AARHUS UNIVERSITY

# Process Tracing methods – an introduction

Ph.D. workshop
University of Konstanz, Germany

March 16, 2012

Derek Beach, PhD
Associate Professor
Department of Political Science
University of Aarhus, Denmark

Email: derek@ps.au.dk

DEPARTMENT OF POLITICAL SCIENCE

# Outline

1. What is Process Tracing?
2. What are causal mechanisms?
3. Three variants of PT
4. Causal inference in PT
5. Studying causal mechanisms?
6. When can PT be used, and not used?

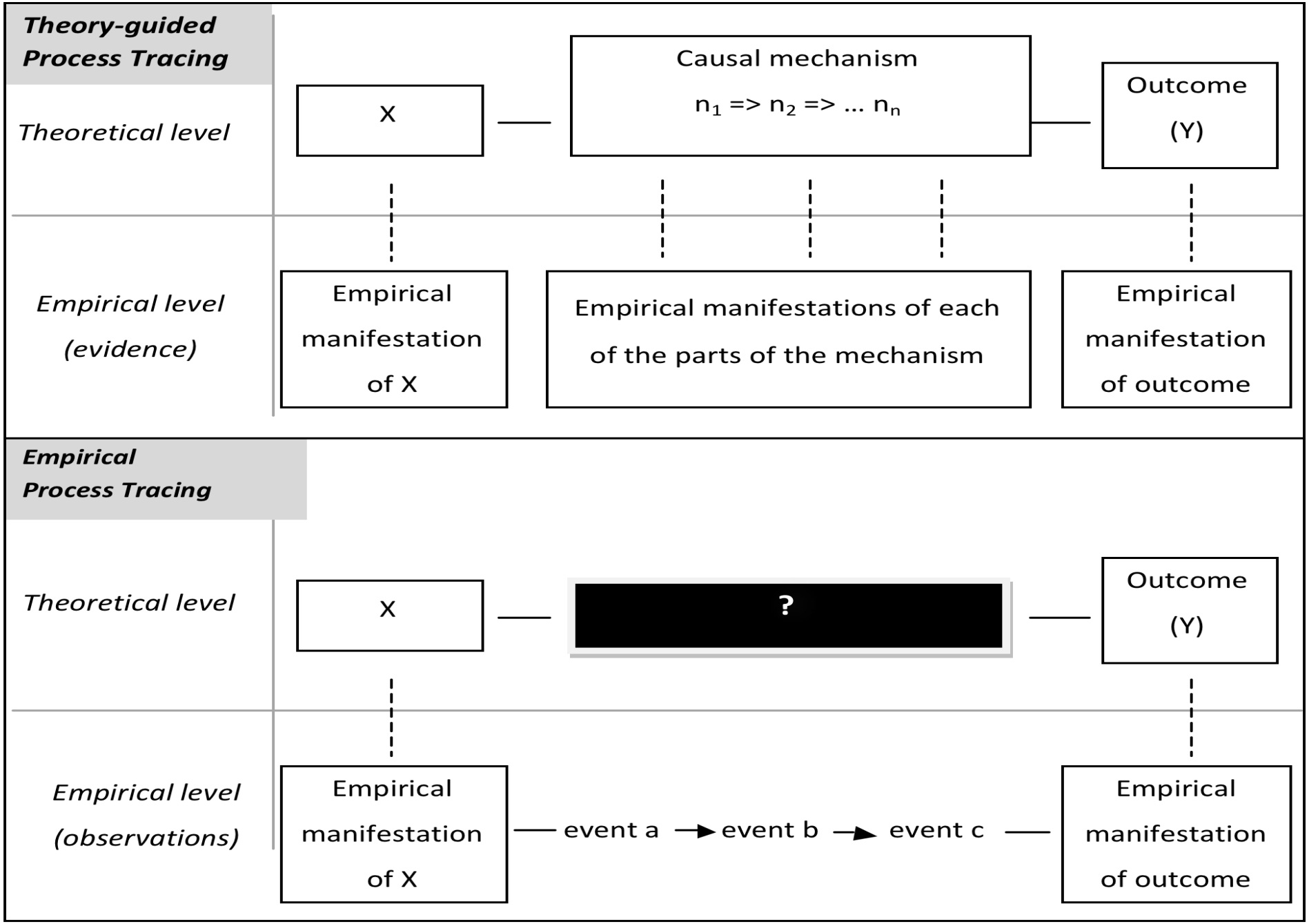
# 1. What is Process tracing?

Single case research method that can be used to make within-case inferences about presence/absence of causal mechanisms

# 1. What is Process tracing?

'the cause-effect link that connects independent variable and outcome is unwrapped and divided into smaller steps; then the investigator looks for observable evidence of each step.' (Van Evera 1997:64).

- focus is on **studying causal mechanisms using in-depth single case study**


Theory-guided Process Tracing
Theoretical level
X
Causal mechanism
$n_1$ => $n_2$ => ... $n_n$
Outcome (Y)
Empirical level (evidence)
Empirical manifestation of X
Empirical manifestations of each of the parts of the mechanism
Empirical manifestation of outcome
Empirical Process Tracing
Theoretical level
X
?
Outcome (Y)
Empirical level (observations)
Empirical manifestation of X
event a
event b
event c
Empirical manifestation of outcome

# 1. What is Process tracing?

KKV, Gerring – case study methods more analogous to medical experiment

- in perfect world measure effect of t and c on same unit ($U_t$ and $U_c$)

- analyze mean causal effects

PT – closer to criminal trial

- evidence assessed for each part of explanation (mechanism) to detect whether it can be concluded beyond reasonable doubt that mechanism existed

# 2. What are causal mechanisms?

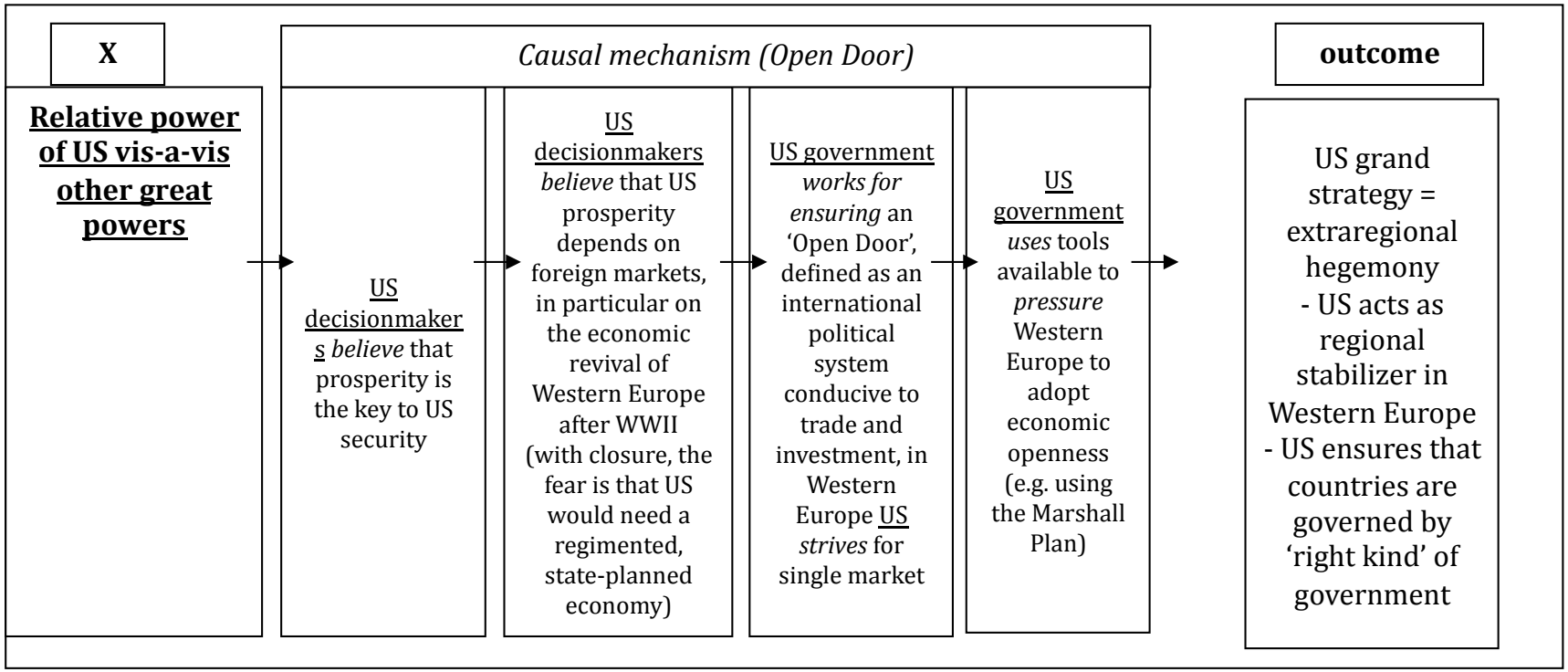
: a theory of a *system* of *interlocking* parts that *transmits causal forces* from X to Y

