

## POL 51: Scientific Study of Politics

James Fowler

[jhfowler@ucdavis.edu](mailto:jhfowler@ucdavis.edu)

TA: Aimee Lodigiani

TA: Nathan Hadley

## Introduction

- Political Science = Scientific Study of Politics
- Politics
  - "who gets what, when, and how" (Harold Lasswell)
- Scientific Study
  - "Accumulation and establishment of knowledge which has been systematized and formulated with reference to the discovery of general truths or the operation of general laws"

## Why research methods?

- Role as voters and/or future leaders must know how to evaluate research
- Role as a student—want to be able to write good papers (and get good grades!)

## Defining Scientific Research in Social Sciences

- Goal of scientific inquiry is *inference* – using observations from the world to learn about other unobserved facts
- Procedures should be *public*
- Conclusions are *uncertain*

## Two kinds of approaches

- Positive – what *is*?
- Normative – what *ought*?
- Normative is harder because you must defend a value judgment

## Three Kinds of Research

- Theoretical – from theory to observation (deductive)
- Empirical – from observation to theory (inductive)
- Applied – from observations and theories to applications and prescriptions

## A Brief History of Political Science as a Discipline

- Traditional Political Science
  - Historical approaches – detailed descriptions of political events
  - Legalistic approaches – study of constitutions and legal codes
  - Institutional approaches – study of powers and functions of political institutions
- Behavioralism – study of individual and group political behavior

## A Brief History of Political Science as a Discipline

- Postbehavioralism – we should focus more on prescriptive and policy issues
- New Institutionalism – mathematical study of powers and functions of political institutions
- Constructivism – related to post-modernism suggests that we can change behavior by redefining it

## Is Political Science a Science?

- Theory: humans act consistently in a discoverable manner
- Theory: human behavior is predictable and susceptible to general laws
- The observability problem: much human behavior is not directly observable

## Is Political Science a Science?

- The external awareness problem: behavior may change when subjects become aware of the researcher's purpose
- The measurement problem: unemployment, poverty, what is it?
- The complexity problem: no  $E=mc^2$
- Ethical problems: experimental approaches subject to preventing physical and emotional risk

## Two Styles of Research

- Quantitative and Qualitative
- One Logic of Inference
  - Descriptive Inference
  - Causal Inference

## "Interpretation" and Inference

- Interpretation involves not only human action but the reason for it
- Twitch or wink? Thin vs. thick description (Geertz)
- Susceptible to the same rules of inference
- Feelings and meanings may be important but science can only apply to their *observable* implications
- Soaking and poking (Fenno)

## Science and Uniqueness

- Many important events are unique – in fact all events are unique
- Distance between reality and thickest description is far greater than distance between thickest and thinnest description
- Must balance between general and specific
- Best way to understand a unique event is to use inference to study similar events

## Characteristics of Scientific Knowledge

- Empirically *verifiable* – objective observation or experience can be used to determine whether a statement is false
  - Non-scientific knowledge is based on inherently subjective observation or experience or not falsifiable
  - “Common sense” can be either
- Humans are good at detecting patterns bad at detecting nonpatterns

## Characteristics of Scientific Knowledge

- *Explanatory*
  - Prediction – “If X then Y”
  - Explanation = prediction + a reason for the prediction – “If X then Y because of Z”

## Characteristics of Scientific Knowledge

- *Transmissible* – others must be able to analyze and replicate the findings
  - Violence and TV – bad methods but helped science anyway
- *Generalizable* – explaining several phenomena is better than explaining just one
- *Provisional* – all knowledge is subject to revision and change – be skeptical!

## Deterministic vs. Probabilistic Explanations

- Deterministic explanations predict all phenomena with 100% accuracy—there are no counter-examples
- Probabilistic explanations predict some phenomena with <100% accuracy—there are potentially very many counter-examples

## Proposing Explanations

- Dependent variable (aka outcome variable)
  - “phenomenon thought to be caused, depend on, or to be a function of an independent variable”
- Independent variables (aka explanatory variables)
  - “measures of the phenomena that are thought to influence, affect, or cause some other phenomenon”

## Proposing Explanations

- Relationship between dependent and independent variable may be direct or indirect
- Antecedent variable – occurs prior to all other variables
  - Adequacy of insurance → attitude towards national insurance → vote choice

## Proposing Explanations

- Intervening variable – occurs closer in time to dependent variable and is affected by other independent variables
  - Education → civic duty → voter turnout
  - Education → political knowledge → turnout
- Key causal variable (aka treatment variable) – causal phenomenon of interest

## Proposing Explanations

- Control variables – phenomena from competing theories that may explain the relationship
- Causal explanations depend on (at least) 3 claims:
  - Dependent and independent variable covary
  - Change in independent variable *precedes* change in dependent variable
  - Covariance is not *spurious*

## Types of Explanations

- Hypothesis – statement that describes a relationship between two phenomena
- Theory – statement that organizes predicts and explains a general class of phenomena
- Law – theory that is very general and which has survived a lot of empirical verification

## Hypotheses Should Be

- *positive* not *normative*
- *general*
- *plausible* (but not necessarily obvious)
- *specific* (identify *direction* of relationship carefully *define* the phenomena)
- *consistent* with the data to be used
- *falsifiable* (testable) – avoid tautology (logical statement in which conclusion = premise)

## Formulating Hypotheses

- Should specify the units of analysis (individuals groups states agencies nations)
- Avoid *ecological fallacy*
- Make sure units are *consistent* (don't mix unless there is a clear link between levels)

## Defining Concepts

- Should be *accurate, precise, informative*
- Problems:
  - based on shared understandings
  - may be abstract
  - may be value-laden
  - definition of concept may alter its relationship to other phenomena

## Defining Concepts

- Solutions:
  - Literature Review (we'll learn how to do this later)
  - Avoid overreaching

## Defining Causality

- Fundamental Problem of Causal Inference
- Counterfactual – how would the dependent variable have been different if we kept all else equal (*ceteris paribus*) and changed the key causal variable?
- The degree of change in the dependent variable is the *realized causal effect*
- We must also include the possibility of randomness

## Clarifying Alternative Definitions of Causality

- "Causal Mechanisms"
- "Multiple Causality"
- "Symmetric" and "Asymmetric" Causality

## Assumptions Required for Estimating Causal Effects

- Unit homogeneity – phenomena should be similar
- Conditional independence – phenomena should be randomly selected

## Criteria for Judging Causal Inferences

- Unbiasedness – correct on average
- Efficiency – use as much information as possible to make the inference

## Campbell (1983)

- Ambiguity – uncertainty in issue positions
- Downsian tradition – based on Downs (1957) *An Economic Theory of Democracy*
- Issue position – left/right liberal/conservative raise/lower taxes
- Issue salience – how much do voters think about an issue and how much does it affect their vote?

## Campbell (1983)

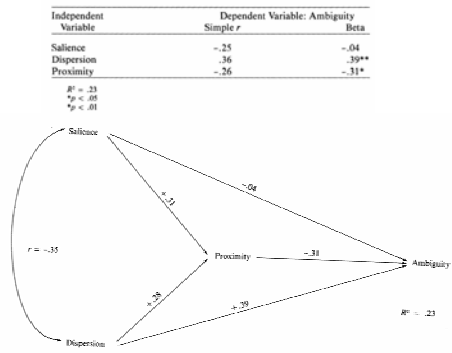
- Dispersion of public opinion – how widely are individuals scattered on an issue?
- Shepsle's theory: candidates are ambiguous to avoid offending voters
- Page's theory: ambiguity results from information constraints

