
- Fuzzy sets are asymmetrical
- Fuzzy sets vs crisp-sets vs dummy variables
- Fuzzy sets vs multi-valued sets

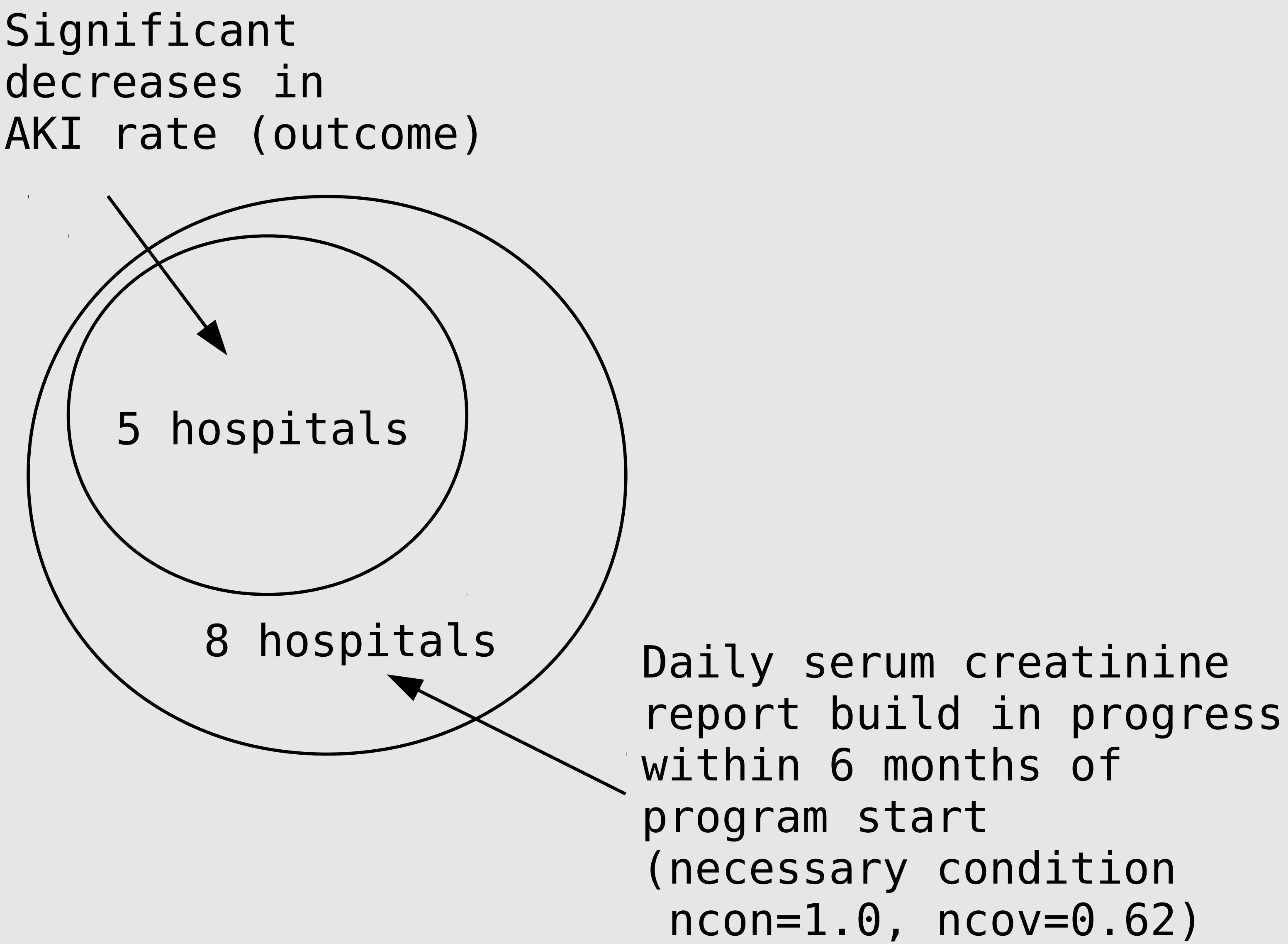
Analyzing Necessary Conditions

Necessity Analysis

- Underdeveloped in the literature; QCA development has focused on sufficiency analysis
- *Kirq* and *acq* have sophisticated necessity testing