(Glennan, 1996, 2002; Bunge, 1997, 2004; Bhaskar, 1979).

**Table 5.1** – The five parts of Owen's causal mechanism whereby democracy produces peace

| *Part of the mechanism* | Conceptualization of mechanism and its parts (entities and activities) |
|---|---|
| *Context* | Crisis between states that can result in war |
| *Condition (X)* | Pair of states where analyzed state is democratic (liberal ideas), and where opponent is either democratic (liberal) or autocratic (illiberal) state |
| *Part 1 ($n_1$ →)* | Liberals will *trust* states they consider liberal and *mistrust* those they *consider* illiberal |
| *Part 2 ($n_2$ →)* | When liberals *observe* a foreign state becoming liberal democratic by their own standards, they will *expect* pacific relations with it |
| *Part 3 ($n_3$ →)* | Liberals will *claim* that fellow liberal democracies share their ends, and that illiberal states do not |
| *Part 4 ($n_4$ →)* | Liberals will *not change* their assessment of foreign states during crises unless those states change their institutions |
| *Part 5 ($n_5$ →)* | Liberal elites will *agitate* for their policies during war-threatening crises |
| *Outcome (Y)* | During crises, statesmen will be constrained to follow liberal elites, thereby not going to war with other liberal states. |

# Layne's case-specific Open Door mechanism



# 2. What are causal mechanisms?

Regularity understanding of causality

'…the *difference* between the *systematic* component of observations made when the explanatory variable takes one value and the systematic component of comparable observations when the explanatory variables takes on *another value*.'

(King, Keohane and Verba, 1994: 81-82, italics added).

# 2. What are causal mechanisms?

Mechanismic understanding of causality

-Open up 'black box' between X and Y

-the dynamic, interactive influence of causes upon outcomes, and in particular how *causal forces* are *transmitted* through a series of interlocking parts of a causal mechanism to *contribute* to produce an outcome.

# 2. What are causal mechanisms?

'… A mechanism is a set of interacting parts – an assembly of *elements producing an effect not inherent in any one of them*. A mechanism is not so much about 'nuts and bolts' as about 'cogs and wheels' – the wheelwork or agency by which an effect is produced.' (Hernes, 1998: 78, italics added)

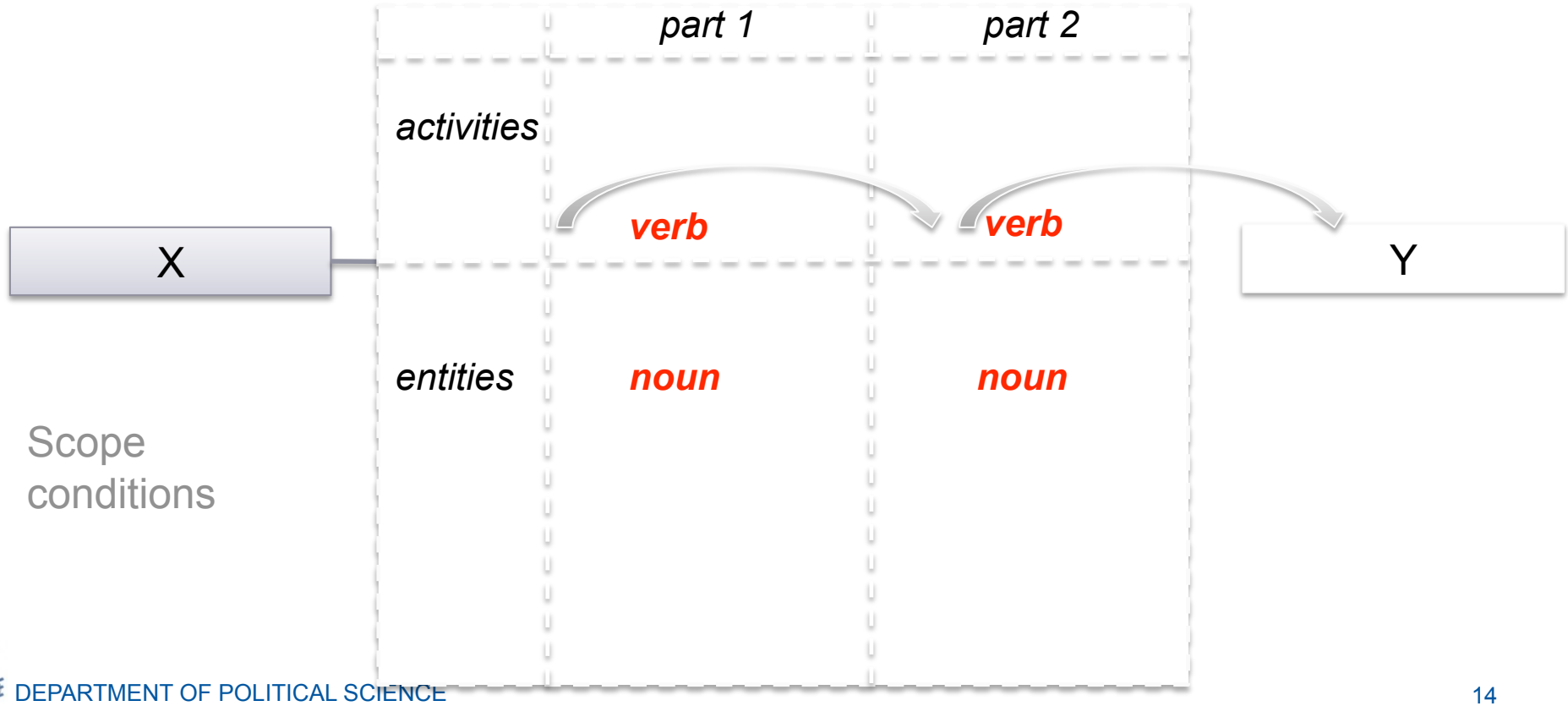
# 2. What are causal mechanisms?

