Exercises for:

Large-N Qualitative Analysis (LNQA): causal generalization in case study and multimethod research
(Princeton 2027)

and

Social science concepts and measurement: new and completely revised edition
(Princeton 2020)

and

Multimethod research, causal mechanisms, and case studies: an integrated approach
(Princeton 2017)

and

A tale of two cultures: qualitative and quantitative research in the social sciences
(Princeton 2012)

and

Politics, gender, and concepts: theory and methodology
(Cambridge 2008)

and

Explaining war and peace: case studies and necessary condition counterfactuals
(Routledge 2007)

Version 9 (2026)

Gary Goertz
ggoertz@nd.edu

Introduction

This represents the ninth edition of my collected exercises. As is my custom, they often represent my new and ongoing research projects. There are about 40 new exercises in this edition. As always these are listed in the first section and then also appear in the relevant substantive section(s).

The main news is Haggard and my Large-N Qualitative Analysis, LNQA for short, book is in production at Princeton University Press and should be out summer 2027. An introduction to the methodology appeared in *Perspectives on Politics* in 2023. Attached to this email is the version that went to Princeton. Working on the book has generated quite a few new exercises last year and this year about 2/3 of the exercises are related to LNQA. The LNQA project is in fact a continuation of the discussions on causal mechanisms, generalization, case studies and causal complexity initiated in the (2017) multimethod book.

LNQA is a research design and causal inference approach that focuses in particular on complex theories and causal mechanisms where causal inference is done within the case as opposed to across cases. Much qualitative and other work involves quite complex theories and so an important part of LNQA is how to portray, and notably diagram, complex causal models-mechanisms. LNQA causal inference involves process tracing these mechanisms within the cases. This might be called the LNQA approach to process tracing. LNQA is very concerned with generalization so the scope of the causal mechanism is critical and is directly related to case selection. Many of the new exercises deal with diagramming complex causal models-mechanisms and/or expressing them in Boolean terms and can be found in the causal models-mechanism and/or Boolean models sections. Many of the exercise deal with doing causal model-mechanism figures which are central to the LNQA approach. In addition, there are some exercises regarding Waldner's (2026) work on "qualitative causal inference" and its connection to within-case causal inference in LNQA.

Since LNQA is fundamentally about causal *generalizations*, issues like the complex relationship between scope, moderators, and external validity are explored in some exercises. How to diagram external validity in causal figures is an important part of the book and appears in exercises.

Again there are exercises dealing with two-way tables; a section is devoted to this topic. As part of the LNQA project we have surveyed a large number of qualitative books published by Princeton and Cornell. It will come as no surprise that many of them contain two-way tables which express the core theory. Working through them it became clear that there are a variety of methodological issues that are not discussed at all because the methodology of two-way tables apparently seems easy and straightforward. This is definitely not the case and some of the exercises point to issues that must faced by all constructing two-way tables. Similarly, a recurrent question is extent to which these two-way tables as well as figures might be expressed as Boolean equations.

As always, there are exercises on concepts and measurement as well as scaling, notably dealing with V-Dem–related datasets and how indicators are conceptualized and coded, typically indicating confusion regarding scale types.

Note that an exercise can appear in multiple sections if it is relevant to more than one topic. The new exercises appear at the top of sections to which they belong, so one can see what is new in any given category.

Brief answers – usually a sketch of an answer – to these exercises (except those that are open-ended and those that I think would be interesting but have not yet had the time to do) are included as well.

If you would like to be informed when these exercises are updated please contact me and I will put you on the email list (ggoertz@nd.edu).

Most of the articles and book chapters referred to in the following exercises are available in PDF format and complete references are found at the end. If possible I have chosen an electronically available article rather than a book.

As always, I welcome comments on these exercises and suggestions for new ones.

Table of contents

New exercises since version 8, also included in relevant sections below, p. 4

Large-N Qualitative Analysis (LNQA), p. 30

Generalization, external validity, scope, etc., p. 45

Process tracing, within case causal inference, counterfactuals, p. 55

Causal issues, causal hypotheses, causal asymmetry, etc., p. 75

Causal models-mechanisms, causal complexity, causal figures, methodology and philosophy of, p. 95

Constraint mechanisms, necessary conditions, norms, permissive causes, p. 131

Concepts, general issues of concept construction, guidelines, and specific concepts. p. 133

Semantic transformations, calibration, p. 183

Terminology, definitions, and semantics of concepts , p. 189

Scaling issues, scale types, p. 192

Concept structure and aggregation, p. 211

Theories and concepts, e.g., two-level theories, p. 233

Latent variables, statistical models, p. 244

Normative issues in concept construction, p. 249

Global indices, barometers, etc., p. 250

Gender and concepts, research design, p. 252

Typologies , p. 260

Multimethod research, p. 267

Case studies, case selection, methodology of, p. 275

The Possibility Principle, p. 284

Two-way tables, construction and analysis, Boolean and linear modeling of, p. 301

Boolean theories, QCA, fuzzy logic, p. 311

New exercises

New exercises also appear in the relevant sections below and can appear more than once if they are relevant to multiple sections.

1. For Waldner (2026) a causal figure consists of variables and is intended to apply to multiple cases. To this correspondence what he calls an "event history map" which is how these variables play out in a given case, in other words the values of the variables for that particular case. This is very similar to the LNQA approach to within-case causal inference. A major difference however is that in LNQA the arrows in the event history map are causal claims about what happened in an individual case. In contrast, for Waldner *these arrows are not causal* (see the next exercise below for more on this). This constitutes a major difference between the Waldner approach and LNQA: in LNQA the causal model is applied directly to the case and the arrows are understood to mean causation in the individual case. Contrast Waldner's position with LNQA on this core topic of inference. For example, in some cases there might be a backdoor path and in other cases there is no evidence for a backdoor path (the same could be true for side door paths).

   > The specific meaning of the nodes in an event-history map implies a specific meaning of the arrows as well. Because these nodes represent realized events rather than random variables, the map itself does not represent a joint probability distribution. The arrows connecting the nodes, therefore, do not represent relationships of probabilistic dependence or causal relationships. Event-history maps constitute the methodologically constrained narrative of the case study, constrained by the conjectured data-generating process represented by the causal graph. It is thus critical to avoid the temptation to interpret an event-history map as representing causal relations. . . . Once this mapping is complete, the causal graph that satisfies our criteria allows us to give a causal interpretation of the event-history map, supporting the claim that each set of events comprising a node in the event-history map generates the set of events comprising the next node via the intentional action comprising the invariant causal mechanism. (Waldner 2026, 132)

2. How to think about and how to diagram evidence in causal model figures is still very much a topic of discussion among process tracing scholars. Waldner (2026) gives an interesting example in the figure below. The lowercase letters indicate some hoop tests that give evidence about whether X is a cause of Y. Notice that he is concerned about the reader being confused between the evidence which are individual events about the case distinguished from the causal factors indicated by the capital letters (see the exercise above on this point). So the question then becomes how is one to interpret the (causal) arrows going to and from these events. This can be contrasted with the smoking gun figure discussed in the exercise below. Describe how LNQA could diagram these three hoop tests as evidence about a causal relationship between X and Y.

   > Consider a hypothesis that X is connected to Y by the intervening events X → a → b → c → Y, where X and Y are the variables and a, b, and c

> are intervening events rather than variables. If careful research does not uncover evidence that a, b, and c were present, the hypothesis has failed three hoop tests and can be rejected. But if each hypothesis passes its hoop test, the intervening events a, b, and c were present. This generates a sequence of events, but as we have repeatedly affirmed, concatenation is not causation. . . . Figure 7.5 is not a causal graph because it mixes variables and events. (Waldner 2026, 262)

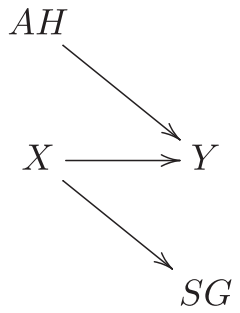
Figure 1: Hoop tests à la Waldner

$$X \longrightarrow a \longrightarrow b \longrightarrow c \longrightarrow Y$$

3. The interpretation of a smoking gun test varies quite significantly among authors. Here I give Waldner's interpretation of it in his figure 2 below. Noticed that in contrast to the LNQA interpretation of evidence, here the evidence consists of the causal relationship between X and SG: this causal relationship is evidence that X was the cause of Y. How would one graph this from the LNQA perspective? Also discuss the importance of *unique* evidence that we have highlighted in the Waldner quote. Why is that critical for this being smoking gun evidence? Would one draw the same graph for a hoop test? Contrast this with how Blair et al. do smoking gun, discussed in chapter 5, appendix B of the LNQA book.

> Figure 1.3 also contains a node denoted SG. This node is not on the front-door path between X and Y; it is what James Mahoney calls an "auxiliary outcome" or "particular occurrences that should occur alongside (or perhaps as a result of) the main outcome of interest if in fact the outcome was caused in the way stipulated by the theory under investigation." This node represents what process-tracing methods call "smoking-gun" evidence. Such evidence conforms to a unique prediction made by only one hypothesis; hence, the graph contains an arrow from X to SG but not from AH [auxiliary hypothesis, e.g., alternative hypothesis] to SG. The hypothesis denoted X thus passes a smoking-gun test and is strongly corroborated because it *uniquely* predicts this evidence. (Waldner 2026, 38, emphasis is mine, see also his discussion around Figure 7.6).

Figure 2: Smoking gun tests à la Waldner



4. Central to Waldner's account of what he calls "qualitative cause inference" is the notion of a "hypothetical intervention." This is an intervention on the key treatment variable in the model or theory. So for example, a hypothetical intervention could set X to a value different from that observed in individual case i. In contrast, he has nothing to say about counterfactuals in the book. Discuss the similarities and differences between his notion of hypothetical intervention and the methodology of counterfactual analysis. Are the criteria – for example, the three described below – for good hypothetical interventions the same or consistent with the criteria for good counterfactual analysis?

> For any hypothetical intervention, we should be more willing to accept the claim if the reasons adduced on its behalf exhibit theoretical plausibility, mechanistic credibility, and the associated descriptive inferences are each deemed valid. If these elements are missing or only weakly established, we should be correspondingly less willing to accept the claim about the results of the hypothetical intervention. In the ideal scenario, a claim about the consequences of a hypothetical intervention has all three qualities: theoretical plausibility, mechanistic credibility, and descriptively valid inferences. Conversely, we should have serious doubts about such interventions when they lack grounding in general theories; when they lack fine-grained mechanistic accounts of the relevant actors, their intentions, their actions, and the consequences of their actions; and when the best available evidence refutes or only weakly supports the derived hypotheses. (Waldner 2026, 216–17)

5. Critical to LNQA is the connection between the causal model and Boolean interpretations of the model. Madrid (2025) provides a very interesting example of a study with close connections to LNQA. He basically looks at democracy-authoritarian outcomes in South America in the 19th century ending in 1929. While the book has democracy in its title most of the countries in South America are not democratic in the 19th or early 20th century. However, he does focus particular attention on the four countries that did achieved democracy by the interwar period. As one can see from the causal model figure below, figure 3, he is not only explaining democracy but also authoritarianism.

Figure 3: Explaining democracy-authoritarianism in 19th century South America

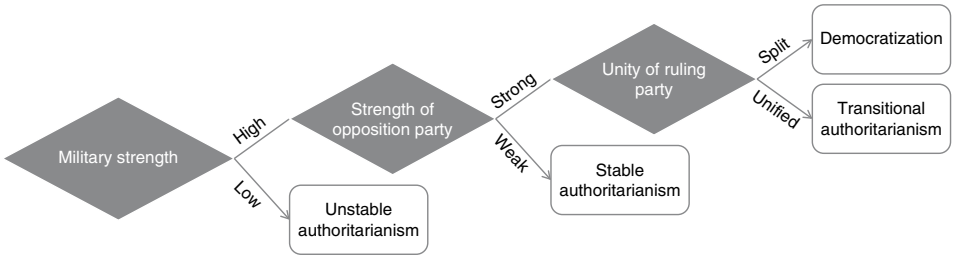


FIGURE 1.1 Determinants of regime outcomes in South America before 1930

The figure has his three main causal factors and he says this:

> [R]egime outcomes in South America during the nineteenth and early twentieth century were shaped by three main variables: the strength of the military, the strength of opposition parties, and the unity of ruling parties (see Figure 1.1). These three variables did not change in a strict sequence, nor were they closely causally related to each other. The strengthening and professionalization of the military did not cause strong opposition parties to emerge, nor did it necessarily precede the rise of strong opposition parties. (Madrid 2025, 38)

There are a number of interesting questions raised by his causal model. For example, would the figure change significantly if the factors were put in a different (temporal) order? Can one give a Boolean interpretation of this causal model, for example, one Boolean model per outcome? For example, to what extent are the three factors sufficient for the outcome of democracy? Are they also necessary for the outcome? The same can be asked of all of the outcomes illustrated in the figure.

6. In most causal model figures, for example DAGs, no specific reference is made to causal claims in the negative: that X is not a cause of Y. However, in the DAG world these noncausal claims are absolutely critical in various contexts. A good example is instrumental variables. To have a valid instrumental variable implies important, and often somewhat implausible causal claims. How might one diagram those in a DAG-like fashion?

   Answer: Figure 4 is how Mellon (2025) nicely does this making very clear via the dashed arrows what causal relationships cannot hold for a valid instrumental analysis, W is the instrument.

Figure 4: Instrumental variables and noncausal relationships

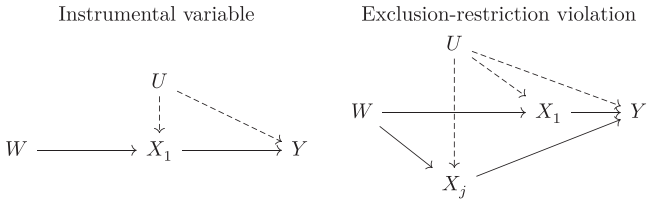


FIGURE 1 Directed Acylic Graphs (DAGs) illustrating instrumental variables (IVs) and endogeneity. *Note*: The figure illustrates the causal structure assumed by IV estimates and how exclusion-restriction violations undermine them.

7. As discussed in chapter 3 of the concept book any concept that involves two words is potentially problematic because one in effect has often two concepts each of which is worthy of its own analysis and then the question about how to combine the two into one hyphenated concept. Helmke and Rath (2025) Illustrate this issue in their discussion of "democratic norms." As illustrated below they discussed at length the concept of a norm from a rational choice perspective as an equilibrium. But they do not discuss is how the adjective democratic modifies a norm. Discuss how this might be done. How does one use a concept of democracy in the process? Does one add definitional features of democracy to those defining norms?

> Specifically, we define political norms as informal behavioral rules that an individual follows in a given political circumstance only if (a) they expect others to follow the rule and (b) they believe others believe they ought to follow the rule. The first condition commonly refers to empirical expectations and suffices to establish purely descriptive norms. The second condition refers to normative expectations and is necessary for creating injunctive or social norms. . . . For any norm to exist, a sufficient number of people must also recognize that a given circumstance (i.e., a particular event in a certain context) calls for a given behavioral response, or what Bicchieri (2006, pp. 11–13) refers to as the "contingency condition." Just as crucial, we argue, is whether a sufficient number of people also believe that the circumstance itself exists. . . . Having roughly pinned down what we mean by norms, what makes a given norm democratic? (Helmke and Rath 2025, 236–37)

8. There is quite a bit of variation and how scholars have thought about moderator, mediator variables and "context" in the discussion of external validity. And their article Geng et al. (2025) take the DAG approach to thinking about these issues. Based on the moderator section of the LNQA book (chapter 4) discuss how they have diagrammed context, see figure 5, how they have described it, and it might be conceived of as a moderator. See in particular the discussion associated with the figure.

Figure 5: Context, moderators and external validity

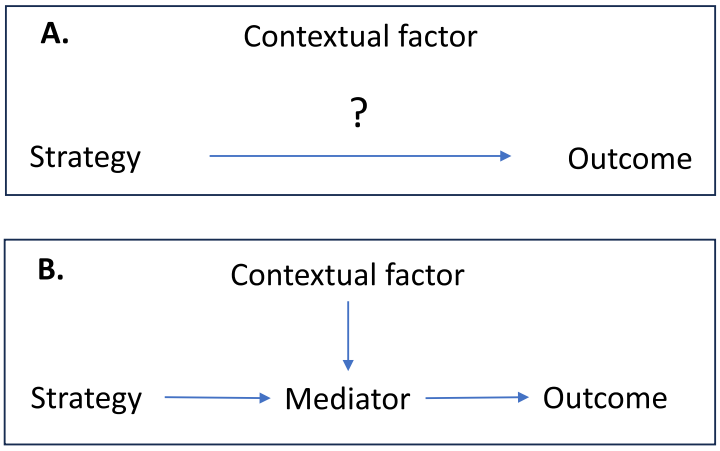


Fig. 3 Mediators enables consideration of context. Absence of a hypothesized mechanism limits analysis of context (Panel **A**). Proposed mechanisms — in this case a single mediator for simplicity — offers a conceptual receptacle for how context might influence a strategy's effects (Panel **B**). In the convention of directed acyclic diagrams, two arrows that point into one node implies that the effect of each arrow is influenced by the other, in this case representing how context effects the effect of the strategy on the outcome

Answer: The figure is fundamentally ambiguous because it is not clear what the arrow to the mediator means. For example it could mean that contextual factor is a cause of the mediator. So context is just another causal factor in the model. In the LNQA approach context does not influence the mediator itself but its causal relationship to the outcome, i.e., β. In general, moderator variables require significant interpretation when presented in DAGs. The causal of model cannot just be read from the DAG, it must be interpreted via the text discussing the DAG and eventually mathematical equations.

9. Strong regularities such as those found in LNQA analysis play an important role in applications of the Possibility Principle. This is because these regularities are used to infer what cannot happen. This "cannot happen" is based on the empirical regularity. Discuss how this is used in the following example:

> To study dispute onset, it is necessary to first identify the set of pairs of states that have the possibility of entering into a maritime boundary dispute. . . . In practice, at least as far as delimited maritime boundaries are concerned, states do not go beyond the boundary lines they have already agreed to and initiate new conflicting claims. Indeed, I could record no instance of a new dispute onset after full delimitation. For these reasons, the analyses exclude dyad-year observations after full delimitation. (Yüksel 2025, 1181)

Answer: The empirical fact that states do not initiate new maritime disputes once boundaries have been established is used to eliminate dyad-years from the dataset after the boundary has been settled. If this empirical regularity were not the case one would need to include all yours after a boundary settlement because there is a potential for a future boundary dispute.

10. Whenever everyone is confronted with a two-way theory table one should always ask whether the row variable is more or less important than the column variable, or equally important. Analyze the table below from Lessing's influential book on state-cartel violence. Determine if row or column is more important in his theory. Hint: Consider the outcomes in the cells of the table to be ordinal, from 1 to 4. Then write a little equation where row and column can take on values of one or two and where the equation combining row and column generates the outcomes in the cells.

Figure 6: Two-way table analysis: Lessing on state-cartel violence

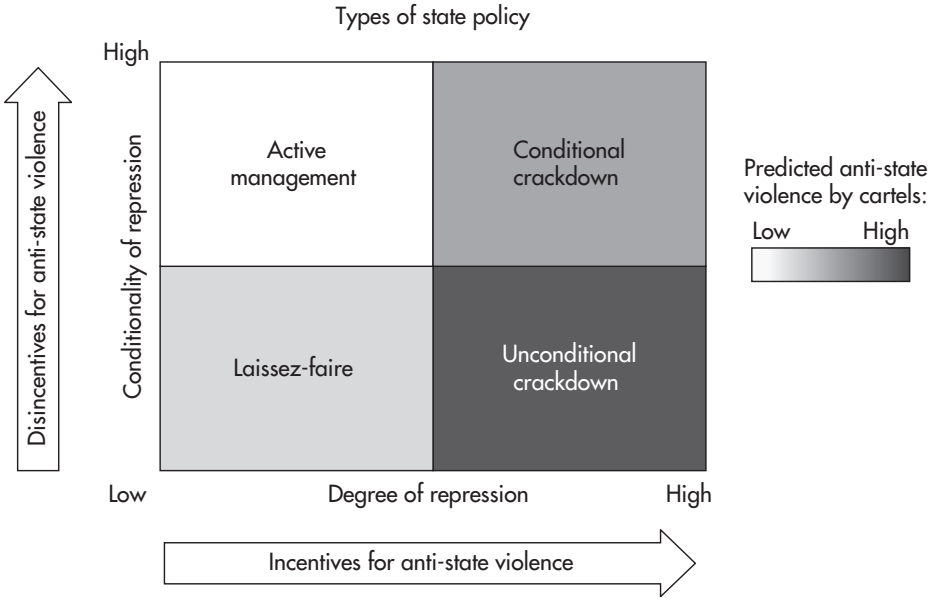


FIGURE 1.1. **Degree vs. Conditionality of Repression: Policy Types and Incentives for Violence.** *Degree of repression* measures the overall amount of force directed at cartels; *conditionality of repression* measures how much additional repression a cartel incurs by engaging in anti-state violence.

Answer: Conditionality is more important than degree of repression. The basic equation is 1*Degree + 2*Conditionality = cartel violence

11. Core to LNQA is the determination of the basic data set and then within that the establishment of LNQA scope. Typically, LNQA scope is a subset of the basic data set. This means there are potentially multiple basic datasets which have the same LNQA scope subset subset. Sechser and Fuhrmann's analysis of the success of nuclear compellence efforts, discussed in more detail in chapter 7 of the LNQA book , provides an example of an X generalization of the form "if nuclear compellence ($X = 1$), then it will not succeed ($Y = 1$, coded as compellence failure). This provides an interesting example where there are multiple potential basic datasets. An X generalization involves a

subset of the basic dataset. It is quite possible that two different basic datasets have the same X generalization subset. This is just a Venn diagram with two circles that have an intersection: the X generalization is that intersection. But the choice of the basic dataset is critical because it determines the boundary between internal and external validity and is hence consequential for the the overall research program.

To construct their basic dataset they start with an existing dataset of "Militarized Compellent Threats" which contains 210 interstate compellent threat initiator-target dyads over the years 1918–2001. In virtually all basic datasets the author must make explicit spatial and temporal boundaries. Here the temporal boundaries are 1918–2001 and the spatial boundaries include all dyadic *interstate comepllent threats*. The LNQA dataset could be derived by coding these cases for whether they involved nuclear compellence (as opposed to conventional compellence for example). The LNQA generalization only involves $X = 1$ cases, however, and out of those compellent threats only 19 are $X = 1$ cases of attempted *nuclear* compellence.

A second option for defining the basic dataset is to take all MIDs involving nuclear powers and code whether or not that MID involved nuclear compellence. This generates the same set of $X = 1$ cases. However, clearly this basic dataset does not include nonnuclear powers and thus has no cases before 1945 among the $X = 0$ cases. Both these basic datasets choose the same $X = 1$ cases, but differ significantly on the $X = 0$ cases.

Discuss whether this potential difference in the basic data set has important implications. For example, typically external validity is determined by the basic data set, so does that matter in this particular case?

Answer: These choices have implications for the overall causal generalization, which involves the whole basic dataset. The second option makes more sense because the causal generalization goal is about the behavior of *nuclear* states in militarized disputes. A good test between the two datasets is the counterfactual of going from $X = 0$ to $X = 1$. In the second basic dataset this makes complete sense: only a nuclear power could have decided on adopting a strategy of nuclear compellence. In the first dataset it does not make sense for the vast majority of states because making a nuclear compellent threat is not possible for states that do not have nuclear weapons to begin with.

12. One way to deal with this confirming observations and also strengthen regularities is to move to the tail of the distribution. At the same time many comparative scholars prefer minimalist definitions of democracy. Discuss the potential trade-offs. Leipziger (2025) illustrates going to the tails to increase regularity:

> Note that the employed minimalist conceptualization of democracy also holds implications for my findings. In Online Appendix G, I show that where democracy had the least impact on reducing ethnic inequality, the quality of democracy was lowest. This fits with a broader pattern regarding the consequences of democracy, showing that the beneficial effects mainly manifest in enduring, high-quality democracies. (Leipziger 2025, 389)

13. When scholars want to dichotomize democracy (or break it into categories) using the V-Dem data they usually use the fairly complicated scheme illustrated in the figure below, figure 7, from Lührmann et al. (2018). Basically, for example, .75 is *not* the threshold for "liberal democracy" Lührmann et al. add three additional, necessary requirements. Does this means that gthe V-Dem scale by itself "overshoots" the concept of liberal democracy? Could you use these ideas to reformulate the V-Dem continuous measure so that .75 is now "liberal democracy", .50 is "electoral democracy", .25 and .25 is "electoral autocracy"? Also, they broke up the continuum into four parts, are there arguments for breaking it into 5 parts?

Figure 7: Dichotomizing democracy: the Lührmann et al.–V-Dem approach

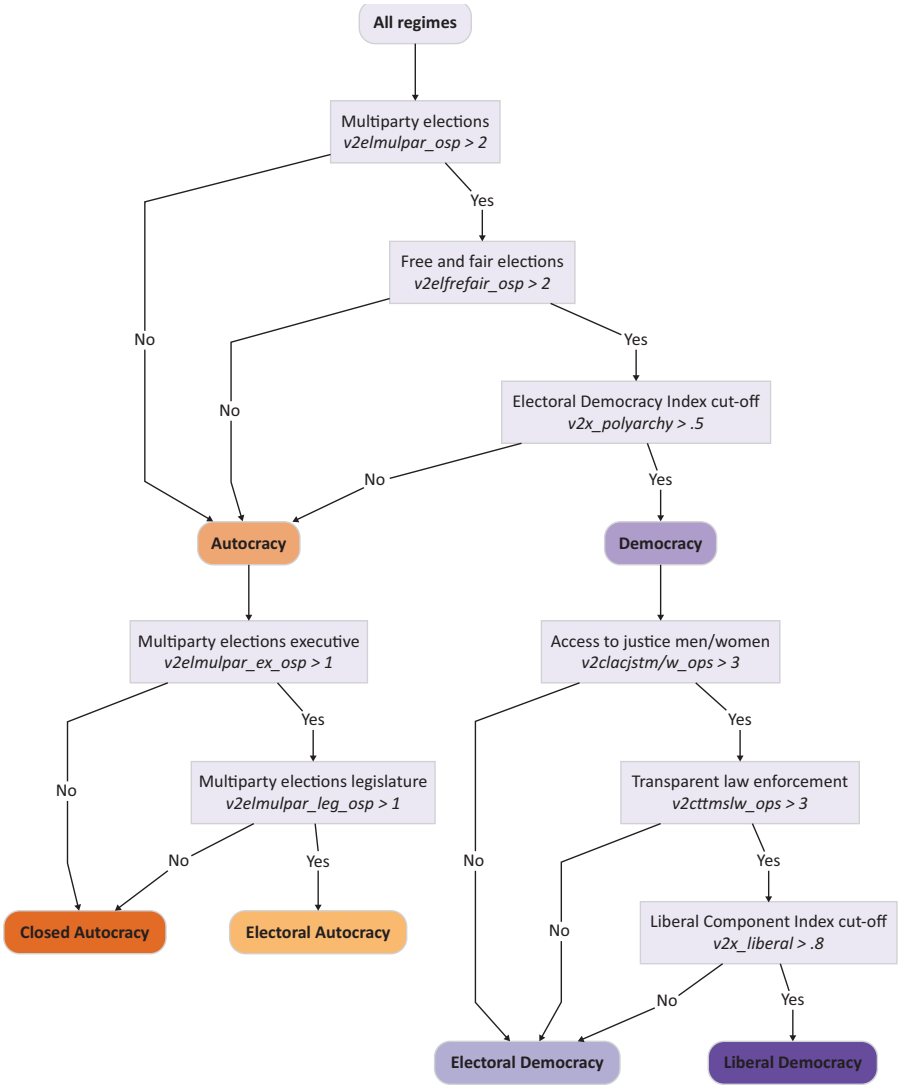


Figure 1. Coding schema for the RoW typologies (for descriptions of variables see Coppedge et al., 2017a).

Answer: For the last question, the concept book recommends odd-number divisions. This permits there to be a zone in the middle. Even number divisions have the middle split between two different categories.

14. Ross (2004) on the relationship between natural resources and civil war has been cited more than a thousand times and has recently been replicated by Becker and colleagues (2024). Like LNQA it falls into the medium-N qualitative research zone with his selection of 13 cases. While published before the explosion of work on qualitative case selection, it nevertheless useful to look at his rationale for selecting cases, notably because the goal is to use these case studies to evaluate various proposed causal mechanisms linking natural resources to civil war (a major civil war topic in the early 2000s due to Collier and Hoeffler's work). He does not spend much time describing and justifying his selection of cases. Basically his whole description of case selection is as follows:

    > The thirteen cases in the sample were selected from the Collier-Hoeffler list of thirty-six civil wars that began or continued between 1990 and 2000 (see Table 1). The cases were chosen on a "most likely" basis: the sample includes *all* civil wars that occurred between 1990 and 2000 in which scholars, nongovernmental organizations, or UN agencies *suggested that natural resource wealth, or natural resource dependence, influenced the war's onset, duration, or casualty rate* (Ross 2004, 46; emphasis is mine)

    In LNQA terms his is choosing a subset of the (1,1) cases, i.e., where $X = 1$ causes $Y = 1$ according to some scholars. Discuss why it might not be *all* the (1,1) cases.

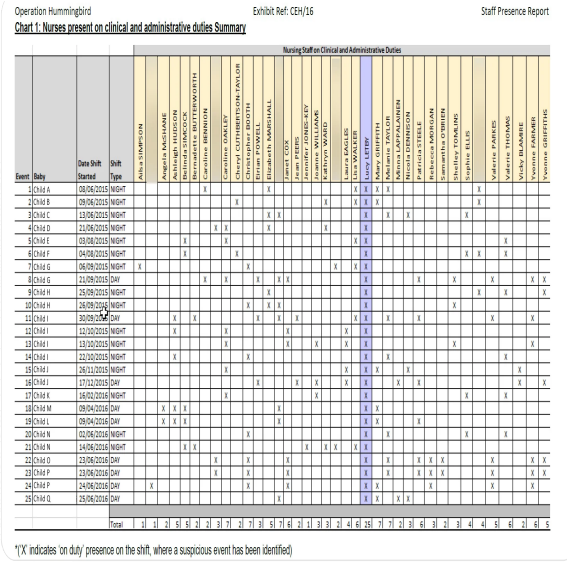
    Answer: In the text emphasized above he is choosing cases where scholars have *already made within-case causal inferences*. So the cases he is missing are those where $X = 1$ and $Y = 1$ but where scholars see no causal connection between the two. Potentially, there are zero of these cases but one would need to look at *all* the (1,1) cases in the Collier-Hoeffler list to verify this.

15. In a long *New Yorker* (2024) article (see also the ITV documentary on the case) Aviv discusses a very famous case in the UK involving a nurse Lucy Letby convicted of killing six newborn babies and attempted killing of five others. The core evidence was a very strong regularity which placed her at the "scene of the crime," see table 12 below. A huge debate revolved about this regularity from a statistical point of view. Apparently the regularity was what convinced many. Discuss the Russo-Williamson hypothesis in this connection, i.e., the mechanism of death of the babies. Was there strong mechanism evidence?

    > Letby had worked on a struggling neonatal unit at the Countess of Chester Hospital,run by the National Health Service, in the West of England, near Wales.The case centered on a cluster of seven deaths, between June, 2015, and June, 2016. All but one of the babies were premature; three of them weighed less than three pounds. No one ever saw Letby harming a child, and the coroner did not find foul play in any of the deaths. . . . The case against her gathered force on the basis of a single diagram shared by the police, which circulated widely in the media [see the table below]. On the vertical axis were twenty four "suspicious events," which included

> the deaths of the seven newborns and seventeen other instances of babies suddenly deteriorating. On the horizontal axis were the names of thirty-eight nurses who had worked on the unit during that time, with X's next to each suspicious event that occurred when they were on shift. Letby was the only nurse with an uninterrupted line of X's below her name. She was the "one common denominator". . . . Letby had become the country's most reviled woman – "the unexpected face of evil," as the British magazine *Prospect* put it – largely because of that unbroken line. (Aviv 2024, 34)

Table 1: Regularities and Letby's conviction of murder of newborn babies



Operation Hummingbird — Exhibit Ref: CEH/16 — Staff Presence Report

Chart 1: Nurses present on clinical and administrative duties Summary

Nursing Staff on Clinical and Administrative Duties

Event	Baby	Date Shift Started	Shift Type	Alba SIMPSON	Angela MASHANE	Ashley HUGHES	Belinda SHACOCK	Bernadette BUTTERWORTH	Caroline BENNION	Caroline OAKLEY	Cheryl CUTHBERTSON-TAYLOR	Christine BOOTH	Elaine DOYLE	Elizabeth MARSHALL	Janet COY	Joan PEERS	Jennifer JONES-KEY	Joanne HALLAM	Kathryn WARD	Laura ANGLES	Lisa MAIR	Lucy LETBY	Mary GRIFFITH	Melanie PAYLOR	Michelle HAMILTON	Nicola DENNISON	Patricia STEELE	Rebecca MORGAN	Samantha CORBIN	Shazia ARIF	Sophie ELLIS	Valerie PARKES	Valerie THOMAS	Vicky BLAMIRE	Yvonne FARMER	Yvonne GRIFFITHS
1	Child A	08/06/2015	NIGHT						X					X							X	X	X	X							X					
2	Child B	09/06/2015	NIGHT							X									X		X	X	X									X				
3	Child C	13/06/2015	NIGHT									X	X									X		X		X				X						
4	Child D	21/06/2015	NIGHT							X	X		X						X			X														
5	Child E	03/08/2015	NIGHT				X				X										X	X											X			
6	Child F	04/08/2015	NIGHT				X					X										X									X	X		X		
7	Child G	06/08/2015	NIGHT	X									X							X		X	X													
8	Child G	21/09/2015	DAY						X		X			X		X	X						X				X				X				X	X
9	Child H	25/09/2015	NIGHT												X								X									X		X		X
10	Child H	26/09/2015	NIGHT										X		X	X							X								X					
11	Child I	30/09/2015	DAY				X		X					X		X	X					X	X			X							X			X
12	Child I	12/10/2015	NIGHT				X				X						X					X	X													
13	Child I	13/10/2015	NIGHT								X						X		X			X	X								X					X
14	Child I	22/10/2015	NIGHT				X						X									X	X			X							X			
15	Child I	26/11/2015	NIGHT								X									X		X	X				X							X		
16	Child I	17/12/2015	DAY										X			X		X		X		X	X			X		X						X		X
17	Child K	16/02/2016	NIGHT								X								X				X									X				
18	Child M	09/04/2016	DAY			X	X	X							X								X	X												
19	Child L	09/04/2016	DAY			X	X	X							X								X	X				X								
20	Child N	02/06/2016	NIGHT										X										X									X				
21	Child N	14/06/2016	NIGHT					X	X											X		X	X		X											
22	Child O	23/06/2016	DAY							X		X				X							X		X			X	X	X		X			X	X
23	Child P	23/06/2016	DAY							X		X				X							X		X			X	X	X		X			X	X
24	Child P	24/06/2016	DAY		X							X				X							X	X				X				X			X	
25	Child Q	29/06/2016	DAY												X								X		X		X									
		Total		1	1	2	5	5	2	2	3	7	2	7	3	5	7	6	2	1	3	3	2	4	6	25	7	7	2	3	6	3	2	3	4	4	5	6	2	6	5

*('X' indicates 'on duty' presence on the shift, where a suspicious event has been identified)

Source: Aviv 2024.

16. Necessary conditions as causes are often connected to scope conditions which are also often causes. Discuss the following claim and whether it should be formulated as a scope condition:

> There is another reason why the notions of difference-maker and cause come apart. This is where we have a necessary condition for something to occur but where it is not really a causal contributor to it. For example, having a liver is a necessary condition for having cirrhosis. No one could have cirrhosis unless they had a liver; therefore having a liver is a difference-maker with respect to that disease. But it would seem wrong to say that the liver caused the disease even though it is a necessary condition for having it. (Anjum and Mumford 2018, 127)

17. A important question in causal model diagraming and statistical models is whether a given factor is a scope condition or a moderator. In many setting a given factor seems like it could be either. For example, discuss this question in the context of Slough and Tyson's discussion of external validity. The same issue arises in Esterling et al.'s discuss of external validity.

> Empirical targets depend on at least three critical ingredients of a constituent study. First, a study takes place in a setting, represented by $\theta \in \Theta$, and captures all the contextual features, as well as characteristics of the study population, that determine the mechanism's effect, or quantitative influence. The setting can include both measured and unmeasured features or characteristics that are relevant to how the mechanism's influence arises and how it is measured. (Slough and Tyson 2024, 12)
>
> We say a claim is externally valid to the extent one has accurately specified why or how the effect generalizes across settings. This means specifying the range of conditions in settings γ across which the effect generalizes. Hence, we define external validity as the correct specification of the conditions that enable/disable or augment/weaken a causal effect. External validity is present when that specification is true – i.e., when γ is correctly labeled. (Esterling et al. 2025, 13)

18. Collins et al. (2024) discuss case selection and research design in case study research. One of the core questions that they pose involves comparative case studies. How does their discussion relate the LNQA which does not involve comparative case studies and how does that relate to generalizability?

> Question F: How Does the Study Structure in Strategic Comparison?
>
> A central tenet in the scholarly enterprise is that generating explanations requires variation. Qualitative researchers seeking conceptual generalizability need to be able to identify, describe, and explain variations in their observations. This requires some form of structured comparison. (Collins et al. 2024, 31)

19. In case study research designs the scope of the theory-empirical analysis is often clear when a few cases are selected for empirical analysis. Discuss what the are scope conditions for Schwartz's analysis. How do the cases selected connect to the scope conditions?

> Case Selection. Within the Central American context, I analyze three institution-level cases: the customs administration and policing institutions in Guatemala and land tenure in Nicaragua. Each of these cases demonstrates the wartime emergence of alternative rules of the game that have similarly undermined state performance. These specific cases were selected for several reasons. First, they each reflect distinct domains of state activity (extractive, coercive, and administrative), allowing me to assess the common building blocks that characterize processes of state institutional change broadly. Second, by comparing the Guatemalan and Nicaraguan contexts, I leverage variation on core political and structural factors, which could plausibly shape processes of institutional development in different ways. Specifically, the two countries differ widely in terms of civil war antecedents, regime ideologies, war-making strategies, and the nature of foreign support. The comparison thus provides a critical opportunity to explore why we observe the emergence of similar kinds of institutional logics amid such stark differences. . . . This book focuses on three cases of undermining rules, above all, because it seeks to identify the mechanisms underlying the relationship between civil war and"weak" states. (Schwartz 2023, 27)

20. Scaling issues are ubiquitous in datasets and indicator construction. Discuss the scaling description of a variable in a new education dataset: is it nominal?

2.1 Existence of compulsory education (compuls_exist)

Question: Is there a compulsory education system in place?

Clarification: The system must require at least one year of compulsory education for children/youth.

Answer categories:

(a) No, there is no legal requirement that education must be compulsory
(b) Yes, there is a legal requirement of compulsory education in place for certain groups of the population (e.g., by ethnicity, gender or geographic region)
(c) Yes, there is a legal requirement of compulsory education in place for all children in the relevant age group

Measurement level: Nominal (Lutscher et al. 2022, 3)

Answer: It is clearly ordinal not nominal.

21. The debate about how to think about necessary and sufficient conditions in a statistical context has been common for many years. Discuss the negative position of Gailmard – widely influential – on this issue. Note that at the end he states that what matters is the data generating process (DGP); what if these are modeled as necessary or sufficient conditions? Does this contradict the notion of stochastic?

> To be (somewhat more) specific, suppose we are interested in explaining a social event or outcome Y that can either be "large" or "small" in value. This might be, for instance, the extent of a civil war in a group of countries. And we want to know whether some factor X that can either be "present" or "absent" helps us to explain the outcome. Suppose that the observed cases line up with the attributes in Table 3.1. In this data we can "observe" that the presence of factor X is a necessary condition for Y to be large; Y is never large otherwise. On the other hand, the absence of factor X is sufficient for Y to be small; whenever X is absent this value of Y ensues. This all makes perfect sense until it is revealed that X and Y are both coins. "Present" means the first one came up heads when flipped and "absent" means it came up tails; "Large" means the second one came up heads when flipped, and "small" means it came up tails. When we know the DGP and we know it is stochastic, it is transparent that methods to derive necessary and/or sufficient conditions for a particular result Y to occur are faulty. . . . In this book, DGPs are formalized as probability distributions. . . . The hunt for necessary and sufficient conditions is misguided in stochastic DGPs. (Gailmard 2014, 80, 82)

Table 2: Necessary and sufficient condition data analysis

TABLE 3.1. *Observations for Outcome Y and Explanatory Factor X*

Case	***X* Attribute**	***Y* Attribute**
1	Absent	Small
2	Present	Large
3	Present	Large
4	Present	Small
5	Absent	Small

Source: Gailmard 2014.

Answer: See Dion's Bayesian probabilistic analysis of necessary conditions reproduced in chapter 5 of LNQA book.

22. Two-way tables are quite ubiquitous in theorizing both for qualitative and quantitative methods. These raise a variety of questions regarding the theory – what is the relationship between the 2–3 independent variables – as well as case selection. Discuss the two-way table below. How many dimensions are there in fact? In fact it is really a multidimensional table; what are the independent variables? How do they interact to produce the dependent variable of welfare provision? What case studies would be the most useful? How would one one do M-LNQA using their data–theory, given the various interaction terms in their theory?

Table 3: Predicted welfare effect of subversion in resource-poor and resource-rich countries

	Resource-poor countries		Resource-rich countries	
	Separatist subversion	Center-seeking subversion	Separatist subversion	**Center-seeking subversion**
Guns vs. butter constraint	Strong	Strong	Weak	**Weak**
Motivation for mass co-optation	Weak	Strong	Weak	**Strong**
Overall effect on welfare provision	Negative	Indeterminate	Indeterminate	**Positive**

Note: Core hypothesis in bold.

Source: Eibl and Hertog 2024.

23. Fairfield and Charman (2022 chapter 4 table 1) decibel scale to measure the strength of evidence. Compare that with the LNQA proposal for strength in chapter 5 of the LNQA book.

Table 4: Qualitative to quantitative correspondences (dB)

dB	Acoustic Perception	Plain Language Description	Equivalent Odds or Likelihood Ratio
3	smallest meaningful difference	very weak	2:1
6	clearly noticeable difference	weak	4:1
10	twice as loud	moderate	10:1
20	four times louder	strong	100:1
30	eight times louder	very strong	1000:1

Source: Fairfield and Charman 2022.

Answer: Notice that both are quite nonlinear. The LNQA scale is S-shaped like a logistic function while Fairfield and Charman is log-linear. One can tell how very nonlinear the the Fairfield and Charman scale is by looking at the likelihood column which starts out fairly slowing then increases very rapidly.

24. Craggs and Kelly (2018) do a meta-ethnographic synthesis of studies of school belonging. They summarize their analysis in the figure below which appears at the end of their article. To put it very briefly, ethnographers, interpretivists, etc., often are hesitant to include causal claims and are more interested in meaning (e.g., concepts) than causal relationships. Using ideas and figures from Goertz (2020) do two different versions of this figure. Do a completely conceptual one, noting that already in the figure there are at least three levels, "Higher order concept," "Main concept" and "Subconcept", from Goertz's perspective a three-level concept. Then do a causal version where the "Higher order concept" becomes the dependent variable, and hence it is a causal model (Hint: do it as a two-level model, chapter 10 of the concept book).

Table 5: Conceptualizing various forms of belonging)

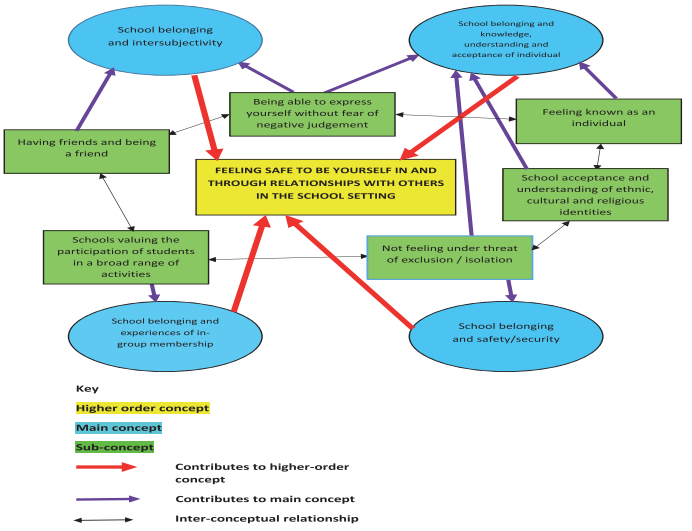


Figure 1. Diagrammatic representation of the higher order concept, main concepts and sub-concepts.

Source: Craggs and Kelly 2018.

25. One of the core guidelines when thinking about conceptualization is to think about terminology. Core to a massive data project is the concept of "social conflict" as defined below. What is one core component of the concept described below that probably should be included in the term "social conflict"? How does this have concrete implications for the data set described in their table 1?

> We used the expression"social conflict" as a generic category to encompass all the expressions used by scholars and/or mentioned in the sources. We followed Nicolas (2002, 75) and considered social conflict to be any event

> involving a group of at least three individuals belonging to different families and which either perpetuates violence or threatens violence against one or more members of a different group or against representatives of political, religious, and economic power; or any event involving an attack on property, buildings, furniture, papers, or other signs symbolizing such powers. We chose to use a low threshold on the number of participants not only to grasp infinitesimal episodes of discontent, but also to allow for the inclusion of other datasets that used different thresholds. (Chambru and Maneuvrier-Hervieu 2024, 1085)

Answer: Central to their concept is the notion of *violent* conflict which lies at the core of their definition above. Instead of being assumed it should be made explicit because there is a whole literature on nonviolent conflict which may or may not be relevant.

26. Chapter 8 of the concept book discusses the standard rule for typologies that they be mutually exclusive and exhaustive and argues that these standard methodological rules are not good ones. Discuss the typology in the large-scale database on social conflicts described in the previous exercise, e.g., food riot, slave revolt, political conflict, etc. Is the typology meant to be mutually exclusive and exhaustive? Are the authors assigning a given conflict only one category? How is that problematic, e.g., conflict fitting multiple categories.

Table 6: List of variables included in the HiSCoD database

TABLE 2. Classification Used in the HiSCoD Database

Category	Numeric values	Description	No. of obs.
Food riot	1	Includes riots protesting against high prices, popular taxation, interception of food transport, and food looting	4,226
Tax riot	2	Includes smuggling-related events, and riots protesting against taxes (local, feudal, or national)	4,301
Religious conflict	3	Includes conflict relating to the practicing of religion	649
Conflict with local or national authorities	4	Includes conflict with the military (including deserters and draft dodgers), police, marshals, parliaments, municipal authorities, militias or any conflict protesting against a decision taken by authorities	3,550
Feudal conflict	5	Includes attacks on castles and conflicts concerning the abolition of feudal rights	940
Slave revolt	6	Includes slave revolts either on boats, at gatherings, and/or maroons	3
Political conflict	7	Includes all conflicts protesting against a political regime during a revolutionary or transitional period, and any opposition to a political regime aiming to change it	1,088
Labor conflict	8	Includes strikes, collective bargaining, conflict against guilds or between different guilds, wage riots, etc.	1,403
Banditry	9	Includes robberies by a group of individuals or events related to organized crime	0
Other forms of conflict	0	Includes all conflicts which do not fit in previous categories	4,458

Note: We categorized events related to seigneurial taxes as tax riots rather than feudal conflict. In 333 instances, the lack of information in the sources pre-empted the attribution to an exact category.

Source: Chambru and Haneuvrier-Hervieu, 2024.

Answer: While not completely clear it seems like the authors are applying the mutually exclusive and exhaustive rule: "We addressed this issue by creating a simplified classification based on 10 categories to encompass a wide range of geographic areas and a long-time span (Table 2). We assigned each event to one category based on the information available and, when relevant, on

the researchers' original typology – which is also included in the database" (Chambru and Maneuvrier-Hervieu 2024, 1086).

27. Apply the Goertz three-level framework and diagram the concept of civil war complexity described by Brosche et al. 2023, illustrated in their triangle figure below. Once diagrammed in this fashion discuss how one might aggregate across the three dimensions. This is probably necessary to answer their "how much" question and would probably be necessary if civil war complexity were to be used in statistical analyses. The elephant in the room is the concept of *complexity*, which is itself not really defined. Given the three corners of the triangle discuss how one might conceptualize and measure complexity.

> Conceptualizing Civil War Complexity. To identify a civil war's base attributes and to meaningfully capture different degrees and types of complexity, we turn to classic conflict theory. Building on Johan Galtung's conflict triangle, Wallensteen defines conflict as consisting of three elements: (1) actors, (2) behaviors, and (3) contradictions/issues at stake. As such, building our conceptualization on the latter's conflict triangle brings nuances to how civil wars vary and identification of its base components. Allowing civil wars to vary in the type and number of actors, behaviors, and issues thus helps define a system's parts and their interlinkages. We define civil war complexity as the degree to which the core components of conflict – actors, behaviors, and issues – are manifold in number and type, and the degree to which these are interlinked. We unpack each component of our concept both theoretically and empirically. Intertwining theory and empirics allows us to move between Giovanni Sartori's "What is?" question regarding the boundaries of our concept, and the"How much?" question regarding empirical measurement. . . . First, to demonstrate the variations of actor complexity, Figure 2A,B displays variations in the number of actors participating in violence in a country between 1989 and 2016. Figure 2A includes state, nonstate, and communal actors, whereas Figure 2B shows rebel nonstate actors only. (Brosch et al. 2023, 143–44, 156)

Figure 8: Conceptualizing and measuring civil war complexity

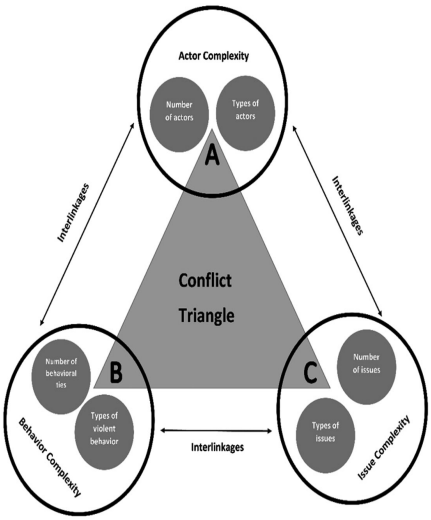


Figure 1. Graphical illustration of conceptual model.

Source: Brosche et al. 2023.

Answer: Since complexity is a continuum one approach would be to start with what configuration of the three corners is the least complex? One might suggest that if one corner scores the maximum while the other two corners are at the minimum that is the least complex conflict. This then suggests that maximum complexity is achieved when all three corners are equal in size. Following this logic then one could use approaches to equality (e.g., income inequality) to conceptualize and measure conflict complexity.

28. Goertz (2017) provides the "Overdetermination rule" for case selection. Discuss Beckley's use of overdetermination as a criterion to exclude cases:

> The History of Peaking Powers since 1870. To assess how rising powers respond to severe and sustained economic growth slowdowns, I analyze every case from 1870 to 2018 in which a great power's per capita gross domestic product (GDP) grew at least twice as fast as the global average for at least seven years and then suffered at least a 50 percent decline in growth rates over the next seven years. . . . This exclusion leaves nine cases, which are displayed in table 1. For each case, I list the year in which the great power's average annual growth rates started to slow. Because they are based on imprecise data, these peak years are rough approximations of when the great power transitioned from rapid to slow growth. I therefore use these years only to identify cases for further research, not to conduct precise quantitative tests. I examine all the cases in detail except interwar France, Germany, and the Soviet Union, because mercantilist expansion was arguably overdetermined in those cases. (Beckley 2023, 24–25)

29. In their study of the impact of nuclear weapons and asymmetry in militarized crises, Fanlo and Sukin find a very strong regularity involving asymmetry and

"high stakes" (i.e., motivation) in these crises. They note that they cannot do statistical analyses because there is no variation on the interaction of asymmetric inferiority and high stakes. How LNQA can be used to do empirical analyses in these "no variation" settings and how can no variation can often mean "strong regularity."

> As our theory predicts, asymmetrically inferior states often have high stakes and prevent their superior opponents from achieving victory. If we consider asymmetric superiority to require an arsenal fifty times larger than the inferior state, then every time an inferior state in an asymmetric crisis dyad had high stakes, the superior state failed to achieve its goals. Moreover, the inferior state had high stakes in all but one asymmetric crisis. . . . Furthermore, we cannot directly test the interaction of asymmetric inferiority and high stakes because there is no meaningful variation. In all cases in our data, with the exception of France in the Berlin Wall Crisis, asymmetrically inferior states have high stakes, as our theory would predict. (Fanlo and Sukin 2023, 463)

Answer: LNQA works exactly with these strong regularity results, and can "directly test" the hypotheses in question.

30. The V-Indoc [Varities of Indoctrination] coding manual (Varieties of Political Indoctrination in Education and the Media (V-Indoc): Codebook 2023), which uses the V-Dem methodology has this for scaling:"Scale: Dichotomous, Nominal, Ordinal, or Interval/Ratio (Extra response options such as N/A or Other, are not counted as part of this classification)." Discuss this list of scale types, particularly "interval/ratio." See the next exercise for one on "dichotomous." For example, is V-Dem or V-Indoc interval or ratio? Since V-Indoc is about education, is "years of education" (a common independent variable) interval or ratio? Also, the coding manual includes "multiple selection" for scale. What scale type is that?

    Answer: As noted in the next exercise dichotomous is not a scale type. It is odd that "interval/ratio" are put together because they are different. V-Dem and V-Indoc are interval, as are all latent variable measurement models. Years of education is ratio because zero is a real zero (zero is arbitrary in interval indicators). "Multiple selection" means nominal.

31. It is always useful to understand the scale type of indicators. Sometimes in coding manuals and elsewhere people will describe the indicator as dichotomous. This is not a scale type. Often it is coded [0,1]. With such a coding scheme describe how the indicator could be nominal, ordinal, interval, or ratio depending on what that [0,1] means.

    Answer: The [0,1] could mean male-female, which is nominal. It could be ordinal when 1 means more than 0 for the indicator. It can be interval when a bunch of [0,1] indicators are added together to form an index. By adding [0,1] indicators one is converting the dichotomous indicators into interval ones. Finally, it could be ratio if the [0,1] indicate the real line segment between zero and one implying there is an underlying continuum between zero and one of real numbers.

32. The quote below talks about a famous article on smoking and cancer. It raises some interesting questions. The first is the exclusion of women, which raises questions about external validity. One general question is whether the strength of the original findings should or does influence our willingness to extrapolate to new populations? What is actually the strength the finding discussed below, i.e., 0.00 versus 1.14? We are accustomed to focusing on the difference between these two numbers, but what if one looked more closely at the 0.00 finding? How might that be seen as a very strong regularity? The second question deals with mechanisms. There are powerful arguments for external validity if the mechanism would be the same in the target population? How might that play out in extrapolating from men to women in this study?

    > THE SMOKING GUN. We close with the one of the most influential articles of the 20th century, to demonstrate how even very narrowly defined study samples may provide widely generalizable results if conducted with an eye to rigorous internal validity. Entitled "The mortality of doctors in relation to their smoking habits: a preliminary report," this 1954 article by Doll and Hill concerned the association between lung cancer and cigarette smoking in British physicians. The investigators excluded physicians who did not return usable responses, and also women physicians, and physicians aged 35 years, because of their low expected frequency of lung cancer deaths. The remaining sample was a male, physician cohort of 24,389, about 40% of those in the Medical Register. During the 29-month follow-up, investigators observed only 36 confirmed lung cancer deaths, occurring at rates of 0.00 per 1,000 in nonsmokers, and 1.14 per 1,000 among smokers of 25 or more grams of tobacco per day. (Kukull and Gangul 2012, 1888–89)

    Answer: For question (1), 0.00 might mean that *no nonsmokers died of lung cancer*. In fact, of the 36 people who died of lung cancer in the study, all were smokers. This could be interpreted as smoking is a necessary condition for lung cancer death. This then means that all mechanisms leading to lung cancer death must include smoking. The second question might be about the extend to which men and women's lungs are different and function differently. If they function the same (i.e., same mechanism) then we might feel that generalizing to women is more valid since we have good reasons to think that the mechanism linking smoking to cancer is will be the same in women.

33. Here is another exercise related to the quote above about smoking and cancer related to selection. The first is the exclusion of women. Given when this study was done they were probably few women physicians to begin with that could be sampled. Is that a reason to exclude them? Given the standard societal norms at the time it is much less likely that women were smokers than men. So is that a selection effect or an antecedent condition that we should worry about? The issue is whether the selection condition has any plausible causal connection to the outcome. Another way to phrase the question is whether we would be in trouble if all we did was just control for gender? Still, we have a real causal chain whereby societal norms lead to women smoking less which leads them to getting less cancer. So do we need to worry about the first link in that causal chain, i.e., endogeneity?

Answer: I think most people would say no. This is because arguing that societal norms about women smoking cause cancer beyond them not smoking seems implausible, i.e., in DAG language there is no direct link between societal norms about smoking and cancer. As discussed in the causal model chapter of the LNQA book, a direct link means that there must be another mechanism leading from societal norms to cancer which does not go through smoking: what would be that mechanism? If there are no really plausible ones then we do not need to worry about a direct path to cancer. So while there is clearly selection there is no other causal mechanism linking selection directly to outcomes hence perhaps no reason to worry about it statistically or otherwise. There might be some other feature of women (e.g., a famous Fisher genetic cause) that produces less cancer but that is separate from the norms causing cancer.

34. Boulanger Martel et al. (2024) conduct a literature review of peace with adjectives. They give numerous examples, 61, of people attaching adjectives to the concept of peace. Discuss how or whether this make sense unless one knows what peace is to begin with, i.e., peace without adjectives.

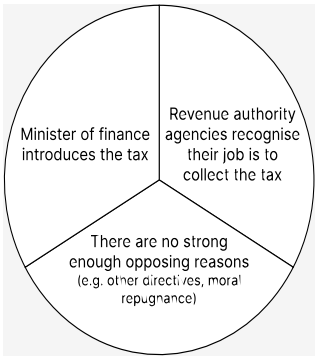
35. Hulme (2026) argues that Congress is a strong constraint on presidential uses of force. He has a nice scatterplot illustrating that congressional approval is a necessary condition. One of the key themes in the qualitative methods literature as well as quantitative is alternative explanations and confounders. Hulme reports (personnel communication) that many suggest that an alternative explanation is the president's view on the use of force. So this could be a control variable or confounder. Goertz (2017) argues that in terms of case selection one should follow the "avoid overdetermination rule". Which cases in Hulme are overdetermined? Which cases then should be chosen when applying the avoid overdetermination rule? Notice the same issues arise in Tannenwald's influential book on nuclear weapons taboo. (Thanks to Hulme for comments on this exercise.)

36. As discussed in chapter 4 on causal models in the LNQA book manuscript, legacy studies involve complex mechanism where factors in the distant past influence current behavior. Do the causal model diagram for Wilfahrt, M. 2022. *Precolonial legacies in postcolonial politics: representation and redistribution in decentralized West Africa*. Cambridge University Press.

37. Cartwright and colleagues (2024) have a different approach from Goertz and Haggard, (causal model chapter), for diagramming necessary conditions which are jointly sufficient using what they call the causal pie and is illustrated in the first figure below. One potential advantage of the causal pie approach is that it could allow one to indicate the relative importance of different factors by the size of the pie slices. In the first figure below each piece of the pie is the same size suggesting equal importance. The second figure from Cartwright et al. diagrams a policy intervention, focus on the first pie in the figure. The policy intervention is graphically indicated by a 1 in a large square. Support factors, i.e., S1.1 and S1.2 which are necessary for the intervention have a causal effect are in elipses and much smaller in size than policy intervention

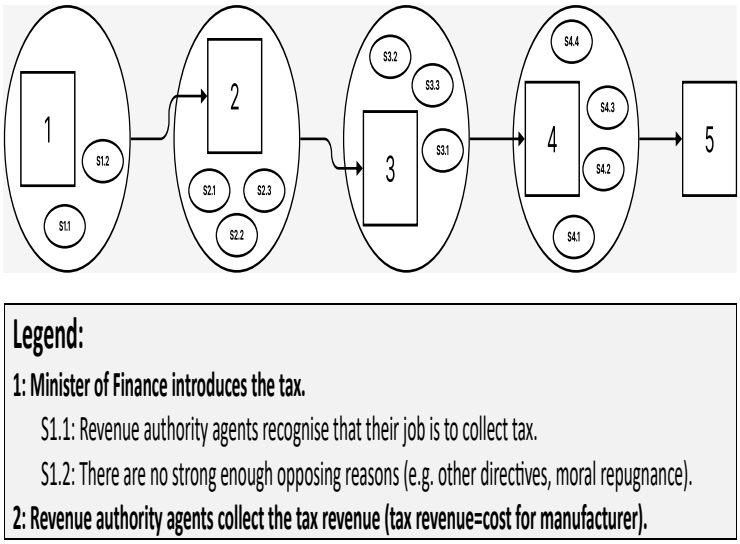
1. This might suggest that these support factors are less important. Might there be situations where the support factors are more important than the policy interventions?

Figure 9: Causal pie: diagramming necessary and sufficient conditions



Source: Cartwright et al. 2024.

Figure 10: Causal pie: relative importance of pieces of the causal pie



Source: Cartwright et al. 2024.

38. Cartwright et al. (2024) discuss the following story from the philosophy literature on causation. Draw a little causal figure to illustrate what is going on. Hint: be sure to put signs above the causal arrows.

> A villain intent on killing a walker pushes a boulder off a bluff to crush the walker. But just as the boulder approaches, the walker notices it and ducks. Pushing the boulder causes the walker to duck. The ducking causes the walker to continue to be alive. (Cartwright et al. 2024, 45)

Answer: Villain pushes boulder toward walker ==(+)==> Walker ducks ==(-)==> Walker stays alive. This is a nice example of the constraint-deterrence model discussed in the causal mechanism-model chapter of the LNQA book. Cartwright likes to call these derailers which is the same basic idea. As in constraint–deterrence examples, the killer is motivated to kill the victim and pushes a boulder, which is a positive cause of death. However, the walker notices that and steps out of the way which is a negative cause. This is a situation where the second cause essentially cancels out the first and one ends up in the status quo ante.

39. Tomz (2007) in his important analysis of international debt and the various means of dealing with potential defaults between countries has a chapter on gunboat diplomacy. This was basically the use or threat of military force by the part of European powers to deal with potentially defaulting governments in Latin America in the 19 century. The statistical analysis nicely illustrates various features of the possibility principal. In his research design gathers data on all the major European and US suppliers capital for 19th-century Latin American countries The basic statistical unit of analysis is the investor-debtor dyad. These are the dyads where it might be possible that the creditor country uses military force against the debtor country who is in default or threat of default. One aspect of the possibility principal he ignores is whether the given investor has has the capability to use military force against a defaulting country. The list of creditor countries is Great Britain, France, Germany, Belgium, the Netherlands, Switzerland, and the United States. So the possibility principal asks if it is possible that a country like Switzerland, which is landlocked and has no Navy could even possibly use gunboat diplomacy. The same can be asked of countries like Belgium. Also interesting are cases where early in the century the country does not have the naval capability, but does have it later in the century such as Germany. So the question is whether one should exclude all of these dyads where the investor country does not have the military capability to conduct gunboat diplomacy?

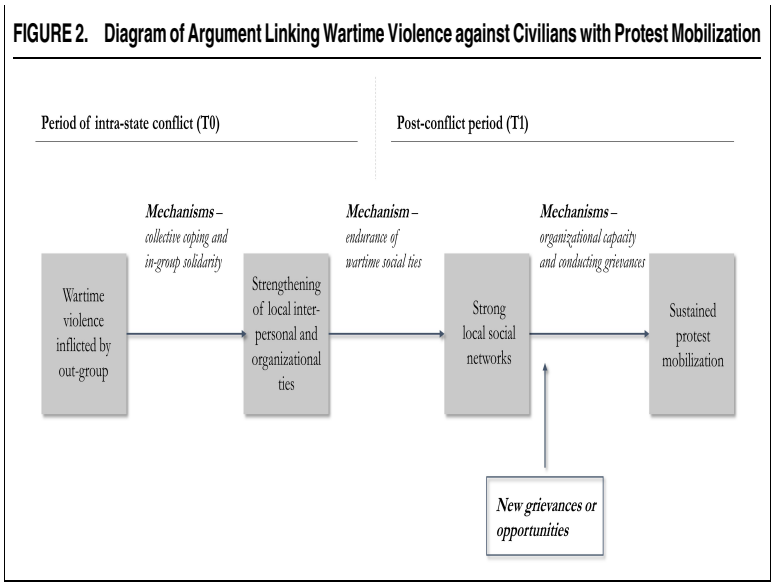
    Answer: Application of the Possibility Principle would involve both (1) the interest in potentially using the gunboat diplomacy as evidenced by potential defaults as well as (2) the capacity to conduct gunboat diplomacy via the possession of a Navy capable of intervening in South America. Tomz uses the first one but not the second one.

40. Another exercise related to the one above deals with the fact that Tomz stops his data analysis in 1920: "All alleged examples of bondholder wars, including the famous Venezuelan crisis, occurred in the century before World War I. By the 1920s, military enforcement was already seen as 'a thing of the past' and a 'matter of history' "(Tomz 2007, 116). Obviously issues of default for Latin American countries continue and become a major issue for example in the 1980s. So evaluate his rationale for stopping in 1920? He is saying is that

there are no examples of the dependent variable from 1920 onward. Is that a good-valid reason to stop the analysis in 1920? Note that this is effectively selecting on the dependent variable (i.e., when it is consistently zero).

41. Berman et al. (2024) provide a causal mechanism in their analysis of popular mobilization in Iraq. In particular discuss the factor "new grievances or opportunities" box in the model. It has an arrow going up to the final arrow in the figure. How are we to interpret what this arrow means? In terms of our modeling guidelines this looks like a moderator variable because it is an arrow to an arrow. One might argue that in fact it indicates an interaction term. What evidence can you provide just that the interaction term interpretation is correct?

Figure 11: Diagram of argument linking wartime violence against civilians with protest mobilization



Source: Berman et al. 2024.

Answer: One bit of evidence would that there is an interaction term in the statistical model.

If it is an interaction term discuss their case selection strategy for their case study in the article which is described as follows:

> Our quantitative results point to a positive relationship between violence legacies and protest, especially when the perpetrator of violence is from an out-group like a foreign invader or an ethnic militia. To further unpack the mechanisms behind this relationship, in this section, we conduct a brief case study of Fallujah, which witnessed some of the fiercest battles during the war and was a site of significant protest activity during 2011. We selected Fallujah to analyze for several reasons. First, it is a district whose protest activity was relatively high and is "well predicted" by our regression models . . . Second, it is a district where our models suggest that

> civilian casualties had a considerable impact on protest; when we remove civilian casualties from the model, the predicted protest levels in Fallujah fall well below observed values. Third, there are abundant historical and qualitative data on Fallujah, since it was a major site of conflict during the war. (Berman et al. 2024, 226)

Should one be selecting on the interaction term not only the individual original causal variable, civilian casualties, in the interaction model term?

Large-N Qualitative Analysis (LNQA)

1. In their study of the impact of nuclear weapons and asymmetry in militarized crises, Fanlo and Sukin find a very strong regularity involving asymmetry and "high stakes" (i.e., motivation) in these crises. They note that they cannot do statistical analyses because there is no variation on the interaction of asymmetric inferiority and high stakes. How LNQA can be used to do empirical analyses in these "no variation" settings and how can no variation can often mean "strong regularity."

   > As our theory predicts, asymmetrically inferior states often have high stakes and prevent their superior opponents from achieving victory. If we consider asymmetric superiority to require an arsenal fifty times larger than the inferior state, then every time an inferior state in an asymmetric crisis dyad had high stakes, the superior state failed to achieve its goals. Moreover, the inferior state had high stakes in all but one asymmetric crisis. . . . Furthermore, we cannot directly test the interaction of asymmetric inferiority and high stakes because there is no meaningful variation. In all cases in our data, with the exception of France in the Berlin Wall Crisis, asymmetrically inferior states have high stakes, as our theory would predict. (Fanlo and Sukin 2023, 463)

   Answer: LNQA works exactly with these strong regularity results, and can "directly test" the hypotheses in question.

2. Fairfield and Charman (2022 chapter 4 table 1) decibel scale to measure the strength of evidence. Compare that with the LNQA proposal for strength in chapter 5 of the LNQA book.

Table 7: Qualitative to quantitative correspondences (dB)

dB	Acoustic Perception	Plain Language Description	Equivalent Odds or Likelihood Ratio
3	smallest meaningful difference	very weak	2:1
6	clearly noticeable difference	weak	4:1
10	twice as loud	moderate	10:1
20	four times louder	strong	100:1
30	eight times louder	very strong	1000:1

Source: Fairfield and Charman 2022.

Answer: Notice that both are quite nonlinear. The LNQA scale is S-shaped like a logistic function while Fairfield and Charman is log-linear. One can tell how very nonlinear the the Fairfield and Charman scale is by looking at the likelihood column which starts out fairly slowing then increases very rapidly.

3. Collins et al. (2024) discuss case selection and research design in case study research. One of the core questions that they pose involves comparative case studies. How does their discussion relate the LNQA which does not involve comparative case studies and how does that relate to generalizability?

> Question F: How Does the Study Structure in Strategic Comparison?
> A central tenet in the scholarly enterprise is that generating explanations requires variation. Qualitative researchers seeking conceptual generalizability need to be able to identify, describe, and explain variations in their observations. This requires some form of structured comparison. (Collins et al. 2024, 31)

4. Ross (2004) on the relationship between natural resources and civil war has been cited more than a thousand times and has recently been replicated by Becker and colleagues (2024). Like LNQA it falls into the medium-N qualitative research zone with his selection of 13 cases. While published before the explosion of work on qualitative case selection, it nevertheless useful to look at his rationale for selecting cases, notably because the goal is to use these case studies to evaluate various proposed causal mechanisms linking natural resources to civil war (a major civil war topic in the early 2000s due to Collier and Hoeffler's work). He does not spend much time describing and justifying his selection of cases. Basically his whole description of case selection is as follows:

> The thirteen cases in the sample were selected from the Collier-Hoeffler list of thirty-six civil wars that began or continued between 1990 and 2000 (see Table 1). The cases were chosen on a "most likely" basis: the sample includes *all* civil wars that occurred between 1990 and 2000 in which scholars, nongovernmental organizations, or UN agencies *suggested that natural resource wealth, or natural resource dependence, influenced the war's onset, duration, or casualty rate* (Ross 2004, 46; emphasis is mine)

   In LNQA terms his is choosing a subset of the (1,1) cases, i.e., where $X = 1$ causes $Y = 1$ according to some scholars. Discuss why it might not be *all* the (1,1) cases.

   Answer: In the text emphasized above he is choosing cases where scholars have *already made within-case causal inferences*. So the cases he is missing are those where $X = 1$ and $Y = 1$ but where scholars see no causal connection between the two. Potentially, there are zero of these cases but one would need to look at *all* the (1,1) cases in the Collier-Hoeffler list to verify this.

5. One way to deal with this confirming observations and also strengthen regularities is to move to the tail of the distribution. At the same time many comparative scholars prefer minimalist definitions of democracy. Discuss the potential trade-offs. Leipziger (2025) illustrates going to the tails to increase regularity:

> Note that the employed minimalist conceptualization of democracy also holds implications for my findings. In Online Appendix G, I show that where democracy had the least impact on reducing ethnic inequality, the quality of democracy was lowest. This fits with a broader pattern regarding the consequences of democracy, showing that the beneficial effects mainly manifest in enduring, high-quality democracies. (Leipziger 2025, 389)

6. Critical to LNQA is the connection between the causal model and Boolean interpretations of the model. Madrid (2025) provides a very interesting example of a study with close connections to LNQA. He basically looks at democracy-authoritarian outcomes in South America in the 19th century ending in 1929. While the book has democracy in its title most of the countries in South America are not democratic in the 19th or early 20th century. However, he does focus particular attention on the four countries that did achieved democracy by the interwar period. As one can see from the causal model figure below, figure 12, he is not only explaining democracy but also authoritarianism.

Figure 12: Explaining democracy-authoritarianism in 19th century South America

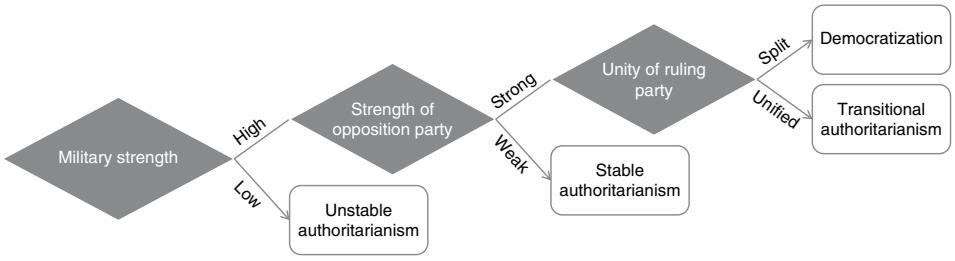


FIGURE 1.1 Determinants of regime outcomes in South America before 1930

The figure has his three main causal factors and he says this:

> [R]egime outcomes in South America during the nineteenth and early twentieth century were shaped by three main variables: the strength of the military, the strength of opposition parties, and the unity of ruling parties (see Figure 1.1). These three variables did not change in a strict sequence, nor were they closely causally related to each other. The strengthening and professionalization of the military did not cause strong opposition parties to emerge, nor did it necessarily precede the rise of strong opposition parties. (Madrid 2025, 38)

There are a number of interesting questions raised by his causal model. For example, would the figure change significantly if the factors were put in a different (temporal) order? Can one give a Boolean interpretation of this causal model, for example, one Boolean model per outcome? For example, to what extent are the three factors sufficient for the outcome of democracy? Are they also necessary for the outcome? The same can be asked of all of the outcomes illustrated in the figure.

7. Y generalizations are closely connected to selecting on the dependent variable concerns because one is selecting cases of Y equals one. Ever since King, Keohane, and Verba (1994) these have been controversial and the standard line is still one should never do that and one cannot learn anything from doing it. For example, Bueno de Mesquita and Fowler (2021) have a whole chapter on this topic in their textbook. Discuss the quote below by Pearl and McKenzie which is using the canonical smoking and lung cancer example. Notice how Pearl and McKenzie state that the probabilities are backwards, probability of X given Y. They claim that one knows nothing based on these data that would be relevant to a decision about smoking. Explain why that is false.

> Of course Hill knew that an RCT was impossible in this case, but he had learned the advantages of comparing a treatment group to a control group. So he proposed to compare patients who had already been diagnosed with cancer to a control group of healthy volunteers. Each group's members were interviewed on their past behaviors and medical histories. To avoid bias, the interviewers were not told who had cancer and who was a control. . . . The probability logic is backward too. The data tell us the probability that a cancer patient is a smoker instead of the probability that a smoker will get cancer. It is the latter probability that really matters to a person who wants to know whether he should smoke or not. The results of the study were shocking: out of 649 lung cancer patients interviewed, all but two had been smokers. (Pearl and Mackenzie 2017, 192)

   Answer: The regularity that the research discovered is extremely strong, .97. So one might conclude that a necessary condition for lung cancer is smoking. Notice that this does not tell you a lot about what Pearl and McKenzie want to know which is what is the probability of lung cancer if one smokes. But it does tell you that if you do not want to get lung cancer you are almost certain not to get it if you do not smoke. Smoking is part of virtually all paths to lung cancer. Certainly a useful bit of information.

8. A core feature of LNQA is the use of "generalization case studies." These are less detailed than than the core case studies but nevertheless examine whether the mechanism question is at work in additional cases within the scope. Sometimes people call these "plausibility probes" as illustrated in the quote below. Discuss the terminological issues between calling them generalization case studies versus plausibility probes. The terminology does suggest different goals.

> This chapter reaches beyond China, South Korea, and Taiwan to examine anti-corruption efforts in a diverse set of authoritarian regimes: Cuba, Malaysia, Rwanda, Singapore, and Vietnam. I use short case studies of these regimes as "plausibility probes" for the broader applicability of my theory in different regions and in different types of authoritarian systems. (Carothers 2022, 197)

9. LNQA uses Y and X generalizations. Comparative methods work often refers back to Mill's classic book which has been a part of the methods discussion for decades. To which of Mill's methods do these most closely seem connected?

Answer: Y generalizations match nicely onto Mill's method of agreement. However, X generalizations fit less well with any of them. Perhaps they fit best with the little-discussed "Method of concomitant variations." From the wikipedia page:

> Whatever phenomenon varies in any manner whenever another phenomenon varies in some particular manner, is either a cause or an effect of that phenomenon, or is connected with it through some fact of causation. . . . If across a range of circumstances leading to a phenomenon, some property of the phenomenon varies in tandem with some factor existing in the circumstances, then the phenomenon can be associated with that factor. For instance, suppose that various samples of water, each containing both salt and lead, were found to be toxic. If the level of toxicity varied in tandem with the level of lead, one could attribute the toxicity to the presence of lead. John Stuart Mill 1843. *A System of Logic*, vol. 1, p. 470.

10. Jenne (2015) provides another example of a complex theory. In particular notice the mixing of X and Y generalizations. Do a causal mechanism figure and/or write a Boolean equation for the theory.

> The argument developed in these pages is that conflict managers of *both* regimes were far more likely to achieve a successful outcome if powerful third parties acted to *nest the domestic disputes in a stable regional environment*. Put another way, cooperative (or soft-power) conflict management requires hard-power backing to achieve success. . . . I argue that effective civil conflict management requires building security from the outside in – that is, stabilizing the external environment before (or at the same time as) reconfiguring domestic institutions. If sectarian conflicts are securely "nested" in a stable regional and hegemonic environment, they are likely to become far more tractable; cooperative conflict management in particular is far likelier to meet with success. . . . *Failing* to securely nest sectarian struggles in a stable external environment may be setting peacemakers and peacekeepers up for an impossible task. . . . This book demonstrates that "nested security" is a key background condition for successful conflict reduction, regardless of the resources of the third party, the intervention strategy, and the attitudes of the government and minority representatives. . . . I demonstrate that nested security is effectively a precondition of success for cooperative conflict management. . . . To establish effective regional security regimes, three features must be present. First, a decision-making body must be established that can cope with internal veto players as well as member-state aggression. Second, a regularized, independent enforcement mechanism must be created with the full backing of regional hegemons and/or great powers. Finally, the regime must engender cooperation between monitors and enforcers to provide effective conflict management at all stages of the conflict – from preventive diplomacy to postwar reconstruction. (Jenne 2015, 9–10, 18, 45, 194)

11. Moravcsik's classic entitled *The choice for Europe: social preferences and state power from Messina to Maastricht*, has a complex design. The topic is European integration 1955–92: he covers 3 countries across 5 separate time periods, so that's 15 macro-cases. He then breaks each macro-case down into major issues, around 3–6 per case – not always the same, because different issues were relevant at different times – which are treated in parallel, analyzed in

the text, and reported in tables. So that is somewhere around 50–100 "mini-cases," with the smaller unit of analysis being the issue-country-period (i.e. one issue in one country across one period). The book then repeats that analysis for three theoretically sequential steps in producing interstate cooperation: national preference formation, interstate bargaining, international institutionalization. The resulting book contains 500 dense pages, much of it containing primary evidence. What are the possibilities of using LNQA to analyze this complex design – keeping in mind that LNQA typically only uses a small subset of the possible cases described above (150 issue-country-periods, across three dependent variables)? Could one do "causal mechanism case studies" and "generalization case studies"? (Thanks to Andy Moravcsik for the idea for this exercise as well as comments on it).

12. Downes and Monton (2013; see Downes (2021) for a fuller development with more dependent variables) explore if foreign imposed regime change results in democratization. Note that the dependent variable is the change in the polity democracy score 10 years later. A case analysis would be able to more clearly evaluate whether the foreign imposed regime change is responsible for this. According to them this typically means changing institutions and not just leaders, which itself is relatively rare in foreign imposed regime change, perhaps only 1/4 of the cases. Their table 2 lists the thirteen countries that experienced an attempted institutional FIRC [foreign-imposed regime change] at the hands of a democracy (p. 123). Discuss this as a potential *X* generalization LNQA analysis.

13. Slater and Wong (2022) provide a complex theory–causal mechanism of democratization and economic development in "developmental Asia." Draw a figure that includes all the critical components and how they combine to produce the outcome of democratization.

    See also various sections below including (1) Boolean models, (2) two-level theories, (3) Generalization–external validity, (4) within-case causal inference, (5) counterfactuals, (6) causal mechanisms–figures.

14. Core to LNQA are regularities. Often these can be complex and expressed as Boolean equations. This appears different guises in Weyland (2020) and Weyland (2023). These are consistent but expressed in different ways (thanks to Weyland for his clarification on this). His basic complex model given in his 2020 article is as follows: (IW AND HDE) OR (IW AND SCR AND ECR) OR (IW AND ECR) $\longrightarrow$ SDE where

    SDE – suffocation of democracy
    HIN – high instability
    PCH – susceptibility to para-legal change
    ECH – comparatively easy changeability in European parliamentarianism
    HWI – huge hydrocarbon windfalls
    ECR – economic crisis
    SCR – security crisis.

    Hence there are three paths to the outcome of democracy suffocation. In the same article he has a table, table **??** with these six factors but grouped into

two categories " Institutional weakness" (IW) and "Exogenous conjuncture" (EC). The LNQA book manuscript discusses this kind of complex model in the chapter on causal complexity (chapter 11) under the rubric of "two-level theories." Work through the the connection of the Boolean equation above and the implicit Boolean models given in the table below. Which possible combinations of these two general factors – IW and EC – does he *not* include in his final model? Why?

Table 8: Conditions for the suffocation of democracy by populist leaders

Table 1
Conditions for the suffocation of democracy by populist leaders

Institutional Weakness			Exogenous Conjuncture			Regime Outcome	Populist Leader
High Instability (HIN)	Paralegal Change (PCH)	Easy Change (ECH)	Economic Crisis (ECR)	Security Crisis (SCR)	Hydrocarbon Windfall (HWI)	Suffocation of Democracy (SDE)	
							Bolivarian Populism—Latin America
1	1	0	1	0	**1**	**1**	**Chávez – VEN**
1	1	0	0	0	**1**	**1**	**Morales – BOL**
1	1	0	0	0	**1**	**1**	**Correa – ECU**
1	1	0	1	0	0	0	(Moreno – ECU)
1	1	0	0	0	0	0	Gutiérrez – ECU
0	1	0	0	0	0	0	Zelaya – HON
0	1	0	0	0	0	0	Lugo – PAR
0	1	0	0	0	0	0	Humala – PER
0	1	0	1	0	0	0	N. Kirchner-ARG
0	1	0	0	0	0	0	C. Kirchner-ARG
1	1	0	0	0	0	0	Colom – GUA
							Neoliberal Populism—Latin America
0	**1**	0	**1**	**1**	0	**1**	**Fujimori – PER**
0	1	0	0	1	0	0	Uribe – COL
0	1	0	1	0	0	0	Menem – ARG
0	1	0	1	0	0	0	García – PER
0	1	0	0	0	0	0	Toledo – PER
0	1	0	0	0	0	0	Bucaram – ECU
0	1	0	1	0	0	0	Collor – BRA
1	1	0	0	0	0	0	Serrano – GUA
							Ethno-National Populism—Europe
0	0	**1**	**1**	0	0	**1**	**Orbán – HUN**
0	0	**1**	**1**	0	0	**1**	**Erdoğan – TUR**
0	0	1	0	0	0	0	Berlusconi – ITA
0	0	1	0	0	0	0	Mečiar – SVK
0	0	1	0	0	0	0	Fico – SVK
0	1	0	0	0	0	0	Borisov – BUL
0	0	1	0	0	0	0	Zeman/Babiš–CZE
0	1	1	0	0	0	0	Băsescu – ROM
0	0	1	0	0	0	0	Kaczyński – POL
0	0	1	0	0	0	0	Papandreou–GRE
0	0	1	0	0	0	0	Tsipras – GRE
0	**0**	**0**	**0**	**0**	**0**	**0**	**Trump—USA**

Source: Weyland 2020.

Answer: If you consider IS to be (HIN OR PCH OR ECH) and EC to be (ECR OR SCR OR HWI) and the high level model to be IW AND EC is sufficient for SDE then there 3*3 possible models combining IW and EC. Presumably the ones he did not include did not appear in any of the cases.

15. Multimethod LNQA becomes an important possibility particularly when the scope of the analysis is a medium-N or when the treatments are relatively rare. One example of this is the analysis of comprehensive peace agreements and civil wars of which there are. Karreth et al. argue that high economic leverage IGO membership contributes to the success of these agreements. The number of high leverage economic IGOs is modest including the World Bank, and other major international institutions ("IGOs with high economic leverage and institutionalized structures are a small subset of international organizations" p. 499). Discuss how one might identify the (1,1) causal mechanism cases based on the dependent variable of high levels of implementation. What would be the critical case studies to support their causal claims?

16. Below is one core figure for Levitsky and Way (2022) regarding social revolutions and the durability of authoritarian regimes. First, could or should one

interpret this via a Boolean equation and put ANDs, ORs, and S (sufficient for) or N (necessary condition) at appropriate places in the figures? Could you write all the relevant Boolean equations?

Figure 13: Ideal-typical path to durable authoritarianism

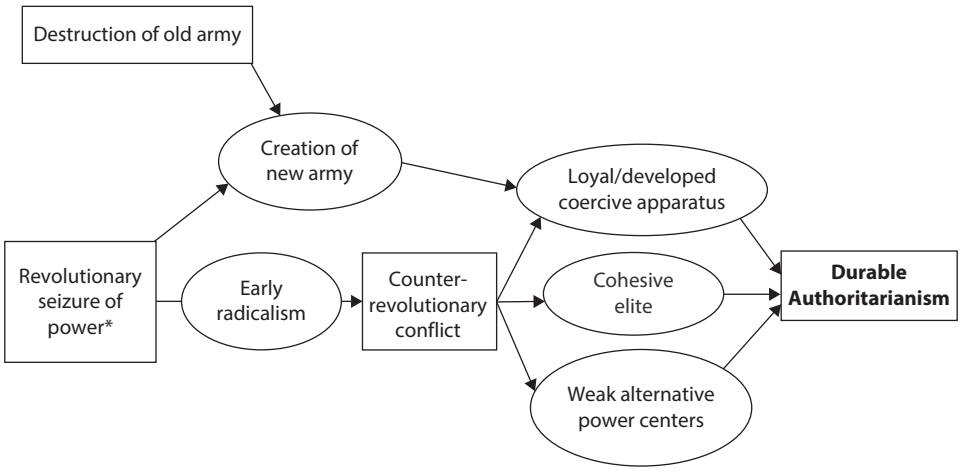


FIGURE 1.2: The Ideal-Typical Revolutionary Reactive Sequence.
* In some cases (e.g., China), the reactive sequence occurs before seizure of power.

Source: Levitsky and Way (2022).

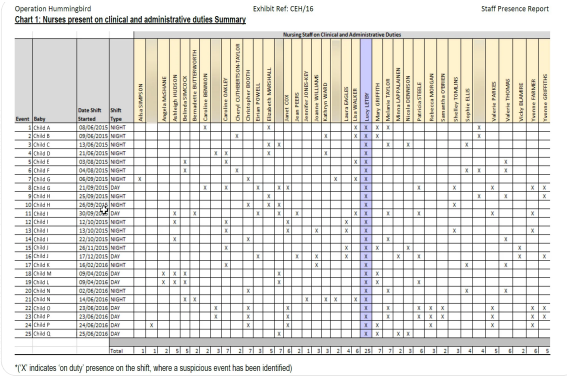
17. LNQA is fundamentally a medium-N methodology. Sometimes it seems that the number of observations excludes using it. But often there are scenarios where the research design reduces the actual number of observations to a manageable size. Pape (2003; 2005) in a controversial article and book talks about suicide terrorism. His data set, basically a Y generalization set up, has about 200 suicide terrorism attacks. Discuss how in practice when doing case studies this would easily fit within the realm of possibility for LNQA.

    Answer: It turns out that about 95 percent of these suicide terrorism attacks occur within about 16 to 17 suicide terrorism campaigns. When doing the case studies one would certainly focus on the campaigns because each of the suicide attacks is really embedded within the campaign. So in fact the unit of analysis is not suicide terrorism attack but rather fundamentally the organization which is using it as a standard tactic. The number of such organizations may well be a medium-N.

18. Core to the methodology of LNQA are regularities and within-case causal inference. In a court case that drew extensive media attention in the United Kingdom a nurse, Letby, was convicted of killing six babies and suspected of killing at least a couple dozen more. A huge part of the prosecution case rested on a regularity presented in table **??**. Note that the nurse was on duty for all the cases, a classic Y regularity. From the LNQA perspective this is

evidence but not conclusive. From the LNQA perspective what additional evidence must be presented to be convinced that she murdered these babies? Note that there was a long article in the *New Yorker* (Aviv 2024) about this case as well is a Royal Statistical Society report (2022) dealing with exactly this sort of legal situation. Aviv says: "She was the 'one common denominator.... Letby had become the country's most reviled woman – the unexpected face of evil,' as the British magazine *Prospect* put it – largely because of that unbroken line [regularity]" (Aviv 2024, 34).

Table 9: Regularities and the Letby conviction of the murder of numerous babies



Operation Hummingbird — Exhibit Ref: CEH/16 — Staff Presence Report

Chart 1: Nurses present on clinical and administrative duties Summary

Event	Baby	Date Shift Started	Shift Type	Nursing Staff on Clinical and Administrative Duties
1	Child A	08/06/2015	NIGHT	
2	Child B	09/06/2015	NIGHT	
3	Child C	13/06/2015	NIGHT	
4	Child D	21/06/2015	NIGHT	
5	Child E	03/08/2015	NIGHT	
6	Child F	04/08/2015	NIGHT	
7	Child G	06/09/2015	NIGHT	
8	Child G	21/09/2015	DAY	
9	Child H	25/09/2015	NIGHT	
10	Child H	26/09/2015	NIGHT	
11	Child I	30/09/2015	DAY	
12	Child I	12/10/2015	NIGHT	
13	Child I	13/10/2015	NIGHT	
14	Child I	22/10/2015	NIGHT	
15	Child J	28/11/2015	NIGHT	
16	Child J	17/12/2015	DAY	
17	Child K	16/02/2016	NIGHT	
18	Child M	09/04/2016	DAY	
19	Child L	09/04/2016	DAY	
20	Child N	02/06/2016	NIGHT	
21	Child N	14/06/2016	NIGHT	
22	Child O	23/06/2016	DAY	
23	Child P	23/06/2016	DAY	
24	Child P	24/06/2016	DAY	
25	Child Q	25/06/2016	DAY	
		Total		

*('X' indicates 'on duty' presence on the shift, where a suspicious event has been identified)

Answer: For LNQA the *causal* generalization rests mostly on the analysis of the death of the individual babies. Is there clear evidence that she killed them, on a case-by-case basis? In the *New Yorker* article, Aviv strongly suggest that in none of the cases was there strong evidence that she killed the baby. From the LNQA perspective causal inference rests on discussion of the six cases she was convicted for murder on.

19. Y generalizations are closely connected to selecting on the dependent variable concerns because one is selecting cases of Y equals one. Ever since King, Keohane, and Verba (1994) these have been controversial and the standard line is still one should never do that and one cannot learn anything from doing it. For example, Bueno de Mesquita Bueno de Mesquita and Fowler (2021) have a whole chapter on this topic in their textbook. Discuss the quote below by Pearl and McKenzie which is using the canonical smoking and lung cancer example. Notice how Pearl and McKenzie state that the probabilities are backwards, probability of X given Y. They claim that one knows nothing based on these data that would be relevant to a decision about smoking. Explain why that is false.

> Of course Hill knew that an RCT was impossible in this case, but he had learned the advantages of comparing a treatment group to a control group. So he proposed to compare patients who had already been diagnosed with cancer to a control group of healthy volunteers. Each group's members were interviewed on their past behaviors and medical histories. To avoid bias, the interviewers were not told who had cancer and who was

> a control. . . . The probability logic is backward too. The data tell us the probability that a cancer patient is a smoker instead of the probability that a smoker will get cancer. It is the latter probability that really matters to a person who wants to know whether he should smoke or not. The results of the study were shocking: out of 649 lung cancer patients interviewed, all but two had been smokers. (Pearl and Mackenzie 2017, 192)

Answer: The regularity that the research discovered is extremely strong, .97. So one might conclude that a necessary condition for lung cancer is smoking. Notice that this does not tell you a lot about what Pearl and McKenzie want to know which is what is the probability of lung cancer if one smokes. But it does tell you that if you do not want to get lung cancer you are almost certain not to get it if you do not smoke. Smoking is part of virtually all paths to lung cancer. Certainly a useful bit of information.

20. LNQA uses *Y* and *X* generalizations. Comparative methods work often refers back to Mill's classic book which has been a part of the methods discussion for decades. To which of Mill's methods do these most closely seem connected?

    Answer: *Y* generalizations match nicely onto Mill's method of agreement. However, *X* generalizations fit less well with any of them. Perhaps they fit best with the little-discussed "Method of concomitant variations." From the wikipedia page:

    "Whatever phenomenon varies in any manner whenever another phenomenon varies in some particular manner, is either a cause or an effect of that phenomenon, or is connected with it through some fact of causation." John Stuart Mill 1843. *A System of Logic*, vol. 1, p. 470.

    If across a range of circumstances leading to a phenomenon, some property of the phenomenon varies in tandem with some factor existing in the circumstances, then the phenomenon can be associated with that factor. For instance, suppose that various samples of water, each containing both salt and lead, were found to be toxic. If the level of toxicity varied in tandem with the level of lead, one could attribute the toxicity to the presence of lead.

21. Moravcsik's classic entitled *The choice for Europe: social preferences and state power from Messina to Maastricht*, has a complex design. The topic is European integration 1955–92: he covers 3 countries across 5 separate time periods, so that's 15 macro-cases. He then breaks each macro-case down into major issues, around 3–6 per case – not always the same, because different issues were relevant at different times – which are treated in parallel, analyzed in the text, and reported in tables. So that is somewhere around 50–100 "mini-cases," with the smaller unit of analysis being the issue-country-period (i.e. one issue in one country across one period). The book then repeats that analysis for three theoretically sequential steps in producing interstate cooperation: national preference formation, interstate bargaining, international institutionalization. The resulting book contains 500 dense pages, much of it containing primary evidence. What are the possibilities of using LNQA to analyze this complex design – keeping in mind that LNQA typically only uses a small subset of the possible cases described above (150 issue-country-periods, across

three dependent variables)? Could one do "causal mechanism case studies" and "generalization case studies"? (Thanks to Andy Moravcsik for the idea for this exercise as well as comments on it).

22. Downes and Monton (2013; see Downes (2021) for a fuller development with more dependent variables) explore if foreign imposed regime change results in democratization. Note that the dependent variable is the change in the polity democracy score 10 years later. A case analysis would be able to more clearly evaluate whether the foreign imposed regime change is responsible for this. According to them this typically means changing institutions and not just leaders, which itself is relatively rare in foreign imposed regime change, perhaps only 1/4 of the cases. Their table 2 lists the thirteen countries that experienced an attempted institutional FIRC [foreign-imposed regime change] at the hands of a democracy (p. 123). Discuss this as a potential *X* generalization LNQA analysis.

23. Core to LNQA are regularities. Often these can be complex and expressed as Boolean equations. This appears different guises in Weyland (2020) and Weyland (2023). These are consistent but expressed in different ways (thanks to Weyland for his clarification on this). His basic complex model given in his 2020 article is as follows: (IW AND HDE) OR (IW AND SCR AND ECR) OR (IW AND ECR) ⟶ SDE where

    SDE – suffocation of democracy
    HIN – high instability
    PCH – susceptibility to para-legal change
    ECH – comparatively easy changeability in European parliamentarianism
    HWI – huge hydrocarbon windfalls
    ECR – economic crisis
    SCR – security crisis.

    Hence there are three paths to the outcome of democracy suffocation. In the same article he has a table, table **??** with these six factors but grouped into two categories " Institutional weakness" (IW) and "Exogenous conjuncture" (EC). The LNQA book manuscript discusses this kind of complex model in chapter on causal complexity (chapter 11) under the rubric of "two-level theories." Work through the the connection of the Boolean equation above and the implicit Boolean models given in the table below. Which possible combinations of these two general factors – IW and EC – does he *not* include in his final model? Why?

Table 10: Conditions for the suffocation of democracy by populist leaders

Table 1
Conditions for the suffocation of democracy by populist leaders

Institutional Weakness			Exogenous Conjuncture			Regime Outcome	Populist Leader
High Instability (HIN)	Paralegal Change (PCH)	Easy Change (ECH)	Economic Crisis (ECR)	Security Crisis (SCR)	Hydrocarbon Windfall (HWI)	Suffocation of Democracy (SDE)	
							Bolivarian Populism—Latin America
1	1	0	1	0	**1**	**1**	**Chávez – VEN**
1	1	0	0	0	**1**	**1**	**Morales – BOL**
1	1	0	0	0	**1**	**1**	**Correa – ECU**
1	1	0	1	0	0	0	(Moreno – ECU)
1	1	0	0	0	0	0	Gutiérrez – ECU
0	1	0	0	0	0	0	Zelaya – HON
0	1	0	0	0	0	0	Lugo – PAR
0	1	0	0	0	0	0	Humala – PER
0	1	0	1	0	0	0	N. Kirchner-ARG
0	1	0	0	0	0	0	C. Kirchner-ARG
1	1	0	0	0	0	0	Colom – GUA
							Neoliberal Populism—Latin America
0	**1**	0	**1**	**1**	0	**1**	**Fujimori – PER**
0	1	0	0	1	0	0	Uribe – COL
0	1	0	1	0	0	0	Menem – ARG
0	1	0	1	0	0	0	García – PER
0	1	0	0	0	0	0	Toledo – PER
0	1	0	0	0	0	0	Bucaram – ECU
0	1	0	1	0	0	0	Collor – BRA
1	1	0	0	0	0	0	Serrano – GUA
							Ethno-National Populism—Europe
0	0	**1**	**1**	0	0	**1**	**Orbán – HUN**
0	0	**1**	**1**	0	0	**1**	**Erdoğan – TUR**
0	0	1	0	0	0	0	Berlusconi – ITA
0	0	1	0	0	0	0	Mečiar – SVK
0	0	1	0	0	0	0	Fico – SVK
0	1	0	0	0	0	0	Borisov – BUL
0	0	1	0	0	0	0	Zeman/Babiš–CZE
0	1	1	0	0	0	0	Băsescu – ROM
0	0	1	0	0	0	0	Kaczyński – POL
0	0	1	0	0	0	0	Papandreou–GRE
0	0	1	0	0	0	0	Tsipras – GRE
0	**0**	**0**	**0**	**0**	**0**	**0**	**Trump—USA**

Source: Weyland 2020.

Answer: If you consider IS to be (HIN OR PCH OR ECH) and EC to be (ECR OR SCR OR HWI) and the high level model to be IW AND EC is sufficient for SDE then there 3*3 possible models combining IW and EC. Presumably the ones he did not include did not appear in any of the cases.

24. Multimethod LNQA becomes an important possibility particularly when the scope of the analysis is a medium-N and when the treatments are relatively rare. One example of this is the analysis of comprehensive peace agreements and civil wars. Karreth et al. argue that high economic leverage IGO membership contributes to the success of these agreements. The number of high leverage economic iGOs is modest including the World Bank, and other major international institutions ("IGOs with high economic leverage and institutionalized structures are a small subset of international organizations" p. 499). Discuss how one might identify the (1,1) causal mechanism cases based on the dependent variable of high levels of implementation. What would be the critical case studies to support their causal claims?

25. For much qualitative work, and for many theories in general, it is useful to give a Boolean equation which represents the structure of the argument. Give a Boolean equation that describes the core argument of the influential Levitsky and Way book on competitive-authoritarian regimes (2010).

    Answer: Here is Mahoney's Boolean description of Levitsky and Way: "For example, Levitsky and Way's (2010) theory of the causes of stable authoritarianism can be summarized as follows: (W AND (O OR L) $\Rightarrow$ S. In the equation, W is high Western linkage, O is high organizational power, L is high Western leverage, S is stable authoritarianism, ... and $\Rightarrow$ is both sufficiency

and temporal order. The equation says, 'not-high Western linkage and either high organizational power or not-Western leverage (or both) is sufficient for stable authoritarianism." (Mahoney and Rodríguez-Caceres 2022, 10–11)

26. Core to LNQA is the crafting of causal mechanism figures. Use the guidelines of the LNQA book (chapter 5 on causal mechanisms) to redo the Levitsky and Way, see the figure below. For example, how to interpret of some of the ambiguities, e.g., one Y variable for the factors on the right in their figure. Another question is whether it is one Y with different values or different Ys.

Figure 14: Revolutionary regime paths

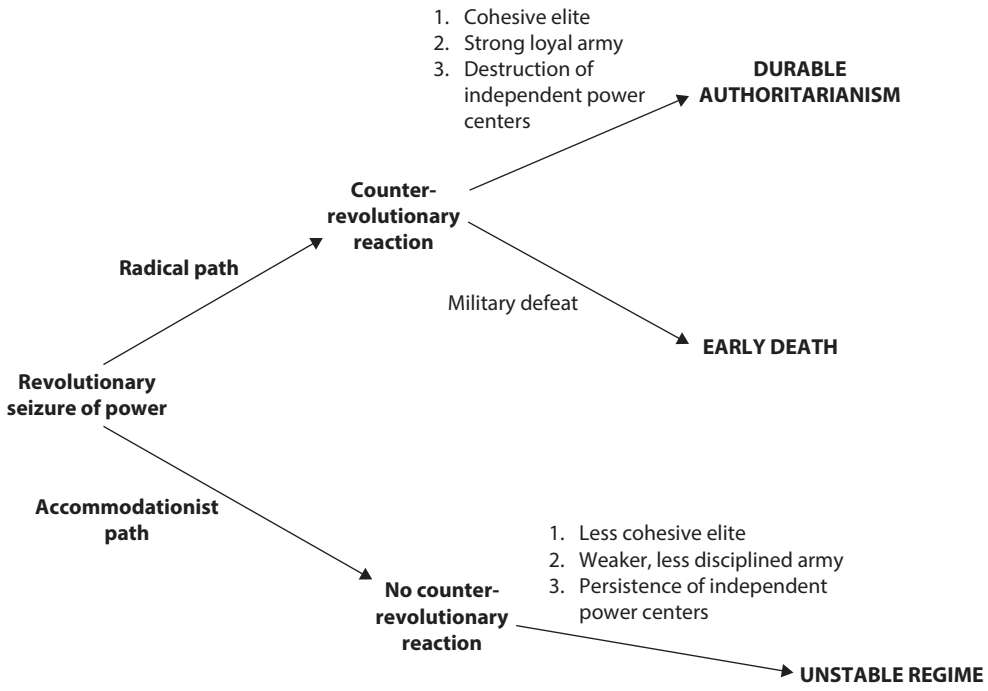


FIGURE 1.3: Three Revolutionary Regime Paths.

Source: Levitsky and Way (2022).

27. LNQA is fundamentally a medium-N methodology. Sometimes it seems that the number of observations excludes using it. But often there are scenarios where the case study research design reduces the actual number of observations to a manageable size. Pape (2003; 2005) in a controversial article and book talks about suicide terrorism. His data set, basically a Y generalization set up, has about 200 suicide terrorism attacks. Discuss how in practice when doing case studies this would easily fit within the realm of possibility for LNQA.

    Answer: It turns out that about 95 percent of these suicide terrorism attacks occur within about 16 to 17 suicide terrorism campaigns. When doing the case studies one would certainly focus on the campaigns because each of

the suicide attacks is really embedded within the campaign. So in fact the unit of analysis is not suicide terrorism attack but rather fundamentally the organization which is using it as a standard tactic. The number of such organizations may well be a medium-N.

28. A key issue in case study research is choosing cases in the context of concerns about generalization. Discuss Alter's rationale for choosing cases given her concern over generalizing beyond northern international courts. Would LNQA be applicable here, if so how?

> Although there are many European cases one could examine, only three of the eighteen case studies in this book focus on European legal institutions because I want to show that the new terrain of international law exists beyond Europe. Because I prefer less likely cases, I focus on human rights courts from Latin America and Africa, and cases where ICs with economic subject matter jurisdiction end up speaking to human rights issues, rather than a case study involving the European Court of Human Rights. (Alter 2014, 25)

29. Simmons and Smith (2021) and contributors to the anthology discussed the concept of "noncontrolled comparison." The authors and contributors typically are working from an interpretivist framework. Discuss the extent to which LNQA would also count as a noncontrolled comparison. See a related question in the concept section, page 133.

    Answer: LNQA is certainly a noncontrolled methodology. Whether it counts as a comparison is debatable since the within case analyses are conducted without explicit comparison to other cases.

30. In a very interesting comparative study Barany (2016) explores the role of the military when faced with revolutionary movements both violent and nonviolent. He claims that the support of the military is a necessary condition, Y generalization, for success in the third chapter but does not specifically address that in the empirical data analysis across the 11 cases. To what extent does his empirical analysis support this general claim?

> No institution matters more to a states survival than its military, and no major uprising within a state can succeed without the support or at least the acquiescence of the armed forces. This is not to say that the armys backing is *sufficient* to make a successful revolution; indeed, revolutions require so many political, social, and economic forces to line up just right, and at just the right moment, that revolutions rarely succeed. But support from a preponderance of the armed forces is a *necessary* condition for revolutionary success. I make two central arguments. The first is that *the response of the regimes regular armed forces to an uprising is critical to the success or failure of that uprising.* .... In sum, though the outcome of a rebellion is nearly always determined by the states coercive agencies – whether they defend the state or support the rebels – few writers on revolutions give the military its due and treat its part in this or that revolution with the attention and sensitivity to nuance it deserves. (Barany 2016, 5, 8)
>
> For easy reference, I have compiled two comprehensive tables to demonstrate the comparative weight of each independent variable as they relate

> to the eleven case studies, and the relative difficulty of predicting the generals responses to revolutions. The values assigned to the individual factors are the inductive result of the case studies. Obviously, the deeper our knowledge about a particular case, the less arbitrary these values will be, but, ultimately, they remain debatable or, put differently, falsifiable. (Barany 2016, 165, systematic within-case causal inferences for a bunch of independent variables; "Top six factors are in bold. Scale (06): 0 = irrelevant, not a factor 1 = of trivial importance 2 = of little importance 3 = somewhat important 4 = quite important 5 = very important 6 = critical, decisive") (Barany 2016, 265)

31. Roterg (2017) provides a nice example of a research project that has strong connections with LNQA. Discuss how he moves to the tails to choose his cases. How does he conduct a set of basically Y generalization analyses? For example,

> The sixth chapter of the book examines the fortunes of the five nation-states that showed the largest scoring gains in reducing corruption between 2004 and 2014. It explains how they changed from horribly corrupt to markedly less corrupt over that decade, according to the Corruption Perceptions Index and the World Bank's Control of Corruption indicator. This "most improved" cohort includes Georgia, Liberia, Rwanda, Macedonia, and Montenegro. (Rotberg 2017, 13)

> The Gift of Political Will and Leadership Prior chapters assert that exertions of political will – the leadership factor – are necessary (but not always sufficient) to curb corruption within nation-states. Developing states (and the Nordic and Antipodean exceptions) that successfully destroyed the incubus of corruption within their borders are (or were) run by strong leaders determined to prevent their governmental colleagues from abusing political or bureaucratic power for private gain. They dramatically altered their countries' inherited political cultures so as to eliminate rent seeking, contract fiddling, and influence peddling. In every case, those leaders are acting (or did act) not because of moral scruples but primarily because they fully understand how corrupt behavior destroys developmental prospects, envelops their nascent nations in a moral haze antithetical to economic growth and social betterment, and undermines a regime's (and their own) legitimacy. (Rotberg 2017, 223)

32. Discuss the extent to which Dunning (2008) constitutes a partial LNQA. Is he looking for X generalizations or Y generalizations. What would constitute the scope of analysis for doing all cases? Explain why the number of cases cannot be too large and hence a possible LNQA.

> It therefore makes sense to select cases for which values on key independent variables predict a relatively democratic effect of resource rents. . . . As just discussed, two such independent variables emerged in the theoretical and empirical analyses of previous chapters as central: the extent of resource dependence and the degree of inequality in non-resource sectors of the economy. In particular, the previous analysis suggests the value of selecting for intensive case-study analysis those resource-rich rentier states in which resource dependence is low and private, non-resource inequality is high. In this chapter, I conduct an in-depth analysis of historical and contemporary evidence from Venezuela, where I conducted the most

> extensive fieldwork; in Chapter Six, I develop complementary analyses of Chile, Bolivia, Ecuador, and Botswana. These cases offer several advantages for my purposes. First and most important, all are resource-rich rentier states and meet the two key conditions of relatively low resource dependence and relatively high private, non-resource inequality. . . . In the next chapter, I extend the purview beyond the Latin American context, discussing case-study evidence from Botswana; in the conclusion, I discuss comparative issues further. (Dunning 2008, 150)

Answer:

It is an *X* generalization: if resource dependence is low =1 AND high inequality in non-resource sectors = 1 then democracy. Scope is resource-rich rentier states. Hence the N for this *X* generalization cannot be large.

Generalization, external validity, scope, etc.

33. In case study research designs the scope of the theory-empirical analysis is often clear when a few cases are selected for empirical analysis. Discuss what the are scope conditions for Schwartz's analysis. How do the cases selected connect to the scope conditions?

> Case Selection. Within the Central American context, I analyze three institution-level cases: the customs administration and policing institutions in Guatemala and land tenure in Nicaragua. Each of these cases demonstrates the wartime emergence of alternative rules of the game that have similarly undermined state performance. These specific cases were selected for several reasons. First, they each reflect distinct domains of state activity (extractive, coercive, and administrative), allowing me to assess the common building blocks that characterize processes of state institutional change broadly. Second, by comparing the Guatemalan and Nicaraguan contexts, I leverage variation on core political and structural factors, which could plausibly shape processes of institutional development in different ways. Specifically, the two countries differ widely in terms of civil war antecedents, regime ideologies, war-making strategies, and the nature of foreign support. The comparison thus provides a critical opportunity to explore why we observe the emergence of similar kinds of institutional logics amid such stark differences. . . . This book focuses on three cases of undermining rules, above all, because it seeks to identify the mechanisms underlying the relationship between civil war and"weak" states. (Schwartz 2023, 27)

34. A important question in causal model diagraming and statistical models is whether a given factor is a scope condition or a moderator. In many setting a given factor seems like it could be either. For example, discuss this question in the context of Slough and Tyson's discussion of external validity. The same issue arises in Esterling et al.'s discuss of external validity.

> Empirical targets depend on at least three critical ingredients of a constituent study. First, a study takes place in a setting, represented by $\theta \in \Theta$, and captures all the contextual features, as well as characteristics of the study population, that determine the mechanism's effect, or quantitative

> influence. The setting can include both measured and unmeasured features or characteristics that are relevant to how the mechanism's influence arises and how it is measured. (Slough and Tyson 2024, 12)
>
> We say a claim is externally valid to the extent one has accurately specified why or how the effect generalizes across settings. This means specifying the range of conditions in settings γ across which the effect generalizes. Hence, we define external validity as the correct specification of the conditions that enable/disable or augment/weaken a causal effect. External validity is present when that specification is true – i.e., when γ is correctly labeled. (Esterling et al. 2025, 13)

35. Necessary conditions as causes are often connected to scope conditions which are also often causes. Discuss the following claim and whether it should be formulated as a scope condition:

> There is another reason why the notions of difference-maker and cause come apart. This is where we have a necessary condition for something to occur but where it is not really a causal contributor to it. For example, having a liver is a necessary condition for having cirrhosis. No one could have cirrhosis unless they had a liver; therefore having a liver is a difference-maker with respect to that disease. But it would seem wrong to say that the liver caused the disease even though it is a necessary condition for having it. (Anjum and Mumford 2018, 127)

36. Core to LNQA is the determination of the basic data set and then within that the establishment of LNQA scope. Typically, LNQA scope is a subset of the basic data set. This means there are potentially multiple basic datasets which have the same LNQA scope subset subset. Sechser and Fuhrmann's analysis of the success of nuclear compellence efforts, discussed in more detail in chapter 7 of the LNQA book , provides an example of an X generalization of the form "if nuclear compellence ($X = 1$), then it will not succeed ($Y = 1$, coded as compellence failure). This provides an interesting example where there are multiple potential basic datasets. An X generalization involves a subset of the basic dataset. It is quite possible that two different basic datasets have the same X generalization subset. This is just a Venn diagram with two circles that have an intersection: the X generalization is that intersection. But the choice of the basic dataset is critical because it determines the boundary between internal and external validity and is hence consequential for the the overall research program.

    To construct their basic dataset they start with an existing dataset of "Militarized Compellent Threats" which contains 210 interstate compellent threat initiator-target dyads over the years 1918–2001. In virtually all basic datasets the author must make explicit spatial and temporal boundaries. Here the temporal boundaries are 1918–2001 and the spatial boundaries include all dyadic *interstate comepllent threats*. The LNQA dataset could be derived by coding these cases for whether they involved nuclear compellence (as opposed to conventional compellence for example). The LNQA generalization only involves $X = 1$ cases, however, and out of those compellent threats only 19 are $X = 1$ cases of attempted *nuclear* compellence.

A second option for defining the basic dataset is to take all MIDs involving nuclear powers and code whether or not that MID involved nuclear compellence. This generates the same set of $X = 1$ cases. However, clearly this basic dataset does not include nonnuclear powers and thus has no cases before 1945 among the $X = 0$ cases. Both these basic datasets choose the same $X = 1$ cases, but differ significantly on the $X = 0$ cases.

Discuss whether this potential difference in the basic data set has important implications. For example, typically external validity is determined by the basic data set, so does that matter in this particular case?

Answer: These choices have implications for the overall causal generalization, which involves the whole basic dataset. The second option makes more sense because the causal generalization goal is about the behavior of *nuclear* states in militarized disputes. A good test between the two datasets is the counterfactual of going from $X = 0$ to $X = 1$. In the second basic dataset this makes complete sense: only a nuclear power could have decided on adopting a strategy of nuclear compellence. In the first dataset it does not make sense for the vast majority of states because making a nuclear compellent threat is not possible for states that do not have nuclear weapons to begin with.

37. Generalization is core to LNQA and usually a central concern for case studies research in general. If one is only going to do a few cases in a case study research design then the selection of cases becomes critical in terms of generalization. In particular what is the role of other variables, Z, the in the selection of these cases (discussed at length in Goertz 2017). The popular paired comparison which follows matching looks for cases with the same values for Z. So implicitly generalization is limited to the range of those Z that match which could be very narrow or problematic if only 2–3 pairs are chosen. Generalization is however not a criterion for paired designs. However, if one is interested in generalization then the logic suggests choosing cases which are very different on Z the to show that X is a cause of Y in a variety of settings. Discuss how Stuart chooses cases with generalization in mind. Here is a sample of her case study selection discussion:

> Beyond the ELF, EPLF, and SPLM/A, the next two cases demonstrate that the imitation of the Chinese model of governance is not simply a function of communist ideology, but a strategy for rebels with more transformative ambitions regardless of ideology. I select two cases that differed in almost all ways but their more transformative goals and their non-communist ideology. Furthermore, I selected cases outside of Africa and East Asia to highlight that the imitation of the Chinese model is not the result of contagion from a small slice of conflicts. . . . The quantitative portion offers a final test of the generalizability of my proposed claims. (Stewart 2021, 19–20)

38. A core feature of LNQA is the use of "generalization case studies." These are less detailed than than the core case studies but nevertheless examine whether the mechanism question is at work in additional cases within the scope. Sometimes people call these "plausibility probes" as illustrated in the quote below. Discuss the terminological issues between calling them

generalization case studies versus plausibility probes. The terminology does suggest different goals.

> This chapter reaches beyond China, South Korea, and Taiwan to examine anti-corruption efforts in a diverse set of authoritarian regimes: Cuba, Malaysia, Rwanda, Singapore, and Vietnam. I use short case studies of these regimes as "plausibility probes" for the broader applicability of my theory in different regions and in different types of authoritarian systems. (Carothers 2022, 197)

39. It appears that frequently the concept of external validity and the concept of generalization are considered synonyms, as in the quote below by Druckman. Discuss the pros and cons of this practice. If they should be conceptually different, if so what are the differences?