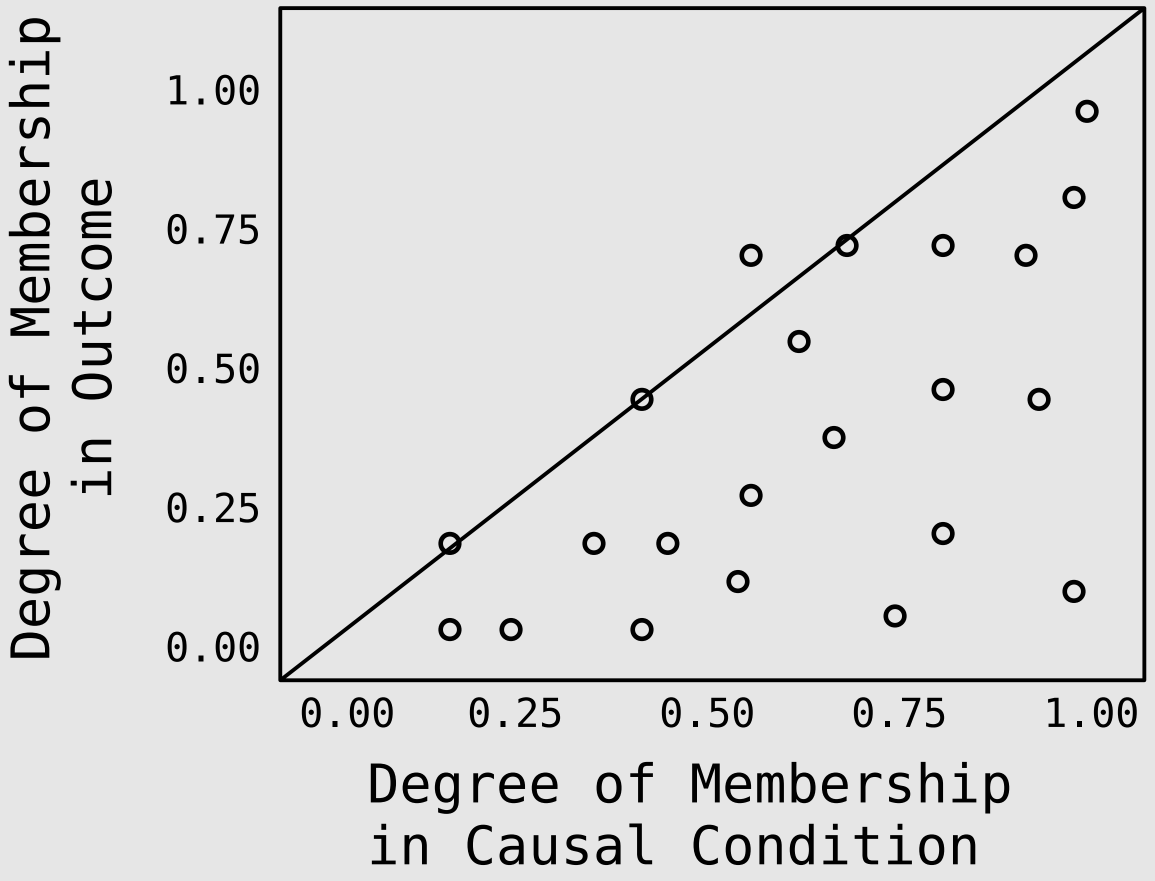
Necessary Conditions

Outcome is subset of cause: Causal condition must (almost always) be present for outcome to occur.



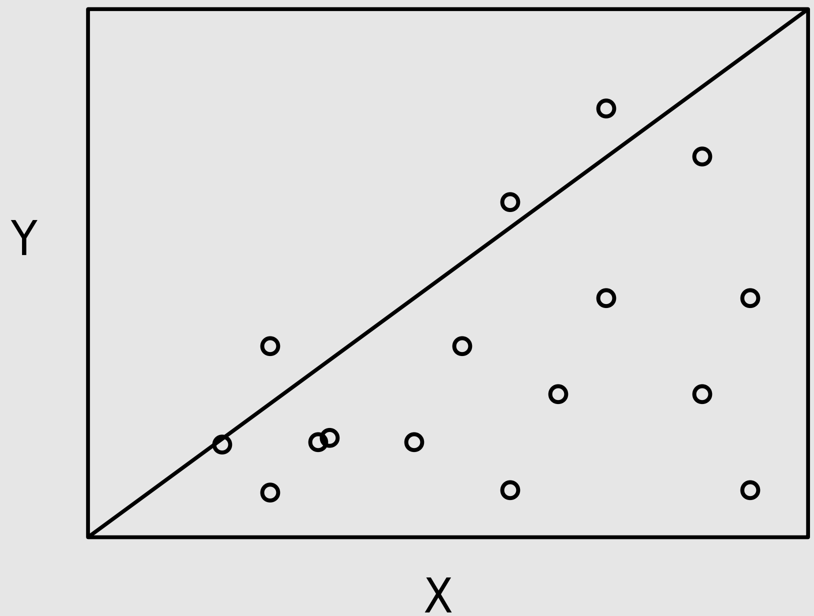
Fuzzy Subset Relationship Consistent with Necessity

Outcome is a subset of Cause ($X \geq Y$)