TABLE 1  
Issues Included in the Analysis

Election	Subject of Issue	Candidates	Issue Salience	Opinion Dispersion	Proximity to		Ambiguity of	
					Democrat	Republican	Democrat	Republican
1968	Urban unrest	Humphrey/Nixon	46	1.95	1.22	0.36	1.54	1.61
1968	Vietnam War	Humphrey/Nixon	70	1.98	0.06	0.29	1.54	1.50
1972	Vietnam War	McGovern/Nixon	19	1.90	2.41	0.88	1.28	1.60
1972	Inflation	McGovern/Nixon	10	1.48	0.80	1.01	1.67	1.60
1972	Jobs	McGovern/Nixon	8	1.97	2.03	0.18	1.68	1.72
1972	Rights of accused	McGovern/Nixon	5	2.12	1.33	0.00	1.67	1.73
1972	Minority groups	McGovern/Nixon	8	1.97	1.82	0.09	1.51	1.55
1972	Taxes	McGovern/Nixon	3	2.29	0.89	0.56	1.89	1.76
1972	Urban unrest	McGovern/Nixon	5	2.09	0.20	1.35	1.59	1.77
1972	Campus unrest	McGovern/Nixon	0	1.86	1.75	0.03	1.71	1.38
1976	Rights of accused	Carter/Ford	0	2.12	0.80	0.61	1.65	1.77
1976	Busing	Carter/Ford	1	1.69	2.85	2.64	1.79	1.94
1976	Minority groups	Carter/Ford	2	1.98	1.16	0.32	1.53	1.62
1976	Health insurance	Carter/Ford	1	2.37	1.18	0.57	1.58	1.72
1976	Jobs	Carter/Ford	46	1.70	1.55	0.14	1.55	1.48
1976	Urban unrest	Carter/Ford	0	1.91	0.25	0.83	1.46	1.50
1976	Marijuana	Carter/Ford	1	2.17	1.05	0.94	1.50	1.48
1976	Taxes	Carter/Ford	6	2.25	0.86	0.25	1.70	1.59
1976	Women's rights	Carter/Ford	0	2.07	0.20	0.49	1.51	1.56
1980	Inflation	Carter/Reagan	72	1.49	0.27	0.44	1.47	1.36
1980	Defense spending	Carter/Reagan	18	1.50	1.67	0.40	1.47	1.50
1980	Government spending	Carter/Reagan	10	1.89	0.78	1.00	1.35	1.52
1980	Jobs	Carter/Reagan	26	1.88	1.29	0.61	1.41	1.49
1980	Russian relations	Carter/Reagan	3	1.83	1.45	0.70	1.47	1.54
1980	Minority groups	Carter/Reagan	1	1.59	1.45	0.55	1.41	1.29
1980	Women's rights	Carter/Reagan	0	1.93	0.27	2.03	1.42	1.83

Note: The issue salience measure is a percentage. The opinion dispersion and ambiguity measures are in terms of standard deviations. The proximity measures reported here are the difference between the median public position and median perception of the candidates' positions. The proximity scores used in this analysis are the negatives of those distances.

The Causes of Issue Ambiguity



## Thornberry & Christenson (1984)

- Etiology of crime and delinquency
  - Social control theory
  - Strain theory
  - Social learning theory
  - Integrated models
- All assume unemployment → crime
- What about the reverse?
- Exogenous vs. Endogenous

Table 1. Effect Parameters for Reciprocal Model of Unemployment and Crime\*

Relationship	Parameter	Unstandardized Coefficient	Standard Error	Standardized Coefficient
Unemployment Stability Coefficients	$a_1$	.621	.042	.559**
	$a_2$	.493	.043	.519**
	$a_3$	.493	.043	.523**
Crime Stability Coefficients	$b_1$	.287	.025	.472**
	$b_2$	.012	.058	.010
	$b_3$	.134	.049	.106**
Unemployment to Crime	$d_1$	.194	.105	.126*
	$d_2$	.685	.150	.351**
	$d_3$	1.261	.144	.485**
Crime to Unemployment	$c_1$	.062	.020	.157**
	$c_2$	.062	.020	.101**
	$c_3$	.062	.020	.128**
	$c_4$	.069	.010	.184**
	$c_5$	.069	.010	.119**
	$c_6$	-.021	.092	-.033
	$c_7$	-.021	.092	-.041
Goodness of Fit	$\chi^2$	38.49**		
	d.f.	14		
	$\chi^2/d.f.$	2.75		
	n	544		

\* For this model and the ones presented in the following tables, correlations among the residuals and the additional goodness-of-fit statistics are presented in the Appendix.

\*  $p < .10$ .

\*\*  $p < .05$ .

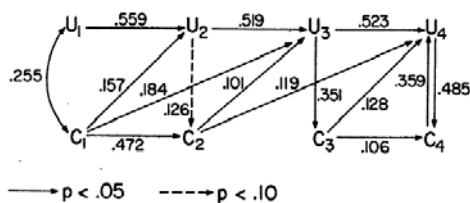


Figure 2. Significant Effects for Reciprocal Model of Unemployment and Crime (Standardized Regression Coefficients)

## Research Questions

- Questions come from
  - personal observation or experience
  - the research of others
  - interest in a broad theory
  - interest in career advancement
- Should be about *politics*
- Should not be too focused on facts – must ask something about a *relationship*
- Should be *positive* instead of *normative*
- The “so what” test

## Improving Research Questions

- Popper – discovery contains ‘an irrational element’ or a ‘creative intuition’
- Project should be consequential for
  - political social or economic life
  - understanding something that significantly affects many people’s lives
  - understanding and predicting events that might be beneficial or harmful

## Improving Research Questions

- Project should make a specific contribution to an identifiable scholarly literature by increasing our collective ability to construct verified scientific explanations of some aspect of the world

## What is a Literature Review?

- A detailed account of what has been published on a specific topic by scholars and researchers
- If separate then it is generally called an annotated bibliography
- More generally it is part of the beginning to a research paper monograph or thesis

## What is a Literature Review?

- General purpose
  - convey to readers current knowledge on the topic
  - Highlight gaps in knowledge
- Key goals
  - Criticism
  - Thoroughness
  - Readability

## Motivations

- See what has been investigated before
- Provide a general explanation for observed behavior or phenomenon
- Identify potential relationships
- See how others have measured concepts
- Identify potential sources of data
- Consider alternative research designs
- See relationships *between* studies

## Additional Motivations

- Note areas of agreement and conflict
- Identify specific gaps in knowledge
- Avoid wasting time

## Three General Phases

1. General scan of the literature
2. Focus on specific questions and issues
3. Organize a detailed hierarchical critique

## Six Detailed Steps

1. Given a particular research question identify the broad questions you need to ask
2. List the resources that can possibly give information on these questions
3. Search by: authors, keywords, phrases, etc.
4. Follow relevant citations to other works
5. Revise and reiterate continuing to narrow or broaden depending on success rate
6. Organize and classify results

## Questions About Your Review

- What is the specific area of interest corresponding to the research question?
- What type of studies are important?
  - Theoretical?
  - Methodological?
  - Cross-cutting?
- What is the scope of the study across fields, time, and focus?
- Is the topic too broad? Too narrow?

## Questions About Your Review

- Is there a *critical analysis* of the literature instead of a simplistic *description*?
- Are linkages between works and groups of works identified?
- Are foundational works included? Are recent works included?
- Is this something that readers can *actually use* to understand the current level of knowledge?



## Questions About Specific Works

- Is the key problem/question identified?
- Is the significance and importance of the project justified?
- Does the author approach the questions from the best theoretical perspective?
- What is the general research approach: review interpretation criticism addition?

## Questions About Specific Works

- Does the literature review place the work in proper context? Does it give multiple perspectives?
- How effective is the study design?
- Is the language persuasive? Subjective? Normative? Dispassionate?
- Is the measurement valid?

## Questions About Specific Works

- Is the methodology appropriate?
- Do the results seem justified by the process?
- How do the conclusions add to our knowledge?
- How does this work enhance my literature review?

## Where to Look

- Campus Only Resources:
  - webofscience.com
  - jstor.org
  - web.lexis-nexis.com/universe
- citeseer.ist.psu.edu & other websites
- government websites (Thomas, Census, BLS, FEC)
- syllabi and scholar's web sites
- books (libraries!)

## Using Citations

- An author's citations indicate relevant connected works
- JR: "pyramid citations"
- More useful analogy: "network of citations"
- Be conscious of dates and places

## Political Science Citation Format

- Aberbach, Joel D. and Bert A. Rockman. 1995. "The Political Views of US Senior Federal Executives 1970-1992." *Journal of Politics* 57, 838-52.
- Bryner, G.C. 1987. *Bureaucratic Discretion*. Elmsford NY: Pergamon Press.
- Blah blah (Bryner 1987)
- (Aberbach and Rockman 1995; Bryner 1987)
- Bryner (1987) suggests that ...
- Aberbach and Rockman (1995 p.840) note ...