> [E]xternal validity concerns the confidence one has that the cause–effect relationship "holds over variation in persons, settings, treatment variables, and measurement variables." Put another way, external validity refers to generalizability. . . . The first question one must ask when assessing external validity is "what is being generalized?". . . External validity concerns then come down to assessing whether features of the experimental sample, context, treatment, and measure preclude generalizing the relationship beyond the single study. . . . Table 3.2. Generalization questions. (Druckman 2021, 61, 62, 68)

Answer: In LNQA generalization is within the scope. Cases outside the scope are considered external validity. Because scope must be well-defined to do LNQA, external validity has a much clearer meaning. But in general, for example, it is not clear if moving along the continuum of X, treatment, means generalization or external validity. Most people seem to think of external validity in terms the subjects or unit of analysis, e.g., in a book looking at all Latin American cases, external validity in the final chapter is discussing countries from other regions.

40. Generalization is core to LNQA and usually a central concern for case studies research in general. If one is only going to do a few cases in a case study research design and selection of cases becomes critical in terms of generalization. In particular what is the role of other variables, Z, the in the selection of these cases (discussed at length in Goertz 2017). The popular paired comparison which follows matching looks for cases with the same values for Z. So implicitly generalization is limited to the range of those Z that match which could be very narrow or problematic if only 2–3 pairs are chosen.z Generalization is however not a criterion for paired designs. However, if one is interested in generalization then the logic suggests choosing cases which are very different on Z the to show that X is a cause of Y in a variety of settings. Discuss how Stuart chooses cases with generalization in mind. Here is a sample of her case study selection discussion:

> Beyond the ELF, EPLF, and SPLM/A, the next two cases demonstrate that the imitation of the Chinese model of governance is not simply a function of communist ideology, but a strategy for rebels with more transformative ambitions regardless of ideology. I select two cases that differed in almost

> all ways but their more transformative goals and their non-communist ideology. Furthermore, I selected cases outside of Africa and East Asia to highlight that the imitation of the Chinese model is not the result of contagion from a small slice of conflicts. . . . The quantitative portion offers a final test of the generalizability of my proposed claims. (Stewart 2021, 19–20)

41. A core feature of LNQA is the use of "generalization case studies." These are less detailed than than the core case studies but nevertheless examine whether the mechanism question is at work in additional cases within the scope. Sometimes people call these "plausibility probes" as illustrated in the quote below. Discuss the terminological issues between calling them generalization case studies versus plausibility probes. The terminology does suggest different goals.

> This chapter reaches beyond China, South Korea, and Taiwan to examine anti-corruption efforts in a diverse set of authoritarian regimes: Cuba, Malaysia, Rwanda, Singapore, and Vietnam. I use short case studies of these regimes as "plausibility probes" for the broader applicability of my theory in different regions and in different types of authoritarian systems. (Carothers 2022, 197)

42. It appears that frequently the concept of external validity and the concept of generalization are considered synonyms, as in the quote below by Druckman. Discuss the pros and cons of this practice. If they should be conceptually different, if so what are the differences?

> [E]xternal validity concerns the confidence one has that the cause–effect relationship "holds over variation in persons, settings, treatment variables, and measurement variables." Put another way, external validity refers to generalizability. . . . The first question one must ask when assessing external validity is "what is being generalized?". . . . External validity concerns then come down to assessing whether features of the experimental sample, context, treatment, and measure preclude generalizing the relationship beyond the single study. . . . Table 3.2. Generalization questions (Druckman 2021, 61, 62, 68)

Answer: In LNQA generalization is within the scope. Cases outside the scope are considered external validity. Because scope must be well-defined to do LNQA, external validity has a much clearer meaning. But in general, for example, it is not clear if moving along the continuum of X, treatment, means generalization or external validity. Most people seem to think of external validity in terms the subjects or unit of analysis, e.g., in a book looking at all Latin American cases, external validity in the final chapter is discussing countries from other regions.

43. A key issue in case study research is choosing cases in the context of concerns about generalization. Discuss Alter's rationale for choosing cases given her concern over generalizing beyond northern international courts. Would LNQA be applicable here, if so how?

> Although there are many European cases one could examine, only three of the eighteen case studies in this book focus on European legal institutions because I want to show that the new terrain of international law exists beyond Europe. Because I prefer less likely cases, I focus on human rights courts from Latin America and Africa, and cases where ICs with economic subject matter jurisdiction end up speaking to human rights issues, rather than a case study involving the European Court of Human Rights. (Alter 2014, 25)

44. Often one claims that case studies should be representative of the population (e.g., Gerring 2006). Discuss what might constitute a " representative" case. Is that what Gavin is choosing? Discuss his view of generalizing from case studies.

> As a historian interested in these questions, my way of assessing Sechser, Furhmann, and Kroenig's arguments is straight-forward. I would identify the most important example where these issues are engaged, look at the primary documents, see how the authors 'coded crucial variables and determine how good a job their analysis does in helping us understand both the specific crisis itself and the larger issues driving nuclear dynamics. Political scientists might describe this as running both a 'strong' and a 'critical' test; in other words, *if the authors' theories don't fully explain the outcomes and causal mechanisms in the most important and most representative case*, how useful are the findings in explaining the broader issues? Is there such a case? Is there such a case? In a speech on November 10th, 1958, Soviet Premier Nikita Khrushchev demanded the Western powers – the United States, Great Britain, and France – remove their military forces from West Berlin within six months. This ultimatum was the start of a tense, four-year period that many believe brought the world closer to thermonuclear war than any time before or since, culminating in the Cuban Missile Crisis of October 1962. According to a leading historian of post-war international politics, "(T)he great Berlin crisis of 1958 to 1962" was "the central episode of the Cold War."16 And as McGeorge Bundy states, "there were more than four years of political tension over the future of Berlin. . . . Khrushchev's Berlin crisis gives us what is otherwise missing in the nuclear age: a genuine nuclear confrontation in Europe." . . . Most importantly, I would want to be convinced that the causal mechanisms identified by the authors did in fact drive the origins, development, and outcome of this crisis. . . . And if these models can't tell us anything about arguably the most important and consequential nuclear standoff in history, should I take comfort that it apparently can explain why the U.S. successfully restored Jean-Bertrand Aristide in 1994 or "won" in Nicaragua in 1984? Or look at the 1983 'Able Archer' affair, perhaps the most recent case where the risk of thermonuclear was possible (if highly unlikely). (Gavin 2014, 16, 24 emphasis is mine)

45. Discuss the claim below. Is it true that matched cases are more representative of the population than unmatched cases? What is the implied definition of "representative" in his argument? What about scenarios where there are many more unmatched cases that matched ones? What if the data are highly skewed (as much social science data are)?

> And since treated cases are dropped from the dataset if no appropriate matches are found, it ensures that selected cases are more representative

> of average belligerents and wars than if outliers with few or no matches were chosen. (Lyall 2020, 32)

46. The notion of "conceptual replication" is wide-spread in psychology. In fact, conceptual replication is in fact theoretical generalization. Discuss the merits of this notion. Chambers provides an example of someone who thinks it is a bad idea because it is not actual replication of the original experiment.

> Although used widely in psychology, the term conceptual replication does not feature in the scientific method of other disciplines. In fact, the term itself is misleading because conceptual replications dont actually replicate previous experiments; they instead assume (rather than test) the truth of the findings revealed in those experiments, infer the underlying cause, and then seek converging evidence for that inferred cause using an entirely different experimental procedure. Viewed within the framework of the HD scientific method, this process can be thought of as extrapolating from a body of findings to refine theory and generate new hypotheses (see figure 3.1). As important as this step is, it depends first and foremost on the reliability of the underlying evidence base. To rely solely on extrapolation at the expense of direct replication is to build a house on sand. (Chambers 2017, 49)

47. Sechser and Fuhrmann 2017 is an excellent example of large-N qualitative analysis. Discuss their use of case studies first the (1,1) cases then the set of cases that potentially falsify their theory:

> The previous chapter analyzed clear failures of nuclear coercion. We showed that in six serious nuclear crises, and three less severe cases, countries failed to coerce their adversaries despite engaging in nuclear brinkmanship. In this chapter, we turn our attention to cases in which nuclear blackmail seemingly worked. These crises are apparent outliers for nuclear skepticism theory and fertile ground for evidence in support of the nuclear coercionist school. Indeed, the leaders involved in these episodes – as well as many scholars who have studied them – asserted that nuclear weapons were useful for military blackmail. . . . We critically assess ten crises that are widely regarded as successful instances of nuclear coercion, focusing our attention on the seven most serious cases. None of these crises provides unequivocal evidence that nuclear coercion works. In each instance, at least one of three factors mitigates the conclusion that a nuclear threat resulted in a coercive victory. First, factors other than nuclear weapons often played a significant role in states' decisions to back down. Second, on close inspection, some crisis outcomes were not truly "victories" for the coercer. Third, when nuclear weapons have helped countries in crises, they have aided in deterrence rather than coercion. (Sechser and Fuhrmann 2017, 172–73)

48. A key issue in many experiments is extrapolation to the population of interest. Discuss Lupton's discussion of these issues:

> Furthermore, our theories about how people think should be broadly applicable across individuals from a variety of backgrounds. Researchers can also employ sample populations that mirror the demographics of policymakers in certain ways. The samples I use skew male, as do world leaders more generally, and are rather highly educated (with almost 90 percent

> of participants in my surveys having some form of a college education-compared with 80 percent of world leaders). The majority of participants in my study are also interested in international affairs and international events. While I do not claim that my sample populations perfectly represent the demographics of policymakers, the participants in my study are not as uneducated or uninterested in politics as one might initially think. Moreover, I randomly assign participants to different treatment groups, which helps address the issue of external validity. (Lupton 2019, 52)

49. Concept generalization across groups is a core methodological issue. Discuss the sense of "meaning" in the abstract and how it relates to generalization of a concept across groups.

> Measurement invariance assesses the (psychometric) equivalence of a construct across groups or measurement occasions and demonstrates that a construct has the same meaning to those groups or across repeated measurements. . . . Appropriate and proper comparison of a construct between groups or across times, therefore, depends first on ensuring equivalence of meaning of the construct. The untoward consequences of measurement noninvariance can be readily illustrated in the study of depression in men and women. Suppose frequency of crying, weight gain, and feelings of hopelessness are indicative of the severity of depression in women, but only feelings of hopelessness are indicative of the severity of depression in men. If the three indicators are combined into a scale to compare depression in women and men, mean differences on the scale may mislead because crying and weight gain have little relation to depression in men. In this example, men may score lower than women on the depression scale because they cry less and gain less weight. (Putnick and Bornstein 2016, 71)

50. Discuss case selection in Stasavage's influential book *Public debt and the birth of the democratic state France and Great Britain, 16881789*. Clearly he is using the very popular paired comparison, in terms the causal mechanism book a (0,0) case compared to a (1,1) case. There are lots of nondemocracies in this. Why choose France? What about choosing more (1,1) cases to generalize?

> Representative political institutions may improve a government's ability to make credible commitments through several different mechanisms. This chapter has used a formal model of legislative bargaining to provide support for my three main arguments. I first demonstrated that if capital owners are in the minority, then party formation can lead to credible commitment, but only if players bargain over multiple issues. In addition, one can expect the perceived credibility of taxation or borrowing to vary according to the partisan composition of government. Both of these observable implications will be considered in detail in subsequent chapters. I next showed that constitutional checks and balances will have little effect on credibility unless there is some mechanism ensuring that capital owners control a veto point. This helps support the argument that multiple veto points may in many cases be insufficient to ensure credible commitment. Finally, I developed my argument about bureaucratic delegation, suggesting that it will improve credibility only if capital owners have the political authority to block any attempt to override bureaucratic decisions. This too is an empirical prediction that is considered in subsequent chapters. (Stasavage 2003, 49–50)

51. Large-N qualitative analysis can become particularly attractive when there is an interaction term and events that are quite rare for both parts of the interaction. Analyze how this might be possible following hypothesis:

> Hypothesis: Salient territorial threats make mass killing more likely to occur only when leaders espouse exclusionary ideologies. Otherwise, territorial threats have no effect on mass killing. (Hong and Kim 2019, 355)

52. Discuss Alter's rationale for her generalization case studies: "Although there are many European cases one could examine, only three of the eighteen case studies in this book focus on European legal institutions because I want to show that the new terrain of international law exists beyond Europe. Because I prefer less likely cases, I focus on human rights courts from Latin America and Africa, and cases where ICs with economic subject matter jurisdiction end up speaking to human rights issues, rather than a case study involving the European Court of Human Rights." (Alter 2014, 25)

53. Shannon 2009 is a case study of the US invasion of Panama in 1989. If one were to do a large-N qualitative analysis starting from this article what would be the basic case selection criteria using figure 2. Should you consider case selection further down the tree? Why is that important?

    Answer:

    The top of the decision tree is those leaders who are motivated to violate the norm. These are the cases where we can see the norm potentially in action.

54. King and Zeng (2007) have presented a large-N argument about counterfactuals. Basically the article deals with the perils of extrapolation and a technique for deciding when a counterfactual is an extrapolation outside the data. Basically, counterfactuals within the data are more reliable than extrapolations outside it: a bit of conventional wisdom. One conclusion from this methodology is that one should collect more data to increase the region covered by the data; a conclusion not too surprising from a statistical methods point of view. Use the Possibility Principle cube to think about these issues.

    (1) It is quite possible for counterfactuals within the boundaries of the data to be quite distant from actual data points. Illustrate this in the cube.

    (2) Use the cube to show how counterfactuals outside the data might be closer to actual data points than points in the data cube.

    (3) Use the democracy and war literature to illustrate how interpolation might be more problematic than extrapolation.

    (4) Applying the Possibility Principle principle usually means reducing the size of the sample space, while King and Zeng implicitly argue for increasing it. However, both argue that causal or counterfactuals are best made with contrasting cases that are "close" to each other. Discuss.

    Answer:

    For (1) and (2) put points at the corners, and a few outside the corners and leave the middle of the cube empty.

The democratic peace literature is paralleled by a small "autocratic" peace literature. But there is a fair amount of evidence that hybrid or transitional regimes are more war prone. Hence interpolation of causal effects may not work in this case. Also, extrapolation to higher quality democracy is probably going to be more valid than interpolation to gray zone democracies.

For (4) the key issue hidden by the King and Zeng article is that one wants to make contrasts with cases that are nearby. Data in the convex hull may be more distant than data outside it. The Possibility Principle says that increasing the convex hull may be counter-productive if the contrasting cases are too far away.

55. Large-N qualitative testing works when there are few cases in the (1,1) cell. One generic situation where this can occur (illustrated by the Mansfield and Synder example in Goertz 2017) is an interaction term where already one or both of the terms is relatively rare. For international conflict scholars this can occur when joint democracy is interacted with virtually any other variable. Discuss this, for example, in Ghatak et al. 2017 who interact joint democracy with territorial conflict.

56. Fearon and Laitin in their oft-cited work (2003) on civil war (in)famously include a variable on amount of mountainous terrain (see Collier and Hoeffler 2004). This is controversial in terms of causal inference because it essentially does not vary within countries. Influential causal inference scholars deny causal status to variables which cannot be manipulated. Discuss how one use large-N qualitative case studies to explore the causal mechanism linking mountainous terrain to civil war. Can this illustrate how a constant can be part of a causal mechanism?

57. The medium-N paradigm means in this context exploring cases in border areas of the (1,1) zone as secondary case studies. For the democratic peace this would be borderline democracies or borderline peace (in the joint democracy is sufficient for peace version of the democratic peace). An interesting exercise to see what the cases scholars have used when critiquing of the democratic peace, e.g., Elman (1997) or Rosato (2003).

58. Large-N qualitative testing. Analyze this example of large-N qualitative testing.

> In the replicated version of Schultz's data set, there are 147 democratic victories – that is, unreciprocated MIDs initiated by democracies. Of these episodes, Table 2 lists the twenty-five most influential cases as measured by the dfbeta statistic, which reports the change in the coefficient of interest when an individual case is excluded from the regression. These twenty-five cases – roughly 2 percent of the overall data set – exert the greatest downward impact on the democratic initiator variable, in effect doing the most work to confirm the predictions of the democratic credibility hypothesis. The empirical analysis in *Democracy and Coercive Diplomacy* relied on the MID data set because, according to the study, the data set is composed of "cases in which states used threats of force, regardless of how prominent or how severe the ensuing crisis eventually became." If true, the data set would indeed be useful for testing the democratic credibility

> hypothesis, because the theory aims to explain the success and failure of militarized threats. However, this characterization of the MID data set is largely incorrect. Our research indicates that none of the twenty-five most influential democratic victories in our replication of Schultz's analysis represents an actual threat made by a democracy. In fact, as Table 2 reports, eight of the twenty-five cases appear to involve no militarized dispute at all. The remaining cases entail unilateral raids, skirmishes, or border violations (twelve cases!, troop movements or exercises without a coercive demand (two cases), and encounters with fishing boats or other civilian vessels (three cases!. Because these cases do not involve coercive threats, they do not belong in an empirical test of democratic credibility theory. (Downes and Sechser 2012, 468)

59. Large-N qualitative testing and multimethod research more generally involves looking for evidence of the causal mechanism in cases. Discuss Simmons and Danner (2011, 235) where the number of (1,1) cases is manageable, but they cannot find much qualitative evidence. Does that invalidate the statistical analysis?

60. Copeland in his analysis of causal importance of economic interdependence and future expectations looks all cases of major power crisis and war: "In this book, I look at the onset of essentially all the significant great power crises and wars from 1790 to 1991" (2015, 76). In the context of large-N qualitative testing, discuss the advantages and disadvantages of this Y-centric approach.

Within case causal inference, counterfactuals, process tracing,

61. For Waldner (2026) a causal figure consists of variables and is intended to apply to multiple cases. To this correspondence what he calls an "event history map" which is how these variables play out in a given case, in other words the values of the variables for that particular case. This is very similar to the LNQA approach to within-case causal inference. A major difference however is that in LNQA the arrows in the event history map are causal claims about what happened in an individual case. In contrast, for Waldner *these arrows are not causal* (see the next exercise below for more on this). This constitutes a major difference between the Waldner approach and LNQA: in LNQA the causal model is applied directly to the case and the arrows are understood to mean causation in the individual case. Contrast Waldner's position with LNQA on this core topic of inference. For example, in some cases there might be a backdoor path and in other cases there is no evidence for a backdoor path (the same could be true for side door paths).

> The specific meaning of the nodes in an event-history map implies a specific meaning of the arrows as well. Because these nodes represent realized events rather than random variables, the map itself does not represent a joint probability distribution. The arrows connecting the nodes, therefore, do not represent relationships of probabilistic dependence or causal relationships. Event-history maps constitute the methodologically

> constrained narrative of the case study, constrained by the conjectured data-generating process represented by the causal graph. It is thus critical to avoid the temptation to interpret an event-history map as representing causal relations. . . . Once this mapping is complete, the causal graph that satisfies our criteria allows us to give a causal interpretation of the event-history map, supporting the claim that each set of events comprising a node in the event-history map generates the set of events comprising the next node via the intentional action comprising the invariant causal mechanism. (Waldner 2026, 132)

62. How to think about and how to diagram evidence in causal model figures is still very much a topic of discussion among process tracing scholars. Waldner (2026) gives an interesting example in the figure below. The lowercase letters indicate some hoop tests that give evidence about whether X is a cause of Y. Notice that he is concerned about the reader being confused between the evidence which are individual events about the case distinguished from the causal factors indicated by the capital letters (see the exercise above on this point). So the question then becomes how is one to interpret the (causal) arrows going to and from these events. This can be contrasted with the smoking gun figure discussed in the exercise below. Describe how LNQA could diagram these three hoop tests as evidence about a causal relationship between X and Y.

> Consider a hypothesis that X is connected to Y by the intervening events X → a → b → c → Y, where X and Y are the variables and a, b, and c are intervening events rather than variables. If careful research does not uncover evidence that a, b, and c were present, the hypothesis has failed three hoop tests and can be rejected. But if each hypothesis passes its hoop test, the intervening events a, b, and c were present. This generates a sequence of events, but as we have repeatedly affirmed, concatenation is not causation. . . . Figure 7.5 is not a causal graph because it mixes variables and events. (Waldner 2026, 262)

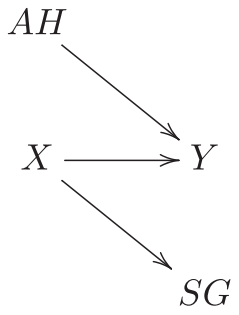
Figure 15: Hoop tests à la Waldner

$$X \longrightarrow a \longrightarrow b \longrightarrow c \longrightarrow Y$$

63. The interpretation of a smoking gun test varies quite significantly among authors. Here I give Waldner's interpretation of it in his figure 16 below. Noticed that in contrast to the LNQA interpretation of evidence, here the evidence consists of the causal relationship between X and SG: this causal relationship is evidence that X was the cause of Y. How would one graph this from the LNQA perspective? Also discuss the importance of *unique* evidence that we have highlighted in the Waldner quote. Why is that critical for this being smoking gun evidence? Would one draw the same graph for a hoop test? Contrast this with how Blair et al. do smoking gun, discussed in chapter 5, appendix B of the LNQA book.

> Figure 1.3 also contains a node denoted SG. This node is not on the front-door path between X and Y; it is what James Mahoney calls an "auxiliary outcome" or "particular occurrences that should occur alongside (or perhaps as a result of) the main outcome of interest if in fact the outcome was caused in the way stipulated by the theory under investigation." This node represents what process-tracing methods call "smoking-gun" evidence. Such evidence conforms to a unique prediction made by only one hypothesis; hence, the graph contains an arrow from X to SG but not from AH [auxiliary hypothesis, e.g., alternative hypothesis] to SG. The hypothesis denoted X thus passes a smoking-gun test and is strongly corroborated because it *uniquely* predicts this evidence. (Waldner 2026, 38, emphasis is mine, see also his discussion around Figure 7.6).

Figure 16: Smoking gun tests à la Waldner



64. Central to Waldner's account of what he calls "qualitative cause inference" is the notion of a "hypothetical intervention." This is an intervention on the key treatment variable in the model or theory. So for example, a hypothetical intervention could set X to a value different from that observed in individual case i. In contrast, he has nothing to say about counterfactuals in the book. Discuss the similarities and differences between his notion of hypothetical intervention and the methodology of counterfactual analysis. Are the criteria – for example, the three described below – for good hypothetical interventions the same or consistent with the criteria for good counterfactual analysis?

> For any hypothetical intervention, we should be more willing to accept the claim if the reasons adduced on its behalf exhibit theoretical plausibility, mechanistic credibility, and the associated descriptive inferences are each deemed valid. If these elements are missing or only weakly established, we should be correspondingly less willing to accept the claim about the results of the hypothetical intervention. In the ideal scenario, a claim about the consequences of a hypothetical intervention has all three qualities: theoretical plausibility, mechanistic credibility, and descriptively valid inferences. Conversely, we should have serious doubts about such interventions when they lack grounding in general theories; when they lack fine-grained mechanistic accounts of the relevant actors, their intentions, their actions, and the consequences of their actions; and when the best available evidence refutes or only weakly supports the derived hypotheses. (Waldner 2026, 216–17)

65. Counterfactuals are core to causal inference in both qualitative and quantitative methodologies. Within Boolean or mathematical logic perspectives there is a difference between a counterfactual about a necessary condition versus for one for a sufficient condition. For example,

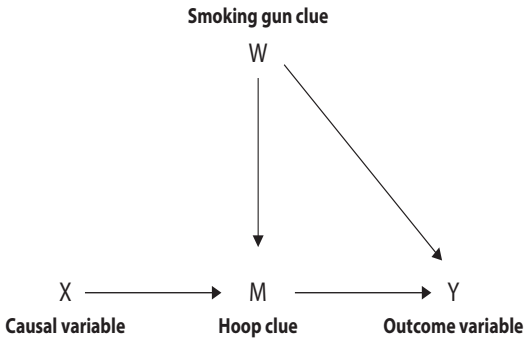
> A causal hypothesis in political science often takes the following simple form: "x was a cause of y," or equivalently, "x caused y." This type of question typically arises when a researcher seeks to explain the reason a past event happened. . . . From the counterfactual viewpoint, x is said to have caused y when y would not have happened had x been absent. (Yamamoto 2012, 239)

Discuss the differences if X is necessary condition for Y versus a sufficient condition for Y. Does the absence of X which is a sufficient condition for Y mean that Y does not occur?

Answer: Because of equifinality the absence of sufficient condition X does not necessarily mean that Y would not have occurred.

66. Discuss the Blair et al. interpretation of smoking gun and hoop tests for individual cases within a DAG process tracing model. In particular, which of the factors in the figure are causal and which are evidence about causes? This is also about interpreting what $\rightarrow$ means in causal model figures.

Figure 17: Process tracing and within-case inference



Source: Based on Blair et al. 2023.

Answer: Some arrows are about *evidence* for a causal relationship and other arrows are about the *causal* relationship itself. Probably one needs different kinds of arrows since these are two quite different aspects of causal inference.

67. Doing within-case causal inference in LNQA is closely connected to doing causal mechanism figures. Do a figure for each phase of Widner's case study regarding the US response to EBOLA in West Africa. See also Widner's book on case study methodology (Widner et al. 2022).

> In 2014, an unprecedented outbreak of Ebola virus disease in Liberia, Sierra Leone, and Guinea shined a harsh spotlight on global capacity to

deal effectively with a fast-moving epidemic that crossed international borders. By the end of July, the outbreak had started to overwhelm health care systems in all three affected countries. In Liberia, health centers began to close, and President Ellen Sirleaf appealed for help from the United States. President Barack Obama tasked USAID's Office of US Foreign Disaster Assistance (OFDA) to lead an interagency response. From early August 2014 to January 2016, an OFDA Disaster Assistance Response Team, or DART, deployed to Liberia to help coordinate efforts to stop the spread of infection. The DART was the first to involve a large-scale partnership with the US Centers for Disease Control and Prevention (CDC) to combat an infectious disease outbreak. Although the deployment, which scaled up earlier assistance, took place five months after the first reported cases and required extensive adaptation of standard practices, it succeeded in helping bring the epidemic under control: the total number of people infected – 28,616 – was well below the potential levels predicted by the CDC's models. This US–focused case study highlights the challenges of making an interagency process work in the context of an infectious disease outbreak in areas where health systems are weak.

68. Within-case causal inference is core to LNQA. The process tracing literature loves the detective metaphor. Discuss Cartwright and Hardie's figures for analyzing a detective story, see the figures below. In particular, focus on issues of aggregation across a level, e.g., motive, opportunity, motive.

> Mystery stories can provide good examples of the importance of each premise to the stability of the overall structure. Consider Freeman Wills Croft's Inspector French, who is always at pains to lay out clearly the arguments that support his conclusions. In *Death On the Way*, Inspector French has become convinced, on the basis of a body of seemingly good evidence, that Carey murdered Ackerly. We reconstruct his warrant for this in the form we advocate – that of an argument with explicit premises and explicit subpremises. Premise 1 in French's argument concerns motive. Ackerly, French argued, had cottoned on to a major fraud that Carey was involved in. This was backed by compelling evidence for three claims: fraud had been committed; Ackerly had raised suspicions about it; and Carey was responsible for the fraud. This last was supported by evidence that Carey was in a position to perpetrate the fraud, that he had income otherwise unaccounted for, and, very importantly, that he had committed suicide when it looked as if the fraud might be revealed despite Ackerly's death. (Cartwright and Hardie 2012, 19)

Figure 18: Within-case causal inference: The argument pyramid supporting French's accusation of Carey

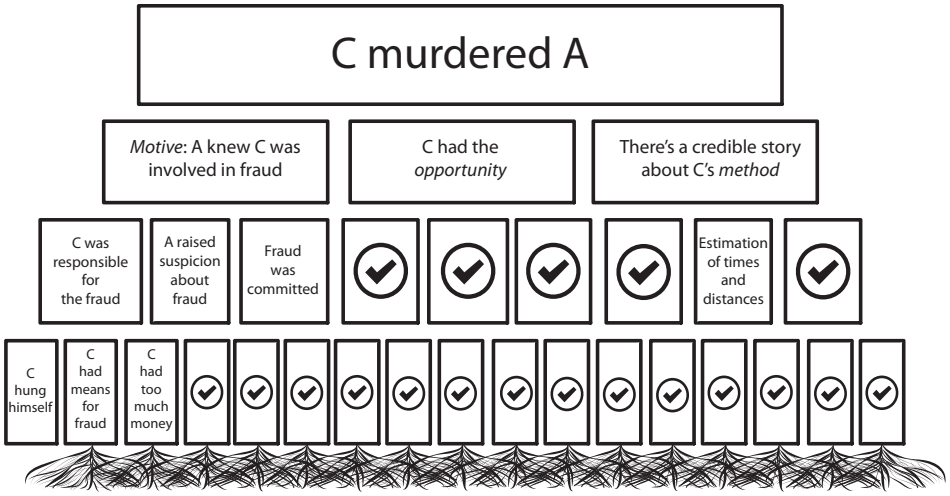


Figure I.3: The argument pyramid supporting French's accusation of Carey

Source: Cartwright and Hardie 2012.

Figure 19: The collapse of a within-case causal inference: the collapse of an accusation again Carey

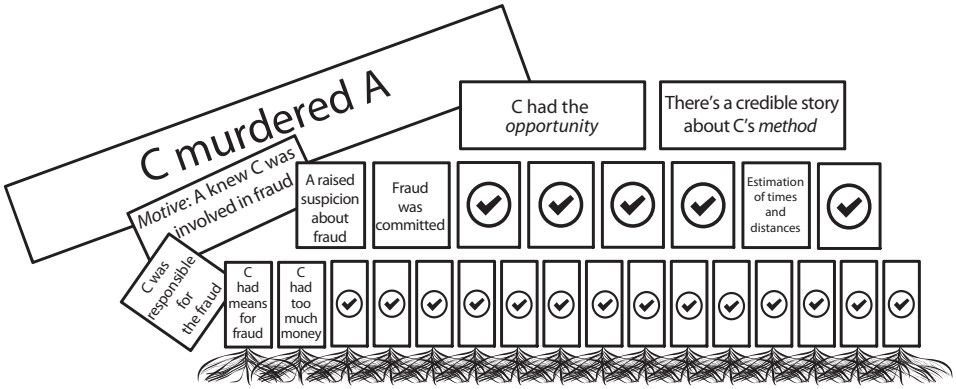


Figure I.4: The collapse of an accusation

Source: Cartwright and Hardie 2012.

Answer: It appears that they use necessary condition aggregation across levels.

69. Within-case causal inference needs to distinguish between "outcomes" and "facts" which support causal claims. In the Holmes story *Silver Blaze* a critical

part of story deals with tendons of sheep which have been cut. This tendon-nicking could be described as a "fact" that is used to support a causal inference, not an "outcome." It is not an outcome to be explained but a fact used to support a central causal claim of the story (Goertz 2019). Others call it an outcome. For example, Fairfield and Charman describing this: "Mahoney (2010, 125–31) and Collier (2011, 828) speak to this point in their discussions of 'auxiliary outcomes' that are not part of the primary causal sequence. Collier (2011) uses the example of Straker's tendon-nicking experiments on sheep in *The Adventure of Silver Blaze*. This clue strengthens Sherlock Holmes' inference that the horse was responsible for the death of Straker, who was trying to throw a race – the horse kicked Straker in the head when it sensed something was wrong." (Fairfield and Charman 2022, 101). How do "facts" aka "clues"play a role in within-case causal inference in the *Silver Blaze* story.

70. Counterfactuals are a core methodology for within-case causal inference. They arise naturally for necessary conditions because the absence of the necessary condition leads to the absence of the outcome. In contrast they are problematic for sufficient conditions. The literature on counterfactuals for sufficient conditions is quite small and many do not notice that there is any difference. The main issue is that the absence of X when X is sufficient does not lead to clear conclusions about the presence or absence of Y. This means that the counterfactual analysis of sufficient conditions is all about equifinality. If a given X is absent is there another X present that can substitute for it? For contrasting view of sufficient condition counterfactuals see Mahoney (2020; Mahoney and Barrenechea 2019). Look for examples of sufficient condition hypotheses in LNQA and other literature and explore the relevant counterfactuals for within-case analysis.

71. Interviews and focus groups often generate "causal claims," e.g., something happened *because* of X. These causal claims are often not considered distinct from "descriptive claims" about facts. Go through Widner's chapter "Descriptive accuracy in interview-based case studies" and locate the causal claims in her discussion. Is there a pattern to where they seem to arise in her discussion? Here are a couple of examples of causal questions to get going with:

> How would you say that joining the deliberation/process/negotiation late colored your view of the issues and shaped your actions, or did it not make much difference?
>
> The military built emergency treatment centers that were never used because the epidemic ended by the time the centers were ready. The US military action was irrelevant. . . . Later the epidemic declined and the ETUs [emergency treatment units] weren't used, but in the end what seemed to matter to the public was the visible sign that a big power cared, which generated a psychological boost. (Widner 2022, 125)

72. Counterfactuals often appear in major historical controversies. One of the biggest ones regarding World War II was what would have happened had the United States not dropped the two atomic bombs on Japan. A major report

after the war endorsed a controversial counterfactual about the dropping of the atomic bomb. Newman (1995) notes:

> The most dogmatic statement of this counterfactual proposition appears in both the *Summary Report* and *Japan's Struggle*: "Based on a detailed investigation of all the facts, and supported by the testimony of the surviving Japanese leaders involved, it is the Survey's opinion that certainly prior to December 31, 1945, and in all probability prior to November 1, 1945, Japan would have surrendered even if the atomic bombs had not been dropped, even if Russia had not entered the war, and even if no invasion had been planned or contemplated." United States Strategic Bombing Survey (USSBS), Summary Report (Pacific War).

Newman vigorously argues against this counterfactual. Explore the various ways one could frame the counterfactual. Are there other interesting counterfactuals he does not contemplate? What kind of evidence does he drawn on to conclude that the atomic bomb was in fact a major cause of the Japanese surrender and that counterfactually they would not have surrendered for a considerable amount of time had the bomb not been dropped.

Answer: One interesting counterfactual he does not completely explore is counterfactually what would it have had taken the Japanese Emperor to sue for peace in the absence of the atomic bomb. As a matter of historical fact he was already trying to talk to the Americans via the Russians about ending the war. Another one, for example, would a land invasion have been necessary to end the war in the absence of the atomic bomb?

73. Counterfactuals often appear in major historical controversies. One of the biggest ones regarding World War II was what would have happened had the United States not dropped the two atomic bombs on Japan. A major report after the war endorsed a controversial counterfactual about the dropping of the atomic bomb. Newman (1995) notes:

    > The most dogmatic statement of this counterfactual proposition appears in both the *Summary Report* and *Japan's Struggle*: "Based on a detailed investigation of all the facts, and supported by the testimony of the surviving Japanese leaders involved, it is the Survey's opinion that certainly prior to December 31, 1945, and in all probability prior to November 1, 1945, Japan would have surrendered even if the atomic bombs had not been dropped, even if Russia had not entered the war, and even if no invasion had been planned or contemplated." United States Strategic Bombing Survey (USSBS), Summary Report (Pacific War).

    Newman vigorously argues against this counterfactual. Explore the various ways one could frame the counterfactual. Are there other interesting counterfactuals he does not contemplate? What kind of evidence does he drawn on to conclude that the atomic bomb was in fact a major cause of the Japanese surrender and that counterfactually they would not have surrendered for a considerable amount of time had the bomb not been dropped.

    Answer: One interesting counterfactual he does not completely explore is counterfactually what would it have had taken the Japanese Emperor to sue for peace in the absence of the atomic bomb. As a matter of historical fact he

was already trying to talk to the Americans via the Russians about ending the war. Another one, for example, would a land invasion have been necessary to end the war in the absence of the atomic bomb?

74. Doing within-case causal inference in LNQA is closely connected to doing causal mechanism figures. Do a figure for each phase of Widner's case study regarding the US response to EBOLA in West Africa. See also Widner's book on case study methodology (Widner et al. 2022).

> In 2014, an unprecedented outbreak of Ebola virus disease in Liberia, Sierra Leone, and Guinea shined a harsh spotlight on global capacity to deal effectively with a fast-moving epidemic that crossed international borders. By the end of July, the outbreak had started to overwhelm health care systems in all three affected countries. In Liberia, health centers began to close, and President Ellen Sirleaf appealed for help from the United States. President Barack Obama tasked USAID's Office of US Foreign Disaster Assistance (OFDA) to lead an interagency response. From early August 2014 to January 2016, an OFDA Disaster Assistance Response Team, or DART, deployed to Liberia to help coordinate efforts to stop the spread of infection. The DART was the first to involve a large-scale partnership with the US Centers for Disease Control and Prevention (CDC) to combat an infectious disease outbreak. Although the deployment, which scaled up earlier assistance, took place five months after the first reported cases and required extensive adaptation of standard practices, it succeeded in helping bring the epidemic under control: the total number of people infected – 28,616 – was well below the potential levels predicted by the CDC's models. This US–focused case study highlights the challenges of making an interagency process work in the context of an infectious disease outbreak in areas where health systems are weak.

75. Within-case causal inference is core to LNQA. The process tracing literature loves the detective metaphor. Discuss Cartwright and Hardie's figures for analyzing a detective story, see the figures below. In particular, focus on issues of aggregation across a level, e.g., motive, opportunity, motive.

> Mystery stories can provide good examples of the importance of each premise to the stability of the overall structure. Consider Freeman Wills Croft's Inspector French, who is always at pains to lay out clearly the arguments that support his conclusions. In *Death On the Way*, Inspector French has become convinced, on the basis of a body of seemingly good evidence, that Carey murdered Ackerly. We reconstruct his warrant for this in the form we advocate – that of an argument with explicit premises and explicit subpremises. Premise 1 in French's argument concerns motive. Ackerly, French argued, had cottoned on to a major fraud that Carey was involved in. This was backed by compelling evidence for three claims: fraud had been committed; Ackerly had raised suspicions about it; and Carey was responsible for the fraud. This last was supported by evidence that Carey was in a position to perpetrate the fraud, that he had income otherwise unaccounted for, and, very importantly, that he had committed suicide when it looked as if the fraud might be revealed despite Ackerly's death. (Cartwright and Hardie 2012, 19; alternative figures)

Figure 20: Within-case causal inference: The argument pyramid supporting French's accusation of Carey

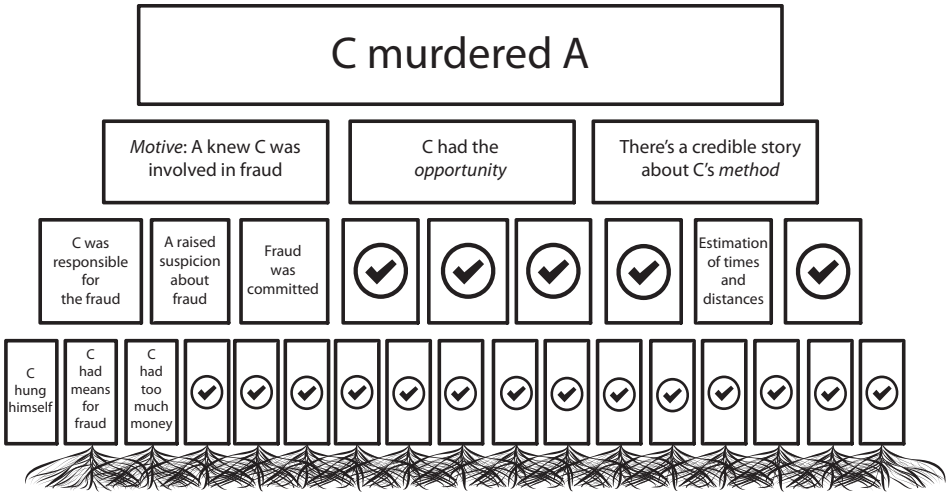


Figure I.3: The argument pyramid supporting French's accusation of Carey

Source: Cartwright and Hardie 2012.

Figure 21: The collapse of a within-case causal inference: the collapse of an accusation again Carey

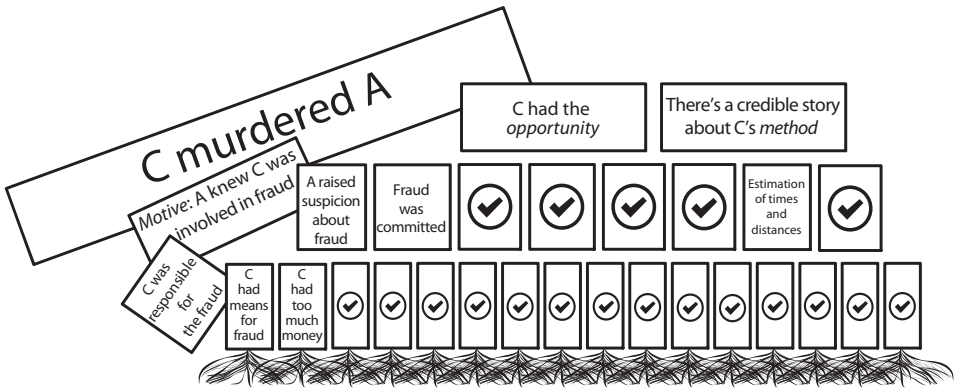


Figure I.4: The collapse of an accusation

Source: Cartwright and Hardie 2012.

Answer: It appears that they use necessary condition aggregation across levels.

76. Within-case causal inference needs to distinguish between "outcomes" and "facts" which support causal claims. In the Holmes story *Silver Blaze* a critical

part of story deals with tendons of sheep which have been cut. This tendon-nicking could be described as a "fact" that is used to support a causal inference, not an "outcome." It is not an outcome to be explained but a fact used to support a central causal claim of the story (Goertz 2019). Others call it an outcome. For example, Fairfield and Charman describing this: "Mahoney (2010, 125–31) and Collier (2011, 828) speak to this point in their discussions of 'auxiliary outcomes' that are not part of the primary causal sequence. Collier (2011) uses the example of Straker's tendon-nicking experiments on sheep in *The Adventure of Silver Blaze*. This clue strengthens Sherlock Holmes' inference that the horse was responsible for the death of Straker, who was trying to throw a race – the horse kicked Straker in the head when it sensed something was wrong." (Fairfield and Charman 2022, 101). How do "facts" aka "clues"play a role in within-case causal inference in the *Silver Blaze* story.

77. Counterfactuals are a core methodology for within-case causal inference. They arise naturally for necessary conditions because the absence of the necessary condition leads to the absence of the outcome. In contrast they are problematic for sufficient conditions. The literature on counterfactuals for sufficient conditions is quite small and many do not notice that there is any difference. The main issue is that the absence of X when X is sufficient does not lead to clear conclusions about the presence or absence of Y. This means that the counterfactual analysis of sufficient conditions is all about equifinality. If a given X is absent is there another X present that can substitute for it? For contrasting view of sufficient condition counterfactuals see Mahoney (2020; Mahoney and Barrenechea 2019). Look for examples of sufficient condition hypotheses in LNQA and other literature and explore the relevant counterfactuals for within-case analysis.

78. Interviews and focus groups often generate "causal claims," e.g., something happened *because* of X. These causal claims are often not considered distinct from "descriptive claims" about facts. Go through Widner's chapter "Descriptive accuracy in interview-based case studies" and locate the causal claims in her discussion. Is there a pattern to where they seem to arise in her discussion? Here are a couple of examples of causal questions to get going with:

> How would you say that joining the deliberation/process/negotiation late colored your view of the issues and shaped your actions, or did it not make much difference?
>
> The military built emergency treatment centers that were never used because the epidemic ended by the time the centers were ready. The US military action was irrelevant. . . . Later the epidemic declined and the ETUs [emergency treatment units] weren't used, but in the end what seemed to matter to the public was the visible sign that a big power cared, which generated a psychological boost. (Widner 2022, 125)

79. Bove and Nisticó (2014)use a relatively new methodology called the synthetic control methodology which has generated much interest. The title of their article is "Coups d'état and defense spending: a counterfactual analysis." Are they actually doing a counterfactual in the traditional sense?

Answer: As the name indicates the methodology generates a *control* comparison observation. One constructs a control group to compare with their military coup cases. A comparison of treatment and control is not obviously a counterfactual in the classic sense of the term.

80. It is not uncommon to confuse necessary condition counterfactuals with sufficient condition ones. See Fenoaltea (1973) for a beautiful discussion of this in the context of economic history. Fischer's classic book (1970) also deals with this. Discuss Fisher's examples.

> One common form of the reductive fallacy is the confusion of necessary with sufficient cause – the confusion of a causal component without which an effect will not occur, with all the other causal components which are required to make it occur. This sort of error appears in causal explanations which are constructed like a single chain and stretched taut across a vast chasm of complexity. The classic example is the legendary battle that was lost for the want of a horseshoe nail; for the the want of a nail the shoe was lost, for the want of a shoe the horse was lost, for the want of a horse the rider was lost, for the want of a rider the message was lost, for the want of the message the regiment was lost, and for the want of a regiment the battle was lost. (Fischer 1970, 172; he then goes on to use the classic example of the battle of Antietam for which the accident loss of confederate orders to the North was a key factor in the North's victory)

81. A key question in the causal analysis of individual cases is the relationship between temporal distance and causal importance. Should events or factors closer to the event to be explained receive greater or less causal importance. Analyze the quote below from Gerring (2005) who argues for increasing impact as temporal distance increases in the causal chain. Does it matter whether the cause is a necessary or sufficient condition?

> Consider the following path diagram. $X_1 \Rightarrow X_2 \Rightarrow X_3 \Rightarrow X_4 \Rightarrow Y$ We are apt to consider X_1 to be the cause and causal factors $X_2 - X_4$ intermediate (and less important) causes, all other things being equal. Of course, all other things are rarely equal. We are likely to lose causal power (accuracy and completeness) as we move further away from the outcome. Yet, if we did not – e.g. if the correlations in this imaginary path diagram were perfect – we would rightly grant priority to X_1. Causes lying close to an effect are not satisfying as causes, precisely because of their proximity. Rather, we search for causes that are 'ultimate' or 'fundamental'.
>
> Consider a quotidian example. To say that an accident was caused because A ran into B is not to say much that is useful about this event. Indeed, this sort of statement is probably better classified as descriptive, rather than explanatory. An X gains causal status as it moves back further in time from the event in question. If, to continue with this story, I claim that the accident was caused by the case of beer consumed by A earlier that evening, I have offered a cause that has greater priority and is, on this account at least, a better explanation. If I can show that the accident in question was actually a re-enactment of a childhood accident that A experienced 20 years ago, then I have offered an even more interesting explanation. Similarly, to say that the Civil War was caused by the attack on Fort Sumter, or that the First World War was caused by the assassination of the Archduke Francis Ferdinand at Sarajevo, is to make a causal

> argument that is almost trivial by virtue of its lack of priority. It does not illumine very much, except perhaps the mechanism that might be at work vis-a-vis some prior cause.
>
> The further away we can get from the outcome in question, the more satisfying (ceteris paribus) our explanation will be. This explains much of the excitement when social scientists find 'structural' variables that seem to impact public policy or political behavior. It is not that they offer more complete or more accurate explanations; indeed, the correlations between X and Y are apt to be much weaker. It is not that they are more relevant; indeed, they are less relevant for most policy purposes, since they are apt to be least amenable to change. Priority often imposes costs on other criterial dimensions. Yet, such explanations will be better insofar as they offer us more power, more leverage on the topic. They are non-obvious. (Gerring 2005: 174–5)

82. There is significant debate about path dependency, historical contingency and the like. Explain why the "for example counterfactual" is different than the historical contingency definition preceding it.

> By historical contingency we mean chance-influenced events with substantial long-term effects, that is, events which clearly take history down a different path than it otherwise would have followed. . . . For example, had Empress Elizabeth died a month later, Prussia might have lost the Seven Years War. The chance event caused a different path to be taken. (Bendor and Shapiro 2019, 131)

83. Discuss the validity and methodology of a core counterfactual from Simmons's very influential book on human rights:

> No available secondary source account argues that Japan would have passed legislation to address some of the most egregious aspects of discrimination against women in employment had not the question of Japan's position on the CEDAW come up. Clearly, this was a case in which an externally negotiated agreement changed the country's legislative agenda, placing issues of women's equality much higher on the list of legislative priorities than would have existed had Japan not been faced with the issue of what to do about the CEDAW. To be sure, women's equality had its domestic supporters, but they were clearly in a minority and somewhat isolated politically and bureaucratically. They were hardly a match for Japan's powerful business interests that wanted to maintain the status quo. Second, the existence of the CEDAW increased the size of the coalition that was to support the EEOL. The Ministry of Foreign Affairs, for example, would not normally weigh in on domestic legislation on equality in women's employment, but they did in this case because they wanted Japan on board with this major UN initiative. (Simmons 2009, 240)

84. Discuss the "light-switch" analogy as a way to think about process tracing and equifinality. The basic analogy is that the light-switch is on and the light is on (essentially the (1,1) cases). Discuss process tracing as following electricity through various junction boxes to the light bulb. Equifinality lies in the fact that there might be multiple switches that can possibly control the same light. Discuss what would be the analogy in case study work of "turning on or off" the switch (within cases or between cases). What about counterfactuals? (Thanks to Mike Desch for suggesting this analogy)

85. In the literature on within-case causal inference researchers often talk about hoop tests and smoking gun tests in terms of "certainty" and "uniqueness": "The first dimension is called certitude or certainty and captures how likely it is to confirm a specific observable implication in process tracing. The second dimension is called uniqueness and asks whether an observable implication can be derived from a single or multiple hypotheses" (Rohlfing 2014 table 1). Rohlfing goes on to claim: "A hoop test is characterized by high certainty and no uniqueness." (Rohlfing 2014, 612) Describe a situation where a hoop test has a high degree of uniqueness.

    Answer:

    As the hoop test gets closer to being a sufficient condition, its uniqueness increases, i.e., when the hoop is very small.

86. Use the democratic peace to propose a specific counterfactual so that World War I does not occur.

87. Discuss the following claim by one of the most prominent historians today. Does it invoke a necessary or sufficient condition claim? Do a counterfactual analysis of it.

    > The argument that British intervention in the war was made inevitable by the violation of Belgian neutrality has been repeated by historians ever since [Lloyd George]. (Ferguson 1999a: 231)

88. The confusion of necessary with sufficient conditions also arises when defining what "cause" means. Explain how this works in Wendt's analysis of cause:

    > In saying that 'X causes Y' we assume that: (1) X and Y exist independent of each other, (2) X precedes Y temporally, and (3) but for X, Y would not have occurred. . . . The logical empiricist model of causal explanation, usually called the deductive-nomological model or D-N model, is rooted in David Hume's seminal discussion of causality. Hume argued that when we see putative causes followed by effects, i.e., when we have met conditions (1) and (2), all we can be certain about is that they stand in relations of constant conjunction. The actual mechanism by which X causes Y is not observable (and thus uncertain), and appeal to it is therefore epistemically illegitimate. Even if there is necessity in nature, we cannot know it. How then to satisfy the third, counterfactual condition for causality, which implies necessity? (Wendt 1999: 79)

89. Quite a few different terms are often used to distinguish between different kinds of causes, e.g., "contributory cause," "remote cause," "intervening/mediating cause." Use Russett's very interesting analysis of World War I (1962) – in particular Fay's account – to discuss what these kinds of cause mean. See also Fischer's classic book (1970) which talks about eight kinds of cause: (1) all antecedents, (2) regularistic antecedents, (3) controllable antecedents, (4) rational and/or motivational antecedents, (5) abnormal antecedents, (6) structural antecedents, (7) contingent-series antecedents, (8) precipitant antecedents.

90. Brooks and Wohlforth argue for a *probabilistic* interpretation of their causal claims in their response to English. For the philosophically inclined, there is a major debate in philosophy about the "sense" that probably statements have for individual cases. Many argue that probability statements only make sense in over repeated events. This is called the "frequentist" position, e.g., "Our probability theory has nothing to do with questions such as: 'is there a probability of Germany being at some time in the future involved in a war with Liberia.' ... probability when it refers to a single event has no meaning at all for us" (von Mises 1957, 9, 11).

91. Van Evera's well-known article (1984) on World War I argues that the "cult of the offensive" was an important cause of the war. Find three key counterfactuals in that article that imply necessary condition hypotheses. One of the three should involve the "window of opportunity" idea.

92. What is the logical form of the hypothesis that U.S. military pressure would put such a strain on the Soviet Union "that the Soviet leadership would have little choice but to make substantial concessions on arms control" (Gaddis 1989, 13)?

93. Many historians feel that events like World War I were "over-determined." Explain why this would generally lead them to discount necessary condition counterfactuals, e.g., "The fact that so many plausible explanations for the outbreak of the war [World War I] have been advanced over the years indicates on the one hand that it was massively overdetermined, and on the other that no effort to analyze the causal factors involved can ever fully succeed" (Schroeder 1972, 320).

94. Define the contrapositive: (1) starting with necessary conditions, (2) starting with sufficient conditions, (3) with multiple necessary conditions, (4) with multiple sufficient conditions.

95. The Schroeder chapter in the Goertz and Levy anthology is a brilliant example of how one can use the contrapositive to make a very provocative claim. (1) Explain what he did. (2) Do the same thing with a causal claim about the end of the Cold War.

96. The following quote summarizes the argument of Sagan (1986) on the causes of World War I:

> Sir Edward Grey's failure to present a clear and credible threat of British intervention early in the July crisis and the specific preemptive aspects of Germany's offensive war plans caused by the slow Russian mobilization and the Liege bottleneck are linked together as an immediate cause of the First World War. (Sagan 1986: 168; this is key thesis of whole article).

In the conclusions Sagan proposes the following counterfactual which is potentially *logically* inconsistent with the basic theory given above. What is that potential logical inconsistency? Can you give an interpretation which makes everything consistent?

> *If* Grey had given a clear warning earlier, *if* the Czar had further delayed Russian mobilization against Austria and then Germany, and *if* the German offensive war plans had not been able to depend upon a preemptive *coup de main* against Liege and the decisive battle in France before Russian mobilization was completed in the East, *then* it is possible, just possible, that Bethmann-Hollweg would have had the time and the courage necessary to apply sufficient pressure on Vienna to accept the "Halt in Belgrade." And if this had occurred, 1914 might today appear as only another one of a series of Balkan crises that almost led to a world war. (Sagan 1986, 169).

97. A key question in the causal analysis of individual cases is the relationship between temporal distance and causal importance. Should events or factors closer to the event to be explained receive greater or less causal importance. Analyze the quote below from Gerring (2005) who argues for increasing impact as temporal distance increases in the causal chain. Does it matter whether the cause is a necessary or sufficient condition?

> Consider the following path diagram. $X_1 \Rightarrow X_2 \Rightarrow X_3 \Rightarrow X_4 \Rightarrow Y$ We are apt to consider X_1 to be the cause and causal factors $X_2 - X_4$ intermediate (and less important) causes, all other things being equal. Of course, all other things are rarely equal. We are likely to lose causal power (accuracy and completeness) as we move further away from the outcome. Yet, if we did not – e.g. if the correlations in this imaginary path diagram were perfect – we would rightly grant priority to X_1. Causes lying close to an effect are not satisfying as causes, precisely because of their proximity. Rather, we search for causes that are 'ultimate' or 'fundamental'.
>
> Consider a quotidian example. To say that an accident was caused because A ran into B is not to say much that is useful about this event. Indeed, this sort of statement is probably better classified as descriptive, rather than explanatory. An X gains causal status as it moves back further in time from the event in question. If, to continue with this story, I claim that the accident was caused by the case of beer consumed by A earlier that evening, I have offered a cause that has greater priority and is, on this account at least, a better explanation. If I can show that the accident in question was actually a re-enactment of a childhood accident that A experienced 20 years ago, then I have offered an even more interesting explanation. Similarly, to say that the Civil War was caused by the attack on Fort Sumter, or that the First World War was caused by the assassination of the Archduke Francis Ferdinand at Sarajevo, is to make a causal argument that is almost trivial by virtue of its lack of priority. It does not illumine very much, except perhaps the mechanism that might be at work vis-a-vis some prior cause.
>
> The further away we can get from the outcome in question, the more satisfying (ceteris paribus) our explanation will be. This explains much of the excitement when social scientists find 'structural' variables that seem to impact public policy or political behavior. It is not that they offer more complete or more accurate explanations; indeed, the correlations between X and Y are apt to be much weaker. It is not that they are more relevant; indeed, they are less relevant for most policy purposes, since they are apt to be least amenable to change. Priority often imposes costs on other criterial dimensions. Yet, such explanations will be better insofar as they

> offer us more power, more leverage on the topic. They are non-obvious. (Gerring 2005: 174–5)

98. The following is a well-known nursery verse that describes King Richard III's fall from power. Analyze the temporal distance causal claims embedded in it.

> For want of a nail the shoe was lost,
> For want of a shoe the horse was lost,
> For want of a horse the rider was lost,
> For want of a rider the battle was lost,
> For want of a battle the kingdom was lost,
> And all for the want of a horseshoe nail.

Answer:

Fischer's classic book on historical explanation (1970) discusses this example. See Lowe (1980) for a more technical philosophical analysis that focuses on counterfactuals and causal explanations.

99. United States law gives particular importance to the person or action that was the last one which could have potentially avoided a bad event (e.g., car accident). What does this mean about the relative causal importance of recent or distant events.

100. How can you use the powder keg model to combine the arguments of English and Brooks/Wohlforth debate about the end of the Cold War?

Answer:

Basically the Soviet decline is the powder keg which needs the spark of new ideas and leadership. This idea is very common in the social constructivist literature, e.g., "fire [powder keg] metaphor serves for the types of interactions that we describe throughout this book. There must be an idea, advocates to spread the idea, and a public ready to receive it" (Keck and Sikkink 1998: 46).

101. A key issue in the concept of a turning point revolves around whether a turning point should be defined in terms of points in history where things changed or should also include "potential" turning points but where things did not change. For example, Herrmann and Lebow do not include potential turn points in their definition: "We define a turning point in terms of two properties. First, it must be a change of significant magnitude, not an incremental adjustment but a substantial departure from previous practice. Second, it must be a change that would be difficult to undo" (Herrmann and Lebow 2004). Discuss the following example from A.J.P. Taylor regarding a turning point in German history where Germany did not turn:

> 1848 was the decisive year of German, and so of European, history: it recapitulated Germany's past and anticipated Germany's future. Echoes of the Holy Roman Empire merged into a prelude of the Nazi 'New Order'; the doctrines of Rousseau and the doctrines of Marx, the shade of Luther and the shadow of Hitler, jostled each other in bewildering succession. Never has there been a revolution so inspired by a limitless faith in the power of ideas; never has a revolution so discredited the power of ideas

> in its result. The success of the revolution discredited conservative ideas; the failure of the revolution discredited liberal ideas. After it, nothing remained but the idea of Force, and this idea stood at the helm of Germany history from then on. For the first time since 1521, the German people stepped on to the centre of the German stage only to miss their cues once more. German history reached its turning-point and failed to turn. This was the fateful essence of 1848. (Taylor 1945: 68; for a different perspective see Blackbourn and Eley 1984)

102. Use the idea of a *window of opportunity* to define all potential turning points. What is the role of necessary versus sufficient conditions?

103. Many historians feel that the "Pleiku" analysis of World War I is basically correct (i.e., that there were many Balkan crises and that one of them would have sparked a world war). Some similar arguments have been made with regard to the impact of the Versaille Treaty on the occurrence of World War II, that is made World War II very likely. What would be one counterfactual argument that strongly contests such views? Are there parallels to the end of the Cold War debate?

    Answer:

    One counterfactual would stress the absolutely central role of Hitler. For example, Mueller has in several places used this idea. For example,

    > Thus it is simply not true that the seeds of another great war were planted at the peace conference of 1919. In order to bring about another major war it was necessary for Germany, first, to desire to expand into non-German areas, second, to be willing to risk and threaten military action in order to get these areas, and third, to be willing to pursue war when these desires were opposed by other major countries. It seems to me that none of these propositions – particularly the last two – were very popular in Hitler's Germany, that almost no one accepted all of them, and that only Hitler, it appears, combined a fanatical acceptance of them with a maniacally determined and effective capacity to carry them out. (Mueller 1991, 22, see also Mueller 1989; Byman and Pollack 2001, 118)

104. Can you make counterfactual arguments about the role of key individual decision-makers in the World War I case? End of Cold War?

    Answer:

    Obviously this is a question about which historians have debated for years.

    Byman and Pollack sees German and Austrian leadership in these terms, "Absent the Iron Chancellor, it is hard to imagine a defeated Austria aligning with Prussia after the humiliations of Sadowa and Königgrätz. Similarly, it is equally hard to imagine a leader other than Wilhelm II repeatedly antagonizing Britain for so little purpose" (2001: 134).

    Brown argues for the essential role of Gorbachev by arguing against the sufficiency of alternative explanations: "None of these interpretations [(1) learning, (2) power, and (3) domestic political pressure] on its own provides an adequate explanation of the changing Soviet political agenda in the second half of the 1980s, although the first and the third – or a combination of the two

– have often been seen as sufficient by their proponents (p. 13) . . . It is tempting to see Gorbachev as the handmaiden of history or the embodiment of social forces which, if Gorbachev rather than Dmitry Ustinov had died in December 1984, would have brought forth an alternative leader in the mid-1980s whose policies would have been broadly the same as Gorbachev's producing similar results. This, however, is a temptation that should be resisted, for it has got little but a 'retrospective determinism' to commend it" (Brown 1996, 317).

105. Here are some claims that constitute core ideas of prominent work (both by political scientists and historians) on World War I (I do not include those defended in various chapters of the Goertz and Levy anthology). In all cases these are central claims made by the author(s). The quotes below are obviously dramatic simplifications of complex arguments but they do capture the core of the scholar's argument and are useful for classroom discussion. Some are expressed in necessary condition counterfactual terms. Others are not; for these generate some necessary condition counterfactuals. More generally, analyze the causal explanations in them. These can be used as the basis for class discussion or additional reading.

    Goemens:
    I have to put the blame for four years of war and over ten million deaths squarely on the shoulders of Germany's regime. If Germany had not been a semirepressive, moderately exclusionary regime, would the war have ended sooner? Such counterfactual claims are notoriously hard to evaluate. . . . An argument can nevertheless be made that a nonrepressive, nonexclusionary Germany would have ended the war before 1918, with 1917 a likely termination year. (Goemans 2000, 314–15, in conclusions)

    Ferguson:
    Rather than joining the Allied war effect, he [Ferguson] said, Britain should have maintained its neutrality and allowed the Germans to win a limited Continental war against the French and the Germans. In that event, he postulated, Germany whose war aims in 1914 were relatively modest, would have respected the territorial integrity of Belgium, France, and Holland and settled for a German-led European federation. Had Britain "stood aside" he continued, it is likely that the century would have been spared the Bolshevik Revolution, the Second World War, and perhaps even the Holocaust. (Boynton 1999, 43; New Yorker piece on Ferguson's book)

    Geiss:
    The events of July and early August 1914 cannot be properly understood without a knowledge of the historical background provided by the preceding decades of Imperialism. On the other hand, that background alone is not sufficient to explain the outbreak of the First World War. Two general historical factors proved to be decisive, and both were fused by a third to produce the explosion known as the First World War. Imperialism, with Wilhelmine *Weltpolitik* as its specifically German version, provided the general framework and the basic tensions; the principle of national self-determination constituted, with its revolutionary potential, a permanent but latent threat to the

old dynastic empires and built-up tensions in south-east Europe. The determination of the German Empire – then the most powerful conservative force in the world after Tsarist Russia – to uphold the conservative and monarchic principles by any means against the rising flood of democracy, plus its *Weltpolitik* made war inevitable. (Geiss 1984, 46; original is the introduction of his book *July 1914*)

Taylor:
But there was only one decision which turned the little Balkan conflict between Austria-Hungary and Serbia into a European war. That was the German decision to start general mobilization on 31st July, and that was in turn decisive because of the academic ingenuity with which Schlieffen, now in his grave, had attempted to solve the problem of a two-front war. (Taylor 1969, 101)

Maier:
The irreversible momentum toward general war in 1914 is usually seen as a result of three factors: the hopeless, long-term instability of the Habsburg empire, the rigid structure of opposing alliances, and the inelectuable pull of military preparations. (Maier 1988, 822)

Williamson:
Of all the central actors in 1914, Conrad alone could have – by saying no to Berchtold or expressing hesitation to Franz Joseph or accepting some modified "Halt in Belgrade" – brought the crisis to a more peaceful conclusion. (Williamson 1988, 815–16)

Van Evera:
The consequences of the cult of the offensive are illuminated by imagining the politics of 1914 had European leaders recognized the actual power of the defense. . . . Thus the logic that led Germany to provoke the 1914 crisis would have been undermined, and the chain reaction by which the war spread outward from the Balkans would have been very improbable. In all likelihood, the Austro-Serbian conflict would have been a minor and soon-forgotten disturbance on the periphery of European politics. (Van Evera 1984, 105)

Snyder:
These war plans and the offensive doctrines behind them were in themselves an important and perhaps decisive cause of the war. (Snyder 1984, 108) In short, the European militaries cannot be blamed for the belligerent diplomacy that set the ball rolling towards World War I. Once the process began, however, their penchant for offense and their quickness to view war as inevitable created a slide towards war that the diplomats did not foresee. (Snyder 1984, 138–39)

Trachtenberg:
If in 1914 everyone understood the system and knew, for example, that a Russian or German general mobilization would lead to war, and if, in addition, the political authorities were free agents – that is, if their hands were not being forced by military imperatives, or by pressure from the generals – then the existence of the system of interlocking mobilization plans could hardly be said in itself to have been a "cause" of war because, once it was set off, the time for negotiation was cut short. But if the working of the system was understood in advance, a decision for general mobilization was a decision for war; statesmen

would be opting for war with their eyes open. To argue that the system was, in such a case, a "cause" of war makes about as much sense as saying that any military operation which marked the effective beginning of hostilities . . . was a real "cause" of an armed conflict (Trachtenberg 1990–91, 122)

Sagan:
If Grey had given a clear warning earlier, *if* the Czar had further delayed Russian mobilization against Austria and then Germany, and *if* the German offensive war plans had not been able to depend upon a preemptive *coup de main* against Liege and the decisive battle in France before Russian mobilization was completed in the East, *then* it is possible, just possible, that Bethmann-Hollweg would have had the time and the courage necessary to apply sufficient pressure on Vienna to accept the "Halt in Belgrade." And if this had occurred, 1914 might today appear as only another one of series of Balkan crises that almost led to a world war. (Sagan 1986, 169)

Riker:
The process of confusion concerning the cause of α [World War I], which, being ambiguous, cannot properly be said to have a cause, starts with the assertion that the cause of B [the Austro-Serbian war] is the assassination at Sarajevo, event D. It can perhaps be demonstrated by valid arguments – or at least by arguments likely to be accepted by many historians – that the assassination of Franz Ferdinand is a sufficient condition of the Austro-Serbian war; and (somewhat dubiously, however) it can even be argued that it is was a necessary condition. (Riker 1957, 63–64)

Causation, causal hypotheses, causal asymmetry, etc.

106. Cartwright et al. (2024) discuss the following story from the philosophy literature on causation. Draw a little causal figure to illustrate what is going on. Hint: be sure to put signs above the causal arrows.

> A villain intent on killing a walker pushes a boulder off a bluff to crush the walker. But just as the boulder approaches, the walker notices it and ducks. Pushing the boulder causes the walker to duck. The ducking causes the walker to continue to be alive. (Cartwright et al. 2024, 45)

Answer: Villain pushes boulder toward walker ==(+)==> Walker ducks ==(-)==> Walker stays alive. This is a nice example of the constraint-deterrence model discussed in the causal mechanism-model chapter of the LNQA book. Cartwright likes to call these derailers which is the same basic idea. As in constraint–deterrence examples, the killer is motivated to kill the victim and pushes a boulder, which is a positive cause of death. However, the walker notices that and steps out of the way which is a negative cause. This is a situation where the second cause essentially cancels out the first and one ends up in the status quo ante.

107. The quote below talks about a famous article on smoking and cancer. It raises some interesting questions. The first is the exclusion of women, which

raises questions about external validity. One general question is whether the strength of the original findings should or does influence our willingness to extrapolate to new populations? What is actually the strength the finding discussed below, i.e., 0.00 versus 1.14? We are accustomed to focusing on the difference between these two numbers, but what if one looked more closely at the 0.00 finding? How might that be seen as a very strong regularity? The second question deals with mechanisms. There are powerful arguments for external validity if the mechanism would be the same in the target population? How might that play out in extrapolating from men to women in this study?

> THE SMOKING GUN. We close with the one of the most influential articles of the 20th century, to demonstrate how even very narrowly defined study samples may provide widely generalizable results if conducted with an eye to rigorous internal validity. Entitled "The mortality of doctors in relation to their smoking habits: a preliminary report," this 1954 article by Doll and Hill concerned the association between lung cancer and cigarette smoking in British physicians. The investigators excluded physicians who did not return usable responses, and also women physicians, and physicians aged 35 years, because of their low expected frequency of lung cancer deaths. The remaining sample was a male, physician cohort of 24,389, about 40% of those in the Medical Register. During the 29-month follow-up, investigators observed only 36 confirmed lung cancer deaths, occurring at rates of 0.00 per 1,000 in nonsmokers, and 1.14 per 1,000 among smokers of 25 or more grams of tobacco per day. (Kukull and Gangul 2012, 1888–89)

Answer: For question (1), 0.00 might mean that *no nonsmokers died of lung cancer*. In fact, of the 36 people who died of lung cancer in the study, all were smokers. This could be interpreted as smoking is a necessary condition for lung cancer death. This then means that all mechanisms leading to lung cancer death must include smoking. The second question might be about the extend to which men and women's lungs are different and function differently. If they function the same (i.e., same mechanism) then we might feel that generalizing to women is more valid since we have good reasons to think that the mechanism linking smoking to cancer is will be the same in women.

108. Here is another exercise related to the quote above about smoking and cancer related to selection. The first is the exclusion of women. Given when this study was done they were probably few women physicians to begin with that could be sampled. Is that a reason to exclude them? Given the standard societal norms at the time it is much less likely that women were smokers than men. So is that a selection effect or an antecedent condition that we should worry about? The issue is whether the selection condition has any plausible causal connection to the outcome. Another way to phrase the question is whether we would be in trouble if all we did was just control for gender? Still, we have a real causal chain whereby societal norms lead to women smoking less which leads them to getting less cancer. So do we need to worry about the first link in that causal chain, i.e., endogeneity?

    Answer: I think most people would say no. This is because arguing that societal norms about women smoking cause cancer beyond them not smoking

seems implausible, i.e., in DAG language there is no direct link between societal norms about smoking and cancer. As discussed in the causal model chapter of the LNQA book, a direct link means that there must be another mechanism leading from societal norms to cancer which does not go through smoking: what would be that mechanism? If there are no really plausible ones then we do not need to worry about a direct path to cancer. So while there is clearly selection there is no other causal mechanism linking selection directly to outcomes hence perhaps no reason to worry about it statistically or otherwise. There might be some other feature of women (e.g., a famous Fisher genetic cause) that produces less cancer but that is separate from the norms causing cancer.

109. The debate about how to think about necessary and sufficient conditions in a statistical context has been common for many years. Discuss the negative position of Gailmard – widely influential – on this issue. Note that at the end he states that what matters is the data generating process (DGP); what if these are modeled as necessary or sufficient conditions? Does this contradict the notion of stochastic?

> To be (somewhat more) specific, suppose we are interested in explaining a social event or outcome Y that can either be "large" or "small" in value. This might be, for instance, the extent of a civil war in a group of countries. And we want to know whether some factor X that can either be "present" or "absent" helps us to explain the outcome. Suppose that the observed cases line up with the attributes in Table 3.1. In this data we can "observe" that the presence of factor X is a necessary condition for Y to be large; Y is never large otherwise. On the other hand, the absence of factor X is sufficient for Y to be small; whenever X is absent this value of Y ensues. This all makes perfect sense until it is revealed that X and Y are both coins. "Present" means the first one came up heads when flipped and "absent" means it came up tails; "Large" means the second one came up heads when flipped, and "small" means it came up tails. When we know the DGP and we know it is stochastic, it is transparent that methods to derive necessary and/or sufficient conditions for a particular result Y to occur are faulty. . . . In this book, DGPs are formalized as probability distributions. . . . The hunt for necessary and sufficient conditions is misguided in stochastic DGPs. (Gailmard 2014, 80, 82)

Table 11: Necessary and sufficient condition data analysis

TABLE 3.1. *Observations for Outcome Y and Explanatory Factor X*

Case	*X* Attribute	*Y* Attribute
1	Absent	Small
2	Present	Large
3	Present	Large
4	Present	Small
5	Absent	Small

Source: Gailmard 2014.

Answer: See Dion's Bayesian probabilistic analysis of necessary conditions reproduced in chapter 5 of LNQA book.

110. Necessary conditions as causes are often connected to scope conditions which are also often causes. Discuss the following claim and whether it should be formulated as a scope condition:

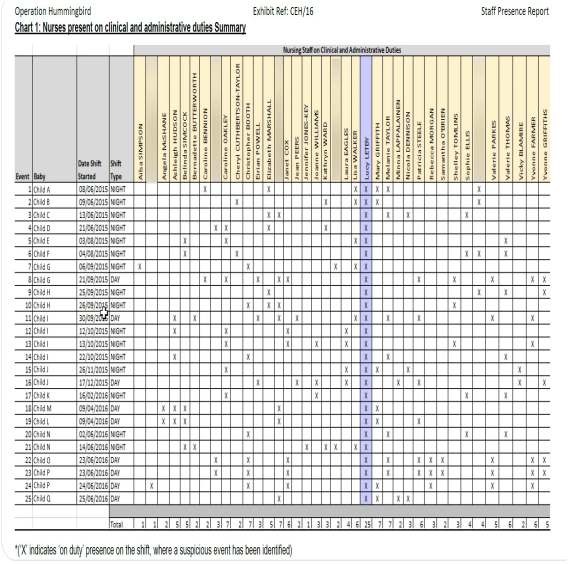
> There is another reason why the notions of difference-maker and cause come apart. This is where we have a necessary condition for something to occur but where it is not really a causal contributor to it. For example, having a liver is a necessary condition for having cirrhosis. No one could have cirrhosis unless they had a liver; therefore having a liver is a difference-maker with respect to that disease. But it would seem wrong to say that the liver caused the disease even though it is a necessary condition for having it. (Anjum and Mumford 2018, 127)

111. In a long *New Yorker* (2024) article (see also the ITV documentary on the case) Aviv discusses a very famous case in the UK involving a nurse Lucy Letby convicted of killing six newborn babies and attempted killing of five others. The core evidence was a very strong regularity which placed her at the "scene of the crime," see table 12 below. A huge debate revolved about this regularity from a statistical point of view. Apparently the regularity was what convinced many. Discuss the Russo-Williamson hypothesis in this connection, i.e., the mechanism of death of the babies. Was there strong mechanism evidence?

> Letby had worked on a struggling neonatal unit at the Countess of Chester Hospital,run by the National Health Service, in the West of England, near Wales.The case centered on a cluster of seven deaths, between June, 2015, and June, 2016. All but one of the babies were premature; three of them weighed less than three pounds. No one ever saw Letby harming a child, and the coroner did not find foul play in any of the deaths. . . . The case against her gathered force on the basis of a single diagram shared by the police, which circulated widely in the media [see the table below]. On the vertical axis were twenty four "suspicious events," which included

the deaths of the seven newborns and seventeen other instances of babies suddenly deteriorating. On the horizontal axis were the names of thirty-eight nurses who had worked on the unit during that time, with X's next to each suspicious event that occurred when they were on shift. Letby was the only nurse with an uninterrupted line of X's below her name. She was the "one common denominator". . . . Letby had become the country's most reviled woman – "the unexpected face of evil," as the British magazine *Prospect* put it – largely because of that unbroken line. (Aviv 2024, 34)

Table 12: Regularities and Letby's conviction of murder of newborn babies



Source: Aviv 2024.

112. One of the fundamental issues in concept analysis which is closely related to causal inference are the twin issues of conceptual and causal asymmetry. In statistical work these are virtually always assumed to go together. But many would claim that for example peace is not the absence of war and that war is not the absence of peace. This connects to the causal asymmetry issue whether the causes of peace are symmetric to the causes of war. While this is typically implicit Blainey in his classic work on war is very explicit about this:

> This book is based on a survey of all the international wars fought since 1700. It argues that in war and peace there are revealing patterns and clues that have been overlooked. As the book is more like an intellectual detective story than a narrative history, the preface may if so desired be ignored now and read as an epilogue. . . . The first part of the book points to weaknesses in well known explanations of peace. . . . The second part of the book examines ingredients which are usually prominent in determining a nation's decision to fight or not to fight; and the overall influence of those ingredients is summarised in the chapter called 'The Abacus of Power'. . . . The last part of the book offers the radical conclusion that the beginning of wars, the prolonging of wars and

> the prolonging or shortening of periods of peace all share the same causal framework. The same explanatory framework and the same factors are vital in understanding each stage in the sequel of war and peace. . . . At first sight the suggestion may seem absurd but this is simply another way of saying that the transition from war to peace is essentially the reverse of the transition from peace to war. . . . For every thousand pages published on the causes of wars there is less than one page directly on the causes of peace. And yet the causes of war and peace, logically, should dovetail into one another. A weak explanation of why Europe was at peace will lead to a weak explanation of why Europe was at war. A valid diagnosis of war will be reflected in a valid diagnoses of peace. (Blainey 1988, ix, 3)

113. As a general rule quantitative scholars are not very interested in the relationships between their various independent variables. However, from a conceptual and causal inference point of view there are important issues which often arise. A classic one is when two independent variables are somehow part of the same concept. Barnhart et al. (2020) illustrate this where they have two core variables one which is democracy, using polity data, and the second women's suffrage. Arguably, see Paxton's work, women's suffrage is part of the concept of democracy. However, many traditional democracy datasets do not include that factor, including polity and others. They make strong claims about the role of women's suffrage in connection with international conflict and more specifically the democratic peace literature:

> Finally, we examine cross-national conflict data to assess the impact of the enfranchisement of women. We find that female suffrage, when coupled with democratic institutions, is not only an important and powerful cause of peace, but that it may even be necessary for the democratic peace. Without female suffrage, democracies (whether examined dyadically or monadically) do not show clear evidence of being more pacific than non-democracies. These findings hold across time periods and do not appear to result from other confounding factors, or from the expansion of the political franchise more generally. By bridging the individual and international levels to show how the gender gap on the use of force matters, we contribute to the growing positivist study of gender and international relations. (Barnhardt et al. 2020, 634)

Someone might argue that it is not women's suffrage that is really the core causal factor but rather *quality of democracy*. Countries with women's suffrage, or more generally, fewer restrictions to suffrage based on race or whatever are higher-quality democracies, *by definition*. So maybe the article should be called the higher-quality democracy peace? (See Goertz and Dul 2023 for an argument and data analysis to this effect).

114. Y generalizations are closely connected to selecting on the dependent variable concerns because one is selecting cases of Y equals one. Ever since King, Keohane, and Verba (1994) these have been controversial and the standard line is still one should never do that and one cannot learn anything from doing it. For example, Bueno de Mesquita Bueno de Mesquita and Fowler (2021) have a whole chapter on this topic in their textbook. Discuss the quote below by Pearl and McKenzie which is using the canonical smoking and lung cancer example. Notice how Pearl and McKenzie state that the probabilities

are backwards, probability of X given Y. They claim that one knows nothing based on these data that would be relevant to a decision about smoking. Explain why that is false.

> Of course Hill knew that an RCT was impossible in this case, but he had learned the advantages of comparing a treatment group to a control group. So he proposed to compare patients who had already been diagnosed with cancer to a control group of healthy volunteers. Each group's members were interviewed on their past behaviors and medical histories. To avoid bias, the interviewers were not told who had cancer and who was a control. . . . The probability logic is backward too. The data tell us the probability that a cancer patient is a smoker instead of the probability that a smoker will get cancer. It is the latter probability that really matters to a person who wants to know whether he should smoke or not. The results of the study were shocking: out of 649 lung cancer patients interviewed, all but two had been smokers. (Pearl and Mackenzie 2017, 192)

Answer: The regularity that the research discovered is extremely strong, .97. So one might conclude that a necessary condition for lung cancer is smoking. Notice that this does not tell you a lot about what Pearl and McKenzie want to know which is what is the probability of lung cancer if one smokes. But it does tell you that if you do not want to get lung cancer you are almost certain not to get it if you do not smoke. Smoking is part of virtually all paths to lung cancer. Certainly a useful bit of information.

115. Y generalizations are closely connected to selecting on the dependent variable concerns because one is selecting cases of Y equals one. Ever since King, Keohane, and Verba (1994) these have been controversial and the standard line is still one should never do that and one cannot learn anything from doing it. For example, Bueno de Mesquita Bueno de Mesquita and Fowler (2021) have a whole chapter on this topic in their textbook. Discuss the quote below by Pearl and McKenzie which is using the canonical smoking and lung cancer example. Notice how Pearl and McKenzie state that the probabilities are backwards, probability of X given Y. They claim that one knows nothing based on these data that would be relevant to a decision about smoking. Explain why that is false.

> Of course Hill knew that an RCT was impossible in this case, but he had learned the advantages of comparing a treatment group to a control group. So he proposed to compare patients who had already been diagnosed with cancer to a control group of healthy volunteers. Each group's members were interviewed on their past behaviors and medical histories. To avoid bias, the interviewers were not told who had cancer and who was a control. . . . The probability logic is backward too. The data tell us the probability that a cancer patient is a smoker instead of the probability that a smoker will get cancer. It is the latter probability that really matters to a person who wants to know whether he should smoke or not. The results of the study were shocking: out of 649 lung cancer patients interviewed, all but two had been smokers. (Pearl and Mackenzie 2017, 192)

116. One of the fundamental issues in concept analysis which is closely related to causal inference are the twin issues of conceptual and causal asymmetry. In

statistical work these are virtually always assumed to go together. But many would claim that for example peace is not the absence of war and that war is not the absence of peace. This connects to the causal asymmetry issue whether the causes of peace are symmetric to the causes of war. While this is typically implicit Blainey in his classic work on war is very explicit about this:

Discussed the extent to which his conceptualization of war–peace is symmetric and the extent to which his explanations of them are also symmetric

> This book is based on a survey of all the international wars fought since 1700. It argues that in war and peace there are revealing patterns and clues that have been overlooked. As the book is more like an intellectual detective story than a narrative history, the preface may if so desired be ignored now and read as an epilogue.... The first part of the book points to weaknesses in well known explanations of peace.... The second part of the book examines ingredients which are usually prominent in determining a nation's decision to fight or not to fight; and the overall influence of those ingredients is summarised in the chapter called 'The Abacus of Power'.... The last part of the book offers the radical conclusion that the beginning of wars, the prolonging of wars, the ending of wars and the prolonging or shortening of periods of peace all share the same causal framework. The same explanatory framework and the same factors are vital in understanding each stage in the sequel of war and peace.... At first sight the suggestion may seem absurd but this is simply another way of saying that the transition from war to peace is essentially the reverse of the transition from peace to war.... For every thousand pages published on the causes of wars there is less than one page directly on the causes of peace. And yet the causes of war and peace, logically, should dove tail into one another. A weak explanation of why Europe was at peace will lead to a weak explanation of why Europe was at war. A valid diagnosis of war will be reflected in a valid diagnoses of peace. (Blainey 1988, ix, 3)

117. Halpern (2016) presents a theory of causation closely connected and developed with Pearl's approach to causal inference. Being a computer scientist it is natural that the "structural equations" he uses are Boolean. He argues that one can include probabilities by making changing the Boolean dichotomous variables to continuous one [0,1]. Discuss whether fuzzy logic or probabilities should be the interpretation of these [0,1] variables.

118. Halpern (2016) presents a theory of causation using Boolean models. These equations have an equal sign (=). How should this be interpreted causally? Is he implicitly treating this to mean "is sufficient for?" How similar or different is "=" from "→"? In addition to what extent do these symbols have causal interpretations? In general do both "=" and "→" receive causal interpretations? Note that the same interpretation issues arise in QCA.

    Answer: Halpern does not really discuss the meaning of = beyond this short comment: "Some variables may have a causal influence on others. This influence is modeled by a set of structural equations. For example, if we want to model the fact that if the arsonist drops a match or lightning strikes then a fire starts, we could use the variables MD, FF, and L as above, with the equation FF = max(L,MD); that is, the value of the variable FF is the maximum of the values of the variables MD and L. This equation says, among other things,

that if MD = 0 and L = 1, then FF = 1. The equality sign in this equation should be thought of more like an assignment statement in programming languages; once we set the values of MD and L, the value of FF is set to their maximum".

119. As a general rule quantitative scholars are not very interested in the relationships between their various independent variables. However, from a conceptual and causal inference point of view there are important issues which often arise. A classic one is when two independent variables are somehow part of the same concept. Barnhart et al. (2020) illustrate this where they have two core variables one which is democracy, using polity data, and women's suffrage. Arguably, see Paxton's work, women's suffrage is part of the concept of democracy. However, many traditional datasets do not include that factor, including polity and others. They make strong claims about the role of women's suffrage in connection with international conflict and more specifically the democratic peace literature:

> Finally, we examine cross-national conflict data to assess the impact of the enfranchisement of women. We find that female suffrage, when coupled with democratic institutions, is not only an important and powerful cause of peace, but that it may even be necessary for the democratic peace. Without female suffrage, democracies (whether examined dyadically or monadically) do not show clear evidence of being more pacific than non-democracies. These findings hold across time periods and do not appear to result from other confounding factors, or from the expansion of the political franchise more generally. By bridging the individual and international levels to show how the gender gap on the use of force matters, we contribute to the growing positivist study of gender and international relations. (Barnhardt et al. 2020, 634)

Someone might argue that it is not women's suffrage that is really the core causal factor but rather *quality of democracy*. Countries with women's suffrage, or more generally, fewer restrictions to suffrage based on race or whatever are higher-quality democracies, *by definition*. So maybe the article should be called the higher-quality democracy peace? (See Goertz and Dul 2023 for an argument and data analysis to this effect).

120. Pearl and McKenzie (2017) spend almost a whole chapter on causal inference in the history of smoking and lung cancer. They note one of the earliest results, and very striking results, which people of the time could not really wrap their heads around:

> Of course Hill knew that an RCT was impossible in this case, but he had learned the advantages of comparing a treatment group to a control group. So he proposed to compare patients who had already been diagnosed with cancer to a control group of healthy volunteers. Each group's members were interviewed on their past behaviors and medical histories. To avoid bias, the interviewers were not told who had cancer and who was a control. The results of the study were shocking: out of 649 lung cancer patients interviewed, all but two had been smokers. . . . The probability logic is backward too. The data tell us the probability that a cancer patient is a smoker instead of the probability that a smoker will get cancer. It is the latter probability that really matters to a person who wants to know whether he should smoke or not. (Pearl and Mackenzie 2017, 192–93)

Why is their question about smoking-cancer framed in the wrong way given these data?

Answer: Note that they say that the probabilities are backward which is exactly the issue with necessary conditions and analytic induction. As in the previous exercise, they are asking what is the probability that I will get lung cancer if I smoke. This is the wrong question given the data in question. As Pearl and MacKenzie note these data are not very useful in answering the question "What is the probability that I will get lung cancer if I smoke. However, if I want to *avoid* lung cancer the advice is clear and the data speak directly to that. The question is "What is the probability I will not get lung cancer if I do not smoke?" While there may be other contributing causes to all the cases of lung cancer in 99 percent of cases smoking is part of the story (an necessary condition).

121. Bernard and his colleagues in their analyzing qualitative data textbook discuss a technique known as analytic induction (Ragin has a book manuscript in progress on this topic). Typically analytic induction is looking at all the $Y = 1$ cases and then inductively trying to find a common cause. One of the earliest and strongest critiques of this was that this procedure did not allow one to "predict" outcomes. This meant that if one knew the common cause one cannot predict the outcome in other cases. Discuss this issue in terms of analytic induction finding necessary conditions and what one can predict if one has a necessary condition. Also discuss extent to which these authors when they talk about "prediction" are really asking about sufficient conditions. If one has a sufficient condition for an outcome then one can predict that it will occur when that sufficient condition occurs.

> Cressey [a classic of analytic induction] could not predict, a priori – i.e., without data about actual embezzlers who had been arrested and jailed for their crime – which bank workers would violate the trust of their employers. Manning could not predict, a priori, which pregnant women would ultimately seek an abortion. And Zeoli et al. couldn't predict which mothers' behavior would not confirm their expectations. . . . Analytic induction fell out of favor after the 1950s because the method accounts for data you've already collected but does not allow prediction about individual cases. While it does not produce perfect knowledge for the prediction of individual cases, it can do as well as statistical induction – the standard in social science – in predicting the outcome in aggregates of cases, and it does so with a relatively small number of cases. (Bernard et al. 2017, 563, 579)

122. Kacowicz claims that permissive or enabling conditions are *not* causes. But how can that be if they have an effect on something? This is perhaps not an uncommon point of view among social constructivists. What is his implicit concept of "cause"?

> I do not claim that international peace and globalization cause illicit transnational flows, but rather that they are permissive conditions that enable their occurrence and proliferation. . . . We premise the occurrence and proliferation of illicit transnational flows across peaceful borders upon two permissive conditions: the preexistence of international peace and the

> impact of globalization. . . . The implied logic here is that we expect more illicit transnational activities to take place under conditions of peace than war. We postulate that the existence of international peace is a permissive condition, (Kacowicz et al. 2021, x, 27)

123. Dion (2003) in a famous article disputes the following claim, at least a specific version of it. Assuming that the sample is random what could one potentially say? Hint: think of each independent variable as a coin being flipped.

> In a frequentist framework, which treats probabilities as constituting the likelihood that a sample drawn from a population is or is not representative of the population, nothing can be said about five or six cases with seven or eight independent variables because of the "degrees of freedom" problem. (Bennett 2022, 67)

Answer: Dion argued using Bayesian setup that if one consistently gets $X = 1$ when $Y = 1$, then with only 5–6 $Y = 1$ cases one can be quite certain, in a Bayesian sense, that X is necessary for Y. This like flipping a coin 6 times and always getting a head? What might you conclude about the coin? Braumoeller and Goertz (1999) find basically the same thing using a frequentist approach.

124. Bove and Nisticó (2014)use a relatively new methodology called the synthetic control methodology which has generated much interest. The title of their article is "Coups d'état and defense spending: a counterfactual analysis." Are they actually doing a counterfactual in the traditional sense?

Answer: As the name indicates the methodology generates a *control* comparison observation. One constructs a control group to compare with their military coup cases. A comparison of treatment and control is not obviously a counterfactual in the classic sense of the term.

125. In general it is not recommended to define via negation. This typically raises the serious possibility of significant heterogeneity in the "not" population. Discuss this issue using the concept of "noncontrolled comparison" which is the central topic of *Rethinking comparison: innovative methods for qualitative political research* (Simmons and Smith 2021) The authors of the anthology are mostly interpretivist but there might be other kinds of methodologies which are noncontrolled but which are also not interpretivist. See the question in the LNQA section, page 30, which raises this possibility.

126. One problematic concept situation is when the conceptualization and coding of the dependent variable involves important causal claims. Discuss this for Chenowith and Stepan's influential work:

> Success and failure are also complex outcomes, about which much has been written. For our study, to be considered a "success" a campaign had to meet two conditions: the full achievement of its stated goals (regime change, anti-occupation, or secession) within a year of the peak of activities and a discernible effect on the outcome, such that the outcome was a direct result of the campaign's activities. The second qualification is important because in some cases the desired outcome occurred mainly because of other conditions. The Greek resistance against the Nazi occupation, for example, is not coded as a full success even though the Nazis ultimately

> withdrew from Greece. Although effective in many respects, the Greek resistance alone cannot be credited with the ultimate outcome of the end of Nazi influence over Greece since the Nazi withdrawal was the result of the Allied victory rather than solely Greek resistance. (Chenowith and Stepan 2011, 10)

127. Often with complex concepts or interaction terms it is not clear what the comparison group might be that one needs for causal analysis. The generic "compared to what?" question. For example, "H1: In states that recently experienced a revolution, personalist dictatorship is associated with greater militarized dispute initiation" (Colgan and Weeks 2015, 170). What is the comparison group since the hypothesis involves the interaction between (1) revolution and (2) personalist dictatorship?

128. "Hypothesis 3: Coups are more likely in states involved in civil conflicts" (Piplani and Talmadge 2016, 1375). Conceptual tautology occurs here because some coups are also civil wars. Discuss how yo might fix this.

129. Causation is frequently built into concepts. Discuss the example of "environmentally displaced persons," which is probably built into international legal conventions.

130. Is "Ethnic conflict" another term where causation is built in?

> Ethnic conflict refers to a very specific phenomenon. Ethnic conflict encompasses all forms of small- and large-scale acts of violence between and among different ethnic groups (Some definitions of ethnic conflict argue that ethnicity must also play a causal role in the conflict – something difficult to determine in practice. See Sambanis 2001; and Lake and Rothchild 1996). (Brancati 2006, 654)

131. Discuss these conceptualizations of electoral violence in terms of a causal relationship:

> At the most abstract level, electoral violence can be understood as any event in which the use of coercive force coincides with the electoral process. A number of more contextually-specific definitions have also been advanced. The United Nations Development Program (UNDP) defines electoral conflict as: "acts or threats of coercion, or physical harm perpetrated to affect an electoral process or that arise in the context of electoral competition." ... There are two key components to all these definitions: the temporal link between violence and elections and the causal link between the two. Electoral violence is conventionally understood as violence that takes place during the electoral cycle, including the pre-electoral, electoral, and post-electoral periods. The causal link, which is often more implicit, limits electoral violence to that which is in some way connected to the electoral process, as opposed to violence that takes place during the electoral process but has no direct bearing on the election. (Birch and Muchlinski 2017, 2–3)
>
> Therefore, the data include events of subnational electoral contention only if subnational elections and national elections took place in close temporal proximity *and* if it is plausible to expect that subnational contentious

> events influenced national elections or responded to both local and national elections. (Daxecker et al. n.d. Electoral Contention and Violence (ECAV) Codebook)

132. How does the adjective "civil" function here? Is it a subsetting one, a causal relationship, or what?

> The next (and possibly most important) strand of the evolution of the understanding of peace is the civil peace. According to this approach, every individual in society has the capacity to mobilize for peace from a variety of different perspectives, whether for disarmament, for international cooperation, or against violence, discrimination, and oppression. It relates to the historical phenomena of social direct action for political, economic, and identity reasons, of citizen advocacy and mobilization, in the attainment or defence of basic human rights and values. It is also related to pacifism in its main forms, where civil action is non-violent in principle. It has been strongly influenced by a wide range of social mobilization dynamics. Without the civil peace and its social forms of mobilization, international and constitutional frameworks would not be able to connect with ordinary people in order to represent their interests, identities, needs, and aspirations. (Richmond 2014, epub 116)

Answer:

It seems to function as a causal relationship, i.e., as *cause* of peace rather than concept of peace. It is about how actions of individuals can create peace.

133. There is often unclarity about whether something is a defining feature of the concept or rather the cause or effect of it. Discuss the extent to which emotions are a defining feature of war, a cause of war, or even ineffective of war.

> The third defining element of war is, according to Clausewitz, the emotions that accompany and drive the extreme activity of waging war. At minimum, there must exist a drive to coerce the enemy using utmost force, as is evident in Clausewitz's core definition of war as "an act of violence intended to compel our opponent to do our will" (75). War is fueled by – and gives rise to – intense emotions of hatred and enmity. Clausewitz argues that both the outbreak and the character of war are influenced by the types and intensity of emotions involved. The prevalence in a society of a warlike drive to dominate others is accordingly the third dimension of quality peace. Clausewitz notes that two sides will not prepare for and wage war against each other unless moved by emotions of enmity. But the type and intensity of emotion will vary, and the character of war will accordingly both react to the emotions and help shape those emotions. So, for example, the same political object may "elicit differing reactions from different peoples, and even from the same people over time. . . . Between two peoples and two states there can be such tensions, such a mass of inflammable material, that the slightest quarrel can produce a wholly disproportionate effect – a real explosion" (81). (Davenport et al. 2018, 262 epub)

134. Two level or SUIN models raise questions about what to include in the statistical analysis. An example on the independent variable side is Uzonyi's concept of a "big political opportunity" which takes three forms that are expressed in

his three hypotheses. Should he include each of the three forms separately in a regression analysis or just the higher variable formed by the logical OR?

> There are generally three types of domestic unrest that may lead to the onset of genocide or politicide. The first, which is captured by those scholars focusing on big opportunities for government mass murder, are militarized threats to a countrys political leadership. . . . Coups provide a similar logic. . . . The third type of threat a government faces is non-militarized unrest. Non-militarized domestic unrest, such as protests, riots and strikes, can spark political violence if the regime perceives that such unrest will grow into greater instability. (Uzonyi 2016, 318–19)

135. Discuss the claim below that it is better to look at the superset than the subset.

> For the sixty-two sovereign states included in the territorial contenders data set, Cunningham et al. identify ninety-seven rebel groups that controlled territory. The territorial contenders data set includes eighty-one of them (83.5 percent). Of the remaining sixteen rebels identified by Cunningham et al. as territorial, five are included among the ancillary materials of the territorial contender data set as near misses. The other eleven are instances where our coders could not identify any clear evidence of territorial control. Consultation with the source documents for the NonState Actor Dataset reveals no citations establishing territorial control for the rebel groups in question. We believe the codings in the territorial contender data set are valid, while accepting that there may be some occasional error in either our or the non-state actor (NSA) datasets. An 84 percent overlap is very high, nonetheless. *Absent an argument specific to a subset, there is considerable evidence in favor of analysis of the entire category of territorial contenders over analysis of any one subset*. (Lemke and Crabtree 2019, 19, emphasis is mine)

136. It is not uncommon that options for case selection form nested subsets. What are the research design and causal inference issues in one like this where genocides are subset of one-sided violence, also one-sided violence occurrence might be almost a subset of civil war occurrence.

> Harff (2003) presents a global examination of genocide in countries that have experienced state failure internal war or regime collapse covering the period 1955–97. Our dataset is different in that it consists of all conflict actors during the period 1989–2004 that are actively involved in an armed conflict resulting in at least 25 battle-deaths in a year. (Eck and Hultman 2007, 242)

137. In the *Two cultures* book a key methodological point is causal asymmetry. Hafner-Burton's (2014) review article does not talk about authoritarian regimes as a source of human rights abuses, while she talk about democracy as a cause of the respect of human rights. Discuss this causal asymmetry.

> *The central insight about state institutions is that, while fully developed democracy may be the best hope for the promotion and protection of human rights, efforts to transition to and institutionalize democracy often incite violence, and the alternatives to democracy are not all equally adverse for human rights*. (Hafner-Burton 2014, 275)

138. Blainey in his oft-cited book (1973) argues for causal symmetry in understanding war and peace. Discuss. For an opposing view see Goertz et al. *The puzzle of peace: the evolution of peace in the international system* where the final chapter argues for asymmetry.

> the causes of war and peace, logically, should dovetail into one another. A weak explanation of why Europe was at peace will lead to a weak explanation of why Europe was at war. A valid diagnosis of war will be reflected in a valid diagnosis of peace. (p. 3)
>
> War and peace appear to share the same framework of causes. The same set of factors should appear in explanations of the outbreak of war; widening of war by the entry of new nations; outbreak of peace; surmounting crises during a period of peace; and, of course, the ending of peace. (p. 293)
>
> Wars usually end when the fighting nations agree on their relative strength, and wars usually begin when fighting nations disagree on their relative strength. Agreement or disagreement is shaped by the same set of causal factors. Thus each factor that is a prominent cause of war can at times be a prominent cause of peace. (p. 122)

139. Geddes (2003) emphasizes the importance of variation on the dependent variable. But, variation on the independent variable is also important. Analyze her test of Skocpol's theory in terms of variation on the two key independent variables.

    Answer:

    She selects *only* cases of agrarian revolt for her analysis so there is no variation on this independent variable. What she is in fact testing is the interaction of agrarian revolt and state breakdown (which does have variation). Typically, statistical interaction terms include variation on both variables.

140. Recently statisticians have become much more concerned with problems of "unit homogeneity." Here is Henry Brady defining the idea:

> We shall make the transformation of $Y_B(1,0)$ into $Y_A(0,0)$ in two steps which are depicted on Table 10.9. If A and B are identical and Z_A and Z_B [Z is the treatment] are identical as well(footnote: By saying that Z_A and Z_B have to be comparable, we mean that $Z_A = 0$ and $Z_B = 0$ are the same thing and $Z_A = 1$ and $Z_B = 1$ are the same thing.) (although we haven t indicated how this might be brought about yet) it might be reasonable to suppose that: $Y_B(1,0) = Y_A(0,1)$, [*Identicality of units and treatment or Unit Homogeneity*]. That is, A and B are mirror images of one another so that the impact of $Z_A = 1$ and $Z = 0$ on B is the same as the impact of $Z = 0$ and $Z = 1$ on A. (Brady 2008, 258)

    Analyze how valid concepts are critical to the existence of unit homogeneity.

    Answer:

    Holland (1986) says "unit homogeneity" means that units are prepared carefully "so that they 'look' identical in all relevant aspects" (Holland 1986, 948). To be "identical" means that the concept must produce identical units, A and B in the Brady analysis. If a study looks at all "states" then the conceptualization of state means that there are no casually important differences in

"stateness." Treatments, Z in the Brady analysis, also are assumed to be homogeneous. So if democracy is an dichotomous independent variable that means that variations in democracy have no causal impact on the dependent variable.

141. As discussed at various points in *Two cultures* the set theoretic approach to causation is asymmetric. A variable might be good at explaining not-Y but weak at explaining Y. Take the following quotes and (1) formulate the strong generalization in terms of a necessary condition, (2) explain how asymmetry solves Hensel's constant independent variable problem to a large degree. Note that correlation between contiguity and militarized disputes and wars is considered one of the strongest in the international conflict literature:

> 5 percent of fatal disputes [aka MIDs] and 5 percent of full-scale wars between minor powers begin between adversaries that do not share a land or sea boundary. (Hensel 2012, 6)
>
> In short, proximity increases the risk of armed conflict .... Proximity between states is basically a constant; countries rarely gain or lose borders over time. As a result the proximity explanation faces difficulties in accounting for the outbreak of rare events such as conflict and war, since a dependent variable cannot account for fluctuations in a dependent variable. (Hensel 2012, 8).

Answer:

Contiguity is a necessary condition for minor power wars.

Asymmetry arises because non-contiguity gives a very good explanation for the lack of severe militarized conflict, but does a poor job of identifying when wars actually break out. Notice that non-war is a almost constant for most dyads. This is of course the converse of the fact that wars are rare events. So both are very constant variables.

142. Discuss the causal figures in *Multimethod research, causal mechanisms, and case studies: an integrated approach*, Weller and Barnes (2014) or Waldner (2014) from a DAG cross-case, large-N, probabilistic perspective. Are these figures identified from a DAG perspective (e.g., Morgan and Winship 2015). Does it matter from a within-case causal inference perspective if they are not?

143. Discuss Pahre's claim that necessary conditions violate common statistical assumptions: "Third, necessary conditions violate the unit homogeneity assumption common in statistics, which states that 'if two units have the same value of the key explanatory variable, the expected value of the dependent variable will be the same' (King et al., 1994: 116). Necessary conditions violate this condition because observations x, y and x, not-y are both consistent with the necessary condition x $\leftarrow$ y; in other words, x may occur with or without y" (Pahre 2005, 131).

144. A key assumption in multidimensional concepts is that the various means of achieving a given aggregate level are equivalent (e.g., 2+3 and 4+1 are really the same). What is one criterion for defending equivalence? Bueno de Mesquita et al. (2005) provides a very good discussion of these issues.

Answer:

One common criterion for equivalence is similar causal effect in some key hypotheses. Bueno de Mesquita et al. (2005) illustrates this.

145. Homogeneity of the zero cases when the dependent variable is coded dichotomously is often very problematic because they include quite different outcomes all coded as zero. Kroenig (2009) is looking at nuclear states supplying sensitive technology to non-nuclear states. It is quite likely that different causal mechanisms are producing these different zero outcomes. The dependent variable is received sensitive nuclear assistance or not. Discuss how the "not" case may contain two (or perhaps more) quite different kinds of zero.

Answer:

The zero cases include those countries who never tried to get such technology and those states who requested it but were turned down. See Kroenig's Appendix C "Selected cases of nonassistance" for some examples, such as "Argentina to Iran (1992). Argentina denied an Iranian request for hot cells (Hibbs 1992)."

146. Sometimes concepts should not build in causal hypotheses. Evaluate the Cohen and Pavoncello (1987) critique of Schmitter's concept of corporatism based on the idea that there are causal hypotheses built into definition.

147. Thomas Franck is well-known for his work (e.g., 1988, 1990, 1992) on the importance of "legitimacy" in international politics. Discuss the theoretical relationship between the concept of legitimacy and four factors related to it. Is the relationship best seen as (1) indicator, (2) ontological, (3) causal, and if causal in which direct do the causal arrows go? For example, Franck says "one may postulate four indicators: *pedigree, determinacy, coherence,* and *adherence*. . . . The content of these four indicators of legitimacy" (Franck 1992, 51).

Answer:

In the 1992 article, Franck uses the term "indicator." However, much of the discussion could lead one to see these four factors as *causes* of legitimacy.

148. A central guideline in building concepts is the consideration of causal relationships involving secondary-level dimensions of a concept. The causal relationships that involve the various secondary-level dimensions open up the possibility that there is significant causal heterogeneity in concepts and hence that causal generalizations are likely to be problematic. Little appears to be making this argument:

> Another way to state my position is to consider whether there are social kinds, analogous to natural kinds. A natural kind is a set of entities that share a common causal structure, and whose behavior can therefore be predicted on the basis of the laws that govern the behavior of such entities (Putnam, 1975b). . . . I deny that any social concepts serve to identify social kinds in the strong sense outlined above. Instead, social concepts function as ideal types or cluster concepts, permitting us to classify a range of diverse phenomena under a single concept. The notion of a cluster

> term captures many scientific concepts-terms that encompass a variety of phenomena that share some among a cluster of properties (Putnam 1975a, 50–54). An ideal-type concept is a complex description of a group of social phenomena that emphasizes some features and abstracts from others (Weber 1949). It is apparent that generalizations and predictions based on cluster concepts and ideal types demand a great deal of care. Since the entities that fall under such concepts do not share a homogeneous causal structure, we cannot infer that instances of the concept will behave in the typical way.... The metaphysical counterpart, then, to my view that there are no governing social regularities among social phenomena, is that there are no genuine social natural kinds. (Little 1993, 190–91).

Discuss Little's claims about "social kinds" versus "natural kinds." Also, explore the extent to which "cluster concepts" are the same as family resemblance concepts.

149. The concept of a welfare state is used as an ongoing example in the concept book. In particular, it is good because the use of expenditure data illustrates the completely fungibility and substitutability of the secondary-level dimensions. It could be that this assumption of substitutability masks causal heterogeneity. See Bonoli (2007) who argues that that causal processes dealing with "new social risks" (e.g., women's participation in the labor force and child care) are quite different from those dealing with classic welfare policies (e.g., unemployment insurance). How should you change the concept of the welfare state? Add an adjective?

150. A very difficult conceptual and methodological problem is dealing with norms, particularly norms that have the logical structure of rights, e.g., human rights. This problem has exercised me a great deal and in fact resulted in a book (Goertz *International norms and decision making*; a large part of that book was devoted precisely to analyzing the concept of a norm). (1) Discuss the concept of a right (contrasting it with prohibitions and obligations). (2) The positive and negative poles are interesting here in terms of behavior: do you focus on the positive exercise of rights (e.g., voting) or on the absence of the exercise of the right. This has theoretical consequences because in the former case you focus on what factors in addition to rights lead to the use of rights, while in the latter you focus on factors that eliminate the right itself in practice. (3) Rights are about *possibility*, relate this to the Possibility Principle of case selection and necessary conditions in general. Dowding and Hees (2003) is a very nice discussion of many of these issues. For measurement see also Rosenbaum (2000) and Carter (2000).

151. One serious issue with dichotomous data is the potential for serious heterogeneity in the zero or one category. This is particularly likely to be a problem in the zero category. The issue of heterogeneity arises almost by definition with nonordered, categorical data. For example, quantitative studies of international conflict that look at the outcome dispute variables that have anywhere from 3 to 9 categories. Often these can be reduced to settlement/compromise, win, draw. If the focus is, say, settlement or compromise then there is a tendency to code that as one and lump win and draw together as zero. Discuss the merits of this procedure. What would be alternatives?

Answer:

Lumping all these together in the zero seems particularly dubious when the causal mechanisms producing these zeros are likely to be quite different.

Within the context of the conflict literature and the militarized disputes dataset, it might make sense to just include draw/stalemate as the contrasting case since this constitutes often 70+% of the zero cases (e.g., Goertz et al. 2005).

This choice could well be brought into the research design and theory if the research were interested in making specific causal contrasts.

152. One of the key guidelines for constructing concepts involves thinking about causal relationships *within* concepts. There is a huge literature on the concept of "ethnicity." Discuss Chandra and Wilkinson's figure 1 (2008, 524) on there concept of ethnicity. Which of the the various factors listed does not really involve any causal relationships with outside variables?

    Answer:

    The "structure" factor is the only which really is *only* and *exclusively* about ethnicity. All the others involve others causal variables. For example, see the discussion of ECI measure and its application to India; ethnic group variable is an interaction between some measure of ethnic "structure" and colonial practices.

    "At the broadest level, we can imagine the term ethnicity as encompassing two families of concepts – the structure of ethnic identities and the practice of ethnic identification. Ethnic structure refers to the distribution of descent-based attributes – and, therefore, the sets of nominal identities – that all individuals in a population possess, whether they identify with them or not. Ethnic practice refers to the act of using one or more identities embedded in this structure to guide behavior. In other words, it refers to the set of activated identities that individuals employ in any given context. The set of activated ethnic identities for any given country is typically a subset of the identities contained in the ethnic structure." Chandra and Wilkinson (2008, 523)

153. State failure is the subject of much academic research as well as interest to policy organizations such as the World Bank and the Carnegie Endowment, not to mention the large state failure project. One of the important risks with complex concepts such as "failed state" (see Iqbal and Starr 2009 for a review) is that some of the secondary-level dimensions of the concept may also be viewed as causes or effects of the phenomenon. Examine the various concepts of state failure and determine which might really be seen as causes or effects of state failure rather than the concept itself.

    Answer:

    For example, as Iqbal and Starr (2009, 316) state: "Specifically, most existing measures or indices of state failure incorporate a number of factors that may, in fact, be determinants of state collapse – such as civil strife and poverty."

154. Birch (2012) in a book on the causes of "Electoral malpractice" includes the Freedom House democracy measure as a lagged (five year) independent variable. Explain why she runs the risk of "conceptual tautology" with her dependent variable of electoral malpractice.

155. Recently statisticians have become much more concerned with problems of "unit homogeneity." Here is Henry Brady defining the idea:

> We shall make the transformation of $Y_B(1,0)$ into $Y_A(0,0)$ in two steps which are depicted on Table 10.9. If A and B are identical and Z_A and Z_B [Z is the treatment] are identical as well(footnote: By saying that Z_A and Z_B have to be comparable, we mean that $Z_A = 0$ and $Z_B = 0$ are the same thing and $Z_A = 1$ and $Z_B = 1$ are the same thing.) (although we haven t indicated how this might be brought about yet) it might be reasonable to suppose that: $Y_B(1,0) = Y_A(0,1)$, [*Identicality of units and treatment or Unit Homogeneity*]. That is, A and B are mirror images of one another so that the impact of $Z_A = 1$ and $Z = 0$ on B is the same as the impact of $Z = 0$ and $Z = 1$ on A. (Brady 2008, 258)

Analyze how valid concepts are critical to the existence of unit homogeneity.

Answer:

Holland (1986) says "unit homogeneity" means that units are prepared carefully "so that they 'look' identical in all relevant aspects" (Holland 1986, 948). To be "identical" means that the concept must produce identical units, A and B in the Brady analysis. If a study looks at all "states" then the conceptualization of state means that there are no casually important differences in "stateness." Treatments, Z in the Brady analysis, also are assumed to be homogeneous. So if democracy is an dichotomous independent variable that means that variations in democracy have no causal impact on the dependent variable.

156. One potentially large issue is what might be called "conceptual endogeneity." This is where a secondary-level dimension is part of the dependent variable and also considered as an independent variable. There has been an explosion of studies of elections in hybrid democratic regimes. Should these elections be considered a *cause* of democratization? Does this conflict with the fact that freer and more competitive elections will change the democracy score of the country by definition? For example, "Scholars also disagree about the role of elections in promoting democratic change. (p. 12) While a familiar indicator of democratic progress, we must remember, electoral turnover does not always support democratic development. (p. 13) These issues lead to a final and more fundamental concern about attributing too much democratic influence to elections" (Bunce and Wolchik 2011, 14).

157. A very important methodological issue is set relationships between independent variables. In an original outline of the book a major part of a chapter was going to be devoted to the methodological issues that this raises. The main methodological concern arises when X_1 is a subset of X_2. For example, in the debate about the territorial peace versus democratic peace, it turns out that the the set of democratic dyads is a subset of the dyads at territorial peace:

"This study compares the conflicting answers of the democratic peace and the territorial peace and examines the empirical record to see which is more accurate. It finds that almost all contiguous dyads settle their borders before they become joint democracies" (Owsiak and Vasquez 2016, 339).

What is the causal interpretation of the subset variable when all of the variation in X_1 is taken up in X_2? Discuss how this is different from multicollinearity, for example, the two might not be very correlated. See Owsiak 2020 for a nice explicit discussion of the subsetting issue and other examples.

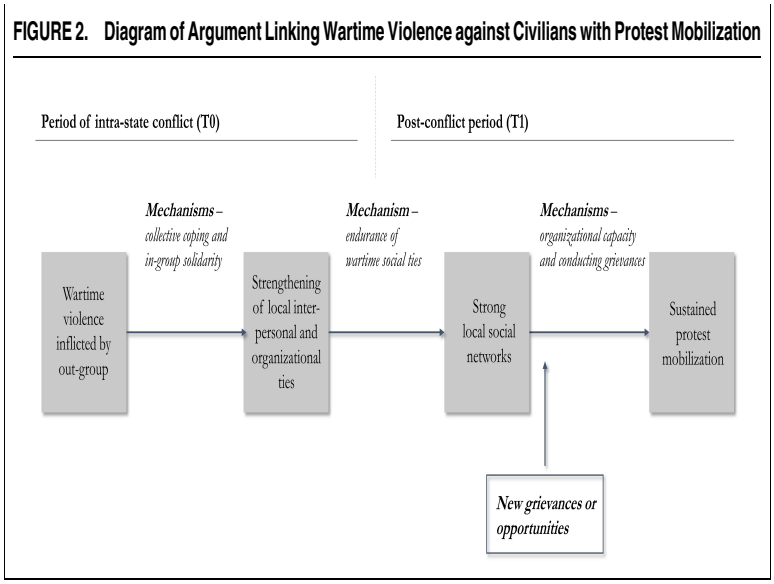
What if X_1 is a perfect subset of X_2 and one introduces an interaction term, $X_1 * X_2$?

This is just with dichotomous variables, but one can find set theoretic relationships among continuous variables (defined as X_1 less than or equal to X_2 for all observation or vice versa).

Causal model-mechanisms, causal model figures

158. Berman et al. (2024) provide a causal mechanism in their analysis of popular mobilization in Iraq. In particular discuss the factor "new grievances or opportunities" box in the model. It has an arrow going up to the final arrow in the figure. How are we to interpret what this arrow means? In terms of our modeling guidelines this looks like a moderator variable because it is an arrow to an arrow. One might argue that in fact it indicates an interaction term. What evidence can you provide just that the interaction term interpretation is correct?

Figure 22: Diagram of argument linking wartime violence against civilians with protest mobilization



Source: Berman et al. 2024.

Answer: One bit of evidence would that there is an interaction term in the statistical model.

If it is an interaction term discuss their case selection strategy for their case study in the article which is described as follows:

> Our quantitative results point to a positive relationship between violence legacies and protest, especially when the perpetrator of violence is from an out-group like a foreign invader or an ethnic militia. To further unpack the mechanisms behind this relationship, in this section, we conduct a brief case study of Fallujah, which witnessed some of the fiercest battles during the war and was a site of significant protest activity during 2011. We selected Fallujah to analyze for several reasons. First, it is a district whose protest activity was relatively high and is "well predicted" by our regression models . . . Second, it is a district where our models suggest that civilian casualties had a considerable impact on protest; when we remove civilian casualties from the model, the predicted protest levels in Fallujah fall well below observed values. Third, there are abundant historical and qualitative data on Fallujah, since it was a major site of conflict during the war. (Berman et al. 2024, 226)

Should one be selecting on the interaction term not only the individual original causal variable, civilian casualties, in the interaction model term?

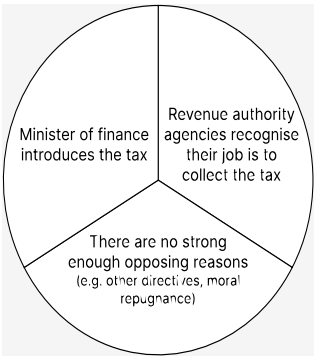
159. Cartwright et al. (2024) discuss the following story from the philosophy literature on causation. Draw a little causal figure to illustrate what is going on. Hint: be sure to put signs above the causal arrows.

