# POLS2044 WEEK 8
# Comparing groups

Australian National University
School of Politics & International Relations
Dr. Richard Frank
https://richardwfrank.com/research_design_2022

In Week 8 of POLS2044 we will be focusing on ways of comparing groups. This is the essence of causal analysis and underscores an often-unstated assumption of much political science research, through comparison of multiple cases we can find similar causal and empirical patterhs.

This week I have two main goals. First, I want to discuss several methods of comparative research. Second, I want to highlight important underlying assumptions and challenges we use and face whenever we theorise and research across cases. In pursuing these goals, we are going to take a step back from Excel and try and reaffirm our understanding of the links between causal theorising and the correlational techniques we started learning in Week 7 and the descriptive statistics we learned in Week 6.

---

## Reading notes and questions

There are three readings for this week. The Posner (2004) article is one of those articles that blew my mind when I first read it. While we will be discussing all three articles this week, please do make sure to give the Posner (2004) article your full attention

**Posner, Dan. 2004. "The Political Salience of Cultural Difference: Why Chewas and Tumbukas Are Allies in Zambia and Adversaries in Malawi."** ***American Political Science Review*** **98(4): 529-546.**

Wow, this is a cool research design. There are a host of possible reasons why this paper has been influential in both the methodological and political literatures as well as being cited over a thousand times. I think it is because the paper highlights a clear and direct puzzle and leverages a plausible natural experiment to present a clear causal argument solving the puzzle.

I would suggest reading this paper for both its comparative analysis of two ethnic groups that span a national border as well as for its overall paper structure.

1. What is the main puzzle and research question?

The paper does not have an explicit hypothesis (that I could find).

2. How would you phrase the expected relationship (i.e., hypothesis) between the cause and effect?

The paper describes why Posner (2004) draws on this natural experiment.

3. What is a natural experiment and how does this comparative case study approximate one?
4. What is Posner's (2004) unit of analysis?

The main evidence presented is a survey in four villages.

5. Who are the homogenous groups in each village?
6. How large is the sample size? How was the sample determined?

The analysis uses both means and standard deviations in describing his results in both the text, Figure 2, and Table 1. He also runs a t-test and reports a p-value (p. 534).

7. Which way (text, figure, or table) did you find the results most intuitive?
8. Given the size of the standard deviations in the text (p. 533-4) do you understand why Posner (2004) did not include standard deviations in Figure 2?

One way this paper does differ from many research papers we have read is in the paper structure. The paper starts with a focus on the "what" descriptive information before switching to the "why" causal story on page 535.

9. Do you find Posner's (2004) consideration (and ultimate dismissal) of alternate explanations convincing? Why/why not?

The paper then runs another survey to show that cultural differences are neither a sufficient nor necessary condition for political cleavages.

10. Is there anything inherently theoretically surprising in the fact that administrative boundaries shape politicians' incentives to highlight certain cultural groups for coalition forming? Or is the paper's main contribution finding and describing a natural experiment that suggests that is the only viable causal pathway to explain the patterns in Zambia and Malawi?

**Lijphart, Arend. 1971. "Comparative Politics and the Comparative Method."** ***American Political Science Review*** **65(3): 682–693.**

This paper has been even more influential than the Posner (2004) article in the study of comparative politics. To date, it has been cited 5993 times.

The paper begins by highlighting the fact that "comparative politics is the only [political science subfield] that carries a methodological instead of a substantive label," (Lijphart 1971: 682). Nevertheless, it is almost as hard to describe how to *do* comparative politics as it was back in week 3 to describe how to conduct a case study.

11. The paper highlights several strengths and weaknesses of the comparative method. What are they? What are his proposed solutions/means of addressing the weaknesses?

12. How does Lijphart (1971) think that intranational analysis aids comparability more than international comparisons?

13. Did you use (and mention) any of the six types of case studies he describes towards the end of the paper?

**Diamond, Jared. 2005. "One Island, Two Peoples, Two Histories: The Dominican Republic and Haiti," in *Collapse: How Societies Choose to Fail or Succeed*. New York: Penguin Books: 329-357.**

You really only need to read the first 12.5 pages of this reading. This reading highlights some converging and mostly diverging trends experienced by two countries sharing the same island—Haiti and the Dominican Republic. The focus is on explaining why two countries that shared a number of environmental factors ended up with divergent environmental and social outcomes.

14. Do you find the historical, ecological, and political factors he highlights convincing?

15. Can any conclusions from this comparative case study be generalised to other cases?

---

## LECTURE PART 1: Introduction

**What should we take away from the previous weeks?**

This class is geared towards developing your knowledge about how to (1) evaluate others' research and (2) how to conduct your own research.

We have developed our understanding of the different steps of the scientific method and realised that this process is often messier and less linear than expected.