## Nelson Review

- Purpose: Given the distinction between sex and gender how do women fit into the polity when gender is itself a complex constituent of the polity?
- How can these patterns be compared across and within nations institutions social movements and groups of people?

## Nelson Review

- Foundational work: Duverger's *The Political Role of Women*
  - First “behavioralist cross-national comparison of women's electoral participation to use both election and survey data”
- Findings
  - women vote less frequently than men
  - were slightly more conservative
  - when married tended to vote like their husbands

## Nelson Review

- Failings
  - Does not develop a general theory of the political and social relationships between men and women
  - Does not consider regime and economic factors
- Nelson reviews six recent works in a literature review format
  - Four from politics
  - Two from economics

## Nelson Review

- Alvarez, *Engendering Democracy in Brazil*
  - Descriptive focus on women's movements during Brazil's transition from military control to democratic government
  - Differentiates
    - feminine groups (focused on immediate needs)
    - feminist groups (focused on more fundamental societal changes)
  - Women's groups faced the challenge of redefining their role in the more open society: internal disagreement about whether to
    - remain independent or
    - define their role in political parties (necessitating compromise)
  - Some groups sought dual roles of militancy and integration

## Nelson Review

- Watson, *Playing the State*
  - Addresses the conflict that occurs when women's groups must decide whether or not to participate in the established political process
  - Since most women's groups come from center-left political perspectives this integration causes internal friction
  - Contends that in Australia women's groups “played the state” better than any other WID by entering government and creating permanent institutions

## Nelson Review

- Chou, Clark, and Clark, *Women in Taiwan Politics*
  - Analyzes female politicians in post-independence Taiwan
  - Describes complex system for allocating seats to women
- Nelson critique
  - Like Duverger authors miss theoretical issue of women's groups integration into the structures of government
  - Do not describe why electoral quota system is needed

## Nelson Review

- West and Blumberg, *Women & Social Protest*
  - A (heterogeneous) collection of more sociological perspectives that conceptualize women's social movements by “goal location and strategy as well as opportunity structures and resources”
  - “a wide variety of units of analysis”
    - For instance (Schmid chapter) the Green Party of Germany practiced equal male/female representation on party lists (those standing for office) which apparently influenced other European party practice

## Nelson Review

- Tinker, *Persistent Inequalities*
  - A “festschrift” to honor the twentieth anniversary of Ester Boserup's *Women's Role in Economic Development* (introduced the study of women in development)
  - Boserup's classic work differentiated between
    - female farming (small-scale and direct)
    - male farming (more oriented towards the technologies of production)
  - notes that when the former is ignored development programs are less likely to succeed

## Nelson Review

- Kardam, *Bringing Women In*
  - Looks at how three major development programs dealt with issues associated with bringing more women into development programs
  - Can institutionalized cultures at these agencies respond to demand for more women-centered program giving?
  - Depends on other political demands such as
    - outside constituencies
    - quantifiable accountability
    - varying sensitivity to local politicians
- Nelson: misses opportunity for broad theorizing

## Discussion Questions

- What is the key problem/question identified?
- Does Nelson explain the significance/importance?
- What is the general research approach?
- Are dissenting perspectives included or is it unitary?
- Is it comprehensive or selective?
- What is the theoretical perspective?
- Does Nelson address the questions that Duverger suggested but could not resolve?
- Does this add to our knowledge? How?

## Experiments

- Research in which researcher controls the independent variable
- Example: Does self interest affect how people allocate resources?
- Ultimatum Game

## Political Science Experiments

- Voting behavior
- Games
  - Prisoner's Dilemma, Public Goods, Coordination, Bargaining, Markets, Auctions, Individual Choice
- International Relations
- Committees
- Media
- Leadership

## Ideal Experiments

- *Experimental group* receives a *test stimulus*
- *Control group* does not receive the stimulus
- Researcher controls *group composition*
- Researcher controls *treatment*

## Ideal Experiments

- Researcher identifies *experimental effect* by measuring dependent variable before and after stimulus
- Researcher controls *extraneous factors* that might affect dependent variable

## Experiments and Validity

- Internal validity – whether variation in the treatment (and not something else) is causing change in the dependent variable

## Threats to Internal Validity

- History
- Maturation
- Testing
- Regression to the mean
- Experimental mortality
- Instrument decay
- Selection bias

## Experiments and External Validity

- External validity – generalizability of the causal effect to other experimental or non-experimental contexts
- Threats to external validity
  - Stakes
  - Unrepresentative subject group (sophomores professionals)
  - Hawthorne effect

## Assignment of Subjects

- Random assignment
- Precision matching
- In practice assignment is usually biased

## Advantages of Experimentation

- Ability to derive causal inferences
- Experimental control
- Precise measurement
- Ability to explore details of experiments
- Relative economy

## Disadvantages of Experimentation

- Artificial environment
- Unrepresentative subject pools
- External validity
- Experimenter bias

## Research Design

- Plan that shows how researcher intends to fulfill goals of planned study
- As important as questions hypotheses and measurement
- Must optimize design on two elements:
  - Theory: must be able to answer the question
  - Practice: must be possible, ethical

## Experimental Designs

- Simple post-test design
  - Assignment → treatment → measurement
  - Sensitive to pre-treatment differences
- Pre-test/post-test (classic) design
  - Assignment → measurement → treatment → measurement
  - Sensitive to *instrument reactivity*

## Experimental Designs

- Time series design
  - Assignment → measurements  $1 \dots m$  → treatment → measurements  $m+1 \dots n$
  - But expensive and time series is messy
- Multigroup design
  - Assignment to control and several treatment groups → measurement → unique treatment for each group → measurement
  - Reduces sample sizes

## Experimental Designs

- Factorial design
  - Assignment to control and several treatment groups → measurement → unique treatment *combination* for each group → measurement
  - Can *quickly* reduce sample size
- Solomon Four-Group Design
  - Assignment groups → measurement of one treatment one control group → treatment for both treatment groups → measurement
  - Controls for instrument reactivity but expensive

## Field Experiments

- Experimental design in a “natural” setting
- Better external validity
- Worse internal validity
  - Measurement problems
  - Uncontrolled environment
- Ethical issues

TABLE 1  
DESCRIPTION OF THE EXPERIMENTAL DISTRICTS

District	Exp. Candidate	Exp. Villages	Treatment	Ethnicity
Kandi	Kerekou	Kassalou	clientelism	Bariba (92%)
		Keten	public policy	Bariba (90%)
Nikki	Kerekou	Oucouou	clientelism	Bariba (89%)
		Kpawolou	public policy	Bariba (88%)
Bembereke	Saka Lafia	Bembereke Est	clientelism	Bariba (86%)
		Wannarou	public policy	Bariba (88%)
Perere	Saka Lafia	Tisserou	clientelism	Bariba (93%)
		Alafiarou	public policy	Bariba (94%)
Abomey-Bohicon	Soglo	Agnangnan	clientelism	Fon (99%)
		Guidjazoun	public policy	Fon (99%)
Ouidah-Pahou	Soglo	Acadjame	clientelism	Fon (99%)
		Ahozon	public policy	Fon (99%)
Aplahoue	Amoussou	Boloume	clientelism	Adja (99%)
		Avetoume	public policy	Adja (96%)
Dogbo-Toviklin	Amoussou	Dekoundji	clientelism	Adja (99%)
		Avedjin	public policy	Adja (99%)
Parakou	Ker./Lafia	Guena	competition	Bariba (80%)
		Thiam	competition	Bariba (82%)
Come	Am./Soglo	Kande	competition	Adja (90%)
		Tokan	competition	Adja (95%)

TABLE 2  
COMPARING VOTING BEHAVIOR OF SURVEY RESPONDENTS WITH  
AGGREGATE VOTING BEHAVIOR IN EXPERIMENTAL DISTRICTS<sup>a</sup>