> A villain intent on killing a walker pushes a boulder off a bluff to crush the walker. But just as the boulder approaches, the walker notices it and ducks. Pushing the boulder causes the walker to duck. The ducking causes the walker to continue to be alive. (Cartwright et al. 2024, 45)

Answer: Villain pushes boulder toward walker ==(+)==> Walker ducks ==(-)==> Walker stays alive. This is a nice example of the constraint-deterrence model discussed in the causal mechanism-model chapter of the LNQA book. Cartwright likes to call these derailers which is the same basic idea. As in constraint–deterrence examples, the killer is motivated to kill the victim and pushes a boulder, which is a positive cause of death. However, the walker notices that and steps out of the way which is a negative cause. This is a situation where the second cause essentially cancels out the first and one ends up in the status quo ante.

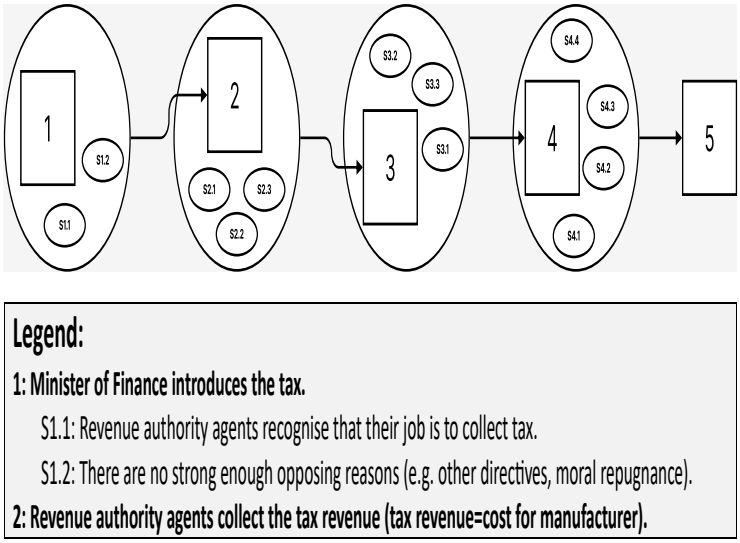
160. Cartwright and colleagues (2024) have a different approach from Goertz and Haggard, (causal model chapter), for diagramming necessary conditions which are jointly sufficient using what they call the causal pie and is illustrated in the first figure below. One potential advantage of the causal pie approach is that it could allow one to indicate the relative importance of different factors by the size of the pie slices. In the first figure below each piece of the pie is the same size suggesting equal importance. The second figure from Cartwright et al. diagrams a policy intervention, focus on the first pie in the figure. The policy intervention is graphically indicated by a 1 in a large square. Support factors, i.e., S1.1 and S1.2 which are necessary for the intervention have a causal effect are in elipses and much smaller in size than policy intervention 1. This might suggest that these support factors are less important. Might there be situations where the support factors are more important than the policy interventions?

Figure 23: Causal pie: diagramming necessary and sufficient conditions



Source: Cartwright et al. 2024.

Figure 24: Causal pie: relative importance of pieces of the causal pie



Source: Cartwright et al. 2024.

161. There is quite a bit of variation and how scholars have thought about moderator, mediator variables and "context" in the discussion of external validity. And their article Geng et al. (2025) take the DAG approach to thinking about these issues. Based on the moderator section of the LNQA book (chapter 4) discuss how they have diagrammed context, see figure 25, how they have described it, and it might be conceived of as a moderator. See in particular the discussion associated with the figure.

Figure 25: Context, moderators and external validity

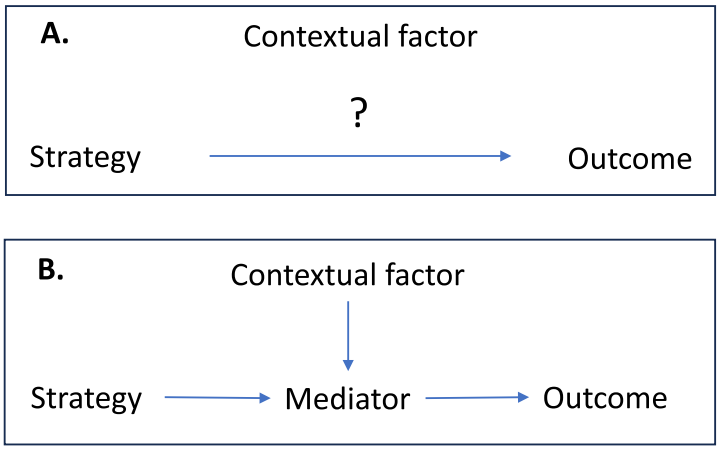


Fig. 3 Mediators enables consideration of context. Absence of a hypothesized mechanism limits analysis of context (Panel **A**). Proposed mechanisms — in this case a single mediator for simplicity — offers a conceptual receptacle for how context might influence a strategy's effects (Panel **B**). In the convention of directed acyclic diagrams, two arrows that point into one node implies that the effect of each arrow is influenced by the other, in this case representing how context effects the effect of the strategy on the outcome

Answer: The figure is fundamentally ambiguous because it is not clear what the arrow to the mediator means. For example it could mean that contextual factor is a cause of the mediator. So context is just another causal factor in the model. In the LNQA approach context does not influence the mediator itself but its causal relationship to the outcome, i.e., β. In general, moderator variables require significant interpretation when presented in DAGs. The causal of model cannot just be read from the DAG, it must be interpreted via the text discussing the DAG and eventually mathematical equations.

162. In most causal model figures, for example DAGs, no specific reference is made to causal claims in the negative: that X is not a cause of Y. However, in the DAG world these noncausal claims are absolutely critical in various contexts. A good example is instrumental variables. To have a valid instrumental variable implies important, and often somewhat implausible causal claims. How might one diagram those in a DAG-like fashion?

    Answer: Figure 26 is how Mellon (2025) nicely does this making very clear via the dashed arrows what causal relationships cannot hold for a valid instrumental analysis, W is the instrument.

Figure 26: Instrumental variables and noncausal relationships

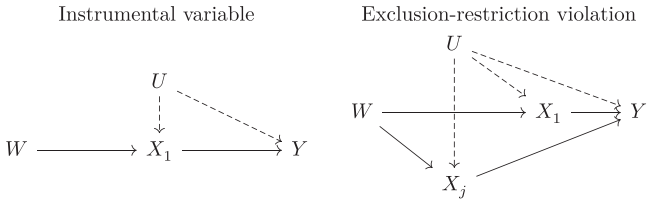


FIGURE 1 Directed Acylic Graphs (DAGs) illustrating instrumental variables (IVs) and endogeneity. *Note*: The figure illustrates the causal structure assumed by IV estimates and how exclusion-restriction violations undermine them.

163. In causal model figures one needs to indicate aggregation decisions. Often there is a choice between "+" and the Boolean AND. Copeland uses "+" the figure below. What in the text justifies the "+"? Are there other parts of the text which would support an AND? Also, he uses two kinds of causal arrows: what is the difference between the two?

Figure 27: Aggregation and arrows in causal model figures: Copeland on US foreign policy and commerce

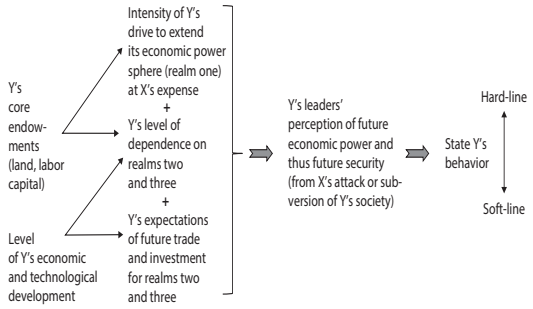


FIGURE 1.2. Three core causal factors

164. Frequently, in fact usually, in causal model figures and DAGs aggregation between factors is not explicit in the figure and must be inferred from the text. Discuss the plausible aggregation options for the various levels in Debre's theory of how regional organizations help authoritarian regimes.

Figure 28: Aggregation in causal models: Debre on regional organizations and authoritarian regimes

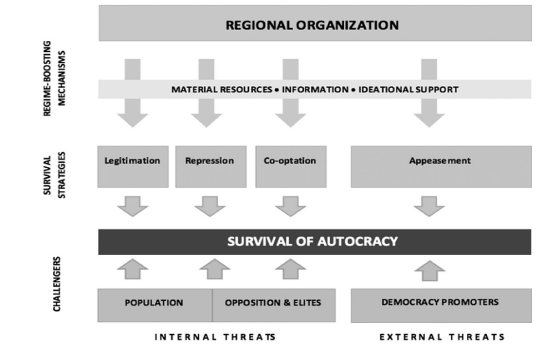
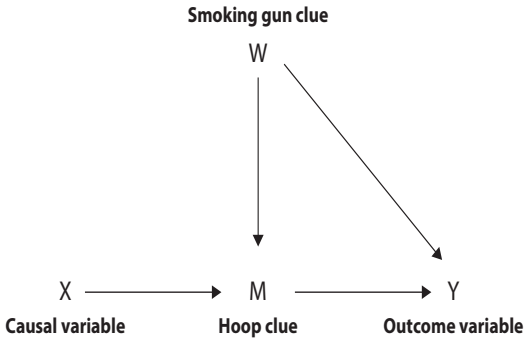


Figure 1. Process of regime-boosting regionalism.

165. Discuss the Blair et al. interpretation of smoking gun and hoop tests for individual cases within a DAG process tracing model. In particular, which of the factors in the figure are causal and which are evidence about causes? This is also about interpreting what → means in causal model figures.

Figure 29: Process tracing and within-case inference

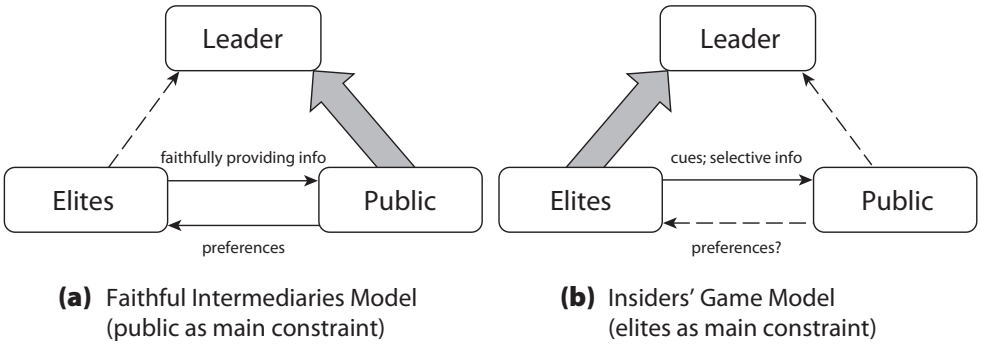


Source: Based on Blair et al. 2023.

Answer: Some arrows are about *evidence* for a causal relationship and other arrows are about the *causal* relationship itself. Probably one needs different kinds of arrows since these are two quite different aspects of causal inference.

166. Saunders (2024) proposes a complex theory of constraints on leader decision-making on war. Below is one of her causal model figures. Discuss the nature of the three different arrows she uses in the figure. Discuss how constraints in particular appear, but also other causal factors which are not constraints, for example motivational factors.

Figure 30: Constraints in causal figures: Saunders on constraints decisionmaking for war



(a) Faithful Intermediaries Model (public as main constraint)

(b) Insiders' Game Model (elites as main constraint)

Source: Saunders 2021.

167. Bulutgil (2022) provides a complex argument for explaining the success of "institutional secularization," see the figure below. Timing and temporal order are important parts of that mechanism. However, that does not appear in her figure very clearly. Could you modify it to make this aspect of her argument clearer?

> The second and critical factor is the relative timing of secularizing parties compared to grassroots religious parties. Briefly put, if pro-religion parties that seek popular support emerge before or at the same time as secularizing parties, they usually crowd out (or fend off) the latter by relying on the preexisting religious infrastructure to recruit regional leaders and grassroots support. Therefore, political parties with a secularizing platform face an uphill battle in these types of contexts. In contrast, if secularizing parties emerge significantly ahead of popular pro-religion parties, then they stand a much better chance of recruiting regional leaders and expanding their network beyond their initial geographical origin. (Bulutgil 2022, 12–13)

Figure 31: Success of secularizing political movements

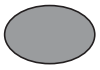
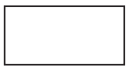
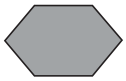
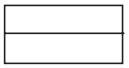
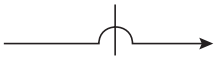
FIGURE 1.3 Success of Secularizing Political Movements: Institutional Secularization

Source: Bulutgil 2022.

168. In a very interesting new book and project James (2022) has developed a system for diagraming international relations theories. For example, in his book he has diagramed virtually all the major realist thinkers in international relations. Below are the elements that he uses to construct his diagrams. Discuss his system. For example, he has an extensive system of different kinds of "boxes," see the figure below. Contrast that with the system presented in the causal model-mechanism chapter of the LNQA book manuscript. See also his diagraming project for many more diagrams, https://visualinternationalrelationsproject.com/about-us/.

Figure 32: Notation system for causal mechanisms and theories

Table 6.3 Systemist Notation

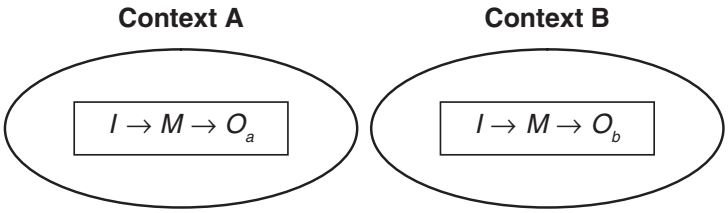
Initial variable		The starting point of a series of relationships
Generic variable		A step in the process being depicted
Divergent variable		Multiple pathways created from a single linkage
Convergent variable		A single pathway created from multiple linkages
Nodal variable		Multiple pathways created from multiple linkages
Co-constitutive variable		Two variables that are mutually contingent upon each other
Terminal variable		The end point of a series of relationships
Connection stated in study		A linkage explicitly made by the author
Connection crossing over		Two separate linkages that do not interact
Connection inferred from study		A linkage inferred by the reader but not made explicit by the author
Interaction effect		Two variables that depend upon the effect of each other

Source: James 2020.

169. In an oft-cited article Falleti and Lynch (2009) discuss the connection between what they call context and causal mechanisms. Their basic idea is illustrated in the figure below. Like antecedent conditions, contextual factors are also causal factors because they changed the nature of causal relationships. Can you redo their figure and introduce causal arrows from context that makes it clear what context is causally influencing?

Figure 33: Context and causal mechanisms

Figure 1
$I \rightarrow M \rightarrow O$ Model in Different Contexts



Note: I = inputs; M = mechanisms; O = outputs.

Source: Falleti and Lynch 2009.

170. One of the core guidelines for creating causal mechanism figures discussed in the LNQA book manuscript is the need to provide what it calls aggregation criteria. This means when there are multiple arrows leading to a given box one needs to discuss whether those should be connected with AND, OR, + or whatever. Discuss what might be the appropriate additions to Loxton's theory that also might make it match his empirical data, see the figure below.

Figure 34: Aggregating factors: conservative party-building

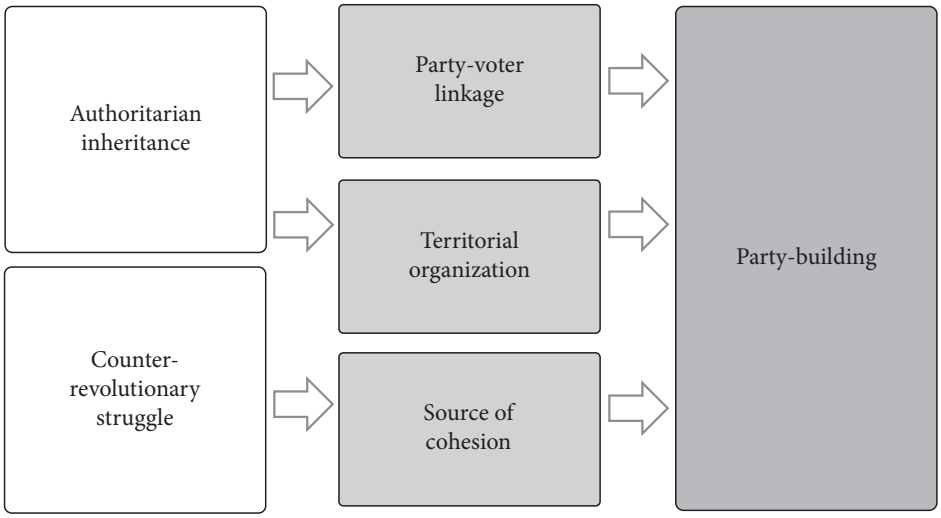


Figure 2.1 Model of the argument.

Source: Loxton 2021.

171. Scholars rarely claim to have identified necessary and sufficient conditions for an outcome. However, sometimes their data tables legitimately support such a conclusion. Discuss whether that is a reasonable interpretation for Cornell et al. to claim given the data in their table, see the figure below. Note also that their table 4.3, would lead to similar conclusions.

> Table 4.2. Democratic legacy and democratic survival in interwar Europe and neo-Europe. . . . Table 4.3 Democratic legacy and interwar democratic spells in Latin America. In logical terms, a pre-interwar democratic legacy was thus sufficient for democratic survival in Europe and the former British settler colonies and for longer democratic endurance in Latin America. These regularities are so manifest that further assessments of this relationship using standard (probabilistic) statistical techniques have little to add.

Figure 35: Necessary and sufficient conditions: democratic survival in the interwar period

Table 4.2 Democratic legacy and democratic survival in interwar Europe and neo-Europe

		Democratic survival in the interwar period	
		Yes	**No**
Pre-1918 democratic legacy	**Yes**	Australia, Belgium, Canada, Denmark, France, Ireland, Netherlands, New Zealand, Norway, Sweden Switzerland, UK, USA	
	No	Czechoslovakia, Finland	Austria, Bulgaria, Estonia, Germany, Greece, Italy, Latvia, Lithuania, Poland, Romania, Spain, Yugoslavia (Albania, Hungary, Portugal, USSR)

Note: The countries in parenthesis had no democratic spell during the interwar period.

Source: Cornell et al. 2020.

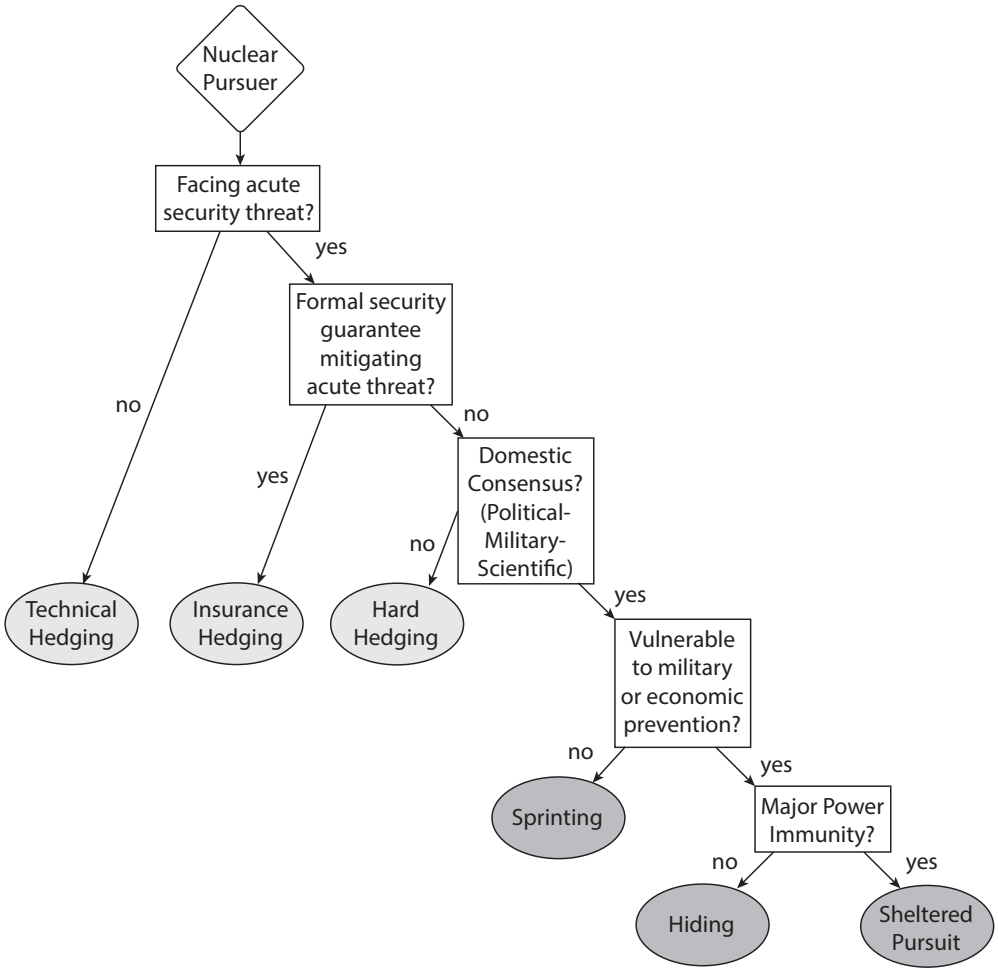
172. Matovski (2021) has a complex theory. Draw the causal mechanism figure. Are there necessary conditions? What about "amplifying" conditions? What about mediator variables? For example,

> This book's key point of departure is that the existing literature correctly identifies many key drivers of electoral authoritarianism but overlooks the most fundamental background factor that ties them all together: the genuine popular appeal of these regimes in troubled societies. In the analytic framework I propose, coercive power still plays an essential role, and pseudo-democratic institutions also help autocracies apply it with greater efficiency and precision. Electoral authoritarian regimes are still propped up by their economic performance, as well as by their resilience to democratic diffusion and external pressures. However, these factors operate in a broader context, defined by the distinct appeal of elected autocracies as guarantors of order and justice in troubled societies. When this appeal is salient and these factors are aligned with it, their effects are greatly amplified. When the strongman appeal is diminished, the use of coercion, high economic performance, and resistance to democratization pressures have weak or negligible impact, at best. At worst, they may be counterproductive and hasten the downfall of electoral authoritarianism. (Matovski 2021, 26–27)

173. As discussed in the causal model-mechanism chapter of the LNQA book manuscript, the → has multiple meanings. Often it means a causal relationship, but sometimes it just indicates temporal order. Discuss how one should interpret the arrows in Narang's main theory, see the figure below. Could the order of the boxes be changed if they were not temporal arrows? For example,

"facing acute security threats" above "nuclear pursuer"? Or could one write a Boolean model (which almost never has temporality in it)?

Figure 36: The ideal-typical revolutionary reactive sequence.



Source: Narang 2020.

174. Jenne (2015) provides another example of a complex theory. In particular notice the mixing of X and Y generalizations. Do a causal mechanism figure and/or write a Boolean equation for the theory.

> The argument developed in these pages is that conflict managers of *both* regimes were far more likely to achieve a successful outcome if powerful third parties acted to *nest the domestic disputes in a stable regional environment*. Put another way, cooperative (or soft-power) conflict management requires hard-power backing to achieve success. . . . I argue that effective civil conflict management requires building security from the outside in – that is, stabilizing the external environment before (or at the same time as) reconfiguring domestic institutions. If sectarian conflicts are securely "nested" in a stable regional and hegemonic environment, they are likely

> to become far more tractable; cooperative conflict management in particular is far likelier to meet with success. . . . *Failing* to securely nest sectarian struggles in a stable external environment may be setting peacemakers and peacekeepers up for an impossible task. . . . This book demonstrates that "nested security" is a key background condition for successful conflict reduction, regardless of the resources of the third party, the intervention strategy, and the attitudes of the government and minority representatives. . . . I demonstrate that nested security is effectively a precondition of success for cooperative conflict management. . . . To establish effective regional security regimes, three features must be present. First, a decision-making body must be established that can cope with internal veto players as well as member-state aggression. Second, a regularized, independent enforcement mechanism must be created with the full backing of regional hegemons and/or great powers. Finally, the regime must engender cooperation between monitors and enforcers to provide effective conflict management at all stages of the conflict – from preventive diplomacy to postwar reconstruction. (Jenne 2015, 9–10, 18, 45, 194)

175. When the author has a complex theory is is useful express it either in Boolean equations or draw causal mechanism figures. Lee has a complicated theory of foreign subversion that is briefly outlined in the quote below. Draw the causal mechanism figure and/or develop a Boolean equation which represents her claims. For example, it seems that there are necessary conditions in her model, e.g., "Only in these situations." She has constraints in her theory e.g., "proxy availability is a key constraint". The loss of strength gradient plays a key role, how would one put that into the figure? Are some factors "facilitating conditions" e.g., the fourth one?

> These observations about costs and constraints have five implications for predicting when subversion will occur. First, as I have already argued, states will be more likely to employ subversion against state authority when confronting adversaries with which they have severe policy disputes. Only in these situations will states be willing to bear the costs and risks of dealing with proxy groups. I call this condition "motive," since states should use subversion when their foreign relations provide for sufficient motivation. Second, subversion will occur only when proxies are available in the target state. Since subversion by definition involves delegation, willing agents, or "means" on the ground, is a necessary condition for subversion. Additionally, treating agent availability as an independent condition implies that proxies can be located by external states but not easily created by them. Without proxies, states cannot engage in subversion. Thus, proxy availability is a key constraint on the use of subversion. Third, as the target's own level of consolidation increases, the more expensive it will be to weaken state administration and control, since these targets are less vulnerable from the outset. Subversion is less likely to succeed against fully consolidated states, which are more informed of local goings-on and will be more likely to defeat efforts to undermine their state authority. Thus, even though subversion of consolidated targets may be attractive, the costs of imposing authority losses may be quite constraining for some states. Knowing this, rational states are unlikely to invest in an expensive strategy that offers little chance of success. We should thus observe few instances of subversion being used against fully consolidated states, which reside primarily in the developed world. Fourth, because material

> resources, training, and protection and safe havens are the most important means through which sponsors empower proxies, sponsors should be more willing to use subversion when they can easily deliver these investments to their proxies. A key impediment to doing so is distance, because distance increases the delivery cost. Similarly, safe havens and training camps are less valuable to proxies if those areas are difficult to reach. Just as conventional force is subject to a "loss-of-strength gradient," in which military strength declines with distance, subversion is subject to an analogous loss-of-interference gradient, in which the ability to deliver investments declines with distance. As with the loss of strength gradient, distance should act as a constraint for all but the most capable states in the international system. This suggests that subversion will be more likely to be observed between adversaries that are contiguous and in areas bordering those adversaries. This key implication informs the statistical analyses in chapter 3. Fifth, states should be more willing to turn to subversion as the costs of other policy instruments increase. The degree to which states will be willing to bear those costs depends in part on the costs and constraints of alternatives. Scholars of international norms have suggested that laws, rules, and norms against the use of force for some ends have made force costly, particularly when would-be violators expect rule enforcement from powerful third-party states. Similarly, defensive military alliances may make conventional force undesirable against some target states. If states anticipate high costs for engaging in conventional force, the primary noncooperative alternative to subversion, then those states will be more likely to favor subversion. The first two implications, about motive and means, are the most important factors that influence the use of subversion, and as I describe later, the primary tests of these propositions appear in chapters 3 and 4. The third and fourth implications, about types of targets and the constraining role of distance, also influence the research design of chapter 3. I provide evidence for the fifth implication, about alternative foreign policy strategies, in chapters 5 and 6. (Lee 2020, 56–57)

Answer: One possible Boolean interpretation of her theory is: R (rivalry) AND RG (at least one rebel group in target) AND not-ED (economically developed) AND [facilitating conditions] (D (distance) + HC (high cost of alternative actions)) is sufficient for subversion. The first two are clearly necessary.

176. Slater and Wong (2022) provide a complex theory–causal mechanism of democratization and economic development in "developmental Asia." Draw a figure that includes all the critical components and how they combine to produce the outcome of democratization.

177. Drawing figures is very useful for contrasting alternative explanations of individual events. In an interesting debate Kocher and Monteiro contest Ferwerda and Miller's natural experiment account of the impact of the line the Germans drew to separate occupied France from so-called free France. Draw the figures that illustrate each party's account of the case. Discuss the various forms of evidence each side uses to support their account. For example,

> Strategic Logic, Political Rule, and Resistance in World War II France In this section, we reconstruct the historical process through which violent

> resistance was produced in the territory that Ferwerda and Miller analyze: the departments of Charente, Cher, Saône-et-Loire, and Vienne. We lay out an alternative causal account that explains variation in the subnational pattern of violent resistance near the LoD [line of demarcation] as a function of the interaction of German and Allied war strategies. We substantiate each element of our argument using qualitative historical evidence. (Kocher and Monteiro 2016, 959)

178. Drawing figures is very useful for contrasting alternative explanations of individual events. In an interesting debate Kocher and Monteiro contest Ferwerda and Miller's natural experiment account of the impact of the line the Germans drew to separate occupied France from so-called free France. Draw the figures that illustrate each party's account of the case. Discuss the various forms of evidence each side uses to support their account. For example,

> Strategic Logic, Political Rule, and Resistance in World War II France In this section, we reconstruct the historical process through which violent resistance was produced in the territory that Ferwerda and Miller analyze: the departments of Charente, Cher, Saône-et-Loire, and Vienne. We lay out an alternative causal account that explains variation in the subnational pattern of violent resistance near the LoD [line of demarcation] as a function of the interaction of German and Allied war strategies. We substantiate each element of our argument using qualitative historical evidence. (Kocher and Monteiro 2016, 959)

179. Bulutgil (2022) provides a complex argument for explaining the success of "institutional secularization," see the figure below. Timing and temporal order are important parts of that mechanism. However, that does not appear in her figure very clearly. Could you modify it to make this aspect of her argument clearer?

> The second and critical factor is the relative timing of secularizing parties compared to grassroots religious parties. Briefly put, if pro-religion parties that seek popular support emerge before or at the same time as secularizing parties, they usually crowd out (or fend off) the latter by relying on the preexisting religious infrastructure to recruit regional leaders and grassroots support. Therefore, political parties with a secularizing platform face an uphill battle in these types of contexts. In contrast, if secularizing parties emerge significantly ahead of popular pro-religion parties, then they stand a much better chance of recruiting regional leaders and expanding their network beyond their initial geographical origin. (Bulutgil 2022, 12–13)

Figure 37: Success of secularizing political movements

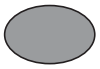
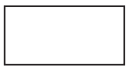
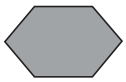
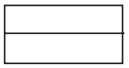
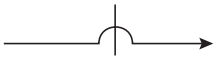
FIGURE 1.3 Success of Secularizing Political Movements: Institutional Secularization

Source: Bulutgil 2022.

180. In a very interesting new book and project James (2022) has developed a system for diagraming international relations theories. For example, in his book he has diagramed virtually all the major realist thinkers in international relations. Below are the elements that he uses to construct his diagrams. Discuss his system. For example, he has an extensive system of different kinds of "boxes," see the figure below. Contrast that with the system presented in the causal mechanism chapter of the LNQA book manuscript. See also his diagraming project for many more diagrams, https://visualinternationalrelationsproject.com/about-us/.

Figure 38: Notation system for causal mechanisms and theories

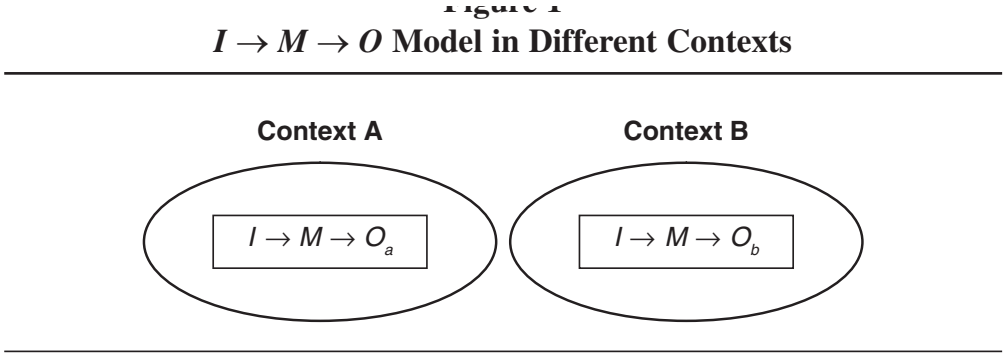
Table 6.3 Systemist Notation

Initial variable		The starting point of a series of relationships
Generic variable		A step in the process being depicted
Divergent variable		Multiple pathways created from a single linkage
Convergent variable		A single pathway created from multiple linkages
Nodal variable		Multiple pathways created from multiple linkages
Co-constitutive variable		Two variables that are mutually contingent upon each other
Terminal variable		The end point of a series of relationships
Connection stated in study		A linkage explicitly made by the author
Connection crossing over		Two separate linkages that do not interact
Connection inferred from study		A linkage inferred by the reader but not made explicit by the author
Interaction effect		Two variables that depend upon the effect of each other

Source: James 2020.

181. In an oft-cited article Colletti and Lynch discuss the connection between what they call context and causal mechanisms. Their basic idea is illustrated in the figure below. Like antecedent conditions, contextual factors are also causal factors because they changed the nature of causal relationships. Can you redo their figure and introduce causal arrows from context that makes it clear what context is causally influencing?

Figure 39: Context and causal mechanisms



Note: I = inputs; M = mechanisms; O = outputs.

Source: Falleti and Lynch 2009.

182. One of the core guidelines for creating causal mechanism figures discussed in the LNQA book manuscript is the need to provide what it calls aggregation criteria. This means when there are multiple arrows leading to a given box one needs to discuss whether those should be connected with AND, OR, + or whatever. Discuss what might be the appropriate additions to Loxton's theory that also might make it match his empirical data, see the figure below.

Figure 40: Aggregating factors: conservative party-building

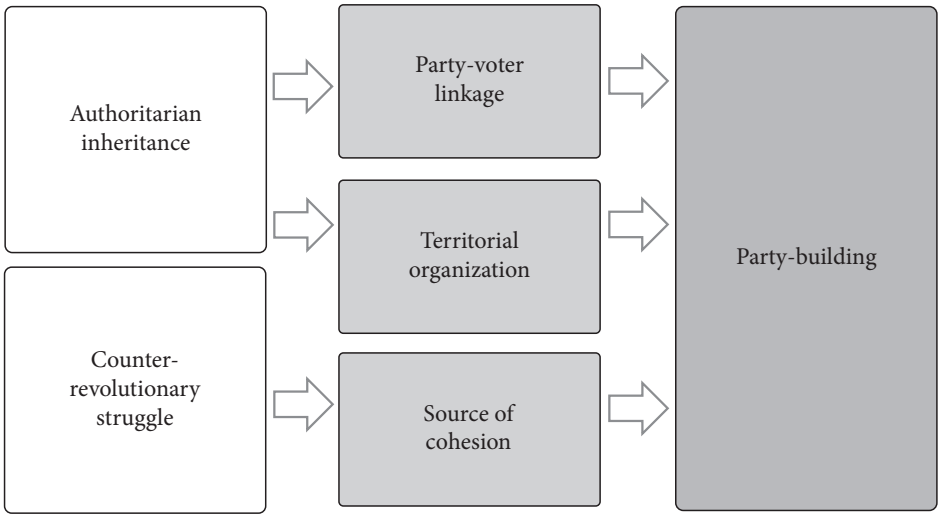


Figure 2.1 Model of the argument.

Source: Loxton 2021.

183. Scholars frequently talk about "mediating factors." This has a pretty clear meaning in statistical work but has much larger relevance, e.g., in causal chains. Draw a figure of Bertrand's theory of violent nationalist mobilizations and how mediating factors might appear.

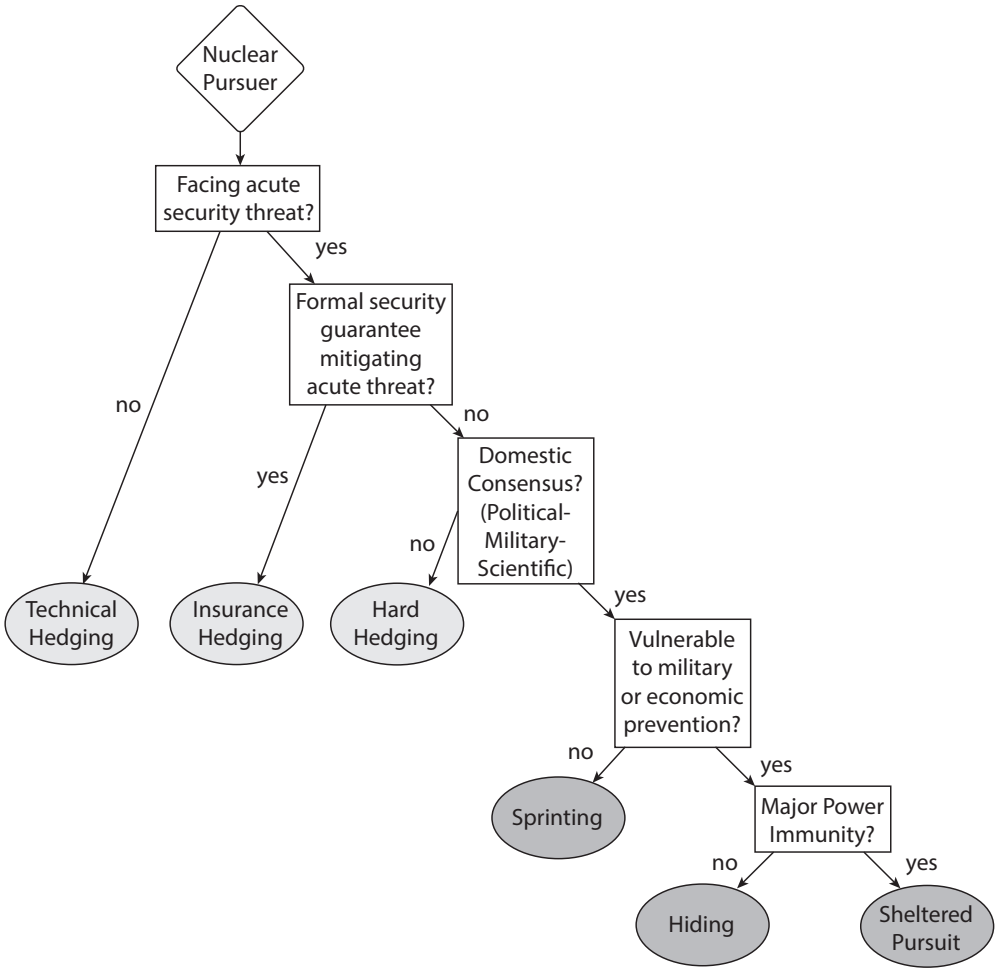
> Figure 2.1 illustrates the mix of mediating factors. Individual effects of grievances are usually difficult to clearly distinguish. Yet, without grievances, there are few reasons why any groups would develop a nationalist agenda.56 I take them as given, and focus primarily on mediating factors. As illustrated, the three factors of mobilizational capacity, state repression and external support contribute to the degree to which a conflict will become violent. By this perspective, it is incorrect, I contend, to equate high levels of violence with deeper grievances. In many cases, the degree of violence is less a predictor of grievances than these structural features that allow a nationalist group to use violent means, and the opportunity structure of doing so, given a state's repressive capacity. Another mediating factor, I suggest, is also important: the availability of alternative institutional channels to express discontent. Violent mobilization is a costly option for nationalist groups. I hypothesize that the more channels are available to express their discontent and negotiate peaceful outcomes to conflict. (Bertrand 2021, 37)

184. Matovski (2021) has a complex theory. Draw the causal mechanism figure. Are there necessary conditions? What about "amplifying" conditions? What about mediator variables? For example,

> This book's key point of departure is that the existing literature correctly identifies many key drivers of electoral authoritarianism but overlooks the most fundamental background factor that ties them all together: the genuine popular appeal of these regimes in troubled societies. In the analytic framework I propose, coercive power still plays an essential role, and pseudo-democratic institutions also help autocracies apply it with greater efficiency and precision. Electoral authoritarian regimes are still propped up by their economic performance, as well as by their resilience to democratic diffusion and external pressures. However, these factors operate in a broader context, defined by the distinct appeal of elected autocracies as guarantors of order and justice in troubled societies. When this appeal is salient and these factors are aligned with it, their effects are greatly amplified. When the strongman appeal is diminished, the use of coercion, high economic performance, and resistance to democratization pressures have weak or negligible impact, at best. At worst, they may be counterproductive and hasten the downfall of electoral authoritarianism. (Matovski 2021, 26–27)

185. As discussed in the causal mechanism chapter of the LNQA book manuscript, the $\rightarrow$ has multiple meanings. Often it means a causal relationship, but sometimes it just indicates temporal order. Discuss how one should interpret the arrows in Narang's main theory, see the figure below. Could the order of the boxes be changed if they were not temporal arrows? For example, "facing acute security threats" above "nuclear pursuer"? Or could one write a Boolean model (which almost never has temporality in it)?

Figure 41: The ideal-typical revolutionary reactive sequence.



Source: Narang 2020.

186. Jenne (2015) provides another example of a complex theory. In particular notice the mixing of X and Y generalizations. Do a causal mechanism figure and/or write a Boolean equation for the theory.

> The argument developed in these pages is that conflict managers of *both* regimes were far more likely to achieve a successful outcome if powerful third parties acted to *nest the domestic disputes in a stable regional environment*. Put another way, cooperative (or soft-power) conflict management requires hard-power backing to achieve success. . . . I argue that effective civil conflict management requires building security from the outside in –that is, stabilizing the external environment before (or at the same time as) reconfiguring domestic institutions. If sectarian conflicts are securely "nested" in a stable regional and hegemonic environment, they are likely to become far more tractable; cooperative conflict management in particular is far likelier to meet with success. . . . *Failing* to securely nest sectarian struggles in a stable external environment may be setting peacemakers

> and peacekeepers up for an impossible task. . . . This book demonstrates that "nested security" is a key background condition for successful conflict reduction, regardless of the resources of the third party, the intervention strategy, and the attitudes of the government and minority representatives. . . . I demonstrate that nested security is effectively a precondition of success for cooperative conflict management. . . . To establish effective regional security regimes, three features must be present. First, a decision-making body must be established that can cope with internal veto players as well as member-state aggression. Second, a regularized, independent enforcement mechanism must be created with the full backing of regional hegemons and/or great powers. Finally, the regime must engender cooperation between monitors and enforcers to provide effective conflict management at all stages of the conflict – from preventive diplomacy to postwar reconstruction.(Jenne 2015, 9–10, 18, 45, 194)

187. When the author has a complex theory is is useful express it either in Boolean equations or draw causal mechanism figures. Lee has a complicated theory of foreign subversion that is briefly outlined in the quote below. Draw the causal mechanism figure and/or develop a Boolean equation which represents her claims. For example, it seems that there are necessary conditions in her model, e.g., "Only in these situations". She has constraints in her theory e.g., "proxy availability is a key constraint". The loss of strength gradient plays a key role, how would one put that into the figure? Are some factors "facilitating conditions" e.g., the fourth one?

> These observations about costs and constraints have five implications for predicting when subversion will occur. First, as I have already argued, states will be more likely to employ subversion against state authority when confronting adversaries with which they have severe policy disputes. Only in these situations will states be willing to bear the costs and risks of dealing with proxy groups. I call this condition "motive," since states should use subversion when their foreign relations provide for sufficient motivation. Second, subversion will occur only when proxies are available in the target state. Since subversion by definition involves delegation, willing agents, or "means" on the ground, is a necessary condition for subversion. Additionally, treating agent availability as an independent condition implies that proxies can be located by external states but not easily created by them. Without proxies, states cannot engage in subversion. Thus, proxy availability is a key constraint on the use of subversion. Third, as the target's own level of consolidation increases, the more expensive it will be to weaken state administration and control, since these targets are less vulnerable from the outset. Subversion is less likely to succeed against fully consolidated states, which are more informed of local goings-on and will be more likely to defeat efforts to undermine their state authority. Thus, even though subversion of consolidated targets may be attractive, the costs of imposing authority losses may be quite constraining for some states. Knowing this, rational states are unlikely to invest in an expensive strategy that offers little chance of success. We should thus observe few instances of subversion being used against fully consolidated states, which reside primarily in the developed world. Fourth, because material resources, training, and protection and safe havens are the most important means through which sponsors empower proxies, sponsors should

> be more willing to use subversion when they can easily deliver these investments to their proxies. A key impediment to doing so is distance, because distance increases the delivery cost. Similarly, safe havens and training camps are less valuable to proxies if those areas are difficult to reach. Just as conventional force is subject to a "loss-of-strength gradient," in which military strength declines with distance, subversion is subject to an analogous loss-of-interference gradient, in which the ability to deliver investments declines with distance. As with the loss of strength gradient, distance should act as a constraint for all but the most capable states in the international system. This suggests that subversion will be more likely to be observed between adversaries that are contiguous and in areas bordering those adversaries. This key implication informs the statistical analyses in chapter 3. Fifth, states should be more willing to turn to subversion as the costs of other policy instruments increase. The degree to which states will be willing to bear those costs depends in part on the costs and constraints of alternatives. Scholars of international norms have suggested that laws, rules, and norms against the use of force for some ends have made force costly, particularly when would-be violators expect rule enforcement from powerful third-party states. Similarly, defensive military alliances may make conventional force undesirable against some target states. If states anticipate high costs for engaging in conventional force, the primary noncooperative alternative to subversion, then those states will be more likely to favor subversion. The first two implications, about motive and means, are the most important factors that influence the use of subversion, and as I describe later, the primary tests of these propositions appear in chapters 3 and 4. The third and fourth implications, about types of targets and the constraining role of distance, also influence the research design of chapter 3. I provide evidence for the fifth implication, about alternative foreign policy strategies, in chapters 5 and 6. (Lee 2020, 56–57)

Answer: One possible Boolean interpretation of her theory is: R (rivalry) AND RG (at least one rebel group in target) AND not-ED (economically developed) AND [facilitating conditions] (D (distance) + HC (high cost of alternative actions)) is sufficient for subversion. The first two are clearly necessary (see the Russia chapter).

188. Slater and Wong (2022) provide a complex theory–causal mechanism of democratization and economic development in "developmental Asia." Draw a figure that includes all the critical components and how they combine to produce the outcome of democratization.

189. Below is one core figure for Levitsky and Way (2022) regarding social revolutions and the durability of authoritarian regimes. First, could or should one interpret this via a Boolean equation and put ANDs, ORs, and S (sufficient for) or N (necessary condition) at appropriate places in the figures? Could you write all the relevant Boolean equations?

Figure 42: Ideal-typical path to durable authoritarianism

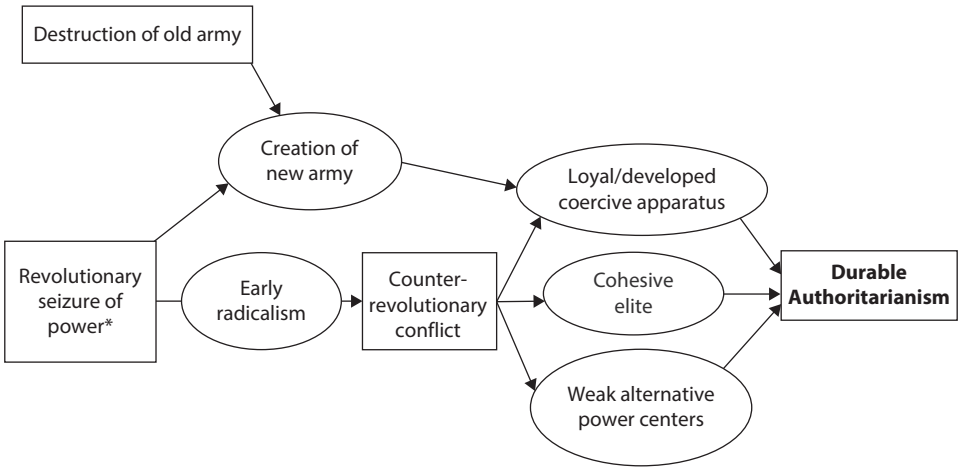


FIGURE 1.2: The Ideal-Typical Revolutionary Reactive Sequence.
* In some cases (e.g., China), the reactive sequence occurs before seizure of power.

Source: Levitsky and Way (2022).

190. Core to LNQA is the crafting of causal mechanism figures. Use the guidelines of the LNQA book (chapter 5 on causal mechanisms) to redo the Levitsky and Way, see the figure below. For example, how to interpret of some of the ambiguities, e.g., one Y variable for the factors on the right in their figure. Another question is whether it is one Y with different values or different Ys.

Figure 43: Revolutionary regime paths

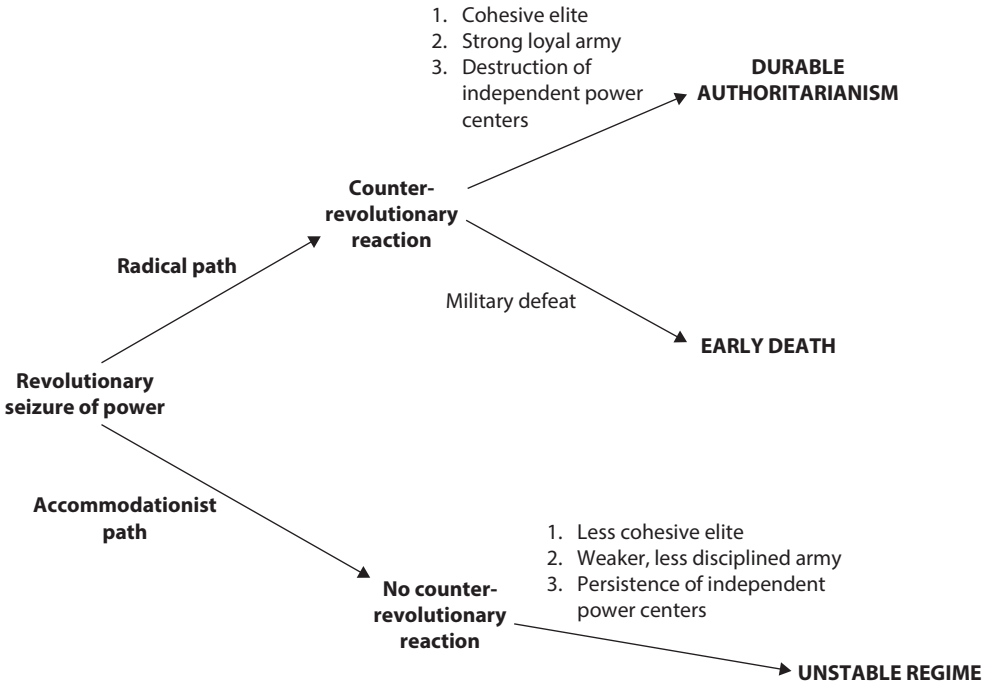


FIGURE 1.3: Three Revolutionary Regime Paths.

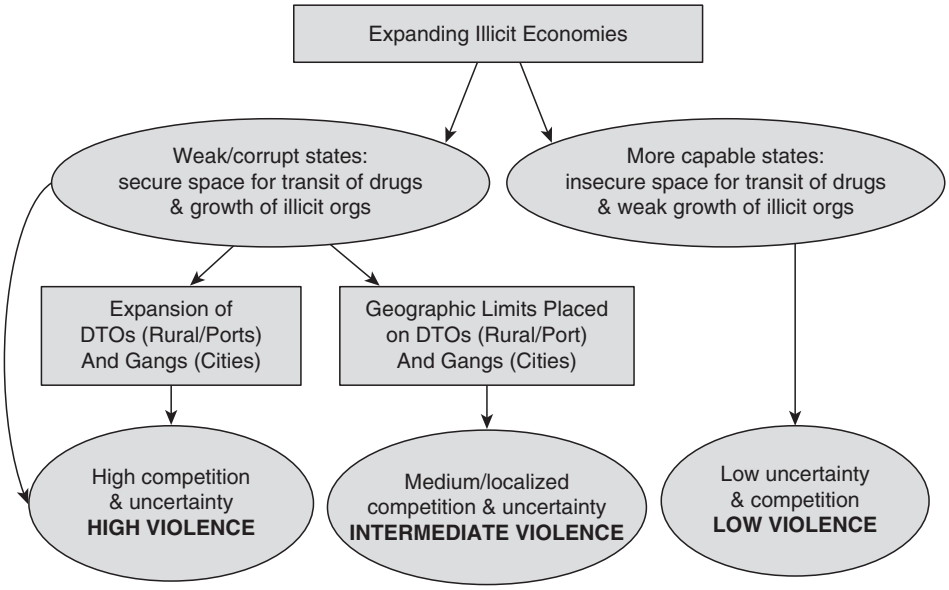
Source: Levitsky and Way (2022).

191. The Sherlock Holmes story "Silver Blaze" has become a stable example in the qualitative methods literature, e.g., Collier (2011), Fairfield and Charman (2020). Draw the causal mechanism figure corresponding to the story, key elements, causal relationships, and "clues."

192. Inverted tree figures are not uncommon in qualitative work. Here is a nice example from Yashar (2018). It also illustrates some issues in drawing causal mechanism figures using just $\rightarrow$. How might some arrows have a different Boolean interpretation than others. Could one write the Boolean equations for this figure?

> In particular, this book emphasizes three factors: the transnational illicit economy, state capacity, and territorial competition among organizations.... First, I contend that the development of a transnational illicit economy and illicit criminal organizations set the stage for the high levels of violence that we now see in Latin America.... Second, I argue that illicit trade and transit is likely to take hold where illicit actors find weak and/or complicit state institutions (particularly law-and-order institutions such as the police and courts).... Third, and finally, I find that the highest levels of violence are emerging particularly where illicit organizations encounter organizational competition (either from other illicit organizations or the state or both) to control previously hegemonic territorial enclaves.... No

single factor determines the outcome. The combination of factors, however, can be deadly. (Yashar 2018, 18–19)



Answer:

The arrows to the bottom level of the tree in a Boolean interpretation mean "is sufficient for." One plausible interpretation of the arrows above are as invoking the logical AND. Here is one Boolean equation: IF (illicit economic AND weak state AND DTO expansion) THEN high violence.

193. A common an issue in causal mechanism figures is aggregation or structure between boxes and circles. Second problem is when there are multiple factors within a given box. Take Tripps prominent theory of conflict and gender policy changes to discuss how one might insert structure both between boxes is with as well as within a box.

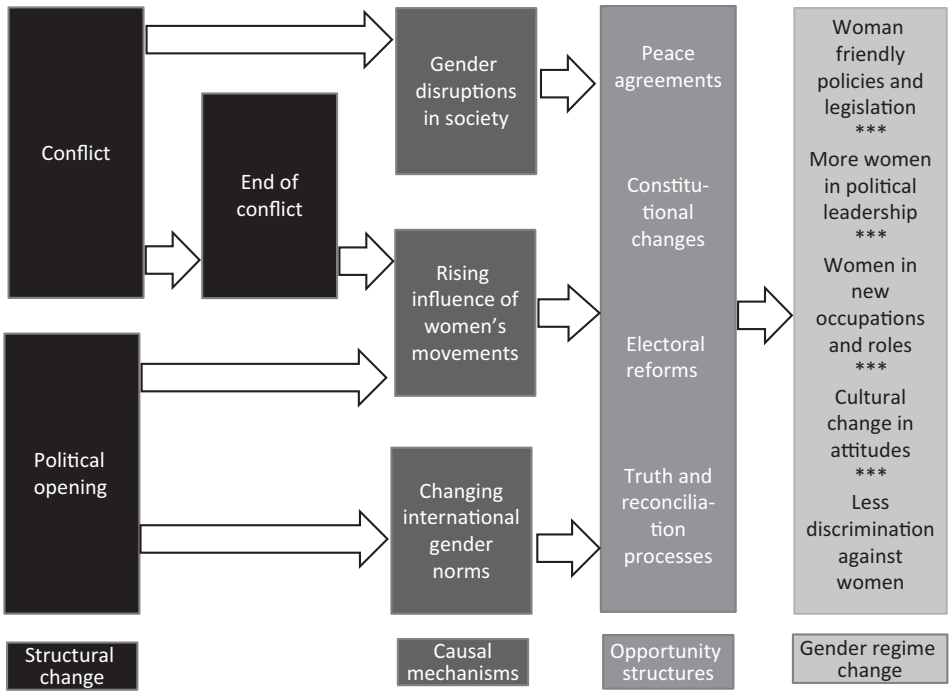
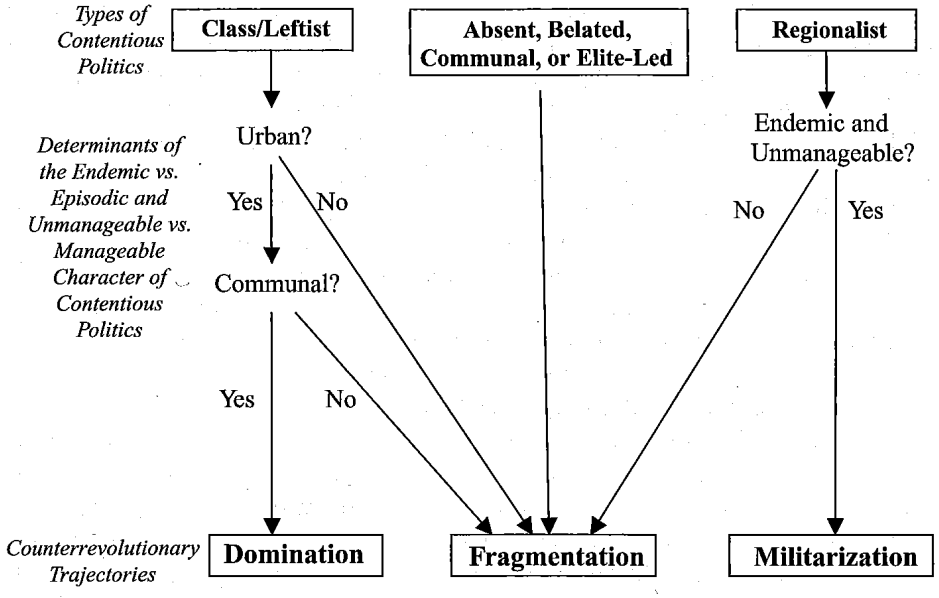


FIGURE 2.1 Model of Process of Postconflict Impacts on Gender Policy

194. Can Slater's (2010) causal mechanisms be expressed in Boolean equations? If so what would they look like?

FIGURE 4. Contentious Politics and Counterrevolutionary Trajectories



195. Debs and Monteiro (2017) in their theory of proliferation of nuclear weapons frequently talk about necessary conditions for proliferation which is quite common in this literature in general. Discuss how one might introduce these

considerations into their causal mechanism figure, indicating which factors are necessary, and eventually the various paths to sufficiency.

> Empirically, proliferation occurs in a limited range of strategic environments. Specifically, we find two sets of strategic circumstances – or pathways – to nuclear acquisition. First, a high level of security threat combined with high relative conventional power on the part of the proliferating state. Second, a high level of security threat combined with the presence of an ally that is deemed unreliable. All other strategic settings result in the maintenance of a state's nonnuclear status. . . . In doing so, the strategic logic of nuclear proliferation sheds light on several hitherto underappreciated historical patterns. First, states that do not face a high-level security threat have not acquired the bomb. The presence of a significant security threat is a necessary condition for nuclearization. Historically, no state has acquired nuclear weapons without perceiving its security environment as highly threatening, regardless of how strong other pressures to acquire the bomb – including considerations of domestic or international prestige, the psychology of leaders, or the economic preferences of ruling elites – may be. Second, among states that are not protected by a great power sponsor, only those that are strong vis-à-vis their adversaries have acquired the bomb. There is no historical case of a relatively weak state ever succeeding in nuclearizing without having a powerful ally committed to retaliating against a preventive counter-proliferation strike. . . . Third, among states that possess a powerful ally, only those whose security goals are not entirely covered by this sponsor have acquired nuclear weapons. Put differently, states whose security goals are subsumed by their powerful allies' own aims do not possess the willingness to acquire the bomb. . . . Fourth, threats of abandonment issued by a security sponsor – what we call a "sticks-based" nonproliferation policy – are effective in curtailing proliferation only by protegees that are relatively weak vis-à-vis their adversaries. If a protegee is strong vis-à-vis its adversaries, it has the opportunity to proliferate on its own, even if its security sponsor were to abandon it. In this case, the sponsor can only effectively deter proliferation by taking away the protegee's willingness to acquire nuclear weapons, which it can do by extending additional security assurances – what we call a "carrots-based" nonproliferation policy. In other words, whereas sticks can deter proliferation by weak protegee, only carrots will prevent stronger protegees from building nuclear weapons. (Debs and Monteiro 2017, 11–13)

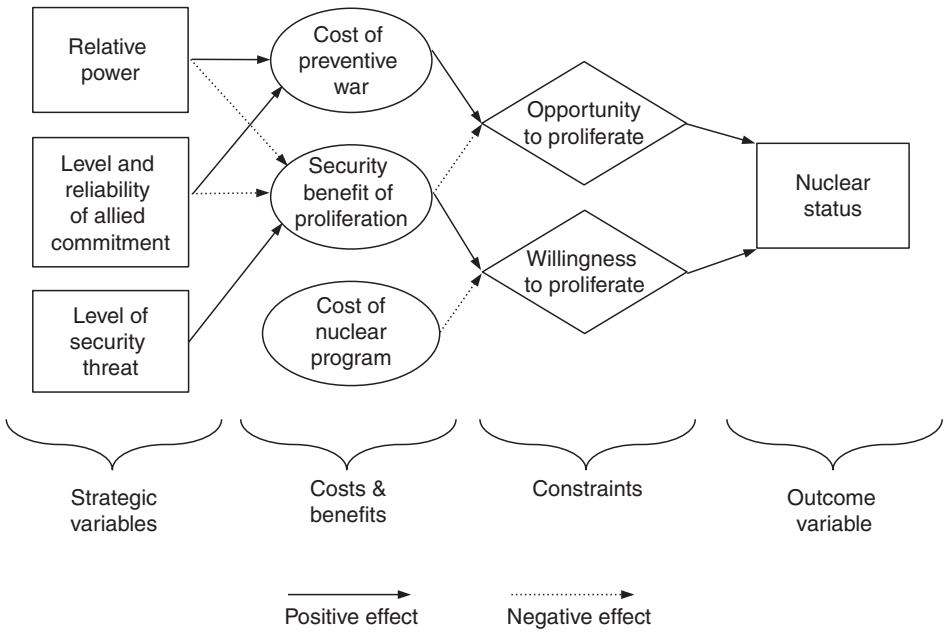


Figure 2.2. The Strategic Logic of Nuclear Proliferation

196. Write the Boolean equation(s) for the causal mechanism figure below. See also his table 1 on page 14, which is a nonstandard 2×2 table.

> The main arguments can be summarized briefly. If a leading state has little reason to fear the takeover of a peripheral region by a local actor or an outside power, then it should not have a strong preference when it comes to the type of order that exists there. In principle, both parity and primacy can prevent disruptive conflicts, the avoidance of which will be its chief objective in these circumstances. Therefore, a leading state should accommodate RRPs [rising regional powers] that can achieve either one because they will enhance local stability over the long run. Conversely, it should oppose RRPs that fall short of this threshold because their rise will contribute to persistent unrest. If a leading state is worried mainly that a local actor might dominate a peripheral region, however, then it will prefer parity to ensure that its own access to the area is not jeopardized. In this case, it should accommodate RRPs that are attempting to weaken local hegemons and oppose RRPs that are trying to gain control over their neighborhoods. Finally, if a leading state is more concerned about an outside power conquering a peripheral region, then it will prefer primacy instead because the strongest local actors are the best barriers to intervention. Thus, it should accommodate RRPs that fully overtake their rivals and oppose RRPs that fail to do so. (Montgomery 2016, 10)

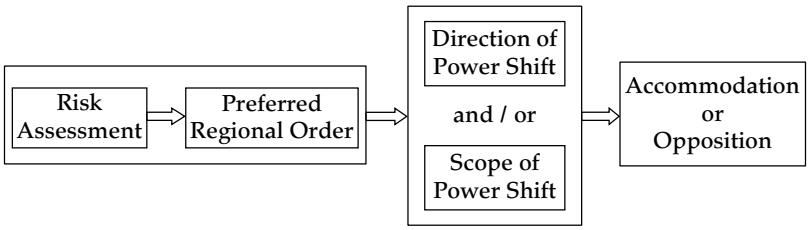


Figure 1. The causal argument

197. The multimethod book espoused the mantra "no causation without causal mechanism." This obviously plays a key role in exceptional studies where there may be randomization treatments but where the result does not match any causal mechanism intuition. The question then is should the journal publish the research when there seems to be no plausible mechanism that would explain the results. Here's a famous example in psychology. The *Journal of Conflict Resolution* had a similar controversial publication regarding the publication of an article "International Peace Project in the Middle East: The Effects of the Maharishi Technology of the Unified Field." (Orme-Johnson et al. 1988). Russett (2017) describes in some detail the controversies and eventual publication of the article.

> History may look back on 2011 as the year that changed psychology forever. It all began when the Journal of Personality and Social Psychology published an article called "Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect." The paper, written by Daryl Bem of Cornell University, reported a series of experiments on psi or "precognition," a supernatural phenomenon that supposedly enables people to see events in the future. Bem, himself a reputable psychologist, took an innovative approach to studying psi. Instead of using discredited parapsychological methods such as card tasks or dice tests, he selected a series of gold- standard psychological techniques and modified them in clever ways. . . . On this basis, Bem argued that the participants were able to benefit in the past from practice they had completed in the future. As you might expect, Bems results generated a flood of confusion and controversy. How could an event in the future possibly influence someones reaction time or memory in the past? If precognition truly did exist, in even a tiny minority of the population, how is it that casinos or stock markets turn profits? And how could such a bizarre conclusion find a home in a reputable scientific journal? (Chambers 2017, 2–3)

198. Draw a causal mechanism figure for Stasavage's influential book *Public debt and the birth of the democratic state France and Great Britain, 16881789*.

> Representative political institutions may improve a government's ability to make credible commitments through several different mechanisms. This chapter has used a formal model of legislative bargaining to provide support for my three main arguments. I first demonstrated that if capital owners are in the minority, then party formation can lead to credible commitment, but only if players bargain over multiple issues. In addition, one can expect the perceived credibility of taxation or borrowing to vary

> according to the partisan composition of government. Both of these observable implications will be considered in detail in subsequent chapters. I next showed that constitutional checks and balances will have little effect on credibility unless there is some mechanism ensuring that capital owners control a veto point. This helps support the argument that multiple veto points may in many cases be insufficient to ensure credible commitment. Finally, I developed my argument about bureaucratic delegation, suggesting that it will improve credibility only if capital owners have the political authority to block any attempt to override bureaucratic decisions. This too is an empirical prediction that is considered in subsequent chapters. (Stasavage 2003, 49–50)

199. The idea of a causal mechanism being a causal chain is very wide spread. Draw a causal mechanism figure for ethnic civil war according to Lindemann and Wimmer. For example, is each link necessary for the following one. How would you diagram the idea that different conditions are relevant in different stages?

> Overall, this suggests that the conditions are not ordered in a historical sequence of causal connections. The historically prior condition for ethnic war does not causally generate the second condition for ethnic war. In other words, the logic here is not that of a necessary chain reaction. Rather, the uniform temporal ordering in which conditions become relevant is due to the logic of the escalation process itself, which makes certain conditions relevant at different points in the process, but does not ensure their emergence. The following stages of this escalation process could be hypothesized: mobilization of ethno-nationalists, their initial organization, subsequent radicalization and mass recruitment, and, finally, the emergence of armed organizations.
>
> Different conditions become relevant during these subsequent stages. Further 'grievance'-inducing conditions are relevant for mobilization. Limited state reach might be relevant for initial organization. Indiscriminate violence becomes relevant for subsequent radicalization and mass recruitment – without indiscriminate violence the process of escalation stops here. And finally, an external sanctuary is often relevant for the actual organization of an armed guerrilla front. (Lindemann and Wimmer 2018, 317)

200. Social constructivists frequently uses the term "permissive cause." Carpenter (2007) discusses agenda-setting in terms of constraints and her causal mechanism figure, figure 1. p. 102. So the absence of permissive condition explains why children of rape is not on the agenda? What do the lines mean (i.e., not arrows). Do they just indicate temporal ordering?

201. The Keck and Sinkink (1998) boomerang model of social movements have been very influential. Think about the figure used to describe the causal mechanism, not an exhaustive list: (1) arrows versus lines what is the difference? (2) some arrows get "pressure" others do not, why? (3) is thickness of arrow an hypothesis about relative causal impact? (4) arrows between NGOs in State B, but not State A, why? (5) what do dotted lines (NGOs to blockage) mean? (6) are there arrows missing, e.g., from NGOs in State A to IGO?

202. Pevehouse is a main example in the discussion of causal mechanisms and multimethod research. One issue is that democratic IGO is relatively constant. For that mechanism to kick in it probably needs some triggering mechanism, e.g., crisis, coup attempt, shock. Add that to the Pevehouse figure to give a more complete mechanism; "What many studies of the liberalization process have in common is their treatment of the impetus for liberalization as an exogenous shock. These shocks may be political or economic in nature, but either can force elites to take some action to restore the legitimacy of their regime. Disagreements then arise within the authoritarian bloc as to the prudent course of action. Some regimes may be able to weather the crisis given a variety of factors, ranging from the nature of the current autocratic regime, economic conditions, to the past performance of the regime. Other regimes may decide to liberalize in an attempt to restore legitimacy." (Pevehouse 2005, 17)

203. Fearon and Laitin in their oft-cited work (2003) on civil war (in)famously include a variable on amount of mountainous terrain (see Collier and Hoeffler 2004). This is controversial in terms of causal inference because it essentially does not vary within countries. Influential causal inference scholars deny causal status to variables which cannot be manipulated. Discuss how one use large-N qualitative case studies to explore the causal mechanism linking mountaineous terrain to civil war. Can this illustrate how a constant can be part of a causal mechanism?

204. Causal mechanisms and scope decisions can be closely related. Bush's main independent variables in her analysis of gender quotas are (1) foreign aid, (2) international election monitoring, and (3) democracy promoting UN peace operation. She could include all countries in her analysis, but she excludes rich democracies.

> A Quantitative Analysis of the Sources of Gender Quotas. The Sample. The unit of analysis for the quantitative analysis of the determinants of quota adoption is the country–year. The sample covers the years from 1970 to 2006 and contains all countries except long-term consolidated and developed democracies. Long-term consolidated and developed democracies were removed because they are subject to different causal processes; they are neither under UN authority nor desirous of signaling their liberalism to the international community, and in fact, they promote democracy abroad. I followed Finkel et al. and removed thirty advanced industrial, long-term consolidated democracies, which resulted in a sample of 165 countries for at least some amount of time. (Bush 2011, 118)

She does include authoritarian regimes, because some do adopt gender quotas. Discuss whether the same causal mechanism works for authoritarian as well as democratizing regimes.

205. Causal mechanisms and scope decisions can be closely related. Bush's main independent variables in her analysis of gender quotas are (1) foreign aid, (2) international election monitoring, and (3) democracy promoting UN peace operation. She could include all countries in her analysis, but she excludes rich democracies.

> A Quantitative Analysis of the Sources of Gender Quotas. The Sample. The unit of analysis for the quantitative analysis of the determinants of quota adoption is the country–year. The sample covers the years from 1970 to 2006 and contains all countries except long-term consolidated and developed democracies. Long-term consolidated and developed democracies were removed because they are subject to different causal processes; they are neither under UN authority nor desirous of signaling their liberalism to the international community, and in fact, they promote democracy abroad. I followed Finkel et al. and removed thirty advanced industrial, long-term consolidated democracies, which resulted in a sample of 165 countries for at least some amount of time. (Bush 2011, 118)

She does include authoritarian regimes, because some do adopt gender quotas. Discuss whether the same causal mechanism works for authoritarian as well as democratizing regimes.

206. A key question in the causal analysis of individual cases is the relationship between temporal distance and causal importance. Should events or factors closer to the event to be explained receive greater or less causal importance. Analyze the quote below from Gerring (2005) who argues for increasing impact as temporal distance increases in the causal chain. Does it matter whether the cause is a necessary or sufficient condition?

> Consider the following path diagram. $X_1 \Rightarrow X_2 \Rightarrow X_3 \Rightarrow X_4 \Rightarrow Y$ We are apt to consider X_1 to be the cause and causal factors $X_2 - X_4$ intermediate (and less important) causes, all other things being equal. Of course, all other things are rarely equal. We are likely to lose causal power (accuracy and completeness) as we move further away from the outcome. Yet, if we did not – e.g. if the correlations in this imaginary path diagram were perfect – we would rightly grant priority to X_1. Causes lying close to an effect are not satisfying as causes, precisely because of their proximity. Rather, we search for causes that are 'ultimate' or 'fundamental'.
>
> Consider a quotidian example. To say that an accident was caused because A ran into B is not to say much that is useful about this event. Indeed, this sort of statement is probably better classified as descriptive, rather than explanatory. An X gains causal status as it moves back further in time from the event in question. If, to continue with this story, I claim that the accident was caused by the case of beer consumed by A earlier that evening, I have offered a cause that has greater priority and is, on this account at least, a better explanation. If I can show that the accident in question was actually a re-enactment of a childhood accident that A experienced 20 years ago, then I have offered an even more interesting explanation. Similarly, to say that the Civil War was caused by the attack on Fort Sumter, or that the First World War was caused by the assassination of the Archduke Francis Ferdinand at Sarajevo, is to make a causal argument that is almost trivial by virtue of its lack of priority. It does not illumine very much, except perhaps the mechanism that might be at work vis-a-vis some prior cause.
>
> The further away we can get from the outcome in question, the more satisfying (ceteris paribus) our explanation will be. This explains much of the excitement when social scientists find 'structural' variables that seem to impact public policy or political behavior. It is not that they offer more complete or more accurate explanations; indeed, the correlations between

> X and Y are apt to be much weaker. It is not that they are more relevant; indeed, they are less relevant for most policy purposes, since they are apt to be least amenable to change. Priority often imposes costs on other criterial dimensions. Yet, such explanations will be better insofar as they offer us more power, more leverage on the topic. They are non-obvious. (Gerring 2005: 174–5)

207. The AMAR "rebellion" scale (Birnir et al. 2018) has as its lowest rebellion level "political banditry and/or sporadic terrorism." Often zero on the scale (like in many dichotomous variables) is potentially a causally heterogeneous category. Discuss to what extent *theoretically* zero is conceived of as "no mobilization" as opposed to nonviolent mobilization such as protests or nonviolent action. Would the causal mechanisms leading to these two different forms of zero be the same or different?

208. Soifer (2012) explores the concept of critical juncture along with critical antecedents in the case of import substitution policies in Latin America.

> The distinct feature of a historical juncture with the potential to be critical is the loosening of the constraints of structure to allow for agency or contingency to shape divergence from the past, or divergence across cases. . . . we must distinguish between two types of causal conditions at work during the critical juncture: the permissive conditions that represent the easing of the constraints of structure and make change possible and the *productive conditions* that, in the presence of the *permissive conditions*, produce the outcome or range of outcomes that are then reproduced after the permissive conditions disappear and the juncture comes to a close. The two types of conditions are nearly always framed as separately necessary and jointly sufficient for divergence to occur. (Soifer 2012, 1573)
>
> They [Slater and Simmons 2010] define a critical antecedent as "factors or conditions preceding a critical juncture that combine in a causal sequence with factors operating during that juncture to produce a divergent outcome" (p. 889) . . . . They write that the critical antecedent "does not produce its causal effect by causing the independent variable to emerge. It does so by helping to determine the differential causal effect of the independent variable across cases when the critical juncture exogenously comes about" (p. 891). Here we see that the critical antecedent is, in their formulation, unrelated to the permissive condition, which (for them) emerges exogenously. Permissive conditions are also distinct from critical antecedents in that although the former mark the temporal bounds of the critical juncture, the latter are operant *before* the juncture emerges. On the other hand, critical antecedents are connected to the productive condition. (Soifer 2012, 1576)

Draw a causal mechanism figure for the table below.

Table 1. Inward-Looking Industrialization as a Critical Juncture

Critical antecedent	Strength of middle class and labor as of 1929
Permissive condition	Collapse of world trade and economic challenges of Great Depression and World War II
Productive condition	Economic ideas of ECLA and more general rise of economic nationalism
Outcome	Inward-looking industrialization implemented to varying degrees in Latin America
End of critical juncture	Recovery of world trade by 1950, and especially after the Korean War
Mechanisms of reproduction	New political coalitions among bureaucrats, domestic elites, and organized labor
Consequences	Crises of populist rule and bureaucratic-authoritarian regimes (O'Donnell, 1973)

209. Slater and Simmons (2010) provide quite a few causal mechanism figures, such as

    Figure 2. McAdam on King Cotton and Black insurgency

    Figure 3. Grzymała-Busse on Communist Party regeneration

    Figure 4. Family law and gender inequality in North Africa.

    Figure 5. The Colliers on labour incorporation in Latin America

    Figure 6. Luebbert on interwar European regime outcomes

    Figure 7. Kalyvas on Christian Democratic parties in Western Europe

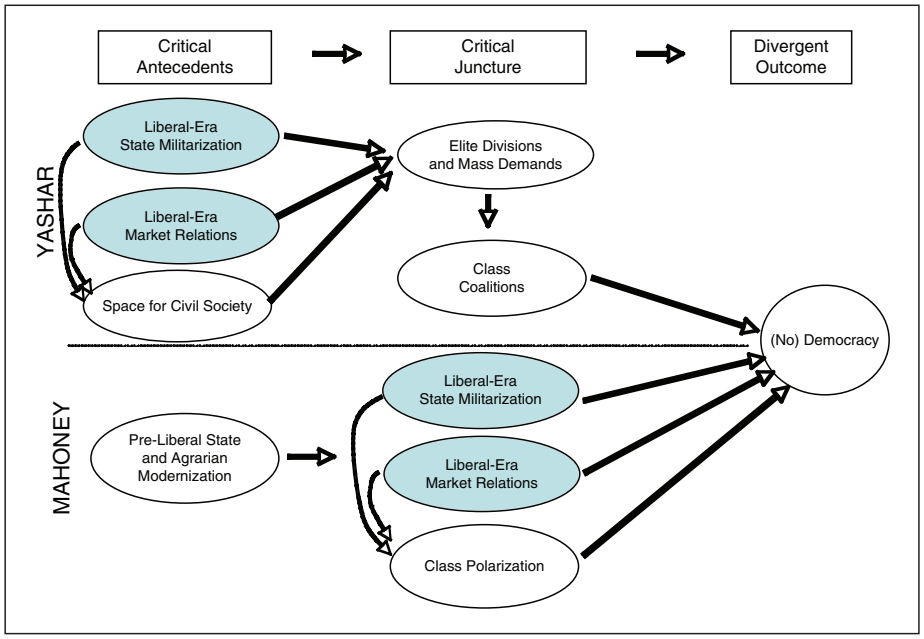
    Figure 8. Regime outcomes in Central America [Yashar, Mahoney]

    Figure 9. Ethnic violence in India [Varshney, Wilkenson, Kohli]

    Figure 10. Racial politics in Brazil and South Africa [Marx, Lieberman]

    Discuss these causal mechanism figures. For example, Do the arrows all mean the same think? Are some of the ellipses necessary conditions? What are the causal connections between text with boxes in front of them and the ellipses?

    For example, here is the McAdams figure.



210. Mansfield and Pevehouse (2000) give a variety of mechanisms that could link preferential trade agreements to reduce military conflict. Do a causal mechanism graph of their theory (like those in chapter 2 of *Multimethod research, causal mechanisms, and case studies*, note that the core hypothesis is an interaction term between trade and PTA membership. A similar exercise can be done with the causal mechanisms leading from common IGO membership to less international conflict: "Indeed, a close look at various international organizations indicates they may serve any of six functions: coercing norm breakers; mediating among conflicting parties; reducing uncertainty by conveying information; problem-solving, including expanding states' conception of their self-interest to be more inclusive and long-term; socialization and shaping norms; and generating narratives of mutual identification. (Russett et al. 1998, 444–45). Discuss whether these "functional" arrows in the causal mechanism diagram should be interpreted as causal.

211. Discuss the causal figures in *Multimethod research, causal mechanisms, and case studies*, Weller and Barnes (2014) or Waldner (2014) from a DAG cross-case, large-N, probabilistic perspective. Are these figures identified from a DAG perspective (e.g., Morgan and Winship 2015). Does it matter from a within-case causal inference perspective if they are not?

212. Grzymała-Busse (2007) is a core example in the causal mechanism chapter of *Multimethod research, causal mechanisms, and case studies: an integrated approach*. In my figure 2.4 of her theory "Robust competition" is constituted by three mechanisms (1) moderation, (2) anticipation, and (3) cooptation. Discuss how on might aggregate or combine these causal mechanisms (e.g., OR, AND, addition). Mikkelsen (2017) uses Grzymała-Busse (2007) a core example in his discussion of fuzzy logic case studies. Compare and discuss his version of the mechanism of competition, e.g., Figure 2 and how he uses fuzzy logic to

combine the three mechanisms with the discussion in chapter 3 of *Multimethod research, causal mechanisms, and case studies: an integrated approach* of the same mechanism.

213. Different causal mechanisms and control variables. Sometimes it is not clear whether something is a separate causal mechanism (e.g., separate X) or whether it is part of some larger causal mechanism, i.e., an M_i factor. For example, Simmons and Elkins (2004) in a widely cited article discuss various mechanisms of policy "diffusion." (I use the scare quotes because it is not clear that they are all diffusion mechanisms, not other kinds of mechanisms, e.g., market competition.) They contrast three theories about why states adopt liberal economic policies. The first is via market competition, if competitor nations are adopting liberal policies and the state adopts to compete in markets. A second mechanism is global norms about economic policies. A third is learning from networks, successful states,or culturally similar states. There are independent variables in their statistical model representing each of these mechanisms. So should these independent variables be considered separate mechanisms, e.g., separate X_i or M_i within a larger theory of diffusion?

214. The $\rightarrow$ in causal mechanism figures is often ambiguous. Sometimes it can mean a noncausal relationship. What would be some situations where this is the case?

    Answer:

    One set of situations is where the arrow is ontological or definitional in nature. Another is game trees or other decisionmaking figures where the arrows indicate different choice options.

Constraint mechanisms

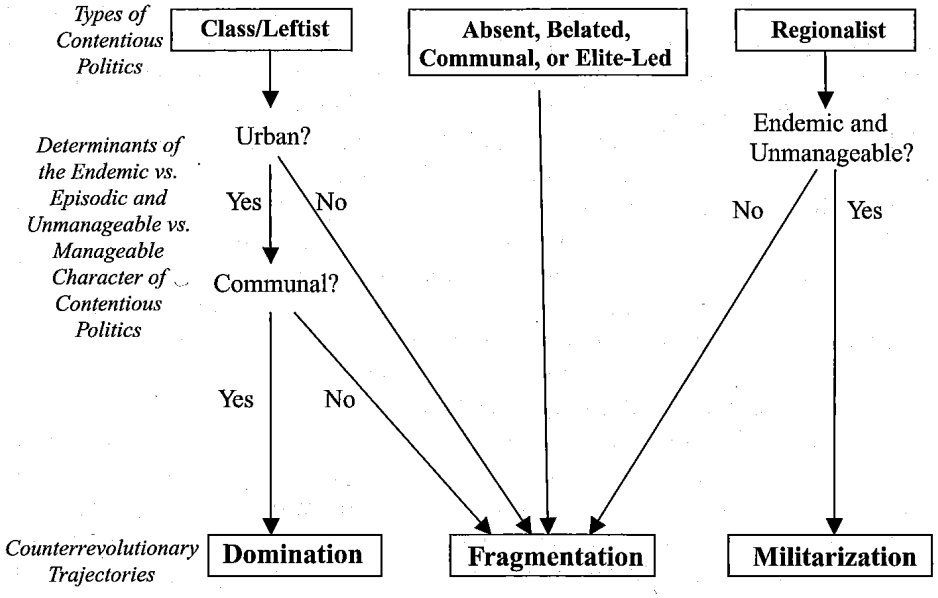
215. Hulme (2026) argues that Congress is a strong constraint on presidential uses of force. He has a nice scatterplot illustrating that congressional approval is a necessary condition. One of the key themes in the qualitative methods literature as well as quantitative is alternative explanations and confounders. Hulme reports (personnel communication) that many suggest that an alternative explanation is the president's view on the use of force. So this could be a control variable or confounder. Goertz (2017) argues that in terms of case selection one should follow the "avoid overdetermination rule". Which cases in Hulme are overdetermined? Which cases then should be chosen when applying the avoid overdetermination rule? Notice the same issues arise in Tannenwald's influential book on nuclear weapons taboo. (Thanks to Hulme for comments on this exercise.)

216. Sufficient condition scatterplots such as the one below suggest that there is potentially a floor below which observations cannot go below. Discuss the two sufficient conditions in the scatterplots and contrast them with the one nonsufficient condition one. Discuss the extent to which popular participation

in the initial stages of constitution making sets a floor for how democratic the country can be later on. Note that this is one of the strongest findings book.

> Next, we disaggregate the participation variable into convening, debating, and ratifying stages in order to test an "origination" hypothesis predicting that the first stage, convening, has the largest impact on democracy. We confirm that participation at this earliest stage is most critical: democracy improved in only 45 percent of cases that incorporated broad consultation at debate and ratification stages, but not at the convening stage. Contrarily, 82 percent of the cases in our data that used popular convening, regardless of popular participation in later stages, show such improvement. (Eisenstadt et al. 2017, 144

FIGURE 4. Contentious Politics and Counterrevolutionary Trajectories



217. Social constructivists frequently uses the term "permissive cause." Typical of social constructivists. Think Carpenter (2007) discusses agenda-setting in terms of constraints and her causal mechanism figure, figure 1. p. 102. So the absence of permissive condition explains why children of rape is not on the agenda? What do the lines mean (i.e., not arrows). Do they just indicate temporal ordering?

218. Discuss the Rodrik (2007) figure below. Note that everything is conceptualized as a constraint (see chapters 2 and 4 of *Multimethod research, causal mechanisms, and case studies: an integrated approach*). Which case studies would be useful? Which ones does he do? Discuss the nature of the arrows – e.g., are they causal – and the two-level structure of the causal mechanism. Could the arrows be casually going up instead of down?

> Moving down the branches of the decision tree is tantamount to discarding candidates for the most binding constraint on growth. The overarching lesson from our theoretical analysis is that it is this constraint, once identified, that deserves the most attention from policymakers. (Rodrik 2007, 66)

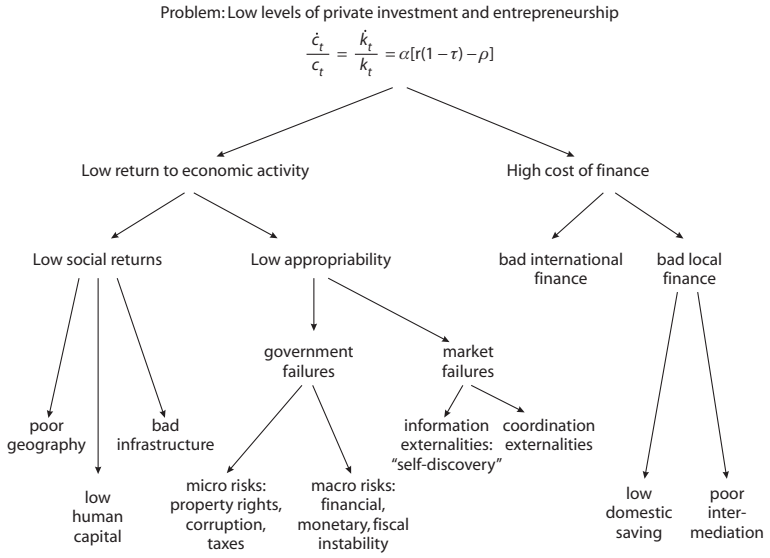


Fig. 2.1. Growth diagnostics

Concepts

219. Boulanger Martel et al. (2024) conduct a literature review of peace with adjectives. They give numerous examples, 61, of people attaching adjectives to the concept of peace. Discuss how or whether this make sense unless one knows what peace is to begin with, i.e., peace without adjectives.

220. Apply the Goertz three-level framework and diagram the concept of civil war complexity described by Brosche et al. 2023, illustrated in their triangle figure below. Once diagrammed in this fashion discuss how one might aggregate across the three dimensions. This is probably necessary to answer their "how much" question and would probably be necessary if civil war complexity were to be used in statistical analyses. The elephant in the room is the concept of *complexity*, which is itself not really defined. Given the three corners of the triangle discuss how one might conceptualize and measure complexity.

> Conceptualizing Civil War Complexity. To identify a civil war's base attributes and to meaningfully capture different degrees and types of complexity, we turn to classic conflict theory. Building on Johan Galtung's conflict triangle, Wallensteen defines conflict as consisting of three elements: (1) actors, (2) behaviors, and (3) contradictions/issues at stake. As such, building our conceptualization on the latter's conflict triangle brings nuances to how civil wars vary and identification of its base components. Allowing civil wars to vary in the type and number of actors, behaviors, and issues thus helps define a system's parts and their interlinkages. We define civil war complexity as the degree to which the core components of conflict – actors, behaviors, and issues – are manifold in number and type, and the degree to which these are interlinked. We unpack each component of our concept both theoretically and empirically. Intertwining

theory and empirics allows us to move between Giovanni Sartori's "What is?" question regarding the boundaries of our concept, and the"How much?" question regarding empirical measurement. . . . First, to demonstrate the variations of actor complexity, Figure 2A,B displays variations in the number of actors participating in violence in a country between 1989 and 2016. Figure 2A includes state, nonstate, and communal actors, whereas Figure 2B shows rebel nonstate actors only. (Brosch et al. 2023, 143–44, 156)

Figure 44: Conceptualizing and measuring civil war complexity

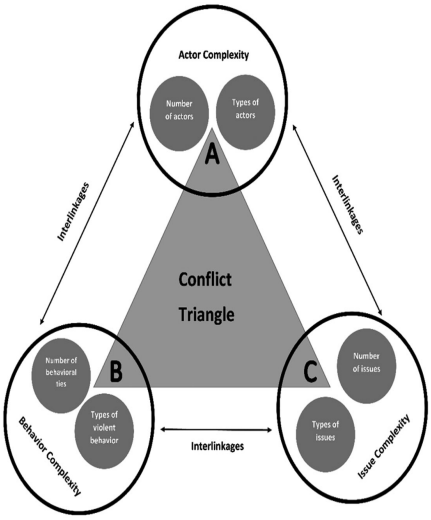


Figure 1. Graphical illustration of conceptual model.

Source: Brosche et al. 2023.

Answer: Since complexity is a continuum one approach would be to start with what configuration of the three corners is the least complex? One might suggest that if one corner scores the maximum while the other two corners are at the minimum that is the least complex conflict. This then suggests that maximum complexity is achieved when all three corners are equal in size. Following this logic then one could use approaches to equality (e.g., income inequality) to conceptualize and measure conflict complexity.

221. One of the core guidelines when thinking about conceptualization is to think about terminology. Core to a massive data project is the concept of "social conflict" as defined below. What is one core component of the concept described below that probably should be included in the term "social conflict"? How does this have concrete implications for the data set described in their table 1?

> We used the expression"social conflict" as a generic category to encompass all the expressions used by scholars and/or mentioned in the sources. We followed Nicolas (2002, 75) and considered social conflict to be any event

> involving a group of at least three individuals belonging to different families and which either perpetuates violence or threatens violence against one or more members of a different group or against representatives of political, religious, and economic power; or any event involving an attack on property, buildings, furniture, papers, or other signs symbolizing such powers. We chose to use a low threshold on the number of participants not only to grasp infinitesimal episodes of discontent, but also to allow for the inclusion of other datasets that used different thresholds. (Chambru and Maneuvrier-Hervieu 2024, 1085)

Answer: Central to their concept is the notion of *violent* conflict which lies at the core of their definition above. Instead of being assumed it should be made explicit because there is a whole literature on nonviolent conflict which may or may not be relevant.

222. As discussed in chapter 3 of the concept book any concept that involves two words is potentially problematic because one in effect has often two concepts each of which is worthy of its own analysis and then the question about how to combine the two into one hyphenated concept. Helmke and Rath (2025) Illustrate this issue in their discussion of "democratic norms." As illustrated below they discussed at length the concept of a norm from a rational choice perspective as an equilibrium. But they do not discuss is how the adjective democratic modifies a norm. Discuss how this might be done. How does one use a concept of democracy in the process? Does one add definitional features of democracy to those defining norms?

> Specifically, we define political norms as informal behavioral rules that an individual follows in a given political circumstance only if (a) they expect others to follow the rule and (b) they believe others believe they ought to follow the rule. The first condition commonly refers to empirical expectations and suffices to establish purely descriptive norms. The second condition refers to normative expectations and is necessary for creating injunctive or social norms. . . . For any norm to exist, a sufficient number of people must also recognize that a given circumstance (i.e., a particular event in a certain context) calls for a given behavioral response, or what Bicchieri (2006, pp. 11–13) refers to as the "contingency condition." Just as crucial, we argue, is whether a sufficient number of people also believe that the circumstance itself exists. . . . Having roughly pinned down what we mean by norms, what makes a given norm democratic? (Helmke and Rath 2025, 236–37)

223. When scholars want to dichotomize democracy (or break it into categories) using the V-Dem data they usually use the fairly complicated scheme illustrated in the figure below, figure 45, from Lührmann et al. (2018). Basically, for example, .75 is *not* the threshold for "liberal democracy" Lührmann et al. add three additional, necessary requirements. Does this means that gthe V-Dem scale by itself "overshoots" the concept of liberal democracy? Could you use these ideas to reformulate the V-Dem continuous measure so that .75 is now "liberal democracy", .50 is "electoral democracy", .25 and .25 is "electoral autoracy"? Also, they broke up the continuum into four parts, are there arguments for breaking it into 5 parts?

Figure 45: Dichotomizing democracy: the Lührmann et al.–V-Dem approach

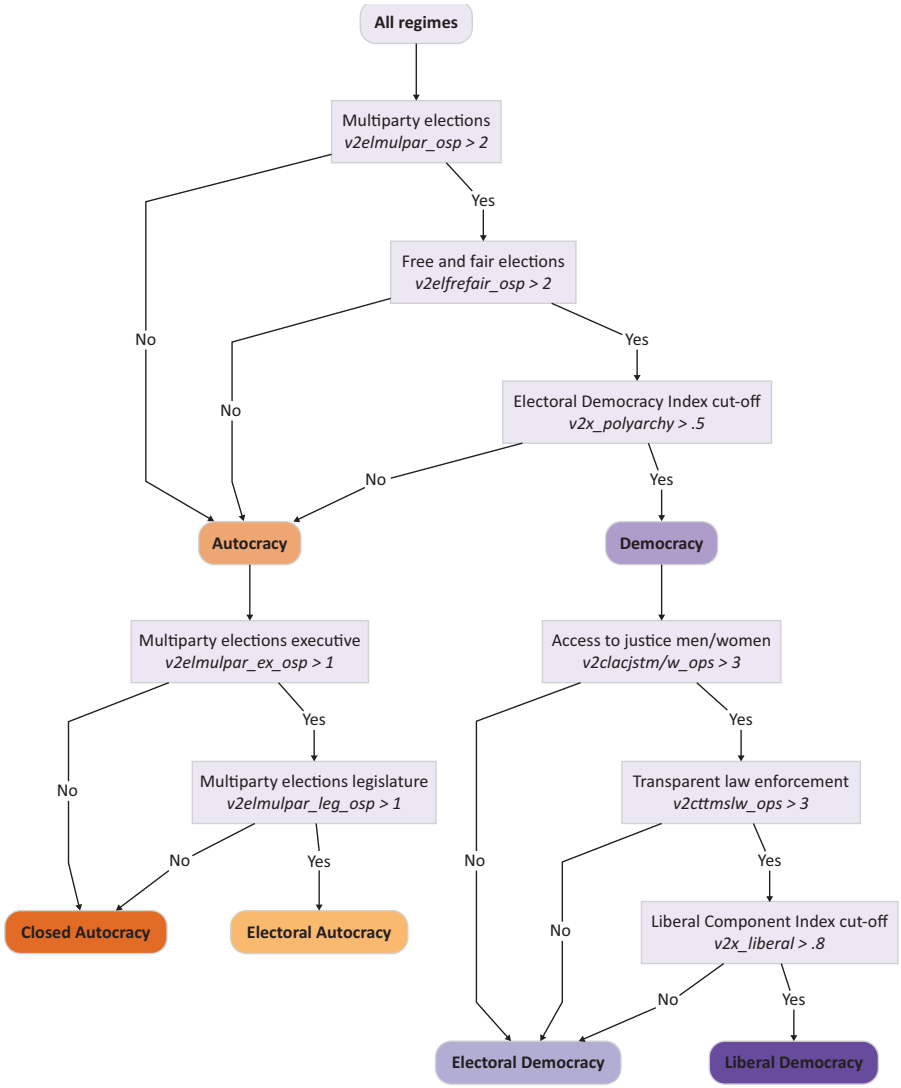


Figure 1. Coding schema for the RoW typologies (for descriptions of variables see Coppedge et al., 2017a).

Answer: For the last question, the concept book recommends odd-number divisions. This permits there to be a zone in the middle. Even number divisions have the middle split between two different categories.

224. As discussed in chapter 3 of the concept book any concept that involves two words is potentially problematic because one in effect has often two concepts each of which is worthy of its own analysis and then the question about how to combine the two into one hyphenated concept. Helmke and Rath (2025) Illustrate this issue in their discussion of "democratic norms." As illustrated below they discussed at length the concept of a norm from a rational choice perspective as an equilibrium. But they do not discuss is how the adjective democratic modifies a norm. Discuss how this might be done. How does

one use a concept of democracy in the process? Does one add definitional features of democracy to those defining norms?

> Specifically, we define political norms as informal behavioral rules that an individual follows in a given political circumstance only if (a) they expect others to follow the rule and (b) they believe others believe they ought to follow the rule. The first condition commonly refers to empirical expectations and suffices to establish purely descriptive norms. The second condition refers to normative expectations and is necessary for creating injunctive or social norms. . . . For any norm to exist, a sufficient number of people must also recognize that a given circumstance (i.e., a particular event in a certain context) calls for a given behavioral response, or what Bicchieri (2006, pp. 11–13) refers to as the "contingency condition." Just as crucial, we argue, is whether a sufficient number of people also believe that the circumstance itself exists. . . . Having roughly pinned down what we mean by norms, what makes a given norm democratic? (Helmke and Rath 2025, 236–37)

225. There are innumerable conceptualizations and measures of democracy but an explicit and separate conceptualization of authoritarianism is very rare. Almost always in theory, as well as in practice, authoritarianism is defined as not-democracy. A core guideline in Goertz (2020) is that if something is a core concept to a research project should be conceptualized on its own and not via a "not." Gerschewski (2023) is a rare example of an author explicitly defining authoritarianism. Discuss his three dimensions of autocratic rule: (1) ideational performance, (2) repression, and (3) cooptation of elites, and the extent to which it is or is not a mirror image of democracy. Notice also his various three-level concepts of related ideas, e.g., repression (table 3.2). See, for example, his figure I.1.

226. Because scaling is virtually never taught in political science methods classes most scholars doing quantitative analysis and creating indices are unfamiliar with how this works. Very common is the adding of ordinal measures or dummy measures to create an index. Discuss the problems with this in this particular example:

> To elucidate the causes and consequences of intervention, MIP measures several national interests-related variables from 1776 to 2019. We apply a National Interests Index across the eras that adds up separate measures on contiguity, colonial history, alliances, and natural resources. This additive index contains relative measures of geopolitical importance between country-dyads, including factors such as geographic continuity, shared alliances, colonial history, and the presence of oil and gas. We calculate the ordinal index, using State B target data, as follows: OilDummy + ColonialDummy + AllianceDummy + ContiguityDummy. (Kushi and Toft 2022, 15)

Answer: They have converted the dummy variables into interval ones in the process of creating the index where now zero and one are not nominal but become interval when put into the final index.

227. One relatively popular way to code the number of deaths in conflict datasets is to use data ranges. This is not a good idea. What would be a better approach and why? Here is one example from a prominent conflict data set.

> The relevant variables from ECAV are deaths and event description. Ideally, one would measure lethality using death estimates. However, such data are not available in ECAV, which codes the death variable on an ordinal scale. The five levels begin with no deaths; fewer than 10; 10 to 100; 100 or more; and unknown, which represents "uncertainty over whether (or how many) deaths occurred" (Daxecker et al. 2019, 16).

Answer: By cutting them as ranges one is implying that all the cases the given range are somehow "the same." This has the additional disadvantage of requiring one to either do some kind of logit on the dependent variable side or some combination of dummy variables on the independent side. A much better approach would be to give whatever estimate the researcher has of the actual deaths with some measure or indication of how confident she is and the range of plausible values for that estimate. For example compare the COW data ranges for the militarized dispute data set with Gibler's approach of giving estimated values and ranges for the estimated value. UCDP often does similar things.

228. It is quite common in conceptualization to give a series of necessary conditions, the same is often true for coding rules for quantitative datasets. What is often not stated – but is often implicit – is whether these necessary conditions are jointly sufficient. Are Fazal's three necessary conditions jointly sufficient for declaration of war?

> International legal scholarship suggests several necessary criteria for declarations of war. First, a declaration of war must be a public proclamation. Second, it must state intent to engage in hostilities with another party. Third, it must be issued according to the laws of the declaring state; it can only be a declaration of war if authored by an actor empowered with the authority to declare war. (Fazal 2012, 562)

229. As a general rule quantitative scholars are not very interested in the relationships between their various independent variables. However, from a conceptual and causal inference point of view there are important issues which often arise. A classic one is when two independent variables are somehow part of the same concept. Barnhart et al. (2020) illustrate this where they have two core variables one which is democracy, using polity data, and the second women's suffrage. Arguably, see Paxton's work, women's suffrage is part of the concept of democracy. However, many traditional democracy datasets do not include that factor, including polity and others. They make strong claims about the role of women's suffrage in connection with international conflict and more specifically the democratic peace literature:

> Finally, we examine cross-national conflict data to assess the impact of the enfranchisement of women. We find that female suffrage, when coupled with democratic institutions, is not only an important and powerful cause of peace, but that it may even be necessary for the democratic peace. Without female suffrage, democracies (whether examined dyadically or

> monadically) do not show clear evidence of being more pacific than non-democracies. These findings hold across time periods and do not appear to result from other confounding factors, or from the expansion of the political franchise more generally. By bridging the individual and international levels to show how the gender gap on the use of force matters, we contribute to the growing positivist study of gender and international relations. (Barnhardt et al. 2020, 634)

Someone might argue that it is not women's suffrage that is really the core causal factor but rather *quality of democracy*. Countries with women's suffrage, or more generally, fewer restrictions to suffrage based on race or whatever are higher-quality democracies, *by definition*. So maybe the article should be called the higher-quality democracy peace? (See Goertz and Dul 2023 for an argument and data analysis to this effect).

230. Those doing domestic armed conflict have the same problem as those defining states in the international system. How to define organizations or groups that are involved in militarized conflict, either with each other or with the state or against individuals.

A recent major data set illustrates the various issues involving in conceptualizing armed, violent, actors at the domestic level.

> The GTD includes attacks that involve "the threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation" (START 2021).... Defining Armed Groups. We define an armed group as an independent organization of non-state actors that justifies the use of violence in the pursuit of political control.... We then use these case studies to exclude approximately 50 percent of candidate groups (1,009 entries) for not meeting the definition of an armed group. In practice, this removes criminal actors, cartels, individuals, pro-government militias, factions, duplicate aliases, and generic names listed in the GTD.... The "pursuit of political control" criteria require evidence of an underlying political cleavage or issue incompatibility. A group must aim to change a policy, regime, or piece of territory through violent means. This decision rule serves two purposes. Because "a terrorist act by definition must have a direct political goal this means that the GTD excludes ordinary criminal violence." Congruently, we exclude criminal actors who perpetrate violence for material rather than political ends.... Table 1. Excluded cases from candidate entries. An armed group may be excluded for multiple reasons (Malone 2022)

Like in the exercise above about the international system, one main issue is veto criteria that eliminate possible actors, e.g., drug cartels. This comes across clearly in the table about exclusion criteria.

The second issue is gray zones: all concepts have significant gray zones. There are at least three gray zones that make it hard to draw a clear line about inclusion or exclusion: (1) The first is the fact that there is a gray zone between armed groups and nonviolent groups. It is well known groups can shift back and forth in many campaigns and use both violent and nonviolent means. (2) The second big gray zone deals with what is often called criminal

or economic armed groups. As literature on organized crime in Latin America shows clearly, these are massively important political actors. They are often the outcome of civil wars and organized crime is very involved in elections and the assassination of various elected officials. (3) Terrorism datasets typically require (wrongly I think) that the organizations be "non-state." But as is well know pro-government militias are an important source of terrorism, e.g., KKK in the US south was very closely connected to the state.

The two conceptual questions mostly revolve around the list of veto criterion and how one deals with important gray zone questions. Discuss how Malone deals with these issues.

231. One large conceptual problem lies with concepts which are defined by negation (Ragin 2023 has a very nice discussion of this). Often the positive concept is clearly defined but the negation of it is a large heterogeneous set of cases. The COW concept of extrasystemic wars falls into this problem. In general the solution involves giving a positive characterization of these negative cases.

    Extrasystemic wars involve wars between "political units" one of which is not a member of the international system. For example, wars with a system member against a non-system actor such as colonial wars and wars against indigenous peoples. The core feature is that one political entity in extrasystemic wars is negatively defined, i.e., not-states.

    So the conceptual question is in some ways exactly the same as for the international system: what constitutes in terms of a positive definition a political unit that can be involved in an extrasystemic war? These can be kingdoms, as in Africa, indigenous groups and tribes, princely kingdoms and the like. It is a tremendous conceptual problem trying to delimit this universe: so does it make any sense at all to talk about extrasystemic wars?

232. Critical to the study of international relations, particularly quantitative approaches is the definition and identification of the states of the international system. The COW project (Russett et al. (1968) along with the related Gleditsch and Ward (1999) provide approaches that have defined states. However, there is still much conceptual and hence operational confusion over how states are to be defined. This can be critical for the study of war. For example, any wars between Latin American political units which are not officially states does not appear in the canonical war data set. These then become the so-called extrasystemic wars. See the next exercise for discussion of this particular concept.

    Overall the COW system might be called a basically a hybrid recognition system for states. The three big ones are recognition by United Kingdom and France in the 19th century and later, membership in the League of Nations, or membership in the United Nations. One can think about this via the existence of "veto" players or criteria. For the 19th century the UK and France are veto players. Should the League of Nations or the United States be veto players? One conceptual question is whether one actually needs any veto criterion at all? The main veto criteria consist of:

- Occupations – COW is inconsistent about whether an occupation kicks the state out of the system. Apparently short occupations are OK, for example, the US in Iraq or Afghanistan. But in general occupations are treated inconsistently, for example see German occupations during World War II.
- Major power vetoes – in the 19th century United Kingdom or France are veto players in the sense that if they do not recognize a country it does not make it into the system. In the post-World War II period a veto question is whether major powers can veto membership of individual states such as in the case of United States–Palestine, Russia–Kosovo, China–Taiwan.
- Population – COW is inconsistent about the minimum population required for membership. This is applied typically in the 19th century, but there are obvious members of the United Nations that are below this population minimum. This is a veto criterion.

Another set of questions deals with dates particularly in the setting of post-colonial or secession states. When do they enter the system? Canada is an interesting case where Gleditsch and Ward have it gaining independence in 1867 while COW has it entering the system in 1920. So the conceptual question is whether one should date the country based on independence or recognition criteria?

Finally, the core conceptual question is whether one could construct a set of criteria for membership in the international system based purely on multiple sufficient recognition criteria, with no veto criteria.

See the exercise below with the same issue arises in defining arms groups in the context of civil war conflict.

233. Healy (2019) has written an important book on effective communication of data in figures. A key point is to clearly communicate in the figure the relationship between X and Y. He uses an example given in the figure below with heavily skewed data for GDP per capita against life expentancy. This kind of skewness in data is quite common in the social sciences. He gives you two different plots and suggest that the logged data are better. Discuss the extent to which this is true.

Figure 46: Log transformations of data: life expentancy and GDP/capita, raw data

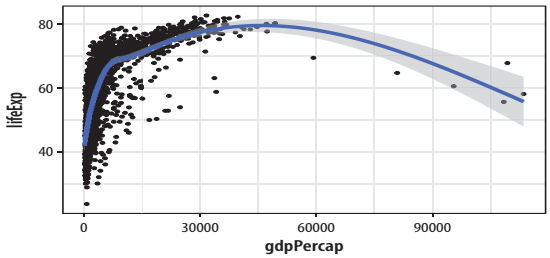


Figure 3.6: Life expectancy vs GDP, showing both points and a GAM smoother.

Figure 47: Log transformations of data: life expentancy and GDP/capita, logged data

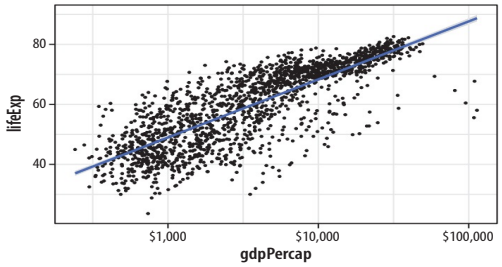


Figure 3.9: Life expectancy vs GDP scatterplot, with a GAM smoother and a log scale on the x-axis, with better labels on the tick-marks.

Answer: As Healy notes:

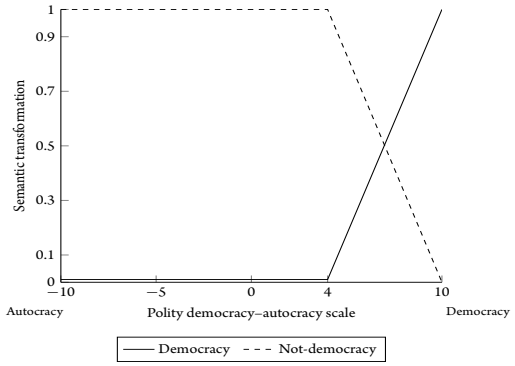
> If the relationship between X and Y is really nonlinear then logging to make it look linear is in fact falsifying in some ways the relationship between the two. In addition, for those interested it is not necessarily clear that logging is the best nonlinear fit for the data. There may be other nonlinear models that fit the data better. (Healy 2019, 61)

234. One of the main guidelines in chapter 3 of the concept book is that if your focus is on a particular concept you should develop a bespoke concept and measure of it. In the whole literature on democratic backsliding the core concept is some concept and measure of democracy. As discussed above, this is then modified by some definition of what constitutes backsliding, either as some movement in the authoritarian direction or maybe passing some threshold. One option is to actually focus on what constitutes backsliding. Most of the contributions to the Little-Meng symposium stress that backsliding is not uniform across all the defining dimensions of democracy. One could start by

asking about the tatics of backsliding executives: what aspects of democracy do they attack first, most often, or most effectively? Then democratic backsliding would be looking at those specific features of democracy which are most involved in backsliding. What are the dimensions of democracy most attacked by backsliders? What are the indicators of them? Could you construct another more specialized backsliding concept and measure using indicators of the V-Dem or Little-Meng? For example, in the Baron et al. contribution to the symposium (2024) on Little and Meng they list the following "symptoms" – in my terminology tatics – of backsliding in their table 1: reduction in horizontal accountability, reduction in judicial independence, reduction in legislative oversight, weakened civil service or integrity institutions relaxation of term limits, reduction in vertical accountability, repression of the opposition, systemic reduction in election freedom and fairness curtailed civil liberties, media repression. Could these be used to create a measure of democratic backsliding?

235. Chapter 4 of the concept book deals with transforming basic raw data to better conform with concepts. Instead of the standard linear model (see the exercises above about scaling and using the average). The book suggests that if one is interested specifically in democracy (and not authoritarianism) one potential transformation using the polity scale looks like figure **??**. Discuss this option in the context of the Little-Meng debate. In particular, (1) how does it deal with the exercise above blacksliding among democratic versus authoritarian states? (2) Is it going to emphasize the degree of backsliding in comparison with the original scale?

Figure 48: A semantic transformation of a democracy measure, focusing on democracy not authoritarianism



Answer: Notice that all countries with a value of less than 4 have zero on this transformed democracy measure. So any change among those does not constitute backsliding because the changes are zero. This illustrates the principle discussed in the concept book whereby some changes in the basic raw data do not involve changes in the concept. Notice as well that this particular transformation stretches out the democracy scale which now runs

from 0 to 1. In the V-Dem scheme of things .5 is often taken as the cut off for the authoritarian regime so the democracy scale runs from .50 to 1.0. So inherently the raw backsliding values have a maximum of .5, since .5 is the cutoff, where is in the transformed data in figure **??** have a maximum backsliding value of 1.0

236. A critical issue in the democracy literature is the concept itself of "democratic backsliding." This is another nice example of a two-concept concept one part of which is "democracy" and the other part of which is "backsliding." Critical in this debate is that measures of democracy typically run from completely authoritarian to completely democratic, e.g., 0 to 1.0 for V-Dem. So does backsliding only apply to those that are mostly democratic or to all countries across the spectrum? Does backsliding simply mean a decrease in your authoritarian-democracy score? This can take place among very authoritarian governments as they move from, say, .30 to .10 on the V-Dem scale. The approach that many take exemplified by Miller's contribution to a symposium (2024) is to dichotomize democracy and authoritarianism and only look at backsliding among the democratic group. This has the standard problem of all the dichotomous measures which is the sharp border between democracy and authoritarianism. So for example if a country moves from democratic to the authoritarian category, e.g., .60 to .40 on the V-Dem scale is backsliding, but .40 to .20 is not backsliding? Does that even constitute "backsliding" to begin with, if backsliding means "not large negative changes?" Does an "autocolpe" or presidential coup constitute backsliding? Discuss how valid it is for Little and Meng to include all countries, authoritarian and democratic, to analyze trends in changes in democratic backsliding.

    Here is Miller on the topic:

    > A specific focus on democracies is warranted for at least two reasons. First, the narrative of democratic decline that L&M critique always has been principally about democracies. . . . Moreover, decline within democracy is almost universally what scholars mean by "democratic backsliding." . . . Instead, the more common claim is that democratic quality is declining in the world's democracies. Consider the three examples that L&M (2023) cite in their introduction as emblematic of scholarly "alarm" around democratic backsliding. Of these examples, the Haggard and Kaufman (2021) study is entirely about democracies, whereas Diamond (2015) and Lührmann and Lindberg (2019) both emphasize that the global decline in democratic quality is minor but nonetheless worrisome because of its concentration in democracies. The democracy sample is defined [by Miller] using the Boix, Miller, and Rosato (2013) dataset [which has a dichotomous conceptualization of democracy]. (Miller 2024, 1–2)

237. One of the most important guidelines in chapter 3 of the concept book is that one should construct ideal type concepts and measures. This plays a critical role in analyzing Little and Meng's new measure of democracy. To begin with they use what they call a "quasi-minimalist" approach to democracy. More important are the scales for each of the indicators. As illustrated with the polity data in chapter 3 it is possible that the scales-conceptualization stops too soon or too low. This means that if many of the cases fall into the

final top category that category is probably not capturing a lot of variation among those countries. In the case of polity this meant that if countries varied significantly from 10 to 15 on an ideal type polity scale: all of those from 10–15 score 10 on polity so important variation is not captured. This is critical in backsliding which means typically a modest reduction in the democracy score. So if the indicator question stops too soon it will not be able to capture significant backsliding, or for that matter increases in democratic quality. In their critique of Little-Meng, Knutson forcefully make this point: "First, across these indicators, the category/value with the greatest spread in democracy values is the top category/value. This demonstrates that these indicators have a very low threshold for considering a country as highly democratic. Would-be democratic backsliders thus have much room for deterioration before they receive a lower score" (Knutsen et al. 2024, 10).

238. In measurement with multiple dimensions, components or indicators *scaling* is a critical, but not often recognized factor. To combine or aggregate a bunch of indicators typically means getting them all on the same scale. In Little and Mang they develop new concept-measure of democracy with a variety of indicators given in table **??**. In this table type indicates the kind of scale for each of these indicators. The note at the bottom of the table says that they are all rescaled in [0,1]. For example this means that dichotomous indicators are rescaled as interval with values between zero and one. Seven-point scales are linearly scaled into the same interval. Take one or two examples and discuss what is at stake by this linear rescaling of the variables. Note that this is required because they are going to take the average of these indicators and for the average to be valid requires interval data.