Everything starts with our theories about how and why some part of the world is (or was) the way it is.

It is crucial to think about how well our theories and our empirical measures are connected.

**We have alternated between description & explanation**

<u>Description</u>
Week 4—Concepts & measures
Week 6—Descriptive statistics

<u>Explanation</u>
Week 3—Qualitative research
Week 5—Surveys and sampling
Week 7—Correlation

**Today's motivating questions**

Why conduct aggregate comparative research?
How can we maximise the probability that our evidence tells us something about the veracity of our causal argument?

**Theoretical arguments often focus on causal (why) explanations.**

**Four hurdles to establishing causality**

1. Is there a **credible mechanism** connecting X and Y?

2. Can we rule out Y causing X (**endogeneity**)?

3. Is there **covariation** between X and Y?

4. Have we controlled for potential **spuriousness** (Z)?

**We evaluate our theories using often limited evidence.**

**Definitions**

> "A population is any group of people, organisations, objects, or events about which we want to draw conclusions; a *case* is any member of such a population." (Brians et al. 2011: 132)

> "A sample is any subgroup of a population of cases that is identified for analysis." (Brians et al. 2011: 132)

> "A representative sample is one in which every major attribute of the larger population from which the sample is drawn is present in roughly the proportion or frequency with which those attributes occur in that larger population." (Brians et al. 2011: 133)

**Challenges to measurement**

> **Conceptual clarity**—Do we know what we want to measure?
> **Operational reliability**—Are the measures repeatable and consistent?
> **Conceptual validity**—Does the measure accurately measure the concept we are trying to measure?

**Reliability and validity**

**Validity**

> Face validity—On its face does a measure appear to be measuring what it says it is measuring?

> Content validity—Does a measure capture all of the systemised concept? Is anything missing? Is anything there that should not be?

> Criterion validity—Does a measure correlate with criterion (i.e. ground truth) variables?

> Construct validity—Do measures behave the way you theoretically expect in the wild?

**This evidence is inherently comparative.**

**Correlation or causal explanation needs comparison.**

Comparison across space
Comparison across time

**Correlation**

A correlation is the statistical association between two variables.

It has five important characteristics (nature, direction, sign, strength, statistical significance).

Calculating a correlation coefficient and its statistical significance is straightforward.

Interpreting what it means is a different thing and requires thinking causally.

**Pearson's correlation coefficient and OLS formulas**

**Random measurement error**

"People are not very good at understanding randomness. There's much more chance out there than we think there is. While we are seeking for patterns and explanations as we look backward, we're not giving a fair shot to the explanation that many things are really just random events." — Lisa Goldberg

**Ceterus paribus assumption**

**Why conduct a significance test?**

We want to be sure that the correlation or regression estimate is not an artefact of random chance or what sample we have.

**How do we conduct a significance test?**

rho is the population correlation coefficient.
Null hypothesis (Ho): rho=0, there is not a significant linear correlation between x and y in the population.
Alternative hypothesis (H1): rho is not equals to 0, there is a significant linear correlation between x and y in the population.

**Student's t-test formula**

**Data visualisation**

Visualising data is a crucial way to understand yourself and describe to others the "what" (descriptive inference) and the "why" (causal inference).
It is often easier to interpret than data tables.
Several popular visualisation types include maps, line charts, bar charts, scatterplots, and histograms.

Visualisations for others are useful only if they are clear, accurate, and are conveying a direct message.

**Conclusions are probabilistic rather than deterministic.**

**Type I and Type II errors**

**Significance tests balance the risk of Type I and Type II errors.**

**Introductory takeaway**

The material we cover in this class builds on itself.
We are trying to develop the skills to both (1) develop descriptive and causal theories and (2) test them.

---

## LECTURE PART 2: Why use comparative research methods?

It enables you to ask questions that can travel.

It uses measures that also travel.

For example, election management bodies' and election monitors' effects on election violence

---

## LECTURE PART 3: What are aggregate data?

Aggregate data are data on **groups**. There are two types:

**Summary indicators** of individual phenomena
Summary individual data is often discrete (can a person read? 0=no; 1=yes)
Summary group data is often a percentage (% that can read).

**Syntality indicators** are measures of the group
For example, democracy and non-democracy

**Two additional categories of aggregate data**

**Areal groups**—groups defined by a geographic area (e.g., a country, city, and village)
**Demographic groups**—groups defined by personal characteristics (e.g. age, education, and occupation)

**Aggregate data help us measure latent phenomena.**

Political communication— # population with TVs or cell phones
Globalisation—imports & exports, international flights
Health—# hospital beds, maternal mortality, children living to the age of five,

**Where can we find aggregate data?**

Censuses—For example, the Australian census run by the ABS
Organisational—World Bank, IMF, European Union, UNESCO, etc.
Sample surveys—From above institutions or Asia Barometer, AES, World Values Survey, Gallup, etc.
Publication content—Newspapers, books, Lexis Nexis, Factiva, Keeping's, etc.
Event data—ACLED, UCDP, SCAD, etc.
Judgement data—V-Dem, Electoral Integrity Project, etc.