District	Candidate	Exp. Villages	Reg. Voters	Sample Size	Sample Mean	Population Mean
Kandi	Kerekou	clientelism	1133	61	1.00 (0)	0.81
		public policy	1169	60	0.49 (4.50)	0.60
Nikki	Kerekou	control	3886	61	0.96 (1.15)	0.75
		clientelism	462	60	0.95 (2.21)	0.90
Bembereke	Lafia	public policy	1090	60	0.93 (2.43)	0.85
		control	2979	60	0.95 (2.00)	0.82
Perere	Lafia	clientelism	999	60	0.92 (2.63)	0.94
		public policy	931	60	0.89 (3.30)	0.93
Abomey	Soglo	control	5204	61	0.91 (2.85)	0.74
		clientelism	657	59	0.76 (4.22)	0.81
Ouidah	Soglo	public policy	442	60	0.13 (3.35)	0.25
		control	4477	61	0.52 (4.00)	0.58
Aplahoue	Amoussou	public policy	1172	60	0.98 (1.13)	0.91
		control	1199	60	0.98 (1.13)	0.90
Dogbo	Amoussou	clientelism	5204	61	0.74 (1.15)	0.86
		public policy	321	60	0.93 (2.5)	0.86
Aplahoue	Amoussou	control	701	61	0.92 (2.26)	0.72
		clientelism	2414	60	0.73 (0.44)	0.64
Dogbo	Amoussou	control	492	59	0.98 (1.13)	0.87
		public policy	511	60	0.91 (2.85)	0.77
Dogbo	Amoussou	control	4037	61	0.98 (2.20)	0.72
		clientelism	1397	60	0.64 (0.8)	0.65
Dogbo	Amoussou	public policy	736	61	0.50 (4.50)	0.47
		control	1161	59	0.45 (0.44)	0.84

<sup>a</sup>The vote choice variable takes the value 1 if the voter of a given type (male or female) chooses a candidate of a given type (for example, northern or southern) running experimental platforms in his or her district and 0 otherwise.

TABLE 3  
DIFFERENCE IN MEANS BETWEEN TREATMENT AND CONTROL  
VILLAGES FOR EACH TYPE OF CANDIDATE<sup>a</sup>

Type of Candidate <sup>b</sup>	Public	Clientelist	Control	Public-Control	Clientelist-Control
Northern	.322 (.032)	.674 (.032)	.565 (.035)	-.243 (.048)***	.109 (.047)**
Southern	.840 (.025)	.890 (.021)	.741 (.029)	.099 (.039)***	.149 (.036)***
Incumbent	.691 (.032)	.897 (.021)	.815 (.027)	-.141 (.042)***	.062 (.033)*
Opposition	.493 (.033)	.681 (.033)	.509 (.033)	-.015 (.047)	.172 (.045)***
Local	.225	.232	.230		
National	.385 (.032)	.603 (.033)	.509 (.033)	-.124 (.046)***	.094 (.047)**
	.226	.224	.230		
	.816 (.027)	.968 (.012)	.835 (.027)	-.019 (.038)	.133 (.028)***
	.201	.222	.194		

\*\*\* significant at 99%; \*\* significant at 95%; \* significant at 90%.  
<sup>a</sup> The first column presents the types of candidates the following three columns present the means of voting behavior in the public policy and the clientelist treatment groups and the control groups, respectively. Sample sizes are indicated right below the means, and standard errors are in parentheses. The last two columns present the public policy treatment effect and the clientelist treatment effect. The vote choice variable takes the value 1 if the voter of a given type (male or female) chooses a candidate of a given type (for example, northern or southern) running experimental platforms in his or her district and 0 otherwise.  
<sup>b</sup> Northern is Kerekou, Lafia; southern is Soglo, Amoussou; incumbent is Kerekou, Amoussou; opposition is Soglo, Lafia; national is Kerekou, Soglo; regional is Amoussou, Lafia.

TABLE 5  
PROBIT ANALYSIS OF VOTE FOR TYPE OF CANDIDATES IN  
TREATMENT VILLAGES

	Southern	Northern	Local	National	Incumbent	Opposition
Constant	-0.946** (0.395)	-0.513 (0.374)	-0.367 (0.306)	-0.741 (0.469)	-0.186 (0.415)	0.222 (0.271)
Sex	-0.513* (0.200)	-0.516*** (0.194)	-0.424 (0.330)	-0.828** (0.332)	-0.415 (0.370)	0.024 (0.231)
Age	0.006 (0.006)	-0.003 (0.005)	-0.009* (0.005)	0.011* (0.006)	0.004 (0.006)	0.002 (0.005)
Past	2.139*** (0.203)	.865*** (0.235)	1.555*** (0.201)	2.057*** (0.271)	1.893*** (0.180)	0.966*** (0.215)
Public policy	0.309*** (0.333)	-0.372*** (0.365)	-0.594* (0.318)	0.429 (0.427)	-0.287 (0.387)	0.512* (0.290)
Clientelist	1.004** (0.447)	0.264 (0.391)	0.444 (0.342)	0.550 (0.457)	0.344 (0.468)	0.754** (0.319)
Sex*Client.	-0.502 (0.505)	-0.191 (0.435)	-0.348 (0.379)	0.489 (0.548)	0.208 (0.539)	-0.324 (0.364)
Sex*Public Pol.	0.167 (0.402)	-1.050** (0.414)	0.147 (0.358)	-0.572 (0.482)	-0.111 (0.450)	-0.773** (0.345)
N	524	543	596	510	472	602
log-L	-145.250	-208.538	-284.0500	-115.986	-146.161	-244.583

\*\*\* significant at 99%; \*\* significant at 95%; \* significant at 90%

## Nonexperimental Designs

- Any research in which you:
  - study a single group
  - lack assignment control
  - lack of control over independent variable
- Not ideal but much data is available
- Better external validity, worse internal validity

## Nonexperimental Design Types

- Time series design
  - Consecutive measures before (during) and after independent variable introduced
- Cross-sectional design
  - Simultaneous measures of dependent and independent variable
- Panel study design
  - Time series + cross-section
- Case study design

## Rules for Constructing Causal Theories

- Construct Falsifiable Theories
- Build Theories That Are Internally Consistent
- Select Dependent Variables Carefully
- Maximize Concreteness
- State Theories in as Encompassing Ways as Feasible

## Improving Theory

- Must generate many observable implications
- Be concrete
- Parsimony – explain as much as possible with as little as possible

## How to Contribute to a Literature

- Test a hypothesis seen as important by scholars but not yet tested
- Test an accepted hypothesis
- Resolve or provide further evidence on one side of a controversy
- Test unquestioned assumptions

## How to Contribute to a Literature

- Study an important topic that has been overlooked
- Show that theories or evidence designed for some purpose in one literature could be applied to another literature to solve an existing but unrelated problem

## Measurement Reliability

- Concerns the extent to which an experiment, test, or any measuring procedure yields the same results on repeated trials
- “the more consistent the results given by repeated measurements the higher the reliability of the measuring procedure” (Carmines and Zeller 1979)

## Methods of obtaining reliability

- Test-retest method – measure, repeat, look for consistency
- Alternative form method – measure using different instruments
- Split-halves method – measure using two different instruments on a split of the data/variables/concepts

## Measurement Validity

- Degree to which a measure approximates the targeted concept

## Types of Validity

- Construct validity – claiming validity of some measure by similarity to another measure with known high validity
- Content validity – the full domain of the concept is measured (all dimensions/aspects)
- Face validity – the superficial type: “it looks like the concept”
- Interim association – correlation analysis across all levels

## Levels of Measurement

- Nominal – categories
- Ordinal – ranks
- Interval – fixed distance between levels
- Ratio – numerical

## Levels of Measurement

- Indices – multi-item measures from accumulating scores to create a composite variable
  - Usually to measure not directly observable concepts: democracy peace access to institutions and so on
- Scales – Multi-item measures that originate at the data source

## Types of Scales

- Likert scale
  - a symmetric attitudinal measure of agreement/disagreement
- Guttman scale
  - respondents are presented in increasingly hard to agree-with statements until they disagree



## Types of Scales

- Semantic differential
  - respondents get a scale between two opposing adjective pairs (idealistic/realistic) for describing a subject
- Factor Analysis
  - a statistical technique that loads variables into new variables (factors) based on serial explanation of variance

## Improving Data Quality

- *Record and report* how data are generated
- Collect data on as many *observable implications* as possible
- Maximize the *validity* of measurements
- Ensure data-collection method is *reliable*
- All data and analysis should be *replicable*

## Improving the Use of Existing Data

- Method should be unbiased
- Method should be efficient

## Niemi Craig Mattei (1991)

- Internal efficacy – beliefs about one’s own ability to understand and participate effectively in politics
- External efficacy – beliefs about responsiveness of government to citizen demands
- How to measure?