    Knutsen et al. 2024 make the point that scale is not linear at all for some indicators: "Second, the ordered categorical variables that L&M use are not interpretable as interval-level values; the categories within each variable scale differently onto the democracy concept. Figure 7 illustrates this problem conceptually, using the codebook entry for DPI's legislative index of electoral competitiveness. The first five categories all relate to highly uncompetitive situations in which opposition parties won no seats; only the top two categories correspond to remotely competitive situations. (Knutsen et al. 2024, 10)

Table 13: Little-Meng indicators of democracy

Table 1

Indicators in L&M's Index

	Source	Concept	Type	%Missing
Proportion Suffrage	V-Dem	Suffrage	Proportion	0%
Presidential Vote	DPI	Exec. comp.	Proportion	64%
Incumbent Party Seat	DPI	Legis. comp.	Proportion	23%
Incumbent Party in Office	DPI	Exec. comp.	Count (max. 20)	33%
Legislative Competitiveness	DPI	Legis. comp.	7-pt scale	18%
Executive Competitiveness	DPI	Exec. comp.	7-pt scale	18%
Incumbent Party Loss	NELDA	Elec. comp.	Dichotomous	16%
Multiparty Competition	L&M/NELDA	Elec. comp.	4-pt scale	6%
Process Violations	L&M/NELDA	Elec. comp.	4-pt scale	6%
Term Limits	Meng	Exec. constraints	Dichotomous	22%
Succession Rules	Meng	Exec. constraints	Dichotomous	24%
Dismissal Rules	Meng	Exec. constraints	Dichotomous	26%

Note: All variables are rescaled 0-1 for the index.

239. Aggregation plays an absolutely central role in conceptualization and measurement. Little and Meng (2024 see the symposium in *PS* regarding their article) have proposed a new measure of democracy and democratic backsliding. Aggregation plays two central roles at least in their project because they are interested in democratic backsliding at the global level which involves aggregation of national democracy scores as well as aggregation within national democracy scores based on 15 indicators. They use the average for both of these without any particular discussion probably, a resort to the standard default: "Taking Stock. To summarize the objective indicators and compare trends among different sets of countries, we construct a simple aggregate objective index by normalizing all individual variables between 0 and 1 when we can do so and taking the average for each country-year" (Little and Meng 2024, 6). Consider one of their key claims which is that backsliding at the global level has not really occurred over the last 10 years in spite of numerous claims to the contrary. Basically backsliding is defined as a decline in the democracy score forgiven country (See other exercises regarding their measure and backsliding in general). At the global level think about hypothetical examples where the value of zero backsliding at the global level could mean (1) significant backsliding or could mean (2) no backsliding at all. Hint: Remember that the average is an aggregation procedure where positive values can compensate for negative values (See Knutsen et al. 2024 in their contribution to the symposium for a discussion of the use of the average).

    Answer: Consider a bimodal world where half the countries experienced significant backsliding and the other half experienced significant increases in democracy: the average is zero. Contrast that with a world where there is essentially no change in democracy levels at all: the average is zero.

240. There are innumerable conceptualizations and measures of democracy but an explicit and separate conceptualization of authoritarianism is very rare. Almost always in theory, as well as in practice authoritarianism, is defined as not-democracy. A core guideline in Goertz (2020) is that if something is a core concept to a research project should be conceptualized on its own and

not via a "not." Gerschewski (2023) is a rare example of an author explicitly defining authoritarianism. Discuss his three dimensions of autocratic rule: (1) ideational performance, (2) repression, and (3) cooptation of elites, and the extent to which it is or is not a mirror image of democracy. Notice also his various three-level concepts of related ideas, e.g., repression (table 3.2). See, for example, his figure I.1.

241. The UCDP project has as its goal to collect data on all forms of "organized violence." It has a partial typology which they claim consists of three mutually exclusive categories. Is this along with other of the UCDP datasets an exhaustive categorization of organized violence? This is significantly more complex than one might imagine if one examines the definitions (https://www.pcr.uu.se/research/ucdp/definitions/). What might be missing from the typology? (Thanks to Peter Wallensteen for comments about this exercise.)

> Organized violence 1989–2017. For the third year running, the annual update from UCDP presents trends in not only state-based armed conflict, but also non-state conflict and one-sided violence. The three categories are mutually exclusive and can be aggregated as 'organized violence'. They also share the same intensity cut-off for inclusion – 25 fatalities in a calendar year. State-based armed conflict includes violence where at least one of the parties is the government of a state, that is, violence between two states and violence between the government and a rebel group. An example of the former is the border conflict between Ethiopia and Eritrea, while the conflict between the Taliban and the Afghan government is an example of the latter. Non-state conflict, on the other hand, is the use of armed force between two organized groups, such as rebel groups or ethnic groups, neither of which is the government of a state. Examples include fighting between the Islamic State (IS) and Tahrir al-Sham in Syria, as well as the interethnic fighting between the nomadic pastoralist Fulani and the mainly agriculturalist Mambila in Nigeria. Finally, one-sided violence covers violence by the government of a state or by a formally organized group targeting unarmed civilians. Recent examples include the CPI-Maoist targeting civilians they consider enemies of their struggle, and the government of Kenya killing protestors following elections. (Pettersson and Eck, 2018, 535)

242. It is quite common in conceptualization to give a series of necessary conditions, the same is often true for coding rules for quantitative datasets. What is often not stated – but is often implicit – is whether these necessary conditions are jointly sufficient. Are Fazal's three necessary conditions jointly sufficient for declaration of war?

> International legal scholarship suggests several necessary criteria for declarations of war. First, a declaration of war must be a public proclamation. Second, it must state intent to engage in hostilities with another party. Third, it must be issued according to the laws of the declaring state; it can only be a declaration of war if authored by an actor empowered with the authority to declare war. (Fazal 2012, 562)

243. One of the fundamental issues in concept analysis which is closely related to causal inference are the twin issues of conceptual and causal asymmetry. In

statistical work these are virtually always assumed to go together. But many would claim that for example peace is not the absence of war and that war is not the absence of peace. This connects to the causal asymmetry issue whether the causes of peace are symmetric to the causes of war. While this is typically implicit Blainey in his classic work on war is very explicit about this:

Discussed the extent to which his conceptualization of war–peace is symmetric and the extent to which his explanations of them are also symmetric

> This book is based on a survey of all the international wars fought since 1700. It argues that in war and peace there are revealing patterns and clues that have been overlooked. As the book is more like an intellectual detective story than a narrative history, the preface may if so desired be ignored now and read as an epilogue. . . . The first part of the book points to weaknesses in well known explanations of peace. . . . The second part of the book examines ingredients which are usually prominent in determining a nation's decision to fight or not to fight; and the overall influence of those ingredients is summarised in the chapter called 'The Abacus of Power'. . . . The last part of the book offers the radical conclusion that the beginning of wars, the prolonging of wars, the ending of wars and the prolonging or shortening of periods of peace all share the same causal framework. The same explanatory framework and the same factors are vital in understanding each stage in the sequel of war and peace. . . . At first sight the suggestion may seem absurd but this is simply another way of saying that the transition from war to peace is essentially the reverse of the transition from peace to war. . . . For every thousand pages published on the causes of wars there is less than one page directly on the causes of peace. And yet the causes of war and peace, logically, should dove tail into one another. A weak explanation of why Europe was at peace will lead to a weak explanation of why Europe was at war. A valid diagnosis of war will be reflected in a valid diagnoses of peace. (Blainey 1988, ix, 3)

244. As a general rule quantitative scholars are not very interested in the relationships between their various independent variables. However, from a conceptual and causal inference point of view there are important issues which often arise. A classic one is when two independent variables are somehow part of the same concept. Barnhart et al. (2020) illustrate this where they have two core variables one which is democracy, using polity data, and women's suffrage. Arguably, see Paxton's work, women's suffrage is part of the concept of democracy. However, many traditional datasets do not include that factor, including polity and others. They make strong claims about the role of women's suffrage in connection with international conflict and more specifically the democratic peace literature:

> Finally, we examine cross-national conflict data to assess the impact of the enfranchisement of women. We find that female suffrage, when coupled with democratic institutions, is not only an important and powerful cause of peace, but that it may even be necessary for the democratic peace. Without female suffrage, democracies (whether examined dyadically or monadically) do not show clear evidence of being more pacific than non-democracies. These findings hold across time periods and do not appear to result from other confounding factors, or from the expansion of the

> political franchise more generally. By bridging the individual and international levels to show how the gender gap on the use of force matters, we contribute to the growing positivist study of gender and international relations. (Barnhardt et al. 2020, 634)

Someone might argue that it is not women's suffrage that is really the core causal factor but rather *quality of democracy*. Countries with women's suffrage, or more generally, fewer restrictions to suffrage based on race or whatever are higher-quality democracies, *by definition*. So maybe the article should be called the higher-quality democracy peace? (See Goertz and Dul 2023 for an argument and data analysis to this effect).

245. Critical to the study of international relations, particularly quantitative approaches is the definition and identification of the states of the international system. The COW project (Russett et al. (1968) along with the related Gleditsch and Ward (1999) provide approaches that have defined states. However, there is still much conceptual and hence operational confusion over how states are to be defined. This can be critical for the study of war. For example, any wars between Latin American political units which are not officially states does not appear in the canonical war data set. These then become the so-called extrasystemic wars. See the next exercise for discussion of this particular concept.

    Overall the COW system might be called the recognition system for states. There is a set of conditions any of which is sufficient to be considered a member of the international system. The three big ones are recognition by United Kingdom and France in the 19th century, membership in the League of Nations, or membership in the United Nations. So one conceptual question is whether there are other sufficient recognition criteria for membership? Then COW has a set of what one might called the veto criterion for membership. One conceptual question is whether one actually needs any veto criterion at all? The main veto criteria consist of:

    - Occupations – COW is inconsistent about whether an occupation kicks the state out of the system. Apparently short occupations are OK, for example, the US in Iraq or Afghanistan. But in general occupations are treated inconsistently, for example see German occupations during World War II.
    - Major power vetoes – in the 19th century United Kingdom or France are veto players in the sense that if they do not recognize a country it does not make it into the system. In the post-World War II period the question is whether major powers can veto membership of individual states such as United States–Palestine, Russia–Kosovo, China–Taiwan.
    - Population – COW is inconsistent about the minimum population required for membership. This is applied typically in the 19th century, but there are obvious members of the United Nations that are below this population minimum. This is a veto criterion.

    Another set of questions deals with dates particularly in the setting of post-colonial or secession states. When do they enter the system? Canada is an

interesting case where Gleditsch and Ward have it gaining independence in 1867 while COW has it entering the system in 1920. So the conceptual question is whether one should date the country based on independence or recognition criteria?

Finally, the core conceptual question is whether one could construct a set of criteria for membership in the international system based purely on multiple recognition criteria, with no veto criteria.

See the exercise below with the same issue arises in defining arms groups in the context of civil war conflict.

246. One large conceptual problem lies with concepts which are defined by negation (Ragin 2023 has a very nice discussion of this). Often the positive concept is clearly defined but the negation of it is a large heterogeneous set of cases. The COW concept of extrasystemic wars falls into this problem. In general the solution involves giving a positive characterization of these negative cases.

    Extrasystemic wars involve wars between "political units" one of which is not a member of the international system. For example, wars with a system member against a non-system actor such as colonial wars and wars against indigenous peoples.

    So the conceptual question is in some ways exactly the same as for the international system: what constitutes a political unit that can be involved in an extrasystemic war? These can be kingdoms, as in Africa, indigenous groups and tribes, princely kingdoms and the like. It is a tremendous conceptual problem trying to delimit this universe: so does it make any sense at all to talk about extrasystemic wars?

247. Those doing domestic armed conflict have the same problem as those defining states in the international system. How to define organizations or groups that are involved in militarized conflict, either with each other or with the state?

    A recent major data set illustrates the various issues involving in conceptualizing arm actors at the domestic level.

    > The GTD includes attacks that involve "the threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation" (START 2021).... Defining Armed Groups. We define an armed group as an independent organization of non-state actors that justifies the use of violence in the pursuit of political control.... We then use these case studies to exclude approximately 50 percent of candidate groups (1,009 entries) for not meeting the definition of an armed group. In practice, this removes criminal actors, cartels, individuals, pro-government militias, factions, duplicate aliases, and generic names listed in the GTD.... The "pursuit of political control" criteria require evidence of an underlying political cleavage or issue incompatibility. A group must aim to change a policy, regime, or piece of territory through violent means. This decision rule serves two purposes. Because "a terrorist act by definition must have a direct political goal this means that the GTD excludes ordinary criminal violence." Congruently, we exclude criminal actors who perpetrate violence for material rather than political ends.... Table 1. Excluded cases

from candidate entries. An armed group may be excluded for multiple reasons (Malone 2022)

In contrast to exercise above about the international system the main issue is not about sufficiency criteria, but veto criteria that eliminate possible actors. This comes across clearly in the table about exclusion criteria.

The second issue is gray zones: all concepts have significant gray zones.

There are at least three gray zones that make it hard to draw a clear line about inclusion or exclusion: (1) The first is the fact that there is a gray zone between armed groups and nonviolent groups. It is well known groups can shift back and forth in many campaigns have both violent and nonviolent dimensions. (2) The second big gray zone deals with what is often called criminal or economic armed groups. As literature on organized crime in Latin America shows clearly, these are massively important political actors. They are often the outcome of civil wars and organized crime is very involved in elections and the assassination of various elected officials. (3) Terrorism datasets typically require (wrongly I think) that the organizations be "non-state." But as is well know pro-government militias are an important source of terrorism, e.g., KKK in the US south.

The two conceptual questions mostly revolve around the list of veto criterion and how one deals with important gray zone questions. Discuss how Malone deals with these issues.

248. Chubb (2022) presents a typology of assertiveness in the context of Chinese militarized activities in East Asia, presented in table below. He presents it as a typology were different levels are *qualitatively different* from each other: "Having distinguished these four qualitatively different types of assertive state behavior in maritime disputes," (Chubb 2020, 88, emphasis is mine). Discuss whether there is in fact some underlying scale to the concept of the assertiveness. He hints there is by saying there is greater escalatory potential as one moves down the types, so they are ordered in that sense. Would it be more accurate to say these are levels of assertiveness. Could one draw a figure of his concept this using the basic framework with four defining dimensions and then multiple indicators for each dimension?

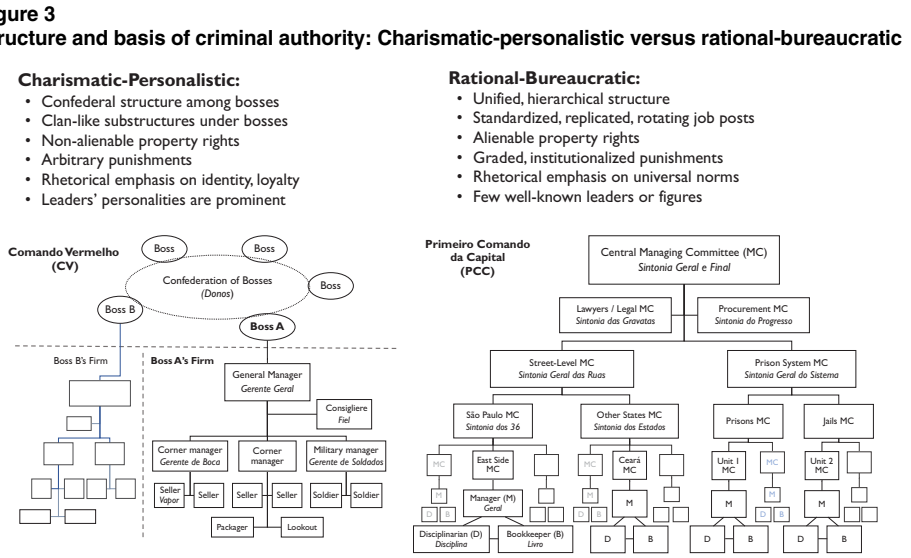
Table 14: Typologies and "qualitatively different": military assertiveness

Table 1. Four-Way Typology of Assertiveness in Maritime and Territorial Disputes, Mapped onto Existing Concepts

	Types of Assertiveness	Existing Concepts
↓ greater escalatory potential	Declarative verbal assertions via non-coercive statements, diplomatic notes, domestic legislation and administrative measures, international legal cases	gray-zone conflict (Green et al.; Peterson); delay (Fravel); minimal conflict (Huth)
	Demonstrative unilateral administration of disputed possession: patrols, surveys, resource development, construction of infrastructure, state-sanctioned tourism or activism, domestic judicial proceedings, and cooperative agreements with third parties	
	Coercive threat or imposition of punishment: may be verbal, diplomatic or administrative, economic punishment, warning shots, physical interference with foreign activities in disputed area	coercion (Schelling; Zhang); political-diplomatic escalation (Huth)
	Use of force application of military force or direct seizure and occupation of disputed possession	use of force / escalation (Fravel); brute force (Schelling); compulsion (Sechser)

249. "Criminal governance" is a huge issue in Latin American politics. Discuss Lessing's various conceptualizations of this concept, particularly within the basic framework of the concept book. As with all two word–concepts, one needs to define "criminal," how does Lessing deal with this? Is it a subset relationship with governance? Analysis his figure, below, using the basic framework for concept analysis.

Figure 49: Criminal governance



250. Concepts often contain causal claims as part of their definitions. This occurs all the time in law. Here is a standard definition of "Discriminatory Harassment" used by the university of Notre Dame. Discuss how it is structured, and also discuss how causal factors are part of the defining features (Thanks to Anibal Pérez-Liñán for this exercise):

> (1) unwelcome conduct
> (2) that is based on an individuals or groups race, color, national origin, ethnicity, religion, genetic information, age, disability, or veteran status and
> (3) that interferes with performance, limits participation in University activities, or creates an intimidating, hostile, or offensive University environment when viewed from the perspective of both the individual and a reasonable person in the same situation.

Answer: Constitutive terms (2) and (3) involve two causal statements (about the motivation for the conduct, and about the effects of the conduct).

251. It appears that the attraction of dichotomization is very strong. The V-Dem project has addressed this particular strong demand (Lührmann et al. 2017; Lindberg 2016). They have a dichotomization procedure which basically divides the V-Dem scale at .5 to separate democracies from non-democracies. They also introduce a five-level ordinal measure that breaks [0,1] into five equal parts. However, they end up adding some additional necessary condition criteria which makes this dichotomization scheme a hybrid, combining the basic linear V-Dem democracy scale with some necessary conditions. Does the addition of these additional necessary condition requirements for democracy imply a critique of the V-Dem measure in the sense that if the

basic measure were working correctly they would not need the additional necessary condition criteria?

> To qualify as a democracy, regimes have to fulfill at least a minimal level of the prerequisites of the electoral democracy as captured by the EDI. Based on Lindberg (2016: 90) we draw the line at a score of above 0.5 on the EDI and introduce two additional necessary conditions: (1) de-facto multiparty elections as indicated by a score above 2 on the V-Dem indicator for multiparty elections (v2elmulpar-osp; 10 and (2) elections have to be free and fair in sense of allowing at least for substantial competition and freedom of participation as indicated by a score above 2 on the V-Dem indicator for free and fair elections (v2elfrfair-osp). These two additional necessary criteria ensure that the electoral core of democracy is at least minimally achieved in all regimes classified as democracies. Compared to merely taking a cut-off point on the EDI, our choice of adding two additional criteria also helps to mitigate concerns that moving from continuous scales to categorical measures is a somewhat arbitrary decision (see for example Bogaards (2010)). We consciously decided not to take additional criteria from Dahl's list of prerequisites as necessary criteria in order to allow for weaknesses in one area to be balanced by strengths in another area. . . . We operationalize liberal democracies by the same criteria as electoral democracies, but they must additionally satisfy the liberal principles of respect for personal liberties, rule of law, and judicial as well as legislative constraints on the executive, as indicated by a score above 0.8 on the V-Dem Liberal Component Index (v2x liberal). Electoral autocracies fail to meet one or more of the above-mentioned criteria of electoral democracies, but subject the chief executive to elections at least a minimal level of multiparty competition as indicated by a score above 1 on the V-Dem multiparty elections indicator (v2elmulpar-osp). Closed autocracies do not satisfy the latter criterion. (Lührmann et al. 2017, 9–10)

252. One of the huge problems with the dichotomization and mutually exclusive typologies is the number of cases that lie on the gray zone between categories. V-Dem has a procedure for doing dichotomous measures (Lührmann et al. 2017). Using the analysis in the concept book, discuss how their analysis misses some key issues and problems, particularly notable is the fact that democracy is quite bimodal. Discuss what one should do with the cases that could fall on either side of a typology boundary (the confidence interval for the democracy value lies on the .50 boundary). Should one include a category "cannot tell" for those country-years? How does the increasing occurrence of democratic backsliding and competitive-authoritarian regimes, make this problem significantly worse?

> The level of ambiguity started increasing from around 1960 and has become worse during the third wave of democratization. By 2016, almost 30 percent of all countries are in one of the ambiguous categories while 12 percent fall in the critical grey zone between democracy and autocracy. . . . Table 2 compares the RIW measure to the other key measures used in the literature. The third column shows that the rate of agreement is relatively high, varying between 88.5 percent (CGV) and 93.1 percent (WTH). Excluding cases that our typology qualifies as ambiguous, the level of agreement varies between 91.7 percent (CGV) and 93.5 percent (GWF). (Lührmann et al. 2017, 13, 15)

253. It is extremely common in the literature on the concept of democracy to use the necessary condition concept structure. The very influential V-Dem project often uses the language necessary conditions such as in the quote below. In their figure 1 discuss the use of the + connecting the essential electoral democracy center to the spokes of the wheel. If electoral democracy were necessary condition what should one use instead of the +? Later in the book they give an additive formula to get the overall democracy score. Use the basic framework to do another version of figure **??**?

> The electoral principle has a special status in the V-Dem conceptual scheme as the *sine qua non* of democracy (see Figure 2.1). We would not want to call a political regime without multiparty elections "democratic".... Because we regard electoral democracy as foundational, the other democracy indices incorporate electoral democracy into their respective indices. Following the conceptual logic presented in his figure 2.1, below. The fact that all varieties of democracy are fairly closely related raises the question of whether we can move even one further level up in the tree of aggregation. Is it possible to speak of "democracy" tout court, or "Big-D," a combination of all five varieties of democracy for which V-Dem has collected empirical measures? One possible such measure of course readily suggests itself, namely to simply take the average across the five component indices: (Coppedge et al. 2020, 32, 39, 127)

Figure 50: V-Dem conceptualization of democracy and necessary condition structures

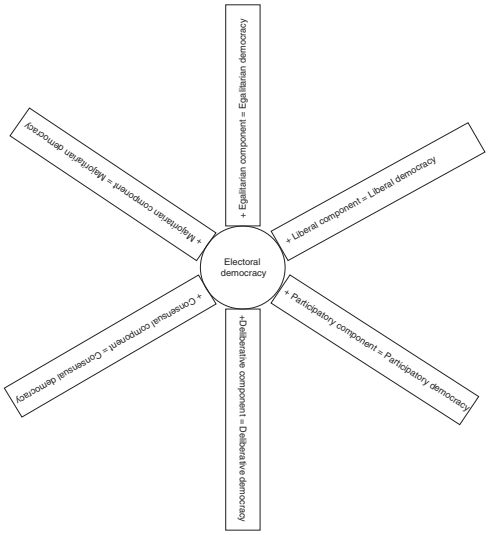


FIGURE 2.1 Principles and components of democracy.

Answer: If it were truly a necessary condition in figure **??** they should use the * to indicate multiplication instead of the +. One can redraw the figure using the basic framework having addition and multiplicaton at the secondary level. It is not intuitive to get from their figure 2.1 to a mean formulation.

254. Campbell et al. (2021) develop a concept and measure of "personal secularism" in their study of secularism in American politics. Draw the basic framework figure this three level concept. Assign the questions that are indicators to the various secondary level dimensions. Discuss the decision to use an additive substructure for both levels. It is the Redundancy Guideline relevant? Contrast the three level concept–measure with their latent variable statistical model.

> To create our Personal Secularism Index we have consulted the expansive body of writing espousing secular social and political thought, including the work of social theorists such as George Jacob Holyoake (1871), Immanuel Kant (1781 [1999]), and David Hume (1777 [2014]); the recent volumes by "new atheists" such as Richard Dawkins (2006), Sam Harris (2005), and Christopher Hitchens (2007); and the statements of belief found on the websites of secular organizations such as the American Humanist Association (AHA) and the International Humanist and Ethical Union (IHEU). Based on these sources, we have distilled three core principles common among Secularists. One core principle is a commitment to science and objective evidence as the basis for understanding the world. For example, Kant contends that "everything in the world happens solely in accordance with laws of nature . . . we have nothing but nature in which we must seek the connection and order of occurrences in the world" (1999, 485). In Humanism and Its Aspirations, published in 2003, the AHA contends that "knowledge of the world is derived by observation, experimentation, and rational analysis.". . . . A second core principle of secularism is the view that only human experience and knowledge provide the proper basis for comprehending reality and making ethical judgments – in other words, "humanism." . . . . A third core principle of secularism is "freethought," the idea that human development and understanding should be based on logic and reason, rather than received authority, dogma, or tradition.
>
> We measure secular beliefs with a series of questions that gauge support for these core secular principles: the degree to which a respondent's perspective is informed by sources other than the supernatural (a term we use nonpejoratively). The core of our secular beliefs scale consists of eight statements, to which our respondents indicated their level of agreement.. . . . To minimize response set bias, the questions were not all worded in the same direction. Five of the statements are worded to affirm secular perspectives: (1) Factual evidence from the natural world is the source of true beliefs, (2) The great works of philosophy and science are the best source of truth, wisdom, and ethics (3) To understand the world, we must free our minds from old traditions and beliefs (4) When I make important decisions in my life, I rely mostly on reason and evidence (5) All of the greatest advances for humanity have come from science and technology The other three statements represent the rejection of secular values: (6) It is hard to live a good life based on reason and facts alone (7) What we believe is right and wrong cannot be based only on human knowledge (8) The world would be a better place if we relied less on science and technology to solve our problems. . . . While the scales of nonreligiosity and personal secularism created using confirmatory factor analysis are the most methodologically sound, they correlate highly with simple additive indices (over .95). Therefore, in most cases we opt to use the additive indices, although in every case the results would be nearly

> identical with the scales created from confirmatory factor scores. The exception is in Chapter 5, where we examine how nonreligiosity, personal secularism, and a variety of political orientations are related to one another over time. For the models in that chapter, we employ the measurement error correction afforded by confirmatory factor analysis. If other scholars use our measures of personal secularism – which we hope they will – an additive index will suffice for nearly all purposes. (Campbell et al. 2021, 27–28, 36)

255. Campbell et al. (2021) develop a concept and measure of "personal secularism" in their study of secularism in American politics. Draw the basic framework figure this three level concept. Assign the questions that are indicators to the various secondary level dimensions. Discuss the decision to use an additive substructure for both levels. It is the Redundancy Guideline relevant?

> To create our Personal Secularism Index we have consulted the expansive body of writing espousing secular social and political thought, including the work of social theorists such as George Jacob Holyoake (1871), Immanuel Kant (1781 [1999]), and David Hume (1777 [2014]); the recent volumes by "new atheists" such as Richard Dawkins (2006), Sam Harris (2005), and Christopher Hitchens (2007); and the statements of belief found on the websites of secular organizations such as the American Humanist Association (AHA) and the International Humanist and Ethical Union (IHEU). Based on these sources, we have distilled three core principles common among Secularists. One core principle is a commitment to science and objective evidence as the basis for understanding the world. For example, Kant contends that "everything in the world happens solely in accordance with laws of nature . . . we have nothing but nature in which we must seek the connection and order of occurrences in the world" (1999, 485). In Humanism and Its Aspirations, published in 2003, the AHA contends that "knowledge of the world is derived by observation, experimentation, and rational analysis.". . . . A second core principle of secularism is the view that only human experience and knowledge provide the proper basis for comprehending reality and making ethical judgments – in other words, "humanism." . . . . A third core principle of secularism is "freethought," the idea that human development and understanding should be based on logic and reason, rather than received authority, dogma, or tradition.
>
> We measure secular beliefs with a series of questions that gauge support for these core secular principles: the degree to which a respondent's perspective is informed by sources other than the supernatural (a term we use nonpejoratively). The core of our secular beliefs scale consists of eight statements, to which our respondents indicated their level of agreement.. . . . To minimize response set bias, the questions were not all worded in the same direction. Five of the statements are worded to affirm secular perspectives: (1) Factual evidence from the natural world is the source of true beliefs, (2) The great works of philosophy and science are the best source of truth, wisdom, and ethics (3) To understand the world, we must free our minds from old traditions and beliefs (4) When I make important decisions in my life, I rely mostly on reason and evidence (5) All of the greatest advances for humanity have come from science and technology The other three statements represent the rejection of secular values: (6) It is hard to live a good life based on reason and facts alone

> (7) What we believe is right and wrong cannot be based only on human knowledge (8) The world would be a better place if we relied less on science and technology to solve our problems. . . . While the scales of nonreligiosity and personal secularism created using confirmatory factor analysis are the most methodologically sound, they correlate highly with simple additive indices (over .95). Therefore, in most cases we opt to use the additive indices, although in every case the results would be nearly identical with the scales created from confirmatory factor scores. The exception is in Chapter 5, where we examine how nonreligiosity, personal secularism, and a variety of political orientations are related to one another over time. For the models in that chapter, we employ the measurement error correction afforded by confirmatory factor analysis. If other scholars use our measures of personal secularism – which we hope they will – an additive index will suffice for nearly all purposes. (Campbell et al. 2021, 27–28, 36)

256. In general it is not recommended to define via negation. This typically raises the serious possibility of significant heterogeneity in the "not" population. Discuss this issue using the concept of "noncontrolled comparison" which is the central topic of *Rethinking comparison: innovative methods for qualitative political research* (Simmons and Smith 2021) The authors of the anthology are mostly interpretivist but there might be other kinds of methodologies which are noncontrolled but which are also not interpretivist. See the question in the LNQA section, page 30, which raises this possibility.

257. A very difficult design decision in constructing concepts is whether to include potential causes or effects of the concept in the concept itself. Diamond and Morlino (2004, 2005) discuss various dimensions of the "quality" of democracy. Some of these might be better thought of as potential causes or effects of democracy. For example, they discuss how equality or education is important in the concept of a high-quality democracy. Discuss the various advantages and disadvantages of including dimensions such as equality or education in the concept of (quality of) democracy. Similarly, one can think about the effects of democracy: do Diamond and Morlino include effects in the concept of quality of democracy that would be independent variables in important analyses where democracy is a key dependent variable?

    Answer:

    One problem with the Diamond and Morlino book is that it is purely about the quality of democracy. There is no sense in which this concept is embedded in some theoretical or empirical program looking at the causes or effects of democracy. If one wants to discuss the causes of democracy (or quality of democracy) one should think hard before including potential independent variables in the concept itself. Also, they define a quality democracy as one with effective rule of law and effective administrative outputs. To what extent does one want to define democracy quality in terms of quality of bureaucracy? Would it be better to keep these as separate concepts? This way we can explore the causal relationships between the two.

258. Discuss how it is possible that Moyn (2012) can survey the historiography of human rights without talking about democracy at all, i.e., the word does not

appear in the review article. What does the absence of democracy in the discussion mean about the concept of human rights guiding Moyn's discussion?

259. There is often overlap between closely connected concepts. Should one worry about this or attempt to make them separate? For example, overlap occurs regularly between various conflict concepts such as coups and civil wars:

> While we have gone to great lengths to assure that coups are not conflated with other forms of anti-regime activity in our dataset, there is no guarantee that scholars coding other events have excluded coups in theirs. For example, we have identified 38 events within the Uppsala/PRIO dataset and 4 events in the Correlates of War (Sarkees, 2000) internal conflict dataset that are best described as coups. Scholars who use these datasets to operationalize civil wars should be wary of including these events. (Powell and Thyne 2011, 256).

260. Negations or negative poles often are tricky, particularly when there is concept asymmetry. Discuss the claim at the end of the following quote about two "direct opposites," particularly in the light of the claim that populism has three core dimensions:

> Beyond the lack of scholarly agreement on the defining attributes of populism, agreement is general that all forms of populism include some kind of appeal to "the people" and a denunciation of "the elite." Accordingly, it is not overly contentious to state that populism always involves a critique of the establishment and an adulation of the common people. More concretely, we define populism as a *thin-centered ideology that considers society to be ultimately separated into two homogeneous and antagonistic camps, "the pure people" versus "the corrupt elite," and which argues that politics should be an expression of the volont gnrale (general will) of the people*.... [T]here are at least two direct opposites of populism: elitism and pluralism.... Populism has three core concepts: the people, the elite, and the general will. (Mudde and Kaltwasser 2017, epub 38, 44)

261. Concepts are frequently asymmetric, e.g., democracy–autocracy. These is true for the concept pair healthy–ill. Discuss this in the context of defining health as the absence of disease:

> In a series of essays published in the 1970s, Christopher Boorse developed a powerful analysis of the theoretical notion of health – that is, the concept of health that is of concern to pathologists and physiologists. In his original essays, Boorse took health to be the absence of "disease," where disease is understood in a wide sense that includes injuries and disabilities. To avoid confusion, he has changed his terminology (but not his view) and now takes health to be the absence of pathology. In Boorse's view, whether a physical or mental state constitutes a pathology does not depend on judgments about how good that state is for people. Boorse defends a naturalistic, nonnormative view of pathology and health.... There is a large literature concerning the concept of health, most of which is critical of Boorse's view. (Hausman 2015, 8)

262. Use the basic concept framework to simplify this complex weighting system.

> Most systems of health measurement define health states in terms of a small number of easily observable levels, along some small number of dimensions. The dimensions have been functional capacities, such as vision or cognition; consequences of multiple physical and mental capacities, such as "self-care"; and subjective states, such as pain or anxiety. So, for example, the Health Utilities Index, Mark 3 (HUI(3)) has eight dimensions: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain. Even with only eight dimensions, if one allows several gradations along each dimension, the number of distinguishable health states becomes very large. The HUI(3) distinguishes five or six levels along each dimension for a total of 972,000 health states. . . . Once health economists have a classification of health states, such as the EQ-5D or the HUI(3), they need to assign numbers to health states that measure the value of the health state and (once one takes into account time) of health. Health economists call these numbers "quality weights." To quantify health states requires assigning quality weights to the health state vectors defined by a health-state classification system or to some set of sequelae of diseases and injuries. The health classification system determines the magnitude of the measurement task. . . . the HUI(3) requires assigning quality weights to 972,000 health states. (Hausman 2015, 36–37)

Answer:

Instead of considering five or six levels for each secondary level dimension create a semantic transformation for each dimension and then decide the weights for each of the five or six secondary level dimensions.

263. The Redundancy Guideline stresses that redundancy can be problematic between concepts particularly closely related concepts. Discuss this solution to the problem of redundancy or conceptual overlaps between transparency and democracy. This arises because by definition good democracies are transparent. Discuss also the relationship between tautology and redundancy.

> The study of transparency and its relationship with democracy is fraught with difficulties, one of which is the possibility of tautology. While we are concerned in our study with the transparency of policymaking and, more specifically, with the dissemination of information, the broad concept of transparency also applies to a full range factors that affect information flow within a society. Transparency thus pertains to questions of who rules, how governments might be replaced, and, indeed, how elections are contested in countries that fill key offices through an electoral process. In other words, transparency may pertain to the very question of whether one can call a political regime "democratic." . . . The analytical tension is obvious: if democracy is, by definition, transparent, then the degree to which governments enjoy and promote obfuscation is simply a measure of their nondemocraticness. . . . We thus employ a minimalist definition of democracy. Following Schumpeter (1942) and more recently Przeworski et al. (2000), we define democracy as a regime in which the executive and the legislature are both filled by "contested elections." . . . . We therefore employ a similarly narrow definition of "transparency." . . . We thus explicitly distinguish the transparency of the electoral system from policy transparency (Hollyer et al. 2011, 1192–93; redundancy solution and conceptual tautology)

264. Discuss the claim about bounded versus unbounded scales in Morris's *The measure of civilization how social development decides the fate of nations*. How does this relate to the Ideal Type Guideline. Clearly fuzzy logic semantic transformation transformations always have a maximum 1.0. How serious critique of that approach is this?

> It [HDI] could, up to a point, be used to measure change through time by simply comparing a single country's score in each annual report, but because the maximum possible score is always 1.0, the HDI does better at charting a nation's relative position within the world at a single point in time than at measuring diachronic changes in development levels. (Morris 2013, 28)

265. Discuss this (dubious) claim in a standard textbook on measurement in medicine:

> Only unobservable constructs require a measurement theory. For observable characteristics, it is usually obvious how the items contribute to the construct being measured and no measurement theory is required. We illustrate this with a few examples. Physical activity can be characterized by frequency, type of activity and intensity. To obtain the total energy expenditure we know how to combine these items. Moreover, for some research questions we are only interested in certain types of physical activity or only in the frequency of physical activity. To assess the severity of diarrhea, a clear example of an observable characteristic, fecal output can be characterized by frequency, amount and consistency. Another example concerns comorbidity, which is characterized by the number of accompanying diseases, the type of diseases or organ systems involved, and the disease severity or the disability or burden they cause. However, if we talk about comorbidity burden, we move in the direction of unobservable constructs. (Vet et al. 2011, 17)

266. The Ideal Type Guideline is a central one for concept construction. This can mean "conceptual creep" as the conceptual scale expands. Discuss the implications, causal and conceptual of such creep as well as the intension–extension implications of such creep. Why might having a clear semantic transformation be critical in this process?

> As it turns out, abstract concepts can creep, too. For example, in 1960, Webster's dictionary defined "aggression" as "an unprovoked attack or invasion," but today that concept can include behaviors such as making insufficient eye contact or asking people where they are from (1). Many other concepts, such as abuse, bullying, mental disorder, trauma, addiction, and prejudice, have expanded of late as well (2). Some take these expansions as signs of political correctness and others as signs of social awakening. We take no position on whether these expansions are good or bad. Rather, we seek to understand what makes them happen. Why do concepts creep? (Levari et al. 2018, 465)

267. Draw the figure that corresponds to this concept of ' international court authority" and discuss the dimensions and structure.

> Our measure of de facto IC [international court] authority has two key components: (1) recognizing an obligation to comply with court rulings

> and (2) engaging in meaningful action pushing toward giving full effect to those rulings. This is a conjunctive standard that is assessed by examining the practices of the relevant actors. A simple public statement that a judgment is legally binding is, without more, inadequate. Equally insufficient is conduct that happens to conform to a judgment. . . . Our framework identifies five types of de facto authority that correspond to the practices of different IC audiences, which we illustrate with examples from the empirical chapters of this book. . . . Figure 2.1 displays the three levels of de facto authority as nested, suggesting that a court first gains narrow, then intermediate, then extensive authority. But as we explain in the discussion that follows, we intend no teleology by this diagram. The three circles of de facto authority may not be nested; an IC could have extensive authority but lack narrow and intermediate authority. (Alter et al. 2018, 29, 31, 33, 34)

268. Do classic Mafia protection rackets count as terrorism?

> Terrorism is the premeditated use or threat of use of extranormal violence or brutality by subnational groups to obtain a political, religious, or ideological objective through intimidation of a huge audience, usually not directly involved with the policy making that the terrorists seek to influence. (Enders and Sandler 2002, 145–46)

Answer:

Because the Mafia has an economic goal it is not terrorism, because it is not "political, religious, or ideological objective." If we added " economic" to the list (sometimes it does appear in definitions of terrorism) then the Mafia would be a terrorist group. Would it be better to define terrorism in terms of "non-economic" goals?

269. The Redundancy Guideline raises the issue of *semantically* redundant versus *empirically correlated*. Discuss this issue in the context of HDI and similar indicators, for example:

> On the other hand, redundancy in the three dimensions of the HDI was also pointed out right after its introduction. In this respect, McGillivray (1991) disclosed a significant and positive correlation between each dimension index. While supporting this finding, Ivanova et al. (1999) further demonstrated that the overall HDI rankings did not change significantly even if measured either by life expectancy only or by the combination of the other indicators. This point was echoed later by Cahill (2005) who thus reached the conclusion that the HDI is not sensitive to the weights employed. By contrast, McGillivray and White (1993) found redundancy in all countries but not within the same group of countries. This result was also proved by Noorbakhsh (1998c). While interpreting the high correlation in all countries as a sign of internal consistency rather than redundancy, Kovacevic (2011) further revealed that the lower two groups in the HDI category differed between themselves and from the upper two in terms of the correlation structure. (Hirai 2017, 85)

270. A key issue in the three-level framework is whether the same data-indicators can be used for different secondary-level dimensions. See how this occurs in "Measuring the rule of law": "A careful examination of the nine factors

reveals two facts. The first is that there is a partial overlap among sub-factors; that is, a sub-factor can simultaneously belong to different factors at once. This is simply to reflect the fact that various rule-of-law dimensions partially overlap in practice" (Botero and Ponce 2011, 16).

271. See the same issue of the same data-indicators in multiple dimensions in Dahl's (1971) famous conceptualization of democracy.

I. Formulate preferences:

A.Freedom to form and join organizations

B. Freedom of expression

C. Right to vote

D. Right of political leaders to compete for support

E. Alternative sources of information

II. Signify preferences

A. Freedom to form and join organizations

B. Freedom of expression

C. Right to vote

D. Eligibility for public office

E. Right of political leaders to compete for support

F. Alternative sources of information

G. Free and fair elections

III. Have preferences weighted equally in conduct of government

A. Freedom to form and join organizations

B. Freedom of expression

C. Right to vote

D. Eligibility for public office

E. Right of political leaders to compete for support

F. Alternative sources of information

G. Free and fair elections

H. Institutions for making government policies depend on votes and other expressions of preference

272. Race is notoriously a problematic concept. Discuss the history and methodology of determining race in the US Census. What are the normative implications of this? What is the aggregation principle behind the "one-drop rule"?

> The constructed, contradictory, and fickle nature of US racial classification is exemplified in the revisions of census categories. Since its inception in 1790, the decennial US census has counted the population by race/color. For much of the censuss history, race was assigned to individuals by an enumerator, characterized as unchangeable, and race mixture was ignored

> entirely. Yet, between 1850 and 1920, racial fluidity was acknowledged, and periodically enumerated with specificity, via fractional mixed-race categories: mulatto (half-black), quadroon (one-quarter black), and octoroon (one-eighth black). By 1930, these categories were removed, and mixed-race blacks, as well as individuals of mixed white-Asian parentage, were subjected to the one-drop rule and categorized with their minority race. (Davenport 2020, 223)

273. The well-known UCDP data projects have the goal of basically mapping all forms of armed conflict, both domestic and international. Discuss the extent to which their typology is mutually exclusive and exhaustive. How is this an application of the Redundancy Guideline? Why is this critical for core UCDP goals.

> Organized violence 1989–2017 For the third year running, the annual update from UCDP presents trends in not only state-based armed conflict, but also non-state conflict and one-sided violence. The three categories are mutually exclusive and can be aggregated as 'organized violence'. They also share the same intensity cut-off for inclusion – 25 fatalities in a calendar year. State-based armed conflict includes violence where at least one of the parties is the government of a state, that is, violence between two states and violence between the government and a rebel group. An example of the former is the border conflict between Ethiopia and Eritrea, while the conflict between the Taliban and the Afghan government is an example of the latter. Non-state conflict, on the other hand, is the use of armed force between two organized groups, such as rebel groups or ethnic groups, neither of which is the government of a state. Examples include fighting between the Islamic State (IS) and Tahrir al-Sham in Syria, as well as the interethnic fighting between the nomadic pastoralist Fulani and the mainly agriculturalist Mambila in Nigeria. Finally, one-sided violence covers violence by the government of a state or by a formally organized group targeting unarmed civilians. Recent examples include the CPI-Maoist targeting civilians they consider enemies of their struggle, and the government of Kenya killing protestors following elections. (Pettersson and Eck, 2018, 535; see https://www.pcr.uu.se/research/ucdp/definitions/)

Answer:

A core UCDP goal is to give an survey the amount of armed conflict in the world. This is essentially adding up all the various kinds of armed conflict into a total. Therefore the categories need to be mutually exclusive so there is no doublecounting. In fact, the addition of new datasets to the UCDP is probably working toward the goal of completeness. This might be signaled by the change in the name of the yearly data articles to "Organized violence" in 2016 " What we now call 'state-based armed conflict' was called just 'armed conflict' in our earlier publications." (Melander et al. 2016)

274. Compare Rawls and his typology of war with UCDP.

> Rawls distinguished between different types of wars in order to define the principles that would best apply in each case. The initial typology found in his lecture notes proposes nine kinds: 1. Wars between existing via states (WW I + II) 2. Civil wars (of social justice) within via states or society (French Rev); 3. Wars of secession of minorities within region: American

> Civil War. 4. Colonial Wars of secession (from Empire): Algerian War; American Rev War? 5. Wars of intervention (humane intervention) 6. Wars of national unification (War of Roses; Tudors) 7. Wars of conquest, of Empire (Wars of Rome). 8. Wars of Crusade, religious or secular 9. Wars of national liberation (in present sense); guerilla wars (Armitage 2017, 335)

275. One of the obvious problems with definitions and conceptualization is when the concept being defined also appears in the definition. This is not uncommon with various "peace" concepts, such as "peacebuilding." Barnett (2006) illustrate this:

> Although peacebuilding is generically defined as external interventions that are designed to prevent the eruption or return of armed conflict, there are critical differences among actors regarding its conceptualization and operationalization. This article surveys and analyzes twenty-four governmental and intergovernmental bodies that are currently active in peace building in order to, first, identify critical differences in how they conceptualize and operationalize their mandate and, second, map areas of potential concern. (Barnett et al. 2006, 36)

See Table 1 with definitions, almost all include peace in the definition.

276. Vabulas and Snidal (2013) define an "Informal IGO":"We thus define an IIGO as: 1. An explicitly shared expectation – rather than a formalized agreement – about purpose 2. With explicitly associated state "members" who 3. Participate in regular meetings but have no independent secretariat or other significant institutionalization such as a headquarters and/or permanent staff."

A core guideline in chapter 3 is to get a list of the defining features. This is closely related to coding rules for generating datasets. Discuss this list and how you might modify it to make it clear more complete. For example, a key feature is no permanent or independent secretariat. This should probably be an independent dimension. Later in the discussion it is clear that in fact there is usually some organizational basis for these informal organizations so that the key feature is in fact that it is not an independent secretariat. Also, one might want to explicitly require that there is no legally binding treaty that constitutes the organization, rather than a mere parenthetical remark. What is the implicit concept structure?

Answer:

Probably as with most datasets the concept structure is the necessary and sufficient condition one. Here is a reformulation:

(a) No legally-binding treaty or agreement that constitutes the organization.

(b) No *independent* secretariat or headquarters.

(c) Regular meetings.

(d) Shared expectation of purpose

(e) States form the core membership

(f) Nonstate actors can be involved (this is important since the COW IGO list does not allow nonstate actors to be members of IGOs)

(g) At least two members.

277. Legalism is an important concept in international international relations. What is the opposite of legalism? For McCall Smith (2000) according to table 1 it is diplomacy? Discuss this negative pole choice. What about "not-legalistic"?

278. It is possible to develop a definition or conceptualization based on cases that individuals or research communities identify as being exemplars of that concept. For example, one might develop a concept of populism based on countries, leaders or parties that are called populist in the literature. "Third, it uses the concept of political style to discern inductively the features of populism as a political style. This is done by examining twenty-eight cases of leaders from across the globe who are generally accepted as populists (that is, labelled as populist by at least six authors within the literature on populism), and identifying what links them in terms of political style." (Moffitt 2016, 27 epub)

Discuss how this might lead to "minimal conceptualizations." Which of Mill's methods is one basically using here?

Answer:

Basically one is using Mill's method of agreement to find the common features of all cases labeled as populist. Then implicitly one will probably assume that those necessary features are also jointly sufficient.

279. Discuss the use of the logical OR the following conceptualization of "extreme group." For example, consider extreme nonviolent action, e.g., Gandi.

> So persons or movements may be called extreme because their views are far out of the mainstream on some issue, *or* because they use violence to further their goals, *or* because they are rigid and intolerant of other points of view. A group can be extremist if it has only one of these features. Some movements – Al Qaeda is a good example – have all of them. Indeed, this is not surprising, because, as we will show, the latter two characteristics, the predilection for violence and the tendency toward rigidity and intolerance, can be derived from the first characteristic. (Wintrobe 2006, 6, my emphasis)

What about "extreme" left nonviolent campaigns? Is violence a requirement for an extreme group, or is ideology sufficient?

280. The Polity measure of democracy has a code of –77 for cases where there is no effective government, e.g., complete state failure. Some conceive of this as a problem of missing data and hence want to estimate the value. For example, sometimes these are given a value of 0 on the polity –10 to 10 scale. Discuss whether coding regime type requires that there actually be a government.

281. Call (2010) discusses the literature on the important and large concept of "failed state." He argues that there are three large secondary-level dimensions

to this concept conceptualized by three "gaps": (1) capacity gap, (2) security gap, and (3) legitimacy gap. He shows that these gaps are not very correlated: "Moreover, the three security gaps produce remarkably different rankings of countries of the world." (p. 309). He sees this as a negative finding for the concept of a failed state. Discuss the underlying correlational view of concepts that he is drawing on. How could this non-correlation be a positive thing from a definitional point of view?

Answer:

He works on the correlational or indicator view that all the dimensions–indicators of a concept should be correlated. A definitional view works on the notion that each definitional part should be different. Hence it is a positive feature that these dimensions are not in fact correlated, because they are in fact different dimensions of the failed state concept.

282. Fordham (2011) logs all independent and dependent variables. The key independent variable is size of country which is a major cause of major power status, the dependent variable. Explore what logging does when the data include both very small and very large countries.

283. The new "Democracy Barometer" (Bühlmann, M. et al. 2012) is a large project whose goal is to conceptualize and measure the "quality of democracy." It is a very complicated concept with at least 4 levels. It is an interesting example of a hybrid concept structure. Discuss where necessary condition ideas are used in the aggregation and when family resemblance is used.

284. One potentially large issue is what might be called "conceptual endogeneity." This is where a secondary-level dimension is part of the dependent variable and also considered as an independent variable. There has been an explosion of studies of elections in hybrid democratic regimes. Should these elections be considered a *cause* of democratization? Does this conflict with the fact that freer and more competitive elections will change the democracy score of the country by definition? For example, "Scholars also disagree about the role of elections in promoting democratic change. (p. 12) While a familiar indicator of democratic progress, we must remember, electoral turnover does not always support democratic development. (p. 13) These issues lead to a final and more fundamental concern about attributing too much democratic influence to elections" (Bunce and Wolchik 2011, 14). See also Gandhi and Lust-Oskar (2009) who discuss at length the causal effects of elections on authoritarianism.

285. Recently statisticians have become much more concerned with problems of "unit homogeneity." Here is Henry Brady defining the idea:

> We shall make the transformation of $Y_B(1,0)$ into $Y_A(0,0)$ in two steps which are depicted on Table 10.9. If A and B are identical and Z_A and Z_B [Z is the treatment] are identical as well(footnote: By saying that Z_A and Z_B have to be comparable, we mean that $Z_A = 0$ and $Z_B = 0$ are the same thing and $Z_A = 1$ and $Z_B = 1$ are the same thing.) (although we haven t indicated how this might be brought about yet) it might be reasonable to suppose that: $Y_B(1,0) = Y_A(0,1)$, [*Identicality of units and treatment or Unit*

> *Homogeneity*]. That is, A and B are mirror images of one another so that the impact of $Z_A = 1$ and $Z = 0$ on B is the same as the impact of $Z = 0$ and $Z = 1$ on A. (Brady 2008, 258)

Analyze how valid concepts are critical to the existence of unit homogeneity.

Answer:

Holland (1986) says "unit homogeneity" means that units are prepared carefully "so that they 'look' identical in all relevant aspects" (Holland 1986, 948). To be "identical" means that the concept must produce identical units, A and B in the Brady analysis. If a study looks at all "states" then the conceptualization of state means that there are no casually important differences in "stateness." Treatments, Z in the Brady analysis, also are assumed to be homogeneous.

286. Use the following quote to discuss the importance of predictive power in terms of evaluating a measurement model or concept. Contrast this with an approach which stresses the semantics and content of the test items.

> The g factor is an artifact of linear correlation analysis. A theorem of Suppes and Zanotti (1981) informs us that for any vector of test scores from an achievement test, it is possible to construct a scalar latent factor such that, conditional on the factor, test scores are independent. The g factor exists for any vector of binary, finite-valued, or countably valued random variables. The g of conventional psychometrics is a product of mathematical conventions in factor analysis. A g also exists to account for correlations among test scores. That is a mathematical theorem of no behavioral consequence for psychometrics or for finance, another field addicted to factor models. The value of g in predicting behavior is the real test of its importance. There is much evidence that it has predictive power. (Heckman 1995, 1105)

287. Cheibub et al. (2010) is a review and updating of the Alvarez et al. and Przeworski et al. well-known measure of democracy. They strongly critique the polity measure and the Freedom House for their aggregation schemes and the large number of ways the scores on the secondary-level dimensions produce the same value on the basic level:

> Regarding aggregation, for each of the ten categories in the political rights checklist and the 15 categories of the civil liberties checklist, coders assign ratings from zero to four and the points are added so that a country can obtain a maximum score of 40 in political rights and 60 in civil rights. With five alternatives for each of ten and 15 categories, there are 5^{10} = 9,765,625 possible ways to obtain a sum of scores between zero and 40 in political rights, and 5^{15} = 30,517,578,125 possible ways to obtain a sum of scores between zero and 60 in civil liberties. . . . In all of these cases, the aggregation rules are arbitrary. (Cheibub et al. 2010, 75)

While their dichotomous coding scheme with three "categories" does not produce the same huge option of possibilities, the same issue arises in their coding scheme. Explain how the aggregation issue arises in their measure. Are their aggregation rules "arbitrary?"

Answer:

While Cheibub et al. do not think of their coding rules as involving an aggregation procedure, they certainly do. They use the necessary and sufficient condition one. While there is only one way to be coded as a democracy, i.e., one on all the rules, there are multiple ways to be coded as a zero. Assuming four rules there are for example, (0,0,1,0), (1,1,0,0), (0,0,0,0), etc. ways to be a dictatorship. In short there are $2^4 - 1 = 15$ ways to be a "dictatorship."

288. Alvarez, Przeworski, Cheibub, Vreeland, and Gandhi who are all active working together on the concept and data on regime type have a clear preference for the term "dictatorship" as the opposite pole to democracy. In contrast, scholars using the polity data prefer the term "autocracy" or "authoritarian." Is there anything of theoretical or conceptual importance at stake in this terminological difference?

289. One of the best ways to find out about concepts is via the codebook for datasets. Sometimes reading the codebook provides some surprises about the concept (and hence the data) that most users are probably not aware of. An example of this phenomenon are the GTD terrorism data (CETIS 2007), the standard dataset for the study of terrorism. If one reads the GTD codebook the problematic nature of the concept is acknowledged in the introduction, but almost all the codebook is about the data. To find out what the GTD concept of terrorism actually is one must read an appendix! As an exercise I recommend that one read the codebook and outline what you think that concept of terrorism used is. I think you will then be surprised by the content of the appendix.

290. In the international relations there is a growing literature on what is called the "capitalist" peace (e.g., see the special issue of *International Interactions* 2010(2)). This literature usually has two goals, (1) so show that capitalist countries are less likely to have militarized conflicts, and (2) capitalism is more important than democracy in reducing conflict. Discuss what is the concept of "capitalism" used in these various studies. Then discuss what are the actual quantitative indicators used. How do they match up? Could you think of better indicators or tests? When left-wing, e.g., socialist or communist, critics discussed "capitalism" how was their meaning of capitalism different from the capitalist peace literature? Discuss McDonald (2010) and the various forms of the concept of capitalism from the beginning of the article and how they are implemented the quantitative analyses at the end.

    Answer:

    If you look at the Gartzke, Mousseau, and McDonald pieces (in the *International Interactions* special issue) "capitalism" is discussed in terms of (1) economic development, (2) (free) markets, (3) contract intensive economies, (4) integration into the global economy, (5) financial openness, (6) degree of property owned by government, (7) free trade and (8) government nontax revenues.

    Most of these are either too narrow (e.g., government nontax revenues, financial openness) or different (e.g., integration into world economy which is not directly about the domestic economy, e.g., oil-exporting states).

One might reasonably define capitalism as a "market-based" economy. There seem to be at least two reasonable ways one might go about operationalizing this. First, one might contrast communist/socialist countries as a dummy variable against all other types. Another way to think about this is to consider the amount of economic activity in a country that goes through the government, which might be operationalized as government expenditures as a percentage of GDP (ideally this would be government at all levels).

McDonald (2010) is a good example of the conceptual and operational problems of this literature. He starts with general claims about the "capitalist" peace. Then he moves to a claim that capitalist economies have less "public property." One might think that public property is property owned by the government. This would then include land, highways, military bases, buildings, etc., not to mention financial capital assets (both at home and abroad). In practice, i.e., quantitative analyses, he uses nontax revenue of the government. This variable seems to be quite distant from the starting point of capitalism.

291. Draw the three-level figure that corresponds to Levitsky and Way's (2010 Appendix I) concept of "competitive authoritarianism."

292. Concepts with adjectives, e.g., competitive authoritarianism, often confuse to two quite different issues. Sometimes they mean regions in the middle of a continuum as is often the case with democracy. They can also two *separate* concepts that are put together, which is the classic use, e.g., "pet fish." He and Warren (2011) illustrate this potential confusion. The concept they propose is "deliberative authoritarianism" which they discuss in the context of hybrid regimes. However, they argue that "deliberation" is separate from "democracy": "If deliberation and democracy are distinct in theory – the one a kind of communication, the other a distribution of powers to decide – they have often been distinct in practice as well.... These observations can be ideal typed. If deliberation is a phenomenon different in kind from democracy, then (in theory) it might combine with non-democratic (authoritarian) distributions of power. We illustrate the ideal types in Table 1, where the terms "authoritarian" and "democratic" refer to the relative dispersion of means of empowerment (dispersion, by implication, provides more opportunities for the affected to exercise power), while communication can vary from "instrumental" to "strategic" and "deliberative.". The combinations produce five familiar types, and one unfamiliar type, deliberative authoritarianism" (He and Warren 2011, 273). Analyze the difference between middle of the continuum and the separate concept approaches in dealing with concepts with adjectives.

293. While scholars often use the "ideal type" to think about concepts there is virtually no methodological literature on building ideal types. He and Warren (2011) illustrate some of the issues of trying to do ideal type concepts with two dimensions:"we then develop the ideal type of deliberative authoritarianism – a regime style that makes frequent use of authoritarian deliberation. In developing this ideal type, we depart from much of the literature on hybrid regimes (He and Warren 2011, 270). If one thinks about about "deliberation" as a separate concept (see the question above) consider their view of the

extreme positive pole of "deliberative authoritarianism" (e.g., Table 1, 273). Is it is really the extreme for deliberation? (Hint: think about the positive extreme for deliberation for democracy).

Answer:

If one thinks of deliberation as a concept in its own right the ideal type, i.e., positive pole, should continue for at least one more column. Certainly good deliberative democracy includes more deliberation than deliberation in authoritarian regimes like China. Since the authors focus on ideal type which is a combination of authoritarianism and deliberation, they naturally want to stop the scale at the end of the authoritarian row, without considering deliberation as a concept in its own right. Hence it would be less natural to have an empty cell at the end of the second row which is high quality deliberation which one would naturally have to include if one wanted to include the positive pole of deliberation for democracy.

294. Adjectives are often used to distinguish among subtypes. Discuss Rathbun's (2011) use of "qualitative" to distinguish between different types of multilateralism; for example, "RH1: State preferences for qualitative multilateralism will be accompanied by preferences for organizations with a smaller number of members." Discuss the other type which is "quantitative" multilateralism.

295. Gandhi and Lust-Oskar (2009) review the literature on "authoritarian elections." Discuss this interesting example of a concepts and adjectives. Fair or competitive elections are of course part of the definition of democracy. Presumably their absence is a requirement for an authoritarian regime. Do such cases fall into the gray zone between democracy and authoritarianism?

296. There is an extensive cognitive psychology literature on the use of *prototypes* in categorization (see Murphy (2002) for a nice review). The basic idea is that people (not social scientists) often think about concepts via prototypes. For example, the concept of a "bird" is based on prototypical birds, e.g., sparrows. It is less clear whether social science concepts should or are based on prototype models of concepts. Shadish, Cook, and Campbell (2002; essentially the 3rd edition of the classic Cook and Campbell) provides an example where concepts (aka "constructs") are discussed using the prototype idea. Discuss the use of prototypes in social science. Are they different from ideal types? Contrast prototypes with concepts based on defining, essential features (a common contrast in the cognitive psychology literature).

297. Colaresi et al. (2007) in their prominent work on the concept of "strategic rivalries" argue (chapter 6) that there are two types of rivalry "positional" (i.e., roughly competition in the regional or global system) or "spatial" (i.e., over territory). Table 6.1 argues that there is a continuum from pure positional to pure spatial, i.e., the "spatial–positional continuum." Discuss whether one continuum with the two combined would be better than two continuums one for each type of rivalry going from presence to absence on each continuum.

298. Abbott, Keohane, Moravcsik, Slaughter, and Snidal (2000) published a very influential article on "The concept of legalization." This article could serve

as a good example for the discussion of concepts. (1) Draw a diagram that represents the three-level structure of their concept. The three secondary-level dimensions are quite clear and the indicators are given in various tables. (2) What is the structure that they implicitly use to connect secondary-level dimensions and indicators? (3) Thinking about the negation of a concept is important; what about the negations given in figure 1 as well as table 1? (4) "Delegation" (table 4) might be separated into two, or even three, separate dimensions? Are "dispute resolution," "rule-making" and "implementation" part of the higher level concept of "delegation?" (5) In the delegation dimension the concept of "binding" appears frequently, but is also central to the "obligation" dimension. Is this a problem?

Answer:

It seems like the implicit structuring used at both levels is the family resemblance one. There is not much sense that any of these are necessary. There are potentially other negations that could be used in figure 1, such as "nonbinding" for obligation. For precision it is interesting that the positive pole uses the term "rule" while the negative pole uses the term "norm." Delegation is problematic because it includes both implementation and arbitration. In terms of figure 1, "international court" suggests third-party dispute settlement mechanism and so the opposite might be bilateral agreements. Much of the problem with delegation is that this is where most of the *organizational* dimensions of legalization appear; but it is not obvious that one would lump administrative, monitoring, and sanctioning organizations together with courts (i.e., one would not do this in an analysis of domestic governance and legal systems.) Also missing is a potential dimension about *who* makes international rules, e.g., states, United Nations, IGOs, etc.

299. Gates et al. (2006) reformulate a new three dimensional concept of democracy using polity and Vanhanen. They use a cube to think about their concept of democracy, anocracy, and autocracy. In their statistical analyses they use a dichotomous coding of democracy, autocracy, and anocracy. How could they have used continuous [0,1] variables for each of these?

> The ideal types also include polities that are close to the corners. As the cube in Figure 1 defines a space, it is possible to examine the distances within this space. In order to classify a regime as either Ideal or Inconsistent, the distance from the point given by the polity's coordinates to the eight corners and the midpoint of the cube is calculated. A regime is defined as Democratic or Autocratic if it is closer to either of the ideal type corners of the cube than to the other corners or the midpoint. The Democratic ideal type will hence include observations that are closer to the corner [1, 1, 1]. Since it is the distance from the democratic corner that defines the ideal type, scores close to 1 on one of the dimensions to some extent offset low scores on the other dimensions. The autocratic ideal type includes all polities that are closer to the autocratic corner than any of the other reference points. All polities that are not coded as Autocratic or Democratic are coded as Inconsistent. For the 1800–2000 and 1900–2000 periods the respective distributions of the three types of polities were as follows: Autocratic (43%, 39%), Democratic (14%, 17%), and Inconsistent

(43%, 44%). (Gates et al. 2006, 898, note that they conceptualize the three dimensions this with the cube).

Answer:

Using the cube they can calculate three distances for each data point in the cube: (1) the distance to (1,1,1) which would be the democracy score, (2) the distance to (0,0,0) which would be the autocracy score, and (3) distance to (.5,.5,.5), this distance will have to be normalized to be in the [0,1] range.

300. Volgy et al. (2008) is a nice example of an implicit three-level concept structure. Draw a figure of the concept. What are the dimensions and the structural principles used at each level?

301. In literature on the concept of ethnicity, perhaps the biggest problem is defining the secondary-level dimensions that are used to conceptualize ethnicity. Some commonly used dimensions are religion, language, and race. Discuss what the complete list should be. Should it include factors like class or ideology?

Answer:

Here is one answer: "What is ethnic identity? Since the publication of Horowitz's (1985) *Ethnic Groups in Conflict*, there has been a convergence among comparative political scientists on which identities we classify as ethnic. For Horowitz, ethnicity is an umbrella concept that "easily embraces groups differentiated by color, language, and religion; it covers 'tribes,' 'races,' 'nationalities,' and castes" (Horowitz 1985, 53). Much of the recent theoretical literature on ethnic politics explicitly follows this umbrella classification (e.g., Varshney 2002, Chandra 2004, Htun 2004, Wilkinson 2004, Posner 2005). Even more importantly, the four principal datasets on ethnic groups that constitute the foundation for cross-national empirical studies of the effect of ethnic identity in comparative politics – the Atlas Narodov Mira (Bruk and Apenchenko 1964), Alesina et al.'s (2003) dataset on ethnic groups in 190 countries, a comparable count of ethnic groups in 160 countries (Fearon 2003), and the Minorities at Risk project (http://www.cidcm.umd.edu/inscr/mar/) – also generally employ this umbrella classification. Only some quibbles remain on the margin about whether castes should be excluded (e.g., Fearon 2003) or retained (e.g., Bruk & Apenchenko 1964, Varshney 2002, Chandra 2004, Htun A 2004, Wilkinson 2004, Posner 2005) and whether region and clan should be included." (Chandra 2008, 397–98)

302. For German speakers. Opp (2005) is a very rare methods textbook that devoted significant attention to concepts – most textbook only look at measurement; it is also a rare textbook that devotes a whole chapter to teaching logic – most research design and methodology textbooks do not discuss this at all. In Opp's discussion of concepts (chapter 4) he distinguishes between (1) "Analytische operationalisierung" and (2) "empirische operationalisierung." Discuss the extent to which (2) corresponds the the standard latent variable model of concepts and (1) corresponds to what the concept book calls ontological concepts.

303. Concept asymmetry appears in various contexts. In survey research there are often major differences depending on the how the question is worded. For example, a classic article by Rugg in 1941 found differences of up to 20 percent of two questions "that seem logically equivalent." (Holleman 1999, 210), "Should the USA forbid public speeches against democracy" versus "Should the USA permit public speeches of democracy." One of the main hypotheses in the literature is that there are different scales for the two responses. Holleman expresses this clearly: "The main issue is whether questions worded differently, although conceptually equivalent, measure similar attitudes (a test of the different attitudes hypothesis) and, if so, whether the similar attitudes are expressed on similar scales (the different scales hypothesis)" (Holleman 1999, 213). She finds empirical support for the two scales hypothesis.

    Using semantic transformations discuss whether these the response of yes to one question is the equivalent to the answer of no on the other. Does your answer suggest that perhaps fuzzy logic scaling or concepts would be appropriate here.

    Answer:

    Concept asymmetry means that "poor" does not necessarily equal "not wealthy." Hence it might be the case that "not forbidden" is not necessarily equivalent to "permitted." The fact that empirically two scales are found might suggest that a fuzzy logic approach might be better.

304. The GTD (global terrorism database) database (CETIS 2007) uses a hybrid concept of terrorism involving 3 necessary conditions and then 2 of 3 other conditions for an incident to be considered terrorist. (1) Discuss whether various kinds of Mafia violence could be considered terrorism under this concept. (2) Discuss if nonviolent demonstrations would count. (3) One important aspect of concept construction is considering the negative pole or phenomenon that do not fall under the concept. The coding manual *implies* that insurgency is different from terrorism, e.g., "Appendix A: guidelines for differentiating between insurgency and terrorism," but it does not define or conceptualize "insurgency"; how might this be an issue in studying terrorism and, say, civil war? (4) Discuss the secondary-level dimension "violates international humanitarian law of war."

    Answer:

    Mafia activities could count because while they fail on one of the 2 out of 3 conditions (economic motives only) they can fufil the other two conditions. Also there are coding categories for "criminal groups" (p. 32).

    "Pacifist"/Anti-War is one possible category of terrorist group.

    In addition to appendix A there is a variable "alternative designation" for incidents which are not considered terrorist, which include categories "Insurgency or guerilla action," and "Internecine conflict action." It is not at all clear what the relationship is between terrorism, insurgency, state breakdown, and civil war.

305. Democracy is a key variable in the theories about civil war. Vreeland (2008) discusses how the polity coding of anocracy is problematic for this use. Explain why these problems make sense given the original goals of the polity concept (Gurr 1974).

Answer:

Gurr (1974) was interested in the stability of regimes, and by consequence regime change or viability. A country in civil war can be considered and coded as nonstable.

306. The concept of "civilian power" engages the nature of foreign policy in countries like the EU and Japan (e.g., see Maull et al. 2006; Smith 2006). Discuss how the adjective "civilian" works in this concept. For example, how does it modify the concept of power? Also discuss the role of the negative pole in this concept.

Answer:

Civilian is clearly discussed as the opposite of "military." Note that many discussions of power would include nonmilitary forms of power.

307. Some key concepts are basically defined in a negative way (e.g., see exercise 306 on "civilian" power). A good example of the issues involved is the debate over the "democratic deficit" in the EU (see Moravcsik 2002; Follesdal and Hixs 2005). One seems to be taking away dimensions from the concept of democracy. In practice "democratic deficit" is often defined without a clear concept of democracy. Discuss the (implicit) concept of democracy used by various authors. Is it defendable or common as a stand-alone concept of democracy?

308. There is a very large literature in American politics on the concept and measurement of racism, and it is a topic that arouses much controversy: "Is white opposition to policies driven by racial prejudice or is it grounded in race-blind ideological principles?" (Feldman and Huddy 2005, 168; this is first sentence of the article). There are many aspects of this debate of interest regarding concepts. (1) One interesting aspect is the variety of terminology, such as "overt prejudice," "new racism," "racial resentment," "old-fashioned racism," and "symbolic racism." In addition there are related concepts such as "prejudice" and "discrimination." In particular the use of adjectives is of interest: discuss the issues and implications of the use of the concept "symbolic racism" (e.g., Tarman and Sears 2005) as a particular use of adjectives. Notice as well the relevance of the negation guideline given in the use of adjectives. (2) The negation guideline suggests that much can be gained by focusing on the opposite concept. Much of the debate revolves around the the concept of a "principled conservatism" which argues against policies such as affirmative action but which is not racist (Sniderman et al. 1996). One could focus on that concept instead of directly on racism. For example, if one is *principled* then those principles should apply across groups by gender and sexual orientation and not just race. (3) It is important to think about causal relationships in concepts. One might suggest that, at least in the case of the south, principled

conservatism is a consequence of racism, given that race was a central component of Southern culture and politics for centuries. See Quillian (2006) for a nice survey of the political science, sociology, and psychology literatures.

309. One serious issue with dichotomous data is the potential for serious heterogeneity in the zero or one category. This is particularly likely to be a problem in zero category. The issue of heterogeneity arises almost by definition with nonordered, categorical data. For example, quantitative studies of international conflict that look at the outcome dispute variables that have anywhere from 3 to 9 categories. Often these can be reduced to settlement/compromise, win, draw. If the focus is, say, settlement or compromise then there is a tendency to code that as one and lump win and draw together as zero. Discuss the merits of this procedure. What would be alternatives?

    Answer:

    Lumping all these together in the zero seems particularly dubious when the causal mechanisms producing these zeros are likely to be quite different.

    Within the context of the conflict literature and the militarized disputes dataset, it might make sense to just include draw/stalemate as the contrasting case since this constitutes often 70+% of the zero cases (e.g., Goertz et al. 2005).

    This choice could well be brought into the research design and theory if the research were interested in making specific causal contrasts.

310. Typically, one attacks theoretically and empirically concept issues from the positive pole first. However, sometimes it might make sense to really focus on the negative. For example, the concept of "judicial independence" is key in the literature on comparative judicial systems and democratization. Evaluate Larkins (1996) proposal to focus on the dimensions of judicial *dependence* as the best way to attack the problem. Also, discuss whether the necessary and sufficient condition or family resemblance approach makes more sense.

    Answer:

    The answer to the second question about concept structure informs the response to the first. Typically, there are many ways in which governments can make judiciaries dependent. Usually you only need one to remove judicial independence; i.e., using the contrapositive of the necessary and sufficient condition structure means that the absence of all the means of judicial dependence is sufficient for judicial independence.

311. Sometimes the "positive" pole of interest is itself defined in negative terms. The concept of an "antisystem party" has a long history in comparative politics, notably Sartori (1976): "a party that would change, if it could, not the government, but the system of government" (cited in Capoccia 2002, 18). Discuss the problems of defining what the "system" is that forms the positive pole. Frequently, "system" means "democracy." Capoccia (2005, 2002) discusses two types of antisystem parties one "ideological" and the other "relational." Relational means distant and extreme parties in ideological space. Discuss the relationship between "anti" in terms of the ideological space of the parties versus the characteristics of the political system itself. How would

a family resemblance perspective on democracy be more problematic than a necessary and sufficient condition one in looking at system as democracy?

Answer:

See Capoccia (2002) for an extended discussion. (1) Inherently it is going to be hard to define what the "system" is. (2) One party could be distant ideologically from the center and not be opposed to the system of government itself. (3) Using the necessary and sufficient condition view of democracy makes it easier to think about antisystem parties since if they are opposed to any secondary-level dimension of democracy then it is an antisystem party.

312. In many debates the role of the negative pole plays a key role. One important point in the debate between Duffield (2003) and Koremenos et al. (2003) is the role of "centralization" in the study of international institutions. What is the opposite pole? What would be the opposite pole in the study of domestic institutions?

    Answer:

    Duffield and Koremenos et al. assume that the opposite pole of centralization is "delegation;" for example, "Duffield makes a good point in stating that greater focus on delegation may provide a way to further refine the concept of centralization, because delegation and centralization are indeed closely related" (Koremenos et al. 2003, 434).

    It is likely that within the literature on domestic institutions design (e.g., constitutions), the opposite pole would be "federalism."

313. A key issue when creating concepts is their eventual theoretical and policy usage. Of great policy importance today is the concept of "indigenous people" (see Corntassel 2003 for an excellent review of different concepts). Clearly, one policy goal is protect indigenous peoples. Discuss how this influences one's definition. Notice that many important international organizations such as the UN, ILO, and the World Bank have developed definitions that have practical and legal import.

314. Describe the three-level structure of Dasgupta's concept of human well-being or destitution (Dasgupta 1990; Dasgupta and Weale 1992). What is its structure? What are the relative weights attached at the indicator and secondary levels?

    Answer:

    The structure is very similar to the one used by Schmitter for corporatism, since it uses the sum of the ranks. The secondary-level level variables such as "health," "education," "wealth" will not be equally weighted in general unless they have the same number of indicators. More generally, is it reasonable to ranking literacy equally with per capita income?

315. Milner and Kubota (2005) discuss a number of important issues regarding concepts and measurement. (1) Discuss their measure of "open/closed trade regimes." What is the strategy used for building the measure? (2) Contrast this with their concept(s) of "economic crisis." What is the strategy used for

building these measures? Would it be better to have one or two economic crisis variables? What impact or role does the information that the two crisis variables are not correlated (i.e., $r = .01$) have on your argument?

Answer:

See Milner and Kubota (2005, 122–24). For the open/closed trade regime variable they use a family resemblance strategy by applying the m-of-n rule: if one of the four secondary-level factors is present then the variable is coded one. Milner and Kubota probably separated out the two economic crisis variables because of their low correlation.

316. Examine Doner et al.'s concept of the "developmental state" (Doner et al. 2005, Table 1) and draw a figure of their three-level concept.

317. Przeworski and his colleagues have proposed an influential concept and measure of democracy:

> Operationally, a regime was classified [dichotomously] as a democracy if none of the four rules listed below applied . . . .
> Rule 1: Executive selection: the chief executive is not elected.
> Rule 2: Legislative selection: the legislature is not elected.
> Rule 3: Party: there is no more than one party. Specifically, this rule applies if (1) there were no parties, or (2) there was only one party, or (3) the current tenure in office ended up in the establishment of a nonparty or one-party rule, or (4) the incumbents unconstitutionally closed the legislature and rewrote the rules in their favor.
> Rule 4: Type II error: a regime passes the previous three rules, the incumbents held office in the immediate past by virtue of elections for more than two terms or without being elected, and until today or the time when they were overthrown they have not lost an election. (Przeworski and Limongi, 1997, 178; see also Alvarez et al. 1996, Przeworski et al. 2000, chapter 1)

Rule 4 (called "Alternation" in Przeworski et al. 2000) only applies if "a regime passes the previous three rules." Would anything change in the final dichotomous democracy codings if we just considered this a fourth rule? Explain.

Answer:

Nothing would change in the final dichotomous democracy coding. Because failure on any rule, including rule 4, eliminates the country as a democracy nothing is gained (except perhaps time) in the overall democracy coding by applying rule 4 only when a country passes rules 1–3.

318. An important concept in the study of international conflict is that of the "similarity" of regime types. One central hypothesis is that international conflict is more likely between dissimilar regimes. One can compare two countries on their level of democracy in at least two ways with a multilevel democracy concept: (1) via the similarity of the secondary-level dimensions or (2) via the similarity of the basic-level measure. Discuss if one procedure is more valid than the other.

Answer:

Similarity comparisons at the level of secondary-level dimensions with necessary and sufficient condition structures are not valid because the claim is that the absence of one dimension makes the a country a nondemocracy, independent of what is happening with other secondary-level dimensions. In general, similarity should be addressed at the basic level not using secondary-level dimensions because using secondary-level dimensions ignores how the concept is structured (see Bueno de Mesquita et al. 2005 for an extensive justification for examining the secondary-level dimensions individually).

319. One important guideline for concept-building is to think about the negative pole. Esping-Anderson (1990; 1999) focused attention on three "worlds" of welfare capitalism, which has informed 15 years of research. This view of the welfare state has three dimensions (1) socialist, (2) liberal, and (3) conservative. Evaluate Hicks and Kenworthy's argument that instead of three dimensions there are two with (2) being the negative pole of (1):

> we can characterize and differentiate welfare states in terms of the "progressive liberalism" and "traditional conservatism" of their policies and programs. The first of these two dimensions is fairly novel. It rearranges Esping-Andersen's separate social democratic and liberal dimensions into two poles of a single dimension. (Hicks and Kenworthy 2003, 51).

320. There exists numerous measures of corporatism (see Siaroff 1999 or Kenworthy 2003 for surveys). Examine Cameron (1984) who uses a hybrid measure. What is the concept model that leads to hybrid measures?

321. Explicitly contrast Schmitter's concept of corporatism with his concept of pluralism. To what extent is pluralism strictly the negation of corporatism?

322. One important issue is redundancy across indicators or dimensions. The CIRI measure and state on human rights violations has four categories of violation: (1) torture, (2) political imprisonment, (3) extrajudicial killings, (4) disappearances. Cingranelli and Richards propose a quantitative measure where each of the four dimensions gets 0–2 on the level of violations in a given country in a given year; their final score is the sum of the scores of the four dimensions. Thus the worst performing states score 8, and the least human rights violating states get 0.

    Woods and Gibney (2009) critique the CIRI scale because it would count the torture, political imprisonment, and disappearance of a single individual three times. How could you reformulate the structure to deal with this critique?

    Answer:

    Rescale the data so that (1) (i.e., torture) gets 0 or 2 points, (2) gets 0 or 4 points, (3) gets zero or 6 points, and (4) (i.e., disappearances) gets zero or 8. If the data fit perfectly the Guttman assumption, then if a country-year gets is scored for disappearances, it would have all levels (1)-(3) and hence a total of either. With the new scoring system it would also have a score of eight.

    The logic is then not to count multiple acts against an individual person separately but to only count the most severe act of physical integrity violation.

323. Often the issue of dichotomous versus continuous lies hidden in conceptual discussions. Vu's discussion of the analysis of the state makes a distinction between "variables" and "attributes." How does the issue of dichotomous dimensions seem to lie in the distinction between the two?

> Weber defines it [the state] as "a human community that (successfully) claims the *monopoly of the legitimate use of physical force within a given territory*." Later students of the state have modified this definition somewhat but the essential elements remain the same. Tilly defines states as "coercion-wielding organizations that are distinct from households and kinship groups and exercise clear priority in some respects over all other organizations within substantial territories." For this definition he means to include city-states, empires, and theocracies and to exclude tribes, lineages, firms, and churches. "National states" are a subset of states that are relatively powerful, centralized, and differentiated sovereign organizations. Levi offers a more restrictive definition than Tilly: "a state is a complex apparatus of centralized and institutionalized power that concentrates violence, establishes property rights, and regulates society within a given territory while being formally recognized as a state by international forums." These later definitions avoid such terms as "monopoly" and "legitimate," perhaps because these are seen as variables, not attributes. (Vu 2010,164–65; Levi 1982, 40 makes the same distinction between "variables" and "elements of a definition.")

Answer:

"Variables" are continuous dimensions, while Vu seems to assume that "attributes" and dichotomous. One could certainly think that all the attributes of a state discussed in this quote are continuous.

324. The topic of concepts with adjectives arises with great force in thinking about the EU as a state. Scholars talk about Westphalian, regulatory, post-modern, regional, international, e.g., (Caporaso 1996). Analyze these various concepts in terms of their coverage of cases the degree to which they are subsets of a Weberian concept of a state. Should one think of the EU in terms of a gray zone? Discuss how the conceptualization of the EU as a state influences your view of the "stateness" of the members.

Answer:

One potential solution is to consider the EU and its members as "quasi-states." Basically, in some areas such as trade the EU acts like a sovereign state. In other areas the members retain sovereign control. Hence both the EU and its members are in the gray zone.

325. For some important concepts like "democracy" and "peace" scholars have attached the adjective "stable." In early work (e.g., 1960s) this was common for democracy, but the "stable" adjective was dropped until people began to work on democratic transitions. In contrast, it is still very commonly used to denote a high level of peace (e.g., Kacowicz and Bar-Siman-Tov 2000). What is a potential problem with using stable as an adjective?

Answer:

In the early literature on democracy and the literature on peace, often stable often means "high level." For many "stable peace" and "positive peace" mean the same thing. It useful to distinguish between the level of a variable and how stable the phenomenon is at any given level. This is not a problem in the literature on democratic transitions since the emphasis is really on stability. See Goertz et al. (2016) for a survey of the literature on the concept of peace.

326. Terrorism is a big policy as well as scientific issue. There are a large variety of different adjectives attached to the concept "terrorism" including international, transnational, state-sponsored, state, domestic, political, revolutionary. In making sense of this one must consider how the adjective is modifying the terrorism concept. Related to this is that many of the adjectives involve the negative pole of a dichotomous concept, e.g., state versus nonstate. Analyze the conceptual issues involved in defining terrorism, with or without adjectives. Discuss how the decisions one makes can have a significant impact on case selection. For a survey of different concepts of terrorism see for example Crenshaw (1972), Gibbs (1989), Held (2003).

327. Sometimes the ideal point is not at the end of the scale but in the middle somewhere. Analyze the concept of "judicial (in)dependence" as an example of this. See Larkins (1996) for a review of the literature.

    Answer:

    Typically, one wants the courts to be not completely unresponsive to elected officials. At the same time the courts cannot be completely submissive to these officials. Hence the ideal point is some degree of independence but also the possibility that judges can be removed by elected officials. Also, there are typically legislative mechanisms for overriding some judicial decisions.

328. It was suggested in the concept book that hybrid concepts – with some necessary and some facilitating conditions – might be good for idea types. Is this what Diamon and Morlino are suggesting?

> The five essays that follow are part of a collaborative effort, launched at a conference at Stanford University, to elaborate and refine the concept of democratic quality and to apply it to a series of six paired country comparisons. We asked each author to discuss a particular dimension of the quality of democracy such as freedom, the rule of law, vertical accountability, responsiveness, and equality (our own list, and by no means exhaustive). We wanted each author to explain how the dimension in question relates to other dimensions in our framework, to suggest possible indicators for measuring the dimension, to identify ways in which this element of democratic quality is subverted in the real world, and to offer (where possible) policy recommendations. Our full framework features eight dimensions: the five outlined above, plus participation, competition, and horizontal accountability. Other dimensions might include transparency and the effectiveness of representation. The different aspects of democratic quality overlap, however, and we choose to treat these latter two as elements of our principal dimensions. We attempt here to identify some of the ways in which the different elements of democracy not only overlap, but also depend upon one another, forming a system in which improvement along one dimension (such as participation) can have beneficial effects along

> others (such as equality and accountability). At the same time, however, there can be trade-offs between the different dimensions of democratic quality, and it is impossible to maximize all of them at once. In this sense at least, every democratic country must make an inherently value-laden choice about what kind of democracy it wishes to be. Talk of a "good" or "better" democracy implies knowing what democracy is. At a minimum, democracy requires: 1) universal, adult suffrage; 2) recurring, free, competitive, and fair elections; 3) more than one serious political party; and 4) alternative sources of information. If elections are to be truly meaningful, free, and fair, there must be some degree of civil and political freedom beyond the electoral arena so that citizens can articulate and organize around their political beliefs and interests. Once a country meets these basic standards, further empirical analysis can ask how well it achieves the three main goals of an ideal democracy–political and civil freedom, popular sovereignty (control over public policies and the officials who make them), and political equality (in these rights and powers)–as well as broader standards of good governance (such as transparency, legality, and responsible rule). (Diamond and Morlino 2004, 20–21)

329. How does Sartori view the question of sufficiency in conceptualization and concept structure? See also Collier and Levitsky (1997) for the same question.

    Answer:

    He focuses his attention on the minimal necessary conditions and does not explicitly make claims about sufficiency. In one footnote he explicitly rejects sufficiency ideas. Collier and Levitsky make no claims about sufficiency and focus only on the minimum requirements for democracy.

330. In the example showing that extension can increase with extension with family resemblance concepts, the book used a constant "half or more" rule. In practice, it seems to be case that as the total number of possible characteristics increases that the threshold does as well. For example, one might use 2 out of 4, but when the total is 10 the rule might be 7 out of 10. Discuss the validity of this tendency. What are the implications in terms of extension?

    Answer:

    It seems like there is some intuition that when the total number of characteristics increases that extension is likely to increase (as demonstrated in the text). To keep the extension in line, the threshold is thus increased.

331. One part of the conventional wisdom is that measures of democracy are highly correlated. It may be the case that this high correlation masks significant disagreement, particularly in the gray zone. Bogaards finds big potential differences between concepts and measures of democracy for African cases. Contrast Bogaards's (2007) examination of African data with Treier and Jackman's (2008) discussion of errors in measurement. Where do each see the biggest problem of measurement error?

    Answer:

    Bogaards almost by definition looks at only gray zone cases, i.e., elections in Africa. His analysis suggests that there can be major differences between democracy measures. In contrast, Treier and Jackman argue that there is most

uncertainty at the ends of the authoritarian–democracy scale; measurements are most accurate in the middle (i.e., the gray zone). Bogaards analyses imply that uncertainty and error are more likely in the middle.

332. A key assumption in multidimensional concepts is that the various means of achieving a given aggregate level are equivalent (e.g., 2+3 and 4+1 are really the same). What is one criterion for defending equivalence? Bueno de Mesquita et al. (2005) provides a very good discussion of these issues.

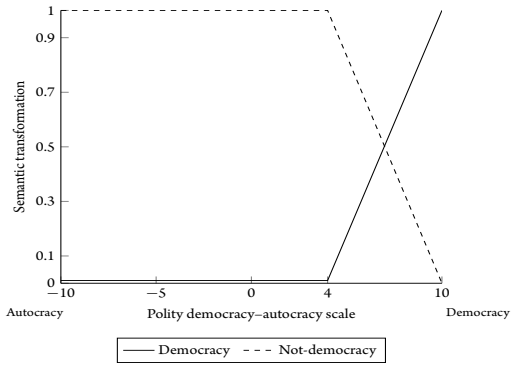
Answer:

One common criterion for equivalence is similar causal effect in some key hypotheses. Bueno de Mesquita et al. (2005) illustrates this.

Semantic transformations, calibration

333. Chapter 4 of the concept book deals with transforming basic raw data to better conform with concepts. Instead of the standard linear model (see the exercises above about scaling and using the average). The book suggests that if one is interested specifically in democracy (and not authoritarianism) one potential transformation using the polity scale looks like figure **??**. Discuss this option in the context of the Little-Meng debate. In particular, (1) how does it deal with the exercise above blacksliding among democratic versus authoritarian states? (2) Is it going to emphasize the degree of backsliding in comparison with the original scale?

Figure 51: A semantic transformation of a democracy measure, focusing on democracy not authoritarianism



Answer: Notice that all countries with a value of less than 4 have zero on this transformed democracy measure. So any change among those does not constitute backsliding because the changes are zero. This illustrates the principle discussed in the concept book whereby some changes in the basic raw data do not involve changes in the concept. Notice as well that this particular transformation stretches out the democracy scale which now runs

from 0 to 1. In the V-Dem scheme of things .5 is often taken as the cut off for the authoritarian regime so the democracy scale runs from .50 to 1.0. So inherently the raw backsliding values have a maximum of .5, since .5 is the cutoff, where is in the transformed data in figure **??** have a maximum backsliding value of 1.0

334. Discuss how one might do semantic transformations of both the independent and dependent variable in Henke's analysis of multilateral cooperation in military coalitions. On the dependent variable side how would that compare to do doing three different dichotomous analyses for three ordinal levels of the dependent variable? How might one use some descriptive statistics to get a sense on how to scale the independent variable of "diplomatic embeddedness"? Would an S-curve transformation make sense?

> I now add to this range of independent variables a new variable: diplomatic embeddedness. I operationalize diplomatic embeddedness by counting all bilateral and multilateral institutional ties the United States has established with a foreign country at the date of intervention. As mentioned in the preceding chapter, I define diplomatic embeddedness as the cumulative number of bilateral and multilateral diplomatic ties that connect a country dyad. Since most of these ties are the result of bilateral or multilateral agreements that a country dyad entertains, such a coding choice seems reasonable. Data on cumulative joint international organization membership come from the International Governmental Organization (IGO) Dataset. Data on cumulative bilateral cooperation treaties come from the World Treaty Index. . . . Lake uses "coalition participation" – a binary variable – as the dependent variable. In this new analysis, I attempt to provide greater detail. Therefore, I use three dependent variables: (1) a binary variable set to one if a country participated at all in a specific U.S.-led coalition; (2) a binary variable set to one if a country deployed at least a company-sized contingent (minimum one hundred troops); and (3) a continuous variable accounting for the exact number of troops deployed by a given country to a specific U.S.-led coalition. While the first dependent variable examines overall participation, the second dependent variable focuses on substantive contributions, thus excluding symbolic or token contributions. The third dependent variable captures the exact level of troop commitment. (Henke 2019, 46, 49)

335. Discuss the structure of the concept of "targeted mass killing" its structure (i.e., aggregation) and threshold. Would a best shot aggregation technique be more appropriate? The basic level concept is an ordinal one with 8 levels. Would a semantic transformation be appropriate here.

> Targeted mass killing is the direct killing of noncombatant members of a group by a formally organized armed force that results in twenty-five or more deaths in an annual period, with the intent of destroying the group or intimidating the group by creating a perception of imminent threat to its survival. A targeted group is defined in terms of political and/or ethnic and/or religious identity. . . . From this broad set of potential cases, a new episode of TMK was coded when there was evidence that an organized armed actor: (1) killed twenty-five or more civilians in a year, (2) these civilians were deliberately targeted by that actor) (3) one (or more) political, ethnic, or religious group(s) were disproportionately targeted, and

> (4) the group was targeted in order to substantially reduce its numbers, expel, or affect the political activity of that specific group. An active TMK episode was coded when all four of these criteria were met. . . . Table 1. A Targeted Mass Killing (TMK) Ordinal Scale. (Butcher et al. 2020, 1528)

336. In her influential book on rape during war Cohen creates a ordinal scale severity of rape with four levels. In later statistical analyses she does an ordered probit on these four levels. Discuss the advantages and disadvantages of doing a semantic transformation on these four levels (like the PTS example in the concept book) and then doing ordinary least squares or something like that instead of an ordered probit.

> In the coding rules, a coding of 0, 1, or 2 is based solely on the description of the magnitude of the reported rape, as is shown in table 2.4 . The highest level of rape, 3, is reserved for those conflict-years that are described in the most dire terms in the source document. These terms include phrases that are used in the policy discourse as synonyms for "severe" but are imprecise in a social science context. (Cohen 2016, 66)

337. Discuss and graph this semantic transformation from the EU statistics bureau:

> Often, the scores are based on the percentiles of the distribution of the indicator across the countries. For example, the top 5% of the units receive a score of 100, the units between the 85th and 95th percentiles receive 80 points, the units between the 65th and the 85th percentiles receive 60 points, the units between the 35th and the 65th percentiles receive 50 points, the units between the 15th and the 35th percentiles receive 40 points, the units between the 5th and the 15th percentiles receive 20 points, and, finally, the bottom 5% of the units receive 0 points (see Table 1). This is a way to prize the most performing countries and penalize the less performing ones. (Eurostat 2005, Tools for Composite Indicators Building, 49)

338. Discuss the claim about bounded versus unbounded scales in Morris's *The measure of civilization how social development decides the fate of nations*. How does this relate to the Ideal Type Guideline. Clearly fuzzy logic semantic transformation transformations always have a maximum 1.0. How a serious critique of that approach is this?

> It [HDI] could, up to a point, be used to measure change through time by simply comparing a single country's score in each annual report, but because the maximum possible score is always 1.0, the HDI does better at charting a nation's relative position within the world at a single point in time than at measuring diachronic changes in development levels. (Morris 2013, 28)

339. A core Ideal Typology guideline is a central one. This can mean "conceptual creep" as the conceptual scale expands. Discuss the implications, causal and conceptual of such creep as well as the intension–extension implications of such creep. Why might having a clear semantic transformation be critical in this process?

> As it turns out, abstract concepts can creep, too. For example, in 1960, Webster's dictionary defined "aggression" as "an unprovoked attack or invasion," but today that concept can include behaviors such as making insufficient eye contact or asking people where they are from (1). Many other concepts, such as abuse, bullying, mental disorder, trauma, addiction, and prejudice, have expanded of late as well (2). Some take these expansions as signs of political correctness and others as signs of social awakening. We take no position on whether these expansions are good or bad. Rather, we seek to understand what makes them happen. Why do concepts creep? (Levari et al. 2018, 465)

340. Discuss the SDG index and semantic transformations, particularly linear versus other options. Describe the semantic transformations for each of the indicators in table 7 using the transformations implicit in the red-yellow-green categorization schemes. How do they differ or not from linear transformations.

> The procedure for calculating the SDG Index has four steps: (i) perform statistical tests for normality and remove extreme values from the distribution; (ii) rescale the data to ensure comparability; (iii) aggregate the indicators within and across SDGs; and (iv) conduct sensitivity and other statistical test. . . . Z-scores are the most commonly used method in constructing composite indices (OECD 2008), but we see several reasons for not applying this approach to data underlying the SDG Index and Dashboards. . . . To remove the effect of extreme values, which can skew the results of a composite index, the OECD (2008) recommends truncating the data by removing the bottom 2.5 percentiles from the distribution. We apply this approach to the lower threshold and truncate data at this level. In this way we attenuate the impact of extreme values at the bottom end of the distribution on the SDG Index scores. . . . Where possible, the thresholds are derived from the SDGs, their targets, or other official sources. All thresholds are specified in absolute terms and described in Table 7. (Sachs et al. 2016, 31)

341. The MPI (Alkire et al. 2018) defines "severely poor" as a deprivation score of 50 percent or higher. What would be a fuzzy logic way of doing this?

    Answer:

    "Severely" means essentially the same as "very" so one can use the strategy for adjectives discussed in the concept book.

342. Key in the massive industry of conceptualizing and measuring well-being is the use of development goals and aspirational goals to perform key scaling operations on individual dimensions. These typically indicate what the 1.0 values are on the fuzzy logic dimension. Once you achieve this threshold then you are good. Discuss this example from Women, Peace, and Security Index: "Other indicators require setting maximum values. We set aspirational maximum values of 15 years for mean years of schooling and 50 percent for parliamentary representation." (Georgetown University's Institute for Women, Peace and Security, PRIO 2017, 57)

343. Many datasets contain ordinal variables. The Minorities at Risk dataset illustrates this nicely. Take variable REPGENCIV and discuss a semantic transformation of it.

> Repression of group civilian populations (those not engaging in violent or nonviolent political activities)
>
> 0 None reported
>
> 1 Surveillance, e.g., domestic spying, wiretapping, etc.
>
> 2 Harassment/containment, e.g., saturation of police/military presence; militarized checkpoints targeting members of groups; curfews, states of emergency
>
> 3 Nonviolent coercion, e.g., arrests, show-trials, property confiscation, exile/deportation
>
> 4 Violent coercion, short of killing, e.g., forced resettlement, torture
>
> 5 Violent coercion, killing, e.g., systematic killings, ethnic cleansing, reprisal killings
>
> -99 No basis for judgment
>
> (All Minorities at Risk (AMAR) Phase I Codebook August 2016, 36)

344. Discuss the semantic transformations of Vahanen's dimensions of his democracy concept. Think carefully about threshold issues.

> The selected threshold values [for democracy] of Competition (30 percent) and Participation (10 percent) are arbitrary, but I believe that they are suitable approximations for distinguishing more or less autocratic systems from political systems that have crossed the minimum threshold of democracy. Because both dimensions of democracy are assumed to be equally important, a country must cross both threshold values if it is to be classified as a democracy. It is not enough to define a threshold value of democracy solely for the ID [Index of Democracy]. In the case of the ID, I have used 5.0 index points as the minimum threshold of democracy, which is clearly higher than the ID value 3.0 produced by the minimum threshold values of Competition and Participation. Countries that have reached all three minimum threshold values (30 percent for Competition, 10 percent for Participation, and 5.0 index points for the ID) can be regarded as democracies. It should emphasized, however, that it is also possible to define threshold values differently, by raising or lowering them. (Vanhanen 2000, 257)

345. The World Bank has developed its own database of "Political Institutions" (Beck et al. 2001). This database of 108 variables includes an index/scale similar to those in concepts and measures of democracy:

> The database also contains two indexes that characterize the competitiveness of elections in countries, one for executive elections and one for legislative elections. The core of the two indexes is the number of parties that could and did compete in the last election. Building on work by Ferree and Singh (1999), we scale countries as follows:
>
> 1. No executive/legislature
> 2. Unelected executive/legislature
> 3. Elected, one candidate
> 4. One party, multiple candidates

5. Multiple parties are legal, but only one won seats (because other parties did not exist, compete, or win seats)

6. Multiple parties competed and won seats (but one party won 75 percent or more of the seats)

7. The largest party received less than 75 percent of the seats. (Beck et al. 2001, 166–67)

Do a semantic transformation of this ordinal scale. What the gray zone would be in this transformation? Should you transform some categories to the same level, e.g., zero? (i.e., the scale is not even ordinal)?

Answer:

As a start, it seems like only levels 6 and 7 represent some form of democracy. One could argue that the gray zone is probably levels 5 and 6, with 5 being dubious. Many concepts of democracy would probably code levels 1–4 as equivalently nondemocratic.

346. One issue in dealing with concepts deals with kinds of scales, ordinal, nominal, interval, and ratio. Is the polity measure of democracy interval or ratio? If it is ratio what is the zero point?

One could argue that –10 is the zero point, completely nondemocratic, like absolute zero as in a semantic transformation. It seems plausible that the polity *concept* of democracy does have a possible zero point (i.e., completely nondemocratic). However, most people treat the measure as interval when changing the scale by adding 10 (which is standard practice).

347. Give two reasons why the following definition of "extreme" case does not fit with a fuzzy logic approach to concepts or one defended by *Social science concepts and measurement*.

> Extremity (E) for the *i*th case can be defined in terms of the sample mean ($\overline{X}$) and the standard deviation (s) for that variable: $E_i = (|X_i - \overline{X}|)/s$. This definition of extremity is the absolute value of the Z-score (Stone 1996, 340) for the *i*th case. This may be understood as a matter of degrees, rather than as a (necessarily arbitrary) threshold. Since extremeness is a unidimensional concept, it may be applied with reference to any dimension of a problem, a choice that is dependent on the scholar's research interest. Let us say that we are principally interested in countries' level of democracy – the dependent variable in the exemplary model that we have been exploring. The mean of our democracy measure is 2.76, suggesting that, on average, the countries in the 1995 data set tend to be somewhat more democratic than autocratic (by Polity's definition). The standard deviation is 6.92, implying that there is a fair amount of scatter around the mean in these data. Extremeness scores for this variable, understood as deviation from the mean, can then be graphed for all countries according to the previous formula. These are displayed in Figure 3. As it happens, two countries share the largest extremeness scores (1.84): Qatar and Saudi Arabia. Both are graded as –10 on Polity's twenty-one-point system (which ranges from –10 to +10). These are the most extreme cases in the population and, as such, pose natural subjects of investigation wherever the researcher's principal question of interest is in regime type. (Seawright and Gerring 2008, 301).

Answer:

(1) This procedure depends on the sample and not on the meaning of the concept. For example, 2.76 is well outside the zone of what polity considers a democracy. Notice that given the positive mean of the data, all the extreme cases are nondemocracy.

(2) An extreme case based on conceptual meaning would typically mean either –10 or positive 10 on the polity scale, i.e., the extreme positive or negative pole.

348. Bühlmann, M. (et al. 2012) use "best practices" to determine the maximum value on the indicators for their Democracy Barometer. However, they claim that they do not need to worry about the minimum values (negative pole) because they are only looking at "established democracies." Can they get by without exploring worst practices?

Terminology, definitions, and semantics

349. Sometimes scholars use a wide array of adjectives to describe what appears to be the same concept. For example, "negative peace" (Boulding 1978), "precarious peace," (George 2000), "adversarial peace," (Bengtsson 2000), or "pre-peace" (Bayer 2005), "conditional peace" (George 2000), or "cold peace," (Miller 2001) all seem to refer to the same phenomenon. Discuss if this is in fact the case. Is one adjective to be preferred to the others? Are there some adjectives that are not good?

350. There is a very large literature in American politics on the concept and measurement of racism, and it is a topic that arouses much controversy: "Is white opposition to policies driven by racial prejudice or is it grounded in race-blind ideological principles?" (Feldman and Huddy 2005, 168; this is first sentence of the article). There are many aspects of this debate of interest regarding concepts. (1) One interesting aspect is the variety of terminology, such as "overt prejudice," "new racism," "racial resentment," "old-fashioned racism," and "symbolic racism." In addition there are related concepts such as "prejudice" and "discrimination." In particular the use of adjectives is of interest: discuss the issues and implications of the use of the concept "symbolic racism" (e.g., Tarman and Sears 2005) as a particular use of adjectives. Notice as well the relevance of the negation guideline given in the use of adjectives. (2) The negation guideline suggests that much can be gained by focusing on the opposite concept. Much of the debate revolves around the the concept of a "principled conservatism" which argues against policies such as affirmative action but which is not racist (Sniderman et al. 1996). One could focus on that concept instead of directly on racism. For example, if one is *principled* then those principles should apply across groups by gender and sexual orientation and not just race. (3) It is important to think about causal relationships in concepts. One might suggest that, at least in the case of the south, principled conservatism is a consequence of racism, given that race was a central component of Southern culture and politics for centuries. See Quillian (2006) for a nice survey of the political science, sociology, and psychology literatures.

351. Alvarez, Przeworski, Cheibub, Vreeland, and Gandhi who are all active working together on the concept and data on regime type have a clear preference for the term "dictatorship" as the opposite pole to democracy. In contrast, scholars using the polity data prefer the term "autocracy" or "authoritarian." Is there anything of theoretical or conceptual importance at stake in this terminological difference?

352. Discuss this definition from Eurostat. What do they really mean?

> Qualitative Indicator: The indicator is derived from a quality assessment, e.g. in the form of a yes/no question or a question about category (e.g. gender, nationality). (Eurostat 2014 Towards a harmonized methodology for statistical indicators, 20)

Answer: They probably mean a nominal indicator. However, "yes/no" can often translate into [0,1] with an underlying scale, which might exist depending on the "category."

353. How does the adjective "civil" function here? Is it a subsetting one, a causal relationship, or what?

> The next (and possibly most important) strand of the evolution of the understanding of peace is the civil peace. According to this approach, every individual in society has the capacity to mobilize for peace from a variety of different perspectives, whether for disarmament, for international cooperation, or against violence, discrimination, and oppression. It relates to the historical phenomena of social direct action for political, economic, and identity reasons, of citizen advocacy and mobilization, in the attainment or defence of basic human rights and values. It is also related to pacifism in its main forms, where civil action is non-violent in principle. It has been strongly influenced by a wide range of social mobilization dynamics. Without the civil peace and its social forms of mobilization, international and constitutional frameworks would not be able to connect with ordinary people in order to represent their interests, identities, needs, and aspirations. (Richmond 2014, epub 116)

Answer:

It seems to function as a causal relationship, i.e., as *cause* of peace rather than concept of peace. It is about how actions of individuals can create peace.

354. Terrorism is a big policy as well as scientific issue. There are a large variety of different adjectives attached to the concept "terrorism" including international, transnational, state-sponsored, state, domestic, political, revolutionary. In making sense of this one must consider how the adjective is modifying the terrorism concept. Related to this is that many of the adjectives involve the negative pole of a dichotomous concept, e.g., state versus nonstate. Analyze the conceptual issues involved in defining terrorism, with or without adjectives. Discuss how the decisions one makes can have a significant impact on case selection. For a survey of different concepts of terrorism see for example Crenshaw (1972), Gibbs (1989), Held (2003).

355. Following the work of Esping-Andersen, much of the work on the welfare state deals with the "commodification" of work. Here one needs to think about the negation: why use de-commodification instead of, for example, noncommodification. Also, the name of the concept implies a process instead of a state, does this matter? For a review see Knijn and Ostner (2002).

356. The word used or the meaning of the word-concept shift can easily shift over time. Discuss how this has been the case of the polity concept of "anocracy," which is part of the three core concepts of the polity measure, democracy, anocracy, and autocracy. Go to the original Gurr (1974) and see what he meant in the 1970s and contrast that with what anocracy means today.

    Answer:

    Anocracy for Gurr (1974) largely meant "anarchy", for example, see his table 2. From the Greek roots anocracy means "absence of rule." For Gurr it means "absence of power or control" (p. 1487, footnote 11). Today anocracy means a hybrid or mixed regime, one with both democratic and authoritarian characteristics.

357. It is quite common, and probably standard practice, in the literature on "quality of democracy" to use only "consolidated democracies." Lijphart expresses the standard view quite clearly. What is the role of the adjective "consolidated"? Does it mean "good" or "high quality" democracy?

    The same question arises in the literature on peace: "Peace consolidation generally refers to a process leading towards a self-sustaining peace" (UN 2010 "Monitoring peace consolidation United Nations practitioners' guide to benchmarking, Annex A")

    > Measuring the quality of democracy is the third step in a three-step judgment. We have to begin by establishing (1) that the countries we are interested in are sufficiently democratic in terms of regime and institutional characteristics to justify being called democracies, and (2) that their democracy is firmly established. . . . I am not sure where exactly the minimum threshold for institutional democracy should be drawn. My main argument is that there must be such a minimum, below which a country does not qualify as a democracy, and that it does not make sense to discuss the quality of democracy in countries whose institutions and rules are not sufficiently democratic. (Lijphart 2011, 18)
    >
    > According to Schedler, for instance, consolidating democracy means reducing the probability of its breakdown to the point where [we] can feel reasonably confident that democracy will persist. (Schedler 1998, 95). In a similar spirit, according to Gasiorowski and Power (1998, 740) a democracy is consolidated when it is sufficiently durable that a return to nondemocratic rule is no longer likely; and according to Acemoglu and Robinson (2005, 30) a democracy is consolidated if the set of institutions that characterize it endure through time. Empirically, questions about consolidation are addressed by examining what distinguishes democracies that achieve such durability from those that do not. In large-N research, this is frequently accomplished by specifying a temporal criterion that is used to identify consolidated democracies. Prominent examples of such criteria are Huntingtons two-turnover test and Gasiorowskis twelve-year threshold. . . . In this article, I propose a new empirical approach to democratic

> consolidation that avoids some of the weaknesses of existing research and leads to new findings about the dynamics and covariates of democratic consolidation and breakdown. . . . More precisely, following existing literature, consolidation is hypothesized to take place when a large, durable, and statistically significant decline in the risk of authoritarian reversals occurs at a well-defined point during the lifetime of a democracy. . . . The results that follow strongly suggest that democratic consolidation indeed occurs: A large and durable decline in the risk of authoritarian reversals is estimated to take place between the seventeenth and twentieth year of a democracys existence. Consolidation reduces the annual risk of breakdown from 1 in 33 for transitional democracies to 1 in 200 for consolidated democracies. In other words, consolidation comes close to eliminating the possibility of a return to dictatorship. (Svolik 2014, 715)

358. Terminology is often very important and leads projects in certain directions. The "Minorities at Risk (MAR) implies that the minority in question is not in the government. Could one pursue a "Majorities at Risk" project. For example, "the MAR coding scheme does not fit countries with ruling minorities or complex coalitions of ethnically defined elites, as for example in Nigeria, India, or Chad, where ethnic conflict will be pursued in the name of excluded majorities (rather than minorities) or ethnic groups that share power (and are thus not "at risk"). (Cederman et al. 2010, 91).

359. Samuel Moyn's published an influential book on *Christian* human rights (2015). What about a book about *Jewish* human rights? Human rights are very closely tied to democracy, so what about a book on *Christian* democracy? How is the adjectives (i.e., human rights with adjectives) working in this case?

360. The minorities at risk project (MAR) has receive much criticism for selecting minorities at risk. The new AMAR deals with this by adding "All" to MAR (Birnir et al. 2014). Discuss why this is bad terminology.

    Answer:

    Presumably the previous version, MAR, had all minorities at risk, what they have done is in fact dropped the "risk" part of the definition, it is now a dataset on minorities *tout court*.

Scaling

361. The V-Indoc [Varities of Indoctrination] coding manual (Varieties of Political Indoctrination in Education and the Media (V-Indoc): Codebook 2023), which uses the V-Dem methodology has this for scaling:"Scale: Dichotomous, Nominal, Ordinal, or Interval/Ratio (Extra response options such as N/A or Other, are not counted as part of this classification)." Discuss this list of scale types, particularly "interval/ratio." See the next exercise for one on "dichotomous." For example, is V-Dem or V-Indoc interval or ratio? Since V-Indoc is about education, is "years of education" (a common independent variable) interval or ratio? Also, the coding manual includes "multiple selection" for scale. What scale type is that?

Answer: As noted in the next exercise dichotomous is not a scale type. It is odd that "interval/ratio" are put together because they are different. V-Dem and V-Indoc are interval, as are all latent variable measurement models. Years of education is ratio because zero is a real zero (zero is arbitrary in interval indicators). "Multiple selection" means nominal.

362. It is always useful to understand the scale type of indicators. Sometimes in coding manuals and elsewhere people will describe the indicator as dichotomous. This is not a scale type. Often it is coded [0,1]. With such a coding scheme describe how the indicator could be nominal, ordinal, interval, or ratio depending on what that [0,1] means.

    Answer: The [0,1] could mean male-female, which is nominal. It could be ordinal when 1 means more than 0 for the indicator. It can be interval when a bunch of [0,1] indicators are added together to form an index. By adding [0,1] indicators one is converting the dichotomous indicators into interval ones. Finally, it could be ratio if the [0,1] indicate the real line segment between zero and one implying there is an underlying continuum between zero and one of real numbers.

363. Scaling issues are ubiquitous in datasets and indicator construction. Discuss the scaling description of a variable in a new education dataset: is it nominal?

    > 2.1 Existence of compulsory education (compuls_exist)
    >
    > Question: Is there a compulsory education system in place?
    >
    > Clarification: The system must require at least one year of compulsory education for children/youth.
    >
    > Answer categories:
    >
    > (a) No, there is no legal requirement that education must be compulsory
    > (b) Yes, there is a legal requirement of compulsory education in place for certain groups of the population (e.g., by ethnicity, gender or geographic region)
    > (c) Yes, there is a legal requirement of compulsory education in place for all children in the relevant age group

    Measurement level: Nominal (Lutscher et al. 2022, 3)

    Answer: It is clearly ordinal not nominal.

364. Because scaling is virtually never taught in political science methods classes most scholars doing quantitative analysis and creating indices are unfamiliar with how this works. Very common is the adding of ordinal measures or dummy measures to create an index. Discuss the problems with this in this particular example:

    > To elucidate the causes and consequences of intervention, MIP measures several national interests-related variables from 1776 to 2019. We apply a National Interests Index across the eras that adds up separate measures on contiguity, colonial history, alliances, and natural resources. This additive index contains relative measures of geopolitical importance between

> country-dyads, including factors such as geographic continuity, shared alliances, colonial history, and the presence of oil and gas. We calculate the ordinal index, using State B target data, as follows: OilDummy + ColonialDummy + AllianceDummy + ContiguityDummy. (Kushi and Toft 2022, 15)

Answer: They have converted the dummy variables into interval ones in the process of creating the index where now zero and one are not nominal but become interval when put into the final index.

365. One relatively popular way to code the number of deaths in conflict datasets is to use data ranges. This is not a good idea. What would be a better approach and why? Here is one example from a prominent conflict data set.

> The relevant variables from ECAV are deaths and event description. Ideally, one would measure lethality using death estimates. However, such data are not available in ECAV, which codes the death variable on an ordinal scale. The five levels begin with no deaths; fewer than 10; 10 to 100; 100 or more; and unknown, which represents "uncertainty over whether (or how many) deaths occurred" (Daxecker et al. 2019, 16).

Answer: By cutting them as ranges one is implying that all the cases the given range are somehow "the same." This has the additional disadvantage of requiring one to either do some kind of logit on the dependent variable side or some combination of dummy variables on the independent side. A much better approach would be to give whatever estimate the researcher has of the actual deaths with some measure or indication of how confident she is and the range of plausible values for that estimate. For example compare the COW data ranges for the militarized dispute data set with Gibler's approach of giving estimated values and ranges for the estimated value. UCDP often does similar things.

366. Healy (2019) has written an important book on effective communication of data in figures. A key point is to clearly communicate in the figure the relationship between X and Y. He uses an example given in the figure below with heavily skewed data for GDP per capita against life expectancy. This kind of skewness in data is quite common in the social sciences. He gives you two different plots and suggest that the logged data are better. Discuss the extent to which this is true.

Figure 52: Log transformations of data: life expectancy and GDP/capita, raw data

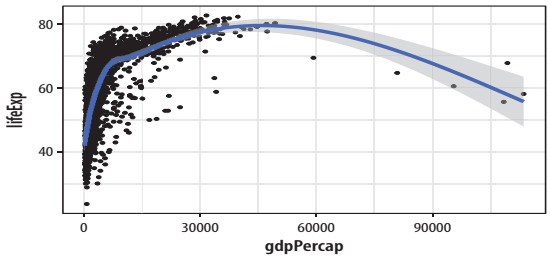


Figure 3.6: Life expectancy vs GDP, showing both points and a GAM smoother.

Figure 53: Log transformations of data: life expectancy and GDP/capita, logged data

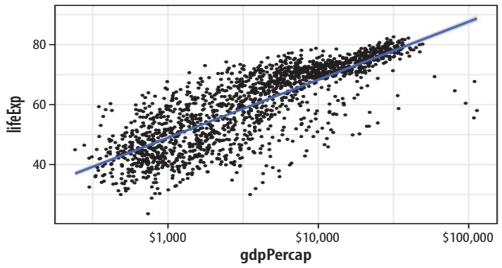


Figure 3.9: Life expectancy vs GDP scatterplot, with a GAM smoother and a log scale on the x-axis, with better labels on the tick-marks.

Answer: As Healy notes:

> If the relationship between X and Y is really nonlinear then logging to make it look linear is in fact falsifying in some ways the relationship between the two. In addition, for those interested it is not necessarily clear that logging is the best nonlinear fit for the data. There may be other nonlinear models that fit the data better. (Healy 2019, 61)

367. A critical issue in the democracy literature is the concept itself of "democratic backsliding." This is another nice example of a two-concept concept one part of which is "democracy" and the other part of which is "backsliding." Critical in this debate is that measures of democracy typically run from completely authoritarian to completely democratic, e.g., 0 to 1.0 for V-Dem. So does backsliding only apply to those that are mostly democratic or to all countries across the spectrum? Does backsliding simply mean a decrease in your authoritarian-democracy score? This can take place among very authoritarian governments as they move from, say, .30 to .10 on the V-Dem scale. The approach that many take exemplified by Miller's contribution to a symposium

(2024) is to dichotomize democracy and authoritarianism and only look at backsliding among the democratic group. This has the standard problem of all the dichotomous measures which is the sharp border between democracy and authoritarianism. So for example if a country moves from democratic to the authoritarian category, e.g., .60 to .40 on the V-Dem scale is backsliding, but .40 to .20 is not backsliding? Does that even constitute "backsliding" to begin with, if backsliding means "not large negative changes?" Does an "autocolpe" or presidential coup constitute backsliding? Discuss how valid it is for Little and Meng to include all countries, authoritarian and democratic, to analyze trends in changes in democratic backsliding.