**Ways of comparing groups**

**An experiment example & results**

**A natural experiment**

Famous examples include Putnam (1993) and Posner (2004).

**A comparative case study**

A single case study would be Geertz's (1972) Bali study.
A comparative case study would be Diamond's (2005) study of Haiti and the Dominican Republic.

"There are far more cultural cleavages in the world than there are conflicts," (Posner 2004: 529).

**Comparative statistics example from my own research**

One final method tries to take us where we started—statistical pseudo-experimentation through a process called **matching**.
Using comparative statistics matching pairs observations that are similar in as many dimensions as possible except for the **treatment** (e.g., election observation).

**Aggregate data important takeaways**

---

## LECTURE PART 4: Group-level data challenges

**Challenges when using group data**

Concept-measurement connection
Ecological fallacy
Creating indicators

**Concept-measure connection**

We have to make sure that we have data at the same level (country, state, individual) as the concept we are studying.
This is in addition to the concept-measurement issues we talked about in Week 4.

**Data options for measuring election violence**

**Ecological fallacy**

An **ecological inference** is an inference about individual behaviour drawn from data about aggregates.
The **ecological fallacy** consists in thinking that relationships observed for groups necessarily hold for individuals.

**Ecological fallacy example**

If we are interested in the relationship between being foreign-born and literacy, we are likely to have aggregate data on both but no individual-level data.
If there is a positive correlation between these two variables at the aggregate level, that does not mean that we can say anything about specific individuals in these areas (Robinson 1950).

**Creating indicators**

Raw data
Additive index (e.g. Perceptions of Electoral Integrity)
Multiplicative indices (e.g. CIRI human rights data)
Transforming (e.g. using the natural log) the raw data
Standardisation helps (e.g. per capita or % GDP)

The more measures you have the better.

**Finding good data**

As much art as science.
Check the types of sources I mentioned earlier.
Data availability is often a function of state capacity (Hollyer et al. 2014) and stability (Schultz & Mankin 2019)

**Example: climate data and conflict**

**Observations are often not independent.**

There is often spatial clustering of the outcomes we study including democracy, violence, development.
However, often our theories and tests assume away neighbourhood effects.

**Two examples of regional clustering**

**Group-level data challenges: important takeaways**

1. We need to ensure that our theory and data operate at the same level (e.g. nation, shire, individual).
2. We should minimise generalising our findings across levels of analysis (avoid ecological inference and fallacies).
3. Creating and using indicators involves both theoretical and statistical considerations (e.g. standardising and transforming).

## LECTURE PART 5: Assessment guidance

Revised problem statement suggested structure

One potential final essay structure

Posner (2004) is a good example of a clear structure.

## WEEK 8 TUTORIALS

In this week's tutorial, we will be discussing and applying the concepts of analysing relationships at the data's aggregation level as discussed in the readings and lecture. It is important to stress that we want to avoid drawing conclusions at a different level of analysis than our data are. For example, Posner's (2004: 530) focus was at the level of the Chewa-Tumbuka "cultural dyad". We would not want to speculate about how his argument would explain Zambia's and Malawi's international relations.

**Part 1: Analysing group data**

In this first part of the tutorial, you will be breaking into your normal groups of three or four students.

Please choose two random countries you would like to learn a bit more about.

For each of your countries locate recent data on (1) crime level and (2) population density. Several possible data sources include the World Bank (https://data.worldbank.org/) and the United Nations Office on Drugs and Crime (https://dataunodc.un.org/). Now compare your two countries:

1. Which one has the higher crime rate?

2. Which one has the higher population density?

3. If you find that the country with the higher density has more crime, could you reasonably conclude that packing people together causes crime? If the lower-density country has more crime, does being farther from others make you anti-social?

4. To properly analyse question 3, what level of analysis would be required?

5. What other potentially relevant factors are also worth studying when trying to explain crime rates?

6. What type of crime did you analyse? Do the same trends you found when answering questions #1 and #2 also apply to other types of crime?

7. In lecture, we reiterated the four hurdles to establishing causality. Can we jump over all four hurdles here? Why or why not?

Overall, did the groups in tutorial find similar answers to these questions?

**Part 2: Revised problem statements**

The remaining tutorial time can usefully be spent discussing your revised problem statement, the challenges you are dealing with when completing the assignments, and any last remaining questions that you may have.