## Niemi Craig Mattei (1991)

1. "I consider myself to be well qualified to participate in politics." (SELQUAL)
2. "I feel that I have a pretty good understanding of the important political issues facing our country." (UNDERSTAND)
3. "I feel that I could do as good a job in public office as most other people." (PUBOFF)
4. "I think that I am better informed about politics and government than most people." (INFORMED)
5. "Sometimes politics and government seem so complicated that a person like me can't really understand what's going on." (COMPLEX)
6. "People like me don't have any say about what the government does." (NOSAY)
7. "I don't think public officials care much what people like me think." (NOCARE)

## Niemi Craig Mattei (1991)

Table 1. Polychoric Correlations among Original Center for Political Studies and New Efficacy Items, by Placement in Questionnaire

	NOCARE	COMPLEX	UNDERSTAND	SELQUAL	PUBOFF	INFORMED
Grouped <sup>a</sup> (N = 849)						
NOSAY	.65	.44	.24	.21	.04*	.20
NOCARE	—	.46	.19	.19	.01*	.12
COMPLEX	—	—	.34	.43	.26	.36
UNDERSTAND	—	—	—	.60	.46	.56
SELQUAL	—	—	—	—	.59	.59
PUBOFF	—	—	—	—	—	.61
Mixed <sup>b</sup> (N = 873)						
NOSAY	.51	.33	.19	.19	.08	.14
NOCARE	—	.31	.13	.12	.00*	.10
COMPLEX	—	—	.32	.39	.25	.37
UNDERSTAND	—	—	—	.52	.50	.63
SELQUAL	—	—	—	—	.56	.64
PUBOFF	—	—	—	—	—	.60

Source: 1988 National Election Study.

Note: Items recoded, as necessary, to account for direction of wording. Listwise deletion of missing data.

\*All efficacy items administered together as a group.

<sup>b</sup>All Likert items randomly "scrambled" and administered in several batteries.

\*p > .05.

## Niemi Craig Mattei (1991)

Internal Efficacy		Political Participation	
COMPLEX	.54		
UNDERSTAND	.72 (.73)	PARTICIP	1.00
SELFQUAL	.77 (.77)	EDUCATION	1.00
PUBOFF	.61 (.64)		
INFORMED	.81 (.83)		
External Efficacy		Score	Percent
NOSAY	.65	0-1	7.5
NOCARE	.70	2-3	9.8
ELECRESP	.66	4-5	15.1
GOVRESP	.68	6-7	17.2
Political Trust		8-9	19.5
TRUSTDC	.72	10-11	15.1
WASTETAX	.51	12-13	9.8
BIGINT	.80	14-16	5.9
CROOKED	.65		
Psychological Involvement			
INTEREST	.87		
TALKPOLIT	.68		
TELEVISION	.75		

## Niemi Craig Mattei (1991)

Table 2. External Validity of New Internal Efficacy Scale

	External efficacy	Political trust	Psychological involvement	Political participation	Education
Internal efficacy	.26	.09	.56	.40	.20
External efficacy	—	.71	.35	.32	.20
Political trust	—	—	.07	.15	.08
Psychological involvement	—	—	—	.44	.19
Political participation	—	—	—	—	.21

Source: 1988 National Election Study.

Note: Figures in this table are correlations with the factors identified in the confirmatory factor analysis ( $p < .05$ ). Positive correlations reflect higher involvement, participation, and education among the more efficacious or trusting. Internal efficacy is based on the four new items only.  $N = 1,271$ . Listwise deletion of missing data.

## Causation

- Conditional Independence
  - assumption that observations are chosen and values assigned to explanatory variables independently of observed outcome variable values: non-endogeneity
- Unit Homogeneity
  - assumption that if two observations have the same value of the key explanatory variable then they have the same expected value of the outcome variable
- Causal inference fails when neither is met

## Selection Bias

- There must be some variation in the outcome variable
- Selection rules correlated with outcome variable lead to biased estimates
- Real world can induce selection bias
  - minor parties may choose not to run in districts where they have no chance of winning
  - support underestimated
- Selection by including all levels of the key causal variables does not introduce bias

## Selection No-Nos

- Bad
  - selecting observations where the key causal variable does not vary
  - Example: Link between occupational prestige and industrialization
  - Rossi (1956) only looked at industrialized countries and inferred a linkage
  - Zerlditch (1971) found the same linkage in non-industrialized countries

## Selection No-Nos

- Very Bad
  - selecting observations where the outcome variable does not vary
  - Examples: coups voting retirements/resignations coalition collapses presidents that lost second election disease presence (retrospective)

## Selection No-Nos

- Even worse
  - selection of observations based on levels of both the explanatory variables and the outcome variable
  - Example: investigating countries with autocratic governments and high levels of political violence

## Sampling

- A *sample* is a subset of a *population*
- An *element* is the unit of analysis
- A *sampling element* or *sampling unit* is one that becomes part of the sample
- Sampling frame
  - the segment of the population who are able to participate given the sampling design
- Sample statistics vs population (descriptive) statistics

## Sampling

- Stratum
  - a population subgroup that shares some characteristic of interest
- Sampling Bias
  - occurs when the sampling frame omits relevant members of the population
- Inference pertains only to samples
  - But we can extend inference to population if we know sample properties

## Sampling Uncertainty

- Margin of error

	Sample Size	Confidence Interval
– Range around a sample statistic within which population estimate is likely to fall	4000	±2%
	1500	±3%
	1000	±4%
- Confidence level

– Probability population estimate falls within margin of error	600	±5%
	400	±6%
	200	±8%
	100	±11%

## Probability Sampling

- Probability Sampling
  - sample units selected from the sampling frame according to some probabilistic scheme
- Simple Random Sample (SRS)
  - each element of the sampling frame has an equal opportunity of being included in the sample

## Probability Sampling

- SRS procedures usually randomized by computer
- SRS advantage: randomizes across subgroups
- SRS disadvantage: difficulty in ensuring randomness
  - Alf Landon vs FDR in 1936 (Readers Digest polled 10M auto owners phone owners and magazine subscribers)

## Systematic Sampling

- Elements in the sample frame are listed and chosen by moving some interval length possibly “looping around”
- Sampling interval chosen to exhaust the sample frame and still produce the desired sample size:
  - size of sampling frame  $N$  desired sample size  $n$  so sample every  $k$ th element where  $k=N/n$
- Starting point is a randomly chosen element
- Advantage: easy
- Disadvantage: alphabetic or other listing biases

## Stratified Sampling

- Group sampling frame elements according to categories of one characteristic and sample from each group separately
- Goal: homogeneity within groups heterogeneity between groups
- Can also stratify by multiple characteristics

## Stratified Sampling

- Proportionate version
  - draw the sampling fraction (desired sample size over the sampling frame size) from each strata (either with SRS or systematic)
- Disproportionate version
  - over-drawing from underrepresented groups in the sample frame (can adjust results with weighting later if desired)
- Advantage: allows for full representation
- Disadvantage: requires full list of sampling frame

## Cluster Sampling

- Suppose we do not have a convenient list of sampling frame elements
- Define natural identifiable groupings (eg clusters) and sample from within each cluster
- Groups are heterogeneous but identifiable

## Cluster Sampling

- Classic example
  - draw random city blocks and sample households from within the blocks
- Advantage: overcomes list problem clustering design very flexible
- Disadvantage: may produce more sampling error due to multi-stage design

## Other Sampling Plans

- Snowball Sample
  - use respondents to generate other respondents
  - Useful for rare phenomena social phenomena
- Quota Sample
  - choose probabilistically by subgroups non-probabilistically
  - 1948 Gallup: Dewey over Truman

## Other Sampling Plans

- Purposive Sample
  - intentional choice of interesting cases
- Convenience Sample
  - easy to select cheap
- Iowa Caucus “cess poll”

## Small $N$ Sampling

- Sampling important for *both* quantitative *and* qualitative work
- In fact sampling is *more* important as  $N \rightarrow 0$ 
  - Easier to find non-random relationships
  - Easier to over-generalize

## Case or Observation?

- The word “case” insufficiently descriptive
  - a single unit with multiple variables (Pakistan) or
  - a single variable (Pakistani GDP in 1999)
- Aggregation issue with cases: India 1950-1990
  - India in a given year
  - regions/states of India
  - Indian voters
  - Which is the right level for a case study of Indian politics?