Parts = factors that are *individually necessary* parts of mechanism, composed of entities that engage in activities (not intervening variables!)

*Entities* = object engaging in activities (noun)

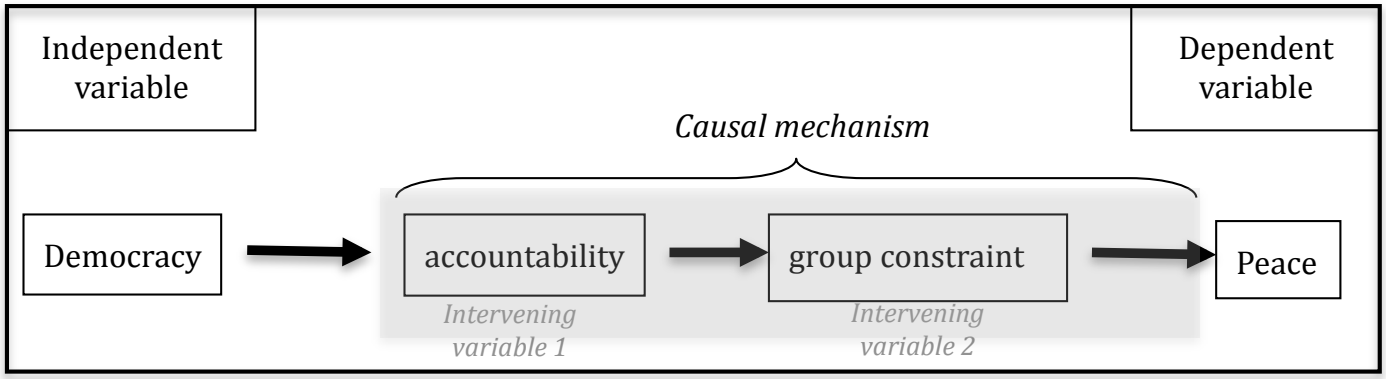
*Activities* = producers of change or what transmits causal forces through CM (verbs)

# 2. What are causal mechanisms?



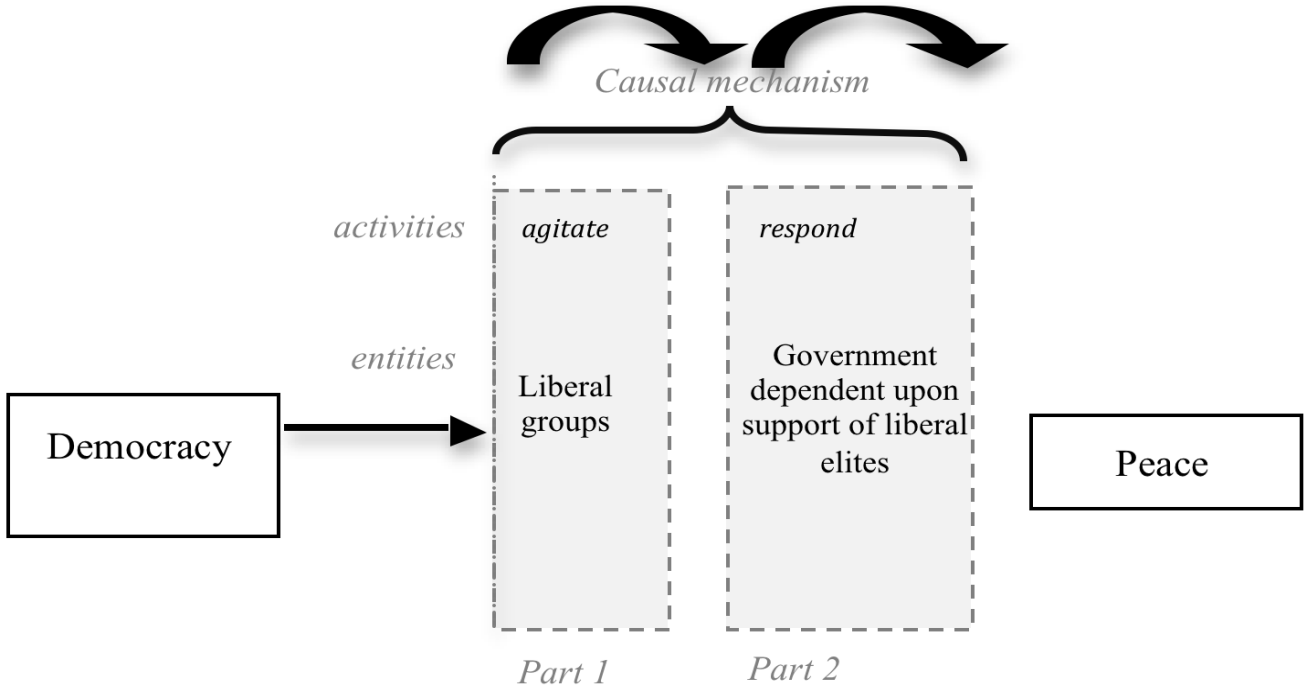
# 2. What are causal mechanisms?

- Mechanisms are NOT a series of intervening variables
  - (example from Rosato, 2003: 585)



# 2. What are causal mechanisms?

**Figure 2.3** – A democratic peace example of a causal mechanism



# Discussion

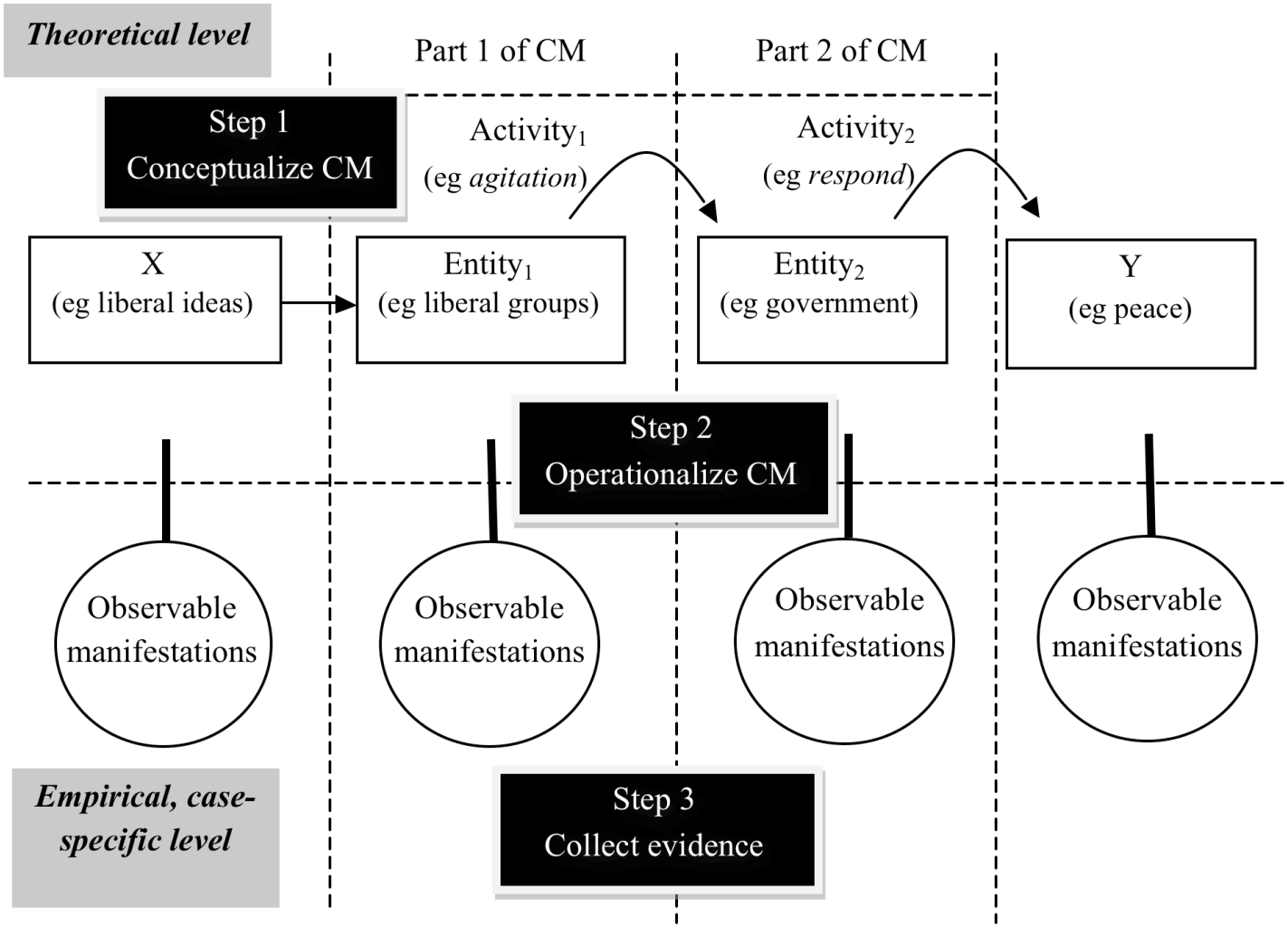
1. Develop a plausible causal mechanism that can explain why economic development (X) contributes to produce democratization (Y) through the creation of an educated middle class.