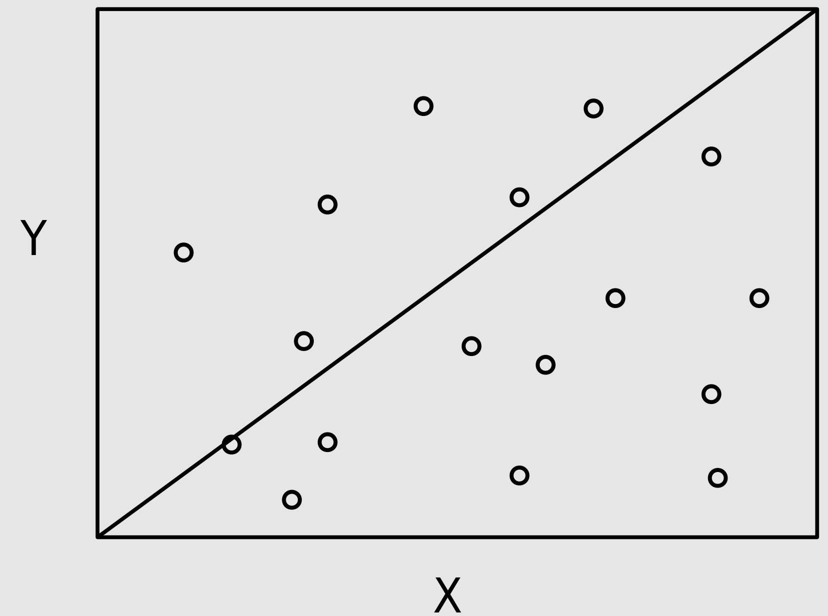


Assessing Necessary Conditions

- *Consistency* measures degree to which subset relationship is "consistent" with necessity



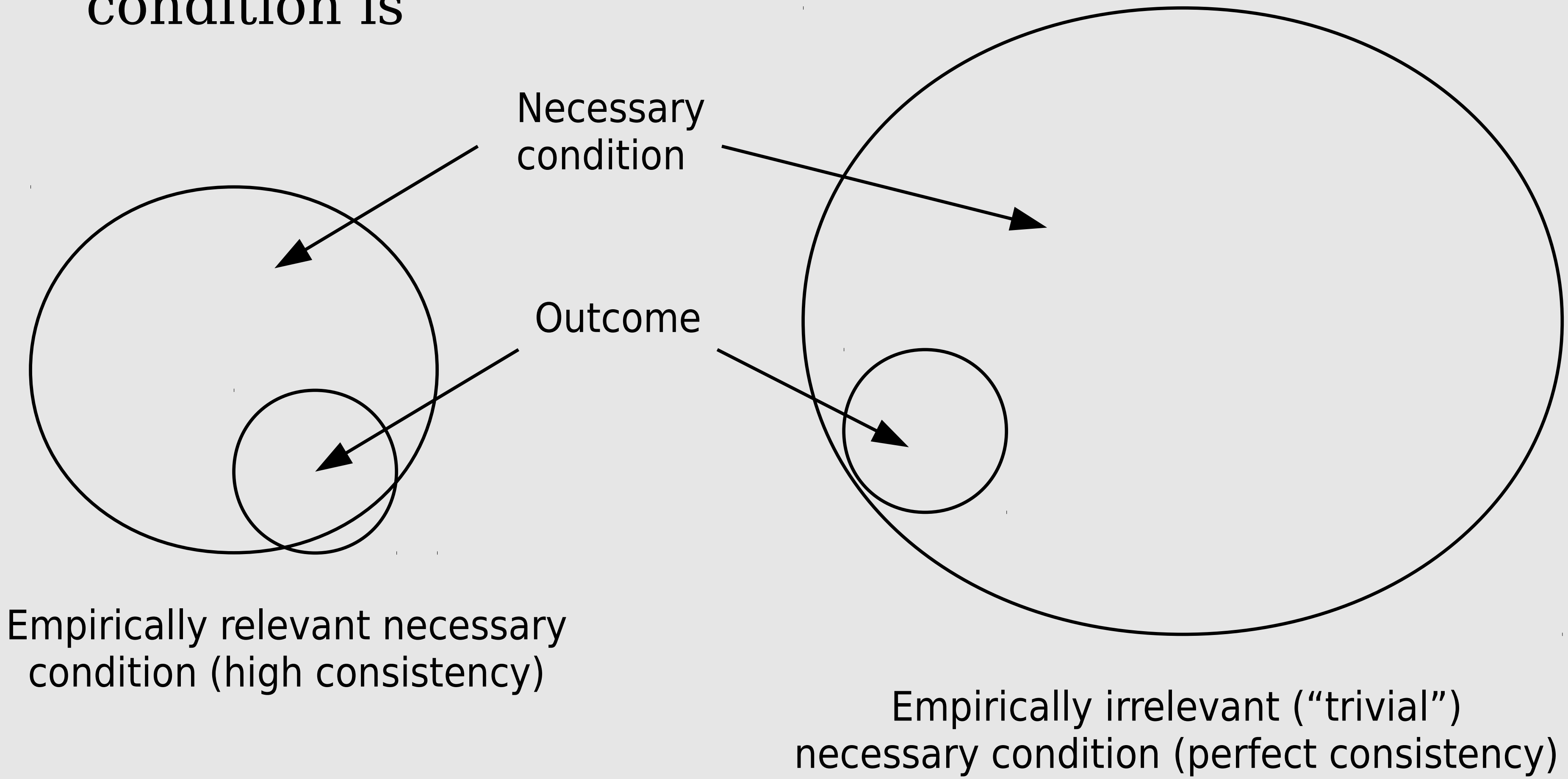
Subset relationship consistent with necessity