## Case or Observation?

- The word “observation” is better
  - one measure on one unit for one outcome variable includes all corresponding measures of the outcome variables
  - Forces us to be specific about what an observation really is

## Selection in Small $N$ Studies

- Suppose
  - an outcome variable  $X$  can take on three different values: low medium high
  - You can only collect two observations
- Three possibilities:
  - (High, Med); (Med, Low); (High, Low)
- Only (High, Low) selection rule is not correlated with levels of the outcome variable

## Indeterminant Research Design #1

- More Inferences Than Observations
- Rule: one fact/datum/observation cannot give information about more than one implication
- Therefore more observations than inferences are required
- Often a problem with comparative politics studies

## Indeterminant Research Design #2

- Multicollinearity
- There always exists overlap between the explanatory power of explanatory variables
- If two explanatory variables are perfectly correlated then the model fails
- If two explanatory variables are highly correlated then variances will be very high
- Problem often exacerbated by IRDN1
- Same problem as more parameters than observations

## Qualitative Diagnostics

- Typically when a qualitative model suffers from IDRD1 or IDRD2 then it is up to the researcher to detect the problem
- Symptoms:
  - Aspects of outcome variable behavior unexplained by levels of explanatory variables
  - Substantial redundancy of effects
  - A new observation ruins existing explanations/model
  - Readers/reviewers can easily think of counterfactuals

## Intentional Selection of Observations

- Best scenario: experiments
  - You control the levels of the explanatory variables through selection and randomization
- More generally values are assigned by history
  - Selection must be made with *intention* to randomize even if absolute randomization is not possible
- Case Control Study
  - retrospective research in which observations with particularly interesting values of the outcome variable are evident
  - Interesting but not able to make inferential claims

## Fenno 1977

- What is the key problem/question identified?
- Does Fenno explain the significance/importance?
- What is the general research approach?
- What is the theoretical perspective?
- Does this add to our knowledge? How?

## Sample Exam Questions

- Short answer
  - What is a reciprocal causal structure? Give an example of this type of structure from your readings
  - What is a theory? Why are competing theories a useful tool in political science?
- Long answer
  - Compare two of the articles from the class syllabus in terms of their research design. What are their hypotheses? How do the authors test those hypotheses? What type of research design do they employ? Suggest an alternative design that may achieve better results

## Counterfactuals and Inference

- Suppose hypothesis  $C \rightarrow E$
- Two strategies to falsify:
  - Counterfactual: *imagine* whether E might have occurred had C been absent
  - Actual: search for similar case in which C is absent to see if E occurs
- Same statistical problem—one observation one explanatory variable so more cases are needed

## Counterfactuals and Inference

- Example: international not domestic factors caused major aspects of Soviet foreign policy
- Range of variation is important so researcher must be explicit about what might have happened
- But what do we know about things that did not happen?

## Counterfactuals and Inference

- Historians and small-*N* researchers dismiss counterfactuals claiming their job is to deal with *reality*
- *But how can they make inferences without implicitly using counterfactuals?*

## “Actuals” and Inference

- Range of variation is known with greater confidence
- But how do we know that the observations are sufficiently similar? (unit homogeneity)
- More observations = more chances for bias
  - omitted variable bias
  - measurement error bias
- Contains implicit counterfactual:
  - switching values of IV would yield causal effect

## Empirical Confirmation

- Actuals: frequency of association
- Counterfactuals: arguments from
  - general principles distinct from hypothesis
  - relevant historical facts
- Example: World War I and the “cult of the offensive” (Van Evera)

## Inference with Multiple Causes

- Suppose hypotheses
  - $AB \rightarrow E$
  - A more important than B
- Actuals: measure difference in causal effects from the *frequencies*
- Counterfactuals: must compare multiple events that did not occur
  - Importance drawn from *plausibility* of the counterfactuals

## Causes and Conditions

- “Cleopatra’s nose” problem (Pascal)
  - If Antony had been less smitten the whole course of history might have been changed...
- To what extent does not A  $\rightarrow$  not B mean A *causes* B?
- *Conditions* are sufficient *causes* are necessary

## Evaluating Counterfactuals

- Counterfactual is not legitimate if we have a theory saying it could not have happened
- The “unimportance of the inevitable” limits counterfactual reasoning (Elster)
- Example: effect of railroad on 19<sup>th</sup> C GNP growth
  - Need strong theory of technical change to disprove counterfactual
  - But theory would predict emergence of railroad!

## Evaluating Counterfactuals

- Counterfactual must be “cotenable” with other initial conditions

## Clouds and Clocks

- Popper asks us to imagine a range of phenomena from deterministic (clocks) to probabilistic (clouds)
  - Where are human individuals?
  - Where is society?
- Newtonian revolution: all clouds are clocks
- Quantum theory: all clocks are clouds
- Popper: we are somewhere in between

## Clouds Clocks and Politics

- Behavioral tradition (clocks)
  - Purpose of science is to discover regularities
  - Individual events should be subsumed by covering laws
  - Only observable relationships are relevant
- Do Almond and Genco agree?

## Milgram Study, 1963

## Tuskegee Study

- Started out with good intentions
- Wanted to demonstrate need for establishing syphilis treatment programs by investigating the effects of untreated disease
- Genuine concern about minority health problems
- Treatment at time was crude and used mercury and arsenic compounds



## Tuskegee Study

- Severe reactions and death were not uncommon
- Suggested untreated patients survival better than treated
- 1932: Follow 300 black syphilitic males (6-8 months only)
- 1933: 300 controls added (to strengthen scientific validity)

## Tuskegee continued

- 1949: Nuremberg Code – no connection made
- 1951: Penicillin widely available – treatment withheld – “never-again” scientific opportunity
- 1972: AP reporter Jean Heller published article exposing study
- 1973: Study stopped, treatment administered still today
- 1997: Clinton apology to **subjects** / families
- Results: 28 deaths, 100 cases of disability, 19 cases of congenital syphilis

## Early Human Research Milestones

- Nuremberg Code - 1947
  - Developed for the Nuremberg Military Tribunal
  - “the voluntary consent of the **human** subject is absolutely essential”
- Declaration of Helsinki
  - Adopted by the 18th World Medical Assembly in Helsinki, Finland in 1964
  - Recommendations similar to Nuremberg Code
  - Basis for the Good Clinical Practices (GCP) used today

## Beecher (1966)

- Published list of 22 studies performed unethically that were published in major journals by respected researchers
- The studies were of questionable study design with no informed consent
- “Until this article, we assumed that unethical research could only occur in a depraved regime like the Nazis”
- Article helped incite debate on research ethics

## Belmont Report

- Three ethical principles:
- Respect for persons
  - Individuals should be treated as autonomous agents, and those with diminished capacity should be protected
- Beneficence
  - Protect people from harm and make efforts to secure their well-being
- Justice
  - Consider the distributional impact of the benefits and burdens of research

## Institutional Review Board / Human Subjects Review Board

- Informed consent
- Exceptions
  - “commonly accepted educational settings” (i.e. classrooms)
  - educational testing
  - public observation of elected or appointed officials and candidates
  - existing documents
  - sanctioned policy studies by agencies
  - consumer “taste and food quality” evaluation

## Observation

- field or local
- direct or indirect (actually seeing the behavior or not)
- participant or nonparticipant
- overt or covert
- structured or unstructured

## Physical Trace Measures

- an indirect observation approach
- what changes result from human activities that can be observed after the subjects have left
- erosion measure
  - observing materials that are worn down
- accretion measure
  - observing materials that are accumulated or built-up
- Are there validity issues with PTM?