Here is Miller on the topic:

> A specific focus on democracies is warranted for at least two reasons. First, the narrative of democratic decline that L&M critique always has been principally about democracies. . . . Moreover, decline within democracy is almost universally what scholars mean by "democratic backsliding." . . . Instead, the more common claim is that democratic quality is declining in the world's democracies. Consider the three examples that L&M (2023) cite in their introduction as emblematic of scholarly "alarm" around democratic backsliding. Of these examples, the Haggard and Kaufman (2021) study is entirely about democracies, whereas Diamond (2015) and Lührmann and Lindberg (2019) both emphasize that the global decline in democratic quality is minor but nonetheless worrisome because of its concentration in democracies. The democracy sample is defined [by Miller] using the Boix, Miller, and Rosato (2013) dataset [which has a dichotomous conceptualization of democracy]. (Miller 2024, 1–2)

368. One of the most important guidelines in chapter 3 of the concept book is that one should construct ideal type concepts and measures. This plays a critical role in analyzing Little and Meng's new measure of democracy. To begin with they use what they call a "quasi-minimalist" approach to democracy. More important are the scales for each of the indicators. As illustrated with the polity data in chapter 3 it is possible that the scales-conceptualization stops too soon or too low. This means that if many of the cases fall into the final top category that category is probably not capturing a lot of variation among those countries. In the case of polity this meant that if countries varied significantly from 10 to 15 on an ideal type polity scale: all of those from 10–15 score 10 on polity so important variation is not captured. This is critical in backsliding which means typically a modest reduction in the democracy score. So if the indicator question stops too soon it will not be able to capture significant backsliding, or for that matter increases in democratic quality. In their critique of Little-Meng, Knutson forcefully make this point: "First, across these indicators, the category/value with the greatest spread in democracy values is the top category/value. This demonstrates that these indicators have a very low threshold for considering a country as highly democratic. Would-be democratic backsliders thus have much room for deterioration before they receive a lower score" (Knutsen et al. 2024, 10).

369. In measurement with multiple dimensions, components or indicators *scaling* is a critical, but not often recognized factor. To combine or aggregate a bunch

of indicators typically means getting them all on the same scale. In Little and Mang they develop new concept-measure of democracy with a variety of indicators given in table **??**. In this table type indicates the kind of scale for each of these indicators. The note at the bottom of the table says that they are all rescaled in [0,1]. For example this means that dichotomous indicators are rescaled as interval with values between zero and one. Seven-point scales are linearly scaled into the same interval. Take one or two examples and discuss what is at stake by this linear rescaling of the variables. Note that this is required because they are going to take the average of these indicators and for the average to be valid requires interval data.

Knutsen et al. 2024 make the point that scale is not linear at all for some indicators: "Second, the ordered categorical variables that L&M use are not interpretable as interval-level values; the categories within each variable scale differently onto the democracy concept. Figure 7 illustrates this problem conceptually, using the codebook entry for DPI's legislative index of electoral competitiveness. The first five categories all relate to highly uncompetitive situations in which opposition parties won no seats; only the top two categories correspond to remotely competitive situations. (Knutsen et al. 2024, 10)

Table 15: Little-Meng indicators of democracy

Table 1

Indicators in L&M's Index

	Source	Concept	Type	%Missing
Proportion Suffrage	V-Dem	Suffrage	Proportion	0%
Presidential Vote	DPI	Exec. comp.	Proportion	64%
Incumbent Party Seat	DPI	Legis. comp.	Proportion	23%
Incumbent Party in Office	DPI	Exec. comp.	Count (max. 20)	33%
Legislative Competitiveness	DPI	Legis. comp.	7-pt scale	18%
Executive Competitiveness	DPI	Exec. comp.	7-pt scale	18%
Incumbent Party Loss	NELDA	Elec. comp.	Dichotomous	16%
Multiparty Competition	L&M/NELDA	Elec. comp.	4-pt scale	6%
Process Violations	L&M/NELDA	Elec. comp.	4-pt scale	6%
Term Limits	Meng	Exec. constraints	Dichotomous	22%
Succession Rules	Meng	Exec. constraints	Dichotomous	24%
Dismissal Rules	Meng	Exec. constraints	Dichotomous	26%

Note: All variables are rescaled 0-1 for the index.

370. Because scaling is virtually never taught in political science methods classes most scholars doing quantitative analysis and creating indices are unfamiliar with how this works. Very common is the adding of ordinal measures or dummy measures to create an index. Discuss the problems with this in this particular example:

> To elucidate the causes and consequences of intervention, MIP measures several national interests-related variables from 1776 to 2019. We apply a National Interests Index across the eras that adds up separate measures on contiguity, colonial history, alliances, and natural resources. This additive index contains relative measures of geopolitical importance between country-dyads, including factors such as geographic continuity, shared alliances, colonial history, and the presence of oil and gas. We calculate the

> ordinal index, using State B target data, as follows: OilDummy + Colonial-Dummy + AllianceDummy + ContiguityDummy. (Kushi and Toft 2022, 15)

Answer: They have converted the dummy variables into interval ones in the process of creating the index where now zero and one are not nominal but become interval when put into the final index.

371. One relatively popular way to code the number of deaths in conflict datasets is to use data ranges. This is not a good idea. What would be a better approach and why? Here is one example from a prominent conflict data set.

> The relevant variables from ECAV are deaths and event description. Ideally, one would measure lethality using death estimates. However, such data are not available in ECAV, which codes the death variable on an ordinal scale. The five levels begin with no deaths; fewer than 10; 10 to 100; 100 or more; and unknown, which represents "uncertainty over whether (or how many) deaths occurred" (Daxecker et al. 2019, 16).

Answer: By cutting them as ranges one is implying that all the cases the given range are somehow "the same." This has the additional disadvantage of requiring one to either do some kind of logit on the dependent variable side or some combination of dummy variables on the independent side. A much better approach would be to give whatever estimate the researcher has of the actual deaths with some measure or indication of how confident she is or the range of plausible values for that estimate (better). For example compare the COW data ranges for the militarized dispute data set with Gibler's approach of giving estimated values and ranges for the estimated value. UCDP often does similar things.

372. It is often the case that when one has an ordinal scale with more than, say, five categories the statistical analysis of this as an independent variable becomes problematic (e.g., lots of dummy variables). One possibly popular solution to this is to just dichotomize at some point in the ordinal scale. In the influential EPR data set discuss whether the seven point scale is ordinal or not? Does the fact that one can dichotomized imply that it is ordinal? Following the discussion in the concept book how can one do a semantic transformation of this seven-level scale into one continuous variable? Can you draw and justify the semantic transformation figure?

> I employ the EPR's [Ethnic Power Relations] main variable, "the degree of access to power enjoyed by political leaders who claimed to represent various" politically relevant ethnic groups. (They exclude any group whose members do not have elites at the center claiming to represent them.) Access to power is coded as a seven point categorical variable. Categories include "monopoly," "dominant," "senior partner," "junior partner," "regional autonomy," "powerless," or "discriminated." To examine the effect of a group's access to power and how it affects whether its members would execute a coup versus rebel, I collapse the seven-point variable into a dichotomous variable: Inclusion. Groupings with access to central state power (i.e. coded as "monopoly," "dominant," "senior partner," or "junior partner") are scored a 1, while all others ("regional autonomy,"

> "powerless," or "discriminated") are scored a 0. (Roessler 2016, 211–12; see http://www.epr.ucla.edu for the coding manual to help you decide whether it is ordinal or not).

Answer: The coding manual says this:

> We categorized all politically relevant ethnic groups according to the degree of access to central state power by those who claimed to represent them. Some held full control of the executive branch with no meaningful participation by members of any other group, some shared power with members of other groups, and some were excluded altogether from decision-making authority. Within each of these three categories, coders differentiated between further subtypes, including absolute power, power sharing regimes, and exclusion from central power.

This suggests an underlying concept which is "access to power." Clearly the three main "types" are ordinal. If one looks at the "subtypes" they also seem ordinal, e.g., senior partner versus junior partner.

373. Matching has become a very popular way to do statistical analyses with observational data. Matching relies on comparing treatment and control groups that are well matched on key control variables or covariates. To do this means putting the covariates on the same scale to allow one to evaluate the "distance" between observations based on the covariates. In the terms of the concept book this is scaling and then aggregation. Discuss the interesting procedure whereby aggregation is fixed by the distance measure and weighting is done via scaling. This contrasts with the procedure proposed in the concept book where scaling is done separately and weighting is done via the aggregation technique. What is the implicit weighting principle for decreasing the importance of a covariate in the quote below? What is the substantive implication in terms of the logging example?

> Adjusting Imbalance Metrics for Relative Importance. Although one can always define a data set that will produce large differences between any two imbalance metrics, in practice the differences among the choice of these metrics are usually not large or at least not the most influential choice in most data analysis problems. Although we describe continuous and discrete metrics below, more important is the choice of how to scale the variables that go into it. Indeed, every imbalance metric is conditional on a definition of the variables in X, and so researchers should think carefully about what variables may be sufficient in their application. . . . For example, consider data designed to predict the Democratic proportion of the two-party vote in a cross-section of Senate elections ("vote"), with incumbency status as the treatment, controlling for covariates population and the vote in the prior election ("lagged vote"). Clearly, lagged vote will be highly predictive of the current vote, whereas the effect of population will be tiny. Thus, in designing an imbalance metric, we want to be sure to match lagged vote very well, and should only be willing to prune observations based on population when we see gross disparities between treatment and control. For example, if we used Euclidean distance, we could code vote and lagged vote on a scale from 0 to 100 and population on a smaller scale, such as log population. In computing the Euclidean distance between two observations, then, population differences would

> count as equivalent to relatively small vote differences. (King et al. 2017, 478, 479)

Answer:

The example suggests that shrinking the scale, e.g., logging, reduces its impact. Logging makes some – not all – big real differences seem much smaller. But note that it also makes some small real differences seem larger.

374. Discuss the claim about bounded versus unbounded scales in Morris's *The measure of civilization how social development decides the fate of nations*. How does this relate to the Ideal Type Guideline. Clearly fuzzy logic semantic transformation transformations always have a maximum 1.0. How a serious critique of that approach is this?

> It [HDI] could, up to a point, be used to measure change through time by simply comparing a single country's score in each annual report, but because the maximum possible score is always 1.0, the HDI does better at charting a nation's relative position within the world at a single point in time than at measuring diachronic changes in development levels. (Morris 2013, 28)

375. Discuss the semantic transformation implied in this nonlinear categorization scheme.

> Following guidelines for ICEWS aggregation, we create seven count variables for conflict and cooperation among various parties: counts of high-hostility events (intensity: −10 to −8); moderate-hostility events (intensity: −7 to −4); low-hostility events (intensity: −3 to −1); neutral events (intensity: 0); low-cooperation events (intensity: 1 to 4); moderate-cooperation events (intensity: 5 to 6), and high-cooperation events (intensity: 7 to 10).6 We use data from all 1,122 municipalities for 20 years from 1993 to 2012. (Campbell et al. 2017)

376. Discuss the UNDP's semantic transformation and aggregation arguments:

> Societies can subsist without formal education, justifying the education minimum of 0 years. The maximum for expected years of schooling, 18, is equivalent to achieving a masters degree in most countries. The maximum for mean years of schooling, 15, is the projected maximum of this indicator for 2025. . . . Having defined the minimum and maximum values, the dimension indices are calculated as: Dimension index = (actual value minimum value)/ (maximum value minimum value) For the education dimension, equation 1 is first applied to each of the two indicators, and then the arithmetic mean of the two resulting indices is taken. Using the arithmetic mean allows *perfect substitutability* between mean years of schooling and expected years of schooling. Many developing countries have low school attainment among adults but are eager to achieve universal primary and secondary school enrollment. (UNDP 2018, technical-notes, 2; emphasis is mine)

377. Should power be considered a ratio variable?

> Power, however, is like love; it is "easier to experience than to define or measure." Just as one cannot say "I love you 3.6 times more than her," (Beckley 2018, 8)

Answer:

Power almost has to be considered a ratio variable because one wants to say that a country is twice as powerful as another. See, for example, the literature on the 3:1 rule of relative power. So one should be able to say that one person is 3.6 times more powerful.

378. Discuss whether the conflict–cooperation scale in this dataset (Boschee, E. et al. 2015) is interval or ratio. The ICEWS dataset is a product of automated text analysis of more than 30 million news stories over the past 13 years, with global coverage. The news stories are classified into event categories based on the CAMEO (Conflict and Mediation Event Observations) taxonomy, which provides information about the type of event and its characteristics relative to other events. Each event category in the CAMEO classification has a numerical value describing its intensity, ranging from negative 10 to positive 10, using Goldstein's (1992) Conflict-Cooperation Scale. The negative values signify the hostility of an event, with –10 capturing higher hostility. The positive values represent the cooperative nature of the event, with 10 capturing higher cooperation and 1 capturing lower cooperation.

Answer:

This depends mostly on whether the zero in the middle is a true zero or not.

379. Discuss how could one use the hierarchy to give weights to each dimension, is each a necessary condition condition (in theory or practice) for the next highest level:

> Even though the features in Table 1 are in theory independent of one another, they tend to cluster in practice, suggesting a hierarchical ordering of four dimensions: third-party review, third-party ruling, judges, and standing. The first question is whether the treaty provides for independent third-party review. Among pacts with some system of review, the next issue is whether rulings are directly binding in international law. Among pacts with binding rulings, those with standing tribunals are more legalistic than those with ad hoc arbitrators. Finally, tribunals with jurisdiction over claims by individuals, treaty organs, and states alike are more legalistic than those accessible only by states. In terms of remedy, the most legalistic pacts provide rulings with direct effect in national law, but the presence or absence of sanctions – though still significant – is a less meaningful indicator of legalism, with unilateral measures always available to states seeking to enforce third-party rulings in the decentralized international system. (McCall Smith 2000, 143)

Answer: This subset hierarchy allows one to use Guttman scaling techniques.

380. There are two questions. (1) Does the concept/measure have a zero point? (2) If so where is the zero point? It is an open question whether many concepts or measures of democracy are interval (i.e., no zero point) or ratio (zero point).

> [W]hile some regimes are more democratic than others, unless offices are contested, they should not be considered democratic. The

> analogy with the proverbial pregnancy is thus that while democracy can be more or less advanced, one cannot be half-democratic: there is a natural zero-point. (Alvarez et al. 1996, 21)

Discuss these two issues considering the quote above.

381. Often scholars pay no attention to the distinction between interval and ratio variables. Much of international conflict work implicitly works with ratio variables because of the contrast of, for example, between friends and enemies or satisfied and dissatisfied states. Often this means scales where –1 is enemy and 1 is friend, and where 0 is the dividing line. For interval measures there is no real zero. One can always rescale interval measures so that they do include zero, but that does not make the measure a ratio one. Discuss whether the widely used S measure (Signorino and Ritter 1999) is interval or ratio.

Answer:

The original scale of the S measure is $[0, 1]$. This is then transformed to $[-1, 1]$ in the data used in applications. The zero in $[-1, 1]$ does not appear to be the zero needed to distinguish between, say, friends and enemies. Also, the data are very skewed towards 1: the median is often in .75 range, hence there are very few negative values in most situations. See Goertz (2008) for an extended discussion. To make the $[-1, 1]$ S into ratio variable would require some theoretical and methodological argument for determining the zero point.

382. A very standard problem in measure creation is to combine the data on two or more secondary-level dimensions. Often this involves dimensions that have different scales and then some way of structuring them together mathematically. For example:

> we devise a new way to use the information of civil war duration and deaths for the two-stage regression analysis. The duration and casualties of civil wars are correlated, and it is theoretically inappropriate to estimate equation (2) for only one of them. Therefore, we combine the two characteristics of civil war into one, Civil War Destructiveness Index, by adding their standardized values. Before standardization, civil war duration is measured by the maximum duration in number of days of the war. Civil war deaths are measured by the total number of deaths experienced by the nation, divided by population. (Kang and Meernik 2005, 101)

Discuss (1) the importance of differences in variances between the two component parts in constructing the overall measure, (2) whether skewness in the data matter (often social science data are often heavily skewed), (3) the alternatives to standardizing the two scale, and (4) whether one should multiply or add the component parts.

383. An important aspect of all measure construction is standardization. Often this is done based on population, such as GNP/capita. Time is another common standardizing device, e.g., actions/year. Discuss the advantages and disadvantages of using time or the log of time to evaluate the extent of violations of the laws of war in Morrow and Jo (2006, figure 3). One might

feel that to compare the number of violations one needs to take into account the length of the war.

384. The left-right scale ("rile") of the Party Manifesto Dataset (Budge et al. 2001, Klingemann et al. 2006) is one of the most widely used measures in comparative politics. The rile scale is used to situate parties on the left-right continuum based on party manifestos. To do this, statements in the party manifesto are categorized according to 56 possible issue categories. The data report the percentage of all statements comprised by each category. If a category is not mentioned, the dataset reports that 0% of the statements is concerned with that category. A party's position is derived by adding the percentages of 13 categories seen as being left and subtracting the percentages of 13 categories seen as being right.

    Analyze how zero has two different uses in the coding and measurement of policy positions. Explain how this produces problems in the final measure. In particular discuss how aggregating the zeros on the different 26 policy positions becomes zero in the final left-to-right policy scale.

    Answer:

    Zero is used to indicate missing data. That is when one of the 56 issues categories is not present in the party manifesto it is given a value of zero. Zero also means the middle of the –100 to +100 scale. These are two fundamentally different things. See Rehm (2008) for an extensive discussion of this and other conceptual and measurement problems.

385. Zero points on scales have important roles. For example, in interpreting Cox regression model the impact of independent variables is often assessed vis-à-vis a "baseline" model where the independent variables have value zero. Discuss how this matters using the democracy and other key variables, such as level of development, in the Hegre et al. (2001) analysis of the "democratic civil peace."

386. Londregan and Poole (1996) transform the polity scale of –10 to 10 to $-\infty$ to ∞. They justify this because

    > The 21-point GOVTYPE [i.e., polity] scale is nearly continuous. However, because it has a "floor" of –10, and a "ceiling" of 10, we need to be careful not to use a statistical mode that predicts values outside this range. A standard way to ensure this does not happen is to apply a logistic transform to the variable. Let S denote a score on the 21-point scale. Take the following variant of the logit transform of S: T(S)=ln(S+10.5)–ln(10.5-S). This converts scores to a truly continuous scale. A value of S at 10.5 would correspond to a T(S) of ∞, an S of –10.5 to a T(S) at $-\infty$, while a value for T(S) of 3.71 corresponds to a regime type of 10. (Londregan and Poole 1996, 7)

    Consider the following questions: (1) does the fact that polity stops at 10 really mean that a country cannot be more democratic? (2) The transformation used implicitly suggests that the main shift from authoritarian to democracy occurs at zero on the –10 to 10 scale. Is this appropriate? When does polity code the

transition to democracy? (3) What the curve look like in the T(S) function? Does that correspond to a good concept of democracy?

Answer:

The concept book suggests that the concept of democracy should extend beyond 10 of the polity scale. Hence, there is perhaps no imperative reason to do the transformation in first place; that model predicts values greater than 10 is not problem since the polity concept extends beyond 10.

The transformation means that the model is predicting increasing marginal effects around zero, which may not be justified. It seems like the transition to democracy really occurs more toward 6 on the polity scale, and that is the official point where scholars dichotomize the polity measure. The Londregan and Poole transformation assumes a strongly decreasing marginal relationship between around 6–7, where in reality this might be where it is strongly increasing.

387. Are there situations where you really need ratio rather than interval data? For example, in conflict studies where one wants to distinguish between friends and enemies.

388. Sometimes the use of zero becomes a critical cut point. Bennett argues for a continuous specification of the democracy–autocracy continuum. He says that "In the specification advanced below, I avoid arbitrary cutoffs for democracy and autocracy" (2006, 320). Then he does the following "The first interactive similarity [of regime type] variable is constructed by multiplying the initiator's dem [polity democracy score] score by the target's dem score. Given dem scores ranging from –10 to +10, this multiplicative similarity variable will be high (large and positive, to a maximum of +100) for jointly autocratic and jointly democratic pairs. Similarity will be in a middle range, around 0, for pairs of mixed regimes (a dem score for either state near 0 produces a joint similarity score near 0). Similarity takes on its lowest values (to a minimum of +100) for pairs of dissimilar states (with a +10 and –10 dem score combination). If it is similar regimes that do not fight rather than jointly democratic pairs, then this variable will have a negative effect on conflict while other variables do not" (2006, 321). How might be argued to have used an arbitrary cut point with the zero?

Answer:

The construction of the interaction terms assumes that the polity measure is ratio. It could be argued that using the polity zero is arbitrary, particularly since the dichotomous coding of democracy for polity uses 6 or 7. One would need to make the argument, particularly in a fuzzy logic perspective, that zero is really the half-in, half-out point on the polity scale.

389. A very common problem is confounding "not applicable," "missing," and "zero," by using the same codes, often zero. Discuss some common examples of this such as: (1) trade where very small levels of trade might be missing or might be zero (e.g., Oneal and Russett (1999) assume missing data to be 0 in the trade relationship; see also Gleditsch 2002); (2) COW militarized dispute

data on fatalities, if you can find no reports on fatalities, should you code that as missing or zero. See also exercises 390, 570, 309, and 384.

390. A problem with zero codings is that it might seem like a real zero, but in fact is used for other purposes. Vreeland discusses in detail the anocracy coding of countries using the polity data: "PARCOMP [one of the five parts of the polity measure] ranges from repressed to competitive (numerically, from –2 to 3). What does the middle capture? PARCOMP is coded 0 for "not applicable situations, or, more specifically, the coding "is used for polities that are coded as Unregulated, or moving to/from that position (Gurr 1989, 14)" (2008, 406). Discuss the different things that zero means for PARCOMP.

    Answer:

    The PARCOMP scale goes from –2 to 3, this suggests that zero is a middle ground. In fact, it is also used to code "not applicable" which is certainly not in that middle ground.

391. Scholars often do not pay much attention to the difference between interval and ratio variables. Often methodologists and philosophers say that the higher the scale type the better, i.e., ratio is better than interval. This can conflict with the desire to log variables because of skewness. Using foreign direct investment data describe how scholars can implicitly convert ratio data into interval data.

    Answer:

    Raw foreign direct investment (FDI) data (e.g., in constant US dollars) can be positive or negative, hence there is a true zero which is meaningful and hence we have ratio data. Since one cannot take the log of negative numbers these most be converted into all positive numbers and thus the true zero in effect disappears making the data interval level.

392. *Social science concepts and measurement* takes a realist view of concepts that emphasizes what the theorist *means* by the concept. Many scholars will log variables because of skewness and many see skewness as a problem. Explain why an realist and one concerned by the meaning of concepts would object to Fearon's comments:

    > Still, as an index of overall ethnic diversity F [$F = 1 - \sum_{i=1}^{n} p_i^2$ where p_i are the population shares of each ethnic group] has much to recommend it.... And its empirical distribution – summarized in Figure 2 – is not highly skewed. (footnote: Cox (1997) and others sometimes prefer to use the "effective number of ethnic groups" (or political parties' vote or seat shares), which is $1 = (1 - F)$. Thus, a country with n equal-sized groups has an "effective number" of n groups, with departures from equal shares shrinking the effective number continuously. Although the interpretation is "nice," this measure is highly skewed, at least for ethnic fractionalization, so that it tends to exaggerate the influence of very diverse countries like Tanzania when used as an explanatory variable). (Fearon 2003, 209)

    Answer:

She would answer that the methodological cart is driving the conceptual and theoretical horse.

393. *The Bell Curve* (Herrnstein and Murray 1994) was a very controversial book. It turns out that the data on "intelligence" used were in fact not bell-shaped, but skewed to the right (i.e., quite a few very high scores). This skewness was "fixed" so that the data actually analyzed were bell-shaped. Discuss how this will influence their statistical results particularly given that they want to show that intelligence is a cause of outcomes such as income. It is important to note that it is completely standard procedure to assume a normal distribution in standardizing IQ data, for example, to make them comparative across nations or time.

    Answer:

    See Fischer et al. (1996) chapter 2 for a discussion.

394. Logging variables means that some variation in the data become much more important than other. When using raw FDI inflows instead of as a percentage of GDP, FDI is usually logged (Chan and Mason 1992; Wei 2000; Globerman and Shapiro 2003), though not always (Oneal 1994; Li and Resnick 2003). Explain which countries become more important with logged FDI.

    Answer:

    Logging compresses large values and expands low values (i.e., values near zero). So when using logged values the model is trying to explain more the variation in low levels of FDI than high levels.

395. Discuss the extent to which (if any) the two secondary-level factors-civil liberties and political liberties-of the Freedom House measure of democracy are equally weighted in the final measure. See http://www.freedomhouse.org/research/survey2005.htm.

    Answer:

    There is slightly unequal weighting because there are different numbers of indicators used in each secondary-level dimension and because the basic level is the sum of the two secondary levels.

396. Discuss how the political rights, civil rights and electoral democracy concepts and measures are made and how they relate to each other in the Freedom House framework. See http://www.freedomhouse.org/research/freeworld/2004/methodology.htm.

    Answer:

    From the Freedom House web site:

    > In addition to providing numerical ratings, the survey assigns the designation "electoral democracy" to countries that have met certain minimum standards. In determining whether a country is an electoral democracy, Freedom House examines several key factors concerning how its national leadership is chosen.
    >
    > To qualify as an electoral democracy, a state must have satisfied the following criteria:

> 1) A competitive, multiparty political system. 2) Universal adult suffrage for all citizens (with exceptions for restrictions that states may legitimately place on citizens as sanctions for criminal offenses). 3) Regularly contested elections conducted in conditions of ballot secrecy, reasonable ballot security, and in the absence of massive voter fraud that yields results that are unrepresentative of the public will. 4) Significant public access of major political parties to the electorate through the media and through generally open political campaigning.
>
> The electoral democracy designation reflects a judgment about the last major national election or elections. In the case of presidential/parliamentary systems, both elections must have been free and fair on the basis of the above criteria; in parliamentary systems, the last nationwide elections for the national legislature must have been free and fair. The presence of certain irregularities during the electoral process does not automatically disqualify a country from being designated an electoral democracy. A country cannot be listed as an electoral democracy if it reflects the ongoing and overwhelming dominance of a single party or movement over the course of numerous national elections; such states are considered to be dominant party states. Nor can a country be an electoral democracy if significant authority for national decisions resides in the hands of an unelected power, whether a monarch or a foreign international authority. A country is removed from the ranks of electoral democracies if its last national election failed to meet the criteria listed above, or if changes in law significantly eroded the public's possibility for electoral choice.
>
> Freedom House's term "electoral democracy" differs from "liberal democracy" in that the latter also implies the presence of a substantial array of civil liberties. In the survey, all Free countries qualify as both electoral and liberal democracies. By contrast, some Partly Free countries qualify as electoral, but not liberal, democracies.

397. Consider the "Competitiveness of political participation" dimension of the polity measure of democracy (Jaggers and Gurr 1995, 472) given in the table below. Consider how the negative pole (autocracy) and positive pole (democracy) are coded. Think about the universally used measure "Democracy minus Autocracy" and the scale values that produces. What is the asymmetry between the positive and negative poles? How would you make it symmetric? The same question can be asked of the "Constraints on the Chief Executive" dimension. Ignore the substantive nature of the dimensions and just look at how things are scored.

Table 16: The Polity numeric measure of democracy

Authority Coding	Democracy Weight	Autocracy Weight
I. Competitiveness of Political Participation		
(a) competitive	3	0
(b) transitional	2	0
(c) factional	1	0
(d) restricted	0	1
(e) suppressed	0	2

Source: Jaggers and Gurr 1995, 472.

Answer:

Once you do the Democracy-Autocracy operation, this dimension then goes, $-2, -1, 1, 2, 3$. Notice then that democracy can achieve a maximum of 3 while autocracy can only get a –2. Since there are five levels one could have done $-2, -1, 0, 1, 2$.

398. The concept book argues that basically all concepts should be considered continuous. However, there is an extensive literature on freedom and rights that argues that they are dichotomous (see Dowding and Hees 2003 for a continuous view). This may seem counter-intuitive since most people talk about "more or less" freedom; so the implicit view of most is that liberty and freedom are continuous. Notice that in the second quote he talks about "possible" and "impossible" as dichotomous. Contrast this with the Possibility Principle which clearly sees possibility as continuous. Notice that in Kramer's view "overall liberty" is continuous, but individual liberties are dichotomous.

> Are particular liberties similar to overall liberty in being scalar or partitive? In other words, can each particular freedom exist to varying extents? Can somebody be free-to-φ to a certain degree, and be less free-to-φ or more free-to-φ than somebody else? With only one minor qualification, the current essay will maintain that the answers to these questions are negative. The existence of any particular liberty, as opposed to the extent of anybody's overall liberty, cannot vary cardinally or ordinally. This essay sides firmly with Ian Carter, then, who writes that '[t]he freedom to do x is not a matter of degree; one either is or is not free *to do* x'. Carter aptly oppugns 'the claim that specific freedoms are a matter of degree (i.e., that one can be more or less free to do x)' (Carter, 1999, 228, 233, emphasis in original). Hillel Steiner robustly espouses a similar view: 'The notion of degrees of freedom to do an action is superfluous, misleading and descriptively imprecise' (Steiner, 1983, 78). (Kramer 2002, 231)
>
> In being non-scalar, 'free' and 'not free' in discussions of particular actions and types of actions are on a par with 'possible' and 'impossible'. When P has been able to perform X and has now become able to perform X in some novel manner, we cannot correctly state that his performance of X is now more possible than before. 'Possible' applies in an all-or-nothing fashion, as does 'impossible'. Instead of declaring that the performance

> of X by P is now more possible than previously, we ought to declare that the performance of X by P is now possible in more ways than previously. (Kramer 2002, 242)

399. Critical to the construction and analysis of concepts is the underlying scale or continuum from the positive to negative poles. One way to argue for an underlying continuum is if the data have the structure of a Guttman scale (Guttman 1944). The CIRI measure of human rights violations has four categories of violation: (1) torture, (2) political imprisonment, (3) extrajudicial killings, (4) disappearances. If the data fit the Guttman requirements then each category is a subset of the lower level categories. For example, if a state does (2) then does (1); if it does (3) then it also does (2) and (1); if it does (4) then it does (1)–(3). Cingranelli and Richards (1997) argue that their data on physical integrity rights has basically this structure. Draw a Venn diagram illustrating what this looks like.

    Cingranelli and Richards propose a quantitative measure where each of the four dimensions gets 0–2 on the level of violations in a given country in a given year; their final score is the sum of the scores of the four dimensions. Thus the worst performing states score 8, and the least human rights violating states get 0.

    Addition is one way to structure the concept. Another is the maximum. To make this simpler, assume that one can only score 0 or two on each dimension. Provide a rescaling of the four dimensions of the CIRI scale so that when you use the maximum to aggregate you would get *exactly* the same final score as the CIRI data *if* the data fit perfectly the Guttman requirements.

    One problem is that the data do not exactly fit the Guttman requirements. For example, there are cases with "extrajudicial killing" but no "political imprisonment." How would your proposed rescaling and use of the maximum deal with these cases? Do you think it is better or worse than using the sum.

    Woods and Gibney (2009) critique the CIRI scale because it would count the torture, political imprisonment, and disappearance of a single individual three times. How does your proposal using the maximum deal with or not this critique?

    Answer:

    Rescale the data so that (1) (i.e., torture) gets 0 or 2 points, (2) gets 0 or 4 points, (3) gets zero or 6 points, and (4) (i.e., disappearances) gets zero or 8. If the data fit perfectly the Guttman assumption, then if a country-year gets is scored for disappearances, it would have all levels (1)-(3) and hence a total of eight. With the new scoring system it would also have a score of eight. Note that things would get more complicated with a 0, 1, or 2 scale, instead of the 0 or 2, but the logic would be the same.

    The logic is then not to count multiple acts against an individual person separately but to only count the most severe act of physical integrity violation.

400. Often the line between nominal, ordinal and interval is blurred once one begins to do mathematical operations like addition or multiplication. Explain why the polity measure is interval, not ordinal. For example, see this discussion:

> One way to categorize this growing corpus of indicators is by the type of scale used to measure the key concept (democracy) – binary, ordinal, or interval. Binary indices include the democracy–dictatorship ("DD") index produced by Przeworski and collaborators (Cheibub, Gandhi, & Vreeland, 2010) and an index produced by Boix, Miller, and Rosato (2013, hereafter "BMR"). Ordinal measures include the Political Rights ("PR") index and the Civil Liberty ("CL") index, both produced by Freedom House (2013), along with the Polity2 index drawn from the Polity IV database (Marshall, Gurr, & Jaggers, 2013)" (Skaaning et al. 2015, 2).

Answer:

Once one begins to add (or subtract) one is implicitly converting the binary or ordinal variables into interval ones. This is classic measurement theory.

401. The AMAR "rebellion" scale (Birnir et al. 2018) has as its lowest rebellion level "political banditry and/or sporadic terrorism." Often zero on the scale (like in many dichotomous variables) is potentially a causally heterogeneous category. Discuss to what extent *theoretically* zero is conceived of as "no mobilization" as opposed to nonviolent mobilization such as protests or nonviolent action. Would the causal mechanisms leading to these two different forms of zero be the same or different?

402. As discussed in the (Goertz 2020) chapter on scaling, it is often not so clear what kind of scale is being used. Most often the problem is most severe between interval and ratio scales. However, it can occur as well between nominal and ordinal concepts.

Table 7.3 A single scale with multiple interpretations: "Electoral contestation"

0	**Authoritarianism**: No elections or elections with only one party or candidate.
1	**Semi-authoritarianism**: Elections in which more than one party or candidate runs, but not all parties and candidates face the possibility of losing.
2	**Semi-democracy**: Elections in which more than one party or candidate runs and all parties and candidates face the possibility of losing, but not all parties or candidates are allowed to participate.
3	**Democracy**: Elections in which only anti-system extremist groups are banned and all parties and candidates face the possibility of losing.

Source: drawn from Munck (2009: 45).

Source: Gerring 2012.

Gerring discusses table 7.3 in his measurement chapter in the section on scaling:

> Limiting himself to the concept of electoral contestation (a dimension of the larger concept of democracy), Gerardo Munck defines a four-part nominal scale including categories for authoritarianism, semi-authoritarianism,

semi-democracy, and democracy. Here, each category is distinguishable and clearly ranked relative to the concept of theoretical interest. . . . The unconstrained nominal-scale indicator is also more likely to approximate the virtues of an interval scale, where neighboring categories are equidistant from each other. . . . A final option for the four-point scale in Table 7.3 may also be envisioned. Insofar as the first category comprises a true zero – no contestation whatsoever – the key dimension of electoral contestation may be redefined as a ratio scale (an option that Munck endorses). (Gerring 2012, 170, 171)

Discuss the tension between a "nominal" scale with no underlying conceptual dimension (by definition of nominal scale) versus the common notion of a autocracy–democracy continuum.

Aggregation, concept structure

403. In causal model figures one needs to indicate aggregation decisions. Often there is a choice between "+" and the Boolean AND. Copeland uses "+" the figure below. What in the text justifies the "+"? Are there other parts of the text which would support an AND? Also, he uses two kinds of causal arrows: what is the difference between the two?

Figure 54: Aggregation and arrows in causal model figures: Copeland on US foreign policy and commerce

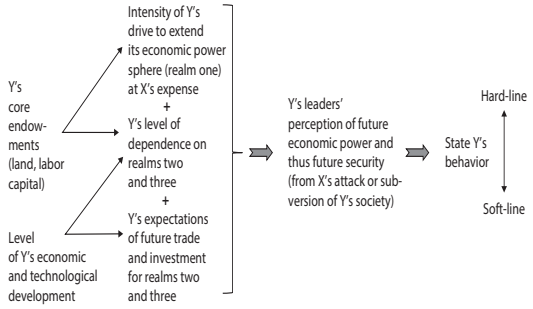


FIGURE 1.2. Three core causal factors

404. One of the core guidelines for creating causal mechanism figures discussed in the LNQA book manuscript is the need to provide what it calls aggregation criteria. This means when there are multiple arrows leading to a given box one needs to discuss whether those should be connected with AND, OR, + or whatever. Discuss what might be the appropriate additions to Loxton's theory that also might make it match his empirical data, see the figure below.

Figure 55: Aggregating factors: conservative party-building

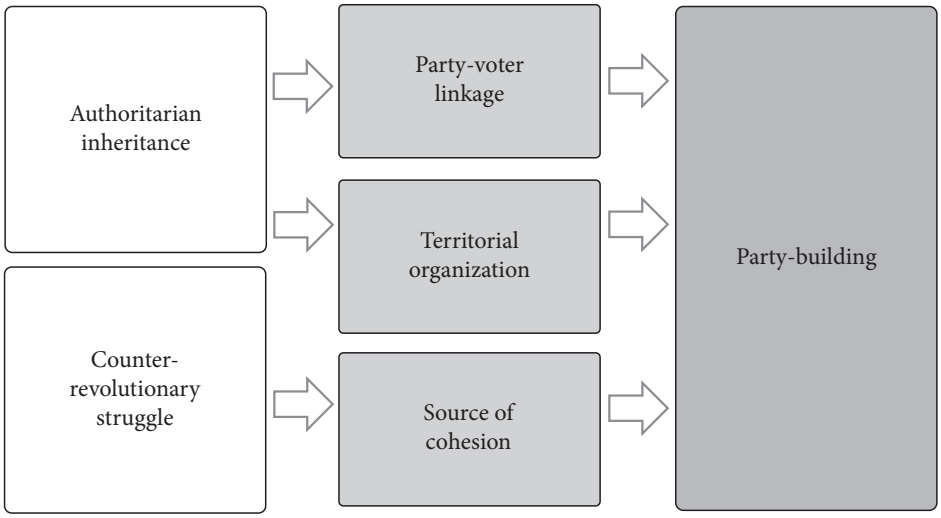


Figure 2.1 Model of the argument.

Source: Loxton 2021.

405. Aggregation plays an absolutely central role in conceptualization and measurement. Little and Meng (2024 see the symposium in *PS* regarding their article) have proposed a new measure of democracy and democratic backsliding. Aggregation plays two central roles at least in their project because they are interested in democratic backsliding at the global level which involves aggregation of national democracy scores as well as aggregation within national democracy scores based on 15 indicators. They use the average for both of these without any particular discussion probably, a resort to the standard default: "Taking Stock. To summarize the objective indicators and compare trends among different sets of countries, we construct a simple aggregate objective index by normalizing all individual variables between 0 and 1 when we can do so and taking the average for each country-year" (Little and Meng 2024, 6). Consider one of their key claims which is that backsliding at the global level has not really occurred over the last 10 years in spite of numerous claims to the contrary. Basically backsliding is defined as a decline in the democracy score forgiven country (See other exercises regarding their measure and backsliding in general). At the global level think about hypothetical examples where the value of zero backsliding at the global level could mean (1) significant backsliding or could mean (2) no backsliding at all. Hint: Remember that the average is an aggregation procedure where positive values can compensate for negative values (See Knutsen et al. 2024 in their contribution to the symposium for a discussion of the use of the average).

    Answer: Consider a bimodal world where half the countries experienced significant backsliding and the other half experienced significant increases in

democracy: the average is zero. Contrast that with a world where there is essentially no change in democracy levels at all: the average is zero.

406. Draw the concept figure for Beliakov's (2022) concept of "civilian control" of the military given in the table below, in particular what are the aggregation principles?

Table 17: Conceptualization of "civilian control"

Table 1. Operationalization of Civilian Control.

Attributes	Indicators (operationalization)	Authors
Subordination of the military to civilian authorities	Compliance with orders issued by civilian authority Timely and accurate implementation of orders The military reports to civilian authorities on major security-related events	Desch (1999), Feaver (2003), Huntington (1957), Pion-Berlin (1992), and Pion-Berlin and Martinez (2017)
Civilian authorities dominate the policy process	The government of the state (especially key ministerial positions) is composed of civilian officials Civilian expertise on security issues informs policy formulation Final decision-making power belongs to civilian authorities	Cohen (2002), Cohn (2011), Cottey et al. (2002), Feaver (1996, 2003), Kuehn et al. (2017), Perlmutter (1969) and Trinkunas (2005)
The military does not compete with the government for political power	Members of the military do not run for offices or compete for seats in the government They do not try to influence politics through blackmailing or challenging the government and affecting public opinion (media appearances, public addresses, etc.) They do not plot, try to perform, or assist a coup	Barany (2012), Brooks (2008, 2009), Croissant et al. (2010), Feaver (2003), Huntington (1957) and Kohn (2002)

407. Campbell et al. (2021) develop a concept and measure of "personal secularism" in their study of secularism in American politics. Draw the basic framework figure this three level concept. Assign the questions that are indicators to the various secondary level dimensions. Discuss the decision to use an additive substructure for both levels. It is the Redundancy Guideline relevant? Contrast the three level concept–measure with their latent variable statistical model.

> To create our Personal Secularism Index we have consulted the expansive body of writing espousing secular social and political thought, including the work of social theorists such as George Jacob Holyoake (1871), Immanuel Kant (1781 [1999]), and David Hume (1777 [2014]); the recent volumes by "new atheists" such as Richard Dawkins (2006), Sam Harris (2005), and Christopher Hitchens (2007); and the statements of belief found on the websites of secular organizations such as the American Humanist Association (AHA) and the International Humanist and Ethical Union (IHEU). Based on these sources, we have distilled three core

> principles common among Secularists. One core principle is a commitment to science and objective evidence as the basis for understanding the world. For example, Kant contends that "everything in the world happens solely in accordance with laws of nature . . . we have nothing but nature in which we must seek the connection and order of occurrences in the world" (1999, 485). In Humanism and Its Aspirations, published in 2003, the AHA contends that "knowledge of the world is derived by observation, experimentation, and rational analysis". . . . A second core principle of secularism is the view that only human experience and knowledge provide the proper basis for comprehending reality and making ethical judgments – in other words, "humanism." . . . . A third core principle of secularism is "freethought," the idea that human development and understanding should be based on logic and reason, rather than received authority, dogma, or tradition.
>
> We measure secular beliefs with a series of questions that gauge support for these core secular principles: the degree to which a respondent's perspective is informed by sources other than the supernatural (a term we use nonpejoratively). The core of our secular beliefs scale consists of eight statements, to which our respondents indicated their level of agreement. . . . To minimize response set bias, the questions were not all worded in the same direction. Five of the statements are worded to affirm secular perspectives: (1) Factual evidence from the natural world is the source of true beliefs, (2) The great works of philosophy and science are the best source of truth, wisdom, and ethics (3) To understand the world, we must free our minds from old traditions and beliefs (4) When I make important decisions in my life, I rely mostly on reason and evidence (5) All of the greatest advances for humanity have come from science and technology The other three statements represent the rejection of secular values: (6) It is hard to live a good life based on reason and facts alone (7) What we believe is right and wrong cannot be based only on human knowledge (8) The world would be a better place if we relied less on science and technology to solve our problems. . . . While the scales of nonreligiosity and personal secularism created using confirmatory factor analysis are the most methodologically sound, they correlate highly with simple additive indices (over .95). Therefore, in most cases we opt to use the additive indices, although in every case the results would be nearly identical with the scales created from confirmatory factor scores. The exception is in Chapter 5, where we examine how nonreligiosity, personal secularism, and a variety of political orientations are related to one another over time. For the models in that chapter, we employ the measurement error correction afforded by confirmatory factor analysis. If other scholars use our measures of personal secularism – which we hope they will – an additive index will suffice for nearly all purposes. (Campbell et al. 2021, 27–28, 36)

408. Often by default scholars linearly aggregate dimensions of a concept. Discuss how scatterplots might inform the aggregation decision and lead to a necessary condition structure. Luna (2014) provides a very nice example with two dimensions of party system institutionalization, electoral volatility and programmatic structure.

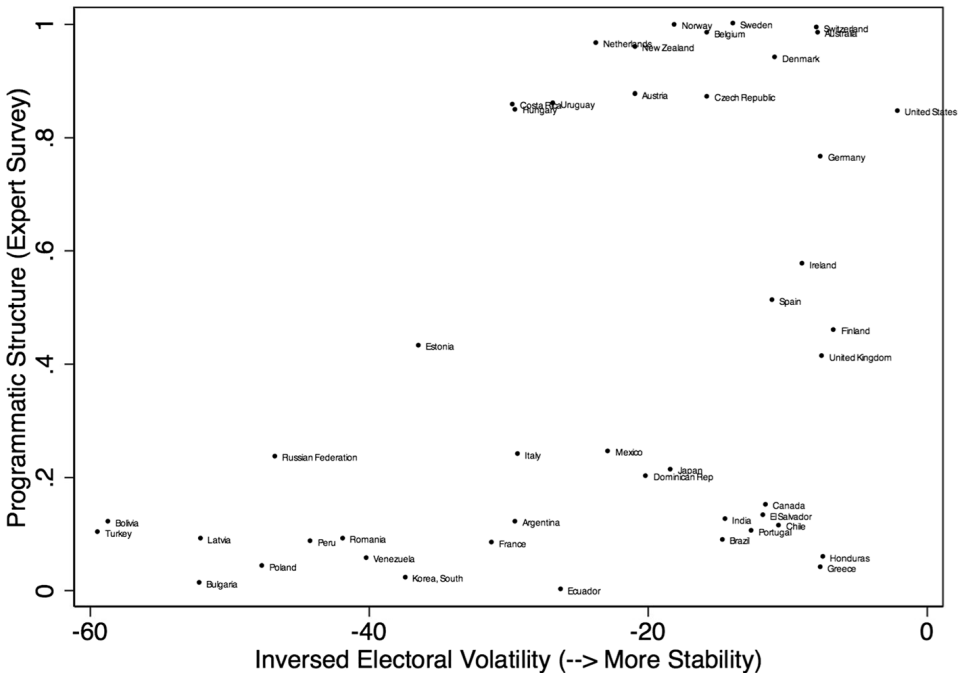


Fig. 2 Electoral volatility and programmatic structure (2000s). Source: Own construction on the basis of Scott Mainwaring, Annabela España, Edurne Zoco, and Carlos Gervasoni's electoral volatility database and data from the Democratic Accountability and Linkages Project (2009)

Answer: As Luna notes "As a result, collapsing the two phenomena in a linear (additive) latent variable measurement is misleading. Although future research is needed to explore the relation between the two dimensions unpacked here and the remaining theoretical dimensions and indicators included in Mainwaring and Scullys conceptualization of PSI, the analysis of the relationship between these first two dimensions presented in this paper already constitutes sufficient ground for claiming that the concepts structure (family resemblance) and aggregation rule (linear addition of dimensions and their indicators into a single index) need to be revised" (Luna 2014, 422; see also Coppedge 2012 chapter for necessary condition scatter plot with a necessary condition relationship between two latent variables)

409. How would could one interpret the clearly triangular relationship, see her figure, between these two conflict variables? Could this be interpreted in terms of constraints?

> Figure 0.1 gives a broader view. Here, I plot the relationship between scores on the Political Terror Scale (PTS, x axis) (Wood and Gibney 2010) and scores on the Sexual Violence in Armed Conflict dataset (SVAC, y axis) (Cohen and Nords 2014). Both the PTS and the SVAC are yearly indices assigned by coding qualitative reports. The PTS measures the overall level of government repression on a scale of 1 (very low violence) to 5 (widespread terror), focusing on lethal violence, torture, and detention. SVAC scores focus exclusively on sexual violence and are measured on a scale from 0 (very few reported cases of sexual violence) to 3 (widespread and/or systematic sexual violence). .... Second, as with the missing northwest quadrant of table 0.1 , there are zero governments whose highest level

of sexual violence exceeded their highest overall level of "terror" during this period. Governments that perpetrate relatively high levels of sexual violence in the context of low overall levels of violence are extraordinarily rare. The reverse situation – high overall levels of violence with low levels of sexual violence – is relatively common. (Hoover Green 2018, 11)

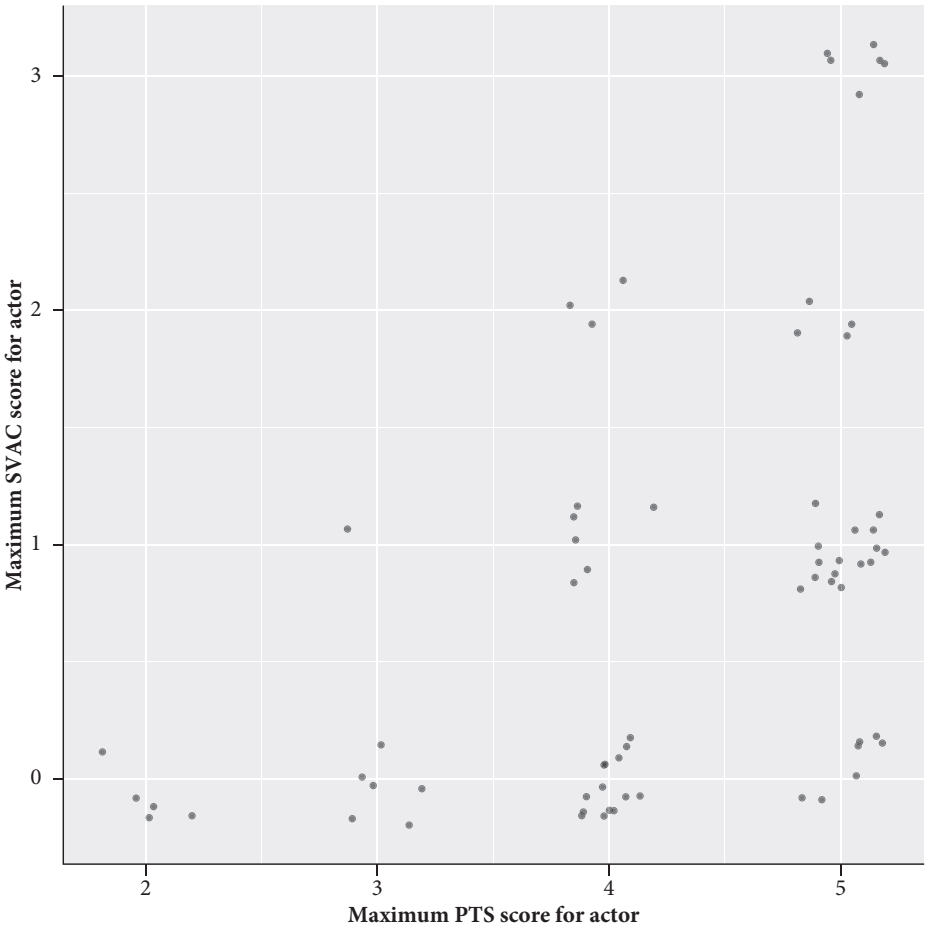


FIGURE 0.1 Comparing PTS and SVAC scores for seventy-five civil wars

Answer:

This illustrates a virtually perfect necessary condition scatterplot. So PTS is a necessary condition for SVAC. So PTS constrains the possible values SVAC can take.

410. Abrahamson and Carter have a core concept which is systemic instability which they argue is closely correlated to the making of territorial claims. They describe five forms that systemic instability can take. They create a composite measure using principal component analysis. Discuss how one might use the fact that clearly it is a logical OR or additive relationship between these five forms. What kind of aggregation operation and scaling might one use

to create this core variable and discuss the advantages and disadvantages of that versus a principal components analysis.

> Now that we have clarified why episodes of systemic and regional crisis are associated with claim making, we outline the main five "types" of events that generate such crises and how they work within the context of our theory: 1. When one or more great powers experience considerable domestic instability, this can generate systemic crisis. . . . 2. Wars involving multiple great powers often generate systemic turmoil. . . . 3. Major reorganizations of the international order, which usually stem from the settlement of major power wars, tend to leave serious bouts of systemic turmoil in their wake. . . . 4. The breakup of empires into smaller units also leads to systemic instability because it often involves a great power or large state breaking into constituent parts. . . . 5. Great power competition and conflict via proxy states can generate systemic turmoil. (Abramson Carter 2021, 113–15)

411. Best shot aggregation is often connected to thresholds. This might appear quite frequently in legal or bureaucratic contacts where the threshold then implies some legal responsibility. Discuss this aggregation technique using examples from international law.

> Following the set of "grave breaches" outlined in the 1949 Geneva Conventions, the treatment of enemy prisoners for capable captors in each war is examined across six separate dimensions: execution, torture, denial of legal rights, compulsory military conscription, hazardous labor, and poor housing and nutrition conditions. Other than a few exceptions, for each type of abuse the captor's treatment is classified according to one of three levels: high, medium, or low. . . . The individual component violations are then used to construct a single summary indicator for the level of prisoner abuse based on the following decision rule: the overall value of prisoner abuse is equal to the highest level of offense across the six dimensions. . . . This rule was adopted to hold captors to the stringent legal standards that prevailed, since any of these violations on their own constitutes a war crime punishable under international law. (Wallace 2015, 29–31)

412. Htun and Weldon (2018) create a series of multidimensional gender indices. They are all additive in terms of aggregation and are treated in linear terms in regression models. Discuss potential semantic transformations for these additive indices. Is there an argument for a semantic transformation is not linear? For example here is the one regarding violence against women:

> We examined each of these dimensions of government response to violence against women for all countries in our study to build an index to facilitate comparison. The index assigns higher values to those policy regimes that address more types of violence and whose actions span these categories of services, legal reform, policy coordination, and prevention of violence. This measure adapts the approach employed by Weldon, in order to take into account the varied types of violence that are salient in different contexts. Assessing this range of policies produces a score out of a total of ten points: Three points for services to victims . . . Three points for legal reform . . . One point for policies or programs targeted to vulnerable populations of women . . . One point for prevention programs . . . Our index sums these elements, so that more points imply more types of

> government response. The most responsive governments score a 10 and those that do nothing score a zero. Responsiveness means addressing as many of these dimensions as possible, both responding to current victims and preventing future violence. (Htun and Weldon 2018, 33–34)

413. Discuss the structure of the concept of "targeted mass killing" its structure (i.e., aggregation) and threshold. Would a best shot aggregation technique be more appropriate? The basic level concept is an ordinal one with 8 levels. Would a semantic transformation be appropriate here.

> Targeted mass killing is the direct killing of noncombatant members of a group by a formally organized armed force that results in twenty-five or more deaths in an annual period, with the intent of destroying the group or intimidating the group by creating a perception of imminent threat to its survival. A targeted group is defined in terms of political and/or ethnic and/or religious identity. . . . From this broad set of potential cases, a new episode of TMK was coded when there was evidence that an organized armed actor: (1) killed twenty-five or more civilians in a year, (2) these civilians were deliberately targeted by that actor) (3) one (or more) political, ethnic, or religious group(s) were disproportionately targeted, and (4) the group was targeted in order to substantially reduce its numbers, expel, or affect the political activity of that specific group. An active TMK episode was coded when all four of these criteria were met. . . . Table 1. A Targeted Mass Killing (TMK) Ordinal Scale. (Butcher et al. 2020, 1528)

414. Morris's wide ranging book *The measure of civilization how social development decides the fate of nations* critically uses a key measure of the concept of " social development." Below is an scaling–aggregation decision he faced. Discuss the advantages and disadvantages of these two aggregation techniques, sum then log versus log then sum. Morris focuses on the differences at the low end of the index, why would in general sum of logs show greater differences? Notice that the log in general accentuates differences on the low end.

> the simplest way to bring out the variations in premodern social development scores is by representing the data on a log-linear graph. I did this at several points in Why the West Rules. For now, adding up the scores on the four traits and then calculating the logarithm of the sum. This has its uses, but if we instead calculate the separate logarithms of the four traits and then add up the logs to produce a single score, we end up with a graph that is no longer dominated by the high energy capture scores. the two upper lines in figure 7.6 show the log of the sum of the traits for West and East, while the two lower lines show the sum of the logs of the traits. as can be seen, summing the logs of the traits produces curves that are more sensitive to even quite small pre-twentieth-century changes in organization, war making, and information technology. . . . The sum of the logs, while less intuitively obvious, is more sensitive to small premodern changes in organization, war making, and information technology. (Morris 2013, 250)

415. Discuss this philosophical claim that "natural kinds" cannot be formed via the logical OR (disjunction). This means that many social science concepts cannot be natural kinds.

> I considered the question of whether a natural kind could be structured disjunctively, and I argued that disjunctive categories do not correspond to natural kinds (see section 1.4). The main consideration I put forward for thinking so was that disjunctive categories are not projectible. It is not that philosophers have privileged access to the a priori or conceptual truth that natural kinds cannot be disjunctively structured. Rather, given that the whole point of discovering natural kinds is to discover sets of co-instantiated properties that are inductively linked to a series of other properties, disjunctive categories do not enable us to achieve this goal. Disjunction undermines projectibility, so it defeats the purpose of having natural kinds in the first place. Moreover, this conclusion is vindicated by the fact that disjunctive kinds are rarely if ever attested in actual scientific practice, and where they are, they appear to be challenged by scientists for sound reasons. (Khalidi 2013, 203)

416. In the literature on HDI and similar measures the question arises about equal weighting of the dimensions. Can making all the dimensions necessary be counted as equal weighting?

> Another way of justifying equal weighting was based on normative ideas. In particular, Haq claimed that the weights be determined 'on the simple premise that all these choices were very important and that there was no a priori rationale for giving a higher weight to one choice than to another' (Haq 1995b: 48). Similarly, Jahan explained: 'The variables [which should be read as "dimensions"] are given equal weights not because of simplicity, but because of the philosophical reasoning that all the components included in the HDI are equally important and desirable in their own rights for building human capabilities' (Jahan 2002: 155). (Hirai 2017, 85)

417. People often push for "integrated" development. Does suggest an additive concept structure or something else? Take the various measures of well-being or poverty as examples. See also the Women's Peace and Security Index: "The Sustainable Development Goals (SDGs) were adopted as an "integrated and indivisible" set of goals, and we sought to respect that principle by giving equal weight to each of the three dimensions in the WPS Index." (Georgetown University's Institute for Women, Peace and Security, PRIO 2017, 57)

418. Discuss the numerous scaling and aggregation decisions in the OECD approach to measuring well-being. For example:

> Box 1.1. The OECD approach to measuring well-being The OECD framework for measuring well-being was introduced in Hows Life? 2011. It builds on a variety of national and international initiatives for measuring the progress of societies, as well as on the recommendations of the Stiglitz, Sen and Fitoussi Report (2009) and the input provided by the National Statistical Offices (NSOs) represented in the OECD Committee on Statistics and Statistical Policy. Conceptually, the framework reflects elements of the capabilities approach (Sen, 1985; Alkire and Sarwar, 2009; Anand, Durand and Heckman, 2011), with many dimensions addressing the factors that can expand peoples choices and opportunities to live the lives that they value including health, education and income (see OECD, 2013a). This approach to measuring current well-being has several important features: It puts people (individuals and households) at the centre of

> the assessment, focusing on their life circumstances and their experiences of well-being. It focuses on well-being outcomes aspects of life that are directly and intrinsically important to people rather than the inputs and outputs that might be used to deliver those outcomes. For example, in the education dimension, measures focus on the skills and competencies achieved, rather than on the money spent on schools or the number of teachers trained. It includes outcomes that are both objective (i.e. observable by a third party) and intrinsically subjective (i.e. those where only the person concerned can report on their inner feelings and states), recognising that objective evidence about peoples life circumstances can be usefully complemented by information about how people experience their lives. . . . Taking these stocks as the primary measurement focus is in line with the recommendations of the Stiglitz, Sen and Fitoussi Report (2009) as well as several other recent measurement initiatives, including the UNECE-Eurostat-OECD Task Force on Measuring Sustainable Development (United Nations, 2009), the UNU-IDHP and UNEPs Inclusive Wealth Report (2012), the Conference of European Statisticians Recommendations on Measuring Sustainable Development (UNECE, 2014) and several country initiatives (e.g. FSO, 2015; Statistics New Zealand, 2011).

419. Discuss the UNDP's semantic transformation and aggregation arguments:

> Societies can subsist without formal education, justifying the education minimum of 0 years. The maximum for expected years of schooling, 18, is equivalent to achieving a masters degree in most countries. The maximum for mean years of schooling, 15, is the projected maximum of this indicator for 2025. . . . Having defined the minimum and maximum values, the dimension indices are calculated as: Dimension index = (actual value minimum value)/ (maximum value minimum value) For the education dimension, equation 1 is first applied to each of the two indicators, and then the arithmetic mean of the two resulting indices is taken. Using the arithmetic mean allows *perfect substitutability* between mean years of schooling and expected years of schooling. Many developing countries have low school attainment among adults but are eager to achieve universal primary and secondary school enrollment. (UNDP 2018, technical-notes, 2; emphasis is mine)

420. Race is notoriously a problematic concept. Discuss the history and methodology of determining race in the US Census. What are the normative implications of this? What is the aggregation principle behind the "one-drop rule"?

> The constructed, contradictory, and fickle nature of US racial classification is exemplified in the revisions of census categories. Since its inception in 1790, the decennial US census has counted the population by race/color. For much of the censuss history, race was assigned to individuals by an enumerator, characterized as unchangeable, and race mixture was ignored entirely. Yet, between 1850 and 1920, racial fluidity was acknowledged, and periodically enumerated with specificity, via fractional mixed-race categories: mulatto (half-black), quadroon (one-quarter black), and octoroon (one-eighth black). By 1930, these categories were removed, and mixed-race blacks, as well as individuals of mixed white-Asian parentage, were subjected to the one-drop rule and categorized with their minority race. (Davenport 2020, 223)

421. Walter (1997, 340–41) gives "at least" three conditions that "must hold" for a third-party guarantee to be effective in ending civil wars. Are each of these necessary? Is there substitutability between them?

422. Examine Hartzell and Hoddie's (2003) concept of "power sharing." How does substitutability play a big role in this concept?

Answer:

"An innovation of this study is to suggest that power sharing should instead be considered a continuous variable that ranges in value from zero to four with each increment representing an additional dimension of power sharing (political, territorial, military, and economic) specified in the rules governing the society. The greater the overall number of power-sharing dimensions specified, the more likely that peaceful relates among collectivities will endure." (Hartzell and Hoddie 2003, 321) This clearly indicates a family resemblance structure.

"A second reason that the inclusion of multiple dimensions of power sharing in a negotiated civil war settlement proves advantageous to the prospects of long-term peace is that it serves as a source of protection against the failure to implement any single power-sharing provision of the settlement." (Hartzell and Hoddie 2003, 321). This clearly indicates that substitutability is core to power-sharing.

One might test to see if any of the four measures is necessary. Also, one might look at different forms of substitutability in addition to the straight additive measure.

423. Substitutability places a key role in a couple of aspects of Gibson and Howard's (2007) analysis of the scapegoating of Jews in Russia in the 1990s. They note that Jews have not been scapegoated in the 1990s. (1) How generally would substitutability play a key role in thinking about scapegoating in general and why some plausible groups might not be targeted. (2) Use substitutability theory to suggest changes in the approach to the psychological variables, dogmatism, xenophobia, etc. used as part of the explanation of scapegoating.

Answer:

(1) It may be that there are multiple possible choices from scapegoats, i.e., they are substitutable. So, for example, failure to scapegoat Jews is because there is some substitute for them: "Finding whom to blame is not necessarily easy or even possible, but for many Russians it likely makes more sense to blame the Communists, the nomenklatura, the legacy of Communism, oligarchs, the West, etc., than to blame Jews." (2) It may be that you need only one of these psychological attitudes, e.g., dogmatism OR xenophobia OR etc. to exhibit scapegoat behavior.

424. Examine substitutability claims in Doyle and Sambanis's analysis (2000) of the effectiveness of UN peacekeeping.

Answer:

"All strategies should address the local sources of hostility, the local capacities for change, and the (net) specific degree of international commitment available

to assist change. One can conceive of these as the three dimensions of a triangle whose area is the 'political space,' or effective capacity, for building peace. This suggests that the dimensions substitute for one another, that is, more of one substitutes for less of another; less extreme hostilities substitute for weak local capacity or minor international commitment." (Doyle and Sambanis 2000, 781; see also the statistical model of table 4, p. 791)

However, "Thus, we theorize that the PB [peacebuilding] process is captured by: PB = IC [International Capabilities]*NLC [Net Local Capacity]." (Doyle and Sambanis 2000, 782)

425. Multiple necessary conditions are sometimes used in a theory to eliminate multiple ways of getting a zero on the dependent variable. Frequently the language of constraints is used, e.g., the absence of the necessary condition constrains the actor into a different path. The multiple necessary conditions thus eliminate substitutable options. Discuss how this works in Doner et al. (2005). Also, explain how the structure of their theory is very similar to Downing's (1992) on the emergence of democracy in early modern Europe?

426. Bennett (2006) and Werner (2000) uses different measures (and hence concepts) of "similarity" between two countries' regime types. Ignoring that Werner uses 5 polity indicators, compare the two similarity measures using the information in equation for the generalized mean:

$$\min(a,b) \leq (a*b)^{1/2} \leq \frac{a+b}{2} \leq \left(\frac{(a^2+b^2)}{2}\right)^{1/2} \leq \max(a,b) \tag{1}$$

In general will the Bennett measure be larger or small than Werner's? What would a table or figure comparing these two measures look like?

427. Brancati (2006) presents a very interesting argument for using the maximum. It combines case selection with the use of the maximum as an aggregation technique. What is the error she makes about the maximum as an integration technique?

> The analysis presented in this article is an ordered logit analysis since the dependent variables in this study are categories of conflict and rebellion ordered from low to high forms of conflict and rebellion. This type of analysis does not assume that the categories of intercommunal conflict and antiregime rebellion are equally spaced, although it does assume that the effects of the explanatory variables are the same for all categories of the dependent variables. For this analysis, I aggregate the group level data to the national level using the maximum level of antiregime rebellion among "at-risk" groups in a country per year, and the maximum level of intercommunal conflict among "at-risk" groups in a country per year as my dependent variables.
>
> I aggregate the data to the national level using the maximum value of conflict or rebellion in a country per year because this method of aggregation overcomes group-based selection bias in the MAR data set. Since the MAR data set does not measure antiregime rebellion and intercommunal conflict for all groups in a country but only "at risk" groups, any measure

> that aggregates across groups, such as the median level of conflict or rebellion in a country, is biased by the absence of groups not deemed "at risk." Using the maximum level of conflict or rebellion in a country overcomes this problem because it does not aggregate across groups. (Brancati 2006, 673)

Answer:

The maximum is an aggregation technique. Using the maximum means that groups not in the dataset are assumed to have lower values than ones in the dataset.

428. The concept book shows that the extension can go up with intension (i.e., adding secondary-level dimensions) for the family resemblance structure. Construct an example where it can go down.

Answer:

To do this create an example where the new dimension is very rare empirically.

429. If the extension of a family resemblance concept can go either up or down with intension (see the previous question) then can you provide a general rule from when each direction (i.e., up or down) is likely to happen?

Answer:

If the added dimension is empirically very common then extension is likely to go up. If the added dimension is empirically rare then extension is likely to go down.

430. Frequently ideal types involve a long list of characteristics. Why does that naturally lead to the use of family resemblance m-of-n procedures when constructing measures? Fearon (2003) might be an example.

Answer:

It is very unlikely that any given object will actually have all the characteristics listed as part of the ideal type. Hence, it is natural that one only requires a certain minimum number. One can perhaps generalize that when long lists of characteristics are given that family resemblance structures will be used, but that when the number is small, necessary and sufficient condition structures are more likely.

431. How does Sartori view the question of sufficiency in conceptualization and concept structure? See also Collier and Levitsky (1997) for the same question.

Answer:

He focuses his attention on the minimal necessary conditions and does not explicitly make claims about sufficiency. In one footnote he explicitly rejects sufficiency ideas. Collier and Levitsky make no claims about sufficiency and focus only on the minimum requirements for democracy.

432. It was suggested in the concept book that hybrid concepts – with some necessary and some facilitating conditions – might be good for idea types. Is this what Diamon and Morlino are suggesting?

> The five essays that follow are part of a collaborative effort, launched at a conference at Stanford University, to elaborate and refine the concept of democratic quality and to apply it to a series of six paired country comparisons. We asked each author to discuss a particular dimension of the quality of democracy such as freedom, the rule of law, vertical accountability, responsiveness, and equality (our own list, and by no means exhaustive). We wanted each author to explain how the dimension in question relates to other dimensions in our framework, to suggest possible indicators for measuring the dimension, to identify ways in which this element of democratic quality is subverted in the real world, and to offer (where possible) policy recommendations. Our full framework features eight dimensions: the five outlined above, plus participation, competition, and horizontal accountability. Other dimensions might include transparency and the effectiveness of representation. The different aspects of democratic quality overlap, however, and we choose to treat these latter two as elements of our principal dimensions. We attempt here to identify some of the ways in which the different elements of democracy not only overlap, but also depend upon one another, forming a system in which improvement along one dimension (such as participation) can have beneficial effects along others (such as equality and accountability). At the same time, however, there can be trade-offs between the different dimensions of democratic quality, and it is impossible to maximize all of them at once. In this sense at least, every democratic country must make an inherently value-laden choice about what kind of democracy it wishes to be. Talk of a "good" or "better" democracy implies knowing what democracy is. At a minimum, democracy requires: 1) universal, adult suffrage; 2) recurring, free, competitive, and fair elections; 3) more than one serious political party; and 4) alternative sources of information. If elections are to be truly meaningful, free, and fair, there must be some degree of civil and political freedom beyond the electoral arena so that citizens can articulate and organize around their political beliefs and interests. Once a country meets these basic standards, further empirical analysis can ask how well it achieves the three main goals of an ideal democracy–political and civil freedom, popular sovereignty (control over public policies and the officials who make them), and political equality (in these rights and powers)–as well as broader standards of good governance (such as transparency, legality, and responsible rule). (Diamond and Morlino 2004, 20–21)

433. The *Diagnostic and statistical manual of mental disorders* (DSM-V) provides many nice examples of how important decisions are made using mostly family resemblance concept structures. Choose some of the diseases on the list and analyze their structure.

434. In *Resort to Arms*, Small and Singer (1982, 210) defined a civil war as "any armed conflict that involves (a) military action internal to the metropole, (b) the active participation of the national government, and (c) effective resistance by both sides." This appears as a list concept. How can you ascertain, given that data are being collected, if the necessary and sufficient condition or family resemblance concept structure is being used? The same question can be asked of Sambanis (2004) who has an "operational definition" of civil war.

    Answer:

Typically it is easy to see the necessary and sufficient condition structure even though it is not explicitly stated. If you examine how the data are actually collected almost always failure on one item of the definition means that the case is excluded. This means that each item is in fact a necessary condition.

435. Examine Hartzell and Hoddie's (2003) concept of "power sharing" and the resolution of civil wars (one only need read pp. 318–21). What is the structure of the concept (necessary and sufficient condition or family resemblance)? How does substitutability play a big role in this concept?

Answer:

"An innovation of this study is to suggest that power sharing should instead be considered a continuous variable that ranges in value from zero to four with each increment representing an additional dimension of power sharing (political, territorial, military, and economic) specified in the rules governing the society. The greater the overall number of power-sharing dimensions specified, the more likely that peaceful relates among collectivities will endure." (Hartzell and Hoddie 2003, 321) This clearly indicates a family resemblance structure.

"A second reason that the inclusion of multiple dimensions of power sharing in a negotiated civil war settlement proves advantageous to the prospects of long-term peace is that it serves as a source of protection against the failure to implement any single power-sharing provision of the settlement." (Hartzell and Hoddie 2003, 321). This clearly indicates that substitutability is core to power-sharing.

436. Discuss the King and Murray's measure of "human security". (1) Compare their analysis to Nussbaum's view of human well-being (2) Evaluate their claim that they do not need to justify their weights because they are not using any.

> According to our definition, a person is in a state of generalized poverty whenever he or she dips below the pre-defined threshold in any of the component areas of well-being. Our dichotomization of each component of well-being is based on the belief that there is a qualitative difference in life experience above and below the threshold. For example, the difference between not having enough food and nutrition to survive and having enough food is fundamentally different from the difference between having enough food to survive and having food that also tastes especially good.
>
> A key advantage of our definition of generalized poverty is that it does not require a set of weights to be developed to equalize the different domains of well-being. Since a person missing even one of these essential elements for any part of a year would be considered impoverished, the only arbitrary element in the definition is the threshold for each domain of well-being. Moreover, the policy world has much experience with choosing threshold values for income and many other areas (such as to decide whether individuals qualify for certain programs). Although these thresholds are arbitrary and can be improved in theory in some ways (at the cost of simplicity), they are frequently used because they are fairly accurate reactions of peoples life experiences and are simple to use. In

> addition, small changes in these thresholds do not always produce as large changes in population-based indexes as weights would in an aggregate well-being index.
>
> For example, we treat both (1) being tortured and (2) being tortured and starving, as impoverished and unacceptable conditions. Condition 2 may be harder to remedy than 1, but we do not have to decide how much worse 2 is than 1 in order to decide that the person is experiencing a state of generalized poverty. Similarly, few would argue that persons to be tortured four times in the next year are secure no matter how high their income. Rather, being tortured in the next year will put them in a state of deprivation or generalized poverty. The prospect of this outcome makes them insecure today. Of course, for analytical purposes other than defining human security, definitions of generalized poverty that include trade-offs between the level of achievement in one domain of well being versus another may be appropriate, particularly when individuals freely choose to balance some domains of well-being against others. Since we do not need to create and justify weights in combining domains, we can include as many other domains as the international community agrees on. For example, we can include education as a domain of well-being, even though it was once not considered an essential element for having a minimal level of well-being. (King and Murray 2002, 594–95)

Answer:

Clearly they take a necessary condition view of the "human security" concept. They are in fact weighting each dimension, and could use fuzzy logic weights if they wanted.

437. Analyze in three-level terms Vanhanen's concepts of "ethnic heterogeneity." In particular, contrast the issue of redundancy (or substitutability) at the indicator level and the secondary level.

> The measurement of ethnic division therefore is based on three types of ethnic groups: (1) ethnic groups based on racial differences, (2) ethnic groups based on linguistic, national or tribal differences, and (3) ethnic groups based on stabilized old religious communities. Thus we have three operationally defined indicators to measure three dimensions of ethnic division. In each dimension, the level of ethnic division will be measured by the percentage of the largest ethnic group of the country's total population. Together the three percentages measure the relative degree of ethnic homogeneity, and the inverse percentages measure the degree of ethnic heterogeneity. The three inverse percentages of ethnic heterogeneity are combined into an Index of Ethnic Heterogeneity (EH) by summing the three percentages. This index is used as the principal operational substitute for the hypothetical concept "ethnic division." (Vanhanen 1999, 59)

Contrast that with his three-level concept of "ethnic conflict." What is the three-level structure here? In particular, in terms of combining the two secondary-level dimensions of "violent ethnic conflict" and "institutionalized ethnic conflict" what is the role of redundancy?

Answer:

Vanhanen justifies his procedure:

> Is it justified to combine the three dimensions of ethnic heterogeneity into an index in this way or do the dimensions differ too greatly from each other? My argument is that different forms of ethnic division strengthen each other and that, it is justified to sum the percentages. For example, if the population of a country is divided into separate ethnic groups not only by language or nationality, but also by clear racial or religious division, the degree of ethnic heterogeneity is higher than in a country in which the population is divided into separate groups only by language or religion. Besides, I am not especially interested in the affects of any particular ethnic division, I am interested in the total level of ethnic heterogeneity. (Vanhanen 1999, 59)

If one thinks that redundancy is acceptable because each factor reinforces the other, one might think of multiplication instead of addition at the secondary level. Also, one wonders why the argument he uses at the secondary level cannot also be used at the indicator level where he applies the maximum, thus signaling complete substitutability.

When combining his two secondary-level factors of ethnic conflict he adds them. But since he argues that they are two indicators of the same thing, the mean would be more appropriate.

> Because the two scales of institutionalized and violent ethnic conflict are intended to measure the same phenomenon from two different perspectives, it is reasonable to combine the scores of the two dimensions into a combined Index of Ethnic Conflict (EC). It is done simply by adding the scores. This means that the value of EC can vary between 0 and 200. This index is used as the principal empirical indicator of ethnic conflict in this study. (Vanhanen 1999, 62)

438. Describe Kramer's (2002) view of the concept of "liberty" or "freedom" as a three-level concept, where the levels are (1) overall liberty, (2) individual liberty, and (3) token liberty. What is the structural principle used at each level?

Answer:

The secondary-level dimensions of individual liberties which combine via family resemblance to constitute overall liberty. This produces a continuous basic-level concept, e.g., 3/5 particular liberties defines .60 as the level of overall liberty. Token liberties – data/ indicator level – are specific occasions where individuals can have or not (dichotomously) liberty. Kramer argues that if an individual token has liberty then the secondary-level dimension has liberty (p. 240).

439. The concept of customary international law has provoked much discussion. Traditionally the concept has two secondary-level dimensions "state practice" and *opinio juris* ("the practice is required by, or consistent with, prevailing international law" (Kirgis 1987, 146). Thus international law faces the same problem as concept-builders in thinking about how to structure the relationship between these two defining dimensions. Discuss Kirgis's approach, it is necessary condition or family resemblance?

Answer:

Kirgis clearly argues that there cases where if the weight of one secondary-level dimension is larger enough it can *substitute* for lack of state practice. Human rights law would be an example.

440. "Any definition that requires a combination of characteristics captures only a subset of most of the identities that comparative political scientists classify as ethnic" (Chandra 2008, 413). What is the implicit assumption about concept structure that Chandra is making if this statement is to be true.

Answer:

She is clearly talking about a necessary condition structure: "Although I have discussed each characteristic individually, most definitions of ethnic identity incorporate various secondary characteristics in addition to the primary characteristic of descent. Any definition that requires a combination of characteristics captures only a subset of most of the identities that comparative political scientists classify as ethnic. This is because, as shown above, each characteristic, taken singly, captures only a subset of these identities. Because any one of the defining characteristics discussed above captures at best a subset of the classification that we started with, any combination of characteristics will capture a still smaller subset. Figure 1 illustrates this point" (Chandra 2008, 412–13).

441. Fearon (2003) has proposed an important concept of "ethnic group." Here is the core of his concept. What is the structure used for this concept?

> Another approach to definition – in several ways more useful for the purpose of constructing a list by countries – is to employ the idea of "radial categories" advanced by linguists and cognitive scientists (Lakoff, 1987; see Collier and Mahon, 1993, for a discussion with respect to political science). In practice, people may understand the meaning of a concept X by reference to prototypical cases. Less prototypical cases may not share all the features of a prototype, and yet still be validly classed as Xs, at least in some circumstances. For example, the prototypical ethnic group has the following features: 1. Membership in the group is reckoned primarily by descent by both members and non- members. 2. Members are conscious of group membership and view it as normatively and psychologically important to them. 3. Members share some distinguishing cultural features, such as common language, religion, and customs. 4. These cultural features are held to be valuable by a large majority of members of the group. 5. The group has a homeland, or at least "remembers" one. 6. The group has a shared and collectively represented history as a group. Further, this history is not wholly manufactured, but has some basis in fact. 7. The group is potentially "stand alone" in a conceptual sense – that is, it is not a caste or caste-like group (e.g., European nobility or commoners). . . . In assembling the list discussed below, I am looking for groups that meet the "prototype" conditions as much as possible. This implies that I allow groups distinguished from others in the same country primarily by religion provided that they meet condition 1 (membership has a strong descent basis) and condition 2 (self-consciousness as group). It also implies that I do not count castes in South Asia as ethnic groups, even though I readily admit that they share an important "family resemblance" to ethnic groups through the descent criterion, and could be validly considered as

> ethnic groups in some research designs (Horowitz, 1985; Chandra, 2000). I believe that a large majority of the groups in the list discussed below meet the conditions for a "prototypical" group fairly well, although for a number of cases, especially in Asia and Africa, the extent to which 2, 4, and 6 are met is unclear. These continents have many groups that are identified by some language commonality, which in most cases does mark some cultural similarity. But the extent of their "groupness," or sense of common identity (conditions 2, 4, and 6) is not clear from the sources I have been able to consult. (Fearon 2003, 201–2).

Answer:

It appears that he is using an implicit family resemblance structure in that not all seven characteristics must be present to be coded as an ethnic group.

Chandra and Wilkinson have this to say about Fearon: "In an improvement over the other two datasets, Fearon (2003) does provide a conceptual justification for his count. He attempts to include groups that fulfill as many of the following prototypical criteria as possible: First, membership is reckoned primarily by descent. Second, members are conscious of group membership. Third, members share distinguishing cultural features. Fourth, these cultural features are valued by a majority of members. Fifth, the group has or remembers a homeland. Sixth, the group has a shared history as a group that is "not wholly manufactured but has some basis in fact. Seventh, the group "is potentially stand alone in a conceptual sense – that is, it is not a caste or caste-like group" (p. 201). . . . Nor is it clear how many prototypical criteria a group must satisfy to be included or how a coder should decide between multiple candidate groups on multiple dimensions that fit the prototypical criteria. Why, for instance, was the category Jat (included in Atlas Narodov Mira but not in Fearon's work [2003]), which appears to meet the first six criteria but not the seventh, not chosen over the category Punjabi, which appears to meet the first, third, fourth, and fifth criteria but not the second, sixth, and arguably, seventh? Some of the groups included in Fearon's count do not meet several of the conditions. Hindi speakers, for instance, are not a group in which members are conscious of group membership, share distinguishing cultural features that are valued by a majority of members, and have or remember a homeland" (2008, 531).

442. Critical to the construction and analysis of concepts is the underlying scale or continuum from the positive to negative poles. One way to argue for an underlying continuum is if the data have the structure of a Guttman scale (Guttman 1944). The CIRI measure of human rights violations has four categories of violation: (1) torture, (2) political imprisonment, (3) extrajudicial killings, (4) disappearances. If the data fit the Guttman requirements then each category is a subset of the lower level categories. For example, if a state does (2) then does (1); if it does (3) then it also does (2) and (1); if it does (4) then it does (1)–(3). Cingranelli and Richards (1997) argue that their data on physical integrity rights has basically this structure. Draw a Venn diagram illustrating what this looks like.

Cingranelli and Richards propose a quantitative measure where each of the four dimensions gets 0–2 on the level of violations in a given country in a

given year; their final score is the sum of the scores of the four dimensions. Thus the worst performing states score 8, and the least human rights violating states get 0.

Addition is one way to structure the concept. Another is the maximum. To make this simpler, assume that one can only score 0 or two on each dimension. Provide a rescaling of the four dimensions of the CIRI scale so that when you use the maximum to aggregate you would get *exactly* the same final score as the CIRI data *if* the data fit perfectly the Guttman requirements.

One problem is that the data do not exactly fit the Guttman requirements. For example, there are cases with "extrajudicial killing" but no "political imprisonment." How would your proposed rescaling and use of the maximum deal with these cases? Do you think it is better or worse than using the sum.

Woods and Gibney (2009) critique the CIRI scale because it would count the torture, political imprisonment, and disappearance of a single individual three times. How does your proposal using the maximum deal with or not this critique?

Answer:

Rescale the data so that (1) (i.e., torture) gets 0 or 2 points, (2) gets 0 or 4 points, (3) gets zero or 6 points, and (4) (i.e., disappearances) gets zero or 8. If the data fit perfectly the Guttman assumption, then if a country-year gets is scored for disappearances, it would have all levels (1)-(3) and hence a total of eight. With the new scoring system it would also have a score of eight. Note that things would get more complicated with a 0, 1, or 2 scale, instead of the 0 or 2, but the logic would be the same.

The logic is then not to count multiple acts against an individual person separately but to only count the most severe act of physical integrity violation.

443. Pevehouse in an important study of the impact of regional organizations on democratization wants to measure the "democraticness" of a regional organization. This is defined as the number of member states that are democracies. Of course a given country is likely a member of numerous regional IGOs, the question is then is how to aggregate to give one number for the variable: "I use only the most democratic organization to measure each state's IO involvement (versus an average of all IOs) since it should take only one membership to supply any of the causal mechanisms posited by my theory" (Pevehouse 2005, 70). Discuss this rationale for using the maximum as an aggregation or structural principle.

Answer:

One might argue that it is the "net effect" of democratic and authoritarian IGO memberships. Also, it is possible that multiple memberships in IGOs might have an increased effect. Hence one could test variables such maximum authoritarian membership in regional IGOs, or sum of democratic IGO memberships.

444. Abbott, Keohane, Moravcsik, Slaughter, and Snidal (2000) published a very influential article on "The concept of legalization." This article could serve

as a good example for the discussion of concepts. (1) Draw a diagram that represents the three-level structure of their concept. The three secondary-level dimensions are quite clear and the indicators are given in various tables. (2) What is the structure that they implicitly use to connect secondary-level dimensions and indicators? (3) Thinking about the negation of a concept is important; what about the negations given in figure 1 as well as table 1? (4) "Delegation" (table 4) might be separated into two, or even three, separate dimensions? Are "dispute resolution," "rule-making" and "implementation" part of the higher level concept of "delegation?" (5) In the delegation dimension the concept of "binding" appears frequently, but is also central to the "obligation" dimension. Is this a problem?

Answer:

It seems like the implicit structuring used at both levels is the family resemblance one. There is not much sense that any of these are necessary. There are potentially other negations that could be used in figure 1, such as "nonbinding" for obligation. For precision it is interesting that the positive pole uses the term "rule" while the negative pole uses the term "norm." Delegation is problematic because it includes both implementation and arbitration. In terms of figure 1, "international court" suggests third-party dispute settlement mechanism and so the opposite might be bilateral agreements. Much of the problem with delegation is that this is where most of the *organizational* dimensions of legalization appear; but it is not obvious that one would lump administrative, monitoring, and sanctioning organizations together with courts (i.e., one would not do this in an analysis of domestic governance and legal systems.) Also missing is a potential dimension about *who* makes international rules, e.g., states, United Nations, IGOs, etc.

445. Brown (2010) proposes a measure for measuring IGO delegation and then applies it to two IGOs. Table 1 has the standard three-level concept structure. He defends a additive index of all the indicator-level variables. Discuss alternative aggregation procedures. Different indicators have different maximum possible values which implies a weighting scheme. The additive index assumes substitutability and equivalence, theoretical and empirical, of the various indicators. Discuss this weighting and aggregation scheme.

446. Bühlmann et al.'s (2012) Democracy Barometer uses an interesting graphical device to show how countries score on the nine key components of the concept, reproduced in the figure below. Propose how one could use this to give a measure from zero to 1.0 of the overall qualitative of democracy. Could you extend this to 9-dimensional space?

Figure 56: Measuring the quality of democracy

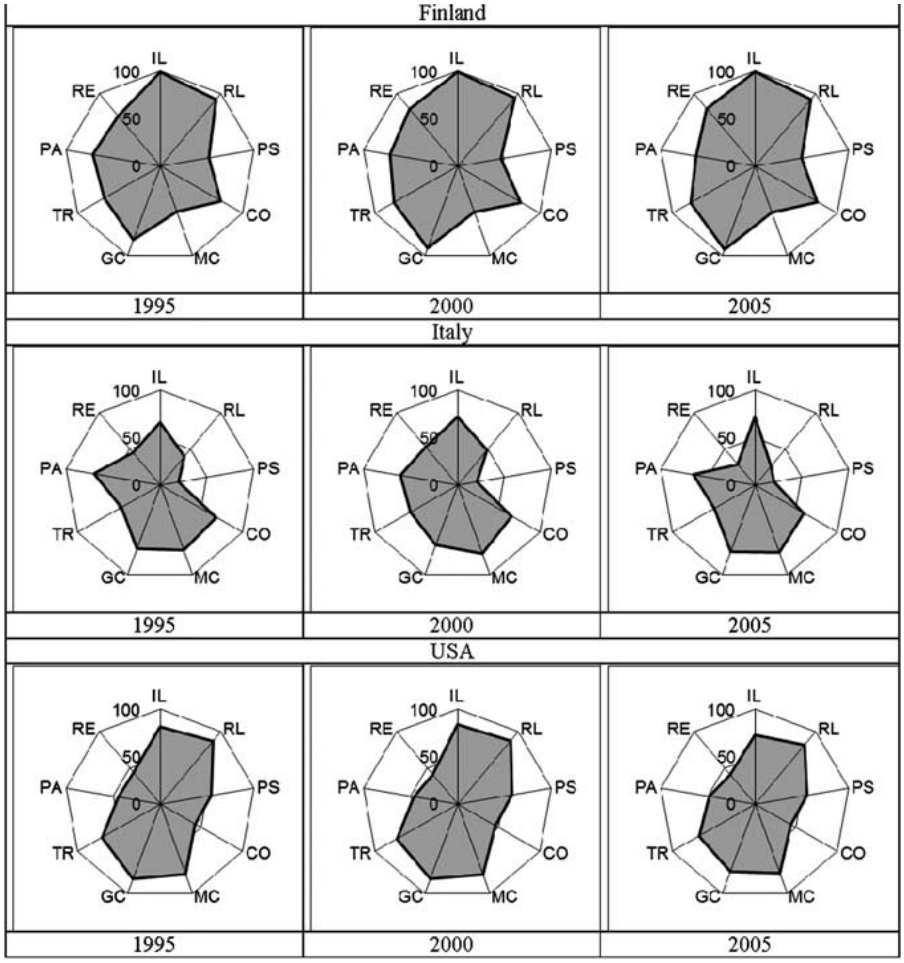


Figure 2 Shapes of democracy, three countries compared. IL: Individual Liberties; RL: Rule of Law; PS: Public Sphere; CO: Competition; MC: Mutual Constraints; GC: Governmental Capability; TR: Transparency; PA: Participation; and RE: Representation.

Source: Bühlmann et al. 2012.

Answer:

The overall measure would be percentage of the polygon covered in gray. An extension to nine-dimensional space would be the volume of the convex surface defined by the nine points (one can visualize this with a cube).

447. Call (2010) discusses the literature on the important and large concept of "failed state." He argues that there are three large secondary-level dimensions to this concept conceptualized by three "gaps": (1) capacity gap, (2) security gap, and (3) legitimacy gap. He illustrates this via a triangle with each dimension at a corner What is an alternative way to illustrate this? What are its advantages and disadvantages vis-à-vis a triangle figure?

Answer:

A perhaps more natural way to illustrate this is via a cube. If one tries to locate a country within a triangle it will be very difficult to illustrate its distance from each corner, because one is trying to force three dimensions into two. Also, the three-dimensional version of a triangle is often a pyramid. This is not good because one cannot be high (or low) on one dimension and low on another.

448. Skaaning et al. make the claim: "Our claim is that once a minimal level of inclusion has been attained – sufficient to constitute an electorate and hence the precondition for an election, as discussed – further increases in suffrage are irrelevant unless and until elections are competitive. This argument will be taken up below." (Skaaning et al. 2015, 2). Discuss how aggregation and substitutability are explicitly or implicitly involved conceptualization and measurement of this "lexical" index.

Theories and concepts, e.g., two-level theories

449. Craggs and Kelly (2018) do a meta-ethnographic synthesis of studies of school belonging. They summarize their analysis in the figure below which appears at the end of their article. To put it very briefly, ethnographers, interpretivists, etc., often are hesitant to include causal claims and are more interested in meaning (e.g., concepts) than causal relationships. Using ideas and figures from Goertz (2020) do two different versions of this figure. Do a completely conceptual one, noting that already in the figure there are at least three levels, "Higher order concept," "Main concept" and "Subconcept", from Goertz's perspective a three-level concept. Then do a causal version where the "Higher order concept" becomes the dependent variable, and hence it is a causal model (Hint: do it as a two-level model, chapter 10 of the concept book).

Table 18: Conceptualizing various forms of belonging)

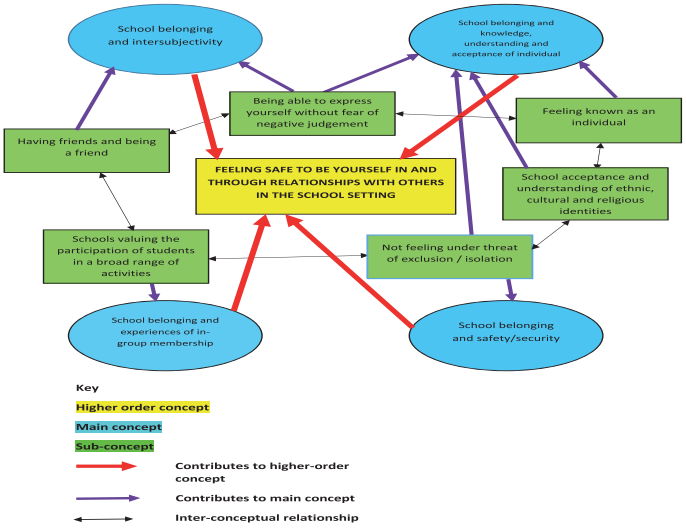


Figure 1. Diagrammatic representation of the higher order concept, main concepts and sub-concepts.

Source: Craggs and Kelly 2018.

450. Two level or SUIN models raise questions about what to include in the statistical analysis. An example on the independent variable side is Uzonyi's concept of a "big political opportunity" which takes three forms that are expressed in his three hypotheses. Should he include each of the three forms separately in a regression analysis or just the higher variable formed by the logical OR?

> There are generally three types of domestic unrest that may lead to the onset of genocide or politicide. The first, which is captured by those scholars focusing on big opportunities for government mass murder, are militarized threats to a country's political leadership. . . . Coups provide a similar logic. . . . The third type of threat a government faces is non-militarized unrest. Non-militarized domestic unrest, such as protests, riots and strikes, can spark political violence if the regime perceives that such unrest will grow into greater instability. (Uzonyi 2016, 318–19)

451. Ohlson (2008) "Understanding causes of war and peace" has implicit two-level theory. One might think that the basic level variables are the same, his "Triple-R" triangle for both causes of war and causes of conflict resolution, but that the secondary-level variables are different for each dependent variable. Discuss.

452. Horowitz (2010) chapter 2 presents a two-level theory. For the two core, basic level, independent variables of organizational capital and financial intensity. Draw a figure with ANDs and ORs (or with *'s and +'s for the linear algebra version) that represent the logical structure of his theory.

453. Discuss the Rodrik (2007) figure below. Note that everything is conceptualized as a constraint (see chapters 2 and 4 of *Multimethod research, causal mechanisms,*

and case studies: an integrated approach). Which case studies would be useful? Which ones does he do? Discuss the nature of the arrows – e.g., are they causal – and the two-level structure of the causal mechanism. Could the arrows be casually going up instead of down?

> Moving down the branches of the decision tree is tantamount to discarding candidates for the most binding constraint on growth. The overarching lesson from our theoretical analysis is that it is this constraint, once identified, that deserves the most attention from policymakers. (Rodrik 2007, 66)

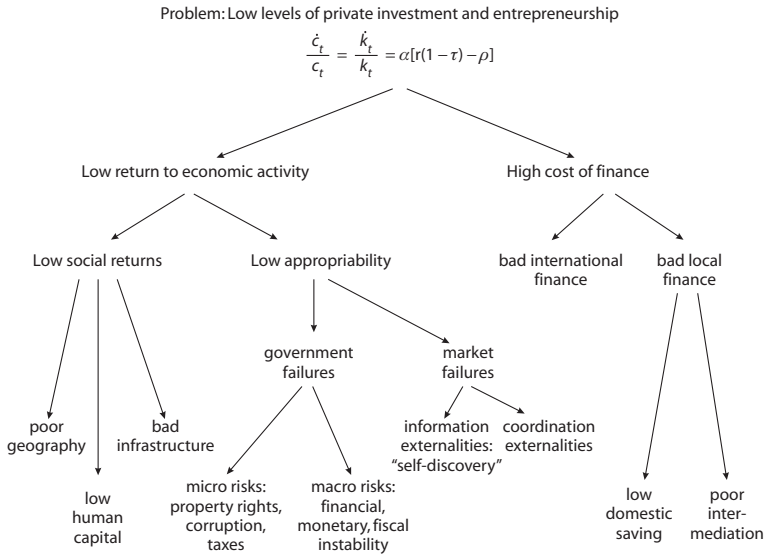


Fig. 2.1. Growth diagnostics

454. Tilly 2004 (particularly chapter 1) provides a rich variety of potential INUS and/or two-level models, for example:

> The book's most general claims follow: ... at least one of the processes under each of the first two headings (categorical inequality and trust networks) and *all* of the processes under the third heading (alterations of public politics) must occur for democratization to ensure. (Tilly 2004, 22).

Formalize his theories in terms of figures and/or INUS equations. This exercise also works well for Tilly's account of other prominent theories (e.g., Rueschemeyer, Stephens, and Stephens 1992).

455. Kaufman (2006) provides an INUS model for explaining extreme ethnic conflict. Write the INUS formula for his model, discussing which variables are necessary condition and how sufficiency is achieved. Use the INUS model to explain why Kaufman argues that "peace," i.e., no extreme ethnic conflict, is overdetermined (e.g., p. 56). See Bennett and George (2005) for a discussion of INUS models; the original source is Mackie (1980); Ragin's methods produce in general INUS models.

Answer:

Here are the basic variables in Kaufman's model:

> Symbolist theory posits the following causal chain, but first, three preconditions for ethnic war are necessary: S1. Widespread group myths exist on both sides that explicitly justify hostility toward, or the need to dominate, the ethnic adversary. S2. Fear of group extinction is strong on both sides at the time violence breaks out. S3. Both sides have a territorial base and the opportunity to mobilize. Next, mobilization for violent conflict is driven by three processes: S4. Extreme mass hostility is expressed in the media and in popular support for the goal of political domination over ethnic rivals (on at least one side) or resistance to such domination. S5. Chauvinist elites use symbolic appeals to group myths, tapping into and promoting fear and mass hostility, to mobilize their groups for conflict. S6. A predation-driven security dilemma arises, in which the growing extremism of the leadership on at least one side results in radicalization of the leadership on the other. These propositions together identify a causal mechanism. For the model to be correct, there must be evidence that in each case of ethnic war, the necessary preconditions not only were present but were causally important in enabling chauvinist mobilization and in causing a predation-driven security dilemma. Additionally, the logic requires that an explicitly genocidal policy must be based on explicitly eliminationist myths. (Kaufman 2006, 58)

One INUS model would be: S1*S2*S3*(S4+S5+S6).

"Peace" is overdetermined in the sense that if any of the necessary conditions, i.e., S1–S3 is absent then there is no conflict. Conflict is also absent if (S4+S5+S6) is too small.

456. Compare Geddes's (1993) figure of Skocpol's theory to that in chapter 10 of *Social science concepts*. Give two major differences between the two.

Answer:

(1) military threat, and substitutability of other causes of state breakdown and (2) problems with interaction term and her interest in "the" effect of military threat.

457. One important question is the extent to which two-level or fuzzy set models/methods differ from statistical ones. Examine Geddes's (1993) figures 3.5 and 3.6. She finds no linear relationship between labor repression and economic growth. Draw a new horizontal and vertical through these figures which gives a 2×2 table that nicely supports the hypothesis that labor repression is a necessary condition for high economic growth.

Answer:

One can easily draw horizontal and vertical lines through her scatter plot to show that her data support the necessary condition hypothesis. Just draw a horizontal line at around 4% GDP growth and the vertical line at 1.5 labor repression.

458. Reformulate Doyle and Sambanis's analysis (2000) of the effectiveness of UN peacekeeping in two-level terms.

Answer:

Basic level concepts are (1) international capacities, (2) hostility, and (3) local capacity. For secondary-level variables see page 783, in the section "The Explanatory Variables."

"Thus, we theorize that the PB [peacebuilding] process is captured by: PB = IC [International Capabilities]*NLC [Net Local Capacity]." (Doyle and Sambanis 2000, 782)

"All strategies should address the local sources of hostility, the local capacities for change, and the (net) specific degree of international commitment available to assist change. One can conceive of these as the three dimensions of a triangle whose area is the 'political space,' or effective capacity, for building peace. This suggests that the dimensions substitute for one another, that is, more of one substitutes for less of another; less extreme hostilities substitute for weak local capacity or minor international commitment." (Doyle and Sambanis 2000, 781; see also the statistical model of table 4, p. 791)

459. Elman (2005) in his article on typological methods discusses the method of "indexing." Using his discussion of Walt's theory of alliance, explain how indexing can be the creation of a higher level variable/concept. How would the indexing he does fit into the varieties of two-level theories?

Answer:

"Threat" is composed of several secondary-level variables such as capability and proximity. By adding up the scores he is clearly using a family resemblance strategy.

460. Combine Walter (1997) and Hartzell and Hoddie (2003) to form a two-level theory of civil war conflict resolution. As basic-level variables use third-party guarantees and power-sharing. See Walter for an argument about the structural relationship between the two. (Note: you might assume the scope conditions that the countries have already started negotiations.)

Answer:

Walter (p. 351) clearly argues that there is substitutability between the two basic-level variables. Walter argues a necessary and sufficient condition view of credible commitments (though she is not completely explicit on this point). Hartzell and Hoddie clearly argue for a family resemblance perspective on power-sharing (p. 321).

461. What are the three "basic level" hypotheses in Skocpol's *States and social revolutions* and which one of these is Geddes (2003) basically (though obviously not exactly) testing in her discussion of Skocpol?

Answer:

The three basic-level hypotheses are: (1) peasant revolt is necessary for social revolution; (2) state breakdown is necessary for social revolution; and (3) the combination of peasant revolt and state breakdown is sufficient for social revolution.

Geddes explores whether "external threat" is correlated with the occurrence of social revolution. She operationalizes external threat in terms of invasion and/or loss of territory. She also briefly explores the hypothesis that external threat is necessary for social revolution. Hence, Geddes examines whether one of Skocpol's secondary-level variables for "state breakdown" is causally related to social revolution. Note that Skocpol herself hypothesizes that external threat affects social revolution by working through the basic-level variable of state breakdown.

462. Sekhon (2004) analyzed Skocpol's theory of social revolution using statistical assumptions. Compare his analysis of Skocpol with the fuzzy-set analysis of Skocpol in chapter 9.

Answer:

Sekhon reassesses Geddes's data from Latin America concerning the relationship between defeat in war and revolution. Contrary to Geddes, he concludes that these data do not rule out the possibility that the two factors are (probabilistically) related to one another. Hence, he concludes that Skocpol's argument may be correct, if stated probabilistically.

By contrast, chapter 10 evaluates Skocpol's theory under the assumption that defeat in war is a secondary-level variable that affects social revolution only by working through the basic-level variable of state breakdown. Hence, from this perspective, Sekhon's reanalysis is not directly testing the theory developed by Skocpol.

463. Burawoy's (1989) critique of Skocpol shows how important, and how difficult, it is to see the complete causal structure of Skocpol's argument. His table 1 (p. 768) clearly divides variables into two levels and he correctly identifies peasant revolt and state crisis as the basic-level variables. However, he does not interpret the relationship between the secondary level and the basic level correctly. What is the nature of his error?

> The task now is to show that both international pressure and an 'organized and independent dominant class with leverage in the state' were necessary ingredients for political crisis . . . So far so good, but note immediately that the contrasting cases [Germany and Japan] do not demonstrate 'international pressure' as necessary for the development of a revolutionary political crisis. In the next chapter Skocpol examines the necessary conditions for the second component of revolution: peasant revolt . . . She now has to demonstrate that both political crisis and peasant autonomy were necessary for peasant revolt. (Burawoy 1989, 766)

Answer:

Burawoy assumes that the secondary-level variables are necessary for the basic-level variables. However, the indicator/data-level variables are better characterized as individually sufficient for the basic-level variable.

464. Most and Starr introduced the influential notion of foreign policy substitutability (Most and Starr 1984; see also Most and Starr 1989). They are also well-known for the idea that opportunity and willingness are individually

necessary and jointly sufficient for foreign policy action (Starr 1996). Put opportunity and willingness together with foreign policy substitutability to get the structure that Ostrom uses in her model of common pool resource institutions.

Answer:

If one puts opportunity and willingness at the basic level and foreign policy substitutability at the secondary level, one arrives at a two-level model. Cioffi-Revilla and Starr (2003) formally present this kind of model in a completely probabilistic fashion.

465. Interpret the following theories in two-level terms. Some may not be complete two-level models, in that the secondary-level may not always be clearly present for all basic-level factors.

(1) Marks, G. 1986. Neocorporatism, and incomes policy in western Europe and North America. *Comparative Politics* 18:253–77.

(2) Diehl, P., C. Ku, and D. Zamora. 2003. The dynamics of international law: the interaction of the normative and operating systems. *International Organization* 57:43–57.

(3) Jacoby, Wade. 2001. *Imitation and politics: redesigning modern Germany.* Ithaca: Cornell University Press.

(4) Ertman, Thomas. 1997. *Birth of the leviathan: building states and regimes in medieval and early modern Europe*. Cambridge: Cambridge University Press.

(5) Kingdon, John. 1984. *Agendas, alternatives, and public policies*. Boston: Little, Brown.

(6) Linz, Juan J., and Alfred Stepan. 1996. *Problems of democratic transition and consolidation: Southern Europe, South America, and Post-Communist Europe*. Baltimore: The Johns Hopkins University Press.

(7) Wickham-Crowley, Timothy. 1996. *Guerillas and revolution in Latin America: a comparative study of insurgents and regimes since 1956*. Princeton: Princeton University Press.

(8) Burgerman, S. 2001. *Moral victories: how activists provoke multilateral action.* Ithaca: Cornell University Press.

(9) Kugler, J., and A. Organski. 1989. The power transition: a retrospective and prospective evaluation. In M. Midlarsky (ed.) *Handbook of war studies.* Ann Arbor: University of Michigan Press. See also Lemke, D., and S. Werner. 1996. Power parity, commitment to change, and war. *International Studies Quarterly* 40:235–60.

(10) Doner, R., et al. 2005. Systematic vulnerability and the origins of developmental states: Northeast and Southeast Asia in comparative perspective. *International Organization* 59:327–61.

(11) Newman, K. 2005 *Rampage: the social roots of school shootings*, chapter 10. New York: Free Press. See also the special issue of *Sociological Methods & Research* in particular Harding, D., C. Fox, and J. Mehta (2002).

Answer:

The answers below are suggestive hints and avenues to pursue and do not constitute a complete answer to the question.

(1) "The requisites of consensual incomes policy that have been discussed here, socialist participation in govern and centralized trade union structure, coincide only in the Nordic countries, the Low Countries, and Austria. . . . Without stable social democratic governance, consensual incomes policy is unlikely to be attempted, and without centralized trade unions, it is unlikely to be a viable policy in the medium term." (Marks 1986, 269)

(2) "A general theoretical argument focuses on four conditions. We argue that the operating system only responds to normative changes when response is "necessary" (stemming from incompatibility, ineffectiveness, or insufficiency to give the norm effect, and when the change is roughly coterminous with a dramatic change in the political environment (that is, "political shock"). We also argue, however, that opposition from leading states and domestic political factors might serve to block or limit such operating system change." (Diehl et al. 2003, 71–72)

(3) "Institutional transfer involves three necessary conditions: first, state elites refer explicitly to a model prominent in another place; second, they try to identify the foreign model's legal framework and the actors that help it function; finally, these elites build a replica of all or part of the model." (Jacoby 2001, 2)

"For institutional transfer to be effective, Jacoby says, two conditions have to be met: First, institutional transfer has to be supported by active segments of civil society, and second, flexibility in adapting institutions to foreign settings is essential . . . . Similarly, the politics of imitation faces higher obstacles when exact transfer is intended. A functional-equivalency approach, with its built-in elasticity, is preferable." (Welsh 2002, 211; review of Jacoby 2001)

(4) "As the patrimonalist fate of Hungary and Poland illustrate, however, the mere presence of participatory local government is in itself not enough to ensure the triumph of bureaucratic constitutionalism. It is only the combination of participatory local government with a strong center equipped with independent capacities of rule that, the British case implies, can assure such an outcome." (Ertman 1997, 324 (last page); note that the other key variable is timing of geopolitical competition, either preor post-1450).

(5) Kingdon's model has three "streams" each of which is necessary and which together are jointly sufficient. His discussion is very clear about the different substitutable ways these streams can be filled.

(6) "Democracy is a form of government of a state. Thus, no modern polity can become democratically consolidated unless it is first a state. . . . If a functioning state exists, five other interconnected and mutually reinforcing conditions must also exist or be crafted for a democracy to be consolidated. First, the conditions must exist for the development of a free and lively civil society. Second, there must be a relatively autonomous and valued political society. Third, there must be a rule of law to ensure legal guarantees for citizens' freedoms and independent associational life. Fourth, there must be a state bureaucracy that is usable by the new democratic government. Fifth, there must be an institutionalized economic society" (Linz and Stepan 1996, 7).

"Having analyzed the necessary conditions for a consolidated democracy and then spelled out the key differences among the four ideal-typical nondemocratic regimes, it should be clear that the characteristics of the previous nondemocratic regime have profound implications for the transition *paths* available . . . Within the logic of our ideal types, it is conceivable that a particular authoritarian regime in its late stages might have a robust civil society, a legal culture supportive of constitutionalism and rule of law, a useable state bureaucracy that operates within professional norms, and a reasonably well-institutionalized economic society. For such a polity, the first and only necessary item on the initial democratization agenda would relate to political society-that is, the creation of the autonomy, authority, power, and legitimacy of democratic institutions. We argue in chapter 6 that Spain, in the early 1970s, approximated this position" (Linz and Stepan 1996, 55).

(7) Wickham-Crowley (1992) has a two-level model with the following structure:

I. Peasant support
A. Agrarian structure
B. Agrarian disruption
C. Rebellious cultures
D. Peasant linkage (guerilla leaders have peasant ties, background)
II. Guerilla/Government military strength
A. Financing
B. Internal solidarity
C. Support from outside (e.g., Cuba, US)
III. Regime weakness (see figure 8.1)
A. Weak cohesion of upper classes
B. Weak cohesion of middle classes
C. US occupation.
IV. Existence of guerilla groups (this is added in Boolean analysis in chapter 12)

(8) Burgerman proposes "an interacting set of necessary conditions" for the success of human rights groups in forging peace agreements: the existence of relevant international norms and transnational activism; the existence of elites in the target state that have control over the armed forces and that have concern about their (country's) international reputation; and the existence of organized domestic groups linking up with transnational activists. Added to these propositions is the sensible if unremarkable condition that "if a major power maintains overriding security or economic interests in the target state, it can inhibit the enforcement of human rights principles and agreements."

(9) "Clearly, the necessary but not sufficient conditions for major war emerge only in the rare instances when power parity is accompanied by a challenger overtaking a dominant nation. The odds of a war in this very reduced subset are 50 percent. No other theoretical statement has, to our knowledge, reduced the number of cases to such a small set, and no other is so parsimonious in its explanatory requirements" (Organski and Kugler 1989, 179, see their table same page).

"More importantly, the power transition suggests that, during are rare periods when a challenger overtakes the dominant nation, war will be waged only if the potential challenger is dissatisfied" (Organski and Kugler 1989, 51).

In particular, it is clear that there are various ways a country can be dissatisfied.

(10) "We contend that developmental states will only emerge when political leaders confront extraordinarily constrained political environments. Specifically, we argue that political elites will only build such institutional arrangements when simultaneously staring down the barrels of three different guns: (1) the credible threat that any deterioration in the living standards of popular sectors could trigger unmanageable mass unrest; (2) the heightened need for foreign exchange and war materiel induced by national insecurity; and (3) the hard budget constraints imposed by a scarcity of easy revenue sources. We call this interactive condition 'systemic vulnerability' " (Doner et al 2005).

(11) Necessary conditions are very explicit. Two-level nature of theories comes out clearly in discussion and data in Appendix A. Substitutability is clear, particularly in footnote 2, page 372.

466. Reinterpret the following QCA analyses in terms of two-level theories:

(1) Stokke, O. 2004. Boolean analysis, mechanisms, and the effectiveness of international regimes. In A. Underdal and O. Young (eds.) *Regime consequences: methodological challenges and research strategies.* London: Kluwer.

(2) Blake, Charles H. and Jessica Adolino. 2001. The enactment of national health insurance: a Boolean analysis of twenty advanced industrial countries. *Journal of Health Politics, Policy and Law* 26:679–708.

(3) Snow, David and Daniel Cress. 2000. The outcomes of homeless mobilization: the influence of organization, disruption, political mediation, and framing. *American Journal of Sociology* 105:1063–1104.

Answers:

(1) In the Boolean equation SUCCESS = A*i + A*S*R one can see "i" and "S*R" being easily interpretable in terms of costs of compliance. Hence there is substitute between them. A is a necessary condition already.

(2) For example, "National health insurance emerged and/or endured in three basic scenarios. First, given a receptive societal culture, a relevant leftist party, and a corporatist pattern of interest group activity, NHI existed regardless of the nature of the major political institutions. Second, unitary states with receptive cultures were capable of producing NHI in countries with weaker executives and in countries without a relevant labor party (irrespective of the nature of the other two independent variables). Third, Anglo-American societies with pluralist interest group environments adopted NHI with a dominant executive (regardless of the federal-unitary distinction or the relevance of labor parties)." (Blake and Adolino 2001, 696)

(3) Snow and Cress (2000) in their analysis of the success of homeless social movements find the same sort of pattern:

> Six SMOs [social movement organizations] obtained positions on boards and task forces that addressed the homeless issue [dependent variable].

> Two pathways led to this outcome. Organizational viability, diagnostic frames, and prognostic frames were necessary conditions for obtaining representation. These conditions were sufficient in combination with either disruptive tactics, where allies were present, or nondisruptive tactics, in the context of responsive city bureaucracies. (p. 1082)

Here too one sees the substitutability of power in the analysis. One needs either "allies" who are influential in the community *or* a friendly city government to begin with. You do not use disruptive tactics with friends, where you do if the city government is unfriendly.

467. Braumoeller (2003) has a large table giving many theories that use ANDs and ORs. Examine this list for other potential two-level theory candidates.

468. The Nussbaum example illustrates that one can think of many theories of (social) welfare and justice in the same structural terms as this book applies to concepts:

> Most theories of justice can also be usefully analyzed in terms of the information used in two different-though interrelated-parts of the exercise, viz. (1) the selection of *relevant personal features*, and (2) the choice of *combining characteristics*. To illustrate, for the standard utilitarian theory, the only intrinsically important "relevant personal features" are *individual utilities*, and the only usable "combining characteristic" is *summation*, yielding the *total* of those utilities. (Sen 1992, 73)

Choose some well-known theories of justice or social welfare and describe the dimensions and the structuring principle.

Answer:

Sen gives the following examples:

> Examples of selection of "relevant personal features" other than utilities include *liberties* and *primary goods* (Rawls 1971), *rights* (Nozick 1974), *resources* (Dworkin 1981), *commodity bundles* (Foley 1967; Pazner and Schmeidler 1974; Varian 1974, 1975; Baumol 1986), and various *mixed* spaces (Suzumura 1983; Wriglesworth 1985; Riley 1987). . . .
>
> The selection of personal features has to be supplemented by the choice of a combining formula, e.g., *sum maximization* (Harsanyi 1955), *lexicographic priorities and maximin* (Rawls 1971; Sen 1977 "On weights and measures"), *equality* (Foley 1967; Nozick 1974; Dworkin 1981), or one of various *other* combining rules (Varian 1975; Suzumura 1983; Wriglesworth 1985; Baumol 1986; Riley 1987). (Sen 1992, 74)

469. Describe the three-level structure of Dasgupta's concept of human well-being or destitution (Dasgupta 1990; Dasgupta and Weale 1992). What is its structure? What are the relative weights attached at the indicator and secondary levels?

Answer:

The structure is very similar to the one used by Schmitter for corporatism, since it uses the sum of the ranks. The secondary-level level variables such as "health," "education," "wealth" will not be equally weighted in general unless

they have the same number of indicators. More generally, is it reasonable to rank literacy equally with per capita income?

470. Chapter 2 stresses that an important part of concept-building is considering the negative pole of the concept. Nussbaum and Sen focus on that concept of human *needs*. What is the opposite pole of a human need?

Answer:

Classically one contrasted the concept of needs with that of *luxury*. For example, "[W]hy is it that all societies make some category distinction between need/necessity and luxury? ... It is not paradoxical to assert that the distinction constitutes a unity, just as it is not mysterious to talk of two gloves, a right-hand and a left-hand, constituting *a* pair.... The relation between necessity and luxury is negative or oppositional so that ... luxury specifies the necessity by indicating what it is not" (Berry 1994, 231–32).

471. When deciding who gets a particular good the weakest link principle, chapter 5, can be used with the need-luxury scale to make a decision. Braybrook (1987) calls this the Principle of Precedence. Explain the moral philosophy involved.

Answer:

The Principle of Precedence means that when Alan needs something that Brenda wants but does not need, then meeting Alan's need is prima facie morally preferable to satisfying Brenda's desire. This suggests that needs have precedence over luxuries. The weakest link principle then suggests that for any good you look at each person's need-luxury score on the good and give it to the person with the minimum (i.e., most needy).

Statistical issues, latent variables, etc.