Subset relationship with substantial inconsistency

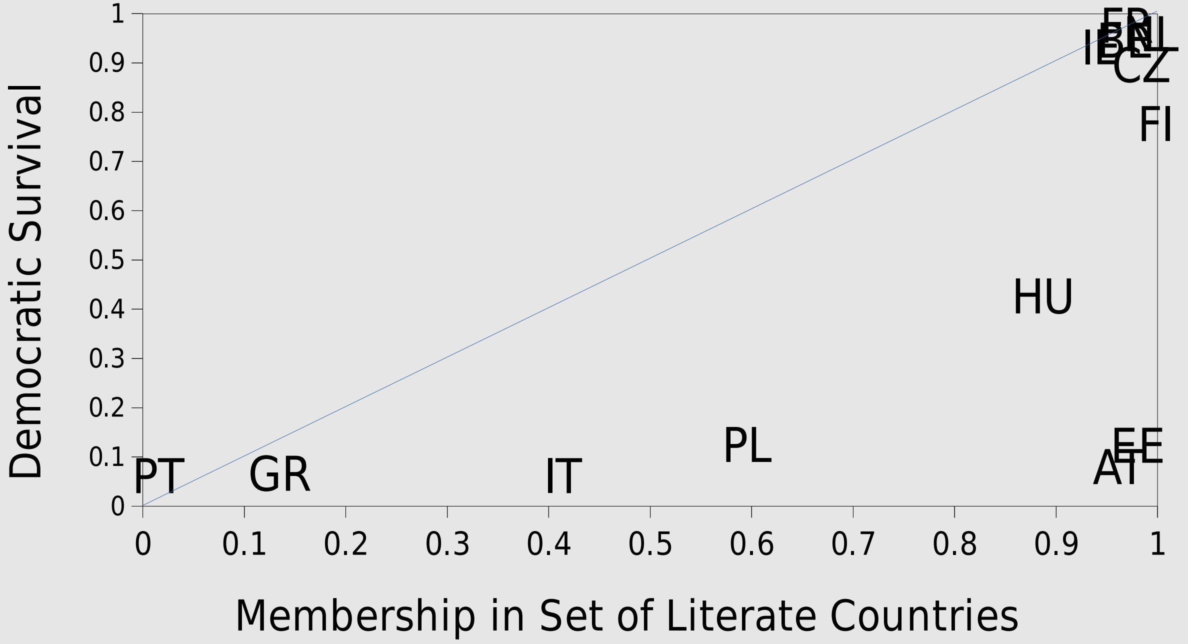
Assessing Necessary Conditions

- *Coverage* measures how “relevant” a necessary condition is



Testing for Necessary Conditions

Obs	Dev	Urb	Lit	Sur
AT	.81	.12	.99	.05
BE	.99	.89	.98	.95
CZ	.58	.98	.98	.89
EE	.16	.07	.98	.12
FI	.58	.03	.99	.77
FR	.98	.03	.99	.95
DE	.89	.79	.99	.05
GR	.04	.09	.13	.06
HU	.07	.16	.88	.42
IE	.72	.05	.98	.92
IT	.34	.10	.41	.05
NL	.98	1.00	.99	.95
PL	.02	.17	.59	.12
PT	.01	.02	.01	.05



Term	Consis	Cov
LIT	0.99	0.58
Solution	0.99	0.58

Testing for Necessary Conditions

- Assess consistency before coverage
- Join terms with logical or (e.g., A+B+C)
- Many solutions are possible
- Use of theory is crucial

Analyzing Sufficient Conditions

Sufficiency Analysis

- More mature than necessity analysis; QCA development—and applications—have focused on sufficiency analysis
- Emphasis on causal complexity (a.k.a., multiple conjunctural causation, “recipes,” or equifinality; also, INUS conditions)

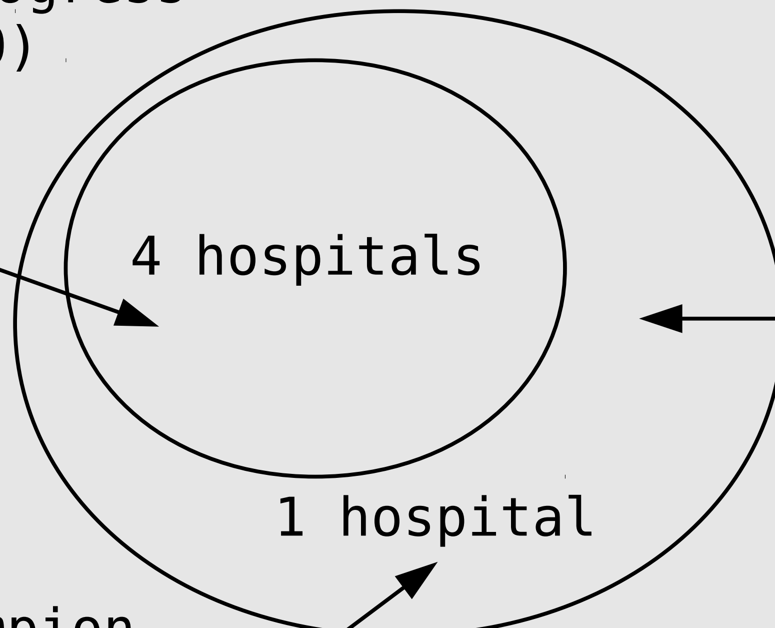
Feature	fs/QCA	Kirq & acq
Based on RSI Algorithms	√	√
Complex Solutions	√	√
Intermediate Solutions	√	
Parsimonious Solutions	√	√
Impossible Conditions		√
Contradictions		√

Sufficient Conditions

Cause is subset of outcome: Outcome (almost) always occurs when causal condition is present.