## Document and Archive Analysis

- Usually historical in nature
- Examples
  - Soviet military expenditures and political decisions
  - coding Supreme Court justices' notes
  - Civil War strategies
  - US Presidential papers

## Episodic Records

- Preserved in a more casual ad hoc manner where survival is sometimes a matter of luck
- These include:
  - personal diaries and memoirs
  - manuscripts correspondence and autobiographies
  - temporary records not intended for preservation
  - commercial material and propaganda
- Shark attacks (Achen and Bartels 2002)

## Episodic Records Issues

- access and availability
- sometimes time-consuming and tedious
- purposeful deception of authors
- general credibility

## Running Record

- Recorded by organizations (government and private) expressly for the purposes of preservation retrieval and historical reference
- Generally carefully stored previously on paper now almost universally in electronic form
- The quantity and quality are usually very high especially from advanced industrialized democracies
- Often have legal ethical and historical significance
  - (ie Nazi documents Swiss bank accounts US slavery records)

## Running Records Examples

- election returns
- Congressional Record Daily Diary and roll calls
- policy implementation records (Congressional intent)
- crime statistics and other DOJ documents
- campaign contributions and spending (FEC)
- political communications: speeches flyers press releases
- diplomatic materials

## Statistical Index Resources

- American Statistics Index
  - annual from 1973
  - Records (nearly) all federal government publications that contain statistical summary data
- Statistical Reference Index
  - annual from 1980
  - Records summary statistical information from selected state government agencies, non-governmental organizations, businesses, and interest groups

## Statistical Index Resources

- Index to International Studies
  - annual from 1983
  - CIS report on publications from UN, WB, IMF, and other NGOs
- Statistical Masterfile combines all above on a CD-ROM

## Advantages of Running Records

- cost usually borne by the collecting agency not the researcher
- often highly accessible
- typically covers a greater period of time than episodic records

## Disadvantages of Running Records

- dependent on the collection and recording practices of someone else
- selective release embargoing and obfuscation
- unknown measurement practices (usually historical)

## Content Analysis

- Extracting numerical information from written records generally through counts of
  - Words, phrases, names, sentences, or titles
- Two approaches
  - human
  - machine
- Coding issue
  - counts vs. intensity
- Problem: inter-coder reliability

## Content Analysis Examples

- News coverage of political events
- Authorship identification
- Speeches
- Party platforms

## Informant Techniques

- How is informant different from interviewee?
- Single informant technique
  - selection of articulate informant critical
  - Does it create bias?
- Multiple informants
- Standardized interviews
- Controlling for structural position

## Informant Techniques

- What is the role of vested interests?
- How should we cope with them?
- Tension between access to information and organizational bias

## Maximizing Leverage

- Find as many observable implications as possible from your theory
- Draw as many observations from those implications as possible
- For example, if theory concerns primary voters in US presidential elections, then study as many state-level primaries as feasible, and draw as many respondents as possible from each

## Single Observation Designs For Causal Inference

- The  $N=1$  scenario is the biggest challenge for causal theories
- Crucial-Case Study (Eckstein): precisely stated theories can be refuted by a single observation

## But Eckstein is Wrong!

- It does not serve useful explanatory purposes:
  - There usually exist alternative explanations and yet only one case at hand and therefore no variation on the outcome
  - There exists measurement error no matter how strongly or specifically the theory is stated
  - There does not exist an unerring deterministic mechanism: the CCS may be due to chance

## Reasoning By Analogy

- Essentially holds all variables but one constant through matching with some well-understood case
- This is still an  $N=1$  study, but matches to another

## Reasoning By Analogy

- So it is still better to be truly comparative and collect as many similar cases as possible, i.e. more observable implications
- Why? We can capture alternative explanations, get varying levels of measurement error, and average across sampling uncertainty

## Making Many Observations From Few

- Same Measurement, New Units
- Look across space: add similar countries, people, organizations, etc
- Look at subunits of the unit of interest: regions, states, counties, cities, etc
- Look across time: go back in time on the unit of interest

## Making Many Observations From Few

- Same Unit, New Measurements
- Look for many effects from the same cause
- Given that the theory is correct, what else would we think that the explanatory variable could affect? I.e. are there other outcomes that would be expected to occur given the observed explanatory variables
- Example: what are the observable implications from just one outcome from social unrest?

## Making Many Observations From Few

- New Measures, New Units
- Look for units of higher or lower level of aggregation
- Look for related units that are easier to collect/observe

## Making Many Observations From Few

- Are there other variables related to, but wholly unlike original?
  - Example: study threats of nuclear war between two nuclear powers
- Link theory and empirical work by using the observable implications of a theory to suggest new observations that should be made

## Elite Interviewing and Survey Research

- Three main data collection activities in political science:
  - interviews/surveys
  - observation
  - document analysis
- Two main types:
  - elites
  - masses

## Examples of Research on Masses

- American National Election Study
- General Social Survey
- Pew Survey Research
- ABC/New York Times
- Measures of Political Attitudes

## Elite Interviewing Details

- Who is an elite?
- Anyone who in terms of the current purposes of the interview is given special, nonstandard treatment -Dexter
- Usually less structured, more open-ended, non-confrontational
- Focused Interview
  - structured set of questions with flexibility

## Elite Interviewing Details

- Generally elites don't like highly standardized forms
- Interviewer is expected to know all pertinent background beforehand
- Recording device may be an issue:
  - tape recorder?
  - hand-written notes?
- Listening and empathy skills are vital

## Survey Research Details

- Phases
  - design
  - implementation
  - analysis
- Construction of device is critical
- Three alternatives
  - mail
  - phone
  - in-person
- hiring/staffing issues
- schedule/cost issues

## Types of Questions

- President Bush has ordered 20,000 more United States Troops be sent to Iraq over the next three months. How do you feel about this—do you think troops should be added at a faster rate or a slower rate?
- In general, do you feel that the pace at which the president is sending troops to Iraq is too fast, too slow, or about right?

## Types of Questions

- Closed-ended vs. Open-ended
- Single-sided (yes/no) vs. double-sided (A/B) vs. response set (A/B/C/...)
- Sensitive vs. routine
- Branching
- Filter
- Demographic
- Interviewer filled (e.g. intelligence assessment)

## Questions to Avoid

- Double-barrel questions
  - “Do you agree that the president is doing a good job on the economy and that his affirmative action position is just?”
- Ambiguous questions
  - “Do you think the election is going well?”

## Questions to Avoid

- Questions that imply particularistic expertise
  - “What is your assessment of the current policy regarding the transportation and storage of high-level civilian radioactive waste?”
- Questions that create attitudes
  - “Campaign finance laws allow PACs unlimited contributions to national political parties. Do you favor imposing limits on these contributions?”

## Questions to Avoid

- Leading questions
  - Don't you think it increases the safety of our children to put up stop signs at all residential intersections?

## Survey Issues to Worry About

- question order
- response set bias from conditioning
- response rate
- sampling bias, representativeness
- termination by interviewer or interviewee
- missing data
- interviewer bias
- cost vs. sample size

## Survey Issues to Worry About

- Determine open-ended vs close-ended questions
- Ensure clarity and precision of question language
- Consider respondent competence, and respondent willingness
- Are questions relevant?