472. Often by default scholars linearly aggregate dimensions of a concept. Discuss how scatterplots might inform the aggregation decision and lead to a necessary condition structure. Luna (2014) provides a very nice example with two dimensions of party system institutionalization, electoral volatility and programmatic structure.

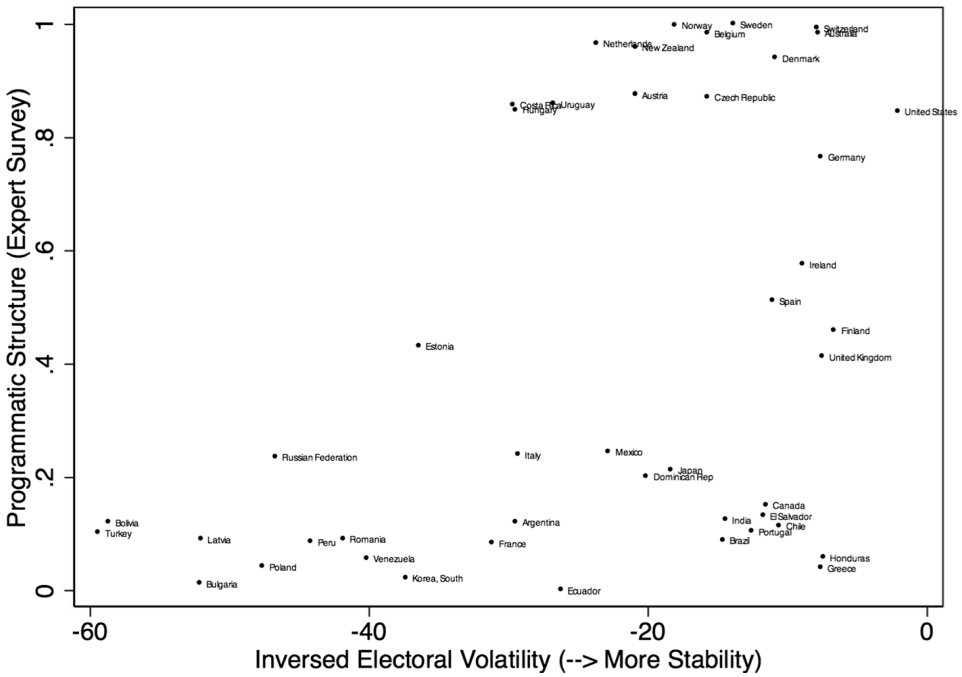


Fig. 2 Electoral volatility and programmatic structure (2000s). Source: Own construction on the basis of Scott Mainwaring, Annabela España, Edurne Zoco, and Carlos Gervasoni's electoral volatility database and data from the Democratic Accountability and Linkages Project (2009)

Answer: As Luna notes "As a result, collapsing the two phenomena in a linear (additive) latent variable measurement is misleading. Although future research is needed to explore the relation between the two dimensions unpacked here and the remaining theoretical dimensions and indicators included in Mainwaring and Scullys conceptualization of PSI, the analysis of the relationship between these first two dimensions presented in this paper already constitutes sufficient ground for claiming that the concepts structure (family resemblance) and aggregation rule (linear addition of dimensions and their indicators into a single index) need to be revised" (Luna 2014, 422; see also Coppedge 2012 chapter for necessary condition scatter plot with a necessary condition relationship between two latent variables)

473. How would could one interpret the clearly triangular relationship, see her figure, between these two conflict variables? Could this be interpreted in terms of constraints?

> Figure 0.1 gives a broader view. Here, I plot the relationship between scores on the Political Terror Scale (PTS, x axis) (Wood and Gibney 2010) and scores on the Sexual Violence in Armed Conflict dataset (SVAC, y axis) (Cohen and Nords 2014). Both the PTS and the SVAC are yearly indices assigned by coding qualitative reports. The PTS measures the overall level of government repression on a scale of 1 (very low violence) to 5 (widespread terror), focusing on lethal violence, torture, and detention. SVAC scores focus exclusively on sexual violence and are measured on a scale from 0 (very few reported cases of sexual violence) to 3 (widespread and/or systematic sexual violence). .... Second, as with the missing northwest quadrant of table 0.1 , there are zero governments whose highest level

of sexual violence exceeded their highest overall level of "terror" during this period. Governments that perpetrate relatively high levels of sexual violence in the context of low overall levels of violence are extraordinarily rare. The reverse situation – high overall levels of violence with low levels of sexual violence – is relatively common. (Hoover Green 2018, 11)

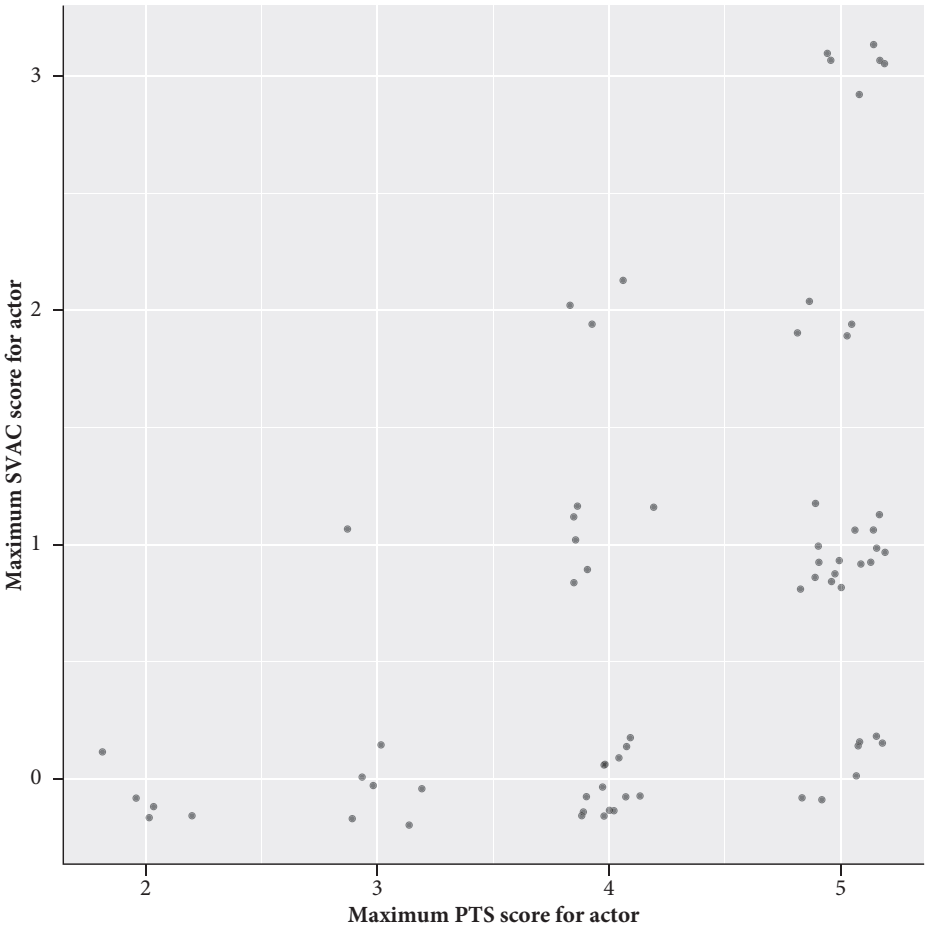


FIGURE 0.1 Comparing PTS and SVAC scores for seventy-five civil wars

Answer:

This illustrates a virtually perfect necessary condition scatterplot. So PTS is a necessary condition for SVAC. So PTS constrains the possible values SVAC can take.

474. A very important methodological issue is set relationships between independent variables. In an original outline of the book a major part of a chapter was going to be devoted to the methodological issues that this raises. The main methodological concern arises when X_1 is a subset of X_2. For example, in the debate about the territorial peace versus democratic peace, it turns out that the the set of democratic dyads is a subset of the dyads at territorial peace:

"This study compares the conflicting answers of the democratic peace and the territorial peace and examines the empirical record to see which is more accurate. It finds that almost all contiguous dyads settle their borders before they become joint democracies" (Owsiak and Vasquez 2016, 339).

What is the causal interpretation of the subset variable when all of the variation in X_1 is taken up in X_2? Discuss how this is different from multicollinearity, for example, the two might not be very correlated. See Owsiak 2020 for a nice explicit discussion of the subsetting issue and other examples.

What if X_1 is a perfect subset of X_2 and one introduces an interaction term, $X_1 * X_2$?

This is just with dichotomous variables, but one can find set theoretic relationships among continuous variables (defined as X_1 less than or equal to X_2 for all observation or vice versa).

475. Discuss the causal figures in *Multimethod research, causal mechanisms, and case studies: an integrated approach*, Weller and Barnes (2014) or Waldner (2014) from a DAG cross-case, large-N, probabilistic perspective. Are these figures identified from a DAG perspective (e.g., Morgan and Winship 2015). Does it matter from a within-case causal inference perspective if they are not?

476. Discuss Pahre's claim that necessary conditions violate common statistical assumptions: "Third, necessary conditions violate the unit homogeneity assumption common in statistics, which states that 'if two units have the same value of the key explanatory variable, the expected value of the dependent variable will be the same' (King et al., 1994: 116). Necessary conditions violate this condition because observations x, y and x, not-y are both consistent with the necessary condition x ← y; in other words, x may occur with or without y" (Pahre 2005, 131).

477. Draw a figure describing the three-level character of the Hicks-Kenworthy (2003) concept of the welfare state. Compare the theoretical discussion of the dimensions of the welfare state with the results of the principal components statistical model. What are the possible options for the secondary-level dimensions?

478. For German speakers. Opp (2005) is a very rare methods textbook that has devoted significant attention to concepts – most textbook only look at measurement. It is also a rare textbook that devotes a whole chapter to teaching logic – most research design and methodology textbooks do not discuss this at all. In Opp's discussion of concepts (chapter 4) he distinguishes between (1) "Analytische operationalisierung" and (2) "empirische operationalisierung." Discuss the extent to which (2) corresponds the the standard latent variable model of concepts and (1) corresponds to what the concept book calls ontological concepts.

479. Use the following quote to discuss the importance of predictive power in terms of evaluating a measurement model or concept. Contrast this with an approach which stresses the semantics and content of the test items.

> The g factor is an artifact of linear correlation analysis. A theorem of Suppes and Zanotti (1981) informs us that for any vector of test scores from an achievement test, it is possible to construct a scalar latent factor such that, conditional on the factor, test scores are independent. The g factor exists for any vector of binary, finite-valued, or countably valued random variables. The g of conventional psychometrics is a product of mathematical conventions in factor analysis. A g also exists to account for correlations among test scores. That is a mathematical theorem of no behavioral consequence for psychometrics or for finance, another field addicted to factor models. The value of g in predicting behavior is the real test of its importance. There is much evidence that it has predictive power. (Heckman 1995, 1105)

480. Recently statisticians have become much more concerned with problems of "unit homogeneity." Here is Henry Brady defining the idea:

> We shall make the transformation of $Y_B(1,0)$ into $Y_A(0,0)$ in two steps which are depicted on Table 10.9. If A and B are identical and Z_A and Z_B [Z is the treatment] are identical as well(footnote: By saying that Z_A and Z_B have to be comparable, we mean that $Z_A = 0$ and $Z_B = 0$ are the same thing and $Z_A = 1$ and $Z_B = 1$ are the same thing.) (although we haven t indicated how this might be brought about yet) it might be reasonable to suppose that: $Y_B(1,0) = Y_A(0,1)$, [*Identicality of units and treatment or Unit Homogeneity*]. That is, A and B are mirror images of one another so that the impact of $Z_A = 1$ and $Z = 0$ on B is the same as the impact of $Z = 0$ and $Z = 1$ on A. (Brady 2008, 258)

Analyze how valid concepts are critical to the existence of unit homogeneity.

Answer:

Holland (1986) says "unit homogeneity" means that units are prepared carefully "so that they 'look' identical in all relevant aspects" (Holland 1986, 948). To be "identical" means that the concept must produce identical units, A and B in the Brady analysis. If a study looks at all "states" then the conceptualization of state means that there are no casually important differences in "stateness." Treatments, Z in the Brady analysis, also are assumed to be homogeneous.

481. A key issue in the three-level framework is whether the same data-indicators can be used for different secondary-level dimensions. See how this occurs in "Measuring the rule of law": "A careful examination of the nine factors reveals two facts. The first is that there is a partial overlap among sub-factors; that is, a sub-factor can simultaneously belong to different factors at once. This is simply to reflect the fact that various rule-of-law dimensions partially overlap in practice" (Botero and Ponce 2011, 16).

482. Concept generalization across groups is a core methodological issue. Discuss the sense of "meaning" in the abstract and how it relates to generalization of a concept across groups.

> Measurement invariance assesses the (psychometric) equivalence of a construct across groups or measurement occasions and demonstrates that a construct has the same meaning to those groups or across repeated measurements. . . . Appropriate and proper comparison of a construct between

groups or across times, therefore, depends first on ensuring equivalence of meaning of the construct. The untoward consequences of measurement noninvariance can be readily illustrated in the study of depression in men and women. Suppose frequency of crying, weight gain, and feelings of hopelessness are indicative of the severity of depression in women, but only feelings of hopelessness are indicative of the severity of depression in men. If the three indicators are combined into a scale to compare depression in women and men, mean differences on the scale may mislead because crying and weight gain have little relation to depression in men. In this example, men may score lower than women on the depression scale because they cry less and gain less weight. (Putnick and Bornstein 2016, 71)

483. Discuss whether a causal interpretation should be given to the latent variable "coup risk."

> My key independent variable is coup risk. . . . But why should we believe that all three items – civil-society strength, political legitimacy and recent coups – tap the coup risk equally well? . . . I then modeled the logit of a latent coup risk, logit(q_{it}), as a linear function of several willingness and capability indicators (z_{it}). Specifically, z_{it} is a vector of country-year characteristics that impact the military's willingness and capability to organize a coup and thus are plausible sources of variation in the probability of coups. (Sudduth 2017, 9)

Normative issues

484. Explain why this is a normatively problematic conceptualization of "disability." Explain why being black might be a disability under this conceptualization.

> In the ICF, problems with human functioning are categorized in three interconnected areas:
>
> impairments are problems in body function or alterations in body structure – for example, paralysis or blindness;
> - activity limitations are difficulties in executing activities – for example, walking or eating;
> - participation restrictions are problems with involvement in any area of life – for example, facing discrimination in employment or transportation.
>
> Disability refers to difficulties encountered in any or all three areas of functioning. . . . Disability arises from the interaction of health conditions with contextual factors – environmental and personal factors as shown in the figure below . . . The ICF is universal because it covers all human functioning and treats disability as a continuum rather than categorizing people with disabilities as a separate group: disability is a matter of more or less, not yes or no. However, policy-making and service delivery might require thresholds to be set for impairment severity, activity limitations, or participation restriction. (WHO-World report on disability 2011, 5)

485. Race is notoriously a problematic concept. Discuss the history and methodology of determining race in the US Census. What are the normative implications of this? What is the aggregation principle behind the "one-drop rule"?

> The constructed, contradictory, and fickle nature of US racial classification is exemplified in the revisions of census categories. Since its inception in 1790, the decennial US census has counted the population by race/color. For much of the censuss history, race was assigned to individuals by an enumerator, characterized as unchangeable, and race mixture was ignored entirely. Yet, between 1850 and 1920, racial fluidity was acknowledged, and periodically enumerated with specificity, via fractional mixed-race categories: mulatto (half-black), quadroon (one-quarter black), and octoroon (one-eighth black). By 1930, these categories were removed, and mixed-race blacks, as well as individuals of mixed white-Asian parentage, were subjected to the one-drop rule and categorized with their minority race. (Davenport 2020, 223)

486. The concept of " transitional justice" must be normative. How are Olsen et al. dodging this issue by including "amnesty" as justice? "Still others value amnesty as a mechanism that acknowledges wrongdoing, but releases societies from the trap of the past. Rather than making normative assumptions about the appropriate form of justice, our dataset recognizes that societies adopt a variety of mechanisms to engage their violent pasts and allows scholars to include or exclude specific mechanisms based upon their own operationalization of transitional justice." (Olsen et al. 2010, 805)

    Answer:

    It might be better then to call the dataset the "Transitional practices" dataset. Also, the justification for amnesty is almost never normative but pragmatic.

487. Braybrook (1987) discusses what he calls the Principle of Precedence. The Principle of Precedence means that when Alan needs something that Brenda wants but does not need, then meeting Alan's need is prima facie morally preferable to satisfying Brenda's desire. This suggests that needs have precedence over luxuries.Explain the moral philosophy involved. What does this suggest in terms of an aggregation principle? Discuss the moral philosophy behind this principle.

    Answer:

    The weakest link principle then suggests that for any good you look at each person's need-luxury score on the good and give it to the person with the minimum (i.e., most needy).

Global indices, barometers, etc.

488. The Possibility Principle is closely related to scope conditions. The welfare state literature only looks at wealthy countries. Much of the literature on well-being focuses on poor or destitute countries (Dasgupta 1990). When

Dasgupta examines the relative destitution of countries he excludes the well off ones. He uses the Borda Rule (Dasgupta and Weale 1992) to aggregate a variety of indicators of well-being, such as infant mortality, literacy, per capita income civil and political rights. The Borda Rule ranks countries on each dimension then sums the ranks (see Fine and Fine 1974 for an extensive and technical discussion of its formal properties). Explain why this concept of human destitution or well-being is sensitive to the exclusion of wealthy countries. For example, the comparison of India and China will significantly be affected by the inclusion or exclusion of nonpoor countries.

Answer:

As Dasgupta and Weale note, the Borda Rule is sensitive to the population chosen because the distance between a pair of countries depends on the population. In particular, if an alternative population puts countries in between a given pair their overall rankings (and hence the comparison between the two) can change. Particularly with regarding to civil and political rights the inclusion of wealthy countries can change the comparison because these are scattered throughout the ranking.

489. A very important theoretical and policy issue is the evaluation and counting of the number of "poor" people in the world. The World Bank is very involved such efforts (e.g., Chen and Ravallion 2001) and concludes, controversially, that amount of poverty has decreased over the last 20–30 years. Typically the World Bank model uses an "international poverty line," typically an income of $1 US per day. Incomes in local currencies must then be converted into US dollars to evaluate the extent of poverty. Purchasing power parities (PPP) are used to do this. These are general figures based on a "basket" of goods and the cost of that basket in different countries, and are often used to make GNP/capita comparisons. These PPP values then provide the conversion rates of local currency into US dollars which make world-wide evaluation of poverty possible. Discuss how when the basket of goods consists of a wide range of goods that it might be problematic when the goal is to evaluate poverty.

Answer:

See Reddy and Pogge (2005) for an extensive discussion. Basically, PPP data use a wide variety of goods that are irrelevant for assessing poverty. PPP depends on the relative price of things like, say, SUVs, which are not relevant for assessing poverty. The relative price of SUVs influences PPP data which may distort the relative costs of things that are really important to poor people, like food. For example, it may be that bread is relatively more expensive in poor countries than in rich (this is often true in poor neighborhoods in the United States compared to middle-class ones). If this is generally true between countries then PPP underestimates the extent of poverty in the world. Hence one would ideally like a PPP that only uses goods that are consumed by poor people, such as food, housing, and medical care.

For the mathematically inclined, one can easily use Reddy and Pogge (2005) to devise an exercise that uses basic calculus (i.e., derivatives) to make the point of the exercise.

Gender

490. As a general rule quantitative scholars are not very interested in the relationships between their various independent variables. However, from a conceptual and causal inference point of view there are important issues which often arise. A classic one is when two independent variables are somehow part of the same concept. Barnhart et al. (2020) illustrate this where they have two core variables one which is democracy, using polity data, and women's suffrage. Arguably, see Paxton's work, women's suffrage is part of the concept of democracy. However, many traditional datasets do not include that factor, including polity and others. They make strong claims about the role of women's suffrage in connection with international conflict and more specifically the democratic peace literature:

    > Finally, we examine cross-national conflict data to assess the impact of the enfranchisement of women. We find that female suffrage, when coupled with democratic institutions, is not only an important and powerful cause of peace, but that it may even be necessary for the democratic peace. Without female suffrage, democracies (whether examined dyadically or monadically) do not show clear evidence of being more pacific than nondemocracies. These findings hold across time periods and do not appear to result from other confounding factors, or from the expansion of the political franchise more generally. By bridging the individual and international levels to show how the gender gap on the use of force matters, we contribute to the growing positivist study of gender and international relations. (Barnhardt et al. 2020, 634)

    Someone might argue that it is not women's suffrage that is really the core causal factor but rather *quality of democracy*. Countries with women's suffrage, or more generally, fewer restrictions to suffrage based on race or whatever are higher-quality democracies, *by definition*. So maybe the article should be called the higher-quality democracy peace? (See Goertz and Dul 2023 for an argument and data analysis to this effect).

491. In the intersectionality literature on often hears that say, black-women, are "qualitatively distinct." Discuss what this might mean. Might this be more a causal claim than a conceptual claim? To what extent is intersectionality about hybrid concepts?

492. Concepts with adjectives often came become typologies. For example, gender quotas come various forms, "party," "legislative," "constitutional," "reserved seat" etc. which could generate "types of gender quotas" (Krook 2014). The goal of gender quotas is to raise the number of women in parliament. As such the typology is really a series of causal mechanisms designed to raised the number of women in parliament. Discuss why the rule of mutually exclusive categories for typologies is particularly a bad idea in this concept.

    Answer:

    Clearly a country could, and maybe should, adopt multiple mechanisms for raising the number of women in parliament. In addition there might be synergies between these mechanisms when used together.

493. Examine a concept–measure for democracy. Which secondary-level dimension is most likely to be affected by gender considerations?

Answer:

Almost all concepts of the democracy have "participation/voting" as a secondary-level dimension. Almost universally, the concept involves principles like participation in general, but at the indicator level it is almost always *white male* participation. Paxton (2000) discusses extensively the mismatch between concept and measure:

> Although there remains some debate about the definition of democracy, most scholars would agree that a democracy has at least three components: competition, participation, and civil liberties. I focus on the second component of democracy – participation – which is the source of the operational mismatch in current research. Participation is central to most definitions of democracy. For example, Diamond, Linz, and Lipset (1990:6–7) explain: "democracy ... denotes ... a highly inclusive' level of political participation in the selection of leaders and policies, at least through regular and fair elections, such that no major (adult) social group is excluded." Modern theoretical definitions of democracy stress numbers of people rather than "types" of people. Most authors today consider the criterion for democratic inclusion to be "all native-born adults" (Paxton 2000, 93–94).

494. When conceptualizing government action through a gendered lens there are two different approaches to identifying policies that take an explicitly feminist approach and assessing the degree to which public policy is gendered. In assessing government action that takes an explicitly feminist frame, in other words that promotes women's rights and strikes down gender hierarchies, a series of subsectors have been identified, e.g., blueprint, political representation, equal employment, reconciliation, family law, body politics, reproductive policy, and sexuality and violence to name some of the major subsections. Conceptualizing gender-specific public policy is a much broader endeavor and involves identifying policies that have a formal gendered content or goal but may not be necessarily feminist, e.g., target men and women's roles but may not be feminist. The study of gendered public policy may also involve conceptualizing how policy implementation and impact may affect men and women's lives and roles.

Questions and Issues:

- How are feminist policies differentiated from gender-specific policy e.g, reconciliation policy versus family policy?
- When is a feminist policy, actually feminist?
- Given the Western-focused nature of feminism, can feminist policies be studied outside of the Western countries?

Selected References:

Bacchi, Carol Lee. 1999. *Women, Policy and Politics: The Construction of Policy Problems*. Sage.

Mazur, Amy G. 2002. *Theorizing Feminist Policy*. Oxford University Press.

Conway, P. M, Ahern, And G.A. Steuernagel. 2004. *Women and Public Policy: A Revolution in Progress*. CQ Press.

Charles, Nickie. 2000. *Feminism, the State and Social Policy*. London: MacMillan.

Pollock, Mark and Amy G. Mazur (eds). 2009. Gender and Public Policy in Europe. special issue of *Comparative European Politics*. 9(1).

495. First introduced as a concept by the United Nations in the 1980s and then through the European Union via the Treaty of Amsterdam in the 1990s, gender mainstreaming has come to be identified with the systematic introduction of gender equality across all policy areas both in terms of formal policy statements and policy implementation. Gender mainstreaming is associated with specific policy instruments that compel governments to systematically introduce gender equality statements and principles in budgets, sectoral plans, etc. Seldom found as a policy approach in the USA, gender mainstreaming tends to mean introducing gender components in curricula in the US context.

    Questions and Issues:

    - To be considered gender mainstreaming does this policy approach have to go beyond formal statements to actually promoting gender-based equality?
    - To what extent does this concept travel outside of the European Union to countries, like the USA, where gender mainstreaming in policy is not a key approach.

    Selected References:

Mazey, Sonia. 2000. *Gender Mainstreaming in the EU*. Kogan Page.

Pollock, Mark and Haffner-Burton, Emilie. 2009. Mainstreaming Gender in the European Union: Getting the Incentives Right. *Comparative European Politics*. 9(1).

Walby, Sylvia. 2005. Gender Mainstreaming: Productive Tensions in Theory and Practice. *Social Politics* 12:321–343.

Woodward, Alison. 2003. European Gender Mainstreaming: Promises and Pitfalls. *Review of Policy Research* 20:65–88.

496. *Thanks to Kara Ellerby for this exercise.* Gender quotas are considered a "fast track" policy aimed at increasing women's representation within the legislative electoral process. They have gained immense popularity over the last thirty years, and are currently present in over one-hundred countries. Conceptually, quotas are best understood by two key characteristics: the degree to which they mandate women's access to public office (is it specified in the Constitution or did political parties create their own quotas?) and at which stage of the electoral process the quota mandates women's representation (does it specify how women are listed on candidate lists or are there reserved seats for women in parliament?). Quotas vary greatly within these dimensions,

but there seem to be patterns among regions. Gender quotas entail important discussions about sex/gender distinctions, descriptive and substantive representation issues and whether or not they really empower women.

Questions and Issues:

- What are the goals of gender quotas? Is it, to get more women in, provide women (and which women?) with more political empowerment, or both? How do secondary dimensions shift if scholars distinguish different goals?
- How does one theorize/operationalize a 'successful' gender quota?
- Are gender quotas meant to be a short-term or long-term policy?
- Is the use of 'gender' rather than 'sex' depoliticizing the very issues quotas are in place to address?

Table 19: Gender quota concept

Quota Type	Key Features
Party Quota	Adopted voluntarily by political parties
	(1) Set out new criteria for party candidate selection.
	(2) Affect composition of party lists in PR electoral systems and candidates eligible for particular seats in majoritarian systems.
	(3) May entail internal party sanctions for non-compliance.
Legislative Quota	Mandated by national parliaments
	(4) Set out new criteria for party candidate selection.
	(5) Affect composition of party lists in PR electoral systems and candidates eligible for particular seats in majoritarian systems.
	(6) Usually entail sanctions for non-compliance.
Soft Quota	Adopted voluntarily by political parties.
	(7) Set out informal targets and recommendations in relation to party candidate selection.
	(8) Set out new criteria for membership of internal party bodies.
	(9) Rarely entail sanctions for non-compliance.

Source: Krook, M. et al. 2014

In the table above discuss how adjectives are used to modify "gender quota." If one were to develop a single concept of gender quota instead of a typology, what would be the secondary-level dimensions based on the key features? Can you draw a diagram of the unified gender quota concept?

Selected References:

Dahlerup, Drude (ed.). 2006. *Women, Quotas and Politics*. Routledge.

Krook, Mona Lena. 2008. Quota Laws for Women in Politics: Implications for Feminist Practice. *Social Politics* 15:345–368.

Krook, Mona Lena. 2006. Reforming Representation: The Diffusion of Candidate Gender Quotas Worldwide. *Politics & Gender* 2:303–327.

Krook, M. et al. forthcoming. Gender Quotas and Models of Political Citizenship. *British Journal of Political Science* 39:781–803.

Caul, Miki. 2001. Political Parties and the Adoption of Candidate Gender Quotas: A Cross-National Analysis. *Journal of Politics* 63:1214–1229.

Squires, Judith. 2007. *The New Politics of Gender Equality*. Palgrave-Macmillian.

497. "Sexual violence," "gender violence" consists of the operation of adding an adjective to a standard concept. The concept focuses on forms of sexual violence that occur particularly often to women, such as rape or domestic violence, although more recently gender violence has been associated with homosexuals and transgendered groups. For this concept there is some variation on names. The United Nations, as well as the feminist literature on the topic, in most of its documents uses "violence against women."

    Questions and Issues:

    - When should the notion gender violence be used as opposed to violence against women?

    Selected References:

Weldon, Laurel W. 2006. Inclusion, solidarity and transnational social movements: the global movement against gender violence." *Perspectives on Politics*, 4:55–74.

Emilia L. Lombardi, Riki Anne Wilchins, Dana Priesing,Diana Malouf 2001. Gender Violence: Transgender Experiences with Violence and Discrimination. *Journal of Homosexuality* 42:89–101.

498. Gendering the concept of institutions is highly complex and involves the premise that institutions as organizations in society and in the state are inherently gender-biased. For some this is a question of research for others it is an underlying assumption. Bringing the study of gender into institutional analysis also implies that feminist agents of change, what some call feminist engagement with the state, will attempt to change the underlying gendered/patriarchal (sic) nature of institutions. The ultimate outcome in the process of these feminist engagements is that institutions reflect cultures of gender equality and account for the improvement of women's right in relation to men. The study of gender and institutions is often conducted in the context of seeking to gender new institutional scholarship. Questions and Issues:

    - To what degree is this a discrete analytical concept – e.g., gendered institutions – rather than an approach to institutional analysis?

    Selected References:

Acker, Joan. 1992. Gendered Institutions: From Sex roles to Gendered Institutions. *Contemporary Sociology*. 1:565–569.

Chappell, Louise A. 2002. *Gendering Government: Feminist Engagement with the State in Australia and Canada*. UBC Press.

Chappell, Louise. 2006. Comparing Political Institutions: Revealing the Gendered "Logic of Appropriateness." *Politics and Gender*. 2:223–225.

Lovenduski, Joni. 1998. Gendering Research in Political Science. *Annual Review of Political Science* 1:333–356.

499. A central concept to feminist analyses of politics, interestingly many empirical feminist analyses conducted in the purview of political science do *not* actually use the term, opting instead for notions of gender-bias, gender-based discrimination etc. In addition, while for many feminist theorists, patriarchy is an operating assumption of any organization or structure, for analysts who seek to conduct empirical studies of gender and politics the presence of patriarchy or not, is a question for future research. Nonetheless, the notion of patriarchy contains conceptual and theoretical ramifications that are important to the systematic study of gender and politics. While there are many articles on patriarchy and works that discusses patriarchy there are few that provide operational research definitions of the concept. Walby (1989) provides some useful perspective on operationalization, measurement and traveling.

    Questions and Issues:

    - If patriarchy is an underlying assumption of every organization, how then do researchers operationalize and measure it?
    - To what degree does the notion of patriarchy travel outside of western contexts?

    Selected References:

Walby, Sylvia. 1989. Theorizing Patriarchy. *Sociology* 23:213–234.

Simon, Duncan. 1995. Theorizing European Gender Systems. *Journal of European Social Policy* 5:263–284.

Randall, Vicky. 1987. *Women and Politics*. University of Chicago Press.

500. Analyzing power is core to the gender and politics literature given the notion that politics is traditionally organized to organize out gendered concerns and power relations are biased against women see patriarchy, for example. At the same time and as a result of its omnipresent nature, it is quite difficult to operationalize power from a gendered perspective for empirical research. For those who are interested in reading further on power and feminist analysis and an excellent overview of feminist theories of power we suggest the classic books by Judith Squires and Connell on the topic. Given the degree to which patriarchy is interlinked to power also see the above entry on patriarchy.

    Selected References:

Connell, Robert. 1987. *Gender and Power*. Cambridge: Polity Press.

Duerst-Lahti, Georgia and Kelly, Rita Mae. 1995. *Gender Power, Leadership, and Goverance*. University of Michigan Press.

Squires, Judith. 1999. *Gender in Political Theory*, chapter 1. Polity Press.

501. An oft-nebulous and contested concept in the study of comparative politics, social capital is related to the study of norms and values about civic engagement, quite often in a cross-national perspective. Gendering the concept, therefore, faces the same issues of imprecision and poor operationalization. Nonetheless, at least for feminist analysts, bringing in the interface between the social construction of men's and women's identities in relation to one another with the development of participatory engagement is an empirical and theoretical imperative. We list several works that place gender on the agenda of comparative work on social capital.

    Selected References:

    Stolle, Dietland and Lewis, Jane. 2002. Social Capital an Emerging Concept. In Hobson, Lewis, Siim (eds.) *Contested Concepts in Gender and Social Politics*, pp. 195–230. Edward Elgar.

    Molyneux, Maxine 2002. Gender and the Silences of Social Capital: Lessons from Latin America. *Development and Change* 3:167–188.

502. Linked to the concept of democracy, this concept is another example of an established concept engaging with gender. Here, the notion of regime transition from and to democracy undergoes gendering. Thus a more complex view of democratization is dealt with here than in the chapter on democracy in the book, with a more global frame that provides conceptualization that goes beyond Western democracies. Women as objects and actors are brought into the conceptual mix as well as taking on seriously gender as an analytical concept in the drivers and dynamics of transition. This gendered take dialogs with the large "transitology" literature in comparative politics. Whether this body of feminist literature provides a structured approach to the concept of gendered transitions remains to be seen given the early stages of the concepts genesis. Nonetheless, the selection of literature cited here provides a road map.

    Selected References:

    Alvarez, Sonia. 1990. *Engendering Democracy in Brazil*. Princeton University Press.

    Waylen, Georgina. 2007. *Engendering Transitions: Women's Mobilization, Institutions and Gender Outcomes*. Oxford: Oxford University Press.

503. For many obvious reasons war has been seen as a highly gendered activity. War, and more generally "national security" have been a focus of feminist scholars, not surprisingly within international relations. The organization of militaries, the conduct of war, discourses of national security, and many other topics have been the focus of a gender analysis.

    It is useful to consider the nongender concepts of war which usually contain two core sets of characteristics (1) nature of the actors, usually one of the actors must be a state, and (2) severity of militarized violence (typically numbers of soldiers killed in battle). One issue is how or should one include gender the basic concept of a war. Much work is concerned with gender and the

causes of war, the conduct of war, or the impact of war. As an exercise it is useful to survey the literature to see which of these has been the subject of a gender analysis, and how the gender analysis is conducted. One development has also been to move away from war to a more general analysis of "human security" (http://www.hsrgroup.org/). This has important implications for a gender analysis since it naturally brings in other issues such as gender violence during wars, health, etc.

Selected References:

Elshtain, Jean Bethke. 1987. *Women and War*. Basic Books.

Reardon, Betty. 1985. *Sexism and the War System*. Teachers College Press.

Cohn, Carol. 1987. Sex and Death in the Rational World of Defense Intellectuals. *Signs: Journal of Women in Culture and Society* 12:687–718.

Enloe, Cyntha. 2000. *Maneuvers: The International Politics of Militarizing Women's Lives*. University of California Press.

Tickner, Ann. 1992. *Gender in International Relations: Feminist Perspectives on Achieving Global Security*. Columbia University Press.

Goldstein, Joshua. 2001. *War and Gender: How Gender Shapes the War System and Vice Versa*. Cambridge University Press.

Young, Iris Marion. 2003. The Logic of Masculinist Protection: Reflections on the Current Security State. *Signs: Journal of Women in Culture and Society* 29:1–25.

Kaufman-Osborn, Timothy. 2005. Gender Trouble at Abu Ghraib? *Politics and Gender* 1:597–619.

Carpenter, R. 2003. 'Women and children first': gender, norms, and humanitarian evacuation in the Balkans 1991–95. *International Organization* 57:661–694.

504. "Women's interests" is central to the analysis of gender and politics. How do researchers define women's interests as opposed to men's; in other words a gendered approach to the notion of interests. More than dialoging with the concept of interests in nonfeminist political science which is quite diffuse, the notion of women's interests, sometimes referred to as gender interests, was developed separately in the literature on gender and politics. From Sapiro's first treatment of the issue to Molyneux's efforts to differentiate between strategic and practical interests, conceptualization of women and/or gender interests involves identifying what, if any are the common political interests of men as opposed to women. A central component of developing operational definition of gendered interests has been the diversity of women and men as groups-across race, class, ethnicity, sexual orientation, religious affiliation etc. Thus, conceptualization of women's interests have increasingly become more sophisticated over-time, taking on the reality of intersectionality in political mobilization.

Selected References:

Diamond, I. and Hartsock, N. 1981. Beyond Interests in Politics: a Comment on Virginia Sapiro's "When are Interests Interesting? The Problem of Political Representation of Women." *American Political Science Review* 75:717–21.

Molyneux, Maxine 1985. Mobilization with Emancipation? Women's Interests, the State and Revolution in Nicargaua. *Feminist Studies*. 11:227–255.

Sapiro, Virginia. 1981. Research Frontier Essay: When are Interests Interesting: The Problem of the Political Representation of Women. *American Political Science Review* 7:701–716.

Jónasdóttir, A.G. 1988. On the Concept of Interests, Women's Interests and the Limitation of Interest Theory. In K.B. Jones and A.G. Jónasdóttir (eds.) *The Political Interests of Gender*, pp. 33–65. Sage Publications.

Typologies

505. Chapter 8 of the concept book discusses the standard rule for typologies that they be mutually exclusive and exhaustive and argues that these standard methodological rules are not good ones. Discuss the typology in the large-scale database on social conflicts described in the previous exercise, e.g., food riot, slave revolt, political conflict, etc. Is the typology meant to be mutually exclusive and exhaustive? Are the authors assigning a given conflict only one category? How is that problematic, e.g., conflict fitting multiple categories.

Table 20: List of variables included in the HiSCoD database

TABLE 2. Classification Used in the HiSCoD Database

Category	Numeric values	Description	No. of obs.
Food riot	1	Includes riots protesting against high prices, popular taxation, interception of food transport, and food looting	4,226
Tax riot	2	Includes smuggling-related events, and riots protesting against taxes (local, feudal, or national)	4,301
Religious conflict	3	Includes conflict relating to the practicing of religion	649
Conflict with local or national authorities	4	Includes conflict with the military (including deserters and draft dodgers), police, marshals, parliaments, municipal authorities, militias or any conflict protesting against a decision taken by authorities	3,550
Feudal conflict	5	Includes attacks on castles and conflicts concerning the abolition of feudal rights	940
Slave revolt	6	Includes slave revolts either on boats, at gatherings, and/or maroons	3
Political conflict	7	Includes all conflicts protesting against a political regime during a revolutionary or transitional period, and any opposition to a political regime aiming to change it	1,088
Labor conflict	8	Includes strikes, collective bargaining, conflict against guilds or between different guilds, wage riots, etc.	1,403
Banditry	9	Includes robberies by a group of individuals or events related to organized crime	0
Other forms of conflict	0	Includes all conflicts which do not fit in previous categories	4,458

Note: We categorized events related to seigneurial taxes as tax riots rather than feudal conflict. In 333 instances, the lack of information in the sources pre-empted the attribution to an exact category.

Source: Chambru and Haneuvrier-Hervieu, 2024.

Answer: While not completely clear it seems like the authors are applying the mutually exclusive and exhaustive rule: "We addressed this issue by creating

a simplified classification based on 10 categories to encompass a wide range of geographic areas and a long-time span (Table 2). We assigned each event to one category based on the information available and, when relevant, on the researchers' original typology – which is also included in the database" (Chambru and Maneuvrier-Hervieu 2024, 1086).

506. The UCDP project has as its goal to collect data on all forms of "organized violence." It has a partial typology which they claim consists of three mutually exclusive categories. Is this along with other of the UCDP datasets an exhaustive categorization of organized violence? This is significantly more complex than one might imagine if one examines the definitions (https://www.pcr.uu.se/research/ucdp/definitions/). What might be missing from the typology? (Thanks to Peter Wallensteen for comments about this exercise.)

> Organized violence 1989–2017. For the third year running, the annual update from UCDP presents trends in not only state-based armed conflict, but also non-state conflict and one-sided violence. The three categories are mutually exclusive and can be aggregated as 'organized violence'. They also share the same intensity cut-off for inclusion – 25 fatalities in a calendar year. State-based armed conflict includes violence where at least one of the parties is the government of a state, that is, violence between two states and violence between the government and a rebel group. An example of the former is the border conflict between Ethiopia and Eritrea, while the conflict between the Taliban and the Afghan government is an example of the latter. Non-state conflict, on the other hand, is the use of armed force between two organized groups, such as rebel groups or ethnic groups, neither of which is the government of a state. Examples include fighting between the Islamic State (IS) and Tahrir al-Sham in Syria, as well as the interethnic fighting between the nomadic pastoralist Fulani and the mainly agriculturalist Mambila in Nigeria. Finally, one-sided violence covers violence by the government of a state or by a formally organized group targeting unarmed civilians. Recent examples include the CPI-Maoist targeting civilians they consider enemies of their struggle, and the government of Kenya killing protestors following elections. (Pettersson and Eck, 2018, 535)

507. The UCDP project has as its goal to collect data on all forms of "organized violence." It has a partial typology which they claim consists of three mutually exclusive categories. Is this along with other of the UCDP datasets an exhaustive categorization of organized violence? This is significantly more complex than one might imagine if one examines the definitions (https://www.pcr.uu.se/research/ucdp/definitions/). What might be missing from the typology? (Thanks to Peter Wallensteen for comments about this exercise.)

> Organized violence 1989–2017. For the third year running, the annual update from UCDP presents trends in not only state-based armed conflict, but also non-state conflict and one-sided violence. The three categories are mutually exclusive and can be aggregated as 'organized violence'. They also share the same intensity cut-off for inclusion – 25 fatalities in a calendar year. State-based armed conflict includes violence where at least

> one of the parties is the government of a state, that is, violence between two states and violence between the government and a rebel group. An example of the former is the border conflict between Ethiopia and Eritrea, while the conflict between the Taliban and the Afghan government is an example of the latter. Non-state conflict, on the other hand, is the use of armed force between two organized groups, such as rebel groups or ethnic groups, neither of which is the government of a state. Examples include fighting between the Islamic State (IS) and Tahrir al-Sham in Syria, as well as the interethnic fighting between the nomadic pastoralist Fulani and the mainly agriculturalist Mambila in Nigeria. Finally, one-sided violence covers violence by the government of a state or by a formally organized group targeting unarmed civilians. Recent examples include the CPI-Maoist targeting civilians they consider enemies of their struggle, and the government of Kenya killing protestors following elections. (Pettersson and Eck, 2018, 535)

508. A constant issue with typologies is the nature of the dependent variable in the cells. Mazzuca's typology, see below, is his core framework: "The typology is the first step in the creation of a more general theory of state formation. It transforms scope conditions implicit in the theory of state formation advanced by Weber and his contemporary followers into explicit causal variables. The typology also helps to enrich neo-Weberian approaches to state formation with valuable insights from neo-Marxist alternatives." (Mazzuca 2021, 42). What is the dependent variable? Is it "state formation" i.e., all are cases of $Y = 1$ but different paths for each cell, or are there four separate Ys? What about $Y = 0$?

Figure 57: Typology of state formation

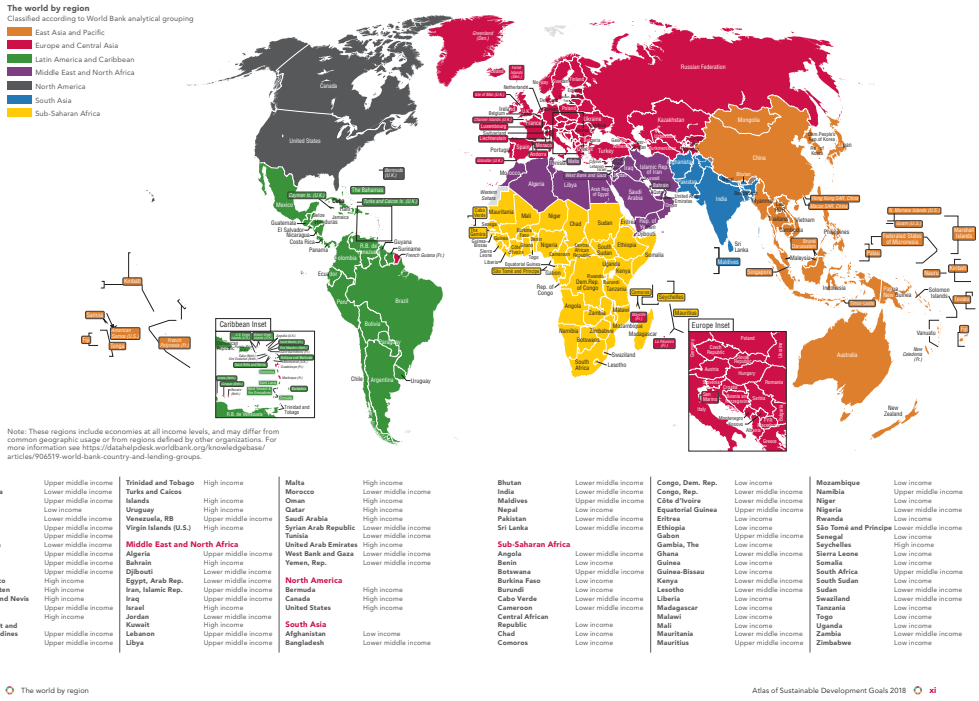
Table 1.1. Geopolitical and International Economic Environments of State Formation

		International economy	
		Feudalism/ mercantilism	*Capitalism/free trade*
Geopolitical context	*Anarchy*	Pioneer state formation	Competing state formation
		War-led (with military innovation)	Trade + war-led
		Western Europe	*United States*
	Hierarchy	Reactive state formation	Latecomer state formation
		War-led (with military imitation)	Trade-led
		Eastern Europe, China, Japan	*Latin America, Africa*

Source: Mazzuca (2021).

509. Chapter 8 the concept book discusses some of the pitfalls of creating typologies using the mutually exclusive and exhaustive rules. The map given in the figure below is the World Bank's typology of regions of the world and is an interesting example to discuss for thinking about mutually exclusive and exhaustive typologies. As far as I can tell any rationale for this particular division of the world into regions is lost in some old World Bank filing cabinet. Here are a few issues to discuss. First, note that it is not completely geographic because French Guyana which is located Latin America is coded red making it part of Europe. However, various colonies and possessions, e.g., US and French, in Oceania (another region?) are considered in the East Asian region. Probably Puerto Rico is considered part of North America. Second, probably the biggest problem is those countries that lie on the boundaries, the gray zone between regions. So for example Turkey is considered part of Europe. Russia is obviously a huge gray zone because it is part of Central Asia, Europe and arguably East Asia. A major gray zone is how to consider North Africa, part of Africa or part of the Middle East? It seems like language and religion are used as well in deciding regions? Finally, why did they consider Europe and Central Asia to be one region?!

Figure 58: World Bank and regions of the world



510. In a very prominent textbook on the art of thinking used in many philosophy classes Kelley discusses the classic rules for constructing a typology, that they are mutually exclusive and exhaustive. Since he uses the example of classes, what principle used by all universities for classes indicates that the mutually exclusive principle does not work in practice?

> Because your categories overlap, we don't know whether to classify this as an introductory course or an art course. The first rule of classification, then, is that the species must not overlap. We express this in logic by saying the species must be mutually exclusive: Each species must exclude all the members of every other species. At the same time, a good classification divides up the genus completely, allowing us to assign every member of the genus to one or another of the species. We express this in logic by saying that the classification must be jointly exhaustive: the species taken together (jointly) must cover (exhaust) all the objects in the genus. (Kelley 2014, 13–14)

Answer:

Cross-listing classes means a course fits into more than one category.

511. Sometimes concepts are basically defined via typologies. This raises the question of there is some kind of basic level concept. The idea of a positive or negative pole makes little sense if the typology is really of different kinds, i.e., apples and oranges. As the following two questions also illustrate, a

key assumption in concepts is that they are causally homogeneous: they have the same causes and effects. The dominant Correlates of War (see also the ATOP alliance data set, Leeds et al. 2002) defines "alliance" (Singer and Small 1966; Gibler and Sarkees 2004) in terms of three types of alliance: (1) defense, (2) neutrality-nonaggression, (3) entente. Discuss whether these three types of alliance have potential problems of causal heterogeneity. For a concrete example of this criticism see Long's (2003) analysis of the alliance-trade relationship.

Answer:

Many have had their doubts about whether the neutrality-nonaggression alliances belong. Typically in the liberal peace literature the alliance variable excludes them. See the ATOP coding manual (Leeds 2005) for additional reservations.

512. Compare Rawls and his typology of war with UCDP.

> Rawls distinguished between different types of wars in order to define the principles that would best apply in each case. The initial typology found in his lecture notes proposes nine kinds: 1. Wars between existing via states (WW I + II) 2. Civil wars (of social justice) within via states or society (French Rev); 3. Wars of secession of minorities within region: American Civil War. 4. Colonial Wars of secession (from Empire): Algerian War; American Rev War? 5. Wars of intervention (humane intervention) 6. Wars of national unification (War of Roses; Tudors) 7. Wars of conquest, of Empire (Wars of Rome). 8. Wars of Crusade, religious or secular 9. Wars of national liberation (in present sense); guerilla wars (Armitage 2017, 335)

513. Concepts with adjectives often came become typologies. For example, gender quotas come various forms, "party," "legislative," "constitutional," "reserved seat" etc. which could generate "types of gender quotas" (Krook 2014). The goal of gender quotas is to raise the number of women in parliament. As such the typology is really a series of causal mechanisms designed to raised the number of women in parliament. Discuss why the rule of mutually exclusive categories for typologies is particularly a bad idea in this concept.

Answer:

Clearly a country could, and maybe should, adopt multiple mechanisms for raising the number of women in parliament. In addition there might be synergies between these mechanisms when used together.

514. The well-known UCDP data projects have the goal of basically mapping all forms of armed conflict, both domestic and international. Discuss the extent to which their typology is mutually exclusive and exhaustive. Are there some forms of armed conflict that are not covered by their scheme? Note that this is a tricky and complicated issue.

> Organized violence 1989–2017 For the third year running, the annual update from UCDP presents trends in not only state-based armed conflict, but also non-state conflict and one-sided violence. The three categories are mutually exclusive and can be aggregated as 'organized violence'. They

also share the same intensity cut-off for inclusion – 25 fatalities in a calendar year. State-based armed conflict includes violence where at least one of the parties is the government of a state, that is, violence between two states and violence between the government and a rebel group. An example of the former is the border conflict between Ethiopia and Eritrea, while the conflict between the Taliban and the Afghan government is an example of the latter. Non-state conflict, on the other hand, is the use of armed force between two organized groups, such as rebel groups or ethnic groups, neither of which is the government of a state. Examples include fighting between the Islamic State (IS) and Tahrir al-Sham in Syria, as well as the interethnic fighting between the nomadic pastoralist Fulani and the mainly agriculturalist Mambila in Nigeria. Finally, one-sided violence covers violence by the government of a state or by a formally organized group targeting unarmed civilians. Recent examples include the CPI-Maoist targeting civilians they consider enemies of their struggle, and the government of Kenya killing protestors following elections. (Pettersson and Eck, 2018, 535; see https://www.pcr.uu.se/research/ucdp/definitions/)

Answer:

For example, one form of conflict not covered by UCDP datasets is that between organized crime and the government. Armed conflict between organized crime groups is covered.

515. Collier et al. 2008 use the following example of a typology. Critical in thinking about typologies is the scale type of the the row and column variables. This table clearly has ordinal variables. How would you change the format to reflect better the ordinal nature of the scale to make it similar to the familiar Cartesian coordinate system?

Figure 59: Typology dimensions and orientation

Table 7.2. Matland's Typology of Policy Implementation

		Conflict: Low	Conflict: High
Ambiguity	Low	Administrative Implementation	Political Implementation
	High	Experimental Implementation	Symbolic Implementation

Source: Adapted from Matland (1995).

Answer:

To make the ordinal nature clearer and to make it similar to the Cartesian system, one should flip "high" and "low" on the Y-axis (ambiguity variable) so that low is near the origin.

516. Traditional typology rules have the core mutually exclusive and exhaustive rules. This is an issue in concepts of ethnic group. Suppose there are two ethic concepts "white" and "black." In fuzzy set one could be a .4 member of white AND .4 member of black. Discuss the extent to which various datasets on ethnicity (e.g., Fearon 2003; Chandra 2008) assume coding of ethnic groups is exclusive. Discuss how this plays a central role in the widely used ethnic fractionalization index: $F = 1 - \sum_{i=1}^{n} p_i^2$ where p_i are the population shares of each ethnic group.

    Answer: Because fuzzy logic does not think strict dividing lines are good practice, it naturally will allow, and even encourage, partial membership in multiple ethnic groups. Because it is not probability based, there is no need for things to add up to one.

    The fractionalization index *assumes* that membership in ethnic groups is (1) zero or one, and (2) that you cannot be a member of two ethnic groups at the same time. In short the sum of the population shares must be one.

Multimethod research

517. Goertz (2017) proposes the Overdetermination Rule for case selection. This suggests that that one should choose cases that have value zero on alternative explanations, Z. Some scholars who like "dueling theories" prefer to choose cases that have one on alternative explanations Z. Bell suggests another criteria. Could you make this more formal?

    > I choose cases based on two primary criteria. First, the three cases I use each provide hard cases for the theory. Picking hard cases allows for more confidence in the broader applicability or "external validity" of the findings – if we find support for the theory despite picking cases that we expect the theory will have difficulty explaining, it increases the likelihood that the theory will have some success in cases we do not examine in detail, or in cases that may emerge in the future. In particular, I look for cases with strong "countervailing conditions" – variables whose presence in a particular case makes it less likely that the outcomes posited by the theory of nuclear opportunism will be observed. (Bell 2021, 33)

    Answer: He is basically suggesting that one should choose cases that have opposite values on alternative explanations. So if X is positively related to the outcome and when should you use cases where Z is expected to have a negative effect and with a large value on Z, or the opposite if X is negatively related.

518. Interviews and focus groups often generate "causal claims," e.g., something happened *because* of X. These causal claims are often not considered distinct from "descriptive claims" about facts. Go through Widner's chapter

"Descriptive accuracy in interview-based case studies" and locate the causal claims in her discussion. Is there a pattern to where they seem to arise in her discussion? Here are a couple of examples of causal questions to get going with:

> How would you say that joining the deliberation/process/negotiation late colored your view of the issues and shaped your actions, or did it not make much difference?
>
> The military built emergency treatment centers that were never used because the epidemic ended by the time the centers were ready. The US military action was irrelevant. . . . Later the epidemic declined and the ETUs [emergency treatment units] weren't used, but in the end what seemed to matter to the public was the visible sign that a big power cared, which generated a psychological boost. (Widner 2022, 125)

519. Cammett (2022) argues that positive deviant cases are good candidates for case studies. Deviance is defined as the residual between the regression line and the observation. Can you give one more specific way to identify these cases based on her discussion below that might be called "surprising success cases"?

> What are Positive Deviant Cases? Positive deviant cases are outliers that exhibit superior performance than the predictions of a model would hold. In technical terms, this refers to cases selected to maximize $|Y_i - \hat{Y}_i|$, or the difference between the actual value and the fitted value in a regression. In order to qualify as cases of positive deviance, however, cases must be more than ephemeral outliers operating in an environment conducive to good performance. Rather, examples of positive deviant organizations or other types of collective actors must exhibit sustained high performance in a context in which good results are uncommon. (Cammett 2022, 220)

Answer: Choose cases (1) score high on Y, (2) score high on X and (3) score high on Zs negatively related to the outcome. These might be called the "surprising success cases."

520. With complex Boolean equations case selection can be quite challenging. Discuss Brookes's case selection strategy. Is she using the Overdetermination Guideline in selecting her cases? Which cases are good for the necessary condition hypothesis and which for sufficiency? With four variables in her model there are 16 possible combinations for case study selection. She choose 6. Which ones does she leave out? Discuss in the context of the Overdetermination Guideline.

> *Case Selection and Methods of Causal Analysis.* To test the causal claims of the CCAP theory, in this book I analyze six TLA campaigns – three failures and three successes – matched into pairs of highly The New Politics of Transnational Labor similar cases with different outcomes: the Liverpool dockers' dispute versus the Australian waterfront conflict; the Tesco Global Union Alliance versus the Alliance for Justice at G4S; and the Shangri-La Hotel campaign versus the Raffles Hotels campaign. This logic of case selection approximates John Stuart Mill's method of difference insofar as it allows one to control for otherwise confounding factors within each pair, including the type of conflict prompting the campaign,

> the institutional context of the country in which the conflict originated, the TLA's goals, the type of employer targeted, whether the campaign was proactive or reactive, and the time period in which the campaign took place. If the CCAP theory is correct, any campaign lacking intraunion coordination, interunion coordination, or context-appropriate power should be unsuccessful. Conversely, all three variables should be present in each successful campaign. . . . Process tracing thus allows me to identify and test for alternative explanations. In sum, across-case comparisons are important, but only within-case analyses can confirm or falsify the theoretical framework proposed in this book. . . . *Causal Relationships as Boolean Equations* $X1 * X2 * X3 = Y$; $\neg X1 + \neg X2 + \neg X3 = \neg Y$ (Brookes 2019, 34–35, 160;)

521. Discuss the same issue in Kaufman's analysis which includes six case studies. Which combinations of independent variable and dependent variables are selected for case studies? Are core configurations omitted?

> To sum up, the symbolic politics theory presented here posits that the way relations between ethnic groups play out in any country depends on four main factors: *symbolic predispositions, perceived threat, leadership, and organization*. . . . . Put more formally, the theory states that threat perceptions, credible chauvinist leadership, and effective organization are all necessary conditions for ethnic mobilization. Significant ethnic mobilization does not happen unless all three factors are present. Increases in any of them – perceptions of stronger threat, more credibility for chauvinist leaders, or better organization – increase the likely scale and intensity of ethnic mobilization. Together, these three factors are also sufficient for ethnic mobilization: if all three are simultaneously present, ethnic mobilization will occur, though violence may not erupt. While prejudice is neither necessary nor sufficient, it is a key motivating factor, at least on one side, in the vast majority of cases. (Kaufman 2015, 12, 22)

522. To achieve clarity in case selection it is often very useful to choose cases that avoid overdetermination on key causal mechanisms. Discuss whether this makes sense in the case selection of the fight against corruption which cannot be the result of external pressure.

> Instead of looking at poorer countries that may have been forced to act against corruption by outside political pressure (e.g., as a condition of World Bank loans), the book focuses on four rich countries where such pressure has played little or no role. The United States, Britain, Switzerland, and Australia have freely made demanding commitments to detect and return looted wealth from abroad, and it is difficult to see these promises reflecting outside coercion. . . . Why concentrate the search for illicit funds on the United States, Switzerland, Britain, and Australia? (Sharman 2017, 7, 17)

523. Ford (2019) in his analysis of labor regimes, to simplify, has three independent variables each with three levels, see his table below (Ford 2019, 17). This would produce eight possible combinations of variables for case selection. Discuss which ones are selected and which ones have no case study.

TABLE 1. Variation in labor migration regimes

	EXTENT AND NATURE OF LABOR MIGRATION		COMPLEXITY OF LABOR MIGRATION SCHEME		RIGIDITY OF LABOR CONTRACTS	
COUNTRY	NUMBER OF REGULAR LABOR MIGRANTS	NUMBER OF IRREGULAR LABOR MIGRANTS	REGULATION OF COUNTRY OF ORIGIN	REGULATION OF SECTOR OF EMPLOYMENT	ABILITY TO RENEW CONTRACT IN-COUNTRY	ABILITY TO CHANGE EMPLOYER
Hong Kong	High	Low	Medium	Low	High	Medium
Japan	Low	Low	Medium	High	Low	Low
Malaysia	High	High	High	High	High	Low
Singapore	High	Low	High	High	High	Medium
South Korea	Medium	Medium	High	High	Medium	Medium
Taiwan	Medium	Low	High	High	Low	Medium
Thailand	Low	High	High	Low	Medium	Medium

524. Often one claims that case studies should be representative of the population (e.g., Gerring 2006). Discuss what might constitute a " representative" case. Is that what Gavin is choosing? Discuss his view of generalizing from case studies.

> As a historian interested in these questions, my way of assessing Sechser, Furhmann, and Kroenig's arguments is straight-forward. I would identify the most important example where these issues are engaged, look at the primary documents, see how the authors 'coded crucial variables and determine how good a job their analysis does in helping us understand both the specific crisis itself and the larger issues driving nuclear dynamics. Political scientists might describe this as running both a 'strong' and a 'critical' test; in other words, *if the authors' theories don't fully explain the outcomes and causal mechanisms in the most important and most representative case*, how useful are the findings in explaining the broader issues? Is there such a case? Is there such a case? In a speech on November 10th, 1958, Soviet Premier Nikita Khrushchev demanded the Western powers – the United States, Great Britain, and France – remove their military forces from West Berlin within six months. This ultimatum was the start of a tense, four-year period that many believe brought the world closer to thermonuclear war than any time before or since, culminating in the Cuban Missile Crisis of October 1962. According to a leading historian of post-war international politics, "(T)he great Berlin crisis of 1958 to 1962" was "the central episode of the Cold War."16 And as McGeorge Bundy states, "there were more than four years of political tension over the future of Berlin. . . . Khrushchev's Berlin crisis gives us what is otherwise missing in the nuclear age: a genuine nuclear confrontation in Europe." . . . Most importantly, I would want to be convinced that the causal mechanisms identified by the authors did in fact drive the origins, development, and outcome of this crisis. . . . And if these models can't tell us anything about arguably the most important and consequential nuclear standoff in history, should I take comfort that it apparently can explain why the U.S. successfully restored Jean-Bertrand Aristide in 1994 or "won" in Nicaragua in 1984? Or look at the 1983 'Able Archer' affair, perhaps the most recent

> case where the risk of thermonuclear was possible (if highly unlikely). (Gavin 2014, 16, 24 emphasis is mine)

525. Discuss how Saylor is applying the Avoid Overdetermination Rule (Goertz 2017) in doing his case selection.

> Above all, I employ process tracing to indicate the viability of a coalitional approach to explaining state building trajectories during resource booms. . . . My focus on weakly institutionalized polities eliminates the possibility that robust institutions account for the way in which a country managed its commodity boom. I furthermore examine these countries' first major commodity booms, as a means to eliminate potential prior "learning effects" that might have affected policymakers during subsequent booms. Third, I limit my analysis to agrarian societies, even though some industrialized countries, like Norway, are noteworthy primary commodity exporters. I exclude industrial states from consideration because they are more likely to possess strong institutions. A focus on developing societies helps to isolate the influence of coalitional politics. (Saylor 2014, 40–41)

526. This case illustrates the importance of the Avoid Overdetermination Rule. An alternative to the importance of nuclear weapons is clearly conventional weapons superiority. In many circumstances it is likely the case that nuclear weapons are accompanied by conventional weapons superiority. How many, if any, of the cases satisfy the Avoid Overdetermination Rule? Discuss what this means in the statistical analyses if essentially the nuclear weapons variable is almost colinear with the conventional weapons variable.

> Conclusion. The chapter critically assessed ten apparently successful cases of nuclear coercion. These cases should have provided the clearest evidence in favor of the view that nuclear blackmail works. However, none of them unequivocally supports the nuclear coercionist school. In each case, there is some doubt – often considerable doubt – about whether nuclear weapons provided states with coercive leverage. Indeed, several of these crises provide strong support for nuclear skepticism theory, highlighting the challenges that countries face when attempting to employ nuclear signals for coercive purposes. It is worth noting that the cases that provide the strongest support for the nuclear coercionist view – including the Cuban missile crisis – happened in the early days of the Cold War. The world has changed dramatically since that time. It is hard to argue that cases from the 1950s and 1960s tell us much about the role of nuclear weapons in world politics today. There is scant evidence that nuclear blackmail has worked since the collapse of the Soviet Union. The nuclear crises of the last quarter-century illustrate the coercive limits, rather than the virtues, of nuclear weapons.

527. An issue in multimethod research is the location of the case study vis-à-vis the independent variable in the statistical analysis. Lee (2018) chooses a case study which scores very low on her independent variable of rivalry (at best a low level rivalry). "I complement the quantitative analysis with a qualitative examination of Malaysias role in undermining domestic sovereignty in the 1970s in Mindanao, a region of the southern Philippines." Given that this is a low level rivalry how might she defend her case selection?

Answer:

She could use an extrapolation justification: if it works for weak rivalries it should work for more serious rivalries. See figure 8.2 in Goertz 2017.

528. Grewal and Kureshi (2019) is an example of multimethod research. The main hypothesis involves an interaction term. Discuss the selection of the two case studies in light of the issues posed by an interaction term in the theory-statistical analysis.

529. Law and norms are often considered constraints. What would be good case studies to study this constraint mechanism?

> The threat that legalistic trade dispute settlement poses to the discretion of political leaders is threefold. First, it may constrain their ability to manage the unforeseen costs of adjustment, making it more costly to provide relief or protection to specific groups injured by trade liberalization. Second, it may limit their general policy autonomy across a range of domestic regulations, which it judges against treaty commitments to eliminate nontariff barriers to trade. A third and final consideration is that the delegation of authority to third parties may constrain their ability to pursue trade policy bilaterally, a strategy with distinct political advantages. On all three counts, political leaders in trading states are risk-averse regarding the impact of dispute settlement on policy discretion. Other things being equal, they do not want to cede veto power over domestic policies to appointed trade law experts or judges, because the political price of doing so may be high. (McCall Smith 2000, 143)

530. Gerring and Cojocaru describe Mansfield and Synder's work as follows. Using the discussion of large-N qualitative testing in *Multimethod research, causal mechanisms, and case studies: an integrated approach* explain what is wrong with their characterization.

> In a setting where the relationship between X and Y is well established – perhaps as a result of cross-case analysis (the researcher's or someone else's) – the pathway case is usually focused specifically on causal mechanisms (M). An example is provided by Mansfield and Snyder's (2005) research on regime transitions and war. The authors find a strong relationship between democratization and bellicose behavior in their large-n cross-national analysis. To ascertain whether their hypothesized causal mechanisms are actually at work in generating this relationship, they look closely at 10 countries where the posited covariational pattern between X and Y clearly holds, that is, where democratization is followed by war. (Gerring and Cojocaru 2016, 406)

Answer:

Their core hypothesis is an interaction between democratization and weak institutions, not democratization alone. They do *not* find a strong relationship between democratization alone and militarized conflict.

531. Should one assume an interaction term or just a straight additive model, or a Boolean model for the data in the table below? The table below is a stripped-down version of their table 2 (p. 88) which gives the two core

independent variables: (1) federal versus centralized state and (2) multiparty system versus two-party system. I use their primary dependent variable "macrocorporatism" (the other one they discuss is "sector coordination"). There is no statistical analysis in this article; the empirical analysis is carried out via four case studies, one from each cell in table **??**.

Table 21: Macrocorporatism

Number of Parties	Level of State Centralization	
	Centralized	Federal
Multiparty	Denmark macrocorporatism=.72	Germany macrocorporatism=.50
Two party	Britain macrocorporatism=–.14	United States macrocorporatism=–.93

Source: Martin and Swank 2011, table 2.

Answer:

As in almost all two-variable models – i.e., X_1 AND $X_2 \rightarrow Y$ – the (1,1,1) cell is clear as is the (0,0,0) cell. The critical feature that will influence the decision between a QCA interpretation against an additive linear one is what happens with the cases where one independent variable is present and the other is absent, i.e., (1,0) or (0,1).

In set theoretic models the value of the (1,0) and (0,1) cells would be close to that of (0,0) because of the necessary condition relationship. The additive model suggests that when one independent variable is present and the other absent we should see intermediate results, somewhere between the extreme values on the dependent variable. Hence, a signal that the additive relationship is at work is when the dependent variable is significantly greater than the (0,0) but yet significantly less than the (1,1) dependent variable value. This is in fact what we find in table **??**. The two off-diagonal cells have values in between those in the on-diagonal cells.

It is pretty clear that a set theoretic interpretation of these data would not be appropriate because of the intermediate values of the dependent variable when only one independent variable is present. However, it is very difficult to make any strong conclusions regarding whether there should be an interaction term.

532. In a very influential article Alesina and Spolaore (1997. On the size and number of nations. *Quarterly Journal of Economics*) give a formal model. Discuss it in terms of game theory multimethod research. For example, how plausible are their assumptions? What is the causal mechanism and how would you explore that in case studies? You can explore this in more detail in their book on the topic.

Answer:

For example, they do not discuss war which is historically probably the most important determinant of the size of nations.

533. Copeland (2015) has at the core of his book an interaction; positive levels of economic interdependence (often trade) interacted with expectations about future economic conditions. Discuss the situation where one of the interaction terms is basically a subset of the other. For example, what case studies we be appropriate? Is the superset variable (positive economic independence) more like a scope condition?

> To determine whether the liberal prediction or realist prediction will prevail, we must introduce an additional causal variable – namely, a state's expectations of the future trade and investment environment. When a dependent state has positive expectations about this future environment, it is more likely to see all the benefits of continuing the current peace and all the opportunity costs of turning to war. Economic interdependence would then be a force for peace. Yet if a dependent state has negative expectations about the future economic environment – seeing itself being cut off from access to foreign trade and investment, or believing that other states will soon cut it off – then the realist logic will kick in. Such a state will tend to believe that without access to the vital raw materials, investments, and export markets needed for its economic health, its economy will start to fall relative to other less vulnerable actors. If this economic decline is anticipated to be severe, the leaders of the dependent state will begin to view war as the rational lesser of two evils – that is, as better than allowing their state to fall to a point where rising states can attack it later or coerce it into submission. (Copeland 2015, 2; basic hypothesis of the book)

534. Seawright (2016) argues for "extreme on X" case selection in multimethod research. Extreme on X could mean very high or very low or distant from the mean. Contrast that with "good" causal mechanism case selection as argued in *Multimethod research, causal mechanisms, and case studies: an integrated approach*, i.e., some extreme cases are good cases and others are not good cases.

535. Bush (2011) is one of the relatively rare statistical multimethod articles, which chooses a case study based on the statistical results. Her case study is Afghanistan:

> A case study can check the results' robustness through process tracing. . . . Since the preliminary large-N analysis generated satisfactory results for my theory, I chose Afghanistan – a country that was well-predicted by the statistical model (footnote: Lieberman 2005, 442–43. The average residuals for Afghanistan from Model 4 in Table 2 were –0.03) (Bush 2011, 126)

Discuss the pro's and con's of this as a good on-line case study – in the qualitative logic sense as well as in the statistical multimethod sense.

536. Collier and Sambanis (2005) conducted a large multimethod project on civil war and case studies. Explore their rationale for selecting cases to analyze since they cannot look at all countries for all time periods.

537. When doing statistical multimethod work scholars often choose "theoretical on-line" cases, i.e., cases where the theory is working (aka (1,1) cases, assuming a positive relationship). Construct an example where the theoretical on-line case is one which is not on the regression line, and which has a very large positive residual vis-à-vis the regression line.

    Answer:

    The ideal theoretical on-line case would lie in the top right hand corner (i.e., $X = 1$ and $Y = 1$, assuming variables [0,1]). Draw a regression line with a modest slope (e.g., .25). Cases in the top right corner will have large positive residuals.

Case studies, case selection

NOTE: the next section on the Possibility Principle is all about case selection as well. See also the multimethod section for case selection exercises.

538. Hulme (2026) argues that Congress is a strong constraint on presidential uses of force. He has a nice scatterplot illustrating that congressional approval is a necessary condition. One of the key themes in the qualitative methods literature as well as quantitative is alternative explanations and confounders. Hulme reports (personnel communication) that many suggest that an alternative explanation is the president's view on the use of force. So this could be a control variable or confounder. Goertz (2017) argues that in terms of case selection one should follow the "avoid overdetermination rule". Which cases in Hulme are overdetermined? Which cases then should be chosen when applying the avoid overdetermination rule? Notice the same issues arise in Tannenwald's influential book on nuclear weapons taboo. (Thanks to Hulme for comments on this exercise.)

539. Goertz (2017) provides the "Overdetermination rule" for case selection. Discuss Beckley's use of overdetermination as a criterion to exclude cases:

> The History of Peaking Powers since 1870. To assess how rising powers respond to severe and sustained economic growth slowdowns, I analyze every case from 1870 to 2018 in which a great power's per capita gross domestic product (GDP) grew at least twice as fast as the global average for at least seven years and then suffered at least a 50 percent decline in growth rates over the next seven years. . . . This exclusion leaves nine cases, which are displayed in table 1. For each case, I list the year in which the great power's average annual growth rates started to slow. Because they are based on imprecise data, these peak years are rough approximations of when the great power transitioned from rapid to slow growth. I therefore use these years only to identify cases for further research, not to conduct precise quantitative tests. I examine all the cases in detail except interwar France, Germany, and the Soviet Union, because mercantilist expansion was arguably overdetermined in those cases. (Beckley 2023, 24–25)

540. Generalization is core to LNQA and usually a central concern for case studies research in general. If one is only going to do a few cases in a case study

research design then the selection of cases becomes critical in terms of generalization. In particular what is the role of other variables, Z, the in the selection of these cases (discussed at length in Goertz 2017). The popular paired comparison which follows matching looks for cases with the same values for Z. So implicitly generalization is limited to the range of those Z that match which could be very narrow or problematic if only 2–3 pairs are chosen. Generalization is however not a criterion for paired designs. However, if one is interested in generalization then the logic suggests choosing cases which are very different on Z the to show that X is a cause of Y in a variety of settings. Discuss how Stuart chooses cases with generalization in mind. Here is a sample of her case study selection discussion:

> Beyond the ELF, EPLF, and SPLM/A, the next two cases demonstrate that the imitation of the Chinese model of governance is not simply a function of communist ideology, but a strategy for rebels with more transformative ambitions regardless of ideology. I select two cases that differed in almost all ways but their more transformative goals and their non-communist ideology. Furthermore, I selected cases outside of Africa and East Asia to highlight that the imitation of the Chinese model is not the result of contagion from a small slice of conflicts. . . . The quantitative portion offers a final test of the generalizability of my proposed claims. (Stewart 2021, 19–20)

541. Y generalizations are closely connected to selecting on the dependent variable concerns because one is selecting cases of Y equals one. Ever since King, Keohane, and Verba (1994) these have been controversial and the standard line is still one should never do that and one cannot learn anything from doing it. For example, Bueno de Mesquita and Fowler (2021) have a whole chapter on this topic in their textbook. Discuss the quote below by Pearl and McKenzie which is using the canonical smoking and lung cancer example. Notice how Pearl and McKenzie state that the probabilities are backwards, probability of X given Y. They claim that one knows nothing based on these data that would be relevant to a decision about smoking. Explain why that is false.

> Of course Hill knew that an RCT was impossible in this case, but he had learned the advantages of comparing a treatment group to a control group. So he proposed to compare patients who had already been diagnosed with cancer to a control group of healthy volunteers. Each group's members were interviewed on their past behaviors and medical histories. To avoid bias, the interviewers were not told who had cancer and who was a control. . . . The probability logic is backward too. The data tell us the probability that a cancer patient is a smoker instead of the probability that a smoker will get cancer. It is the latter probability that really matters to a person who wants to know whether he should smoke or not. The results of the study were shocking: out of 649 lung cancer patients interviewed, all but two had been smokers. (Pearl and Mackenzie 2017, 192)

Answer: The regularity that the research discovered is extremely strong, .97. So one might conclude that a necessary condition for lung cancer is smoking. Notice that this does not tell you a lot about what Pearl and McKenzie want to know which is what is the probability of lung cancer if one smokes. But it

does tell you that if you do not want to get lung cancer you are almost certain not to get it if you do not smoke. Smoking is part of virtually all paths to lung cancer. Certainly a useful bit of information.

542. Goertz (2017) proposes the Overdetermination Rule for case selection. This suggests that that one should choose cases that have value zero on alternative explanations, Z. Some scholars who like "dueling theories" prefer to choose cases that have one on alternative explanations Z. Bell suggests another criteria. Could you make this more formal?

> I choose cases based on two primary criteria. First, the three cases I use each provide hard cases for the theory. Picking hard cases allows for more confidence in the broader applicability or "external validity" of the findings – if we find support for the theory despite picking cases that we expect the theory will have difficulty explaining, it increases the likelihood that the theory will have some success in cases we do not examine in detail, or in cases that may emerge in the future. In particular, I look for cases with strong "countervailing conditions" – variables whose presence in a particular case makes it less likely that the outcomes posited by the theory of nuclear opportunism will be observed. (Bell 2021, 33)

Answer: He is basically suggesting that one should choose cases that have opposite values on alternative explanations. So if X is positively related to the outcome and when should you use cases where Z is expected to have a negative effect and with a large value on Z, or the opposite if X is negatively related.

543. Cammett (2022) argues that positive deviant cases are good candidates for case studies. Deviance is defined as the residual between the regression line and the observation. Can you give one more specific way to identify these cases based on her discussion below that might be called "surprising success cases"?

> What are Positive Deviant Cases? Positive deviant cases are outliers that exhibit superior performance than the predictions of a model would hold. In technical terms, this refers to cases selected to maximize $|Y_i - \hat{Y}_i|$, or the difference between the actual value and the fitted value in a regression. In order to qualify as cases of positive deviance, however, cases must be more than ephemeral outliers operating in an environment conducive to good performance. Rather, examples of positive deviant organizations or other types of collective actors must exhibit sustained high performance in a context in which good results are uncommon. (Cammett 2022, 220)

Answer: Choose cases (1) score high on Y, (2) score high on X and (3) score high on Zs negatively related to the outcome. These might be called the "surprising success cases."

544. A key issue in case study research is choosing cases in the context of concerns about generalization. Discuss Alter's rationale for choosing cases given her concern over generalizing beyond northern international courts. Would LNQA be applicable here, if so how?

> Although there are many European cases one could examine, only three of the eighteen case studies in this book focus on European legal institutions because I want to show that the new terrain of international law exists beyond Europe. Because I prefer less likely cases, I focus on human rights courts from Latin America and Africa, and cases where ICs with economic subject matter jurisdiction end up speaking to human rights issues, rather than a case study involving the European Court of Human Rights. (Alter 2014, 25)

545. Slater and Simmons provide a nice in class example of case selection based on (1,1) cases that they argue are quite diverse. Both cases have the main causal variable of political uncertainty and the same outcome variable of promiscuous power-sharing. However, they then discuss some additional Z factors that these cases have in common (see quote below). Here they argue that they are similar in both cases. So this seems like a more similar systems design. Discussed the tension between the most different systems versus most similar systems in their case selection.

> In the most general terms, Indonesia and Bolivia serve as a "most different systems" pairing (Przeworski & Teune, 1970). Given the countless ways that these cases vary, it is intriguing to uncover parallel patterns of promiscuous powersharing during times of severe uncertainty. These pronounced cross-case differences are methodologically useful because they allow us to control for a variety of alternative explanations specific to each case (e.g., purportedly collusive political culture). In the next section we argue that both Indonesia and Bolivia saw promiscuous powersharing originate in highly uncertain democratic transitions. . . . The first important parallel [Z_1] is a revolutionary past. This is relevant for several reasons. First, histories of contentious mass politics almost certainly made it harder for party leaders to sustain exclusionary elitist pacts, while also probably helping to convince them that such elitism was essential for political stability in the first place. The strength of popular leftist revolutions also nationalized political life, making subnational ethnicity and regionalism less salient than one would expect by looking at these countries pluralistic and geographically fragmented social structures. Most important for our purposes, revolutions had profound consequences for which political cleavages would be most salient at the national level. This determined what would count as promiscuous as opposed to merely inclusionary in the powersharing arrangements attending democratization. . . . Powersharing is not shaped only by political cleavages, however. It is also a function of electoral rules. Like cleavages, these rules tend to be historically shaped, and to become among the most important parameters influencing democratic interactions. Here we find an additional parallel between the Indonesian and Bolivian cases. When each country democratized, it did so against the backdrop of electoral rules imposed by conservative military regimes [Z_2]. (Slater and Simmons 2012, 1372, 1373)

546. Baum and Potter (2015) in their analysis of democracy and war in chapter 7 do a series of case study analyses from the coalition built before the 2003 Iraq war. This involves decisions to join or not to join the coalition. Their main theory is democratic "constraints" on war initiation. "we have attempted to select cases that vary meaningfully on the dimensions that we argue are key to democratic constraint: partisan opposition and a robust and accessible media.

The conditional nature of this argument means that we anticipate that when either (or both) of these conditions is lacking, it will be possible for leaders to discount public opinion and pursue their independent policy preferences, if they have them. Only when both are present will public opinion truly constrain leaders. Table 7.1 locates our cases, in relative terms, along these two key structural dimensions: opposition and media access." (Baum and Potter 2015, 195). Discuss the (0,0) cell where there are no constraints on the decision to join the coalition, how would one do case studies for this particular cell? Give the Boolean equation for their theory. How do the cases selected line up with the Boolean equation?

547. There exists a lot of debate about rules for case selection in multiple case study research. Discuss these principles of case selection:

> For the purposes of differentiating different theories and visualizing the operation of mechanisms singled out in our theory, we have chosen our cases according to seven principles. Three of the seven principles (2, 3, 7) are drawn from Goertz and Mahoney, Gerring, and Weller and Barnes. purposes here (1, 4, 5, 6).
>
> The seven principles are as follows: (1) Cases must have been well documented, so that their basics are not subject to disputes allowing potential accusations of cherry-picking within-case facts. The exception here is the true negative case of Gabon as a "dog that did not bark." This case has not received much scholarly attention. (2) Cases must include both positive and negative cases with real possibility of positive outcomes (for example, conflict) according to one's own or others' competing theories. (3) Cases must include pathway cases (or typical cases) that exemplify the variations of the key explanatory variable(s), thus allowing clear differentiation of competing theories regarding them. (4) Cases must include pathway cases that exemplify the variations within the operation of the core mechanisms, thus allowing clear differentiation of competing theories regarding core mechanisms. (5) To compare the strengths and weaknesses of quantitative and qualitative methods, cases should include those that can be easily miscoded or misidentified in quantitative exercises, even with GIS datasets. (6) Cases should come from different geographical and cultural backgrounds. (7) Together, the cases examined should provide a fairly generable picture about factors and mechanisms across the population or universe of cases. (Tang et al. 2017, 369–70)

548. Explain how Howard is using the Avoid Overdetermination Rule in selecting her cases:

> Choosing the population based on American-initiated solutions is the best way to test for internally driven hypotheses of US foreign policy, since such decisions are a result of dynamics within the United States and not of multilateral deliberation. The vast majority of US foreign policy literature assumes or argues that internal US dynamics drive decisions. In other mediation efforts in ethnic conflicts (e.g. Kosovo, Macedonia, Afghanistan, and Kenya), the United States worked very closely with international partners. Thus, multilateral, coalition behavior, not internal, domestic decisions, often drove foreign policy. It is necessary to choose cases where the leading hypotheses "most likely" determine the outcomes.

> In other words, the domestic audience-focused, unilaterally driven population would have the most likely cases of decision points where the leading hypotheses would hold. (Howard 2017, 725)

549. Discuss case selection in Stasavage's influential book *Public debt and the birth of the democratic state France and Great Britain, 16881789*. Clearly he is using the very popular paired comparison, in terms the causal mechanism book a (0,0) case compared to a (1,1) case. There are lots of nondemocracies in this. Why choose France? What about choosing more (1,1) cases to generalize?

> Representative political institutions may improve a government's ability to make credible commitments through several different mechanisms. This chapter has used a formal model of legislative bargaining to provide support for my three main arguments. I first demonstrated that if capital owners are in the minority, then party formation can lead to credible commitment, but only if players bargain over multiple issues. In addition, one can expect the perceived credibility of taxation or borrowing to vary according to the partisan composition of government. Both of these observable implications will be considered in detail in subsequent chapters. I next showed that constitutional checks and balances will have little effect on credibility unless there is some mechanism ensuring that capital owners control a veto point. This helps support the argument that multiple veto points may in many cases be insufficient to ensure credible commitment. Finally, I developed my argument about bureaucratic delegation, suggesting that it will improve credibility only if capital owners have the political authority to block any attempt to override bureaucratic decisions. This too is an empirical prediction that is considered in subsequent chapters. (Stasavage 2003, 49–50)

550. Selecting on the dependent variable is very common in trying to figure out the causes of something. Discuss how the World Bank is selecting on the dependent variable to figure out the causes of violent conflict.

> This chapter explores the accumulation and intensification of risks and opportunities in critical spaces, called arenas of contestation. These arenas involve what groups care about in their relationships with each other and with the state and thus what they tend to fight over – access to power, land, and resources, equitable delivery of services, and responsive justice and security. These four broad arenas are by no means an exhaustive list, but they have been selected because they have consistently recurred in violent conflict in various contexts. ((World Bank 2018, 141, *Pathways for peace*, chapter 5)

551. Gerring and Cojocaru 2017 describe what they call "algorithmic" case selection, often using a regression model. Discuss if the italicized elements below are *necessary* and *required* or not.

> Algorithmic case selection follows a set of rules executed in a sequence of steps, which we envision as follows. 1. Define the research question and the population of theoretical interest.
>
> 2. Identify a sample of potential cases. Ideally, this sampling frame should be representative of the population of interest.

3. Measure relevant features of the cases – for example, D, X, Y, and/or Z – across the sample.

4. Combine diverse indicators of D, X, Y, and/or Z into indices, *if necessary*.

5. Construct a causal model, *if required*. (Gerring and Cojocaru 2016, 411; emphasis is mine)

552. When doing multiple or comparative case studies when there are multiple hypotheses involves complicated and potentially tricky case selection decisions. Rocabert et al. 2019 have three hypotheses regarding international parliamentary institutions and IGOs and ends up doing two case studies. Draw the table of all the possible values, presence–absence for all variables in the three hypotheses and then locate the case studies within this list of all possible cases one might choose based on this table. They chose (1) ECOWAS and (2) PIF. Are those the best choices?

553. Discuss the extent to which Alter is using the Avoid Overdetermination Guideline in her selection of cases: "I generally pick hard cases, situations where important interests are at stake, where the policies leading to law violations are politically entrenched, and where law operates in places and ways that are counter to the expectations of international relations scholars and domestically focused lawyers." (Alter 2014, 24)

554. Shannon 2009 is a case study of the US invasion of Panama in 1989. If one were to do a large-N qualitative analysis starting from this article what would be the basic case selection criteria using figure 2. Should you consider case selection further down the tree? Why is that important?

Answer:

The top of the decision tree is those leaders who are motivated to violate the norm is the minimum case selection rule. These are the cases where we can see the norm potentially in action.

555. Selecting on the dependent variable is legitimate for necessary condition hypotheses. Explain why.

Answer:

For a clear discussion see Dion (2003) or Harvey (2003). The short answer is that a necessary condition hypothesis says that we should see X present when Y occurs. Hence the test starts by looking at when Y occurs, i.e., selecting on the dependent variable. See Harvey for a discussion of how this has been ignored in the literature on selection and deterrence (e.g., Achen and Snidal 1987). Dion (2003) uses Skocpol as one of his main examples.

556. Causal mechanisms and scope decisions can be closely related. Bush's main independent variables in her analysis of gender quotas are (1) foreign aid, (2) international election monitoring, and (3) democracy promoting UN peace operation. She could include all countries in her analysis, but she excludes rich democracies.

> A Quantitative Analysis of the Sources of Gender Quotas. The Sample. The unit of analysis for the quantitative analysis of the determinants of quota adoption is the country–year. The sample covers the years from 1970 to 2006 and contains all countries except long-term consolidated and developed democracies. Long-term consolidated and developed democracies were removed because they are subject to different causal processes; they are neither under UN authority nor desirous of signaling their liberalism to the international community, and in fact, they promote democracy abroad. I followed Finkel et al. and removed thirty advanced industrial, long-term consolidated democracies, which resulted in a sample of 165 countries for at least some amount of time. (Bush 2011, 118)

She does include authoritarian regimes, because some do adopt gender quotas. Discuss whether the same causal mechanism works for authoritarian as well as democratizing regimes.

557. Collier and Collier (1991) is a classic of comparative historical analysis. Central to their work is the concept of labor incorporation. Discuss their case selection on this variable and whether it should be a partially a scope condition.

> Because their interest is in the effects of labor incorporation, they [Collier and Collier 1991] consider only the effects of variation in the productive condition: all eight of their cases saw labor incorporated into politics after a major social crisis in the early 20th century. Their interest is in the ways in which different "values" of the productive condition shape divergent regime outcomes over subsequent decades. . . . Since the permissive condition of the rise of labor bounds the theory of regime dynamics that they develop, it acts as a scope condition on the proposed relationship between the mode of labor incorporation and regime outcomes. (Soifer 2012, 1581–82)

558. Discuss Huber and Stephens's (2012) case selection strategy, particularly its longitudinal variation on X and Y. Did they choose cases that were (1,1) cases at some point in their history? Note as well that their core theory involves an interaction term.

> Of the various purposive strategies, we follow Ragin's (1987, 2000, 2008) advice to focus on the positive outcome cases, so we select all cases that had high values on the dependent variable as of 1980. This fits Goertz's (2008) admonition to choose diversely among positive outcome cases and Lieberman's (2005, 444) to select for wide variation in the explanatory variables among the cases that fit the theory according to the quantitative analysis. (pp. 49–50)

559. Ostermann (2007) looks at compliance with policy regarding wood harvesting in protected areas in a park which lies on the border between India and Nepal. She is interested in the effect of state policies on compliance. Discuss her case selection discussion in light of the Avoid Overdetermination Guideline. Discuss in particular the key confounding variables of "weak state presence" and "poverty" which can explain compliance or noncompliance.

> The Terai region in the South of Nepal and the North-Indian state of Bihar is an excellent place to test compliance theories in adverse circumstance for two main reasons: (1) it fits squarely within a least favorable case

> research design – the states on both sides of the border are weakly institutionalized, at least in this region; and (2) the populations on either side of the border are generally poor and uneducated. Indeed, for much of the period leading up to this study, the states on both sides of the border had largely retreated from this region in response to armed Maoist movements. . . . Case selection was also part of an effort to use design to gain methodological traction on persistent questions in the socio-legal compliance literature. Chitwan and Valmiki are contiguous and divided only by an open international border that was not defined by geography or culture. The same culture and dominant ethnic group – the Tharus – prevail on both sides of the border, and most people in this region are poor, fairly uneducated, and speak the same language. (Ostermann 2006, 107)

560. The medium-N paradigm means exploring cases in border areas of the (1,1) zone as secondary case studies. For the democratic peace this would be borderline democracies or borderline peace (in the joint democracy is sufficient for peace version of the democratic peace). An interesting exercise to see what the cases scholars have used when critiquing of the democratic peace, e.g., Elman (1997) or Rosato (2003).

561. Norris in *Making democratic governance work* is interested in the impact of governance and/or democracy on a variety of outcomes, e.g., economic growth, welfare, and conflict. She also uses case studies to illustrate her findings and approach. In the various chapters on various dependent variables she uses paired comparisons. Discuss how she decides about these pairs focusing on the problem that she always has two independent variables – governance and democracy – in each chapter. For example, does she use the Avoid Overdetermination Guideline?

562. Copeland in an influential study of trade and militarized conflict makes claims about the democratic peace. Explain how his case selection (all major power crises and wars) does not include key cases for evaluating the democratic peace.

> The trade expectations approach also helps us understand exactly what is at stake in the ongoing debate about the democratic peace. As we saw in chapter 2, it may well be the case that the correlation between mutual democracy and peace is really reflective of an economic peace as opposed to a political one. Democracies are unlikely to fight each other, in other words, not because they respect each other's normative values or because their legislatures pull illiberal leaders back from the brink, but rather because democracies generally have open liberal economic foundations and thus are able to signal their commitment to open-door economic policies into the future. (Copeland 2015, 434)

Answer:

By selecing only major-power crises and wars he does not include the non-barking dogs, where democracies are at high levels of peace.

563. Seawright (2016) argues for "extreme on X" case selection in multimethod research. Extreme on X could mean very high or very low or distant from

the mean. Contrast that with "good" causal mechanism case selection as argued in *Multimethod research, causal mechanisms, and case studies: an integrated approach*, i.e., some extreme cases are good cases and others are not good cases.

564. Horowitz table 2.4 gives a list of all possible case studies: "The major military innovation cases evaluated in this study are the following: early twentieth-century battle fleet warfare, mid-twentieth-century carrier warfare, nuclear weapons, and suicide terrorism. As table 2.5 shows, this selection of [*four*, one per chapter] cases maximizes variation on the two key independent variables: financial intensity and organizational capital. It also allows for significant time variation, over a period of a century, and cases that focus on both nation-states and nonstate actors." (Horowitz 2010, 61–62).

    If you were going to do only *two* case studies and followed the Avoid Overdetermination Guideline which cases would you choose?

The Possibility Principle

NOTE: this refers to the Mahoney and Goertz article and a chapter of the first edition of the concept book, Mahoney, J., and G. Goertz. 2004. The Possibility Principle: choosing negative cases in comparative research. *American Political Science Review* 98:653–69.

565. Tomz (2007) in his important analysis of international debt and the various means of dealing with potential defaults between countries has a chapter on gunboat diplomacy. This was basically the use or threat of military force by the part of European powers to deal with potentially defaulting governments in Latin America in the 19 century. The statistical analysis nicely illustrates various features of the possibility principal. In his research design gathers data on all the major European and US suppliers capital for 19th-century Latin American countries The basic statistical unit of analysis is the investor-debtor dyad. These are the dyads where it might be possible that the creditor country uses military force against the debtor country who is in default or threat of default. One aspect of the possibility principal he ignores is whether the given investor has has the capability to use military force against a defaulting country. The list of creditor countries is Great Britain, France, Germany, Belgium, the Netherlands, Switzerland, and the United States. So the possibility principal asks if it is possible that a country like Switzerland, which is landlocked and has no Navy could even possibly use gunboat diplomacy. The same can be asked of countries like Belgium. Also interesting are cases where early in the century the country does not have the naval capability, but does have it later in the century such as Germany. So the question is whether one should exclude all of these dyads where the investor country does not have the military capability to conduct gunboat diplomacy?

    Answer: Application of the Possibility Principle would involve both (1) the interest in potentially using the gunboat diplomacy as evidenced by potential defaults as well as (2) the capacity to conduct gunboat diplomacy via the

possession of a Navy capable of intervening in South America. Tomz uses the first one but not the second one.

566. Another exercise related to the one above deals with the fact that Tomz stops his data analysis in 1920: "All alleged examples of bondholder wars, including the famous Venezuelan crisis, occurred in the century before World War I. By the 1920s, military enforcement was already seen as 'a thing of the past' and a 'matter of history' "(Tomz 2007, 116). Obviously issues of default for Latin American countries continue and become a major issue for example in the 1980s. So evaluate his rationale for stopping in 1920? He is saying is that there are no examples of the dependent variable from 1920 onward. Is that a good-valid reason to stop the analysis in 1920? Note that this is effectively selecting on the dependent variable (i.e., when it is consistently zero).

567. Strong regularities such as those found in LNQA analysis play an important role in applications of the Possibility Principle. This is because these regularities are used to infer what cannot happen. This "cannot happen" is based on the empirical regularity. Discuss how this is used in the following example:

> To study dispute onset, it is necessary to first identify the set of pairs of states that have the possibility of entering into a maritime boundary dispute. . . . In practice, at least as far as delimited maritime boundaries are concerned, states do not go beyond the boundary lines they have already agreed to and initiate new conflicting claims. Indeed, I could record no instance of a new dispute onset after full delimitation. For these reasons, the analyses exclude dyad-year observations after full delimitation. (Yüksel 2025, 1181)

568. Should one assume that all countries are at risk of a coup or apply the Possibility Principle in the study of coups? See for example"The data are in time-series cross-sectional format and the unit of analysis is the country-year. The data include 200 countries for the period 1968–2003. (Sudduth 2016, 7)

569. A very difficult conceptual and methodological problem is dealing with norms, particularly norms that have the logical structure of rights, e.g., human rights. This problem has exercised me a great deal and in fact resulted in a book (Goertz *International norms and decision making*; a large part of that book was devoted precisely to analyzing the concept of a norm). (1) Discuss the concept of a right (contrasting it with prohibitions and obligations). (2) The positive and negative poles are interesting here in terms of behavior: do you focus on the positive exercise of rights (e.g., voting) or on the absence of the exercise of the right. This has theoretical consequences because in the former case you focus on what factors in addition to rights lead to the use of rights, while in the latter you focus on factors that eliminate the right itself in practice. (3) Rights are about *possibility*, relate this to the Possibility Principle of case selection and necessary conditions in general. Dowding and Hees (2003) is a very nice discussion of many of these issues. For measurement see also Rosenbaum (2000) and Carter (2000).

570. Morrow and Jo illustrate how the Possibility Principle can lead to specific coding decisions: "Violations are not possible for some issues in some cases

because the sides either lacked the capability to carry out such violations or no fighting of the type in question occurred during the war. All the fighting in the Hungarian-Allies War of 1919 was on land, making violations of conduct on the high seas impossible. . . . These cases receive missing value codes of –9 for all five measures of compliance" (Morrow and Jo 2006, 99). Discuss whether these cases should be coded as "missing" or something else.

Answer:

It would seem to confuse cases where the data is really missing with cases where the coding scheme does not apply.

571. Explain how Chandra and Wilkinson's (2008) apply Posner's (2004) principle of *politically relevant* ethnic groups in their two measures ECI and EVOTE.

Answer:

Both these measures involve how ethnic group distinctions are politically mobilized in some manner other another, i.e., they have been made politically relevant by some actor(s): "At the broadest level, we can imagine the term ethnicity as encompassing two families of concepts – the structure of ethnic identities and the practice of ethnic identification. Ethnic structure refers to the distribution of descent-based attributes – and,therefore,the sets of nominal identities – that all individuals in a population possess,whether they identify with them or not. Ethnic practice refers to the act of using one or more identities embedded in this structure to guide behavior. In other words, it refers to the set of activated identities that individuals employ in any given context. The set of activated ethnic identities for any given country is typically a subset of the identities contained in the ethnic structure" Chandra and Wilkinson (2008, 523).

572. McAdam and Boudet explore why environmental social movements arise. Discuss how they use the Possibility Principle to select communities "at risk" of a an environmental social movement.

> That is, instead of selecting successful movements for study, we chose to research communities communities "at risk" for mobilization by virtue of their shared exposure to the "threat" of environmentally risky energy projects. (McAdam and Boudet 2012, 179)
>
> Because all large, potentially controversial, infrastructure projects are required to file an EIS, these records provide a population of communities at risk for mobilization. Thus our cases were drawn from the CSA Illumina's Digests of Environmental Impact Statements, which contains all EISs from the Federal Register. The population of cases was limited to proposals for new energy infrastructure projects in which a Final EIS was completed between 2004 and 2007 (N = 49). We chose projects that had completed a Final EIS in to ensure that the window for potential collective action would be closed once the study commenced. After drawing a random sample of twenty cases, three of them (15% of the sampling frame) were dropped because it was not feasible to conduct subsequent data collection due to the difficulty in attaining local newspaper records. The sample was then supplemented with three California projects for which data had already been The final sample consists of twenty communities responding to eighteen projects in twelve states regarding proposals for LNG terminals (13

cases), nuclear-related projects (2), a hydroelectric project, a wind farm, and a cogeneration project to supply electricity to an existing oil refinery. (McAdam and Boudet 2012, 36).

573. Howell and Pevehouse in their analysis of Congressional impact on decisions to use military force rely an a dataset of "opportunities to use force" (2007, see Appendix B). For example, "we identify a wide variety of opportunities to exercise force abroad. The database inventories the following kinds of crises: attacks on United States embassies and consulates; instances when United States ambassadors, consuls, or military personnel are killed; hijackings that include human casualties; stateside attacks perpetrated by foreign groups; civil wars; interstate armed clashes" (p. 248). Discuss the extent to which the Possibility Principle is used to constitute a "possible use of military force by the President of the USA."

574. Fordham (2011) looks at the creation of major powers. Explain how using the Possibility Principle would significantly change the case selection research design.

    Answer:

    He uses all states as potential major powers, clearly there are only a small number of states that are potential major powers.

575. One case selection issue deals with cases some individuals – in this case states – are eligible to sign a treaty (dependent variable) while others are not. Only countries in the region are eligible to sign those treaties. The question is what do in this situation. Alcañiz includes in her study global treaties and regional treaties. Here is Alcañiz's solution (2012, 325): "the dummy variable Regional that takes the value of 1 if a treaty is open to a restricted menu of members or 0 if it is open to all countries. it is important to highlight that the universe of countries for regional treaties is restricted to eligible members. Valid subjects for the Tlatelolco treaty, for example, will include only countries of Latin America to which the treaty was open to signature." Is this the best or correct solution?

    Answer:

    The dummy variable is not correct. For example, all survival models assume that $0 < P(Y = 1) < 1$. Notice that the probability is assumed to be greater than zero, however, we know that the probability that European countries will sign a Latin American treaty is zero.

    The simple solution is just to remove the observations of countries not eligible to sign a treaty.

576. While the Possibility Principle is most often used to choose cases, it can also be used to create measures and variables. Describe how Mansfield and Pevehouse use the Possibility Principle in their study of the impact of IGOs on democratization. What role is the Possibility Principle playing in indicator or variable construction?

> Our argument centers on the demand for IO membership by democratizing countries. We have implicitly assumed that the supply of IOs is constant across countries. However, not all countries are eligible to join every IO. Consequently, we reestimate the three original models after adding a variable measuring the number of possible organizations that state i could enter in year t: This variable is the ratio of the number of IOs that state i participates in to the number of IOs it could participate in, where we assume that this state has the opportunity to join every universal organization (e.g., the United Nations) and every organization in its geographical region (Manseld and Pevehouse 2006, 154). In no case is this new variable statistically significant, nor does including it alter the coefficients of the other variables in the model. (Mansfield and Pevehouse 2008, 279).

Answer:

They assume that every country *can* joint every universal IGO, similarly they assume that every country in a region *can* join a regional organization. In practice they may not really be true. For example, can Israel join the Arab League?

The role of the Possibility Principle is to standardize the IGO variable. So instead of the raw number of organization that a country is a member of, it is the number of memberships divided by the number of possible memberships.

577. One potentially interesting new variant on the Possibility Principle uses negative cases on the *dependent variable* to defining negative cases. If the dependent variable is defined using the necessary condition structure then one can covert the AND to OR as per the Possibility Principle. Then take as negative cases all cases with at least one necessary condition present on the dependent variable. Discuss this idea. (Thanks to Sverre Bodung for suggesting this idea.)

578. Just before he died David Freedman wrote a critique of Possibility Principle (Freedman 2009). Freedman discusses the Rule of Inclusion in terms of a bivariate, 2×2 table. He states that the Rule of Inclusion is bad practice because it leads to including no cases in the $(X_1 = 0, Y = 0)$ cell. The Rule of Exclusion is typically a bivariate case selection rule, but the Rule of Inclusion is a multivariate principle. Show, using the Skocpol data, how when one has multiple independent variables (i.e., 2 or more) that in general there will be cases of $(X_1 = 0, Y = 0)$ and $(X_2 = 0, Y = 0)$. Explain which cases one will in fact not be included.

Answer:

Using the Rule of Inclusion in the multivariate setting will almost always give cases in the $(X_1 = 0, Y = 0)$ and $(X_2 = 0, Y = 0)$ cells. This occurs because cases of, say, $(X_1 = 1, Y = 0)$ will sometimes have value 0 for X_2. In short, following the Rule of Inclusion one typically has cases of $(X_1 = 1, X_2 = 0, Y = 0)$ and $(X_1 = 0, X_2 = 1, Y = 0)$. What one does not have are cases of $(X_1 = 0, X_2 = 0, Y = 0)$. You can see this by looking at the universe of all cases used to test Skocpol's theory which clearly do have cases of $(X_1 = 0, Y = 0)$ and $(X_2 = 0, Y = 0)$.

579. Kroenig has as his main hypothesis "Hypothesis 1: The more powerful a state is relative to a potential nuclear recipient state, the less likely it will be to provide sensitive nuclear assistance to that state" (2009, 116). He uses an "all cases" research design to test this: "Potential nuclear recipients are all nonnuclear weapon states in the international system" (p. 177). Using the Possibility Principle discuss how this research design decision will include many cases which support his hypothesis, but which might be consider irrelevant for such a test.

    Answer:

    Very small and poor countries might be considered as irrelevant for this test. Since the main hypothesis is about relative power, these cases combined with nuclear suppliers (i.e., powerful countries) will be cases in support of his hypothesis, e.g., France-Fiji.

580. What constitutes "potential mediators" in the study of international conflict mediation, e.g., Frazier 2006?

581. Describe how the Possibility Principle is used in a number of studies that test the influence of rules of war, such as Morrow (2007), Valentino et al. (2006), Downes (2006).

    Answer:

    Only cases where the state had the capability to violate the rules of war are used. For example, Downes (see the Appendix of his article) only includes states that had the capability to target civilians of the adversary during the war.

582. Collier and Sambanis's (2005) *Understanding civil war: evidence and analysis* chose negative cases based on "high risk." What is definition of high risk? Could you choose cases in gray zone on the dependent variable (they use a dichotomous dependent variable)?

583. Describe the concept structure used to construct politically relevant dyads as a population in conflict research. Once you understand the structure could you easily expand the definition of politically relevant dyads?

    Answer:

    The family resemblance structure is used because it uses a rule of the general form "If any of the criteria is met then include the observation." In this case the list is (1) major power status or (2) contiguity. You could just extend the list. For example, Quackenbush (2006) adds alliances to the list.

584. There exists a large literature on diversionary war (see Levy 1989 for a survey). Discuss how the Possibility Principle is used or not to (or should be and is not) to decide on the population of cases to analyze. Notice that some use the concept "opportunities" to use force as the unit of analysis (e.g., Meernik 1994, 2000). See Foster (2006) for an explicit look at "opportunities" as a function of major power status. See also the question regarding Andy Bennett's decisions about opportunities to intervene in his analysis of USSR behavior. Also, contrast the case selection in Ostrom and Job (1986) with that of Meernik

(1994) and make the analogy to the politically relevant versus all dyad debate. See Mitchell and Prins (2004) for another way to think about opportunities for diversion.

Answer:

Foster wants to explore the impact of major power status on diversionary use of force. It is hard to find it but a footnote in table 1 indicates he used politically relevant dyads. Hence, he is using the same variable to choose populations as he is looking at in the main statistical analysis.

Ostrom and Job base their analyses on time periods, i.e., quarterly data setup. Underlying this is the view that the US always has an opportunity to use force. They propose and Meernik follows up on the idea that a dataset on opportunities to use force would be a better research design. Notice that the impact of domestic politics (the key finding in Ostrom and Job) disappears in the Meernik study.

585. The problem of relevant case selection arises in the literature on the intervention of third parties in ongoing conflicts. If it is unlikely that minor powers get involved in conflicts in distant regions it is probably even less likely that they intervene in a partisan fashion in ongoing international or civil war. The Possibility Principle eliminates cases from the population under (statistical) analysis. An alternative approach is to model-in a separate equation-those states that are likely to intervene at all and then look at issues of which side. One way to think about this is "mixture cure" models:

> Thus we estimate, instead of standard survival models, "mixture cure" models (e.g., Sposto 2002; see the online appendix at http://journalofpolitics.org/ for technical details and additional references). These models contain an additional binary regression component to model the probability of the event occurring eventually and to correct the hazard/duration model portion. Estimating the mixture cure model with competing risks confers two major advantages: (1) We avoid having to address the above-mentioned problematic assumption of standard survival models by first defining, then analyzing only "relevant" states. This approach would invariably have resulted in exclusion of cases of intervention by "nonrelevant" states. (2) We are able to ascertain simultaneously what factors affect the probability of intervention, the hazard/timing of the intervention (given the decision to intervene), and the side on which the intervention occurs. (Findley and Teo 2006, 833)

Note that this is similar in spirit to the zero-inflated Poisson models (Chin and Quddus 2003) which attempt to separately model the zero occurrences from the nonzero ones. Discuss these statistical alternatives which separately model the "irrelevant" cases.

Answer:

A key thing to keep in mind in the application of the mixture cure model is any given variable which appears in both equations must considered carefully since the "total" impact is the combination of the two. This may be hard to disentangle because in the variable only influences the hazard rate in the selection equation.

Notice that the Findley and Teo argument also implicitly refers to the problem in the conflict literature where using politically relevant dyads as a population selection device (as opposed to a zero-case selection device) excludes the "impossible happens" cases.

It is important to realize that these statistical approaches as well as the Possibility Principle apply a theoretical or empirical model. One needs to consider which model is better. When the impossible happens is clearly error in the Possibility Principle model. However, it may be the case that zero-case model in the mixture cure performs poorly in modeling the hazard rate.

586. King and Zeng (2007) have presented a large-N argument about counterfactuals. Basically the article deals with the perils of extrapolation and a technique for deciding when a counterfactual is an extrapolation outside the data. Basically, counterfactuals within the data are more reliable than extrapolations outside it: a bit of conventional wisdom. One conclusion from this methodology is that one should collect more data to increase the region covered by the data; a conclusion not too surprising from a statistical point of view. Use the Possibility Principle cube to think about these issues.

(1) It is quite possible for counterfactuals within the boundaries of the data to be quite distant from actual data points. Illustrate this in the cube.

(2) Use the cube to show how counterfactuals outside the data might be closer to actual data points than points in the data cube.

(3) Use the democracy and war literature to illustrate how interpolation might be more problematic than extrapolation.

(4) Applying the Possibility Principle principle usually means reducing the size of the sample space, while King and Zeng implicitly argue for increasing it. However, both argue that causal or counterfactuals are best made with contrasting cases that are "close" to each other. Discuss.

Answer:

For (1) and (2) put points at the corners, and a few outside the corners and leave the middle of the cube empty.

The democratic peace literature is paralleled by a small "autocratic" peace literature. But there is a fair amount of evidence that hybrid or transitional regimes are more war prone. Hence interpolation of causal effects may not work in this case. Also, extrapolation to higher quality democracy is probably going to be more valid than interpolation to gray zone democracies.

For (4) the key issue hidden by the King and Zeng article is that one wants to make contrasts with cases that are nearby. Data in the convex hull may be more distant than data outside it. The Possibility Principle says that increasing the convex hull may be counter-productive if the contrasting cases are too far away.

587. The Possibility Principle can play an important role in concept/measure construction when the number of zeros has a large impact on the resulting measure. Explain how the S measure of dyadic interstate preferences (Signorino and Ritter 1999) is sensitive to this when using all dyads versus politically

relevant dyads. Why do all the many politically irrelevant cases have an influence?

Answer:

See Sweeney and Keshk (2005) for an extensive discussion. Basically, all the politically irrelevant dyads have similar preferences and hence one gets a huge number of cases with high preference similarity (e.g., the vast majority of cases have a value near the maximum of 1.00).

588. Sometimes the Possibility Principle is related to truncation on the dependent variable which appears in case selection. For example, the dependent variable could be US foreign direct investment (Oneal 1994) or terrorist actions (Li 2005; Enders and Sandler 2005). Sometimes the cross-sectional cases are only those countries with a positive value on the dependent variable, for example, those that actually got foreign direct investment or had a terrorist act over the period in study. Explain how the Possibility Principle relates to this kind of potential selection bias. In practice which would be worse the US foreign direct investment case or the terrorism case?

Answer:

Implicitly the Possibility Principle is used because countries that, for example, did not receive US direct investment are perhaps assumed to be those where it was in some sense impossible. US legislation forbids investment in some countries, e.g., Cuba. Of course, there might be cases where there is no economic value in such investments but those cases one would want to include. (Note that FDI raw data often have positive and negative values for inflows and outflows. Some authors take the log which then typically means throwing out all the negative and zero cases.)

The issue is probably less severe in the literature on terrorism. As a matter of history (or historical accident if you prefer) there have been terrorist incidents in almost all countries of any size (e.g., about 120 according to Li 2005). Thus the exclusion of countries who have not had a terrorist incident is not likely to have much practical effect. However, if it were the case that only 60 countries had had terrorist incidents then this issue would become important. Some authors include all cases and then try model the zero cases separately with techniques like zero-inflated Poisson (Chin and Quddus 2003).

589. The Possibility Principle can be used in indicator construction as well. A major issue in the political and economic analysis of countries is their degree of ethnic fractionalization. A large literature in economics (e.g., starting with Easterly and Levine 1997) deals with its impact on economic growth. Clearly ethnic variables have played a large role in the analysis of civil wars and genocide. How might the Possibility Principle be used to think about which ethnic groups to include in an index of ethnic fractionalization in ways similar to the way politically relevant dyads are used to study international conflict?

Answer:

Posner (2004) develops a measure of ethnic fractionalization-what he calls Politically Relevant Ethnic Groups (PREG)-based on the logic of the Possibility Principle and politically relevant dyads:

> The crux of the problem is that the *Atlas* data from which the ELF index was calculated enumerates dozens of groups in each country that may be culturally or linguistically distinct from their neighbors but that are irrelevant as independent political actors. . . . Let me be clear: my assertion is not that the many ethnic groups included in the *Atlas* (and, often, in the newer measures) are unimportant *per se*. Rather, my claim is that these groups are unimportant for the explicitly political mechanism that the researchers are trying to test. To capture the contribution that a country's ethnic heterogeneity makes to its policymaking process requires an index of fractionalization that reflects the groups *that are actually doing the competing over policy*, not the ones that an ethnographer happens to identify as representing distinct cultural units. (Posner 2004, 853)

590. The Possibility Principle is closely related to scope conditions. The welfare state literature only looks at wealthy countries. Much of the literature on well-being focuses on poor or destitute countries (Dasgupta 1990). When Dasgupta examines the relative destitution of countries he excludes the well off ones. He uses the Borda Rule (Dasgupta and Weale 1992) to aggregate a variety of indicators of well-being, such as infant mortality, literacy, per capita income civil and political rights. The Borda Rule ranks countries on each dimension then sums the ranks (see Fine and Fine 1974 for an extensive and technical discussion of its formal properties). Explain why this concept of human destitution or well-being is sensitive to the exclusion of wealthy countries. For example, the comparison of India and China will significantly be affected by the inclusion or exclusion of nonpoor countries.

    Answer:

    As Dasgupta and Weale note, the Borda Rule is sensitive to the population chosen because the distance between a pair of countries depends on the population. In particular, if an alternative population puts countries in between a given pair their overall rankings (and hence the comparison between the two) can change. Particularly with regarding to civil and political rights the inclusion of wealthy countries can change the comparison because these are scattered throughout the ranking.

591. Fox and Lawless (2004, 2005) study the phenomenon of "nascent political ambition" in American politics. They need to gather data on those who "might" run for office. Explain how they use the Possibility Principle in designing their large-N survey (see Maetas et al. 2006 for another approach).

    Answer:

    To form their population for sampling they look at the professions of people *who have already run for office or already hold office*. In accordance with the Possibility Principle they use those that have run or hold office to determine the population of those who might run for office: "Despite the theoretical importance of studying nascent political ambition, a number of methodological and sample design issues make conducting an empirical investigation quite difficult. The complexity of assembling a national sample of potential candidates, alone, explains why most research on political ambition and candidate emergence focuses on declared candidates and office holders. In an attempt to overcome these difficulties, we developed the "eligibility pool approach,"

which we carried out in the Citizen Political Ambition Study. We drew a national sample of 6,800 individuals from the professions and backgrounds that tend to yield the highest proportion of political candidacies: law, business, education, and political/community activism." (Fox and Lawless 2005, 647)

Note also that the other approach to sampling in this area the "reputational approach" starts by starting with office-holders (i.e., people who have been candidates AND successful) as a place to get names of "prospective" or "viable" candidates (e.g., Stone and Maisel 2003).

592. Many theories of coalition government formation rely on formal models where a key variable is the bargaining power of the potential coalition partners. Ansolabehere et al. (2005) describe a method for calculating "voting weights" based on a calculation of all possible minimal winning coalitions (p. 555; see also Warwick and Druckman 2006). Explain how you could use the Possibility Principle to improve this method of calculating weights? (See Waters 1998 for a philosophical discussion and example-using the analogy with copper as used in the book-of how one would think about the about the minimum winning coalition problem)

    Answer:

    "All possible minimal winning coalitions" includes many coalitions that are very unlikely to happen (e.g., between extreme left and right parties). It would be better to include only those minimal winning coalitions with any real likelihood of forming. Notice that doing so would require using theory and data about what kinds of coalitions actually do form (i.e., a theory of the positive outcome). For example, Warwick and Druckman mention this common critique: "A common critique leveled against measures such as these [of bargaining power] is that they assume that all coalitions are equally likely to occur." (Warwick and Druckman 2006, 645).

593. Explain the two variables in Singh and Way's quantitative analysis of nuclear proliferation (2004) that could be used with the Possibility Principle. What in their discussion and empirical results would support such a view? Note that the IAEA has category of "nuclear latent states"; does this concept invoke the Possibility Principle?