- Pharmacist champion w/hours,
- 2+ pharmacists assigned, and
- Report build in progress

(scon=1.0, scov=0.80)



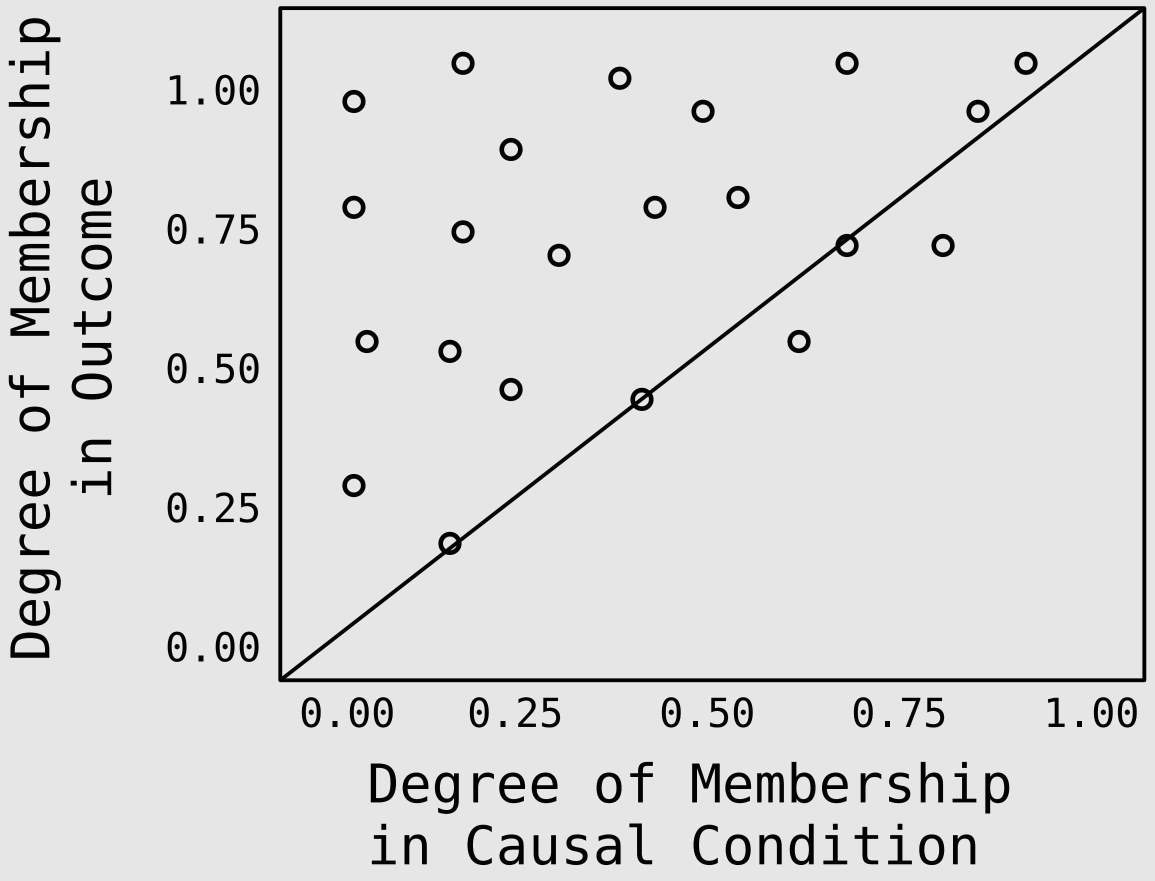
Significant decreases in AKI rate (outcome)

- No pharmacist champion,
- No assigned pharmacists,
- No QI or data champion,
- No consistent report build,
- But: No other organizational priorities

(scon=1.0, scov=0.20)

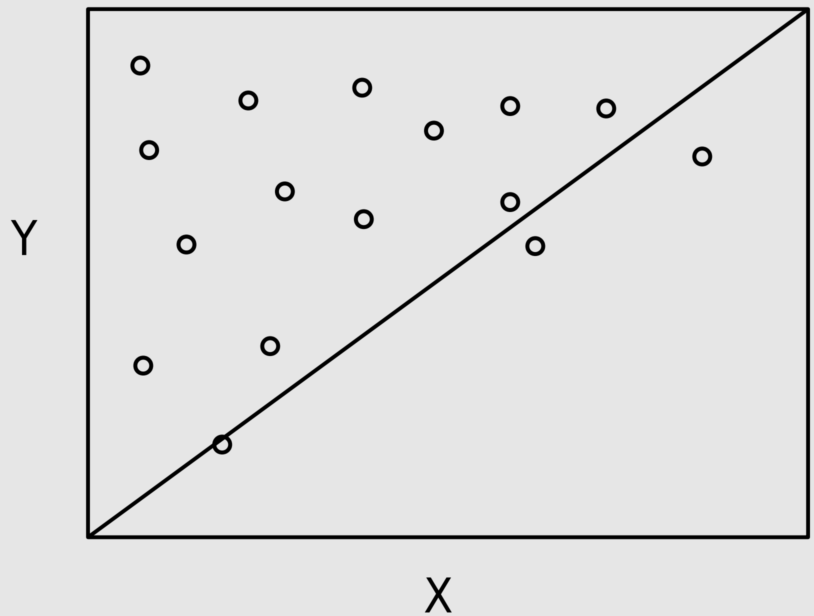
Fuzzy Subset Relationship Consistent with Sufficiency