# 3. Three variants of Process Tracing

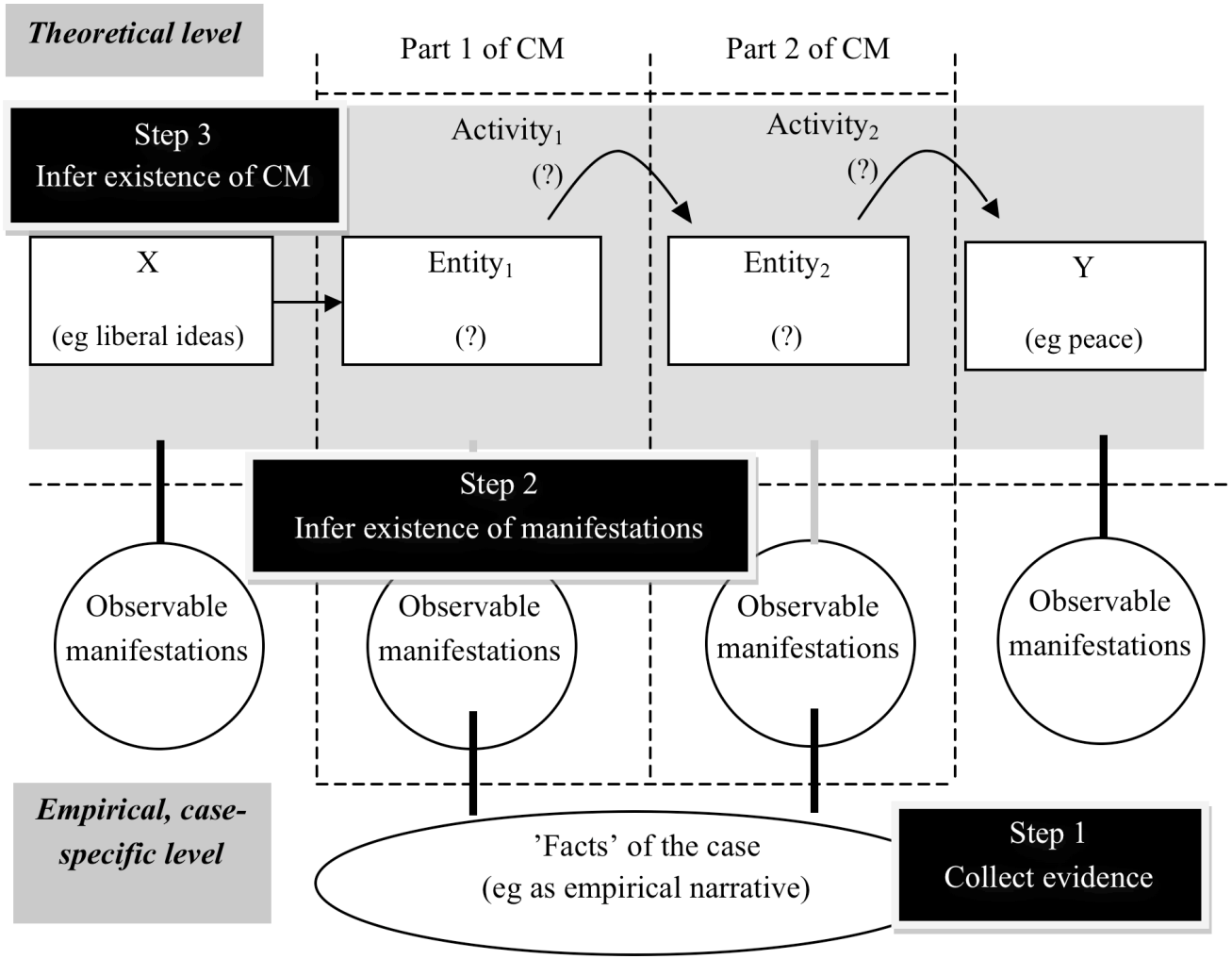
1. Theory-testing
2. Theory-building
3. Explaining outcome

| | Theory-testing | Theory-building | Explaining outcome |
|---|---|---|---|
| Purpose of analysis – research situation | <u>Situation one</u><br>- correlation has been found between X and Y, but is there evidence that there exists a causal mechanism linking X and Y? | <u>Situation two</u><br>Build a plausible causal mechanism linking X:Y based on evidence in case. | <u>Situation three</u><br>Explain particularly puzzling historical outcome by building minimally sufficient explanation in case study. |
| Ambitions of study | Theory-centric | Theory-centric | Case-centric |
| Understanding of causal mechanisms | Systematic (generalizable within context) | Systematic (generalizable within context) | Systematic, non-systematic (case-specific) mechanisms and case-specific conglomerates |
| What are we actually tracing? | Single, generalizable mechanism | Single, generalizable mechanism | Case-specific, composite mechanism that explains the case |
| Types of inferences made | 1) parts of causal mechanism present/absent<br>2) causal mechanism is present/absent in case | Observable manifestations reflect underlying mechanism | Minimal sufficiency of explanation |