    Answer:

    They clearly argue that technological/industrial capacity is a necessary condition for nuclear weapons acquisition: "there are numerous examples of states that have had the technical capacity to build nuclear arms for several decades but have never attempted to do so. However, we argue that it is an important starting point because no nation can build nuclear weapons without attaining a minimal economic/technological capacity" (p. 862)

    "Every country to acquire nuclear weapons, with the exception of Pakistan,was above the threshold embodied in the [industrial capability] index." (p. 875 footnote 30)

    They also note empirically that the existence of a rivalry is basically a necessary condition: "This is not surprising when one recalls that of the countries to

acquire nuclear weapons, all but two (South Africa and France) are coded as participating in an enduring rivalry at the time of acquisition. By alternate but defensible coding rules, one could make the case that both of these exceptions were involved in enduring rivalries." (p. 875)

Also, the empirical results show that the industrial variable and the enduring rivalry variable have extremely large effects, with rivalry having from a 380% to 740% impact and industrial capacity 560% to 2,340% effect. No other variables are even close to this.

594. In the large-N study of US foreign aid – or foreign aid in general – one must choose a population of cases. This involves deciding whether to use all countries or some subset that "might" receive foreign aid. What might be a criterion to exclude a set of countries that almost never get foreign aid (e.g., Lai 2003; see also Kosack and Tobin 2006 who contrast development aid with foreign direct investment in a world-wide sample). See also Carey (2007) for another approach to choosing cases based on who already receives aid.

    The same question can be asked in studies of U.S. arms exports. Examine Blanton (2005). Are developed countries excluded? Should they be included or excluded? What effect might that have on the results? Where can you see the logic of the Possibility Principle in her argument?

    Answer:

    OECD countries almost never receive foreign aid. More generally, aid is not given to rich countries. One does not give foreign aid to Sweden and Switzerland.

    In Blanton all developed countries are excluded. So this means including Saudi Arabia but excluding Israel. Arms transfer certainly do occur to developed countries so it is not completely clear (other than the convention of excluding them in these studies) why they should be excluded. The section "A two-stage model" clearly uses the logic of the Possibility Principle. This also illustrates the close connection of the Possibility Principle to Heckman selection models.

595. In the large-N quantitative conflict literature some researchers prefer to use directed dyads instead of the more common undirected dyad approach. The non-directed dyad only includes one case per dyad-year, e.g., US-USSR. The directed dyad includes two, US-USSR and USSR-US. One advantage of the directed dyad is that it allows you to look at dispute/war initiation. For example, Bennett and Stam (2004) argue that "Given that we want to study initiations, the correct population to employ includes cases where there could have been MID initiation." (p. 55). One problem they then have is what to do in with directed dyad BA when A has already initiated a dispute against B. Discuss whether it is "possible" then for B to initiate a dispute with A (in the same year) and thus whether one should keep (or not) the BA directed dyad in the population.

    Answer:

    It turns out that there are about 150 cases where two countries have multiple disputes within the same year (Goertz, Jones and Diehl 2005). Bennett and

Stam argue (p. 55–56) that one should *exclude* the BA directed dyad when a AB dispute exists. However, if there does happen to be a BA MID they do include it in the analysis. Notice that this is the opposite of the way politically relevant dyads work, since the impossible-but-happens cases are excluded.

596. Use Licklider's findings (1995, table 3, p. 687) in conjunction with Harff (2003) to think about case selection in the analysis of civil wars and genocide/politicide.

Answer:

Negotiated settlement is sufficient condition for no genocide. Military victory is necessary condition for genocide.

597. In the geometric representation of the Possibility Principle Mahoney and Goertz used a cut-off of .50. We think that in practice, qualitative researchers will tend to chose much high cutoffs. Discuss whether you think this is true. Can you think of relevant examples? If it is true why do qualitative researchers do this?

598. Some researchers believe that one should look at "revolutionary situations" instead of revolutions:

> [W]e can clear a good deal of conceptual ground by means of a simple distinction between *revolutionary situations* and *revolutionary outcomes*. Most significant disagreement about the proper definition of revolution falls somewhere along these two dimensions. (Tilly 1978, cited in Walton 1984, 11; see figure in Tilly 1978)

Explain how the idea of a revolutionary situation relates to the Possibility Principle.

Answer:

The quotation from Tilly suggests that one might use the concept of "revolutionary situation" to define cases where the outcome of (successful) revolution is possible. Hence, in a study of revolutionary outcomes, appropriate negative cases might be revolutionary situations that did not culminate in revolutionary outcomes.

See Stinchcombe for a related point: "Rather than explaining the occurrence of revolution, a sociological theory ought to try to explain the occurrence of a 'revolutionary situation.' Whether or not a change in the ruling powers of a society takes place by means of violence depends both on the predisposing characteristics of the social structure and on concrete military and political situations at given historical times" (1965, 169).

599. The cube in Figure 2 (Mahoney and Goertz 2004) presents three zones with disconfirming cases. But in qualitative research the zone near the right-back-bottom (1,1,0) corner is often the most important one for falsifying theories. Why?

Answer:

Necessary condition hypotheses are disconfirmed by cases where the outcome takes place (located in the top half of the cube). Sufficient condition

hypotheses are disconfirmed by cases where the outcome does not happen (located in the bottom half of the cube).

In qualitative research, analysts are often quite familiar with many or all positive cases. As a result, they may be less likely to build theories that are obviously disconfirmed by positive cases. By contrast, they may be less knowledgeable about nonpositive cases, such that these cases may be more likely to disconfirm their theories. The right-back-bottom (1,1,0) corner is where all of these disconfirming nonpositive cases will be located.

600. Explain how Andy Bennett in his *Condemned to repetition? The rise, fall, and reprise of Soviet-Russian military interventionism, 1973–1996,* (1999, chapter 1) faced the Raven Paradox problem and how he implicitly used the Possibility Principle to solve it.

Answer:

Bennett's outcome is Soviet/Russian intervention, and he must decide which nonpositive cases are relevant for theory testing. He uses the Rule of Inclusion to inform this choice. Bennett (1999) uses this logic when selecting cases to test a theory of the causes of Soviet and Russian military intervention. He identifies five factors that provide "opportunities" for intervention, such as the presence of a pro-Soviet/Russian faction or a low level of U.S. threat. He then considers as relevant only those countries or territories where one or more of these factors provided the opportunity for intervention; he does not consider the vast majority of countries or territories, because Soviet or Russian intervention was simply not possible in these cases.

601. Explain how the the Possibility Principle is crucial to the large-N study of war. For example, Vasquez in his ISA presidential address: "Of course, many such dyads do not go to war simply because they do not have any conflict with each other or the opportunity to engage in militarized confrontation. Therefore, the analysis will be confined to only dyads that actually engage in the threat or use of force, what has been called involvement in a militarized interstate dispute (MID)" (Vasquez 2004, 2).

Answer:

The militarized dispute dataset (Jones et al. 1996) was originally developed so that scholars would be able to study a set of cases where war was possible. The Singer and Small war data set existed but there was no set of negative cases to compare it against. Hence, the militarized dispute data set helped scholars identify appropriate negative cases.

602. Read Barbara Harff's (2003) article on genocide. Explain how she (informally) uses the Possibility Principle to select her cases. Does she primarily use the Rule of Inclusion or the Rule of Exclusion to eliminate irrelevant cases? See also later work using the same basic ideas, e.g.:

> We follow Harff (2003) and Colaresi & Carey (2008) in restricting the sample to country-years experiencing political upheaval since almost all mass killings occur during or immediately after episodes of political upheaval.

> Like Colaresi & Carey (2008), we only include country-years with ongoing state failure, defined as revolutionary wars, ethnic civil wars, adverse regime transitions, or genocides or politicides, as coded in the Political Instability Task Force (PITF) data (Marshall, Gurr & Harff, 2013). To ensure that our estimation results are not simply an artifact of this decision, we also examine the full sample including all country-years during the sample period for those countries for which data exist (19552010). Results remain similar. ((Hong and Kim 2019, 535)

Answer:

Harff uses the rule of exclusion to remove cases where genocide is impossible. In particular, she excludes all cases that lack a causal factor (political upheaval) that is almost always necessary for genocide.

Harff (2003) notes that almost all genocides (i.e., 36 out of 37) occur during or immediately after political upheavals. Accordingly, she excludes cases like France and Canada that lack political upheaval when testing her theory of genocide. These politically stable cases have such a low probability of experiencing genocide that their inclusion would distort inferences about other cases where the outcome of interest is possible. For example, one may have a theory of genocide that highlights ethnic divisions as a key independent variable. Under the Rule of Inclusion, contemporary Canada could therefore be considered a relevant case. However, under Harff's (2003) exclusion criterion, Canada is irrelevant because its value on the political upheaval variable eliminates it from the analysis.

603. The quantitative conflict literature uses "politically relevant dyads" to apply the Possibility Principle to select cases. Mahoney and Goertz suggest that the total population is the union of negative cases (selected with Possibility Principle) and the positive cases. However, the conflict literature uses a different procedure to determine the total population of cases. What is that procedure? How does that influence what the researcher does when the "impossible happens?"

Answer:

The quantitative literature on politically relevant dyads adopts the following procedure: (1) use the Possibility Principle to distinguish relevant and irrelevant cases, (2) treat the relevant cases as the whole population of interest, and (3) select the positive cases as a subset of this population. As a result of this procedure, a mistake in step one can cause positive cases to be coded as irrelevant, such that the impossible happens.

604. Does the state failure (http://www.cidcm.umd.edu/inscr/stfail/) project use the Possibility Principle to select cases? If so, how? If not, should it use that principle?

Answer:

The project does not exclude cases where state failure is impossible, and thus the project does not employ the Possibility Principle. The project might reach different results if the Possibility Principle were applied. For example, the application of this principle might lead to the exclusion of many or most of

the advanced industrial democracies. In turn, with the new population of relevant cases, different results could emerge.

605. What would be the set of appropriate negative cases for the following Boolean theory?

$$Y = A + B + C + D \tag{2}$$

Does this contradict the Possibility Principle? Why does this example not fit the general kind of problem that typifies case selection in small-N studies? Hint: Which combinations of A, B, C, and D predict that Y does *not* occur? (Thanks to Charles Ragin for this exercise.)

Answer:

This is a question dealing with the logic of the Possibility Principle. The problem here is that if we followed the Possibility Principle for selecting negative cases, we would not have any negative cases at all since the Possibility Principle *excludes* all cases where A–D are absent (i.e., not-A AND not-B AND not-C AND not-D, this is a hint in the question). But this is exactly the set of cases we need to test the Boolean theory $Y = A + B + C + D$. This situation arises because A–D are individually sufficient for Y. However, models like $Y = A + B + C + D$ are extremely rare in the qualitative-comparative literature.

606. Skocpol considers the Cuban Revolution (1959) to be irrelevant for directly testing her theory. Do you think she excludes this case using the Possibility Principle or scope conditions? Why?

Answer:

Skocpol must be using scope conditions – not the Possibility Principle – to exclude the Cuban Revolution, given that the Possibility Principle excludes only cases where the outcome of interest is absent.

607. Explain how the cube (Mahoney and Goertz 2004) provides a visual summary of disconfirming observations for necessary and/or sufficient condition hypothesis.

Answer:

With a necessary condition hypothesis, theory disconfirmation occurs when the necessary condition is absent, but the outcome of interest is present. Thus, disconfirming cases are located in the top portion of the cube where only one necessary condition is present (i.e., the (1,0,1) and (0,1,1) corners).

With a sufficient condition hypothesis, theory disconfirmation occurs when the sufficient condition is present, but the outcome of interest is absent. Thus, disconfirming cases are located in the bottom portion of the cube where the two jointly sufficient causes are present (i.e., the (1,1,0) corner).

608. Seawright (2002; see Clarke 2002 and Braumoeller and Goertz 2002 for responses) argues that one should use "all cases" for testing hypotheses concerning necessary or sufficient conditions. Explain how the Possibility Principle and the Raven paradox enter into this debate.

Answer:

Much of the debate centers around the issue of which nonpositive cases are relevant for theory testing. Seawright suggests that all cases in an "appropriately defined universe" are relevant to testing a proposition about causal sufficiency, even negative cases that lack the hypothesized sufficient cause. He shows that the inclusion of all cases can substantially enhance statistical significance by increasing the number of confirming observations. By contrast, Clarke argues that including all cases will lead one to confirm a proposition through irrelevant observations, in much the same way that "most ravens are white" might be confirmed by observing yellow pencils and blue books. Braumoeller and Goertz's argument likewise suggests that, when testing a hypothesis about a sufficient cause, cases that lack both the cause and the outcome are irrelevant, since the hypothesis does not imply anything about the number or proportion of these cases that should be present.

609. Find the parameter that can change dramatically depending on whether one applies the Possibility Principle or not in King and Zeng's (King and Zeng 2001a) discussion of statistical estimation in rare events data. For a similar problem see also King and Zeng (2001b).

Answer:

"Prior correction requires knowledge of the fraction of ones in the population, τ. Fortunately, τ is straightforward to determine in international conflict data since the number of conflicts is the subject of the study and the denominator, the population of countries or dyads, is easy to count even if not entirely in the analysis." (King and Zeng 2001a, 144).

In King and Zeng (2001b) one should note similarly that relative risk involves calculations using baseline probabilities ($Pr(Y = 0)$ or $Pr(Y = 1)$) so that $Pr(Y = 1)$ will be much lower with all dyad design.

610. The welfare state literature typically excludes less-developed countries. Is this because of scope conditions or the Possibility Principle? If it is the Possibility Principle, is it the Rule of Inclusion or Rule of Exclusion?

Answer:

To avoid these problems, many analysts of welfare states include only OECD countries (see Pierson 2000; Amenta 2003 for recent reviews). Typically, they justify the exclusion of poorer countries through the use of scope conditions. However, they exclude *only* negative cases, and so they are really employing the Possibility Principle, not scope conditions. In particular, they use the Rule of Exclusion to eliminate countries that possess a condition sufficient for the absence of welfare state development-namely, poverty. Indeed, the finding that economic wealth is related to welfare state development among all countries but not among rich countries is what we would expect if all cases are homogeneous (i.e., if scope conditions do not apply).

611. Fuzzy-set analysis is often problematic when there are many observations near the origin (i.e., observations whose values on all variables approach zero). When observations are near the origin it is extremely easy for them to pass tests of sufficiency since X is likely to be less than Y. How does the Possibility Principle aid fuzzy-set methodology by dealing with these cases?

Answer:

This geometric interpretation of the Possibility Principle offers a solution to an important problem that arises in fuzzy-set analysis. The problem involves what to do with cases that are near the origin (i.e., cases near the (0,0,0) corner). As Ragin notes (2000, 250–51), when testing whether variables are causally sufficient for an outcome, observations with a zero for all the independent variables will always satisfy causal sufficiency and thus artificially inflate the number of cases where the theory works (this dilemma is the Raven Paradox). See also the exchange between Osherson and Smith (1981) and Zadeh (1982). The former is a classic in the cognitive psychology literature on categorization and Zadeh is the inventor of fuzzy-set logic.

612. Do Foran (1997) or Goodwin (2001) apply the Possibility Principle in their studies of (social) revolution?

Answer:

Foran's (1997) Boolean analysis of social revolution selects as negative cases only state-periods that have a positive value on at least one of his five major independent variables. Goodwin (2001) likewise selects as negative cases only state-periods where at least one key independent variable is present.

Two-way tables

613. Two-way tables are quite ubiquitous in theorizing both for qualitative and quantitative methods. These raise a variety of questions regarding the theory – what is the relationship between the 2–3 independent variables – as well as case selection. Discuss the two-way table below. How many dimensions are there in fact? In fact it is really a multidimensional table; what are the independent variables? How do they interact to produce the dependent variable of welfare provision? What case studies would be the most useful? How would one one do M-LNQA using their data–theory, given the various interaction terms in their theory?

Table 22: Predicted welfare effect of subversion in resource-poor and resource-rich countries

	Resource-poor countries		Resource-rich countries	
	Separatist subversion	Center-seeking subversion	Separatist subversion	**Center-seeking subversion**
Guns vs. butter constraint	Strong	Strong	Weak	**Weak**
Motivation for mass co-optation	Weak	Strong	Weak	**Strong**
Overall effect on welfare provision	Negative	Indeterminate	Indeterminate	**Positive**

Note: Core hypothesis in bold.

Source: Eibl and Hertog 2024.

614. Whenever everyone is confronted with a two-way theory table one should always ask whether the row variable is more or less important than the column variable, or equally important. Analyze the table below from Lessing's influential book on state-cartel violence. Determine if row or column is more important in his theory. Hint: Consider the outcomes in the cells of the table to be ordinal, from 1 to 4. Then write a little equation where row and column can take on values of one or two and where the equation combining row and column generates the outcomes in the cells.

Figure 60: Two-way table analysis: Lessing on state-cartel violence

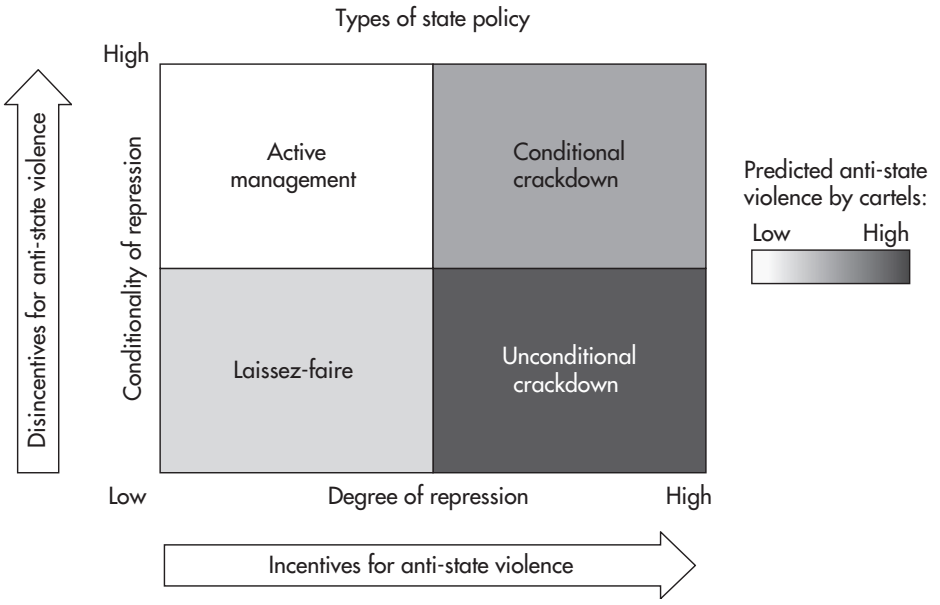


FIGURE 1.1. **Degree vs. Conditionality of Repression: Policy Types and Incentives for Violence.** *Degree of repression* measures the overall amount of force directed at cartels; *conditionality of repression* measures how much additional repression a cartel incurs by engaging in anti-state violence.

Answer: Conditionality is more important than degree of repression. The basic equation is 1*Degree + 2*Conditionality = cartel violence

615. A constant issue with typologies is the nature of the dependent variable in the cells. Mazzuca's typology, see below, is his core framework: "The typology is the first step in the creation of a more general theory of state formation. It transforms scope conditions implicit in the theory of state formation advanced by Weber and his contemporary followers into explicit causal variables. The typology also helps to enrich neo-Weberian approaches to state formation with valuable insights from neo-Marxist alternatives" (Mazzuca 2021, 42). What is the dependent variable? Is it "state formation" i.e., all are cases of $Y = 1$ but different paths for each cell, or are there four separate Ys? What about $Y = 0$?

Figure 61: Typology of state formation

Table 1.1. Geopolitical and International Economic Environments of State Formation

		International economy	
		Feudalism/ mercantilism	*Capitalism/free trade*
Geopolitical context	*Anarchy*	Pioneer state formation	Competing state formation
		War-led (with military innovation)	Trade + war-led
		Western Europe	*United States*
	Hierarchy	Reactive state formation	Latecomer state formation
		War-led (with military imitation)	Trade-led
		Eastern Europe, China, Japan	*Latin America, Africa*

Source: Mazzuca (2021).

616. A constant issue with typologies is the nature of the dependent variable in the cells. Mazzuca's typology, see below, is his core framework: "The typology is the first step in the creation of a more general theory of state formation. It transforms scope conditions implicit in the theory of state formation advanced by Weber and his contemporary followers into explicit causal variables. The typology also helps to enrich neo-Weberian approaches to state formation with valuable insights from neo-Marxist alternatives." (Mazzuca 2021, 42). What is the dependent variable? Is it "state formation" i.e., all are cases of $Y = 1$ but different paths for each cell, or are there four separate Ys? What about $Y = 0$?

Figure 62: Typology of state formation

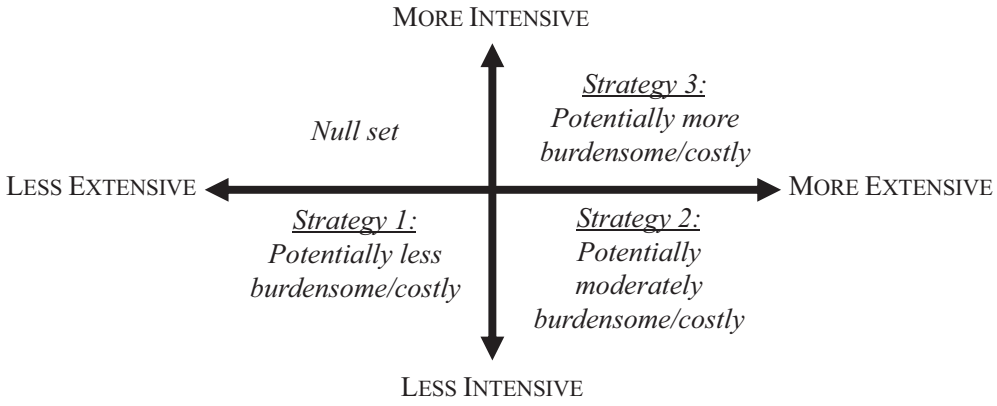
Table 1.1. Geopolitical and International Economic Environments of State Formation

		International economy	
		Feudalism/ mercantilism	*Capitalism/free trade*
Geopolitical context	*Anarchy*	Pioneer state formation	Competing state formation
		War-led (with military innovation)	Trade + war-led
		Western Europe	*United States*
	Hierarchy	Reactive state formation	Latecomer state formation
		War-led (with military imitation)	Trade-led
		Eastern Europe, China, Japan	*Latin America, Africa*

Source: Mazzuca (2021).

617. One of the things that one should always ask when confronted with a 2×2 table is whether the row factor is more important than the column factor. Stewart's analysis of rebel governance is presented in the figure below. Which is the most important factor, row or column?

Figure 63: Intensiveness and extensiveness of rebel governance: which is more important?



Source: Stewart 2021.

Answer: it is the extensiveness factor, the two highest levels in the two-way table occur when this factor is high which makes it more important than the column, intensiveness, factor.

618. In some two-way tables the various cells are ordered by how they influence a particular dependent variable. Weeks (2014) does exactly this in the table below (discussed as some length in the concept book). The various combinations of the row and column factors are predicted to have an ordered impact on the likelihood of a militarized dispute. If this were done as a cube, which would be my recommendation, what would it look like with the third, vertical, axis as the likelihood a militarized dispute?

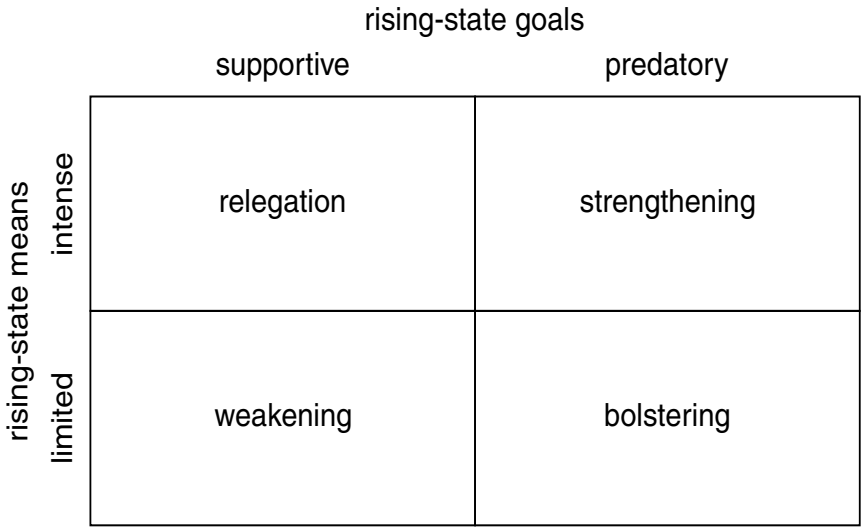
Table 23: Two-way tables: varieties of authoritarian regimes and militarized dispute initiation

	Civilian audience or leader	Military audience or leader
	"Machine" 626 country-years	"Junta" 294 country-years
Elite-constrained leaders	China (after Mao) Kenya Malaysia Mexico (until 1997) Poland Senegal Tanzania USSR (after Stalin) (North) Vietnam	Algeria Argentina Brazil Greece Myanmar (after 1988) Nigeria Rwanda South Korea Thailand
	"Boss" 547 country-years	"Strongman" 584 country-years
Personalist leaders	China (Mao) Cuba (Castro) Indonesia (Sukarno) Iraq (Saddam) Libya (Gaddafi) North Korea (Kims) Portugal (Salazar) Romania (Ceauşescu) USSR (Stalin)	Chile (Pinochet) Egypt (Nasser) Indonesia (Suharto) Iraq (Qasim, al-Bakr) Myanmar (until 1988) Pakistan (Ayub Khan) Somalia (Siad Barre) Spain (Franco) Uganda (Idi Amin)

Source: Weeks 2014.

619. The table below gives a theory table involving two independent variables. One of the first questions one needs to ask is whether there is some underlying variable or the dependent variables in the four cells. In this case, there is a later figure (figure 1) where the four outcomes are stacked from top to bottom as follows: (1) strengthening, (2) bolstering, (3) relegation, and (4) weakening. Does that mean there's an ordinal relationship between these four dependent variables? If we code weak military as 1 (robust as 0) and high strategic as 1 (low as 0), as implied in the later figure, does that suggest a partial ordering?

Table 24: Two-way tables: rising state goals and rising state means



Source: Shifrinson 2020.

620. Often two-way tables are best interpreted via Boolean logic. The table below provides a core summary of the Carnegie and Carson game theoretic model predictions with the two core independent variables on the row and the column of the table. Give the Boolean equation that generates the table. Analyze and the connection between this table and the formal game theoretic model described in the article. Discuss the cases which are most relevant for analyzing various parts of the Boolean equation. Discuss the hypotheses the relationship between the Boolean equations that can generate table **??** and the implicit linear, additive hypotheses in their hypotheses H1 and H2.

> We argue that international organizations constitute a potential solution to disclosure dilemmas. . . . We claim that an IO can serve this purpose under two primary conditions. First, an IO must have the capacity to credibly review, assess, and act on sensitive information, which requires a reputation for technical expertise and relatively unbiased judgment; otherwise, it cannot provide added legitimacy to address the intelligence holder's credibility problem. Second, it must be designed to receive and protect sensitive information by limiting its dissemination within the IO and preventing unauthorized leaks. . . . These [game theoretic] claims

> lead to two testable hypotheses, which are summarized in Table 1 [the table below]. H1: The greater the intelligence reception and protection capabilities of an IO, the more intelligence an informed state discloses to it about states it is not friendly with. H2: The greater the intelligence reception and protection capabilities of an IO, the fewer violations of international law that occur among states that are not friendly with the informed state. (Carnegie and Carson 2019, 271–72)

Table 25: IOs and friends in nuclear weapons regimes: Boolean equations versus linear equations

	IO Unable to Protect Intel	IO Able to Protect Intel
Intel about Friends	Rare Disclosures	Rare Disclosures
	Violations Common	Violations Common
Intel about Non-Friends	Rare Disclosures	Frequent Disclosures
	Violations Common	Violations Less Common

Answer: The Boolean equation which can generate table **??** is: (IO protect) AND (Intel about non-friends) are individually necessary and jointly sufficient for Frequent disclosures. No additive, linear equation can do this. The only linear algebra equation which would work is Y= (IO protect)*(Intel about non-friends), which notably includes no individual variables involving the interaction term. This equation is also nonadditive and nonlinear.

621. Two-way tables are extremely common to express theoretical frameworks in security studies. A core question in the analysis of these tables is the extent to which the content of each cell is the same dependent variable but at different levels versus four different dependent variables. If it is four different dependent variables the causal analysis becomes very complicated, particularly regarding negative cases for these four different dependent variables. Often, however, one can argue there is one underlying dependent variable and each cell is a different level on the dependent variable. Can one make this argument in Ward's (2017) analysis of the behavior of major powers: is there one underlying dependent variable in the table below and can you rank the cells ordinally on this underlying variable?

Table 26: Two-way tables and the dependent variable: status quo norms, distribution and revisionism

	Accept Status Quo Norms, Rules, and Institutions	Reject Status Quo Norms, Rules, and Institutions
Defend (or Decline to Challenge) Distribution of Resources	Satisfied	Normative Revisionist
Examples	*(Pre–World War I Great Britain)*	*(Tokugawa Japan; early Soviet Union)*
Challenge Distribution of Resources	Distributive Revisionist	Radical Revisionist
Examples	*(Prussia/Germany under Bismarck; post-Meiji Japan; Germany under Stresemann)*	*(Revolutionary France; pre–World War I Germany; post-1933 Japan; Nazi Germany)*

Source: Ward 2017.

Answer: It seems like the underlying dependent variable is "level of revisionism." The ordinal levels are (1) satisfied, (2) normative revisionist, (3) distributive revisionist, (4) radical revisionist. As is often the case the ranking between (2) and (3) is perhaps not obvious, a look at the cases in these cells might help decide.

622. Key in constructing and evaluating two-way tables is giving the predicted level of the dependent variable for each cell. What would be the level of the dependent variable Y each cell given in an ordinal fashion from one to four. Can one conclude that the row or column variable is more important and if so why? Could you write a linear equation model that would fit this table's theory (i.e., generate the four Y values in the cells)?

> The more certain threats become, the harder it is to resist acting to stop their emergence. A growing belief that a rising power harbors malign intentions accounts for the change away from cooperation to more competitive strategies. . . . Back to table 1.1, when an existing power prefers short-term cooperation and a rising power avoids provocation, then cooperation between an existing and a rising power becomes likely. As I argued above, I expect this particular configuration of time horizons to be relatively common. (Edelstein 2017, 25–27)

Table 1.1 The interaction of rising and declining power time horizons

		Rising Power	
		Long	*Short*
Declining Power	Long	Increased possibility of hegemonic war	Increased possibility of preventive war
	Short	Conditions conducive to cooperation	Mixture of skirmishes and pragmatic cooperation

Source: Edelstein 2017, 28.

Answer:

The table below provides one interpretation the levels of the dependent variable. If this is correct then Declining power-Long is more powerful than Short. A linear equation which produces this could look like: $Y = 1 + X_1 + 2 * X_2$.

	Rising power–Long	Rising power–Short
Declining power–Long	4	3
Declining power–Short	1	2

623. Steinberg (2015) illustrates perhaps the most popular two-way table structure. Give the values of Y (over-under evaluation of currency) for each cell. Write the Boolean equations for this table. "Exchange rates tend to be more undervalued in developing countries that combine a (1) powerful manufacturing sector *with* (2) state control of labor and finance than in other types of developing countries." (Steinberg 2015, 53; emphasis is mine)

TABLE 1.4. Summary of the argument

LABOR AND FINANCIAL MARKET INSTITUTIONS	MANUFACTURING SECTOR	
	WEAK	***POWERFUL***
PRIVATELY CONTROLLED	Overvalued exchange rate *Preferences*: Manufacturing sector does not prefer undervaluation *Political influence*: Manufacturing sector has limited influence	Overvalued exchange rate *Preferences*: Manufacturing sector does not prefer undervaluation *Political influence*: Manufacturing sector has strong influence
STATE-CONTROLLED	Overvalued exchange rate *Preferences*: Manufacturing sector prefers undervaluation *Political influence*: Manufacturing sector has limited influence	Undervalued exchange rate *Preferences*: Manufacturing sector prefers undervaluation *Political influence*: Manufacturing sector has strong influence

Peters (2017) provides and another example of this very common 2×2 table structure:

TABLE 2.1
Sources of Firms' Heterogeneity of Preferences on Low-Skill Immigration.

		Mobility: Low	Mobility: High
Skill intensity	Low-skill	Supports immigration	Indifferent
	High-skill	Indifferent	Indifferent

Answer:

"With" appears to mean AND and that is confirmed by cell values of 0 (overvalued) in three of the four cells and 1 in the cell with both (1) powerful manufacturing sector *with* (2) state control of labor and finance. The Boolean models are thus: IF X_1 (Manufacturing sector is strong) AND X_2 (Institutions are state controlled) THEN $Y = 1$ (undervalued); IF X_1 (Manufacturing sector is weak) OR X_2 (Institutions are privately controlled) THEN $Y = 0$ (overvalued).

Boolean theories, necessary conditions, QCA, mathematical logic

624. Critical to LNQA is the connection between the causal model and Boolean interpretations of the model. Madrid (2025) provides a very interesting example of a study with close connections to LNQA. He basically looks at democracy-authoritarian outcomes in South America in the 19th century ending in 1929. While the book has democracy in its title most of the countries in South America are not democratic in the 19th or early 20th century. However, he does focus particular attention on the four countries that did achieved democracy by the interwar period. As one can see from the causal model figure below, figure 64, he is not only explaining democracy but also authoritarianism.

Figure 64: Explaining democracy-authoritarianism in 19th century South America

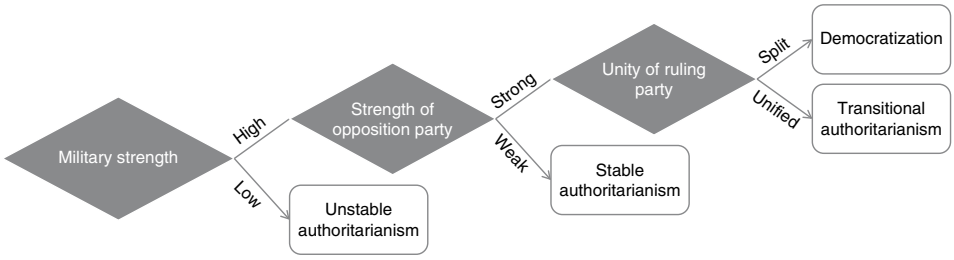


FIGURE 1.1 Determinants of regime outcomes in South America before 1930

The figure has his three main causal factors and he says this:

> [R]egime outcomes in South America during the nineteenth and early twentieth century were shaped by three main variables: the strength of the military, the strength of opposition parties, and the unity of ruling parties (see Figure 1.1). These three variables did not change in a strict sequence, nor were they closely causally related to each other. The strengthening and professionalization of the military did not cause strong opposition parties to emerge, nor did it necessarily precede the rise of strong opposition parties. (Madrid 2025, 38)

There are a number of interesting questions raised by his causal model. For example, would the figure change significantly if the factors were put in a different (temporal) order? Can one give a Boolean interpretation of this causal model, for example, one Boolean model per outcome? For example, to what extent are the three factors sufficient for the outcome of democracy? Are they also necessary for the outcome? The same can be asked of all of the outcomes illustrated in the figure.

625. In classic comparative historical research it is often useful to try to formulate is the author's argument in terms of Boolean equations or figures. There are a variety of potential Boolean equations that arise in Stephen's (1989) classic discussion of Barrington Moore. Analyze the text and see what Boolean equations might correspond to core discussions of the explanatory framework.

626. There is often tension when scholars combine the Boolean algebra of necessary and sufficient conditions in a linear algebra model. Discuss whether T and W are INUS conditions in Cartwright's linear algebra model:

> The SCEM framework is an adaptation for variables with more than two values of J. L. Mackie's (1965) famous account in which causes are INUS conditions for their effects. In the adaptation, causes are INUS conditions for contributions to the effect, where an INUS condition for a contribution to O(i) is an Insufficient but Necessary part of an Unnecessary but Sufficient condition for a contribution to it. Each of the additive terms (a(i)T(i) and W(i)) on the right of the equation O(i) c= a(i)T(i) + W(i) represents a set of conditions that together are sufficient for a contribution to

> O(i) but they are unnecessary since many things can contribute to O; and each component of an additive term (e.g., a(i) and T(i)) is an insufficient but necessary part of it – both are needed and neither is enough alone. (Cartwright 2022, 42–43)
>
> Answer: Neither a(i) nor T(i) is necessary for O(i) and individually they could be sufficient. Hence, they are not INUS conditions.

627. It is quite common in conceptualization to give a series of necessary conditions, the same is often true for coding rules for quantitative datasets. What is often not stated – but is often implicit – is whether these necessary conditions are jointly sufficient. Are Fazal's three necessary conditions jointly sufficient for declaration of war?

> International legal scholarship suggests several necessary criteria for declarations of war. First, a declaration of war must be a public proclamation. Second, it must state intent to engage in hostilities with another party. Third, it must be issued according to the laws of the declaring state; it can only be a declaration of war if authored by an actor empowered with the authority to declare war. (Fazal 2012, 562)

628. The same issue about the sufficiency of multiple necessary conditions arises often in theories as well. Are the multiple necessary conditions sufficient for the outcome? This is a core issue in Boolean theories. Here are two examples where this question arises:

> We expect compensatory layering to occur only when three conditions are met: when states cooperate within a focal institution, when power asymmetries among states are large, and when preferences over organizational design diverge. The first condition, focalness, refers to the convergence of actors' expectations that an institution represents the natural site for cooperation. Schelling introduced the idea of focal points to explain how people select among multiple equilibria in their daily lives. When faced with several similarly beneficial possibilities, actors tend to choose an option based on its "prominence" or "conspicuousness." Schelling is agnostic about the sources of focalness, suggesting that in practice it "may depend on imagination more than on logic." We follow Jupille, Mattli, and Snidal, who argue that two factors shape an institution's focalness. First, actors' past decisions and beliefs create shared expectations that an institution is a "default" or "natural" site of cooperation. Second, focalness depends on an organization's institutional environment; all else being equal, an institution's focalness increases as the number of viable outside options decreases. Institutions with a high degree of focalness will pull in actors even when they are dissatisfied with its design. Second, power asymmetries stem from the material and positional disparities in relationships between two or more states. On the one hand, power asymmetries create tensions over institutional control and the distribution of costs and benefits from cooperation. On the other, they allow powerful states to act as paymasters, bearing the costs of institutional setup and side payments. Asymmetries make it difficult for weaker states to create viable alternatives. In this way, asymmetry also can contribute to focalness – especially the (non)existence of exit options – well beyond the moment of institutional creation. Weaker states may struggle to entice the powerful state away from status quo institutions, or to offer substantial benefits to other weaker states. Third, for compensatory layering to occur, there must

be preference divergence across two or more dimensions of institutional design. There are many reasons why states may disagree on the terms of cooperation. The divergence of preferences can be exacerbated by power asymmetries, given that weaker states may fear powerful states' control of institutions. In a focal organization, actors who are dissatisfied with the status quo will seek to revise the institution rather than abandoning it. The result is a sequence of bargains that is repeated until actors are satisfied, or until the institutional arrangement becomes fundamentally transformed or superseded. The process of compensatory layering only occurs when all three conditions are present. (Long and Schulz 2023, 9–10)

This chapter makes the case that only where these three factors (changing citizenship regimes that challenged local autonomy, transcommunity networks, and political associational space) came together did indigenous movements emerge. As Tables 3.1 and 3.2 lay out, these three factors concatenated in Ecuador, Bolivia, Guatemala, and Mexico by the end of the twentieth century and gave rise to significant indigenous movements in each country. These three factors did not appear together in Peru, where political associational spaces were elusive and community networks were weak – thereby working against indigenous organizing beyond the local level. This chapter develops this three-pronged comparative historical argument against the five most densely populated indigenous countries in Latin America. (Yashar 2005, 23)

629. Jenne (2015) provides another example of a complex theory. In particular notice the mixing of X and Y generalizations. Do a causal mechanism figure and/or write a Boolean equation for the theory.

The argument developed in these pages is that conflict managers of *both* regimes were far more likely to achieve a successful outcome if powerful third parties acted to *nest the domestic disputes in a stable regional environment*. Put another way, cooperative (or soft-power) conflict management requires hard-power backing to achieve success. . . . I argue that effective civil conflict management requires building security from the outside in – that is, stabilizing the external environment before (or at the same time as) reconfiguring domestic institutions. If sectarian conflicts are securely "nested" in a stable regional and hegemonic environment, they are likely to become far more tractable; cooperative conflict management in particular is far likelier to meet with success. . . . *Failing* to securely nest sectarian struggles in a stable external environment may be setting peacemakers and peacekeepers up for an impossible task. . . . This book demonstrates that "nested security" is a key background condition for successful conflict reduction, regardless of the resources of the third party, the intervention strategy, and the attitudes of the government and minority representatives. . . . I demonstrate that nested security is effectively a precondition of success for cooperative conflict management. . . . To establish effective regional security regimes, three features must be present. First, a decision-making body must be established that can cope with internal veto players as well as member-state aggression. Second, a regularized, independent enforcement mechanism must be created with the full backing of regional hegemons and/or great powers. Finally, the regime must engender cooperation between monitors and enforcers to provide effective conflict management at all stages of the conflict – from preventive diplomacy to postwar reconstruction. (Jenne 2015, 9–10, 18, 45, 194)

630. When the author has a complex theory is is useful express it either in Boolean equations or draw causal mechanism figures. Lee has a complicated theory of foreign subversion that is briefly outlined in the quote below. Draw the causal mechanism figure and/or develop a Boolean equation which represents her claims. For example, it seems that there are necessary conditions in her model, e.g., "Only in these situations." She has constraints in her theory e.g., "proxy availability is a key constraint". The loss of strength gradient plays a key role, how would one put that into the figure? Are some factors "facilitating conditions" e.g., the fourth one?

> These observations about costs and constraints have five implications for predicting when subversion will occur. First, as I have already argued, states will be more likely to employ subversion against state authority when confronting adversaries with which they have severe policy disputes. Only in these situations will states be willing to bear the costs and risks of dealing with proxy groups. I call this condition "motive," since states should use subversion when their foreign relations provide for sufficient motivation. Second, subversion will occur only when proxies are available in the target state. Since subversion by definition involves delegation, willing agents, or "means" on the ground, is a necessary condition for subversion. Additionally, treating agent availability as an independent condition implies that proxies can be located by external states but not easily created by them. Without proxies, states cannot engage in subversion. Thus, proxy availability is a key constraint on the use of subversion. Third, as the target's own level of consolidation increases, the more expensive it will be to weaken state administration and control, since these targets are less vulnerable from the outset. Subversion is less likely to succeed against fully consolidated states, which are more informed of local goings-on and will be more likely to defeat efforts to undermine their state authority. Thus, even though subversion of consolidated targets may be attractive, the costs of imposing authority losses may be quite constraining for some states. Knowing this, rational states are unlikely to invest in an expensive strategy that offers little chance of success. We should thus observe few instances of subversion being used against fully consolidated states, which reside primarily in the developed world. Fourth, because material resources, training, and protection and safe havens are the most important means through which sponsors empower proxies, sponsors should be more willing to use subversion when they can easily deliver these investments to their proxies. A key impediment to doing so is distance, because distance increases the delivery cost. Similarly, safe havens and training camps are less valuable to proxies if those areas are difficult to reach. Just as conventional force is subject to a "loss-of-strength gradient," in which military strength declines with distance, subversion is subject to an analogous loss-of-interference gradient, in which the ability to deliver investments declines with distance. As with the loss of strength gradient, distance should act as a constraint for all but the most capable states in the international system. This suggests that subversion will be more likely to be observed between adversaries that are contiguous and in areas bordering those adversaries. This key implication informs the statistical analyses in chapter 3. Fifth, states should be more willing to turn to subversion as the costs of other policy instruments increase. The degree to which states will be willing to bear those costs depends in part on the costs and constraints of alternatives. Scholars of international norms have suggested

> that laws, rules, and norms against the use of force for some ends have made force costly, particularly when would-be violators expect rule enforcement from powerful third-party states. Similarly, defensive military alliances may make conventional force undesirable against some target states. If states anticipate high costs for engaging in conventional force, the primary noncooperative alternative to subversion, then those states will be more likely to favor subversion. The first two implications, about motive and means, are the most important factors that influence the use of subversion, and as I describe later, the primary tests of these propositions appear in chapters 3 and 4. The third and fourth implications, about types of targets and the constraining role of distance, also influence the research design of chapter 3. I provide evidence for the fifth implication, about alternative foreign policy strategies, in chapters 5 and 6. (Lee 2020, 56–57)

Answer: One possible Boolean interpretation of her theory is: R (rivalry) AND RG (at least one rebel group in target) AND not-ED (economically developed) AND [facilitating conditions] (D (distance) + HC (high cost of alternative actions)) is sufficient for subversion. The first two are clearly necessary.

631. Halpern (2016) presents a theory of causation closely connected and developed with Pearl's approach to causal inference. Being a computer scientist it is natural that the "structural equations" he uses are Boolean. He argues that one can include probabilities by making changing the Boolean dichotomous variables to continuous one [0,1]. Discuss whether fuzzy logic or probabilities should be the interpretation of these [0,1] variables.

632. Halpern (2016) presents a theory of causation using Boolean models. These equations have an equal sign (=). How should this be interpreted causally? Is he implicitly treating this to mean "is sufficient for?" How similar or different is "=" from "→"? In addition to what extent do these symbols have causal interpretations? In general do both "=" and "→" receive causal interpretations? Note that the same interpretation issues arise in QCA.

    Answer: Halpern does not really discuss the meaning of = beyond this short comment: "Some variables may have a causal influence on others. This influence is modeled by a set of structural equations. For example, if we want to model the fact that if the arsonist drops a match or lightning strikes then a fire starts, we could use the variables MD, FF, and L as above, with the equation FF = max(L,MD); that is, the value of the variable FF is the maximum of the values of the variables MD and L. This equation says, among other things, that if MD = 0 and L = 1, then FF = 1. The equality sign in this equation should be thought of more like an assignment statement in programming languages; once we set the values of MD and L, the value of FF is set to their maximum."

633. Monty Python's film *Monty Python and the Holy Grail* (1975) is considered one of the classic comedy movies of all times. There is a famous scene where they are deciding whether to burn a woman as a suspected witch which is quite hilarious. Discuss the various logical procedures and errors that they make in debating whether to burn her or not. I suggest you do it this analysis before before looking at the solution on the website: https://cs.uwaterloo.ca/~jhoey/montypythonwitchsceneanalysis.html

This site also provides a video of the clip or you can easily find it on YouTube. There are also versions with subtitles useful because the dialogue does not always come over very clearly. If you like this kind of humor I recommend *A fish called Wanda* one of their best movies.

634. Core to LNQA are regularities. Often these can be complex and expressed as Boolean equations. This appears different guises in Weyland (2020) and Weyland (2023). These are consistent but expressed in different ways (thanks to Weyland for his clarification on this). His basic complex model given in his 2020 article is as follows: (IW AND HDE) OR (IW AND SCR AND ECR) OR (IW AND ECR) ⟶ SDE where

    SDE – suffocation of democracy
    HIN – high instability
    PCH – susceptibility to para-legal change
    ECH – comparatively easy changeability in European parliamentarianism
    HWI – huge hydrocarbon windfalls
    ECR – economic crisis
    SCR – security crisis.

    Hence there are three paths to the outcome of democracy suffocation. In the same article he has a table, table **??** with these six factors but grouped into two categories " Institutional weakness" (IW) and "Exogenous conjuncture" (EC). The LNQA book manuscript discusses this kind of complex model in the chapter on causal complexity (chapter 11) under the rubric of "two-level theories." Work through the the connection of the Boolean equation above and the implicit Boolean models given in the table below. Which possible combinations of these two general factors – IW and EC – does he *not* include in his final model? Why?

Table 27: Conditions for the suffocation of democracy by populist leaders

Table 1
Conditions for the suffocation of democracy by populist leaders

Institutional Weakness			Exogenous Conjuncture			Regime Outcome	Populist Leader
High Instability (HIN)	Paralegal Change (PCH)	Easy Change (ECH)	Economic Crisis (ECR)	Security Crisis (SCR)	Hydrocarbon Windfall (HWI)	Suffocation of Democracy (SDE)	
							Bolivarian Populism—Latin America
1	1	0	1	0	**1**	**1**	**Chávez – VEN**
1	1	0	0	0	**1**	**1**	**Morales – BOL**
1	1	0	0	0	**1**	**1**	**Correa – ECU**
1	1	0	1	0	0	0	(Moreno – ECU)
1	1	0	0	0	0	0	Gutiérrez – ECU
0	1	0	0	0	0	0	Zelaya – HON
0	1	0	0	0	0	0	Lugo – PAR
0	1	0	0	0	0	0	Humala – PER
0	1	0	1	0	0	0	N. Kirchner-ARG
0	1	0	0	0	0	0	C. Kirchner-ARG
1	1	0	0	0	0	0	Colom – GUA
							Neoliberal Populism—Latin America
0	**1**	0	**1**	**1**	0	**1**	**Fujimori – PER**
0	1	0	0	1	0	0	Uribe – COL
0	1	0	1	0	0	0	Menem – ARG
0	1	0	1	0	0	0	García – PER
0	1	0	0	0	0	0	Toledo – PER
0	1	0	0	0	0	0	Bucaram – ECU
0	1	0	1	0	0	0	Collor – BRA
1	1	0	0	0	0	0	Serrano – GUA
							Ethno-National Populism—Europe
0	0	**1**	**1**	0	0	**1**	**Orbán – HUN**
0	0	**1**	**1**	0	0	**1**	**Erdoğan – TUR**
0	0	1	0	0	0	0	Berlusconi – ITA
0	0	1	0	0	0	0	Mečiar – SVK
0	0	1	0	0	0	0	Fico – SVK
0	1	0	0	0	0	0	Borisov – BUL
0	0	1	0	0	0	0	Zeman/Babiš–CZE
0	1	1	0	0	0	0	Băsescu – ROM
0	0	1	0	0	0	0	Kaczyński – POL
0	0	1	0	0	0	0	Papandreou–GRE
0	0	1	0	0	0	0	Tsipras – GRE
0	**0**	**0**	**0**	**0**	**0**	**0**	**Trump—USA**

Source: Weyland 2020.

Answer: If you consider IS to be (HIN OR PCH OR ECH) and EC to be (ECR OR SCR OR HWI) and the high level model to be IW AND EC is sufficient for SDE then there 3*3 possible models combining IW and EC. Presumably the ones he did not include did not appear in any of the cases.

635. Below is one core figure for Levitsky and Way (2022) regarding social revolutions and the durability of authoritarian regimes. First, could or should one interpret this via a Boolean equation and put ANDs, ORs, and S (sufficient for) or N (necessary condition) at appropriate places in the figures? Could you write all the relevant Boolean equations?

Figure 65: Ideal-typical path to durable authoritarianism

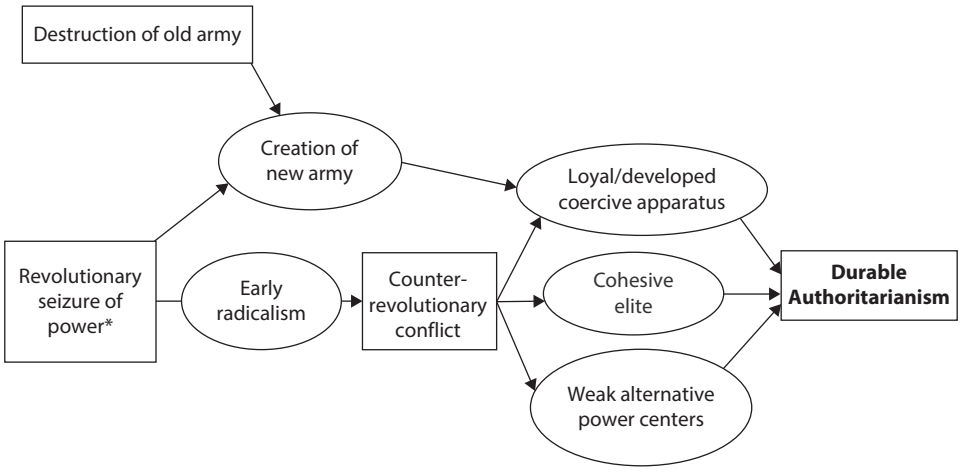


FIGURE 1.2: The Ideal-Typical Revolutionary Reactive Sequence.
* In some cases (e.g., China), the reactive sequence occurs before seizure of power.

Source: Levitsky and Way (2022).

636. Give a Boolean equation representation of Levitsky and Way 2010 account of authoritarian regimes.

Answer: Mahoney and Rodríguez-Caceres give this one: "For example, Levitsky and Way's (2010) theory of the causes of stable authoritarianism can be summarized as follows: (¬W AND (O OR ¬L)) → S. In the equation, W is high Western linkage, O is high organizational power, L is high Western leverage, S is stable authoritarianism, The equation says, "not-high Western linkage and either high organizational power or not-Western leverage (or both) is sufficient for stable authoritarianism." While the equation reads deterministically, Levitsky and Way are clear that their theory does not work for all cases. They report that their theory is fully consistent with outcomes in 28 of 35 cases, with 6 of the 7 non-matches coming close to matching). Hence, only one case is serious outlier." (Mahoney and Rodríguez-Caceres 2022, 10; formated by my Boolean conventions)

637. One of the core guidelines for creating causal mechanism figures discussed in the LNQA book manuscript is the need to provide what it calls aggregation criteria. This means when there are multiple arrows leading to a given box one needs to discuss whether those should be connected with AND, OR, + or whatever. Discuss what might be the appropriate additions to Loxton's theory that also might make it match his empirical data, see the figure below.

Figure 66: Aggregating factors: conservative party-building

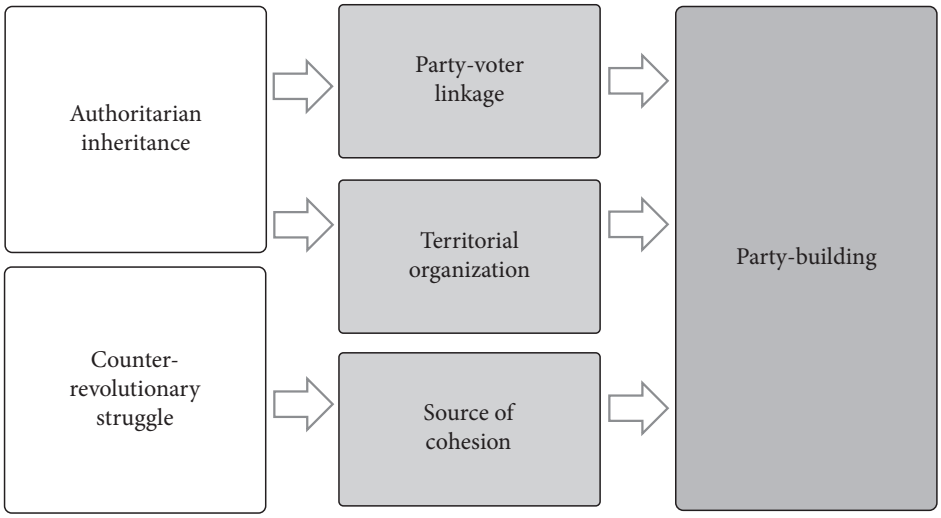


Figure 2.1 Model of the argument.

Source: Loxton 2021.

638. Scholars rarely claim to have identified necessary and sufficient conditions for an outcome. However, sometimes their data tables legitimately support such a conclusion. Discuss whether that is a reasonable thing for Cornell et al. to claim given the data in their table, see the figure below. Note also that their table 4.3, would lead to similar conclusions.

> Table 4.2. Democratic legacy and democratic survival in interwar Europe and neo-Europe . . . . Table 4.3 Democratic legacy and interwar democratic spells in Latin America. . . . In logical terms, a pre-interwar democratic legacy was thus sufficient for democratic survival in Europe and the former British settler colonies and for longer democratic endurance in Latin America. These regularities are so manifest that further assessments of this relationship using standard (probabilistic) statistical techniques have little to add.

Figure 67: Necessary and sufficient conditions: democratic survival in the interwar period

Table 4.2 Democratic legacy and democratic survival in interwar Europe and neo-Europe

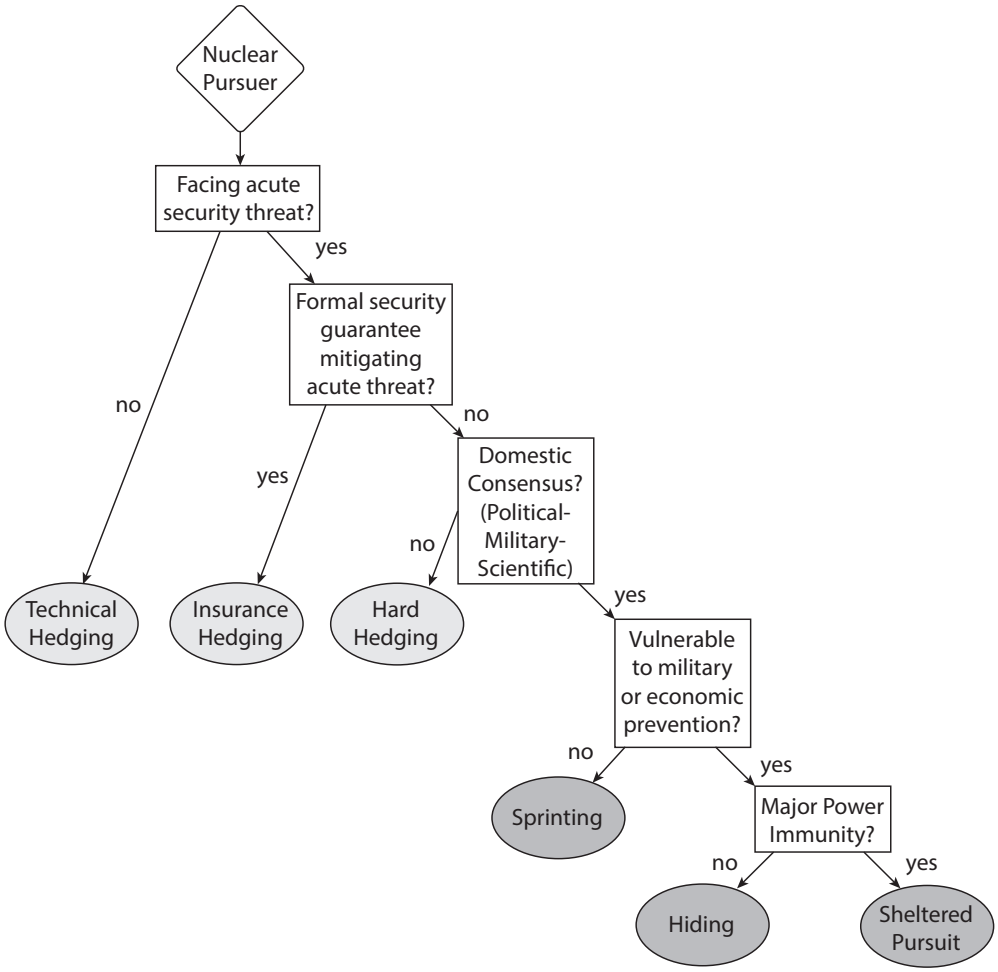
		Democratic survival in the interwar period	
		Yes	**No**
Pre-1918 democratic legacy	**Yes**	Australia, Belgium, Canada, Denmark, France, Ireland, Netherlands, New Zealand, Norway, Sweden Switzerland, UK, USA	
	No	Czechoslovakia, Finland	Austria, Bulgaria, Estonia, Germany, Greece, Italy, Latvia, Lithuania, Poland, Romania, Spain, Yugoslavia (Albania, Hungary, Portugal, USSR)

Note: The countries in parenthesis had no democratic spell during the interwar period.

Source: Cornell et al. 2020.

639. As discussed in the causal mechanism chapter of the LNQA book manuscript, the → has multiple meanings. Often it means a causal relationship, but sometimes it just indicates temporal order. Discuss how one should interpret the arrows in Narang's main theory, see the figure below. Could the order of the boxes be changed if they were not temporal arrows? For example, "facing acute security threats" above "nuclear pursuer"? Or could one write a Boolean model (which almost never has temporality in it)?

Figure 68: The ideal-typical revolutionary reactive sequence.



Source: Narang 2020.

640. In classic comparative historical research it is often useful to try to formulate is the author's argument in terms of Boolean equations or figures. There are a variety of potential Boolean equations that arise in Stephen's (1989) classic discussion of Barrington Moore. Analyze the text and see what Boolean equations might correspond to core discussions of the explanatory framework.

641. There is often tension when scholars combine the Boolean algebra of necessary and sufficient conditions in a linear algebra model. Discuss whether T and W are INUS conditions in Cartwright's linear algebra model:

> The SCEM framework is an adaptation for variables with more than two values of J. L. Mackie's (1965) famous account in which causes are INUS conditions for their effects. In the adaptation, causes are INUS conditions for contributions to the effect, where an INUS condition for a contribution to O(i) is an Insufficient but Necessary part of an Unnecessary but Sufficient condition for a contribution to it. Each of the additive terms

(a(i)T(i) and W(i)) on the right of the equation O(i) c= a(i)T(i) + W(i) represents a set of conditions that together are sufficient for a contribution to O(i) but they are unnecessary since many things can contribute to O; and each component of an additive term (e.g., a(i) and T(i)) is an insufficient but necessary part of it – both are needed and neither is enough alone. (Cartwright 2022, 42–43)

Answer: Neither a(i) nor T(i) is necessary for O(i) and individually they could be sufficient. Hence, they are not INUS conditions.

642. The issue of the sufficiency of multiple necessary conditions arises often in theories. Are the multiple necessary conditions sufficient for the outcome? This is a core issue in Boolean theories. Here are two examples where this question arises:

> We expect compensatory layering to occur only when three conditions are met: when states cooperate within a focal institution, when power asymmetries among states are large, and when preferences over organizational design diverge. The first condition, focalness, refers to the convergence of actors' expectations that an institution represents the natural site for cooperation. Schelling introduced the idea of focal points to explain how people select among multiple equilibria in their daily lives. When faced with several similarly beneficial possibilities, actors tend to choose an option based on its "prominence" or "conspicuousness." Schelling is agnostic about the sources of focalness, suggesting that in practice it "may depend on imagination more than on logic." We follow Jupille, Mattli, and Snidal, who argue that two factors shape an institution's focalness. First, actors' past decisions and beliefs create shared expectations that an institution is a "default" or "natural" site of cooperation. Second, focalness depends on an organization's institutional environment; all else being equal, an institution's focalness increases as the number of viable outside options decreases. Institutions with a high degree of focalness will pull in actors even when they are dissatisfied with its design. Second, power asymmetries stem from the material and positional disparities in relationships between two or more states. On the one hand, power asymmetries create tensions over institutional control and the distribution of costs and benefits from cooperation. On the other, they allow powerful states to act as paymasters, bearing the costs of institutional setup and side payments. Asymmetries make it difficult for weaker states to create viable alternatives. In this way, asymmetry also can contribute to focalness – especially the (non)existence of exit options – well beyond the moment of institutional creation. Weaker states may struggle to entice the powerful state away from status quo institutions, or to offer substantial benefits to other weaker states. Third, for compensatory layering to occur, there must be preference divergence across two or more dimensions of institutional design. There are many reasons why states may disagree on the terms of cooperation. The divergence of preferences can be exacerbated by power asymmetries, given that weaker states may fear powerful states' control of institutions. In a focal organization, actors who are dissatisfied with the status quo will seek to revise the institution rather than abandoning it. The result is a sequence of bargains that is repeated until actors are satisfied, or until the institutional arrangement becomes fundamentally transformed or superseded. The process of compensatory layering only occurs when all three conditions are present. (Long and Schulz 2023, 9–10)

> This chapter makes the case that only where these three factors (changing citizenship regimes that challenged local autonomy, transcommunity networks, and political associational space) came together did indigenous movements emerge. As Tables 3.1 and 3.2 lay out, these three factors concatenated in Ecuador, Bolivia, Guatemala, and Mexico by the end of the twentieth century and gave rise to significant indigenous movements in each country. These three factors did not appear together in Peru, where political associational spaces were elusive and community networks were weak – thereby working against indigenous organizing beyond the local level. This chapter develops this three-pronged comparative historical argument against the five most densely populated indigenous countries in Latin America. (Yashar 2005, 23)

643. It is quite common in conceptualization to give a series of necessary conditions, the same is often true for coding rules for quantitative datasets. What is often not stated – but is often implicit – is whether these necessary conditions are jointly sufficient. Are Fazal's three necessary conditions jointly sufficient for declaration of war?

> International legal scholarship suggests several necessary criteria for declarations of war. First, a declaration of war must be a public proclamation. Second, it must state intent to engage in hostilities with another party. Third, it must be issued according to the laws of the declaring state; it can only be a declaration of war if authored by an actor empowered with the authority to declare war. (Fazal 2012, 562)

644. Jenne (2015) provides another example of a complex theory. In particular notice the mixing of X and Y generalizations. Do a causal mechanism figure and/or write a Boolean equation for the theory.

> The argument developed in these pages is that conflict managers of *both* regimes were far more likely to achieve a successful outcome if powerful third parties acted to *nest the domestic disputes in a stable regional environment*. Put another way, cooperative (or soft-power) conflict management requires hard-power backing to achieve success. . . . I argue that effective civil conflict management requires building security from the outside in –that is, stabilizing the external environment before (or at the same time as) reconfiguring domestic institutions. If sectarian conflicts are securely "nested" in a stable regional and hegemonic environment, they are likely to become far more tractable; cooperative conflict management in particular is far likelier to meet with success. . . . *Failing* to securely nest sectarian struggles in a stable external environment may be setting peacemakers and peacekeepers up for an impossible task. . . . This book demonstrates that "nested security" is a key background condition for successful conflict reduction, regardless of the resources of the third party, the intervention strategy, and the attitudes of the government and minority representatives. . . . I demonstrate that nested security is effectively a precondition of success for cooperative conflict management. . . . To establish effective regional security regimes, three features must be present. First, a decision-making body must be established that can cope with internal veto players as well as member-state aggression. Second, a regularized, independent enforcement mechanism must be created with the full backing of regional hegemons and/or great powers. Finally, the regime must engender cooperation between monitors and enforcers to provide effective conflict

management at all stages of the conflict – from preventive diplomacy to postwar reconstruction.(Jenne 2015, 9–10, 18, 45, 194)

645. When the author has a complex theory is is useful express it either in Boolean equations or draw causal mechanism figures. Lee has a complicated theory of foreign subversion that is briefly outlined in the quote below. Draw the causal mechanism figure and/or develop a Boolean equation which represents her claims. For example, it seems that there are necessary conditions in her model, e.g., "Only in these situations". She has constraints in her theory e.g., "proxy availability is a key constraint". The loss of strength gradient plays a key role, how would one put that into the figure? Are some factors "facilitating conditions" e.g., the fourth one?

> These observations about costs and constraints have five implications for predicting when subversion will occur. First, as I have already argued, states will be more likely to employ subversion against state authority when confronting adversaries with which they have severe policy disputes. Only in these situations will states be willing to bear the costs and risks of dealing with proxy groups. I call this condition "motive," since states should use subversion when their foreign relations provide for sufficient motivation. Second, subversion will occur only when proxies are available in the target state. Since subversion by definition involves delegation, willing agents, or "means" on the ground, is a necessary condition for subversion. Additionally, treating agent availability as an independent condition implies that proxies can be located by external states but not easily created by them. Without proxies, states cannot engage in subversion. Thus, proxy availability is a key constraint on the use of subversion. Third, as the target's own level of consolidation increases, the more expensive it will be to weaken state administration and control, since these targets are less vulnerable from the outset. Subversion is less likely to succeed against fully consolidated states, which are more informed of local goings-on and will be more likely to defeat efforts to undermine their state authority. Thus, even though subversion of consolidated targets may be attractive, the costs of imposing authority losses may be quite constraining for some states. Knowing this, rational states are unlikely to invest in an expensive strategy that offers little chance of success. We should thus observe few instances of subversion being used against fully consolidated states, which reside primarily in the developed world. Fourth, because material resources, training, and protection and safe havens are the most important means through which sponsors empower proxies, sponsors should be more willing to use subversion when they can easily deliver these investments to their proxies. A key impediment to doing so is distance, because distance increases the delivery cost. Similarly, safe havens and training camps are less valuable to proxies if those areas are difficult to reach. Just as conventional force is subject to a "loss-of-strength gradient," in which military strength declines with distance, subversion is subject to an analogous loss-of-interference gradient, in which the ability to deliver investments declines with distance. As with the loss of strength gradient, distance should act as a constraint for all but the most capable states in the international system. This suggests that subversion will be more likely to be observed between adversaries that are contiguous and in areas bordering those adversaries. This key implication informs the statistical analyses in chapter 3. Fifth, states should be more willing to turn to subversion as

> the costs of other policy instruments increase. The degree to which states will be willing to bear those costs depends in part on the costs and constraints of alternatives. Scholars of international norms have suggested that laws, rules, and norms against the use of force for some ends have made force costly, particularly when would-be violators expect rule enforcement from powerful third-party states. Similarly, defensive military alliances may make conventional force undesirable against some target states. If states anticipate high costs for engaging in conventional force, the primary noncooperative alternative to subversion, then those states will be more likely to favor subversion. The first two implications, about motive and means, are the most important factors that influence the use of subversion, and as I describe later, the primary tests of these propositions appear in chapters 3 and 4. The third and fourth implications, about types of targets and the constraining role of distance, also influence the research design of chapter 3. I provide evidence for the fifth implication, about alternative foreign policy strategies, in chapters 5 and 6. (Lee 2020, 56–57)

Answer: One possible Boolean interpretation of her theory is: R (rivalry) AND RG (at least one rebel group in target) AND not-ED (economically developed) AND [facilitating conditions] (D (distance) + HC (high cost of alternative actions)) is sufficient for subversion. The first two are clearly necessary (see the Russia chapter).

646. Halpern (2016) presents a theory of causation closely connected and developed with Pearl's approach to causal inference. Being a computer scientist it is natural that the "structural equations" he uses are Boolean. He argues that one can include probabilities by making changing the Boolean dichotomous variables to continuous one [0,1]. Discuss whether fuzzy logic or probabilities should be the interpretation of these [0,1] variables.

647. Halpern (2016) presents a theory of causation using Boolean models. These equations have an equal sign (=). How should this be interpreted causally? Is he implicitly treating this to mean "is sufficient for?" How similar or different is "=" from "→"? In addition to what extent do these symbols have causal interpretations? In general do both "=" and "→" receive causal interpretations? Note that the same interpretation issues arise in QCA.

    Answer: Halpern does not really discuss the meaning of = beyond this short comment: "Some variables may have a causal influence on others. This influence is modeled by a set of structural equations. For example, if we want to model the fact that if the arsonist drops a match or lightning strikes then a fire starts, we could use the variables MD, FF, and L as above, with the equation FF = max(L,MD); that is, the value of the variable FF is the maximum of the values of the variables MD and L. This equation says, among other things, that if MD = 0 and L = 1, then FF = 1. The equality sign in this equation should be thought of more like an assignment statement in programming languages; once we set the values of MD and L, the value of FF is set to their maximum".

648. Monty Python's film *Monty Python and the Holy Grail* (1975) is considered one of the classic comedy movies of all times. There is a famous scene where they are deciding whether to burn a woman as a suspected witch which is quite hilarious. Discuss the various logical procedures and errors

that they make in debating whether to burn her or not. I suggest you do it this analysis before before looking at the solution on the website: https://cs.uwaterloo.ca/~jhoey/montypythonwitchsceneanalysis.html

This site also provides a video of the clip or you can easily find it on YouTube. There are also versions with subtitles to be useful because the dialogue does not always come over very clearly. If you like this humor I recommend *A fish called Wanda* one of their best movies.

649. Core to LNQA are regularities. Often these can be complex and expressed as Boolean equations. This appears different guises in Weyland (2020) and Weyland (2023). These are consistent but expressed in different ways (thanks to Weyland for his clarification on this). His basic complex model given in his 2020 article is as follows: (IW AND HDE) OR (IW AND SCR AND ECR) OR (IW AND ECR) $\longrightarrow$ SDE where

    SDE – suffocation of democracy
    HIN – high instability
    PCH – susceptibility to para-legal change
    ECH – comparatively easy changeability in European parliamentarianism
    HWI – huge hydrocarbon windfalls
    ECR – economic crisis
    SCR – security crisis.

    Hence there are three paths to the outcome of democracy suffocation. In the same article he has a table, table **??** with these six factors but grouped into two categories " Institutional weakness" (IW) and "Exogenous conjuncture" (EC). The LNQA book manuscript discusses this kind of complex model in chapter on causal complexity (chapter 11) under the rubric of "two-level theories." Work through the the connection of the Boolean equation above and the implicit Boolean models given in the table below. Which possible combinations of these two general factors – IW and EC – does he *not* include in his final model? Why?

Table 28: Conditions for the suffocation of democracy by populist leaders

Table 1
Conditions for the suffocation of democracy by populist leaders

Institutional Weakness			Exogenous Conjuncture			Regime Outcome	Populist Leader
High Instability (HIN)	Paralegal Change (PCH)	Easy Change (ECH)	Economic Crisis (ECR)	Security Crisis (SCR)	Hydrocarbon Windfall (HWI)	Suffocation of Democracy (SDE)	
							Bolivarian Populism—Latin America
1	1	0	1	0	**1**	**1**	**Chávez – VEN**
1	1	0	0	0	**1**	**1**	**Morales – BOL**
1	1	0	0	0	**1**	**1**	**Correa – ECU**
1	1	0	1	0	0	0	(Moreno – ECU)
1	1	0	0	0	0	0	Gutiérrez – ECU
0	1	0	0	0	0	0	Zelaya – HON
0	1	0	0	0	0	0	Lugo – PAR
0	1	0	0	0	0	0	Humala – PER
0	1	0	1	0	0	0	N. Kirchner-ARG
0	1	0	0	0	0	0	C. Kirchner-ARG
1	1	0	0	0	0	0	Colom – GUA
							Neoliberal Populism—Latin America
0	**1**	0	**1**	**1**	0	**1**	**Fujimori – PER**
0	1	0	0	1	0	0	Uribe – COL
0	1	0	1	0	0	0	Menem – ARG
0	1	0	1	0	0	0	García – PER
0	1	0	0	0	0	0	Toledo – PER
0	1	0	0	0	0	0	Bucaram – ECU
0	1	0	1	0	0	0	Collor – BRA
1	1	0	0	0	0	0	Serrano – GUA
							Ethno-National Populism—Europe
0	0	**1**	**1**	0	0	**1**	**Orbán – HUN**
0	0	**1**	**1**	0	0	**1**	**Erdoğan – TUR**
0	0	1	0	0	0	0	Berlusconi – ITA
0	0	1	0	0	0	0	Mečiar – SVK
0	0	1	0	0	0	0	Fico – SVK
0	1	0	0	0	0	0	Borisov – BUL
0	0	1	0	0	0	0	Zeman/Babiš–CZE
0	1	1	0	0	0	0	Băsescu – ROM
0	0	1	0	0	0	0	Kaczyński – POL
0	0	1	0	0	0	0	Papandreou–GRE
0	0	1	0	0	0	0	Tsipras – GRE
0	**0**	**0**	**0**	**0**	**0**	**0**	**Trump—USA**

Source: Weyland 2020.

Answer: If you consider IS to be (HIN OR PCH OR ECH) and EC to be (ECR OR SCR OR HWI) and the high level model to be IW AND EC is sufficient for SDE then there 3*3 possible models combining IW and EC. Presumably the ones he did not include did not appear in any of the cases.

650. For much qualitative work, and for many theories in general, it is useful to give a Boolean equation which represents the structure of the argument. Give a Boolean equation that describes the core argument of the influential Levitsky and Way book on competitive-authoritarian regimes (2010).

Answer: Here is Mahoney's Boolean description of Levitsky and Way: "For example, Levitsky and Way's (2010) theory of the causes of stable authoritarianism can be summarized as follows: (W AND (O OR L) $\Rightarrow$ S. In the equation, W is high Western linkage, O is high organizational power, L is high Western leverage, S is stable authoritarianism, ... and $\Rightarrow$ is both sufficiency and temporal order. The equation says, 'not-high Western linkage and either high organizational power or not-Western leverage (or both) is sufficient for stable authoritarianism." (Mahoney and Rodríguez-Caceres 2022, 10–11)

651. Below is one core figure for Levitsky and Way (2022) regarding social revolutions and the durability of authoritarian regimes. First, could or should one interpret this via a Boolean equation and put ANDs, ORs, and S (sufficient for) or N (necessary condition) at appropriate places in the figures? Could you write all the relevant Boolean equations?

Figure 69: Ideal-typical path to durable authoritarianism

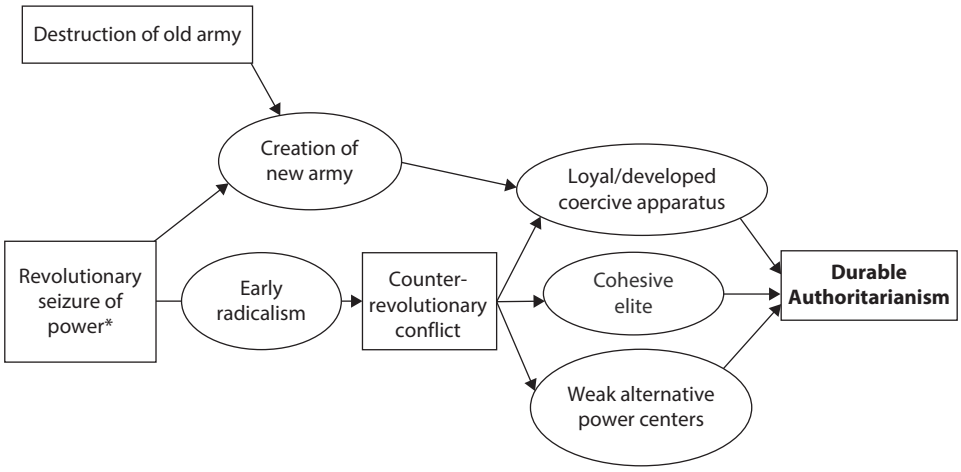


FIGURE 1.2: The Ideal-Typical Revolutionary Reactive Sequence.
* In some cases (e.g., China), the reactive sequence occurs before seizure of power.

Source: Levitsky and Way (2022).

NOTE: For calibration exercises see the semantic transformation section above. See also the section "Two-level theories" for many examples involving QCA, INUS, and SUIN. See also the two-way table section, p. 301 for analyzing two-way tables in Boolean terms.

652. Pearl and McKenzie (2017) spend almost a whole chapter on causal inference in the history of smoking and lung cancer. They note one of the earliest results, and very striking results, which people of the time could not really wrap their heads around:

> Of course Hill knew that an RCT was impossible in this case, but he had learned the advantages of comparing a treatment group to a control group. So he proposed to compare patients who had already been diagnosed with cancer to a control group of healthy volunteers. Each group's members were interviewed on their past behaviors and medical histories. To avoid bias, the interviewers were not told who had cancer and who was a control. The results of the study were shocking: out of 649 lung cancer patients interviewed, all but two had been smokers.... The probability logic is backward too. The data tell us the probability that a cancer patient is a smoker instead of the probability that a smoker will get cancer. It is the latter probability that really matters to a person who wants to know whether he should smoke or not. (Pearl and Mackenzie 2017, 192–93)

Why is their question about smoking-cancer framed in the wrong way given these data?

Answer: Note that they say that the probabilities are backward which is exactly the issue with necessary conditions and analytic induction. As in the

previous exercise, they are asking what is the probability that I will get lung cancer if I smoke. This is the wrong question given the data in question. As Pearl and MacKenzie note these data are not very useful in answering the question "What is the probability that I will get lung cancer if I smoke. However, if I want to *avoid* lung cancer the advice is clear and the data speak directly to that. The question is "What is the probability I will not get lung cancer if I do not smoke?" While there may be other contributing causes to all the cases of lung cancer in 99 percent of cases smoking is part of the story (an necessary condition).

653. Bernard and his colleagues in their analyzing qualitative data textbook discuss a technique known as analytic induction (Ragin has a book manuscript in progress on this topic). Typically analytic induction is looking at all the $Y = 1$ cases and then inductively trying to find a common cause. One of the earliest and strongest critiques of this was that this procedure did not allow one to "predict" outcomes. This meant that if one knew the common cause one cannot predict the outcome in other cases. Discuss this issue in terms of analytic induction finding necessary conditions and what one can predict if one has a necessary condition. Also discuss extent to which these authors when they talk about "prediction" are really asking about sufficient conditions. If one has a sufficient condition for an outcome then one can predict that it will occur when that sufficient condition occurs.

> Cressey [a classic of analytic induction] could not predict, a priori – i.e., without data about actual embezzlers who had been arrested and jailed for their crime – which bank workers would violate the trust of their employers. Manning could not predict, a priori, which pregnant women would ultimately seek an abortion. And Zeoli et al. couldn't predict which mothers' behavior would not confirm their expectations. . . . Analytic induction fell out of favor after the 1950s because the method accounts for data you've already collected but does not allow prediction about individual cases. While it does not produce perfect knowledge for the prediction of individual cases, it can do as well as statistical induction – the standard in social science – in predicting the outcome in aggregates of cases, and it does so with a relatively small number of cases. (Bernard et al. 2017, 563, 579)

654. A typical statistical reflex when faced with a scatterplot is to draw a line through it. One of the classic debates in political science deals with the relationship between level of wealth and democracy. Discuss the extent to which scatterplot in figure below (Norris 2017) is best described as a sufficient condition, based on fuzzy logic, that establishes a floor for the level democracy above which there can be significant variation. There are only four cases which violate the sufficient condition–floor interpretation, these would merit special attention. Note that many of the other countries with large residuals on the quadratic curve are not problematic in this fuzzy logic interpretation.

Figure 70: Wealth as a floor for the level of democracy

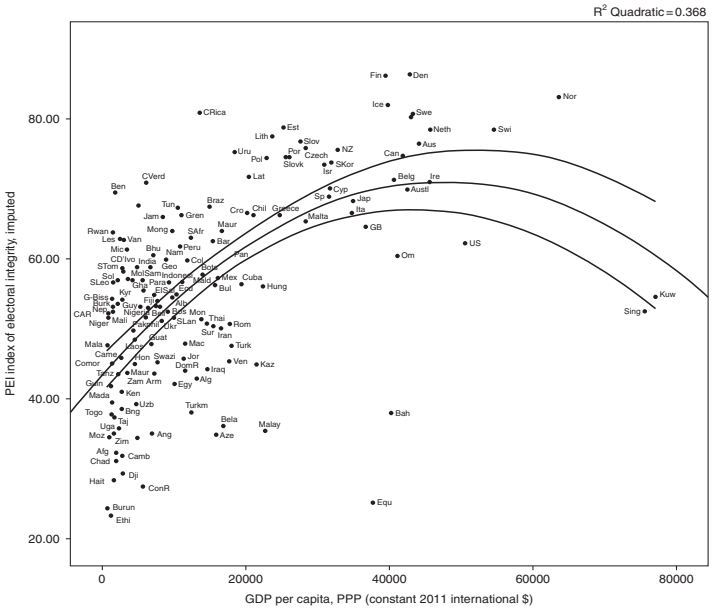


FIGURE 4.5. Electoral integrity and economic development.
Note: The Perception of Electoral Integrity Index (PEI-4.5); GDP per capita (in purchasing parity power) World Bank Development Indicators.
Source: Pippa Norris, Ferran Martinez i Coma, Alessandro Nai and Max Groemping. *The Expert Survey of Perceptions of Electoral Integrity*. (PEI-4.5). Available at www .electoralintegrityproject.com.

655. Often two-way tables are best interpreted via Boolean logic. Table **??** below provides a core summary of the Carnegie and Carson game theoretic model predictions with the two core independent variables on the row and the column of the table. Give the Boolean equation that generates the table. Analyze and the connection between this table and the formal game theoretic model described in the article. Discuss the cases which are most relevant for analyzing various parts of the Boolean equation. Discuss the hypotheses the relationship between the Boolean equations that can generate the table below and the implicit linear, additive hypotheses in their hypotheses H1 and H2.

> We argue that international organizations constitute a potential solution to disclosure dilemmas. . . . We claim that an IO can serve this purpose under two primary conditions. First, an IO must have the capacity to credibly review, assess, and act on sensitive information, which requires a reputation for technical expertise and relatively unbiased judgment; otherwise, it cannot provide added legitimacy to address the intelligence holder's credibility problem. Second, it must be designed to receive and protect sensitive information by limiting its dissemination within the IO and preventing unauthorized leaks. . . . These [game theoretic] claims lead to two testable hypotheses, which are summarized in Table 1 [table **??**]. H1: The greater the intelligence reception and protection capabilities of an IO, the more intelligence an informed state discloses to it about states it is not friendly with. H2: The greater the intelligence reception and protection capabilities of an IO, the fewer violations of international law that occur

among states that are not friendly with the informed state. (Carnegie and Carson 2019, 271–72)

Table 29: IOs and friends in nuclear weapons regimes: Boolean equations versus linear equations

	IO Unable to Protect Intel	IO Able to Protect Intel
Intel about Friends	Rare Disclosures	Rare Disclosures
	Violations Common	Violations Common
Intel about Non-Friends	Rare Disclosures	Frequent Disclosures
	Violations Common	Violations Less Common

Answer: The Boolean equation which can generate table **??** is: (IO protect) AND (Intel about non-friends) are individually necessary and jointly sufficient for Frequent disclosures. No additive, linear equation can do this. The only linear algebra equation which would work is Y= (IO protect)*(Intel about non-friends), which notably includes no individual variables involving the interaction term. This equation is also nonadditive and nonlinear.

656. In the punctuated equilibrium framework shocks are often conceived of as a necessary condition for rapid punctuated change. Discuss the extent to which a military hegemonic shock is a necessary condition for regime change in Gunitsky's work as illustrated in the figure below.

> If these arguments are correct, military hegemonic shocks should increase the likelihood of great powers temporarily choosing to promote their own regimes rather than securing mere loyalty. This can be tested empirically by looking at the rates and types of regime promotions over the twentieth century. As Figures 4 and 5 show, the likelihood of great powers imposing their own regimes increases significantly in the wake of military hegemonic shocks. During the twentieth century, great powers are responsible for seventy-two of the 121 external impositions (about 60 percent). However, great powers nearly monopolize regime promotion in the wake of military hegemonic transitions, when they are promoters in thirty-one of thirty-four cases. (Gunitsky 2014, 568, 570)

Figure 71: Hegemonic shocks as a necessary condition for regime imposition

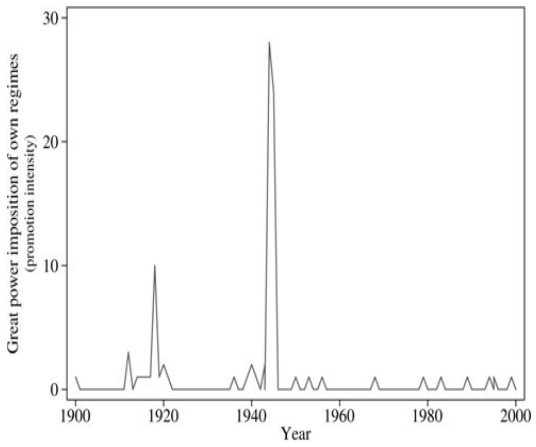


FIGURE 4. *Impositions of their own regime by great powers, 1900–2000*
Notes: The figure measures the rate of hegemonic impositions of their own regime (measured as the product of the number of great powers imposing their own regimes and the number of states experiencing such impositions in a given year). See Appendix 2 for the full list of impositions and their classifications, and an expanded discussion of the measures.

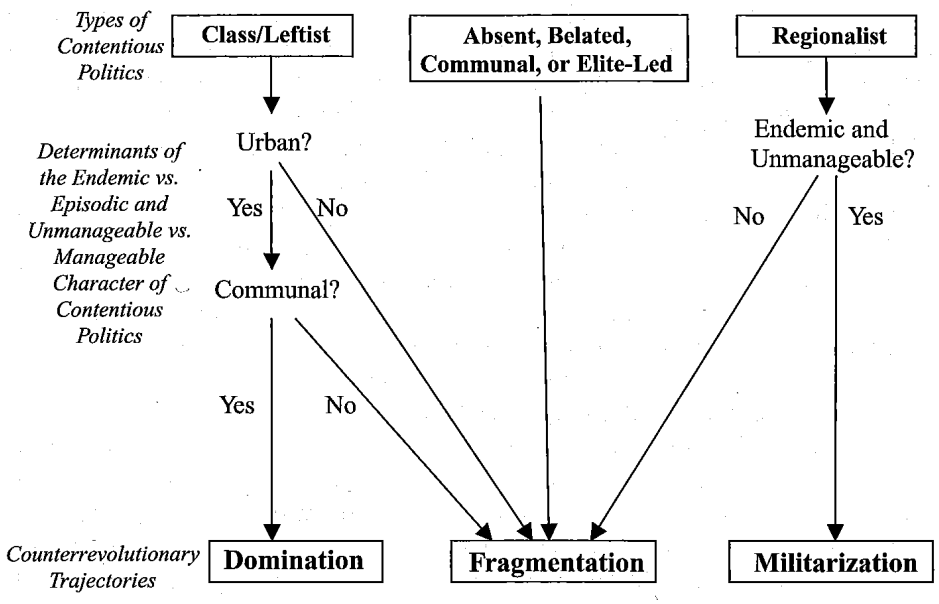
657. Scholars frequently use the language of necessary and sufficient conditions. It is sometimes the case it is not really clear if they really mean that, do the data or theory support it. Discuss this example:

> For the 500 US denomination game, we find qualitative support for the claim that players exhibit in-group bias if and only if they are observed, although the results do not reach significance at conventional levels (row 3; columns 3 and 4). In the 100 US denomination game, however, we find strong evidence that players discriminate in favor of co-ethnics *if and only if* they can be seen to be doing so (row 3; columns 1 and 2). Taken together, these findings offer strong support for this strategy selection mechanism as an important source of the variation we observe in public goods provision across ethnically homogeneous and heterogeneous settings. ((Habyarimana et al. 2007, 721, emphasis is the authors')

658. Sufficient condition scatterplots such as the one below suggest that there is potentially a floor below which observations cannot go below. Discuss the two sufficient conditions in the scatterplots and contrast them with the one nonsufficient condition one. Discuss the extent to which popular participation in the initial stages of constitution making sets a floor for how democratic the country can be later on. Note that this is one of the strongest findings book.

> Next, we disaggregate the participation variable into convening, debating, and ratifying stages in order to test an "origination" hypothesis predicting that the first stage, convening, has the largest impact on democracy. We confirm that participation at this earliest stage is most critical: democracy improved in only 45 percent of cases that incorporated broad consultation at debate and ratification stages, but not at the convening stage. Contrarily, 82 percent of the cases in our data that used popular convening, regardless of popular participation in later stages, show such improvement. (Eisenstadt et al. 2017, 144

FIGURE 4. Contentious Politics and Counterrevolutionary Trajectories



659. Poast's (2019) basic theory is A OR (not-A and not-B) then Y (successful alliance agreement), A is compatible war plans and B is outside options He also explicitly claims not-A and B is sufficient for failure. What are the case studies for studying all parts of this theory, in particular the falsifying ones? Which ones does he include?

660. Write the Boolean equation for the text below. Compare that to the inverted tree causal mechanism figure, reproduced below.

> As such, if a state fears future war, it will adopt coercion against the secessionists to foreclose the possibility of such threats. Conversely, sanguinity about the future is *necessary* for the state to consider peaceful concessions, including the granting of full independence. Whether a state coerces separatists, then, depends on whether it believes it will face future war, which in turn depends on two factors. With respect to the seceded ethnic group, the state concludes future war is likely if there is a deep identity division between the group and the central state. With respect to the existing rivals, the state assesses future war as likely if its regional neighborhood has a militarized history, marked by conflict and war. If the state chooses coercion based on either of these "trip wires," the extent of third-party support for the secessionists determines how much violence the state employs, for both materialist and emotional reasons. Materially, external backing makes the rebel movement stronger, increasing the amount of violence required to defeat it. Emotionally, deep alliances with rivals of the state can lead to pathological violence, fueled by a sense of betrayal. External security, then, is key to understanding both whether, and how much, states coerce secessionists. (Butt 2017, 2)

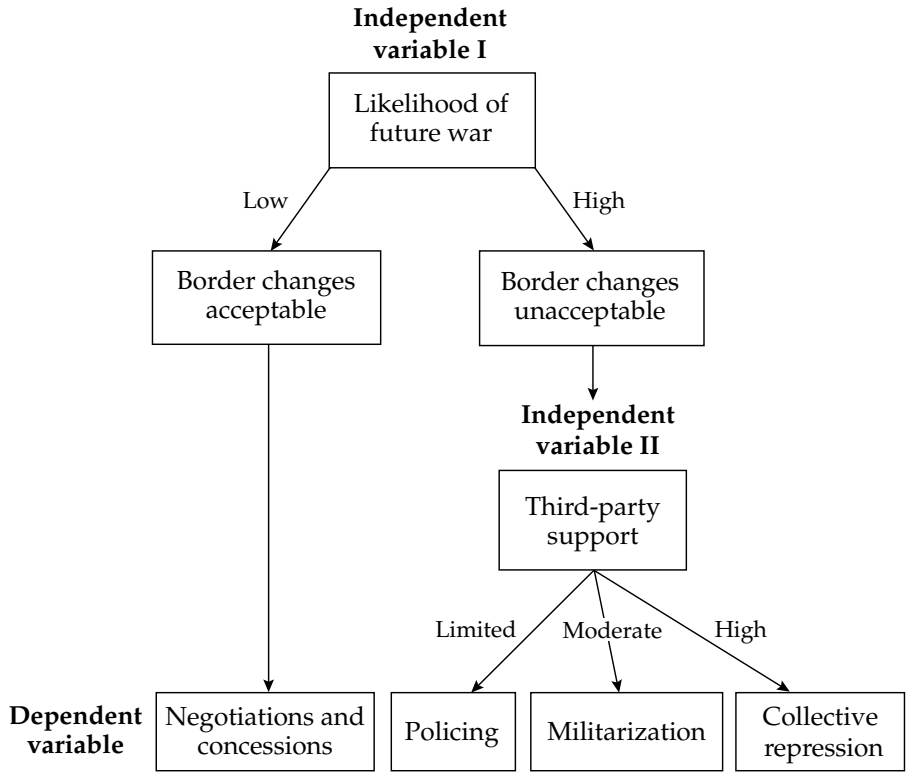


Figure 1. State decision-making when confronted by separatists

Source: Butt 2017, 40.

661. SUIN appears in Boolean theories on a regular basis. Discuss Maestro's theory as SUIN. Write a Boolean equation for the quoted text.

> In the next chapter I present the costly conversations thesis, which argues that decision makers are primarily concerned with the strategic costs of conversation; these costs need to be low before decision makers are willing to engage diplomatically with their enemy. Specifically, leaders look at two factors when determining the costs of conversation: the likelihood the enemy will interpret weakness from an open diplomatic posture, and how the enemy may change its strategy in response to such an interpretation. As leaders' perceptions about these two factors evolve, so too does the cost-benefit analysis of different diplomatic postures. *Only if* a state thinks that it has adequately demonstrated strength and resiliency to avoid adverse inference and that its enemy does not have the capacity to prolong, escalate, or intensify the war in response will it choose an open diplomatic posture. (Maestro 2019, 7; emphasis is mine)

662. Baum and Potter (2015) in their analysis of democracy and war in chapter 7 do a series of case study analyses from the coalition built before the 2003 Iraq war. This involves decisions to join or not to join the coalition. Their main theory is democratic "constraints" on war initiation. "we have attempted to select cases that vary meaningfully on the dimensions that we argue are key to democratic constraint: partisan opposition and a robust and accessible media.

The conditional nature of this argument means that we anticipate that when either (or both) of these conditions is lacking, it will be possible for leaders to discount public opinion and pursue their independent policy preferences, if they have them. Only when both are present will public opinion truly constrain leaders. Table 7.1 locates our cases, in relative terms, along these two key structural dimensions: opposition and media access." (Baum and Potter 2015, 195). Discuss the (0,0) cell where there are no constraints on the decision to join the coalition, how would one do case studies for this particular cell? Give the Boolean equation for their theory. How do the cases selected line up with the Boolean equation?

663. Key in constructing and evaluating two-way tables is giving the predicted level of the dependent variable for each cell. What would be the level of the dependent variable Y each cell given in an ordinal fashion from one to four. Can one conclude that the row or column variable is more important and if so why? Could you write a linear equation model that would fit this table's theory (i.e., generate the four Y values in the cells)?

> The more certain threats become, the harder it is to resist acting to stop their emergence. A growing belief that a rising power harbors malign intentions accounts for the change away from cooperation to more competitive strategies. . . . Back to table 1.1, when an existing power prefers short-term cooperation and a rising power avoids provocation, then cooperation between an existing and a rising power becomes likely. As I argued above, I expect this particular configuration of time horizons to be relatively common. (Edelstein 2017, 25–27)

Table 1.1 The interaction of rising and declining power time horizons

		Rising Power	
		Long	*Short*
Declining Power	Long	Increased possibility of hegemonic war	Increased possibility of preventive war
	Short	Conditions conducive to cooperation	Mixture of skirmishes and pragmatic cooperation

Source: Edelstein 2017, 28.

Answer:

The table below provides one interpretation the levels of the dependent variable. If this is correct then Declining power-Long is more powerful than Short. A linear equation which produces this could look like: $Y = 1 + X_1 + 2 * X_2$.

	Rising power–Long	Rising power–Short
Declining power–Long	4	3
Declining power–Short	1	2

664. Steinberg (2015) illustrates perhaps the most popular two-way table structure. Give the values of Y (over-under evaluation of currency) for each cell. Write the Boolean equations for this table. "Exchange rates tend to be more undervalued in developing countries that combine a (1) powerful manufacturing sector *with* (2) state control of labor and finance than in other types of developing countries." (Steinberg 2015, 53; emphasis is mine)

TABLE 1.4. Summary of the argument

LABOR AND FINANCIAL MARKET INSTITUTIONS	MANUFACTURING SECTOR	
	WEAK	***POWERFUL***
PRIVATELY CONTROLLED	Overvalued exchange rate *Preferences*: Manufacturing sector does not prefer undervaluation *Political influence*: Manufacturing sector has limited influence	Overvalued exchange rate *Preferences*: Manufacturing sector does not prefer undervaluation *Political influence*: Manufacturing sector has strong influence
STATE-CONTROLLED	Overvalued exchange rate *Preferences*: Manufacturing sector prefers undervaluation *Political influence*: Manufacturing sector has limited influence	Undervalued exchange rate *Preferences*: Manufacturing sector prefers undervaluation *Political influence*: Manufacturing sector has strong influence

Peters (2017) provides and another example of this very common 2×2 table structure:

TABLE 2.1
Sources of Firms' Heterogeneity of Preferences on Low-Skill Immigration.

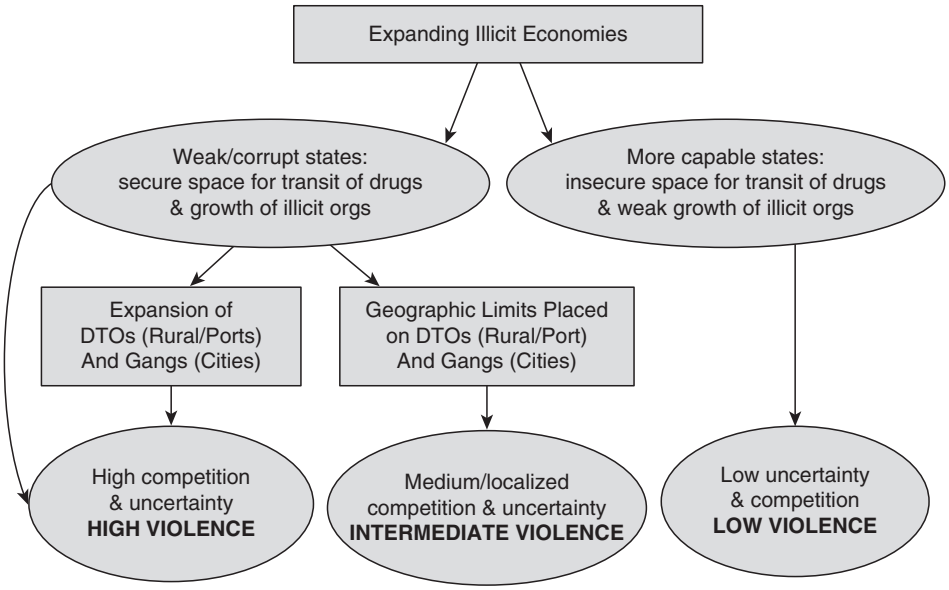
		Mobility: Low	Mobility: High
Skill intensity	Low-skill	Supports immigration	Indifferent
	High-skill	Indifferent	Indifferent

Answer:

"With" appears to mean AND and that is confirmed by cell values of 0 (overvalued) in three of the four cells and 1 in the cell with both (1) powerful manufacturing sector *with* (2) state control of labor and finance. The Boolean models are thus: IF X_1 (Manufacturing sector is strong) AND X_2 (Institutions are state controlled) THEN $Y = 1$ (undervalued); IF X_1 (Manufacturing sector is weak) OR X_2 (Institutions are privately controlled) THEN $Y = 0$ (overvalued).

665. Inverted tree figures are not uncommon in qualitative work. Here is a nice example from Yashar (2018). It also illustrates some issues in drawing causal mechanism figures using just →. How might some arrows have a different Boolean interpretation than others. Could one write the Boolean equations for this figure?

> In particular, this book emphasizes three factors: the transnational illicit economy, state capacity, and territorial competition among organizations. . . . First, I contend that the development of a transnational illicit economy and illicit criminal organizations set the stage for the high levels of violence that we now see in Latin America. . . . Second, I argue that illicit trade and transit is likely to take hold where illicit actors find weak and/or complicit state institutions (particularly law-and-order institutions such as the police and courts). . . . Third, and finally, I find that the highest levels of violence are emerging particularly where illicit organizations encounter organizational competition (either from other illicit organizations or the state or both) to control previously hegemonic territorial enclaves. . . . No single factor determines the outcome. The combination of factors, however, can be deadly. (Yashar 2018, 18–19)



Answer:

The arrows to the bottom level of the tree in a Boolean interpretation mean "is sufficient for." One plausible interpretation of the arrows above are as invoking the logical AND. Here is one Boolean equation: IF (illicit economic AND weak state AND DTO expansion) THEN high violence.

666. With complex Boolean equations case selection can be quite challenging. Discuss Brookes's case selection strategy. Is she using the Overdetermination Guideline in selecting her cases? Which cases are good for the necessary condition hypothesis and which for sufficiency? With four variables in her model there are 16 possible combinations for case study selection. She choose

6. Which ones does she leave out? Discuss in the context of the Overdetermination Guideline.

> *Case Selection and Methods of Causal Analysis.* To test the causal claims of the CCAP theory, in this book I analyze six TLA campaigns – three failures and three successes – matched into pairs of highly The New Politics of Transnational Labor similar cases with different outcomes: the Liverpool dockers' dispute versus the Australian waterfront conflict; the Tesco Global Union Alliance versus the Alliance for Justice at G4S; and the Shangri-La Hotel campaign versus the Raffles Hotels campaign. This logic of case selection approximates John Stuart Mill's method of difference insofar as it allows one to control for otherwise confounding factors within each pair, including the type of conflict prompting the campaign, the institutional context of the country in which the conflict originated, the TLA's goals, the type of employer targeted, whether the campaign was proactive or reactive, and the time period in which the campaign took place. If the CCAP theory is correct, any campaign lacking intraunion coordination, interunion coordination, or context-appropriate power should be unsuccessful. Conversely, all three variables should be present in each successful campaign. . . . Process tracing thus allows me to identify and test for alternative explanations. In sum, across-case comparisons are important, but only within-case analyses can confirm or falsify the theoretical framework proposed in this book. . . . *Causal Relationships as Boolean Equations* $X1 * X2 * X3 = Y$; $\neg X1 + \neg X2 + \neg X3 = \neg Y$ (Brookes 2019, 34–35, 160;)

667. When one combines necessary condition language with increases the probability language is not always clear that the author is consistent. Discuss this claim in the connection between the necessary condition statement and the probability one.

> With regard to the roles of leadership and organization, the causal connections are different. It is certainly possible to say, as I do, that some degree of organization is necessary for sustained political action or mobilization. However, it is also accurate, and more useful, to state the effect in terms of relative probability: the better the group's organization, the greater the likely degree of mobilization. (Kaufman 2015, 242)

668. Debs and Monteiro (2017) in their theory of proliferation of nuclear weapons frequently talk about necessary conditions for proliferation which is quite common in this literature in general. Discuss how one might introduce these considerations into their causal mechanism figure, indicating which factors are necessary, and eventually the various paths to sufficiency.

> Empirically, proliferation occurs in a limited range of strategic environments. Specifically, we find two sets of strategic circumstances – or pathways – to nuclear acquisition. First, a high level of security threat combined with high relative conventional power on the part of the proliferating state. Second, a high level of security threat combined with the presence of an ally that is deemed unreliable. All other strategic settings result in the maintenance of a state's nonnuclear status. . . . In doing so, the strategic logic of nuclear proliferation sheds light on several hitherto underappreciated historical patterns. First, states that do not face a high-level security threat have not acquired the bomb. The presence of a significant

security threat is a necessary condition for nuclearization. Historically, no state has acquired nuclear weapons without perceiving its security environment as highly threatening, regardless of how strong other pressures to acquire the bomb – including considerations of domestic or international prestige, the psychology of leaders, or the economic preferences of ruling elites – may be. Second, among states that are not protected by a great power sponsor, only those that are strong vis-à-vis their adversaries have acquired the bomb. There is no historical case of a relatively weak state ever succeeding in nuclearizing without having a powerful ally committed to retaliating against a preventive counter-proliferation strike. . . . Third, among states that possess a powerful ally, only those whose security goals are not entirely covered by this sponsor have acquired nuclear weapons. Put differently, states whose security goals are subsumed by their powerful allies' own aims do not possess the willingness to acquire the bomb. . . . Fourth, threats of abandonment issued by a security sponsor – what we call a "sticks-based" nonproliferation policy – are effective in curtailing proliferation only by protegees that are relatively weak vis-à-vis their adversaries. If a protegee is strong vis-à-vis its adversaries, it has the opportunity to proliferate on its own, even if its security sponsor were to abandon it. In this case, the sponsor can only effectively deter proliferation by taking away the protegee's willingness to acquire nuclear weapons, which it can do by extending additional security assurances – what we call a "carrots-based" nonproliferation policy. In other words, whereas sticks can deter proliferation by weak protegee, only carrots will prevent stronger protegees from building nuclear weapons. (Debs and Monteiro 2017, 11–13)

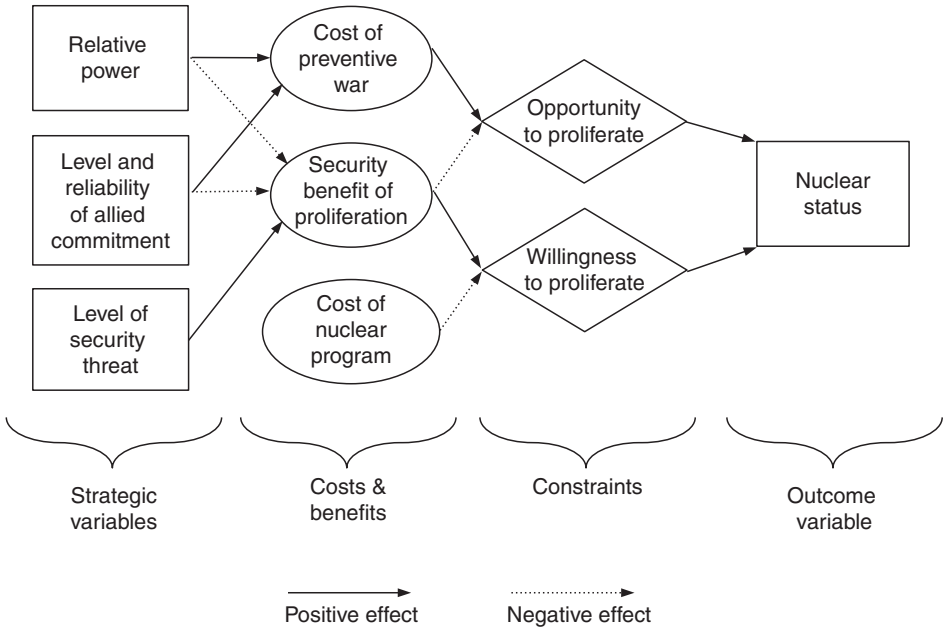
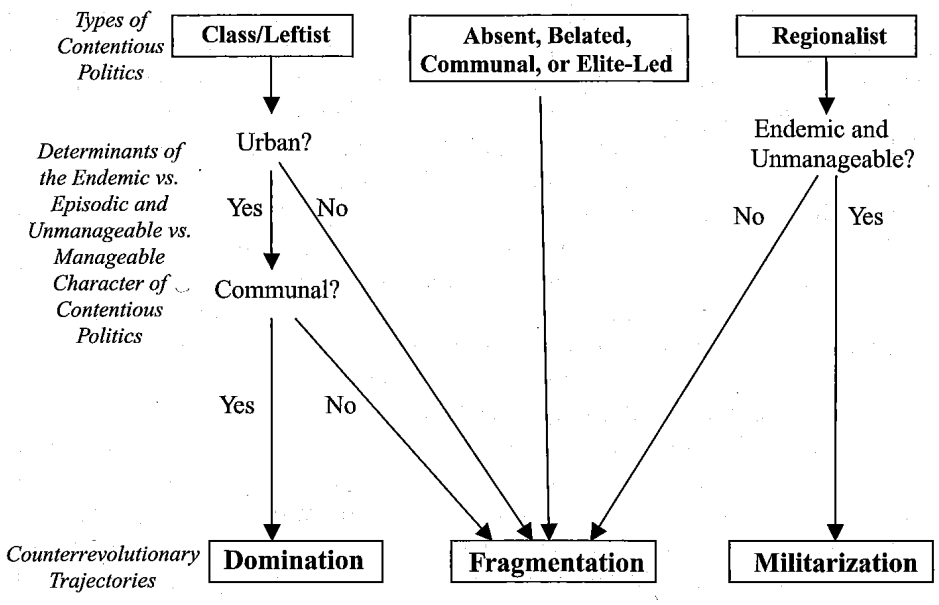


Figure 2.2. The Strategic Logic of Nuclear Proliferation

669. Can Slater's (2010) causal mechanisms be expressed in Boolean equations? If so what would they look like?

FIGURE 4. Contentious Politics and Counterrevolutionary Trajectories



670. Luebbert is a classic of comparative historical research. He basically has three dependent variables as illustrated by the title of his book *Liberalism, fascism, or social democracy*. Can you write a Boolean equation each of the pathways? Discuss the relationship between your three Boolean equations. Here is a quote to get you started.

> Societies that became social democracies or fascist dictatorships between the world wars were distinguished from Britain, France, and Switzerland by the failure of their prewar experiments with Lib-Labism. The failure of Lib-Labism was rooted in the failure of liberalism itself. Liberalism's failure derived from its inability to rally sufficient middle-class support. The critical question is why the middle classes were less supportive of liberal parties in most European countries than they were in Britain, France, and Switzerland. To put the question in a more tractable form, we will ask why liberal parties in most of Europe were less effective in establishing their dominance – that is, in surmounting the obstacles to an extended run of society-shaping political power in the years before 1914. . . . The lack of liberal dominance was crucial to the pattern of working-class mobilization and organization that followed, that is, for class formation. Workers could pursue their political and economic interests in only two ways: through an alliance with reforming liberals or through self-organization. The balance between the two routes workers followed was contingent on the availability of a politically effective liberal ally. The pattern of class formation in turn had essential consequences for the later course of interwar politics. (Luebbert 1991, 55)

671. Write the Boolean equation(s) for the causal mechanism figure below. See also his table 1 on page 14, which is a nonstandard 2×2 table.

> The main arguments can be summarized briefly. If a leading state has little reason to fear the takeover of a peripheral region by a local actor or an outside power, then it should not have a strong preference when

it comes to the type of order that exists there. In principle, both parity and primacy can prevent disruptive conflicts, the avoidance of which will be its chief objective in these circumstances. Therefore, a leading state should accommodate RRPs [rising regional powers] that can achieve either one because they will enhance local stability over the long run. Conversely, it should oppose RRPs that fall short of this threshold because their rise will contribute to persistent unrest. If a leading state is worried mainly that a local actor might dominate a peripheral region, however, then it will prefer parity to ensure that its own access to the area is not jeopardized. In this case, it should accommodate RRPs that are attempting to weaken local hegemons and oppose RRPs that are trying to gain control over their neighborhoods. Finally, if a leading state is more concerned about an outside power conquering a peripheral region, then it will prefer primacy instead because the strongest local actors are the best barriers to intervention. Thus, it should accommodate RRPs that fully overtake their rivals and oppose RRPs that fail to do so. (Montgomery 2016, 10)

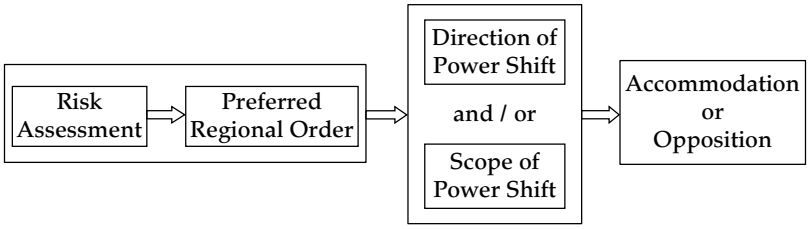


Figure 1. The causal argument

672. Poast provides an interesting possible Boolean model. Write a Boolean equation that expresses his theory.

The core of my argument is that the items at the heart of negotiations for an alliance treaty are also the key features of a war plan: "who" and "where" are articulated in the strategic component of a war plan, while "how" is captured in the operational component of a war plan. This suggests that one can view an alliance treaty negotiation as an effort in joint war planning. This conceptualization of alliance negotiations evokes two variables that likely determine whether the negotiations lead to a signed treaty: the compatibility of ideal war plans and the attractiveness of outside options. The first variable is the compatibility of ideal war plans. At the beginning of a negotiation, each participant reveals its ideal war plan, meaning the plan that reflects that state's preferred strategic and operational components. Ideal war plans are compatible if they articulate similar strategic and operational components. The strategic component of a war plan refers to the target of military force, which reflects a state's perception of threats. Hence, strategic compatibility means the participants in the negotiations have similar perceptions of possible threats. This lies at the heart of alliance treaty negotiations: without agreement on who to attack and where, it makes little sense to negotiate the details of military cooperation. The operational component of a plan refers to the general approach for addressing the identified threat(s), which is reflected by the state's military doctrine. A military doctrine can be either offensive, meaning the fight should be taken to the territory of the perceived threat, or defensive, meaning the objective is to stop the enemy's advance

> by fighting on home territory (either a state's own territory or the territory of an ally). Hence, operational compatibility means the states do not have contradictory military doctrines. The second variable is the attractiveness of outside options. An outside option is the policy each participant will pursue if the negotiation ends in nonagreement. Outside options include unilateral action, an alliance with another state, or buck-passing (i.e., leaving it to other states to attack or deter the threat). The attractiveness of an outside option is the extent to which the participant perceives the outside option as offering a benefit similar to that of an alliance that follows its ideal war plan. The more attractive the outside option, the less willing a participant will be to deviate from its ideal plan. Since attractiveness is a matter of perception, it is private information and is known only to that participant. Moreover, each participant has an incentive to misrepresent this attractiveness, in order to have the final treaty more closely reflect its ideal war plan. Of particular importance is the number of participants that perceive themselves as having an attractive outside option. For example, in a negotiation between two participants, if both perceive themselves as having an attractive outside option, neither will make concessions to secure an agreement. In contrast, if neither participant perceives itself as having an attractive outside option, the participants will not want the negotiation to end in nonagreement. In this case, they are more likely to make concessions to secure an agreement. (Poast 2019, 4–5)

Answer:

Poast basic theory seems to be: A OR (not-A and not-B) then Y (successful alliance agreement), A is compatible war plans and B is outside options He also explicitly claims not-A and B is sufficient for failure.

673. Grzymała-Busse (2007) is a core example in the causal mechanism chapter of *Multimethod research, causal mechanisms, and case studies: an integrated approach*. In my figure 2.4 of her theory "Robust competition" is constituted by three mechanisms (1) moderation, (2) anticipation, and (3) cooptation. Discuss how on might aggregate or combine these causal mechanisms (e.g., OR, AND, addition). Mikkelsen (2017) uses Grzymała-Busse (2007) a core example in his discussion of fuzzy logic case studies. Compare and discuss his version of the mechanism of competition, e.g., Figure 2 and how he uses fuzzy logic to combine the three mechanisms with the discussion in chapter 3 of *Multimethod research, causal mechanisms, and case studies: an integrated approach* of the same mechanism.