## Survey Issues to Worry About

- Size of survey instrument (differs by media)
- Avoid negative questions
  - e.g. “Do you think that people should not support a Congress-member who doesn't believe in balancing the budget?”
- Mail survey: clean layout, easy to answer with boxes/checks/circles

## Survey Issues to Worry About

- Be careful with contingency questions
- Clear instructions important
- Pretest
- Ordering: framing, randomizing, sensitive questions, early questions interesting

## Randomized Response Technique

- For dealing with sensitive questions
- Respondent gets two questions: one sensitive and one routine
- Probability of yes to routine question must be known:
  - $\Pr(\text{Born in July}) = N_{\text{yes}}$

## Randomized Response Technique

- Respondent flips a coin without showing interviewer, determines which question gets answered,
  - $P$  = probability of getting sensitive question
- Calculation of yes answers to sensitive question is
  - $S_{\text{yes}} = (R_{\text{yes}} - (1 - P) N_{\text{yes}}) / P$
  - where  $R_{\text{yes}}$  = proportion of total saying yes

## Formal Models

- Relatively new approach
  - Richardson (1939)
  - Black (1948)
  - Shapley and Shubick (1957)
- Downs (1957)
  - Not a description of a real-world democracy
  - Not a portrait of an ideal democracy
- Attempt to model central processes of *any* democratic system

## What is a model?

- Theory? Analogy? Metaphor?
- Three elements:
  - Set of *primitives* (undefined terms)
  - Set of *additional concepts* derived from primitives
  - *Assumptions* (including language of model)
- Can be verbal, logical, or mathematical

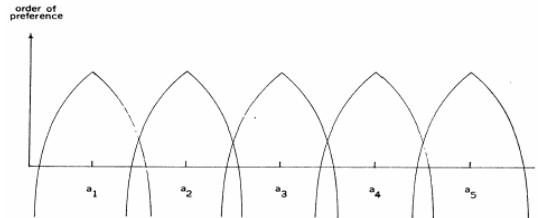


### *Theory of Committees and Elections* (Black 1958)

- Primitives
  - *committee member, proposal, preference*
- Defined concepts
  - *committee, motion, single-peaked curve, optimum*
- Assumptions
  - *Infinite motions, single issue dimension, sincere voting, single-peaked preferences, no abstention*

### *Theory of Committees and Elections* (Black 1958)

- Proposition (theory)
  - *Optimum of median member defeats any other alternative in pairwise voting under majority rule*



### *Theory of Committees and Elections* (Black 1958)

- Prediction (hypothesis)
  - *Median committee member's preference will be the decision of the committee*
- What does it mean if this prediction is falsified?
- Has this added to our understanding of this problem?

### Models Aid Precision

- Many terms in political science are ambiguous
  - “Group strength” or “issue salience”
- BUT, if you cannot model it, you may not understand it
- Examples of improvement from precision:
  - Downs 1957 (parties), Riker 1962 (coalition), Axelrod 1970 (conflict of interest), .....

### Models Force Transparency

- Assumptions “drive” the model, determining what can be deduced
- Informal arguments either loosely state or ignore assumptions
- Examples
  - Single-peaked assumption evident in formal model, not in informal model (Black 1958)
  - “Centrist” American politics depends on many assumptions (Riker and Ordeshook 1973)

### Models Force Consistency

- Formal argument is easy to check for logical (internal) validity
- Formal modelers frequently revise their common sense as they develop a model
  - Needs different assumptions
  - Needs more precise concepts
- Often easier to find invalid part of the argument with math than with English

## Models Aid Cumulation

- Assumptions yield conclusions which can become the assumptions of another model
- Examples:
  - Theory of committees easily becomes part of theory of the Congress
  - Theory of voters easily becomes part of theory of party competition

## Costs and Benefits of Abstraction

- D’Israeli’s (1791) complaint about Aquinas:
  - “Like counting the number of angels that can dance on the point of a needle (head of a pin)”
- Real world is more complex than the model
- But, complexity hinders understanding
  - Most realistic model of a cat brain is another cat brain—but can what we learn from it?
- Main criterion of model building: *utility*

## Costs and Benefits of Abstraction

- Physical scientists know that many functions fit any set of observed data
- Each function may match a different explanation
- Models are not touted as *the* explanation, but one of many

## Prediction vs. Explanation

- Prediction is the critical determinant of success
- Friedman’s (1953) “as if” argument
  - Does not matter if assumptions are true if model predicts well
- Counter-example: Ptolemaic vs. Copernican solar system
- All else equal, realism of assumptions matters

## Formal vs. Other Approaches

- Formal models subject to criticism of assumptions, but so are empirical or verbal models
- Difference: transparency of formal models makes it easier to falsify them = better science

## “Retroduction”

- A relevant known empirical fact X must either be an assumption or an outcome of the model
- The more facts consistent with the model, the better
- New movement in political science: Empirical Implications of Theoretical Models (EITM)

## Normative vs. Positive Theories

- Most positive theorists believe there is a difference:
  - Formal theories are declarative
  - Normative theories are imperative
- Fiorina (1975) claims there is no difference
- BUT, he misses one important difference— values are not falsifiable

## Other Differences

- Micro vs. Macro
  - Is demand best analyzed by modeling individual decisions or assuming an aggregate demand function?
  - Is political support best analyzed by modeling individual vote choices or group support?
- Static vs. Dynamic
  - Is it best to think of a single election or a sequence of elections (finite or infinite?)

## Example: Paradox of Turnout

- Actors: voters
- Actions:
  - vote for Left candidate
  - vote for Right Candidate
  - Abstain
- Outcomes:
  - Left candidate wins
  - Right candidate wins
  - Tie

## Example: Paradox of Turnout

- Assumption: Rationality
  - Preferences are Complete
  - Preferences are Consistent
- Benefit of voting: B
  - (Aldrich normalizes this to 1)
- Cost of voting: C
- Probability of affecting outcome: P
- Vote if  $PB > C$

## Example: Paradox of Turnout

- $P = 0$  unless total vote is within one vote of a tie!

Figure 1. Decision Table for Rational Choice Models of Turnout

Actions:	States of the World				
	Preferred Candidate is				
	(1) Winning by More than One Vote	(2) Winning by Exactly One Vote	(3) Tied	(4) Losing by Exactly One Vote	(5) Losing by More than One Vote
(1) Vote for Preferred Candidate	$1 - C^*$	$1 - C$	$1 - C$	$1/2 - C$	$0 - C$
(2) Vote for Other Candidate	$1 - C$	$1/2 - C$	$0 - C$	$0 - C$	$0 - C$
(3) Abstain from Voting	1	1	1/2	0	0

\*Entry is payoff in utilities to decision maker. It is assumed that  $0 < C < 1/2$ , where  $C$  is costs of voting. A tie is assumed to be broken by the flip of a fair coin. Utilities are normalized, so that the value of the preferred candidate winning is 1; the value of the opponent winning is 0.

## Example: Paradox of Turnout

- As  $N$  gets large,  $P$  gets very small even for close elections
  - Bush v. Gore 2000, even if you knew the difference in Florida would be 1000 votes or less, then  $P \sim 1/2000$
- “True”  $P$  in US Presidential elections is estimated to be about 1 in 10 million
- Suggests voting is “irrational” or there are other unmodeled benefits to voting

## Example: Paradox of Turnout

- Formal theory predicts many relationships that have been observed
  - Turnout increases as benefit increases
  - Turnout decreases as cost increases
  - Turnout increases with closeness of election
- Challenge: identify the incorrect assumptions and change them to get these results *and* significant aggregate turnout

## Recommendations

- Do not skip over technical parts of formal models!
  - Examine definitions and assumptions
- Examine interpretive link between model and conclusions
- Examine link between empirical regularities and model's conclusions and assumptions

## The Comparative Method

- Applies to studies of two or more countries
- Several scholars have argued that the comparative method is unique
- Lijphart: no, the comparative method is just a small-*N* version of a normal research problem

## Weaknesses and Strengths

- Many variables, small number of cases
- Lijphart: if at all possible, use statistical approach
- However, small-*N* comparisons may yield
  - more depth
  - with fewer resources
- Especially useful in initial stages of research

## Lijphart's Recommendations

- Increase the number of cases as much as possible
- Reduce the "property-space" of the analysis
- Focus the comparative analysis on "comparable" cases
- Focus the comparative analysis on "key" variables

## The Case Study Method

- Atheoretical case studies
- Interpretive case studies
- Hypothesis-generating case studies
- Theory-confirming case studies
- Deviant case studies

## The Scientific Study of Politics

- Think critically
- Express uncertainty
- There are many paths to the mountain
- Never be intimidated
  - there are millions of talented people like you
  - but there are an infinite number of ideas