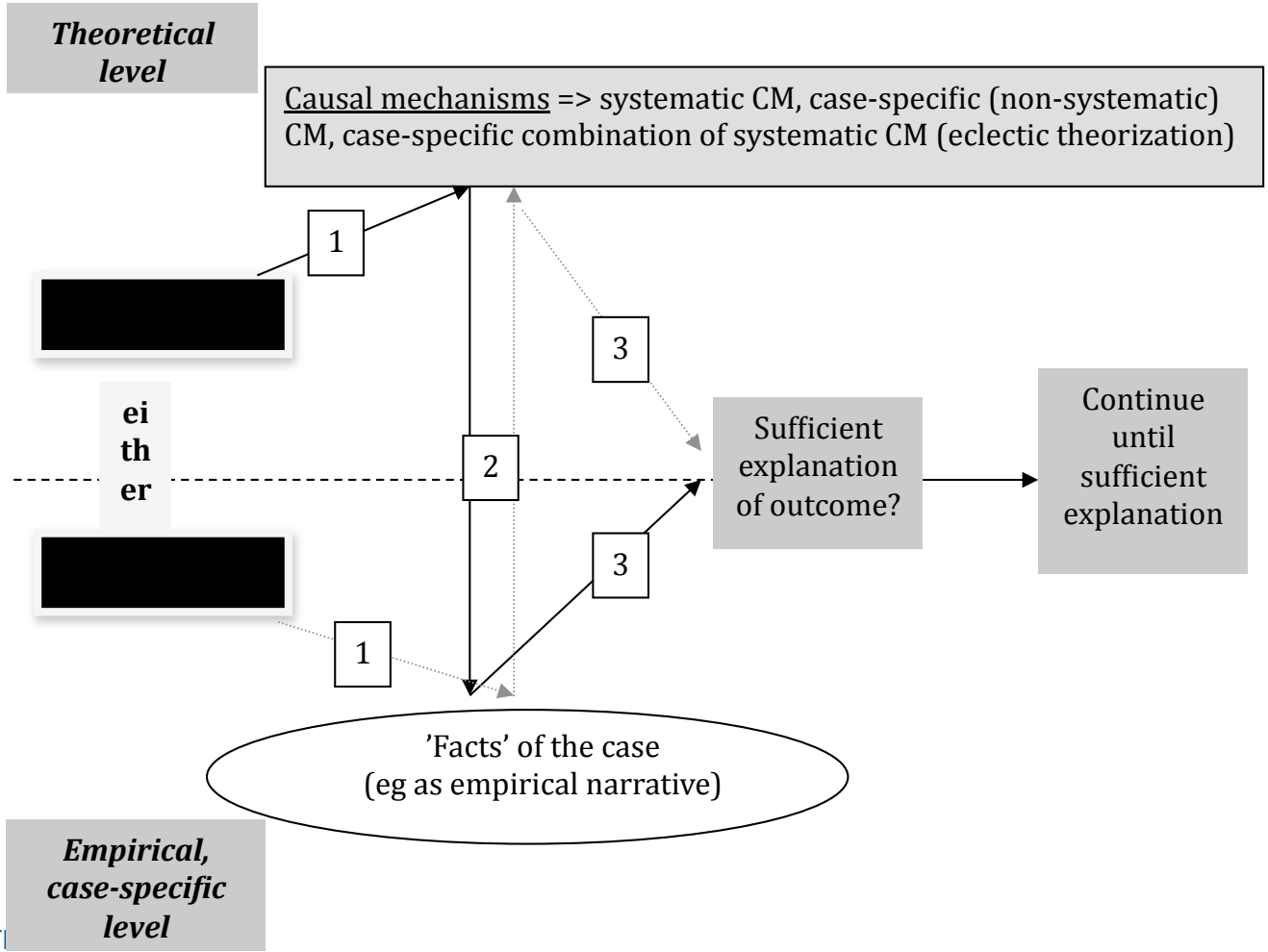
# Theory-testing



# Theory-building



# Explaining outcome

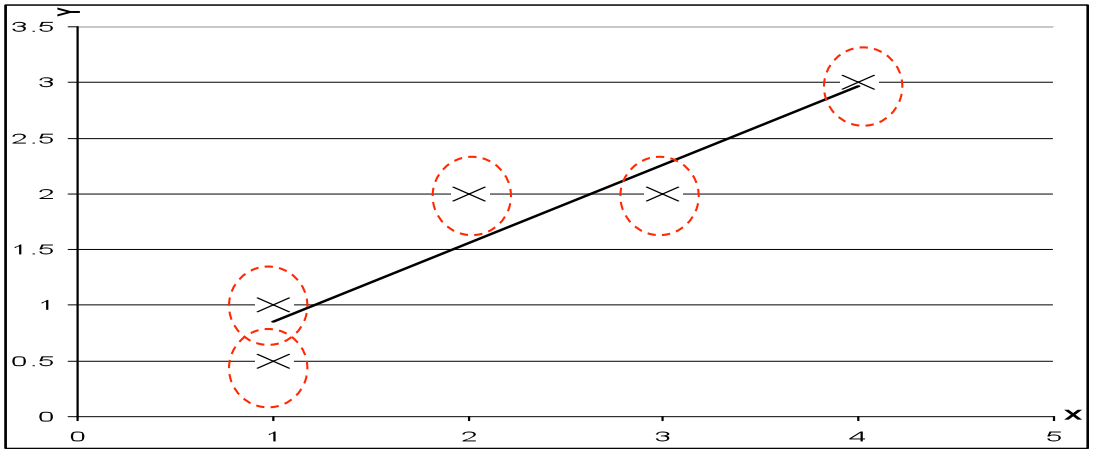


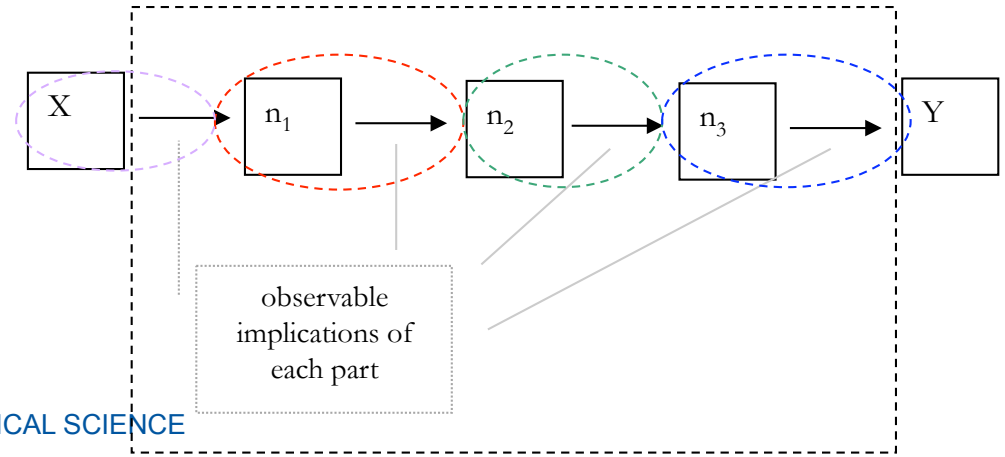
# 4. Causal inference in PT

- KKV, Gerring suggest that there is one logic of inference in all political science

'the differences between the quantitative and qualitative traditions are only stylistic and are methodologically and substantively unimportant. All good research can be understood – indeed, is best understood – to derive from the same underlying logic of inference.' (King, Keohane and Verba, 1994: 4).

| X | Y |
|---|---|
| 1 | 1 |
| 2 | 2 |
| 1 | 0 |
| 4 | 3 |
| 3 | 2 |



# 4. Causal inference in PT

| | **Frequentist logic in qualitative case study research (KKV)** | **Comparativist logic of elimination** | **Bayesian logic of subjective probability (process tracing)** |
|---|---|---|---|
| Ontological understanding of causality | Regularity and probabilistic | Regularity and deterministic | Mechanismic and deterministic |
| Inferences made using: | Classic probability theory and predicted probability that a found association is random or systematic | Mill's methods of agreement and difference and variants of them. | Bayes' theorem about the expected likelihood of finding specific evidence in light of prior knowledge |
| Types of causality assessed | Mean causal effect of X's upon Y | Necessary and/or sufficient conditions that result in Y | Presence/absence of causal mechanism (i.e. transmission of causal forces from X to produce Y) |
| Types of inferences made | Cross-case inferences (to population of phenomenon) | Cross-case inferences (but smaller scope population (contextualized)) | Within-case inferences |

# 4. Causal inference in PT

- Bayesian logic of inference = analyst gives greater weight to evidence that is expected a priori to be less probable based upon our previous knowledge of phenomenon.