674. A very important methodological issue is set relationships between independent variables. In an original outline of the book a major part of a chapter was going to be devoted to the methodological issues that this raises. The main methodological concern arises when X_1 is a subset of X_2. For example, in the debate about the territorial peace versus democratic peace, it turns out that the the set of democratic dyads is a subset of the dyads at territorial peace: "This study compares the conflicting answers of the democratic peace and the territorial peace and examines the empirical record to see which is more accurate. It finds that almost all contiguous dyads settle their borders before they become joint democracies" (Owsiak and Vasquez 2016, 339).

What is the causal interpretation of the subset variable when all of the variation in X_1 is taken up in X_2? Discuss how this is different from multicollinearity, for example, the two might not be very correlated. See Owsiak 2020 for a nice explicit discussion of the subsetting issue and other examples.

What if X_1 is a perfect subset of X_2 and one introduces an interaction term, $X_1 * X_2$?

This is just with dichotomous variables, but one can find set theoretic relationships among continuous variables (defined as X_1 less than or equal to X_2 for all observation or vice versa).

675. It is not uncommon that options for case selection form nested subsets. What are the research design and causal inference issues in one like this where genocides are subset of one-sided violence, also one-sided violence occurrence might be almost a subset of civil war occurrence.

> Harff (2003) presents a global examination of genocide in countries that have experienced state failure internal war or regime collapse covering the period 1955–97. Our dataset is different in that it consists of all conflict actors during the period 1989–2004 that are actively involved in an armed conflict resulting in at least 25 battle-deaths in a year. (Eck and Hultman 2007, 242)

676. Give a fuzzy logic interpretation of the relationship between the two democracy dimensions in Coppedge's (2012) analysis.

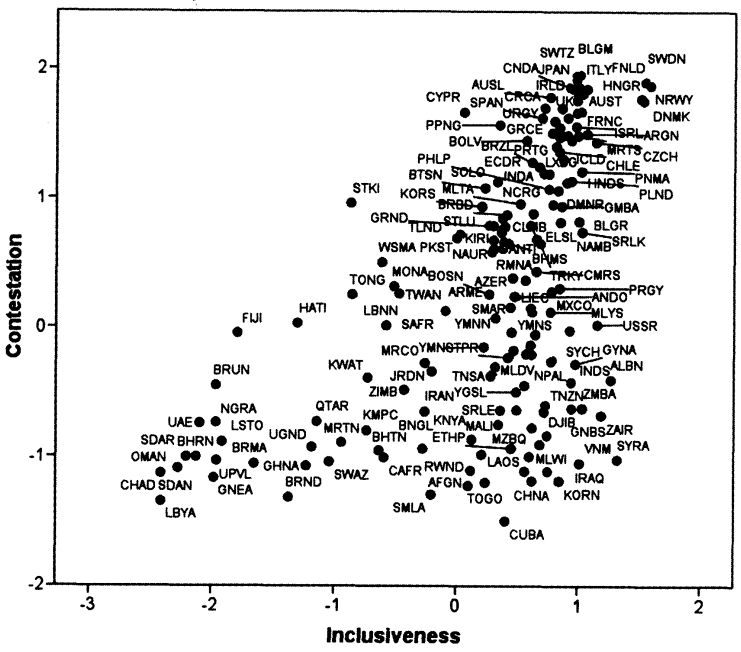


Figure 2.3 Distribution of Countries on Two Dimensions of Democracy, 2000. Both indicators were estimated by exploratory Principal Components Analysis of fourteen democracy indicators. The units of measurement are standard deviations. The estimation procedure and justifications for interpreting these as indicators of Dahl's two dimensions of polyarchy are in Coppedge, Alvarez, and Maldonado (2008).

Answer:

The Coppedge figure would be interpreted in a set theoretic fashion that inclusiveness is necessary for contestation.

677. Sometimes the concept may have no necessary conditions in terms of its definition and structure but there can be set theoretic relationships empirically between dimensions. For example, this can happen when one dimension is empirically a subset of another. Discuss the implications of this in terms of the polity data set and Bueno de Mesquita et al.'s discussion of it.

Answer:

Bueno de Mesquita et al. (2005) provides an example. "For instance, only 26 percent of the times when countries have open executive recruitment have they also institutionalized multiparty competition. But, if they have multiparty competition, then 97 percent of the time they also have open executive recruitment. Open recruitment appears to be necessary but not sufficient for multiparty competition. . . . Clearly, it is very hard to have multiparty competition without the other dimensions, but it is not difficult to achieve thresholds on other dimensions without having multiparty competition." (pp. 453–55)

678. Since people in general and social scientists in particular are not taught mathematical logic is not surprising that they have issues regarding questions of logic. Below Chambers describes an extremely famous and classic experiment – that while he discusses it in terms of confirmation bias – is really a question about logic. Give the two correct ways to test this hypotheses. Why is one much less common practice than the other?

> Since then, many studies have explored the basis of confirmation bias in a range of laboratory controlled situations. Perhaps the most famous of these is the ingenious Selection Task, which was also developed by Wason in 1968. The Selection Task works like this. Suppose I were to show you four cards on a table, labeled D, B, 3, and 7 (see figure 1.1). I tell you that if the card shows a letter on one side then it will have a number on the other side, and I provide you with a more specific rule (hypothesis) that may be true or false: "*If there is a D on one side of any card, then there is a 3 on its other side.*" Finally, I ask you to tell me which cards you would need to turn over in order to determine whether this rule is true or false. Leaving an informative card unturned or turning over an uninformative card (i.e., one that doesn't test the rule) would be considered an incorrect response. Before reading further, take a moment and ask yourself, which cards would you choose and which would you avoid? (Chambers 2017, 5)

Answer:

Here is Chambers:

> If you chose D and avoided B then you're in good company. Both responses are correct and are made by the majority of participants. Selecting D seeks to test the rule by confirming it, whereas avoiding B is correct because the flip side would be uninformative regardless of the outcome. Did you choose 3? Wason found that most participants did, even though 3 should be avoided. This is because if the flip side isn't a D, we learn nothing – the

> rule states that cards with D on one side are paired a 3 on the other, not that D is the only letter to be paired with a 3 (drawing such a conclusion would be a logical fallacy known as "affirming the consequent"). And even if the flip side is a D then the outcome would be consistent with the rule but wouldn't confirm it, for exactly the same reason. Finally, did you choose 7 or avoid it? Interestingly, Wason found that few participants selected 7, even though doing so is correct – in fact, it is just as correct as selecting D. If the flip side to 7 were discovered to be a D then the rule would be categorically disproven – a logical test of what's known as the "contrapositive." And herein lies the key result: the fact that most participants correctly select D but fail to select 7 provides evidence that people seek to test rules or hypotheses by confirming them rather than by falsifying them.

It is not surprising that few people thought of the contrapositive test given its nonintuitive nature. The sufficient condition test is straight forward.

679. One often sees scatterplots like the Tilly one in the study of the dimensions of democracy. What would be the QCA interpretation of these? What might that suggest about aggregation?

Answer:

Empirically the triangular scatterplot makes it look like civil liberties are sufficient for political rights. More substantively, they place a floor under the level of political rights.

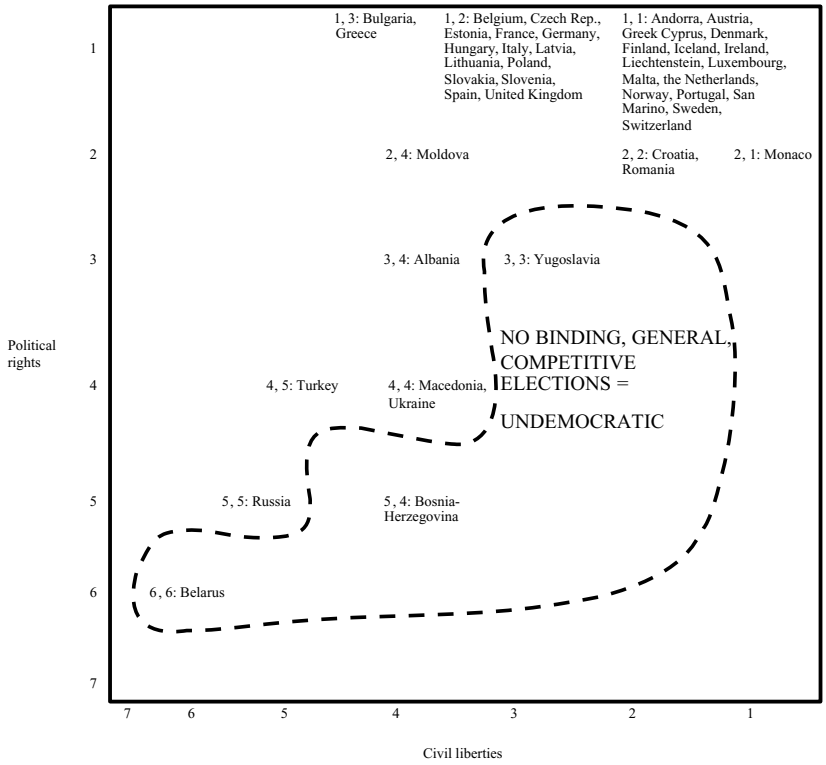


Figure 7.1: Freedom House ratings of European countries on political rights and civil liberties, 2001. *Source:* Compiled from Freedom House 2002.

680. Joshi and Wallensteen (2018) give five dimensions [(1) security provisions, (2) governance, (3) economic reconstruction, (4) Reconciliation and transitional justice, (5) Civil society] for quality peace: "This volume explores five dimensions that both theoretically and empirically necessary for quality peace in a post-acord society" (p. 5). Often scholars give a list of necessary conditions without stating if they are jointly sufficient. Would it be reasonable to also assume that they are jointly sufficient?

681. Discuss the claim below that it is better to look at the superset than the subset.

> For the sixty-two sovereign states included in the territorial contenders data set, Cunningham et al. identify ninety-seven rebel groups that controlled territory. The territorial contenders data set includes eighty-one of them (83.5 percent). Of the remaining sixteen rebels identified by Cunningham et al. as territorial, five are included among the ancillary materials of the territorial contender data set as near misses. The other eleven are instances where our coders could not identify any clear evidence of territorial control. Consultation with the source documents for the NonState Actor Dataset reveals no citations establishing territorial control for the rebel groups in question. We believe the codings in the territorial contender data set are valid, while accepting that there may be some occasional error in either our or the non-state actor (NSA) datasets. An 84 percent overlap is very high, nonetheless. *Absent an argument specific*

> *to a subset, there is considerable evidence in favor of analysis of the entire category of territorial contenders over analysis of any one subset*. (Lemke and Crabtree 2019, 19, emphasis is mine)

682. It is not uncommon that options for case selection form nested subsets. What are the research design and causal inference issues in one like this where genocides are subset of one-sided violence, also one-sided violence occurrence might be almost a subset of civil war occurrence.

> Harff (2003) presents a global examination of genocide in countries that have experienced state failure internal war or regime collapse covering the period 1955–97. Our dataset is different in that it consists of all conflict actors during the period 1989–2004 that are actively involved in an armed conflict resulting in at least 25 battle-deaths in a year. (Eck and Hultman 2007, 242)

683. Discuss the very strong Boolean claim that they have found a necessary and sufficient condition: "For the 500 USh denomination game, we find qualitative support for the claim that players exhibit in-group bias if and only if they are observed, although the results do not reach significance at conventional levels (row 3; columns 3 and 4). In the 100 USh denomination game, however, we find strong evidence that players discriminate in favor of co-ethnics *if and only if* they can be seen to be doing so (row 3; columns 1 and 2). Taken together, these findings offer strong support for this strategy selection mechanism as an important source of the variation we observe in public goods provision across ethnically homogeneous and heterogeneous settings." (Habyarimana et al. 2007, 721, emphasis is the authors')

684. Fuzzy-set analysis is often problematic when there are many observations near the origin (i.e., observations whose values on all variables approach zero). When observations are near the origin it is extremely easy for them to pass tests of sufficiency since X is likely to be less than Y. How does the Possibility Principle aid fuzzy-set methodology by dealing with these cases?

    Answer:

    This geometric interpretation of the Possibility Principle offers a solution to an important problem that arises in fuzzy-set analysis. The problem involves what to do with cases that are near the origin (i.e., cases near the (0,0,0) corner). As Ragin notes (2000, 250–51), when testing whether variables are causally sufficient for an outcome, observations with a zero for all the independent variables will always satisfy causal sufficiency and thus artificially inflate the number of cases where the theory works (this dilemma is the Raven Paradox). See also the exchange between Osherson and Smith (1981) and Zadeh (1982). The former is a classic in the cognitive psychology literature on categorization and Zadeh is the inventor of fuzzy-set logic.

685. Explain how the cube (Mahoney and Goertz 2004) provides a visual summary of disconfirming observations for necessary and/or sufficient condition hypothesis.

    Answer:

With a necessary condition hypothesis, theory disconfirmation occurs when the necessary condition is absent, but the outcome of interest is present. Thus, disconfirming cases are located in the top portion of the cube where only one necessary condition is present (i.e., the (1,0,1) and (0,1,1) corners).

With a sufficient condition hypothesis, theory disconfirmation occurs when the sufficient condition is present, but the outcome of interest is absent. Thus, disconfirming cases are located in the bottom portion of the cube where the two jointly sufficient causes are present (i.e., the (1,1,0) corner).

686. What would be the set of appropriate negative cases for the following Boolean theory?

$$Y = A + B + C + D \tag{3}$$

Does this contradict the Possibility Principle? Why does this example not fit the general kind of problem that typifies case selection in small-N studies? Hint: Which combinations of A, B, C, and D predict that Y does *not* occur? (Thanks to Charles Ragin for this exercise.)

Answer:

This is a question dealing with the logic of the Possibility Principle. The problem here is that if we followed the Possibility Principle for selecting negative cases, we would not have any negative cases at all since the Possibility Principle *excludes* all cases where A–D are absent (i.e., not-A AND not-B AND not-C AND not-D, this is a hint in the question). But this is exactly the set of cases we need to test the Boolean theory $Y = A + B + C + D$. This situation arises because A–D are individually sufficient for Y. However, models like $Y = A + B + C + D$ are extremely rare in the qualitative-comparative literature.

687. Use Licklider's findings (1995, table 3, p. 687) in conjunction with Harff (2003) to think about case selection in the analysis of civil wars and genocide/politicide.

Answer:

Negotiated settlement is sufficient condition for no genocide. Military victory is necessary condition for genocide.

688. In the literature on within-case causal inference researchers often talk about hoop tests and smoking gun tests in terms of "certainty" and "uniqueness": "The first dimension is called certitude or certainty and captures how likely it is to confirm a specific observable implication in process tracing. The second dimension is called uniqueness and asks whether an observable implication can be derived from a single or multiple hypotheses" (Rohlfing 2014 table 1). Rohlfing goes on to claim: "A hoop test is characterized by high certainty and no uniqueness." (Rohlfing 2014, 612) Describe a situation where a hoop test has a high degree of uniqueness.

Answer:

As the hoop test gets closer to being a sufficient condition, its uniqueness increases, i.e., when the hoop is very small.

689. Should one assume an interaction term or just a straight additive model, or a Boolean model for the data in the table below? The table below is a stripped-down version of their table 2 (p. 88) which gives the two core independent variables: (1) federal versus centralized state and (2) multiparty system versus two-party system. I use their primary dependent variable "macrocorporatism" (the other one they discuss is "sector coordination"). There is no statistical analysis in this article; the empirical analysis is carried out via four case studies, one from each cell in the table below.

Table 30: Macrocorporatism

Number of Parties	Level of State Centralization	
	Centralized	Federal
Multiparty	Denmark macrocorporatism=.72	Germany macrocorporatism=.50
Two party	Britain macrocorporatism=–.14	United States macrocorporatism=–.93

Source: Martin and Swank 2011, table 2.

Answer:

As in almost all two-variable models – i.e., X_1 AND $X_2 \rightarrow Y$ – the (1,1,1) cell is clear as is the (0,0,0) cell. The critical feature that will influence the decision between a QCA interpretation against an additive linear one is what happens with the cases where one independent variable is present and the other is absent, i.e., (1,0) or (0,1).

In set theoretic models the value of the (1,0) and (0,1) cells would be close to that of (0,0) because of the necessary condition relationship. The additive model suggests that when one independent variable is present and the other absent we should see intermediate results, somewhere between the extreme values on the dependent variable. Hence, a signal that the additive relationship is at work is when the dependent variable is significantly greater than the (0,0) but yet significantly less than the (1,1) dependent variable value. This is in fact what we find in table **??**. The two off-diagonal cells have values in between those in the on-diagonal cells.

It is pretty clear that a set theoretic interpretation of these data would not be appropriate because of the intermediate values of the dependent variable when only one independent variable is present. However, it is very difficult to make any strong conclusions regarding whether there should be an interaction term.

690. Discuss Pahre's claim that necessary conditions violate common statistical assumptions: "Third, necessary conditions violate the unit homogeneity assumption common in statistics, which states that 'if two units have the same value of the key explanatory variable, the expected value of the dependent

variable will be the same' (King et al., 1994: 116). Necessary conditions violate this condition because observations x, y and x, not-y are both consistent with the necessary condition x ← y; in other words, x may occur with or without y" (Pahre 2005, 131).

691. A very important methodological issue is set relationships between independent variables. In an original outline of the book a major part of a chapter was going to be devoted to the methodological issues that this raises. The main methodological concern arises when X_1 is a subset of X_2. For example, in the debate about the territorial peace versus democratic peace, it turns out that the the set of democratic dyads is a subset of the dyads at territorial peace: "This study compares the conflicting answers of the democratic peace and the territorial peace and examines the empirical record to see which is more accurate. It finds that almost all contiguous dyads settle their borders before they become joint democracies" (Owsiak and Vasquez 2016, 339).

    What is the causal interpretation of the subset variable when all of the variation in X_1 is taken up in X_2? Discuss how this is different from multicollinearity, for example, the two might not be very correlated. See Owsiak 2020 for a nice explicit discussion of the subsetting issue and other examples.

    What if X_1 is a perfect subset of X_2 and one introduces an interaction term, $X_1 * X_2$?

    This is just with dichotomous variables, but one can find set theoretic relationships among continuous variables (defined as X_1 less than or equal to X_2 for all observation or vice versa).

692. Walter's oft-cited book on the peaceful settlement of civil wars makes much use of necessary or sufficient conditions both in hypotheses and the data analysis. Discuss how she mixes the two. For example, "three sections follow. The first section tests whether third-party security guarantees and power-sharing pacts are necessary in combination to bring a peaceful solution to war. The data reveal that civil wars are significantly more likely to end in a successfully implemented settlement if both types of guarantees are present" (Walter 2003, 92).

693. Discuss the "light-switch" analogy as a way to think about process tracing and equifinality. The basic analogy is that the light-switch is on and the light is on (essentially the (1,1) cases). Discuss process tracing as following electricity through various junction boxes to the light bulb. Equifinality lies in the fact that there might be multiple switches that can possibly control the same light. Discuss what would be the analogy in case study work of "turning on or off" the switch (within cases or between cases). What about counterfactuals? (Thanks to Mike Desch for suggesting this analogy)

694. Copeland (2015) has at the core of his book an interaction; positive levels of economic interdepedence (often trade) interacted with expectations about future economic conditions. Discuss the situation where one of the interaction terms is basically a subset of the other. For example, what case studies we be appropriate? Is the superset variable (positive economic independence) more like a scope condition?

> To determine whether the liberal prediction or realist prediction will prevail, we must introduce an additional causal variable – namely, a state's expectations of the future trade and investment environment. When a dependent state has positive expectations about this future environment, it is more likely to see all the benefits of continuing the current peace and all the opportunity costs of turning to war. Economic interdependence would then be a force for peace. Yet if a dependent state has negative expectations about the future economic environment – seeing itself being cut off from access to foreign trade and investment, or believing that other states will soon cut it off – then the realist logic will kick in. Such a state will tend to believe that without access to the vital raw materials, investments, and export markets needed for its economic health, its economy will start to fall relative to other less vulnerable actors. If this economic decline is anticipated to be severe, the leaders of the dependent state will begin to view war as the rational lesser of two evils – that is, as better than allowing their state to fall to a point where rising states can attack it later or coerce it into submission. (Copeland 2015, 2; basic hypothesis of the book)

695. Levy and Thompson (2005) discuss the perhaps most classic hypotheses in international security: "threats of hegemony generate great-power balancing coalitions." They define "balancing coalitions" as an forming an alliance. The sufficient condition hypothesis: "If threat of hegemony then great powers form balancing coalitions." Make the data in the table below such that they would pass the QCA criterion for sufficient condition hypotheses. Then change the data so that the data (1) are a trivial sufficient condition hypothesis, (2) nontrivial sufficient condition hypothesis, and (3) nontrivial but with the relationship in the wrong direction.

Table 2
CAPABILITY CONCENTRATION AND ALLIANCE RESPONSE AGAINST THE LEADING POWER, 1495–1999

	Weaker relative position (less than 33%)	Stronger relative position (33% or greater)
No alliance response	151 (.702)	104 (.452)
Alliance response	64 (.298)	126 (.548)
N = 445	215 (1.000)	230 (1.000)

$\chi^2 = 28.42$; p = .000

Note: The years in which the state exceeded 33 percent of capability shares are indicated in parentheses in the "leading land power" column.

Answer:

A percentage of 80 percent or higher in the alliance formation column, would pass the QCA bar which is usually around 70–80 percent. For a trivial sufficient condition, make the percentage in the left-column the same as in the right. For nontrivial making it significantly lower, e.g., 50 percent. For nontrivial in the wrong direction make the percentage in the left-hand column around 95 percent.

696. Clarke gives an excellent discussion of the logic and methodology, in particular the common logical error of affirming the antecedent. Explain how the Avoid Overdetermination Guideline tries to deal with this by eliminating "other mortal creatures."

> The invalid argument is a logical fallacy known as "affirming the consequent." Consider the following invalid deduction: If x is human, then x is mortal [therefore] If x is mortal then x is human. Obviously, not all mortal creatures are humans. Despite the fact that x is mortal, we cannot conclude that x is human. . . . The problem with affirming the consequent is that there are numerous conditions that imply that a thing is mortal, and being human is simply one of them. In the example given above, for instance, x could be a cat, a dog, or even a dandelion, and observing that "x is mortal" does not give us warrant to conclude that any particular one is true. (Clarke 2007, 888)

Answer:

The Avoid Overdetermination Guideline selects cases where alternative "mortals" variables are zero. If one had a complete list of mortal beings and applied the Avoid Overdetermination Guideline ($Z = 0$) then the only possible case selection options left are humans.

697. One reason why qualitative methods (based on logic and set theory) are different from statistical methods is because most statistics use linear algebra while set theory is based on Boolean algebra. One difference between Boolean and linear algebra is that Boolean algebra requires a second distributive law: $a + (b * c) = (a + b) * (a + c)$. Explain why this does not hold for linear algebra. Explain via Venn diagrams why this works for sets (see Hohn, F. E. (1966). *Applied Boolean Algebra: An Elementary Introduction* for an introduction to Boolean algebra). Thanks to Alrik Thiem for this exercise.

Answer:

The second distributive law can be neatly visualized for sets with two Venn diagrams of three overlapping circles each. The law says . The left hand-side of the equation marks the whole of "a" plus the intersection of "b" and "c," and the right hand-side will mark exactly the same region.

698. Since students in the social sciences are not taught basic mathematical logic it is easy to make errors. Explain the logical error in the following descriptions of the democratic peace:

> Immanual Kants argument that democratic institutions . . . are a necessary condition for peace has been empirically substantiated. (Risse-Kappen 1996, 366)
>
> The recent flurry of studies of the theory of the "democratic peace" follows upon Kant's argument that a necessary condition for peace between states is constitutional republics. (Holsti 1996, 180)

Answer:

The authors have confused a necessary condition with a sufficient condition. The democratic peace says that joint democracy is sufficient for peace.

699. It is not too hard to make mistakes when manipulating necessary and sufficient conditions. Schweller (1992) manages with the hypotheses below to be both redundant and contradictory. Explain.

> 1. A power transition involving a declining democratic leader is both a necessary and sufficient condition for the absence of preventive war. 1a. When a declining democratic leader confronts a rising democratic challenger, accommodation results. 1b. When a declining democratic leader confronts a rising nondemocratic challenger, the leader tries to forma a defensive alliance system to counterbalance the threat. 2. A power transition involving a declining nondemocratic state is a necessary but not sufficient condition for a preventive war, regardless of the regime type of the challenger. (Schweller 1992, 248–49)

Answer:

Hypothesis 1 stipulates that, given a power transition and a declining leader, (a) the leader's democracy is necessary for the absence of war and (b) the leader's democracy is sufficient for the absence of war. Hypothesis 2 stipulates that the leader's nondemocracy is necessary for war. Since there are only two kinds of leaders, democratic and nondemocratic, and since necessary and sufficient conditions are fungible via the contrapositive operation, stating that the leader's democracy is sufficient for the absence of war is logically equivalent to stating that the leader's nondemocracy is necessary for the presence of war: if $P(Y|X) = 0$, then $P(\neg Y|\neg X) = 1$. Therefore, the assertion in Hypothesis 2 that nondemocracy is necessary for war is already implied by Hypothesis 1. On the other hand, Hypothesis 2 specifically states that, given a power transition and a declining leader, the leader's nondemocracy is not sufficient for preventive war. Translating this statement into necessary-condition terms yields the statement, "The leader's democracy is not necessary for the absence of preventive war." Hypothesis 1 makes the claim that democracy is necessary for the absence of war. Hypothesis 2, therefore, merely restates one half of Hypothesis 1 and contradicts the other.

700. Express Przeworski et al.'s (2000) main conclusion regarding GDP/capita and democracy using the logic of necessary and sufficient conditions with (1) two independent variables, (2) as a sufficient condition, (3) with time coefficients, and (4) dependent variable involves authoritarianism.

Answer:

Minimum level of GNP/capita (time t) AND Democracy (time t) are jointly sufficient for no transition to authoritarianism (from time t onwards).

Notice how this orients theory toward to question of transitions to nondemocracy instead of transitions to democracy. What theory is there about transitions to authoritarianism?

701. Researchers often use set theoretic or necessary (or sufficient) condition language. For example, "all sociologists are Democrats" means sociologists are a *subset* of all democrats; which also means being a Democrat is a necessary condition for being a sociologist. Such language can be common but it is often not recognized as such. Take Lipset's 1993 American Sociological Association

Presidential address and find all examples of the use of logic (e.g., necessary condition) or set theory in hypotheses or discussions of empirical results (e.g., all sociologists are Democrats). Express these descriptive statements in terms of hypotheses about necessary or sufficient conditions.

Answer:

There are quite a few. For example, (1) "every country with a population of at least 1 million that has emerged from colonial rule and has had a continuous democratic experience is a former British colony" (Lipset 1993, 5, citing Weiner 1987), (2) "there has been no case of political democracy that has *not* been a market economy" (Lipset 1993, 5, citing Berger 1992). Others can easily be found.

Here is how the logic works using a descriptive statement from Dahl (1971): "all highest-level [developed] countries are polyarchies" (Cited by Diamond 1992, 97).

All highly developed (*Wealth*) countries are polyarchies (*Democracy*).	(1)
All W are D.	(2)
D is necessary for W.	(3)
$\neg D$ is sufficient for $\neg W$ ("$\neg$" means "not").	(4)

Thus (3) or (4) are equivalent logical statements the express Dahl's empirical findings.

702. The "democratic peace" idea is that democracies never fight wars with each other. This can be expressed as a either a necessary condition hypothesis or a sufficient condition hypothesis. Give both hypotheses.

703. It is not uncommon to confuse necessary condition counterfactuals with sufficient condition ones. See Fenoaltea (1973) for a beautiful discussion of this in the context of economic history. Fischer's classic book (1970) also deals with this. Discuss Fisher's examples.

> One common form of the reductive fallacy is the confusion of necessary with sufficient cause – the confusion of a causal component without which an effect will not occur, with all the other causal components which are required to make it occur. This sort of error appears in causal explanations which are constructed like a single chain and stretched taut across a vast chasm of complexity. The classic example is the legendary battle that was lost for the want of a horseshoe nail; for the the want of a nail the shoe was lost, for the want of a shoe the horse was lost, for the want of a horse the rider was lost, for the want of a rider the message was lost, for the want of the message the regiment was lost, and for the want of a regiment the battle was lost. (Fischer 1970, 172; he then goes on to use the classic example of the battle of Antietam for which the accident loss of confederate orders to the North was a key factor in the North's victory)

References

Abbott, K., et al. 2000. The concept of legalization. *International Organization* 54:401–19.

Abebe, D., and T. Ginsburg. 2019. The dejudicialization of international politics? *International Studies Quarterly* 63:521–30.

Abramson, S., and D. Carter. 2021. Systemic instability and the emergence of border disputes. *International Organization* 75:103–46.

Acharya, A., et al. 2016. Explaining causal findings without bias: detecting and assessing direct effects. *American Political Science Review* 110:512–29.

Achen, C., and D. Snidal. 1989. Rational deterrence theory and comparative case studies. *World Politics* 41:143–69.

Agnafors, M. 2013. Quality of government: toward a more complex definition. *American Political Science Review* 107:433–45.

Ahmed, F. 2020. *The perils of international capital.* Princeton: Princeton University Press.

Alacer, J., and P. Ingram. 2013. Spanning the institutional abyss: the intergovernmental network and the governance of foreign direct investment. *American Journal of Sociology* 118:1055–98.

Alcañiz, I. 2012. Democratization and multilateral security. *World Politics* 64:306–40.

Alesina, A., and E. Spolaore. 1997. On the size and number of nations. *Quarterly Journal of Economics* 112:1027–56.

Alkire, S., et al. 2015. *Multidimensional poverty measurement and analysis.* Oxford: Oxford University Press.

Alter, K. 2014. *The new terrain of international law.* Princeton: Princeton University Press.

Alter, K., et al. (eds.). 2018. *International court authority.* Oxford: Oxford University Press.

Altman, D. 2020. The evolution of territorial conquest after 1945 and the limits of the territorial integrity norm. *International Organization* 74:490–522.

Alvarez, M., et al. 1996. Classifying political regimes. *Studies in Comparative International Development* 31:3–36.

Amenta, E. 2003. What we know about the development of social policy: comparative and historical research in comparative and historical perspective. In J. Mahoney and D. Rueschemeyer (eds.) *Comparative historical analysis in the social sciences.* Cambridge: Cambridge University Press.

American Psychiatric Association. 1994. *Diagnostic and statistical manual of mental disorders (DSM-IV),* 4th edition. Washington D.C.: American Psychiatric Association.

American Psychiatric Association. 2013. *DSM-5.* Washington, DC: American Psychiatric Publishing.

Amorim Neto, O., and G. Cox. 1997. Electoral institutions, cleavage structures, and the number of parties. *American Journal of Political Science* 41:149–74.

Andersen, D. 2024. Impartial administration and peaceful agrarian reform: the foundations for democracy in scandinavia. *American Political Science Review* 118:54–68.

Anselmi, M. 2018. *Populism: an introduction.* London: Routledge.

Ansolabehere, S., et al. 2005. Voting weights and formateur advantages in the formation of coalition governments. *American Journal of Political Science* 49:550–63.

Arjona, A. 2016. *Rebelocracy: social order in the Colombian civil war.* Cambridge: Cambridge University Press.

Armitage, D. 2017. *Civil wars: a history in ideas.* New York: Alfred Knopf.

Aronson, J., and M. Hauben. 2006. Drug safety: anecdotes that provide definitive evidence. *British Medical Journal* 333:1267–69.

Aviv, R. 2024. CONVICTION: Lucy Letby, the most notorious nurse in Britain, was found guilty of killing seven babies. Did she do it? *The New Yorker* VOLUME:34–51.

Bühlmann, M., et al. 2012. The democracy barometer: a new instrument to measure the quality of democracy and its potential for comparative research. *European Political Science* 11:519–36.

Bagozzi, B., et al. 2015. Modeling two types of peace: the zero-inflated ordered probit (ZiOP) model in conflict research. *Journal of Conflict Resolution* 59:728–52.
Barany, Z. 2002. Ethnic mobilization with prerequisites: the east European gypsies. *World Politics* 54:277–307.
Barany, Z. 2016. *How armies respond to revolutions and why.* Princeton: Princeton University Press.
Barnhart, J., et al. 2020. The suffragist peace. *International Organization* 74:633–70.
Baron, H., et al. 2024. An events-based approach to understanding democratic erosion. *PS: Political Science & Politics* 57:208–15.
Bassan-Nygate, L., et al. 2025. he Generalizability of IR experiments beyond the United States. *American Political Science Review* 119:1649–69.
Bates, G., et al. 2020. Accountability by numbers: a new global transitional justice dataset (1946–2016). *Perspectives on Politics* 18:161–84.
Baum, M. A., and P. B. K. Potter. 2015. *War and democratic constraint: how the public influences foreign policy.* Princeton: Princeton University Press.
Baumol, W. 1986. *Superfairness.* Cambridge: MIT Press.
Bayer, R. 2005. Reconciliation, democracy, and external threats. Paper presented at the annual meetings of the First Global International Studies Conference, Bilgi University, Istanbul.
Beck, N., and S. Jackman. 1998. Beyond linearity by default: generalized additive models. *American Journal of Political Science* 42:596–627.
Beck, T., et al. 2001. New tools and new tests in comparative political economy: the database of political institutions. *World Bank Economic Review* 15:165–76.
Becker, M., et al. 2024. Replicating the resource curse: a qualitative replication of Ross 2004. *International Studies Quarterly* 68.
Beckley, M. 2018. The power of nations: measuring what matters. *World Politics* 43:7–44.
Beckley, M. 2023. The peril of peaking powers: economic slowdowns and implications for China's next decade. *International Security* 48:7–46.
Beckley, M. 2023. The peril of peaking powers: economic slowdowns and implications for chinaís next decade. *International Security* 48:7–46.
Beliakova, P. 2021. Erosion of civilian control in democracies: a comprehensive framework for comparative analysis. *Comparative Political Studies* 54:1393–423.
Belkin, A., and E. Schofer. 2003. Toward a structural understanding of coup risk. *Journal of Conflict Resolution* 47:594–620.
Belkin, A., and E. Schofer. 2005. Coup risk, counterbalancing, and international conflict. *Security Studies* 14:140–77.
Bell, M. S. 2021. *Nuclear reactions: how nuclear-armed states behave.* Ithaca: Cornell University Press.
Bendor, J., and J. Shapiro. 2019. Historical contingencies in the evolution of states and their militaries. *World Politics* 71:126–60.
Bengtsson, R. 2000. The cognitive dimension of stable peace. In A. Kacowicz et al. (eds.) *Stable peace among nations.* New York: Rowman & Littlefield.
Bennett, A. 1999. *Condemned to repetition? The rise, fall, and reprise of Soviet-Russian military interventionism, 1973–1996.* Cambridge: MIT Press.
Bennett, A., and J. Checkel. 2015. Process tracing: from philosophical roots to best practices. In A. Bennett and J. Checkel (eds.) *Process tracing in the social sciences: from metaphor to analytic tool.* Cambridge: Cambridge University Press.
Bennett, D. 2006. Towards a continuous specification of the democracy-autocracy connection. *International Studies Quarterly* 50:513–37.
Bennett, D., and A. Stam. 2004. *The behavioral origins of war.* Ann Arbor: University of Michigan Press.
Berger, P. 1992. The uncertain triumph of democratic capitalism. *Journal of Democracy* 3:7–17.
Berman, C., et al. 2024. From victims to dissidents: legacies of violence and popular mobilization in iraq (2003–2018). *American Political Science Review* 118:213–34.

Berman, E., J. H. Felter, and Shapiro; J. N. 2018. *Small wars, big data: the information revolution in modern conflict*. Princeton: Princeton University Press.
Berman, E., J. H. Felter, and Shapiro; J. N. 2018. *Small wars, big data: the information revolution in modern conflict*. Princeton: Princeton University Press.
Berry, C. 1994. *The idea of luxury: a conceptual and historical investigation*. Cambridge: Cambridge University Press.
Bertrand, J. 2021. *Democracy and nationalism in southeast Asia: from secessionist mobilization to conflict resolution*. Cambridge: Cambridge University Press.
Birch, S. 2012. *Electoral malpractice*. Oxford: Oxford University Press.
Birch, S. 2020. *Electoral violence, corruption and political order*. Princeton: Princeton University Press.
Birnir, J. 2015. Socially relevant ethnic groups, ethnic structure and AMAR. *Journal of Peace Research* 52:110–15.
Birnir, J., et al. 2016. Political transition constructed: ethnicity, religion and conflict. Manuscript. University of Maryland.
Birnir, J., et al. 2018. Introducing the AMAR (All Minorities at Risk) Data. *Journal of Conflict Resolution* 62:203–26.
Blackbourn, D., and G. Eley. 1984. *The peculiarities of German history: bourgeois politics and society in nineteenth century Germany*. Oxford: Oxford University Press.
Blainey, G. 1988 (1973). *The causes of war,* 3rd edition. New York: Free Press.
Blake, C., and J. Adolino. 2001. The enactment of national health insurance: a Boolean analysis of twenty advanced industrial countries. *Journal of Health Politics, Policy and Law* 26:679–708.
Blanton, S. 2005. Foreign policy in transition? Human rights, democracy, and U.S. arms exports. *International Studies Quarterly* 49:647–67.
Bodea, C., et al. 2017. Do civil wars, coups and riots have the same structural determinants? *International Interactions* 43:537–61.
Bogaards, M. 2007. Measuring democracy through election outcomes: a critique with African data. *Comparative Political Studies* 40:1211–37.
Boix, C., et al. 2012. A complete data set of political regimes, 1800–2007. *Comparative Political Studies* 46:1523–54.
Bonell, C., et al. 2022. Methodological reflections on using qualitative research to explore the causal mechanisms of complex health interventions. *Evaluation* 28:166–81.
Bonoli, G. 2007. Time matters: postindustrialization, new social risks, and welfare state adaption in advanced industrial democracies. *Comparative Political Studies* 40:495–520.
Botero, J., and A. Ponce. 2010. Measuring the rule of law. Manuscript. World Justice Project.
Boulanger Martel, S., et al. 2024. Peace with adjectives: conceptual fragmentation or conceptual innovation? *International Studies Review* 16.
Boulding, K. 1978. *Stable peace*. Austin: University of Texas Press.
Bove, V., and M. Rivera. 2015. Elite co-optation, repression, and coups in autocracies. *International Interactions* 41:453–79.
Bove, V., and R. Nisticó. 2014. Coups détat and defense spending: a counterfactual analysis. *Public Choice* 161:321–44.
Brambor, T., W. Clark, and M. Golder. 2006. Understanding interaction models: improving empirical analyses. *Political Analysis* 14:63–82.
Brancati, D. 2006. Decentralization: feuling the fire or dampening the flames of ethnic conflict and secessionism. *International Organization* 60:651–85.
Braumoeller, B. 2003. Causal complexity and the study of politics. *Political Analysis* 11:209–33.
Braumoeller, B. 2004. Hypothesis testing and multiplicative interaction terms. *International Organization* 58:807–20.
Braumoeller, B., and G. Goertz. 2002. Watching your posterior: comment on Seawright. *Political Analysis* 10:198–203.

Braybrooke, D. 1987. *Meeting needs.* Princeton: Princeton University Press.
Bromley, P., et al. 2024. World Education Reform Database (WERD): a global dataset on education reforms, version 3. Manuscript. Stanford University.
Brookes, M. 2018. *The new politics of transnational labor: why some alliances succeed.* Ithaca: Cornell University Press.
Brosché, J., et al. 2023. Conceptualizing civil war complexity. *Security Studies* 32:137–65.
Brown, R. 2010. Measuring delegation. *Review of International Organizations* 5:141–75.
Budge, I., et al. 2001. *Mapping policy preferences: estimates for parties, electors and governments 1945–1998.* Oxford: Oxford University Press.
Bueno de Mesquita, B., et al. 2005. Thinking inside the box: a closer look at democracy and human rights. *International Studies Quarterly* 49:439–57.
Bueno de Mesquita, E., and A. Fowler. 2021. *Thinking clearly with data: a guide to quantitative reasoning and analysis.* Princeton: Princeton University Press.
Bulutgil, H. 2021. *The origins of secular institutions ideas, timing, and organization.* Oxford: Oxford University Press.
Bunce, V., and S. Wolchik. 2011. *Defeating authoritarian leaders in post-communist countries.* Cambridge: Cambridge University Press.
Burawoy, M. 1989. Two methods in search of science: Skocpol versus Trotsky. *Theory and Society* 18:759–805.
Burgerman, S. 2001. *Moral victories: how activists provoke multilateral action.* Ithaca: Cornell University Press.
Burke, M., et al. 2015. Climate and conflict. *Annual Review of Economics* 7:577–617.
Bush, S. 2011. International politics and the spread of quotas for women in legislatures. *International Organization* 65:103–37.
Busse, M., et al. 2007. Political risk, institutions and foreign direct investment. *European Journal of Political Economy* 23:397–415.
Busse, M., and C. Hefeker. 2007. Political risk, institutions and foreign direct investment. *European Journal of Political Economy* 23:397–415.
Butcher, C., et al. 2020. Introducing the targeted mass killing data set for the study and forecasting of mass atrocities. *Journal of Conflict Resolution* 64:1524–47.
Byman, D. 2021. White supremacy, terrorism, and the failure of reconstruction in the united states. *International Security* 46:53–103.
CETIS. 2007. GTD2 (1998–2004): global terrorism database, draft 1.0. Manuscript. University of Maryland.
CICM (Center for International Development and Conflict Management). 2016. All Minorities at Risk (AMAR) Phase I Codebook, August 2016. Manuscript.
Call, C. 2010. Beyond the 'failed state': toward conceptual alternatives. *European Journal of International Relations* 17:303–26.
Cameron, D. 1984. Social democracy, corporatism, labor quiescence, and the representation of economic interests in advanced capitalist societies. In J. Goldthorpe (ed.) *Order and conflict in contemporary capitalism.* Oxford: Oxford University Press.
Cammett, M. 2022. Positive deviance cases: their value for development research, policy, and practice. In J. Widner and et al. (eds.) *The case for case studies: methods and applications in international development.* Cambridge: Cambridge University Press.
Campbell, D., et al. 2021. *Secular surge: a new fault line in American politics.* Cambridge: Cambridge University Press.
Campbell, D., et al. 20210919. Harvard's atheist chaplain: it's another sign of America's growing secularism. *USA Today.*
Campbell, S., et al. 2017. An ontology of peace: landscapes of conflict and cooperation with application to Colombia. *International Studies Review* 19:92–113.
Campos, N., and C. Frederico. 2026. A new measure of affective polarization. *American Political Science Review* 120:160–78.
Caplan, R. 2019. *Measuring peace: principles, practices, and politics.* Oxford: Oxford University Press.

Capoccia, G. 2002. Anti-system parties: a conceptual reassessment. *Journal of Theoretical Politics* 14:9–34.
Capoccia, G. 2005. *Defending democracy: reactions to extremism in interwar Europe.* Oxford: Oxford University Press.
Caporaso, J. 1996. The European Union and forms of state: Westphalian, regulatory, or post-modern. *Journal of Common Market Studies* 34:31–52.
Carey, S. 2007. European aid: human rights versus bureaucratic inertia. *Journal of Peace Research* 44:447–64.
Carey, S., and N. Mitchell. 2017. Progovernment militias. *Annual Review of Political Science* 20:127–47.
Carnegie, A., and A. Carson. 2018. The spotlights harsh glare: rethinking publicity and international order. *International Organization* 72:627–57.
Carnegie, A., and A. Carson. 2019. The disclosure dilemma: nuclear intelligence and international organizations. *American Journal of Political Science* 63:269–85.
Carothers, C. 2022. *Corruption control in authoritarian regimes: lessons from East Asia.* Cambridge: Cambridge University Press.
Carpenter, R. 2007. Setting the advocacy agenda: theorizing issue emergence and nonemergence in transnational advocacy networks. *International Studies Quarterly* 51:99–120.
Cartwright, N. 2022. How to learn about causes in the single case. In J. Widner and et al. (eds.) *The case for case studies: methods and applications in international development.* Cambridge: Cambridge University Press.
Casal Bértoa, F. 2017. Its been mostly about money! A multi-method research approach to the sources of institutionalization. *Sociological Methods & Research* 46:683–714.
Cebotari, V., and M. Vink. 2013. A configurational analysis of ethnic protest in Europe. *International Journal of Comparative Sociology* 54:298–324.
Cederman, L.-E., et al. 2010. Why do ethnic groups rebel? New data and analysis. *World Politics* 61:87–119.
Celis, K. 2005. Reconciling theory and empirical research: methodological reflections on "women MP's representing women's interests". Paper presented at the annual meetings of the American Political Science Association.
Chambers, C. 2017. *The seven deadly sins of psychology:a manifesto for reforming the culture of scientific practice.* Princeton: Princeton University Press.
Chambru, C., and P. Maneuvrier-Hervieu. 2024. Introducing HiSCoD: a new gateway for the study of historical social conflict. *American Political Science Review* 118:1084–91.
Chan, S., and M. Mason. 1994. Foreign direct investment and host country conditions: looking from the other side now. *International Interactions* 17:215–32.
Chandra, K. 2008. What is ethnic identity and does it matter. *Annual Review of Political Science* 9:397–424.
Chandra, K., and S. Wilkinson. 2008. Measuring the effect of "ethnicity". *Comparative Political Studies* 41:515–63.
Checkel, J. 1993. Ideas, institutions, and the Gorbachev foreign policy revolution. *World Politics* 45:271–300.
Cheibub, J., J. Gandhi, and J. Vreeland. 2010. Democracy and dictatorship revisited. *Public Choice* 143:67–101.
Cheibub, J., et al. 2013. Beyond presidentialism and parliamentarism. *British Journal of Political Science* 44:514–44.
Chen, S., and M. Ravallion. 2001. How did the world's poorest fare in the 1990s. *Review of Income and Wealth* 47:283–300.
Chenoweth, E., and M. Stephan. 2011. *Why civil resistance works the strategic logic of nonviolent conflict.* New York: Columbia University Press.
Chin, H., and M. Quddus. 2003. Modeling count data with excess zeroes: an empirical application to traffic accidents. *Sociological Methods & Research* 32:90–116.
Chirot, D. 1994. *Modern tyrants.* New York: Free Press.
Chronic Poverty Research Centre. 2009. The chronic poverty report 2008-09. Manuscript.

Chubb, A. 2020. PRC Assertiveness in the South China Sea: measuring continuity and change, 1970–2015. *International Security* 45:79–121.
Cingranelli, D., and D. Richards. 2010. The Cingranelli and Richards (CIRI) Human Rights Data Project. *Human Rights Quarterly* 32:401–24.
Clarke, K. 2002. The reverend and the ravens: comment on Seawright. *Political Analysis* 10:194–97.
Clarke, K. 2007. The necessity of being comparative: theory confirmation in quantitative political science. *Comparative Political Studies* 40:886–908.
Cohen, D. K. 2016. *Rape during civil war.* Ithaca: Cornell University Press.
Cohen, Y., and F. Pavoncello. 1987. Corporatism and pluralism: a critique of Schmitter's typology. *British Journal of Political Science* 17:117–22.
Colaresi, M., et al. 2007. *Strategic rivalries in world politics.* Cambridge: Cambridge University Press.
Colaresi, M., and W. Thompson. 2002. Strategic rivalries, protracted conflict, and crisis escalation. *Journal of Peace Research* 39:263–87.
Colgan, J. 2013. *Petro-aggression: when oil causes war.* Cambridge: Cambridge University Press.
Colgan, J., and J. Weeks. 2015. Revolution, personalist dictatorships, and international conflict. *International Organization* 69:163–94.
Collier, D., et al. 2008. Typologies: forming concepts and creating categorical variables. In J. Box-Steffensmier, H. Brady, and D. Collier (eds.) *The Oxford handbook of political methodology.* Oxford: Oxford University Press.
Collier, D., and J. Mahoney. 1996. Insights and pitfalls: selection bias in qualitative research. *World Politics* 49:56–91.
Collier, D., and R. Adcock. 1999. Democracy and dichotomies. *Annual Review of Political Science* 2:537–65.
Collier, P., and A. Hoeffler. 2004. Greed and grievance in civil war. *Oxford Economic Papers* 56:563–95.
Collier, P., and N. Sambanis (eds.). 2005. *Understanding civil war: evidence and analysis, volume I: Africa; volume II: Europe, Central Asia, and other regions.* Washington, DC: World Bank.
Collins, C., et al. 2024. "Which cases do I need?" Constructing cases and observations in qualitative research. *Annual Review of Sociology* 50:21–40.
Copeland, D. 2015. *Economic interdependence and war.* Princeton: Princeton University Press.
Copeland, D. 2024. *A world safe for commerce: American foreign policy from the revolution to the rise of China.* Princeton: Princeton University Press.
Copeland, D. C. 2015. *Economic interdependence and war.* Princeton: Princeton University Press.
Copestake, J., G. Goertz, and S. Haggard. 2020. The Veil of Ignorance Process Tracing (VoiPT) methodology. *Qualitative & Multi-Method Research* 18:19–54.
Coppedge, M. 2012. *Democratization and research methods.* Cambridge: Cambridge University Press.
Coppedge, M., et al. 2020. *Varieties of democracy: measuring two centuries of political change.* Cambridge: Cambridge University Press.
Cornell, A., et al. 2020. *Democratic stability in an age of crisis: reassessing the interwar period.* Oxford: Oxford University Press.
Corntassel, J. 2003. Who is indigenous? "Peoplehood" and ethnonationalist approaches to rearticulating indigenous identity. *Nationalism and Ethnic Politics* 9:75–100.
Cottiero, C., and S. Haggard. 2023. Stabilizing authoritarian rule: the role of international institutions. *International Studies Quarterly* 67:1–15.
Crenshaw, M. 1972. The concept of revolutionary terrorism. *Journal of Conflict Resolution* 16:383–96.
Crenshaw, M. 1981. The causes of terrorism. *Comparative Politics* 13:379–99.

Cunningham, D. 2006. Veto players and civil war duration. *American Journal of Political Science* 50:875–92.
Dafoe, A., et al. 2013. The democratic peace: weighing the evidence and cautious inference. *International Studies Quarterly* 56:201–14.
Dahl, R. 1971. *Polyarchy: participation and opposition.* New Haven: Yale University Press.
Daly, M., and C. Saraceno. 2002. Social exclusion and gender relations. In B. Hobson, J. Lewis, and B. Siim (eds.) *Contested concepts in gender and social politics.* Cheltenham: Edward Elgar.
Dasgupta, P. 1990. Well-being and the extent of its realization in poor countries. *Economic Journal* 100:1–32.
Dasgupta, P. 1993. *An inquiry into well-being and destitution.* Oxford: Oxford University Press.
Dasgupta, P., and M. Weale. 1992. Well-being and the extent of its realization in poor countries. *World Development* 20:119–31.
Davenport, L. 2020. The fluidity of racial classifications. *Annual Review of Political Science* 23:221–40.
Davis, K. 2014. Legal indicators: the power of quantitative measures of law. *Annual Review of Law and Social Science* 10:37–52.
Daxecker, U., et al. 2019. Electoral contention and violence (ECAV): a new dataset. *Journal of Peace Research* 56:714–23.
Debs, A., and N. P. Monteiro. 2017. *Nuclear politics: the strategic causes of proliferation.* Cambridge: Cambridge University Press.
Diamond, L. 1992. Economic development and democracy reconsidered. In G. Marks and L. Diamond (eds.) *Reexamining democracy: essays in honor of Seymour Martin Lipset.* Newbury Park: Sage Publications.
Diamond, L., and L. Morlino (eds.). 2005. *Assessing the quality of democracy.* Baltimore: Johns Hopkins University Press.
Diamond, L., and L. Morlino. 2004. The quality of democracy: an overview. *Journal of Democracy* 15:20–31.
Diehl, P., C. Ku, and D. Zamora. 2003. The dynamics of international law: the interaction of the normative and operating systems. *International Organization* 57:43–57.
Dimitrov, R., et al. 2007. International nonregimes: a research agenda. *International Studies Review* 9:230–58.
Dion, D. 2003. Evidence and inference in the comparative case study. In G. Goertz and H. Starr (eds.) *Necessary conditions: theory, methodology, and applications.* New York: Rowman & Littlefield.
Doi, K., T. Fairfield, and Y. Wang. 2025. Convergence between chinaís and japanís african cooperation summits: the role of african agency. Manuscript.
Doll, R., and Hill A. 1954. The mortality of doctors in relation to their smoking habits: a preliminary report. *British Medical Journal* 1:1451–55.
Doner, R., et al. 2005. Systematic vulnerability and the origins of developmental states: Northeast and Southeast Asia in comparative perspective. *International Organization* 59:327–61.
Donno, D. 2010. Who is punished? Regional intergovernmental organizations and the enforcement of democratic norms. *International Organization* 64:593–625.
Donno, D. 2013. *Defending democratic norms: international actors and the politics of electoral misconduct.* Oxford: Oxford University Press.
Donno, D., et al. 2015. Screening out risk: IGOs, member state selection, and interstate conflict, 1951–2000. *International Studies Quarterly* 59:251–63.
Donno, D., and A.-K. Kreft. 2019. Authoritarian institutions and womens rights. *Comparative Political Studies* 52:720–53.
Dowding, K., and M. van Hees. 2003. The construction of rights. *American Political Science Review* 97:281–93.
Downes, A. 2006. Desperate times, desparate measures: the causes of civilian victimization in war. *International Security* 30:152–95.

Downes, A. B. 2021. *Catastrophic success: why foreign-imposed regime change goes wrong.* Ithaca: Cornell University Press.
Downes, A., and J. Monten. 2013. Forced to be free: why foreign-imposed regime change rarely leads to democratization. *International Security* 37:90–131.
Downes, A., and T. Sechser. 2012. The illusion of democratic credibility. *International Organization* 66:457–89.
Downing, B. 1992. *The military revolution and political change: origins of democracy and autocracy in early modern Europe.* Princeton: Princeton University Press.
Doyle, M., and N. Sambanis. 2000. International peacebuilding: a theoretical and quantitative analysis. *American Political Science Review* 94:779–801.
Druckman, J. 2021. *Experimental thinking: a primer on social science experiments.* Cambridge: Cambridge University Press.
Duffield, J. 2003. The limits of 'rational design'. *International Organization* 57:411–30.
Dunning, T. 2008. *Crude democracy: natural resource wealth and political regimes.* Cambridge: Cambridge University Press.
Dworkin, R. 1981. What is equality? Part 1: equality of welfare, Part 2: equality of resources. *Philosophy and Public Affairs* 10:185–246, 283–.
EGAP. 2017. 10-things-know-about-mechanisms. Manuscript.
EU Social Protection Committee. 2012. *Social protection performance monitor (sppm) methodological report by the indicators sub-group of the social protection committee.* European Union.
Easterly, W., and R. Levine. 1997. Africa's growth tragedy: policies and ethnic divisions. *Quarterly Journal of Economics* 112:1203–50.
Edelstein, D. M. 2017. *Over the horizon: time, uncertainty, and the rise of great powers.* Ithaca: Cornell University Press.
Eibl, F., and S. Hertog. 2024. From rents to welfare: why are some oil-rich states generous to their people? *American Political Science Review* 118:1324–43.
Eisenstadt, T. A., et al. 2017. *Constituents before assembly: participation, deliberation, and representation in the crafting of new constitutions.* Cambridge: Cambridge University Press.
Elman, C. 2005. Explanatory typologies in qualitative studies of international politics. *International Organization* 59:293–326.
Elman, M. (ed.). 1997. *Paths to peace: is democracy the answer?* Cambridge: MIT Press.
Enders, W., and T. Sandler. 2002. Patterns of international terrorism. *International Studies Quarterly* 46:145–65.
Enders, W., and T. Sandler. 2005. After 9/11: Is it all different now? *Journal of Conflict Resolution* 49:259–77.
Ertman, T. 1997. *Birth of the leviathan: building states and regimes in medieval and early modern Europe.* Cambridge: Cambridge University Press.
Esping-Andersen, G. 1990. *The three worlds of welfare capitalism.* Cambridge: Polity Press.
Esping-Andersen, G. 1999. *Social foundations of postindustrial economics.* Oxford: Oxford University Press.
Esterling, K., et al. 2024. The necessity of construct and external validity for deductive causal inference. *Journal of Causal Inference* 21:45–62.
European Commission. 2009. *Portfolio of indicators for the monitoring of the European strategy for social protection and social inclusion, 2009 update.* .
Falleti, T., and J. Lynch. 2009. Context and causal mechanisms in political analysis. *Comparative Political Studies* 42:1143–66.
Fanlo, A., and L. Sukin. 2023. The disadvantage of nuclear superiority. *Security Studies* 32:446–75.
Fazal, T. 2012. Why states no longer declare war. *Security Studies* 21:557–93.
Fearon, J. 2003. Ethnic and cultural diversity by country. *Journal of Economic Growth* 8:195–222.
Fearon, J., and D. Laitin. 2003. Ethnicity, insurgency and civil war. *American Political Science Review* 97:75–90.

Feldman, S., and L. Huddy. 2005. Racial resentment and white opposition to race-conscious programs: principles or prejudice? *American Journal of Political Science* 49:168–83.
Findley, M., and T. Teo. 2006. Rethinking third-party interventions into civil wars: an actor-centric approach. *Journal of Politics* 68:828–37.
Fine, B., and K. Fine. 1974. Social choice and individual rankings, I and II. *Review of Economic Studies* 44:303–22, 459–7.
Fischer, C., et al. 1996. *Inequality by design: cracking the bell curve myth.* Princeton: Princeton University Press.
Foley, J. 1967. Resource allocation in the public sector. *Yale Economic Essays* 7:73–76.
Follesdal, A., and S. Hix. 2005. Why there is a democratic deficit in the EU: a response to Majone and Moravcsik. Manuscript.
Foran, J. 1997. The comparative-historical sociology of Third World social revolutions: why a few succeed, why most fail. In J. Foran (ed.) *Theorizing revolution.* London: Routledge.
Foran, J. 1997. The future of revolutions at the *fin-de-siècle*. *Third World Quarterly* 18:791–820.
Fordham, B. 2011. Who wants to be a major power? Explaining the expansion of foreign policy ambition. *Journal of Peace Research* 48:587–603.
Foster, D. 2006. State power, linkage mechanisms, and diversion against nonrivals. *Conflict Management and Peace Science* 23:1–23.
Fox, R., and J. Lawless. 2004. Entering the arena? Gender and the decision to run for office. *American Journal of Political Science* 48:264–80.
Fox, R., and J. Lawless. 2005. To run or not to run for office: explaining nascent political ambition. *American Journal of Political Science* 49:642–59.
Franck, T. 1988. Legitimacy in the international system. *American Journal of International Law* 82:705–59.
Franck, T. 1990. *Power of legitimacy among nations.* Oxford: Oxford University Press.
Franck, T. 1999. The emerging right to democratic governance. *American Journal of International Law* 86:75–80.
Frazier, D. 2006. Third-party characteristics, territory, and the mediation of militarized interstate disputes. *Conflict Management and Peace Science* 23:1–18.
Freedman, D. 2010. Statistical models and causal inference: a dialogue with the social sciences. In *Black ravens, white shoes, and case selection.* Cambridge: Cambridge University Press.
Freedman, D., et al. 2010. *Statistical models and causal inference: a dialogue with the social sciences.* Cambridge: Cambridge University Press.
Fukuda-Parr, S., et al. 2009. An index of economic and social rights fufillment: concept and methodology. *Journal of Human Rights* 8:195–221.
Fuller, J., et al. 2024. The argument framework is a flexible approach to evidence in healthcare. *Nature Medicine* 30 8:2113–116.
Gandhi, J., and E. Lust-Okar. 2009. Elections under authoritarianism. *Annual Review of Political Science* 12:403–22.
Garrity, M. 2023. What enables or constrains mass expulsion? A new decision-making framework. *Security Studies* 32:537–67.
Gates, S., et al. 2006. Institutional inconsistency and political instability: polity duration, 1800–2000. *American Journal of Political Science* 50:893–908.
Gaus, G. 2000. *Political concepts and political theories.* Boulder: Westview Press.
Gavin, F. 2014. What we talk about when we talk about nuclear weapons. *H-Diplo, ISSF Forum, No. 2* .
Geddes, B. 1999. What do we know about democratization after twenty years? *Annual Review of Political Science* 2:115–44.
Geddes, B. 2003. *Paradigms and sand castles: theory building and research design in comparative politics.* Ann Arbor: University of Michigan Press.
Geddes, B., et al. n.d. Autocratic regimes code book version 1.2. Manuscript.
Gelman, A., and G. Imbens. 2013. Why ask why? forward causal inference and reverse causal questions. Manuscript. NBER Working Paper 19614.

Geng, E., et al. When the parts are greater than the whole: a mechanism based approach to generalizing implementation strategies. *Implementation Science* 20:1–12.
George, A. 2000. Foreword. In A. Kacowicz et al. (eds.) *Stable peace among nations.* New York: Rowman & Littlefield.
George, A., and A. Bennett. 2005. *Case studies and theory development.* Cambridge: MIT Press.
Georgetown University's Institute for Women, Peace and Security, and PRIO. 2017. Women peace and security index, 2017/18. Manuscript.
Gerring, J. 2012. *Social science methodology: a unified framework,* 2nd edition. Cambridge: Cambridge University Press.
Gerschewski, J. 2023. *The two logics of autocratic rule.* Cambridge: Cambridge University Press.
Ghatak, S., et al. 2017. External threat and the limits of democratic pacifism. *Conflict Management and Peace Science* 34:141–59.
Gibbs, J. 1989. Conceptualization of terrorism. *American Sociological Review* 54:329–40.
Gibler, D., and M. Sarkees. 2004. Measuring alliances: the Correlates of War formal interstate alliance dataset, 1816–2000. *Journal of Peace Research* 41:211–22.
Gibson, J., and M. Howard. 2007. Russian anti-semitism and the scapegoating of jews: the dog that didnt bark? *British Journal of Political Science* 37:193–223.
Gleditsch, K. 2002. Expanded trade and GDP data. *Journal of Conflict Resolution* 46:712–24.
Gleditsch, K., and M. Ward. 1999. A revised list of independent states since the Congress of Vienna. *International Interactions* 25:393–413.
Gleditsch, N., and G. Schneider (eds.). 2010. Special issue on the capitalist peace. *International Interactions* 36:107–205.
Globerman, S., and D. Shapiro. 2003. Governance infrastructure and US foreign direct investment. *Journal of International Business Studies* 34::9–39.
Goertz, G. 2003. *International norms and decision making: a punctuated equilibrium model.* New York: Rowman & Littlefield.
Goertz, G. 2008. A checklist for constructing, evaluating, and using concepts or quantitative measures. In J. Box-Steffensmier, H. Brady, and D. Collier (eds.) *The Oxford handbook of political methodology.* Oxford: Oxford University Press.
Goertz, G., B. Jones, and P. Diehl. 2005. Maintenance processes in international rivalries. *Journal of Conflict Resolution* 49:742–69.
Goertz, G., P. Diehl, and A. Balas. 2016. *The puzzle of peace: the evolution of peace in the international system.* Oxford: Oxford University Press.
Goertz, G., and A. Mazur (eds.). 2008. *Politics, gender, and concepts: theory and methodology.* Cambridge: Cambridge University Press.
Goertz, G., and J. Dul. 2023. The democratic peace: generalization, the confounder test and Necessary Condition Analysis. Manuscript.
Goertz, G., and J. Levy (eds.). 2007. *Explaining war and peace: case studies and necessary condition counterfactuals.* London: Routledge.
Goertz, G., and J. Mahoney. 2012. *A tale of two cultures: qualitative and quantitative research in the social sciences.* Princeton: Princeton University Press.
Goertz, G., and S. Haggard. 2023. Generalization, case studies, and within-case causal inference: Large-N Qualitative Analysis (LNQA). In H. Kincaid and J. Van Bouwel (eds.) *The Oxford handbook on the philosophy of political science.* Oxford: Oxford University Press.
Goettlich, K. 2025. Territorial integrity as an etiquette of thieves: non-conquest in nineteenth-century imperialism. *International Organization* 79:712–38.
Goldberg, E. 20210821. The new chief chaplain at Harvard? An atheist. *New York Times*.
Goldstone, J. 1990 (2016). *Revolution and rebellion in the early modern world.* Berkeley: University of California Press.
Goldstone, J., et al. 2000. State failure task force report: phase III findings. Manuscript. University of Maryland.

Gonzlez, Y. M. 2020. *Authoritarian police in democracy: contested security in Latin America.* Cambridge: Cambridge University Press.
Goodwin, J. 2001. *No other way out: states and revolutionary movements, 1945–1991.* Cambridge: Cambridge University Press.
Grävingholt, J., and et al,. 2015. Disaggregating state fragility: a method to establish a multidimensional empirical typology. *Third World Quarterly* 36:1281–98.
Guardado, J., et al. 2018. Quality peace in post–civil war settings: the role of local institutions. In M. Joshi and P. Wallensteen (eds.) *Understanding quality peace: peacebuilding after civil war.* London: Routledge.
Gunitsky, S. 2014. From shocks to waves: hegemonic transitions and democratization in the twentieth century. *International Organization* 68:561–97.
Gunitsky, S. 2018. Democratic waves in historical perspective. *Perspectives on Politics* 16:634–51.
Gurr, T. 1974. Persistence and change in political systems 1800–1971. *American Political Science Review* 68:1482–504.
Guttman, L. 1944. A basis for scaling qualitative data. *American Sociological Review* 9:139–50.
Habyarimana, J., et al. 2007. Why does ethnic diversity undermine public goods provision? *American Political Science Review* 101:709–25.
Hafner-Burton, E., and J. Ron. 2009. Seeing double: human rights impact through qualitative and quantitative eyes. *World Politics* 61:360–401.
Halpern, J. 2016. *Actual causality.* Cambridge: MIT Press.
Handlin, S. 2017. *State crisis in fragile democracies: polarization and political regimes in South America.* Cambridge: Cambridge University Press.
Harding, D., C. Fox, and J. Mehta. 2002. Studying rare events through qualitative case studies: lessons from a study of rampage school shootings. *Sociological Methods & Research* 31:174–217.
Harff, B. 2003. No lessons learned from the Holocaust? Assessing risks of genocide and political mass murder since 1955. *American Political Science Review* 97:57–73.
Harkness, K. 2016. The ethnic army and the state: explaining coup traps and the difficulties of democratization in Africa. *Journal of Conflict Resolution* 60:587–616.
Harkness, K. A. 2018. *When soldiers rebel: ethnic armies and political instability in Africa.* Cambridge: Cambridge University Press.
Harsanyi, J. 1955. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy* 63:309–21.
Hartzell, C. A., and M. Hoddie. 2020. *Power sharing and democracy in post-civil war states: the art of the possible.* Cambridge: Cambridge University Press.
Hartzell, C., and M. Hoddie. 2003. Institutionalizing peace: power sharing and post–civil war conflict management. *American Journal of Political Science* 47:318–32.
Harvey, F. 2003. Practicing coercion: revisiting successes and failures using Boolean logic and comparative methods. In G. Goertz and H. Starr (eds.) *Necessary conditions: theory, methodology, and applications.* New York: Rowman & Littlefield.
Haslanger, S. 2015. Social structure, narrative and explanation. *Canadian Journal of Philosophy* 45:1–15.
Hathaway, O. 2002. Do human rights treaties make a difference? *Yale Law Journal* 111:1935–2042.
Hausman, D. 2015. *Valuing health: well-being, freedom, and suffering.* Oxford: Oxford University Press.
He, B., and M. Warren. 2011. Authoritarian deliberation: the delibarative turn in Chinese political development. *Perspectives on Politics* 9:269–90.
Healy, K. 2019. *Data visualization: a practical introduction.* Princeton: Princeton University Press.
Hegre, H., et al. 2001. Toward a democratic civil peace? Democracy, political change, and civil war. *American Political Science Review* 95:33–48.
Held, V. 2004. Terrorism and war. *Journal of Ethics* 8:59–75.

Helmke, G., and J. Rath. 2025. Defining and measuring democratic norms. *Annual Review of Political Science* 28:233–51.
Henke, M. E. 2019. *Constructing allied cooperation: diplomacy, payments, and power in multilateral military coalitions.* Ithaca: Cornell University Press.
Hensel, P. 2012. Territory: geography, contentious issues, and world politics. In J. Vasquez (ed.) *What do we know about war?* 2nd edition. New York: Rowman & Littlefield.
Hicks, A., and L. Kenworthy. 2003. Varieties of welfare capitalism. *Socio-Economic Review* 1:27–61.
Hirai, T. 2017. *The creation of the human development approach.* London: Palgrave.
Holleman, B. 1999. The nature of the forbid/allow asymmetry: two correlational studies. *Sociological Methods & Research* 28:209–44.
Hollyer, J. R., B. P. Rosendorff, and J. R. Vreeland. 2011. Democracy and transparency. *Journal of Politics* 73:1191–205.
Homola, J., et al. 2024. Fixed effects and post-treatment bias in legacy studies. *American Political Science Review* 118:537–44.
Hong, M., and G. Uzonyi. 2018. Deeper commitment to human rights treaties: signaling and investment risk perception. *International Interactions* 44:1040–64.
Hooghe, L., and G. Marks. 2003. Unraveling the central state, but how? Types of multi-level governance. *American Political Science Review* 97:233–43.
Hoover Green, A. 2018. *The commander's dilemma: violence and restraint in wartime.* Ithaca: Cornell University Press.
Horowitz, M. 2010. *The diffusion of military power: causes and consequences for international politics.* Princeton: Princeton University Press.
Howard, L. 2015. US foreign policy habits in ethnic conflict. *International Studies Quarterly* 59:721–34.
Howell, W., and J. Pevehouse. 2007. *While dangers gather: Congressional checks on presidential war powers.* Princeton: Princeton University Press.
Htun, M., and S. L. Weldon. 2018. *The logics of gender justice: state action on women's rights around the world.* Cambridge: Cambridge University Press.
Huber, E., and J. Stephens. 2012. *Democracy and the left: social policy and inequality in Latin America.* Chicago: University of Chicago Press.
Hudson, V., et al. 2015. Clan governance and state stability: the relationship between female subordination and political order. *American Political Science Review* 109:535–55.
Hug, S. 2013. The use and misuse of the "minorities at risk" project. *Annual Review of Political Science* 16:191–208.
Humphreys, M., and A. Jacobs. 2015. Mixing methods: a Bayesian approach. *American Political Science Review* 109:653–73.
Irajpanah, K., and K. Schultz. 2021. Off the menu: post-1945 norms and the end of war declarations. *Security Studies* 30:485–516.
Jacoby, W. 2000. *Imitation and politics: redesigning modern Germany.* Ithaca: Cornell University Press.
Jakobsen, T., et al. 2013. Why do poor countries suffer costly conflict? *Conflict Management and Peace Science* 30:140–60.
James, P. 2022. *Realism and international relations: a graphic turn toward scientific progress.* Oxford: Oxford University Press.
Jones, D., S. Bremer, and J. Singer. 1996. Militarized interstate disputes, 1816–1992: rationale, coding rules, and empirical patterns. *Conflict Management and Peace Science* 15:163–212.
Joshi, M., et al. 2017. Sequencing the peace: how the order of peace agreement implementation can reduce the destabilizing effects of post-accord elections. *Journal of Conflict Resolution* 61:4–28.
Joshi, M., and P. Wallensteen (eds.). 2018. *Understanding quality peace: peacebuilding after civil war.* London: Routledge.

Joshi, M., and P. Wallensteen. 2018. Understanding quality peace: introducing the five dimensions. In M. Joshi and P. Wallensteen (eds.) *Understanding quality peace: peacebuilding after civil war.* London: Routledge.
Jost, T. 2023. The institutional origins of miscalculation in chinaís international crises. *International Security* 48:47–90.
Kacowicz, A., et al. 2021. *The unintended consequences of peace: peaceful borders and illicit transnational flows.* Cambridge: Cambridge University Press.
Kahwati, L., and H. Kane. 2019. *Qualitative comparative analysis in mixed methods research and evaluation.* Newbury Park: Sage Publications.
Kalyvas, S. 2015. How civil wars help explain organized crime—and how they do not. *Journal of Conflict Resolution* 59:1517–40.
Kang, S., and J. Meernik. 2005. Civil war destruction and the prospects for economic growth. *Journal of Politics* 67:88–109.
Karantycky, A. 2000. *The comparative study of freedom – 1999–2000: survey methodology.* Freedom House.
Karreth, J., et al. 2023. International third parties and the implementation of comprehensive peace agreements after civil war. *Journal of Conflict Resolution* 67:494–521.
Kasuya, Y., and K. Mori. 2019. Better regime cutoffs for continuous democracy measures. Manuscript. V-Dem Users Working Paper Series, 25.
Kaufman, S. 2006. Symbolic politics or rational choice? Testing theories of extreme ethnic violence. *International Security* 4:45–86.
Kaufman, S. J. 2015. *Nationalist passions.* Ithaca: Cornell University Press.
Keck, M., and K. Sikkink. 1998. *Activists beyond borders: advocacy networks in international politics.* Ithaca: Cornell University Press.
Kelley, D. 2014. *The art of reasoning,* 4th edition. New York: W.W. Norton.
Kenkel, B., and J. Paine. 2023. A theory of external wars and European parliaments. *International Organization* 77:102–43.
Kenworthy, L. 2003. Quantitative indicators of corporatism. *International Journal of Sociology* 33:10–44.
Ker-Lindsay, J., and F. Mikulas. 2023. *Secession and state creation: what everyone needs to know.* Oxford: Oxford University Press.
Khalidi, M. 2013. *Natural categories and human kinds: classification in the natural and social sciences.* Cambridge: Cambridge University Press.
Khosrowi, D. 2023. Extrapolating from experiments, confidently. *European Journal for the Philosophy of Science* 13.
Kier, E. 2021. *War and democracy: labor and the politics of peace.* Ithaca: Cornell University Press.
Kim, N., and G. Uzonyi. 2020. Government mass killing and post-conflict domestic trials. *Studies in Conflict & Terrorism* 43:396–413.
King, G., et al. 2017. The balance-sample size frontier in matching methods for causal inference. *American Journal of Political Science* 61:473–89.
King, G., and C. Murray. 2002. Rethinking human security. *Political Science Quarterly* 116:585–610.
King, G., and L. Zeng. 2001. Logistic regression in rare events data. *Political Analysis* 9:137–63.
King, G., and L. Zeng. 2001a. Improving forecasts of state failure. *World Politics* 53:623–58.
King, G., and L. Zeng. 2007. When can history be our guide? The pitfalls of counterfactual inference. *International Studies Quarterly* 51:183–210.
Kingdon, J. 1984. *Agendas, alternatives, and public policies.* Boston: Little, Brown.
Kirgis, F. 1987. Custom on a sliding scale. *American Journal of International Law* 81:146–51.
Klein, J., G. Goertz, and P. Diehl. 2006. The new rivalry data set: procedures and patterns. *Journal of Peace Research* 43:331–48.
Klein, J., G. Goertz, and P. Diehl. 2008. The peace scale: conceptualizing and operationalizing nonrivalry and peace. *Conflict Management and Peace Science* 25:67–80.

Klingemann, H.-D., et al. 2006. *Mapping policy preferences II: estimates for parties, electors and governments in central and eastern Europe, European Union and OECD 1990–2003.* Oxford: Oxford University Press.
Klosek, K., et al. 2021. Frozen conflicts in world politics: a new dataset. *Journal of Peace Research* 58:849–58.
Knijn, T., and I. Ostner. 2002. Commodification and de-commodification. In B. Hobson, J. Lewis, and B. Siim (eds.) *Contested concepts in gender and social politics.* Cheltenham: Edward Elgar.
Knutsen, C., et al. 2024. Conceptual and measurement issues in assessing democratic backsliding. *PS: Political Science & Politics* 57:162–77.
Koch, L. 2023. Military regimes and resistance to nuclear weapons development. *Security Studies* 32:239–70.
Kocher, M., and N. Monteiro. 2016. Lines of demarcation: causation, design-based inference, and historical research. *Perspectives on Politics* 14:952–75.
Koivu, K. 2016. In the shadow of the state: mafias and illicit markets. *Comparative Political Studies* 49:155–83.
Kosack, S., and J. Tobin. 2006. Funding self-sustaining development: the role of aid, FDI and government in economic success. *International Organization* 60:205–43.
Kramer, M. 2002. Why freedoms do not exist in degrees. *Political Studies* 50:230–43.
Krause, P. 2017. *Rebel power: why national movements compete, fight, and win.* Ithaca: Cornell University Press.
Krebs, R., and R. Licklider. 2015. United they fall: why the international community should not promote military integration after civil war. *International Security* 40:93–138.
Kreuzer, M. 2023. *The grammar of time: a toolbox for comparative historical analysis.* Cambridge: Cambridge University Press.
Kroenig, M. 2009. Exporting the bomb: why states provide sensitive nuclear assistance. *American Political Science Review* 103:113–33.
Krook, M. 2014. Electoral gender quotas: a conceptual analysis. *Comparative Political Studies* 47:1268–93.
Krook, M., et al. 2009. Gender quotas and models of political citizenship. *British Journal of Political Science* 39:781–803.
Krygier, M. 2016. The rule of law: pasts, presents, and two possible futures. *Annual Review of Law and Social Science* 12:199–229.
Kugler, J., and A. Organski. 1989. The power transition: a retrospective and prospective evaluation. In M. Midlarsky (ed.) *Handbook of war studies.* Ann Arbor: University of Michigan Press.
Kukull, W., and M. Ganguli. 2012. Generalizability: the trees, the forest, and the low-hanging fruit. *Neurology* 78:1886–91.
Kushi, S., and M. Toft. 2023. Introducing the military intervention project: a new dataset on us military interventions, 1776–2019. *Journal of Conflict Resolution* 67:752–79.
Lührmann, A., et al. 2017. Regimes in the world (RIW): a robust regime type measure based on V-Dem. Paper presented at the annual meetings of the The Varities of Democracy Institute.
Lührmann, A., et al. 2018. Regimes of the World (RoW): opening new avenues for the comparative study of political regimes. *Politics and Governance* 6:60–77.
Lührmann, A., et al. 2020. Constraining governments: new indices of vertical, horizontal, and diagonal accountability. *American Political Science Review* 114:811–20.
Lai, B. 2003. Examining the goals of US foreign assistance in the post–Cold War period, 1991–96. *Journal of Peace Research* 40:103–28.
Lake, D. A. 2016. *The statebuilder's dilemma: on the limits of foreign intervention.* Ithaca: Cornell University Press.
Lake, M. 2017. Building the rule of war: postconflict institutions and the micro-dynamics of conflict in Eastern DR Congo. *International Organization* 71:281–315.

Lakoff, G. 1987. *Women, fire and dangerous things: what categories reveal about the mind.* Chicago: University of Chicago Press.
Laporte, J., D. Collier, and . 2007. Choices about levels of measurement: placing nominal scales in perspective. Manuscript.
Larkins, C. 1996. Judicial independence and democratization: a theoretical and conceptual analysis. *American Journal of Comparative Law* 44:605–26.
Lawson, G. 2019. *Anatomies of revolution.* Cambridge: Cambridge University Press.
Lee, M. 2018. The international politics of incomplete sovereignty: how hostile neighbors weaken the state. *International Organization* 72:283–315.
Lee, M. M. 2020. *Crippling Leviathan: how foreign subversion weakens the state.* Ithaca: Cornell University Press.
Leeds, B. 2005. Alliance treaty obligations and provisions (ATOP) codebook (version 3.0). Manuscript. Rice University.
Leipziger, L. 2025. Democracy and ethnic inequality: a comparative case study. *Democratization* 32:373–96.
Leipziger, L. E., et al. 2025. Does ethnopolitical exclusion cause civil war onset via grievances? Evidence from 15 case studies. *Perspectives on Politics*.
Lemke, D., and S. Werner. 1996. Power parity, commitment to change, and war. *International Studies Quarterly* 40:235–60.
Lemke, D., and W. Reed. 2001. The relevance of politically relevant dyads. *Journal of Conflict Resolution* 45:126–45.
Lessing, B. 2021. Conceptualizing criminal governance. *Perspectives on Politics* 19:854–73.
Levari, M., et al. 2018. Prevalence-induced concept change in human judgment. *Science* 360:1465–67.
Levi, M. 2002. The state of the study of the state. In I. Katznelson and H. Milner (eds.) *Political science: the state of the discipline.* New York: W.W. Norton.
Levin, D. 2019. Partisan electoral interventions by the great powers: Introducing the PEIG Dataset. *Conflict Management and Peace Science* 36:88–106.
Levitsky, S., and J. Loxton. 2013. Populism and competitive authoritarianism in the Andes. *Democratization* 20:107–36.
Levitsky, S., and L. Way. 2010. *Competitive authoritarianism: hybrid regimes after the Cold War.* Cambridge: Cambridge University Press.
Levitsky, S., and L. Way. 2022. *Revolution and dictatorship: the violent origins of durable authoritarianism.* Princeton: Princeton University Press.
Levy, J. S. 1989. The diversionary theory of war: a critique. In M. Midlarsky (ed.) *Handbook of war studies.* Boston: Unwin Hyman.
Levy, J., and W. Thompson. 2005. Hegemonic threats and great power balancing in Europe, 1495–1999. *Security Studies* 14:1–30.
Li, Q. 2005. Does democracy promote or reduce transnational terrorist incidents? *Journal of Conflict Resolution* 49:278–97.
Li, Q., and A. Resnick. 2003. Reversal of fortunes: democratic institutions and foreign direct investment inflows to developing countries. *International Organization* 57:175–211.
Licklider, R. 1995. The consequences of negotiated settlements in civil wars, 1945–1993. *American Political Science Review* 89:681–90.
Lijphart, A. 2011. Democratic quality in stable democracies. *Society* 48:17–18.
Lindberg, S. 2016. Ordinal versions of V-Dem's indices: When interval measures are not useful for classification, description, and sequencing analysis purposes. *Geopolitics, History, and International Relations* 8:76–111.
Lindemann, S., and A. Wimmer. 2018. Repression and refuge: why only some politically excluded ethnic groups rebel. *Journal of Peace Research* 55:305–19.
Linos, K., and T. Pegram. 2016. The language of compromise in international agreements. *International Organization* 70:587–621.

Linz, J., and A. Stepan. 1996. *Problems of democratic transition and consolidation: Southern Europe, South America, and post-communist Europe.* Baltimore: Johns Hopkins University Press.
Little, A., and A. Meng. 2024. Measuring democratic backsliding. *PS: Political Science & Politics* 57:149–61.
Little, D. 1993. On the scope and limits of generalization in the social sciences. *Synthese* 97:183–207.
Londregan, J., and K. Poole. 1996. Does high income promote democracy? *World Politics* 49:1–31.
Long, A. 2003. Defense pacts and international trade. *Journal of Peace Research* 40:537–52.
Long, T., and C.-A. Schulz. 2023. Compensatory layering and the birth of the multipurpose multilateral IGO in the Americas. *International Organization* 77:1–32.
Lowe, E. 1980. For want of a nail. *Analysis* 40:50–52.
Loxton, J. 2021. *Conservative party-building in Latin America authoritarian inheritance and counterrevolutionary struggle.* Oxford: Oxford University Press.
Loyle, C., and H. Binninsbø. 2018. Justice during armed conflict: a new dataset on government and rebel strategies. *Journal of Conflict Resolution* 62:442–66.
Luebbert, G. 1991. *Liberalism, fascism, or social democracy: social classes and the political origins of regimes in interwar Europe.* Oxford: Oxford University Press.
Luna, J. 2014. Party system institutionalization: do we need a new concept? *Studies in Comparative International Development* 49:403–25.
Lupton, D. L. 2020. *Reputation for resolve: how leaders signal determination in international politics.* Ithaca: Cornell University Press.
Lupton, D., and V. Morkevicius. 2019. The fog of war: violence, coercion and *jus ad vim*. In J. Galliott (ed.) *Force short of war in modern conflict: jus ad vim.* Edinburgh: Edinburgh University Press.
Lutscher, P., et al. 2022. Education policies and systems across modern history, codebook v2.0. Manuscript.
Lutscher, P., et al. 2022. Education policies and systems across modern history, codebook v2.0. Manuscript.
Lyall, J. 2020. *Divided armies: inequality and battlefield performance in modern war.* Princeton: Princeton University Press.
Maat, E. 2024. *Elite rivalry, mass killing and genocide in authoritarian regimes: why autocrats kill.* London: Routledge.
MacDonald, P. 2023. Civilized barbarism: what we miss when we ignore colonial violence. *International Organization* 77:21–53.
Mackie, J. 1980 (1974). *The cement of the universe: a study of causation.* Oxford: Oxford University Press.
Maerz, S. 2020. The many faces of authoritarian persistence: a set-theory perspective on the survival strategies of authoritarian regimes. *Government and Opposition* 55:64–87.
Maetas, C., et al. 2006. When to risk it? Institutions, ambitions, the decision to run for the U.S. House. *American Political Science Review* 100:195–208.
Mahoney, J. 2021. *The logic of social science.* Princeton: Princeton University Press.
Mahoney, J., and R. Barrenechea. 2019. The logic of counterfactual analysis in historical explanation. *British Journal of Sociology* 70:306–38.
Mainwaring, S., and A. Prez-Linán. 2013. *The emergence and fall of democracies and dictatorships: Latin America since 1900.* Cambridge: Cambridge University Press.
Malone, I. 2022. Unmasking militants: organizational trends in armed groups, 1970–2012. *International Studies Quarterly* 66.
Mann, M. 2005. *The dark side of democracy: explaining ethnic cleansing.* Cambridge: Cambridge University Press.
Mansfield, E., and J. Pevehouse. 2000. Trade blocs, trade flows, and international conflict. *International Organization* 54:775–808.

Mansfield, E., and J. Pevehouse. 2006. Democratization and international organization. *International Organization* 60:137–67.
Mansfield, E., and J. Pevehouse. 2008. Democratization and the varieties of international organization. *Journal of Conflict Resolution* 52:269–94.
Mansfield, E., and R. Bronson. 1997. Alliances, preferential trade arrangements and international trade. *American Political Science Review* 91:94–107.
Marks, G. 1986. Neocorporatism, and incomes policy in western Europe and north America. *Comparative Politics* 18:253–77.
Matovski, A. 2021. *Popular dictatorships: crises, mass opinion, and the rise of electoral authoritarianism.* Cambridge: Cambridge University Press.
Maull, H., et al. 2006. Civilian powers in international relations. Manuscript.
Mazur, A., and D. Stetson. 2003. Quantifying complex concepts: the case of the *women's movement* in the RNGS project. *APSA-CP: Newsletter of the APSA Comparative Politics Section* 14:11–14.
Mazzuca, S. 2021. *Latecomer state formation: political geography and capacity failure in Latin America.* New Haven: Yale University Press.
McAdams, D., and H. Boudet. 2012. *Putting social movements in their place explaining opposition to energy projects in the United States, 2000–2005.* Cambridge: Cambridge University Press.
McCall Smith, J. 2000. The politics of dispute settlement design: explaining legalism in regional trade pacts. *International Organization* 54:137–80.
Meernik, J. 1994. Presidential decision making and the political use of military force. *International Studies Quarterly* 38:121–138.
Meernik, J. 2000. Modeling international crises and the political use of force by the USA. *Journal of Peace Research* 38:547–62.
Mellon, J. 2025. Rain, rain, go away: 194 potential exclusion-restriction violations for studies using weather as an instrumental variable. *American Journal of Political Science* 69:881–98.
Merry, S. 2016. *The seductions of quantification: measuring human rights, gender violence, and sex trafficking.* Chicago: University of Chicago Press.
Michel, V.,, and K. Sikkink. 2013. Human rights prosecutions and the participation rights of victims in Latin America. *Law and Society Review* 47:873–907.
Mikkelsen, K. 2015. Fuzzy-set case studies. *Sociological Methods & Research* 46:422–55.
Miller, B. 2001. The global sources of regional transitions from war to peace. *Journal of Peace Research* 38:199–225.
Miller, M. 2024. How Little and Meng's objective approach fails in democracies. *PS: Political Science & Politics* 57:202–7.
Miller, M. K. 2021. *Shock to the system: coups, elections, and war on the road to democratization.* Princeton: Princeton University Press.
Milner, H., and K. Kubota. 2005. Why the move to free trade? democracy and trade policy in the developing countries. *International Organization* 59:107–43.
Mitchell, S., and B. Prins. 2004. Rivalry and diversionary uses of force. *Journal of Conflict Resolution* 48:937–81.
Moffitt, B. 2016. *The global rise of populism: performance, political style, and representation.* Stanford: Stanford University Press.
Moncada, E. 2021. *Resisting extortion.* Cambridge: Cambridge University Press.
Montgomery, E. B. 2016. *In the hegemon's shadow: leading states and the rise of regional powers.* Ithaca: Cornell University Press.
Mookherji, S, and LaFond A. 2013. Strategies to maximize generalization from multiple case studies: lessons from the Africa routine immunization system essentials (ARISE) project. *Evaluation* 19:284–303.
Moravcsik, A. 1998. *The choice for Europe.* Ithaca: Cornell University Press.
Moravcsik, A. 2002. In defence of the 'democratic deficit': reassessing the legitimacy of the European Union. *Journal of Common Market Studies* 40:603–24.

Morgan, M. 2019. Exemplification and the use-values of cases and case studies. *Studies in History and Philosophy of Science* 78.
Morgan, S., and C. Winship. 2015. *Counterfactuals and causal inference: methods and principles for social research,* 2nd edition. Cambridge: Cambridge University Press.
Morgenbesser, L., and T. Pepinsky. 2019. Elections as causes of democratization: southeast Asia in comparative perspective. *Comparative Political Studies* 52:3–35.
Morris, I. 2013. *The measure of civilization how social development decides the fate of nations.* Princeton: Princeton University Press.
Morrow, J. 2007. When do states follow the rules of war? *American Political Science Review* 101:559–72.
Morrow, J., and H. Jo. 2006. Compliance with the laws of war: dataset and coding rules. *Conflict Management and Peace Science* 23:91–113.
Most, B., and H. Starr. 1984. International relations theory, foreign policy substitutability, and "nice" laws. *World Politics* 36:383–406.
Most, B., and H. Starr. 1989. *Inquiry, logic, and international politics.* Columbia: University of South Carolina Press.
Moyn, S. 2012. Substance, scale, and salience: the recent historiography of human rights. *Annual Review of Law and Social Science* 8:123–40.
Moyn, S. 2015. *Christian human rights.* Philadelphia: University of Pennsylvania Press.
Mudde, C., and C. Kaltwasser. 2017. *Populism : a very short introduction.* Oxford: Oxford University Press.
Mukherjee, R. 2022. *Ascending order: rising powers and the politics of status in international institutions.* Cambridge: Cambridge University Press.
Munck, G. 2009. *Measuring democracy: a bridge between scholarship and politics.* Baltimore: Johns Hopkins University Press.
Munro, E., N. Cartwright, and et al. 2016. *Improving child safety: deliberation, judgement and empirical research.* University of Durham: CHESS.
Murphy, G. 2002. *The big book of concepts.* Cambridge: MIT Press.
Nalepa, M., and E. Powell. 2016. The role of domestic opposition and international justice regimes in peaceful transitions of power. *Journal of Conflict Resolution* 60:1191–218.
Narziny, K. 2012. Anglo-American primacy and the global spread of democracy: an international genealogy. *World Politics* 64:341–73.
Nathan, L. 2006. Domestic instability and security communities. *European Journal of International Relations* 12:275–99.
Neundorf, A., et al. 2023. Varieties of political indoctrination in education and the media (V-Indoc): codebook. Manuscript.
Neundorf, A., et al. 2024. Varieties of indoctrination: the politicization of education and the media around the world. Manuscript.
Newman, K., et al. 2005. *Rampage: the social roots of school shootings.* New York: Free Press.
Newman, R. P. 1995. Ending the war with Japan: Paul Nitze's "Early Surrender" counterfactual. *Pacific Historical Review* 63:167–94.
Nielsen, S., et al. 2022. Unpacking context in realist evaluations: Findings from a comprehensive review. *Evaluation* 28:91–112.
Norris, P. 2012. *Making democratic governance work.* Cambridge: Cambridge University Press.
Nozick, R. 1974. *Anarchy, state and utopia.* Oxford: Basil Blackwell.
Nygård, H. 2017. The role of international organizations in regime transitions: how IGOs can tie a dictator's hands. *Conflict Management and Peace Science* 34:406–30.
Ohlson, T. 2008. Understanding causes of war and peace. *European Journal of International Relations* 14:133–60.
Olsen, T., et al. 2010. Transitional justice in the world, 1970–2007: insights from a new dataset. *Journal of Peace Research* 47:803–9.
Oneal, J. 1994. The affinity of foreign investors for authoritarian regimes. *Political Research Quarterly* 47:565–88.

Oneal, J., and B. Russett. 1999. Assessing the liberal peace with alternative specifications: trade still reduces conflict. *Journal of Peace Research* 36:423–42.
Osherson, D., and E. Smith. 1981. On the adequacy of prototype theory as a theory of concepts. *Cognition* 9:35–58.
Ostermann, S. 2016. Rule of law against the odds: overcoming poverty and the high cost of compliance in the developing world. *Law & Policy* 38:101–23.
Ostrom, C., and B. Job. 1986. The president and the political use of force. *American Political Science Review* 80:541–66.
Outhwaite, W., and S. Turner (eds.). YEAR. *The SAGE handbook of political sociology, vol. 1.* Newbury Park: Sage Publications.
Owsiak, A. 2020. ??Summary chapter ?? In S. Mitchell and J. Vasquez (eds.) *What do we know about war,* ?? edition. New York: Rowman & Littlefield.
Owsiak, A., and J. Vasquez. 2016. The cart and the horse redux: the timing of border settlement and joint democracy. *British Journal of Political Science* 49:339–54.
Palestini, S. 2021. Regional organizations and the politics of sanctions against undemocratic behaviour in the Americas. *International Political Science Review* 42:469–83.
Pape, R. 2003. The strategic logic of suicide terrorism. *American Political Science Review* 97:343–61.
Pape, R. 2005. *Dying to win: the strategic logic of suicide terrorism.* New York: Random House.
Pape, R., and J. Feldman. 2010. *Cutting the fuse: the explosion of global suicide terrorism and how to stop it.* Chicago: University of Chicago Press.
Parkkinen, V.-P., et al. 2018. *Evaluating evidence of mechanisms in medicine: principles and procedures.* New York: Springer Verlag.
Paxton, P. 2000. Women in the measurement of democracy: problems of operationalization. *Studies in Comparative International Development* 35:92–111.
Pazner, E., and D. Schmeidler. 1974. Difficulty in the concept of fairness. *Review of Economic Studies* 41:441–43.
Pearl, J., and D. MacKenzie. 2017. *The book of why: the new science of cause and effect.* New York: Basic Books.
Percy, S. 2007. Mercenaries: strong norm, weak law. *International Organization* 61:367–97.
Pettersson, T., and K. Eck. 2018. Organized violence, 1989–2017. *Journal of Peace Research* 55:535–47.
Pevehouse, J. 2002. Democracy from the outside-in? International organization and democratization. *International Organization* 56:515–49.
Pevehouse, J. 2005. *Democracy from above: regional organizations and democratization.* Cambridge: Cambridge University Press.
Pevehouse, J., T. Nordstrom, and K. Warnke. 2004. The Correlates of War international governmental organizations data version 2.0. *Conflict Management and Peace Science* 21:101–20.
Pierson, P. 2000. Three worlds of welfare state research. *Comparative Political Studies* 33:791–821.
Pion-Berlin, D., and R. Martnez. 2017. *Soldiers, politicians, and civilians: reforming civil-military relations in democratic Latin America.* Cambridge: Cambridge University Press.
Piplani V., and Talmadge C. 2016. When war helps civilmilitary relations: prolonged interstate conflict and the reduced risk of coups. *Journal of Conflict Resolution* 60:1368–94.
Posner, D. 2004. Measuring ethnic fractionalization in Africa. *American Journal of Political Science* 48:849–63.
Powell, J., et al. 2016. Combating coups détat in Africa, 1950–2014. *Studies in Comparative International Development* 51:482–502.
Powell, J., and C. Thyne. 2011. Global instances of coups from 1950 to 2010: a new dataset. *Journal of Peace Research* 48:249–59.
Powell, J., and M. Chacha. 2016. Investing in stability: economic interdependence, coups d'état and the capitalist peace. *Journal of Peace Research* 53:525–38.
Przeworski, A. 2019. *Crises of democracy.* Cambridge: Cambridge University Press.

Przeworski, A., et al. 2000. *Democracy and development: political institutions and well-being in the world, 1950–1990.* Cambridge: Cambridge University Press.
Przeworski, A., and F. Limongi. 1997. Modernization: theories and facts. *World Politics* 49:155–83.
Putnam, H. 1975b. Probability, explanation, and information. In (ed.) *Mind, language and reality: philosophical papers*, vol. 2. Cambridge: Cambridge University Press.
Putnick, D., and M. Bornstein. 2016. Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Developmental Review* 41:71–90.
Quackenbush, S. 2006. Identifying opportunity for conflict: politically active dyads. *Conflict Management and Peace Science* 23:37–52.
Quillian, L. 2006. New approaches to understanding racial prejudice and discrimination. *Annual Review of Sociology* 32:299–328.
Ragin, C. 2000. *Fuzzy-set social science.* Chicago: University of Chicago Press.
Ragin, C. 2008. *Redesigning social inquiry: fuzzy sets and beyond.* Chicago: University of Chicago Press.
Randolph, S., et al. 2010. Economic and social rights fufillment index: country scores and rankings. *Journal of Human Rights* 9:230–261.
Rathbun, B. 2011. Before hegemony: generalized trust and the creation and design of international security organizations. *International Organization* 65:243–73.
Ravallion, M. 2016. *The economics of poverty: history, measurement, and policy.* Oxford: Oxford University Press.
Rawls, J. 1971. *A theory of justice.* Cambridge: Harvard University Press.
Ray, J. 2013. War on democratic peace. *International Studies Quarterly* 57:198–200.
Reddy, S., and T. Pogge. 2005. How *not* to count the poor. Manuscript. Columbia University.
Rehm, P. 2008. The party manifesto data set: some critical remarks. Manuscript. Duke University.
Richards, R., and R. Smith. 2015. Statebuilding and the politics of non-recognition. In C. Daase et al. (eds.) *Recognition in international relations: rethinking a political concept in a global context.* London: Palgrave.
Riley, J. 1987. *Liberal utilitarianism: social choice theory and J.S. Mill's philosophy.* Cambridge: Cambridge University Press.
Ringen, S. 2011. The measurement of democracy: towards a new paradigm. *Society* 48:12–16.
Rio, A., et al. 2024. Education policies and systems across modern history: a global dataset. *Comparative Political Studies* 5:851–89.
Rocabert, J., et al. 2019. The rise of international parliamentary institutions: purpose and legitimation. *Review of International Organizations* 14:607–31.
Rodríguez-Franco, D. 2016. Internal wars, taxation, and state building. *American Sociological Review* 81:190–213.
Rodrik, D. 2007. *One economics, many recipes: globalization, institutions, and economic growth.* Princeton: Princeton University Press.
Roger, C., and S. Rowan. 2022. The new terrain of global governance: mapping membership in informal international organizations: codebook. Manuscript.
Roger, C., and S. Rowan. 2022. The new terrain of global governance: mapping membership in informal international organizations. *Journal of Conflict Resolution* 67:1248–269.
Rosato, S. 2003. The flawed logic of the democratic peace theory. *American Political Science Review* 97:585–602.
Rosenbaum, E. 2000. On measuring freedom. *Journal of Theoretical Politics* 12:205–27.
Rosenbaum, P. 2005. Heterogeneity and causality: unit heterogeneity and design sensitivity in observational studies. *American Statistician* 59:147–52.
Ross, J. 1993. Structural causes of oppositional political terrorism: towards a causal model. *Journal of Peace Research* 30:317–29.
Ross, M. 2004. How do natural resources influence civil war? Evidence from thirteen cases. *International Organization* 58:35–67.

Ross, M. 2012. *The oil curse: how petroleum wealth shapes the development of nations.* Princeton: Princeton University Press.
Royal Statistical Society. 2022. *Healthcare serial killer or coincidence? Statistical issues in investigation of suspected medical misconduct.* London: Royal Statistical Society.
Rueschemeyer, D., E. Stephens, and J. Stephens. 1992. *Capitalist development and democracy.* Chicago: University of Chicago Press.
Russett, B. 2017. A history of the *Journal of Conflict Resolution*. *Journal of Conflict Resolution* 61:1844–52.
Russett, B., et al. 1968. National political units in the twentieth century: a standardized list. *American Political Science Review* 62:932–51.
Russett, B., et al. 1998. The third leg of the Kantian tripod for peace: international organizations and militarized disputes, 1950–1985. *International Organization* 52:441–68.
Russo, F. 2011. Correlational data, causal hypotheses, and validity. *Journal for General Philosophy of Science* 42:85–107.
Sachs, J. et al. 2016. *SDG Index & dashboards.* New York: Bertelsmann Stiftung.
Sambanis, N., et al. 2020. External intervention, identity, and civil war. *Journal of Conflict Resolution* 53:2155–82.
Samii, C. 2016. Causal empiricism in quantitative research. *Journal of Politics* 78:941–55.
Sartori, G. 1976. *Parties and party systems: a framework for analysis.* New York: Cambridge University Press.
Saylor, R. 2014. *State building in boom times: commodities and coalitions in Latin America and Africa.* Oxford: Oxford University Press.
Schaeffer, N., and J. Dykema. 2020. Advances in the science of asking questions. *Annual Review of Sociology* 46:37–60.
Schauer, R. 2015. *The force of law.* Cambridge: Harvard University Press.
Schedler, A. 1998. What is democratic consolidation? *Journal of Democracy* 9:91–107.
Schmidt, R. 2024. A graphical method for causal program attribution in theory-based evaluation. *Evaluation* 30:380–407.
Schneider, C. 2024. *Set-theoretic multi-method research: a guide to combining QCA and case studies.* Cambridge: Cambridge University Press.
Schrank, A. 2013. Quantitative cross-national sociology and the methodological abyss: comment on Alcacer and Ingram. *American Journal of Sociology* 118:1099–111.
Schultz, K. 2001. *Democracy and coercive diplomacy.* Princeton: Princeton University Press.
Schweller, R. 1992. Domestic structure and preventive war: are democracies more pacific? *World Politics* 44:235–69.
Seawright, J. 2002a. Testing for necessary and/or sufficient causation: which cases are relevant? *Political Analysis* 10:178–193.
Seawright, J. 2002b. What counts as evidence? Reply. *Political Analysis* 20:204–7.
Seawright, J. 2016. *Multi-method social science: combining qualitative and quantitative tools.* Cambridge: Cambridge University Press.
Seawright, J., and J. Gerring. 2008. Case selection techniques in case study research: a menu of qualitative and quantitative options. *Political Research Quarterly* 61:294–308.
Sechser, T., and M. Fuhrmann. 2017. *Nuclear weapons and coercive diplomacy.* Cambridge: Cambridge University Press.
Sekhon, J. 2004. Quality meets quantity: case studies, conditional probability, and counterfactuals. *Perspectives on Politics* 2:281–93.
Sekhon, J. 2008. The Neyman-Rubin model of causal inference and estimation via matching methods. In J. Box-Steffensmier, H. Brady, and D. Collier (eds.) *The Oxford handbook of political methodology.* Oxford: Oxford University Press.
Sen, A. 1977. On weights and measures: informational constraints in social welfare analysis. *Econometrica* 45:1539–72.
Sen, A. 1992. *Inequality reexamined.* Cambridge: Harvard University Press.
Sen, A. 1999. *A measure of freedom.* Oxford: Oxford University Press.

Shadish, W., T. Cook, and D. Campbell. 2002. *Experimental and quasi-experimental designs for general causal inference.* Boston: Houghton Mifflin.
Shanks, C., et al. 1996. Inertia and change in the constellation of IGOs, 1981–92. *International Organization* 50:593–628.
Shannon, V. 2000. Norms are what states make of them: the political psychology of norm violation. *International Studies Quarterly* 44:293–316.
Sharman, J. 2015. War, selection, and micro- states: economic and sociological perspectives on the international system. *European Journal of International Relations* 21:194–214.
Siaroff, A. 1999. Corporatism in 24 industrial countries: meaning and measurement. *European Journal of Political Research* 36:175–205.
Siaroff, A. 2003. Comparative presidencies: the inadequacy of the presidential, semi-presidential and parliamentary distinction. *European Journal of Political Research* 42:287–312.
Signorino, C., and J. Ritter. 1999. Tau-b or not tau-b: measuring the similarity of foreign policy positions. *International Studies Quarterly* 43:115–44.
Simmons, B. 2009. *Mobilizing for human rights.* Cambridge: Cambridge University Press.
Simmons, B., and A. Danner. 2010. Credible commitments and the International Criminal Court. *International Organization* 64:225–56.
Simmons, B., and Z. Elkins. 2004. The globalization of liberalization: policy diffusion in the international political economy. *American Political Science Review* 98:171–90.
Simmons, E., and N. Smith (eds.). 2021. *Rethinking comparison: innovative methods for qualitative political research.* Cambridge: Cambridge University Press.
Singer, J. D., and M. Small. 1966. Formal alliances 1815–1939: a quantitative description. *Journal of Peace Research* 3:1–32.
Singh, S., and C. Way. 2004. The correlates of nuclear proliferation: a quantitative test. *Journal of Conflict Resolution* 48:859–85.
Skaaning, S.-E., et al. 2015. A lexical index of democracy. *Comparative Political Studies* 48:1491-525.
Slater, D., and E. Simmons. 2010. Informative regress: critical antecedents in comparative politics. *Comparative Political Studies* 43:886–917.
Slater, D., and E. Simmons. 2012. Coping by colluding: political uncertainty and promiscuous powersharing in Indonesia and Bolivia. *Comparative Political Studies* 46:1366–93.
Small, M., and J. D. Singer. 1982. *Resort to arms: international and civil wars, 1816–1980,* 2nd edition. Beverly Hills: Sage Publications.
Snow, D., and D. Cress. 2000. The outcomes of homeless mobilization: the influence of organization, disruption, political mediation, and framing. *American Journal of Sociology* 105:1063–1104.
Soifer, H. 2012. The causal logic of critical junctures. *Comparative Political Studies* 45:1572–97.
Staniland, P. 2021. Leftist insurgency in democracies. *Comparative Political Studies* 54:518–552.
Starr, H. 1996. *Anarchy, order, and integration: how to manage interdependence?* Ann Arbor: University of Michigan Press.
Stasavage, D. 2003. *Public debt and the birth of the democratic state: France and Great Britain 1688–1789.* Cambridge: Cambridge University Press.
Steele, L., et al. 2022. Measuring ethnic diversity. *Annual Review of Sociology* 48:43–63.
Stein, D., et al. 2010. What is a mental/psychiatric disorder? From DSM-IV to DSM-V. *Psychological Medicine* 40:1759–65.
Steinberg, D. A. 2015. *Demanding devaluation: exchange rate politics in the developing world.* Ithaca: Cornell University Press.
Stephens, E. 1989. Capitalist development and democracy in South America. *Politics and Society* 17:281–352.
Stephens, J. 1989. Democratic transition and breakdown in western Europe 1870–1939: a test of the Moore hypothesis. *American Journal of Sociology* 94:1019–77.
Stewart, M. 2021. *Governing for revolution: social transformation in civil war.* Cambridge: Cambridge University Press.

Stinchcombe, A. 1965. Social structure and organizations. In J. March (ed.) *Handbook of organizations.* Chicago: Rand McNally.
Stokke, O. 2004. Boolean analysis, mechanisms, and the effectiveness of international regimes. In A. Underdal and O. Young (eds.) *Regime consequences: methodological challenges and research strategies.* Boston: Kluwer Academic Publishers.
Stone, W., and L. Maisel. 2003. The not-so-simple calculus of winning: potential U.S. House candidates' nominations and general election results. *Journal of Politics* 65:951–77.
Suzumura, K. 1983. *Rational choice, collective decisions and social welfare.* Cambridge: Cambridge University Press.
Svolik, M. 2014. Which democracies will last? Coups, incumbent takeovers, and the dynamic of democratic consolidation. *British Journal of Political Science* 45:715–38.
Sweeney, K., and O. Keshk. 2005. The similarity of states: using S to compute dyadic interest similarity. *Conflict Management and Peace Science* 22:165–87.
Tallberg, J., et al. 2016. Democratic memberships in international organizations: sources of institutional design. *Review of International Organizations* 11:59–87.
Tamir, Y. 2019. *Why nationalism.* Princeton: Princeton University Press.
Tang, S., et al. 2017. Does oil cause ethnic war? Comparing evidence from process-tracing with quantitative results. *Security Studies* 26:359–90.
Tansey, O. 2017. The fading of the anti-coup norm. *Journal of Democracy* 28:145–56.
Tarman, C., and D. Sears. 2005. The conceptualization and measurement of symbolic racism. *Journal of Politics* 67:731–61.
Tarrow, S. 2021. Progress outside of paradise: old and new comparative approaches to contentious politics. *Comparative Political Studies* 54:1885–901.
Taylor, A. 1945. *The course of German history: a survey of the development of Germany since 1815.* London: Hamish Hamilton.
Teele, D. 2014. Ordinary democratization: the electoral strategy that won British women the vote. *Politics and Society* 42:537–61.
Teital, R. 2000. *Transitional justice.* Oxford: Oxford University Press.
Thiem, A., M. Baumgartner, and D. Bol. 2016. Still lost in translation: a correction of three misunderstandings between configurational comparativists and regressional analysts. *Comparative Political Studies* 49:742–74.
Thompson, W. 2001. Identifying rivals and rivalries in world politics. *International Studies Quarterly* 45:557–86.
Thyne, C. 2017. The impact of coups d'état on civil war duration. *Conflict Management and Peace Science* 34:287–307.
Tilly, C. 1978. *From mobilization to revolution.* Reading: Addison-Wesley.
Tilly, C. 2004. *Contention and democracy in Europe, 1650–2000.* Cambridge: Cambridge University Press.
Tilly, C., et al. 2004. *Contention and democracy in Europe, 1650–2000.* Cambridge: Cambridge University Press.
Tomz, M. 2007. *Reputation and international cooperation: sovereign debt across three centuries.* Princeton: Princeton University Press.
Treier, S., and S. Jackman. 2008. Democracy as a latent variable. *American Journal of Political Science* 52:201–17.
Tripp, A. 2015. *Women and power in postconflict Africa.* Cambridge: Cambridge University Press.
UNESCO Institute of Statistics. 2011. *International Standard Classification of Education ISCED 2011.* Montreal: UNESCO.
United Nations. 2010. *Monitoring peace consolidation United Nations practitioners' guide to benchmarking.* New York: United Nations.
Vabulas, F., and D. Snidal. 2013. Organization without delegation: Informal intergovernmental organizations (IIGOs) and the spectrum of intergovernmental arrangements. *Review of International Organizations* 8:193–220.

Valentino, B., et al. 2006. Covenants without the sword: international law and the protection of civilians in times of war. *World Politics* 58:339–77.
Valenzuela A. 1997. Paraguay: the coup that didn't happen. *Journal of Democracy* 8:43–55.
Vanhanen, T. 1999. Domestic ethnic conflict and ethnic nepotism: a comparative analysis. *Journal of Peace Research* 36:55–73.
Vanhanen, T. 2000. A new dataset for measuring democracy, 1810–1998. *Journal of Peace Research* 37:251–65.
Varian, H. 1974. Equity, envy, and efficiency. *Journal of Economic Theory* 9:63–91.
Varian, H. 1975. Distributive justice, welfare economics and the theory of fairness. *Philosophy and Public Affairs* 4:223–47.
Varshney, A. 2002. *Ethnic conflict and civic life: Hindus and Muslims in India.* New Haven: Yale University Press.
Vasquez, J. 2004. The probability of war, 1816–1992. *International Studies Quarterly* 48:1–27.
Vasquez, J., and A. Rundlett. 2016. Alliances as a necessary condition of multiparty wars. *Journal of Conflict Resolution* 60:1395–418.
Velleman, P., and L. Wilkensen. 1993. Nominal, ordinal, interval, and ratio typologies are misleading. *American Statistician* 47:65–72.
Vet, H. de, et al. 2011. *Measurement in medicine: a practical guide.* Cambridge: Cambridge University Press.
Vinjamuri, L., and J. Snyder. 2004. Advocacy and scholarship in the study of war crime tribunals and transitional justice. *Annual Review of Political Science* 7:345–62.
Volgy, T., et al. 2008. Identifying formal intergovermental organizations. *Journal of Peace Research* 45:849–62.
Vu, T. 2010. Studying the state through state formation. *World Politics* 62:148–75.
WHO. 2011. *World disability report.* WHO.
Waldner, D. 2015. What makes process tracing good? Causal mechanisms, causal inference, and the completeness standard in comparative politics. In A. Bennett and J. Checkel (eds.) *Process tracing in the social sciences: from metaphor to analytic tool.* Cambridge: Cambridge University Press.
Waldner, D. 2019. Causal mechanisms and qualitative causal inference in the social sciences. In M. Nagatsu and A. Ruzzene (eds.) *Contemporary philosophy and social science: an interdisciplinary dialogue.* London: Bloomsbury.
Waldner, D., and E. Lust. 2018. Unwelcome change: coming to terms with democratic backsliding. *Annual Review of Political Science* 21:93–113.
Wallace, G. P. R. 2015. *Life and death in captivity: the abuse of prisoners during war.* Ithaca: Cornell University Press.
Walter, B. 1997. The critical barrier to civil war settlement. *International Organization* 51:335–64.
Walter, B. 2003. *Committing to peace: the successful settlement of civil wars.* Princeton: Princeton University Press.
Walter, B. 2017. The new new civil wars. *Annual Review of Political Science* 20:469–86.
Walton, J. 1984. *Reluctant rebels: comparative studies of revolution and underdevelopment.* New York: Columbia University Press.
Warwick, P., and J. Druckman. 2006. The paradox of portfolio allocation: an investigation into the nature of a very strong but puzzling relationship. *European Journal of Political Research* 45:635–65.
Waters, C. 1998. Causal regularities in the biological world of contingent distributions. *Biology and Philosophy* 13:5–36.
Weatherson, B. 2002. Intrinsic vs. extrinsic properties. In E. Zalta (ed.) *Stanford Encyclopedia of Philosophy.*
Weeks, J. 2014. *Dictators at war and peace.* Ithaca: Cornell University Press.
Wei, S.-J. 2000. How taxing is corruption on international investors. *Review of Economic Studies* 81:1–11.

Weiner, M. 1987. Empirical democratic theory. In M. Weiner and E. Ozbudun (eds.) *Competitive elections in developing countries.* Durham: Duke University Press.
Weitzer, R. 2015. Human trafficking and contemporary slavery. *Annual Review of Sociology* 41:223–42.
Weyland, K. 2023. *Democracy's resilience to populisms threat: countering global alarmism.* Cambridge: Cambridge University Press.
Wickham-Crowley, T. 1991. *Guerrillas and revolution in Latin America: a comparative study of insurgents and regimes since 1956.* Princeton: Princeton University Press.
Widner, J. 2018. All hands on deck: the US response to west Africas ebola crisis, 2014–2015. Manuscript. Princeton University.
Widner, J. 2022. Descriptive accuracy in interview-based case studies. In J. Widner and et al. (eds.) *The case for case studies: methods and applications in international development.* Cambridge: Cambridge University Press.
Widner, J., and et al. (eds.). 2022. *The case for case studies: methods and applications in international development.* Cambridge: Cambridge University Press.
Wilfahrt, M. 2022. *Precolonial legacies in postcolonial politics: representation and redistribution in decentralized west Africa.* Cambridge: Cambridge University Press.
Wimmer, A. 2018. *Nation building: why some countries come together while others fall apart.* Princeton: Princeton University Press.
Wimmer, A., and B. Min. 2009. The location and purpose of wars around the world: a new global dataset, 1816–2001. *International Interactions* 35:390–417.
Wintrobe, R. 2006. *Rational extremism: the political economy of radicalism.* Cambridge: Cambridge University Press.
Woodward, J. 2004. *Making things happen: a theory of causal explanation.* Oxford: Oxford University Press.
World Justice Project. 2011–. *Rule of law index.* World Justice Project.
Wriglesworth, J. 1985. *Libertarian conflicts in social choice.* Cambridge: Cambridge University Press.
Yashar, D. 2005. *Contesting citizenship in Latin America: the rise of indigenous movements and the post liberal challenge.* Cambridge: Cambridge University Press.
Yashar, D. J. 2018. *Homicidal ecologies: illicit economies and complicit states in Latin America.* Cambridge: Cambridge University Press.
Zadeh, L. 1982. A note on prototype theory and fuzzy sets. *Cognition* 12:291–97.
Zellman, A., and J. Fox. 2024. Under God, indivisible? Religious salience and interstate territorial conflict. *Journal of Peace Research* 61:825–41.