Cause is a subset of Outcome ($Y \geq X$)

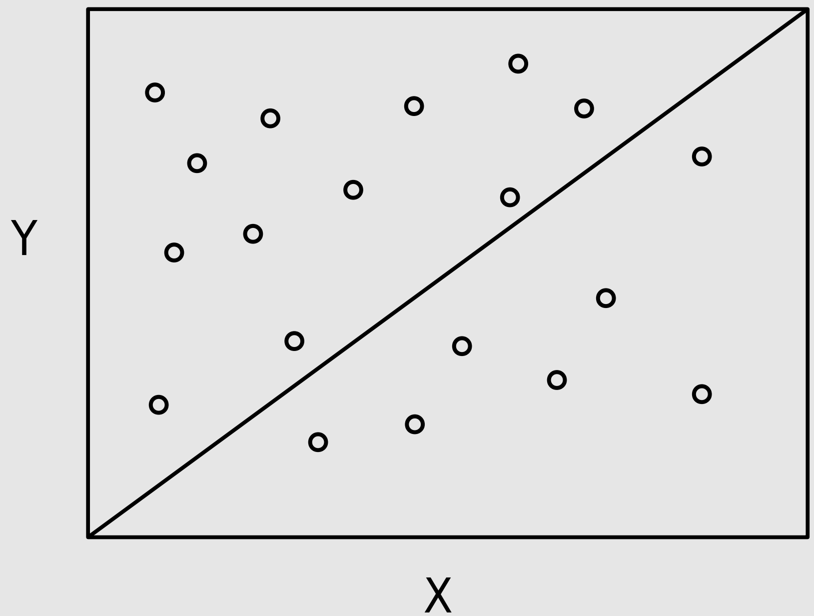


Assessing Sufficient Conditions

- *Consistency* measures degree to which subset relationship is "consistent" with sufficiency



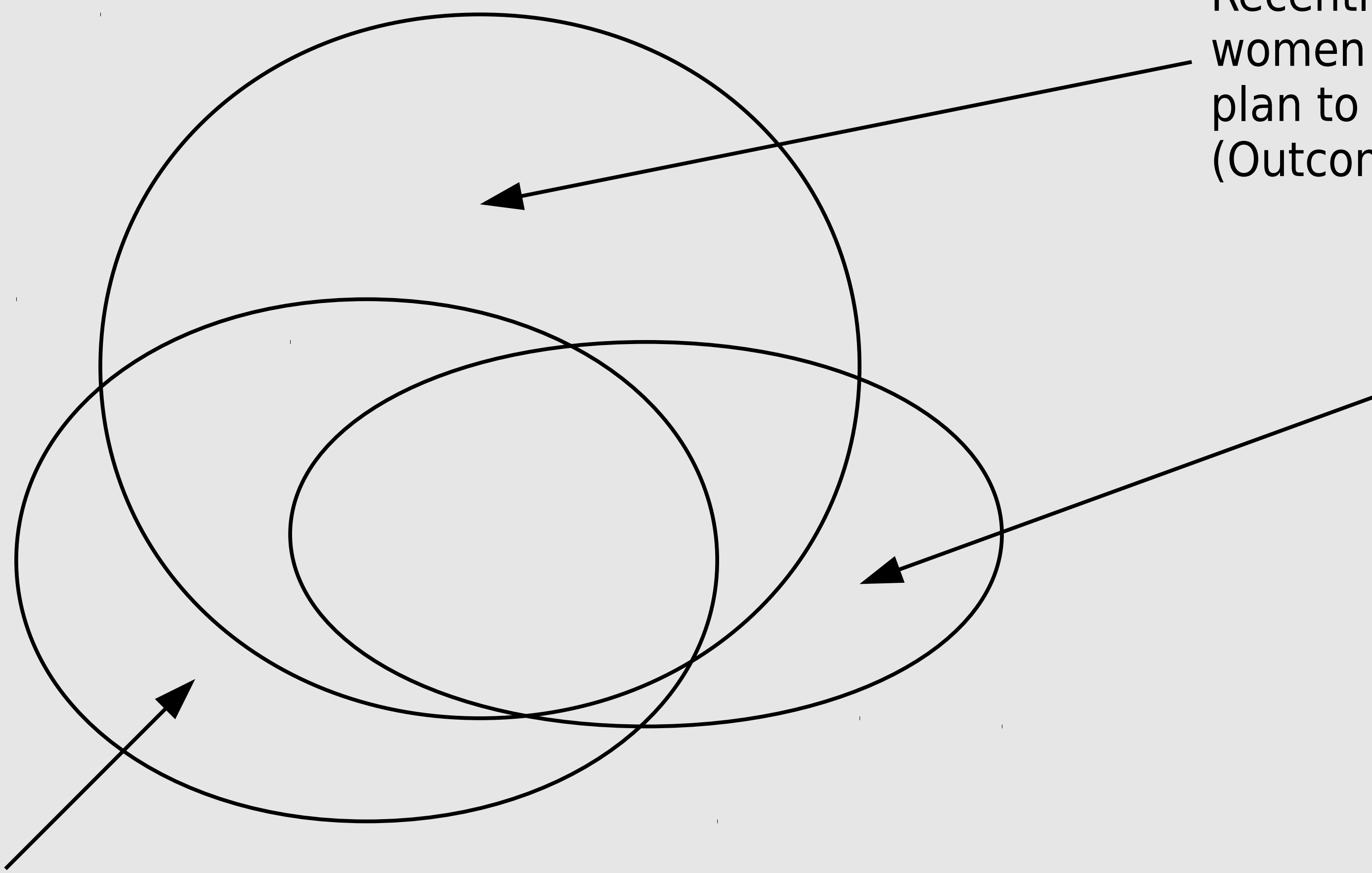
Subset relationship consistent with sufficiency



