POLS2044 WEEK 5
Surveys and Sampling

Australian National University
School of Politics & International Relations
Dr. Richard Frank

In Week 5 of POLS2044 we will be focusing on sampling and surveys. What do these terms mean in the context of political science research? Put simply, the world is too big and complex (and costly) to gather perfect and complete data on many outcomes (Y) and the causal mechanisms (X) that (may) affect them.

For instance, if we want to know how Australians view the Russia-Ukraine war, we do not have the time and money to call or visit all 25,948,285 Australians and ask. Instead, if we were a local polling firm, we would survey 1,618 people online and ask them. This is what the *Sydney Morning Herald* did back in April 2023 (link to the article here), and they found (with a margin of error of 2.5%) that "81 per cent of Australians back the federal government's decision to send medical equipment and other supplies to Ukraine, while 75 per cent want to send non-lethal military equipment and 77 per cent favour personal sanctions on Russian leaders."

Surveys are a popular way to *sample* a larger population's feelings or perspectives quickly and relatively cheaply. However, there are many other ways that researchers gather samples and try and theorise about them and generalise to the broader population. The three readings for this week include one overview article (Schaeffer and Presser 2003) that describes the myriad considerations that go into formulating a survey questionnaire and two examples of creative and quite unique ways of looking at samples of some hard to quantify outcomes. I recommend that you start with Schaeffer and Presser (2003) before reading Levitt and Venkatesh (2000) and Page (1997). Remember to read strategically. The Amanda Hoover Green article from week 1 is worth revisiting. Focus on getting out the main parts of each paper. What is the research question? What literature is this question based in? What is the main argument? What is the main hypothesized relationship? What is the sample? What are the main findings?

I. Reading notes and questions

Schaeffer, Nora Cate, and Stanley Presser. 2003. "The Science of Asking Questions." ***Annual Review of Sociology*** **29: 65–88.**

To be honest, this article is like a chocolate nut brownie—dense but filling. The focus is on survey methodology, and this review article summarizes the decisions involved when developing two common types of survey questions.

1. What are these two common types of questions?
2. The article highlights many ways that question wording, how the question is asked, the multiple-choice scales, and what is being asked about can affect the answers given. How might this link to the types of reliability and validity concerns we learned about over the last few weeks?
3. The article highlights the challenges when asking people about things they had not thought about before. Why should we think about people who answer "don't know" to survey questions?

4. Why might asking an agree/disagree question be problematic?
5. Does this overview of research on the many ways that survey questions can shape survey responses make you more or less likely to believe individual survey results?

Levitt, Steven, and Sudhir Venkatesh. 2000. "An Economic Analysis of a Drug-Selling Gang's Finances." Q*uarterly Journal of Economics* 115: 755- 789.

This article is co-authored by one of the creators of Freakonomics, Steven Levitt, an economist. The main driving force behind the article, however, is Sudhir Venkatesh, a sociologist. He was the one who spent the better part of a decade spending time with members of The Black Kings, a Chicago street gang. He wrote about his experience in the book "Gang Leader for a Day." The article we are reading focuses on the economic analysis of the gang's finances.

1. This is a very detailed analysis of one gang during one four-year window. What potential biases do the authors recognise as possible limitations of their work?
2. The authors highlight the relative low wages of the rank-and-file members. Why do the street-level sellers work for these wages and what do they hope to achieve in the future?
3. What lessons can we draw from this work about the possibilities, difficulties, and limitations of researching hard to study populations?
4. Are there other hard to study populations that you think could also be usefully analysed?
5. How was the gang structure comparable to a fast-food franchise?
6. What was the main puzzle they found in the data and how did they try to explain it?

Page, Stewart. 1997. "An Unobtrusive Measure of Racial Behavior in a University Cafeteria." *Journal of Applied Social Psychology* 27: 2,172-2,176.

This article is short and to the point. The focus is on gathering and analysing observational data on a socially undesirable behaviour. Racist behaviour, like the criminal behaviour in the previous article, is a real outcome in the world, but it is hard to study because people have incentives to hide or downplay the behaviour.

1. Do you think this article provides evidence of the behaviour he says is being observed?
2. Did the author adequately address or control for alternate explanations?
3. Focus on the statistic used to reject the null hypothesis in favour of the "speculative hypothesis." What is it? Can you conduct a web search to see how it is calculated? We will be working through this statistic later in the term, but it is worth highlighting now when we see it in the wild.

Lecture PART 1: INTRODUCTION

Recapping the last month

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.

We have realised that this process is often messier and less linear than published research might lead you to believe.

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 highlighted some types of qualitative research methods.
Today we talk about a quantitative research method—surveys.
Before that, however, we continue our discussion about how we can chose our sample to increase its representativeness of the larger population.
This speaks to both qualitative and quantitative research approaches and connects to our discussion in Weeks 2 and 3.