- 'What is important is not the number of pieces of evidence within a case that fit one explanation or another, but the likelihood of finding certain evidence if a theory is true versus the likelihood of finding this evidence if the alternative explanation is true.' (Bennett 2006:341).

# 4. Causal inference in PT

Bayes' formula

posterior probability = prior probability x likelihood ratio

# 4. Causal inference in PT

posterior probability = the posterior probability of the degree of confidence we have in the validity of a hypothesis (h) about the existence of a part of a causal mechanism *after* collecting evidence (e).

$$p(h \mid e)$$

# 4. Causal inference in PT

*Prior* = degree of confidence that the researcher has in the validity of a hypothesis prior to gathering evidence, based upon existing theorization, empirical studies and other forms of expert knowledge.

$$p(h)$$

# 4. Causal inference in PT

*Likelihood ratio* = expected probability of finding evidence supporting a hypothesis based upon the researcher's interpretation of the probability of finding it in relation to the hypothesis and background knowledge informed by previous studies ($p(e \mid h)$, compared with the expected probability of finding the evidence if the hypothesis is not true ($p(e \mid \sim h)$.

$$p(e \mid \sim h)/p(e \mid h)$$

# 4. Causal inference in PT

## Bayes' formula

$$p(h|e) = \frac{p(h)}{p(h) + \frac{p(e|\sim h)}{p(e|h)} * p(\sim h)}$$

# 4. Causal inference in PT

Silver Blaze example – testing part of mechanism (whether horse abducted by insider)

- Prior = low (why would insider kidnap own horse!) = 20% (p(~h) = 80%)

- Likelihood of test = p(e|h) = 90%, p(e|~h) = 10%

$$0.692 = \frac{0.2}{0.2 + (0.1/0.9) * 0.8}$$

# 4. Causal inference in PT

What if 50-50 test?

- Prior = low = 20% (p(~h) = 80%)

- Likelihood of test = p(e|h) = 50%, p(e|~h) = 50%

$$p(h|e) = \frac{p(h)}{p(h) + \frac{p(e|\sim h)}{p(e|h)} * p(\sim h)}$$

# 4. Causal inference in PT

What if high confidence in prior?

- Prior = low = 70% (p(~h) = 30%)

- Likelihood of test = p(e|h) = 80%, p(e|~h) = 20%

$$p(h|e) = \frac{p(h)}{p(h) + \frac{p(e|\sim h)}{p(e|h)} * p(\sim h)}$$

# 5. Studying causal mechanisms

-develop strong empirical tests for whether all parts of causal mechanism are present or not

- logic of empirical testing in process tracing => if we expected X to cause Y, each part of the mechanism between X and Y should leave the predicted empirical manifestations which can be observed in the empirical material.

# 5. Studying causal mechanisms

- Detecting these manifestations => development of carefully formulated predictions of what evidence we should expect to see if the hypothesized part of the mechanism exists

- Predictions translate theoretical concepts of the causal mechanism into *case-specific* observable manifestations (expected evidence).

# 5. Studying causal mechanisms

Empirical predictions - 4 different types of evidence

1. *Pattern evidence* = statistical patterns in the evidence.
2. *Sequence evidence* = temporal and spatial chronology of events
3. *Trace evidence* = mere existence provides proof
4. *Account evidence* = content of empirical material

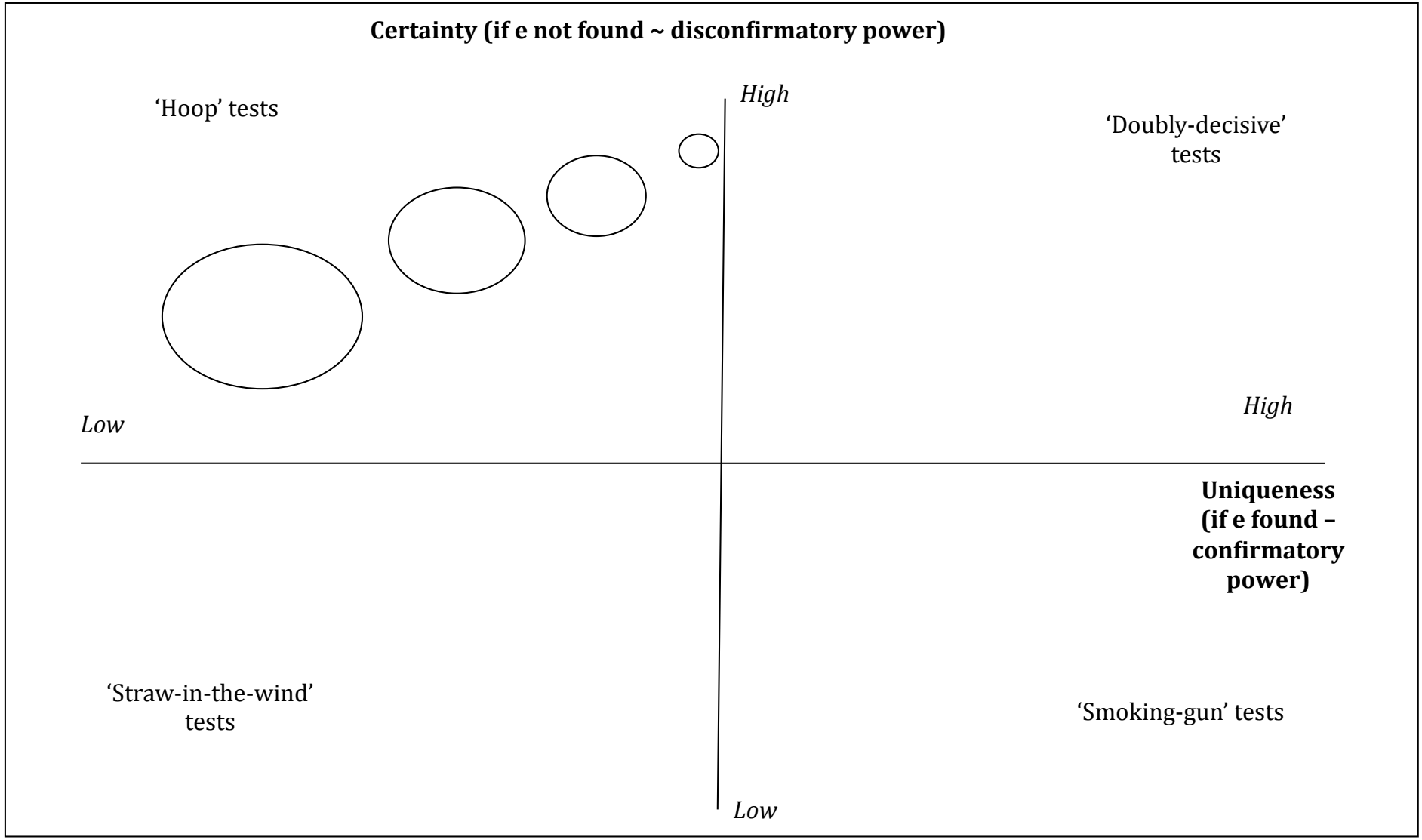
# 5. Studying causal mechanisms

- *unique predictions* => empirical predictions that do *not* overlap with those of other theories => confirmatory power if e found

- Uniqueness corresponds to the likelihood ratio, where predictions are developed that maximize the value of p(e|h) in relation to p(e|~h).