Subset relationship with substantial inconsistency

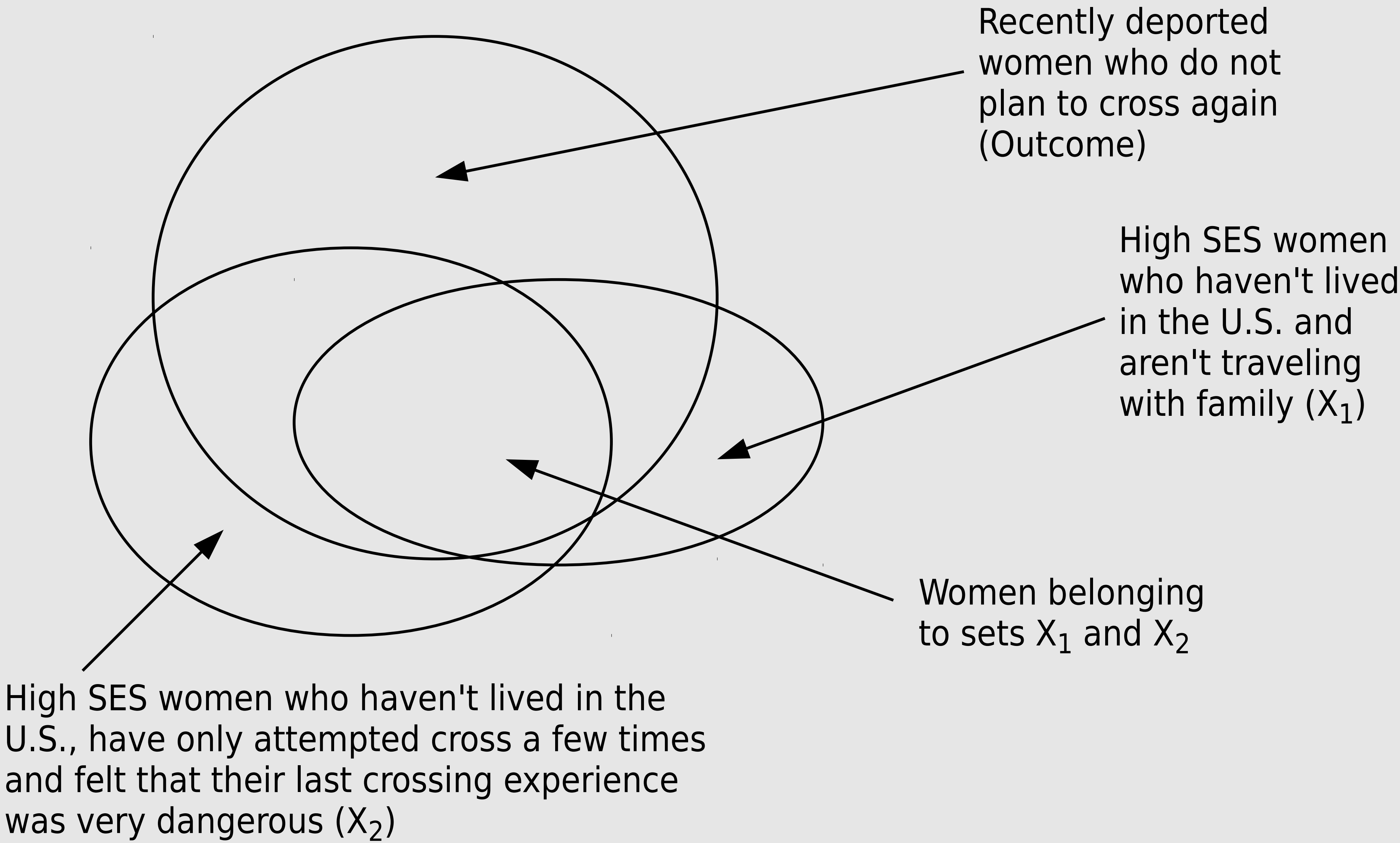
Assessing Sufficient Conditions

- *Coverage* measures the relative "importance" of each solution



Assessing Sufficient Conditions

- *Coverage* measures the relative "importance" of each solution



Testing for Sufficient Conditions

Term	Consis	Raw Cov	Uniq Cov
HISES*liveus*travfam +	0.90	0.32	0.13
HISES*liveus*numcross*DANGER	0.82	0.46	0.26
Solution	0.86	0.58	