# 5. Studying causal mechanisms

- *certain prediction* => prediction is unequivocal and the prediction (e) must be observed or else the theory fails the empirical test => disconfirmatory power if e not found


Certainty (if e not found ~ disconfirmatory power)
'Hoop' tests
High
'Doubly-decisive' tests
Low
High
Uniqueness (if e found – confirmatory power)
'Straw-in-the-wind' tests
'Smoking-gun' tests
Low

# 5. Studying causal mechanisms

*straw-in-the-wind* test = empirical predictions that have a low level of uniqueness and a low level of certainty (low confirmatory and disconfirmatory power)

- do little to update our confidence in a hypothesis irrespective of whether we find e or ~e, as both passed and failed tests are of little if any inferential relevance for us.

# 5. Studying causal mechanisms

*Hoop tests* = predictions that are certain but not unique (low confirmatory and high disconfirmatory power)

- failure of test (finding ~e) reduces our confidence in the hypothesis but finding e does *not* enable updating.

# 5. Studying causal mechanisms

*Smoking gun tests* = highly unique but have low or no certainty in their predictions (high confirmatory and low disconfirmatory power)

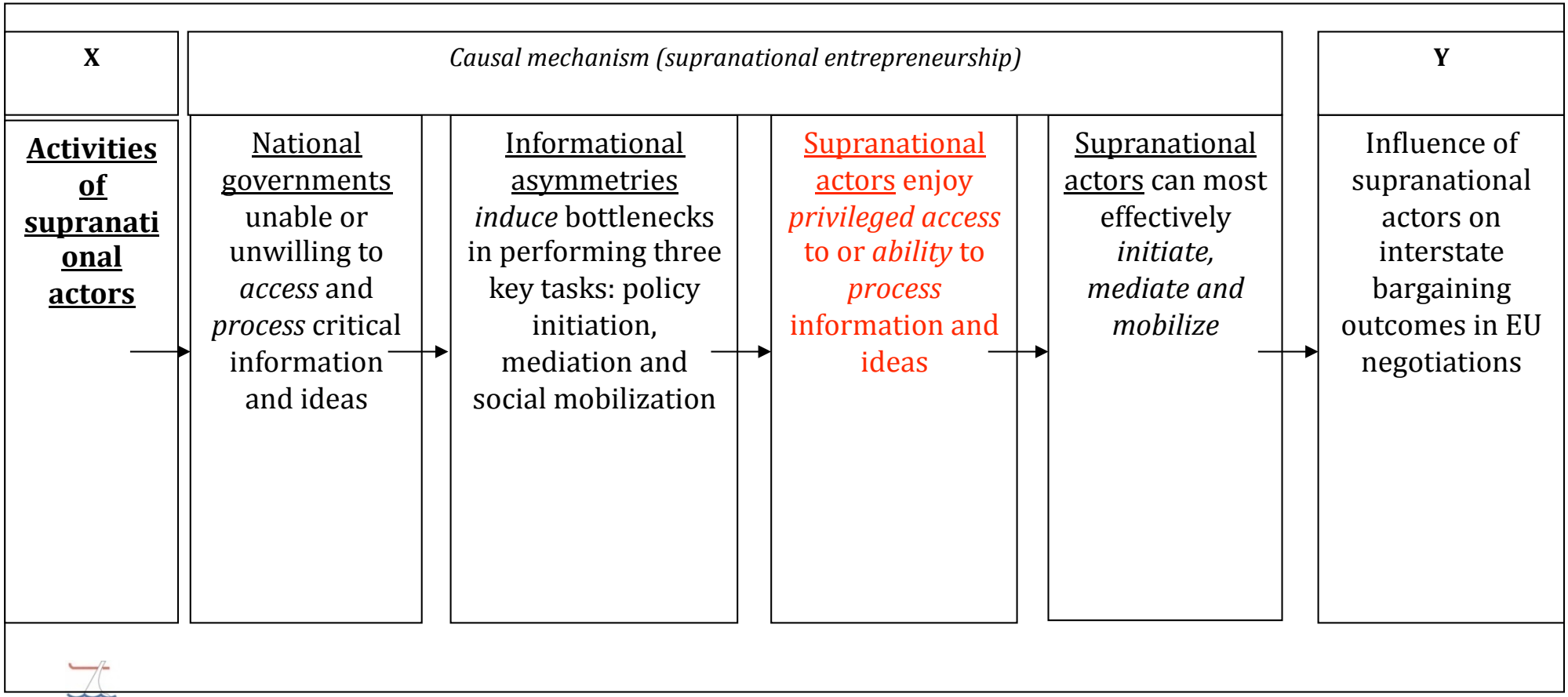
- Likelihood ratio is small (finding e given h highly probable whereas ~h is highly improbable), thereby greatly increasing our confidence in the validity of h if we find e. If not find e => no updating.

# 5. Studying causal mechanisms

*Doubly decisive tests* => both certainty and unique (high confirmatory and disconfirmatory power)

- evidence has to be found or our confidence in the validity of the hypothesis is reduced (updating when ~e)
- at the same time the test is able to discriminate strongly between evidence that supports the hypothesis and alternatives (small likelihood ratio), enabling updating when we find e.

# 5. Studying causal mechanisms



# 5. Studying causal mechanisms

Moravcsik example : test of 'the Commission has privileged access to information'.

- straw in the wind = 'expect to see that the Commission has many civil servants'
- stronger test = 'expect to see that the Commission in the most sensitive areas of negotiations was much better informed about the content and state-of-play of the negotiations than governments, *possessing more detailed substantive issue briefs and more accurate and updated information on the state of play*'

# Discussion

1. Operationalize an empirical test drawn from your own research, describing the uniqueness and certainty.

# 6. The uses of PT

| Testing for the necessity of the *parts* of a causal mechanism in single case | Testing for the necessity of a mechanism as a *whole* at the population-level | Testing for the sufficiency of a mechanism in single case | Testing for the sufficiency of a condition at the population-level |
|---|---|---|---|
| Process tracing (theory-testing or explaining outcome) | Comparative cross-case methods (e.g. positive on outcome design – all Y's) | Explaining outcome process tracing | Comparative cross-case methods (e.g. positive on cause design, all X's) |

# 6. The uses of PT – nesting?

- *Systematic* factors only in cross-case (Rohlfing)
- Deterministic theory
  - LNA when traditional statistical analysis = probabilistic (mean causal effects across population)
  - SNA (PT) = deterministic ontology

- Divorcing X from X + CM
  - can X be meaningfully divorced from CM if we PT studies are to communicate with other methods?
  - Are we studying two different things : LNA = **X:Y** / PT = X+**CM**=>Y
  - One solution = use configurational theories
    - Fx X = liberal ideas or X1 (liberal ideas) + X2 (liberal groups) + X3 (responsive gov)

# 6. The uses of PT – nesting?

- Explaining outcome PT *cannot* be nested for several reasons:

1.Use of non-systematic factors in accounting for Y (minimal sufficiency)

2.Eclectic, non-systematic (case-specifiic) combination of theories, with theories used in pragmatic fashion as heuristic tools to account for outcome (more idiographic focus)

** deeply interested in the case

** however EO PT can have some exportable findings – 'lessons'

# 6. The uses of PT – nesting?

- Theory-testing studies can be nested in two situations

  1) have strong X:Y correlation from LNA research

  - Does X cause Y in manner predicted by theory? (Owen)
  - Is there a causal relationship, or is it spurious?

  2) well-developed theory but is there empirical support (when small scope of N)

  ** problem with probabilistic / deterministic theorization + what we are studying…

# 6. The uses of PT – nesting?

- Theory-building studies can be nested in two situations

  1) have strong X:Y correlation from prior research but no idea how X caused Y

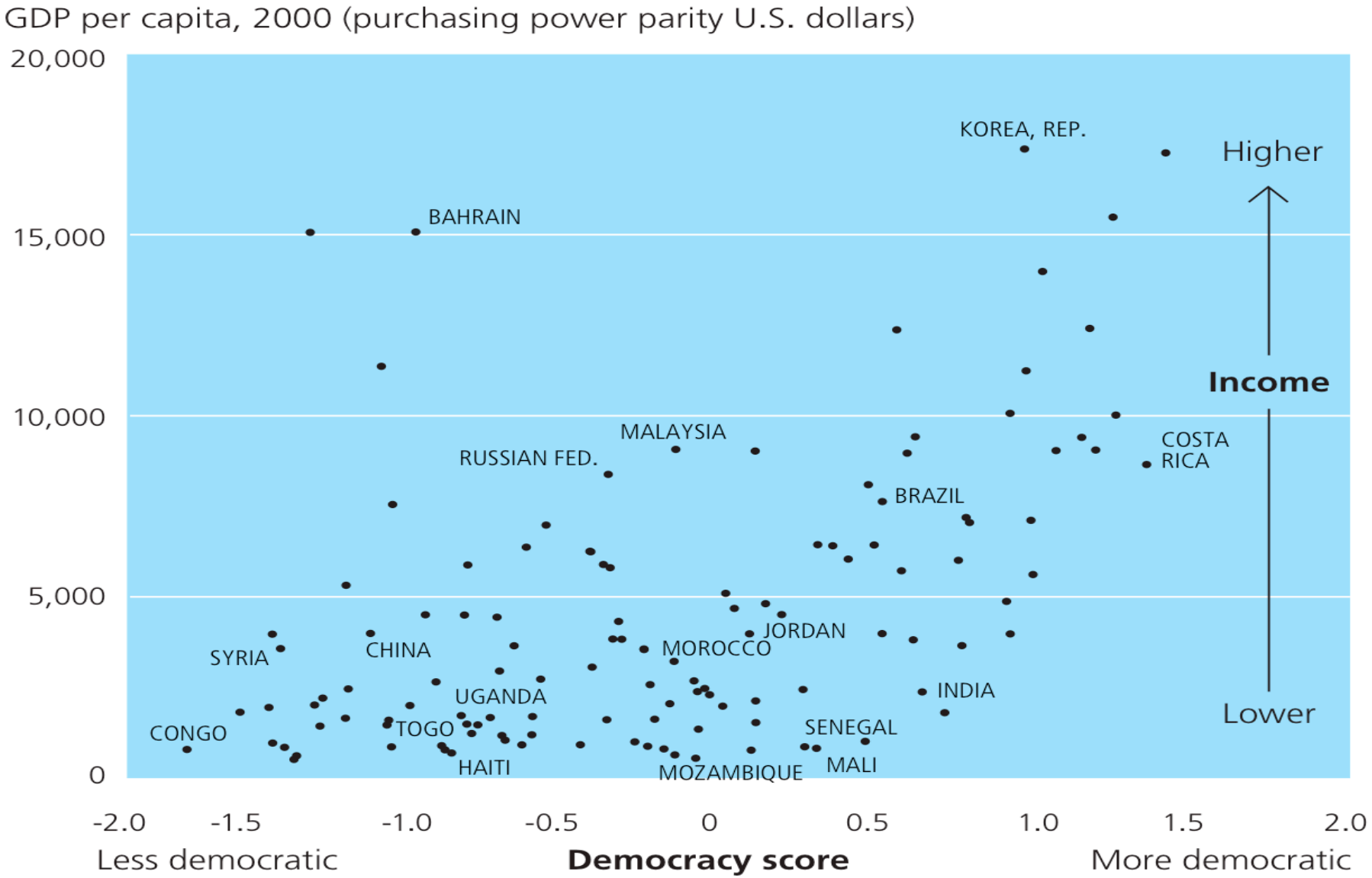
  2) Know Y but unclear about what caused it (what is X?)

  ** challenge of identifying non-systematic factors in single case study

# What cases are relevant for:

- Theory-testing of economic development -> democracy

- Theory-building explaining why low income countries can become democratic




*Note:* Democracy score is the voice and accountability indicator from World Bank 2001c.
*Source:* World Bank 2001c, 2002e.

| | **Theory-testing** | **Theory-building** | **Explaining outcome** |
|---|---|---|---|
| Case chosen because ambition is to … | Test the necessity of the parts in the causal mechanism in empirical test | Theorize a plausible causal mechanism based upon the empirical evidence | Prove minimal sufficiency of causal mechanism (or set of mechanisms) in a single important case |
| Uses of variant of process tracing in a broader mixed-method design | 1) an X:Y correlation has been found but we are unsure of causality<br>2) a well-developed theory exists but we are unsure whether there is empirical support | 1) an X:Y correlation has been found but we are unsure of the mechanism whereby X produces Y<br>2) we are unable to explain what caused Y with existing theories, resulting in the building of a new theorized mechanism that can account for the deviant case | Not possible due to the inclusion of non-systematic parts, although limited 'lessons' can be drawn about potential systematic parts that merit further research in other cases |

| | **Theory-testing** | **Theory-building** | **Explaining outcome** |
|---|---|---|---|
| Case chosen because ambition is to … | Test the necessity of the parts in the causal mechanism in empirical test | Theorize a plausible causal mechanism based upon the empirical evidence | Prove minimal sufficiency of causal mechanism (or set of mechanisms) in a single important case |
| Uses of variant of process tracing in a broader mixed-method design | 1) an X:Y correlation has been found but we are unsure of causality<br>2) a well-developed theory exists but we are unsure whether there is empirical support | 1) an X:Y correlation has been found but we are unsure of the mechanism whereby X produces Y<br>2) we are unable to explain what caused Y with existing theories, resulting in the building of a new theorized mechanism that can account for the deviant case | Not possible due to the inclusion of non-systematic parts, although limited 'lessons' can be drawn about potential systematic parts that merit further research in other cases |

# 6. The uses of PT

- Strong *within-case* inferences can be made using in-depth single case study
- No *cross-case* inferences can be made with PT
- Whether PT can be used in conjunction with other methods depends upon the variant of PT (yes for theory-testing and building, no for explaining outcome)