Truth Table Construction

Truth table algorithm sorts observations into types

Obs	Dev	Urb	Lit	Brk
AT	.81	.12	.99	.95
BE	.99	.89	.98	.05
CZ	.58	.98	.98	.11
EE	.16	.07	.98	.88
FI	.58	.03	.99	.23
FR	.98	.03	.99	.05
DE	.89	.79	.99	.95
GR	.04	.09	.13	.94
HU	.07	.16	.88	.58
IE	.72	.05	.98	.08
IT	.34	.10	.41	.95
NL	.98	1.00	.99	.05
PL	.02	.17	.59	.88
PT	.01	.02	.01	.95

	Dev	Urb	Lit	Consis	Y	Consis Obs	Inconsis Obs
1	T	T	T	0.41	F	DE	BE, CZ, NL
2	T	T	F	—	—		
3	T	F	T	0.51	F	AT	FI, FR, IE
4	T	F	F	—	—		
5	F	T	T	—	—		
6	F	T	F	—	—		
7	F	F	T	0.83	T	EE, PL	HU
8	F	F	F	0.99	T	GR, IT, PT	

Reading Truth Tables

Truth table assesses consistency between types and outcome

Democracy usually did not break down in countries that were
(a) developed, urbanized, and literate (row 1) or
(b) developed, not urbanized, and literate (row 3).

Democracy usually did break down in countries that were
(c) not developed, not urbanized, and literate (row 7) or (d) not developed, not urbanized, and not literate (row 8)

	Dev	Urb	Lit	Consis	Y	Consis Obs	Inconsis Obs
1	T	T	T	0.41	F	DE	BE, CZ, NL
2	T	T	F	—	—		
3	T	F	T	0.51	F	AT	FI, FR, IE
4	T	F	F	—	—		
5	F	T	T	—	—		
6	F	T	F	—	—		
7	F	F	T	0.83	T	EE, PL	HU
8	F	F	F	0.99	T	GR, IT, PT	

Reading Truth Tables

Remainders are logically possible conditions lacking empirical instances

Remainders

	Dev	Urb	Lit	Consis	Y	Consis Obs	Inconsis Obs
1	T	T	T	0.41	F	DE	BE, CZ, NL
2	T	T	F	—	—		
3	T	F	T	0.51	F	AT	FI, FR, IE
4	T	F	F	—	—		
5	F	T	T	—	—		
6	F	T	F	—	—		
7	F	F	T	0.83	T	EE, PL	HU
8	F	F	F	0.99	T	GR, IT, PT	

Invariance in Truth Tables

	Dev	Urb	Consis	Y	Consis Obs	Inconsis Obs
1	T	T	0.41	F	DE	BE, CZ, NL
2	T	F	0.51	F	AT	FI, FR, IE
3	F	T	—	—		
4	F	F	0.89	T	EE, GR, IT, PL, PT	HU

	Dev	Urb	Lit	Consis	Y	Consis Obs	Inconsis Obs
1	T	T	T	0.41	F	DE	BE, CZ, NL
2	T	T	F	—	—		
3	T	F	T	0.51	F	AT	FI, FR, IE
4	T	F	F	—	—		
5	F	T	T	—	—		
6	F	T	F	—	—		
7	F	F	T	0.83	T	EE, PL	HU
8	F	F	F	0.99	T	GR, IT, PT	

Reducing Truth Tables to Boolean Equations

To Primitive Expressions:

Term	Consis	Raw Cov	Uniq Cov	Observations
dev*urb*LIT +	0.83	0.42	0.27	EE, PL, [HU]
dev*urb*lit	0.99	0.40	0.24	GR, IT, PT
Solution	0.88	0.66		

Reducing Truth Tables to Boolean Equations

To Primitive Expressions:

Term	Consis	Raw Cov	Uniq Cov	Observations
dev*urb*LIT +	0.83	0.42	0.27	EE, PL, [HU]
dev*urb*lit	0.99	0.40	0.24	GR, IT, PT
Solution	0.88	0.66		

To Prime Implicants:

Term	Consis	Raw Cov	Uniq Cov	Observations
dev*urb	0.89	0.71	0.71	EE, PL, GR, IT, PT, [HU]
Solution	0.89	0.71		

Reducing Truth Tables to Boolean Equations

Reduce Prime Implicants (Complex Solution):

Term	Consis	Raw Cov	Uniq Cov	Observations
dev*urb	0.89	0.71	0.71	EE, PL, GR, IT, PT, [HU]
Solution	0.89	0.71		

Reducing Truth Tables to Boolean Equations

Reduce Prime Implicants (Complex Solution):

Term	Consis	Raw Cov	Uniq Cov	Observations
dev*urb	0.89	0.71	0.71	EE, PL, GR, IT, PT, [HU]
Solution	0.89	0.71		

Reduce Prime Implicants Using Remainders (Parsimonious Solution):

Term	Consis	Raw Cov	Uniq Cov	Observations
dev	0.82	0.73	0.73	EE, PL, GR, IT, PT, [HU]
Solution	0.82	0.73		

Constructing Intermediate Solutions

Manually, or via directional expectations

Complex Solution		Parsimonious Solution
Acsir + ACSir + ASIR	→	i + SR

Multiple intermediate solutions are possible:

Air + ACSi + ASIR	→	Air + ASIR	→	Ai + ASR