Today's motivating questions

What is a sample? When is it representative? What does it represent?
What types of surveys are there?
What are their strengths and weaknesses?

Motivating puzzle

Surveys are some of the most frequently reported forms of research reported in the media.

However, most people do not understand how to interpret both (1) the questions being asked and (2) how we can evaluate whether the sample represents the population being studied.

How reliable, representative, or valid is this survey?

Jimmy Kimmel people on the street survey about world geography
Source: https://youtu.be/umpalMtQE50

LECTURE PART 2—Sampling

Research often samples ideas from the population of explanations provided by the existing literature. Example: Carey et al. (2022) sampling Collier and Hoeffler (2004)

Sampling is akin to only being able to take one slice of pie.

As we saw last week population measures matter.

Article by David Everatt in The Conversation about South Africa's 2022 census.

Canberra Times article about Census 2021

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)

Generalisability

> "A truly representative sample is a microcosm—a smaller, but accurate, mode—of the larger population from which it is taken. To the extent that a sample is truly representative, conclusions based on a study of that sample may be safely regarded as applying to the original population."
> (Brian et al. 2011: 133)

An epic sampling fail—The Literary Digest Presidential Poll of 1936

Photo of Harry Truman holding up newspaper saying he lost the 1948 election

> Both election outcomes show the connection between some surveys and the ground truth (criterion validity)

How can we select a representative sample?

> Probability sampling using random samples: randomly selected cases so all cases (and all combinations of cases) have an equal chance of being selected.

Screenshot of random.org

Screenshot of how to generate random numbers in Excel

A graduate student recently came up with a way of generating true random numbers from the radioactive decay of a banana's potassium.

Variations on random sampling

<u>Systematic</u> random samples:

1. Count the number of cases you have and divide by the number of cases you want in your sample. We can call the result *k*.

For example, we have 1,000 names and we want 50 random names. *k* would be 20 (1,000/50).

2. With some random selection process (e.g. a random number table), we select a number, *j*, between 1 and *k*.

For example, my banana random number generator selects *j*=7.

3. We then go through case list and chose the cases from j to (j + (n-1)k).

In my example, we would chose names 7, 27, 47, 67,…., 987.

Cluster or multistage sampling

Note in systematic random sampling any bias is determined by the case list we use.

With cluster or multi-stage sampling the focus is not on individual cases (e.g. people) but on groups of cases (e.g. a house or dorm).

Example of ways we can sample the Australian population using AEC data, Excel, and QGIS.

Stratified random sampling

Used when a population subgroup you want to study is too small that a random sample is likely to not gather enough cases.

This involves two (stratified) samples:

A simple or systematic random sample of a **smaller group** that is **larger than the expected occurrence** (e.g. a group is 5% of population but twice that is surveyed).

A simple or systematic random sample of a **larger group** that is **smaller than the expected occurrence** (e.g. 95% of population but 90% is surveyed).

This is possible only if you know ***ex ante*** the size of each group in the population.

Reaching conclusions given selection bias

1. Weight each group by their overall distribution in the population.

2. Compare across samples.

3. Matching

Example of stratified random sampling using Australian age and income distribution

Example from New York Times about how one panelist skewed 2016 Los Angeles Times survey results

There was a trade-off between bias and variance.

Reliability and validity bullseye example.

Non-probability samples

These are samples that are not selected randomly.

Convenience samples
Volunteer samples
Purposive samples
Snowball samples

Convenience samples

Examples of drink walking and me getting surveyed on the streets of Kathmandu

Volunteer samples

Examples of Boston sleep studies and ANU psychology studies

Purposive samples

McDonalds, SELTS, Elon Musk on Twitter, Eurovision

Snowball samples

Cohen and Arieli (2011) on field research in conflict environments, Sudhir Venkatesh on Chicago gangs

Considerations when determining sample size

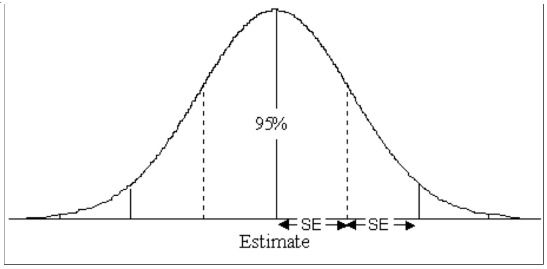
Population homogeneity
The number of variables, treatment options, or response categories
Desired sampling error & confidence interval

Sampling error and confidence intervals

The more cases we have in our sample:
(1) the more likely we are to have a random sample,
(2) the smaller the sampling error, and
(3) the smaller the confidence intervals.

The Australian National Election Study's methodology

Confidence intervals explained



There is a 95% chance that the confidence interval which extends to two standard errors on either side of the estimate contains the "true value".

This interval is called the 95% confidence interval and is the most commonly used confidence interval. The 95% confidence interval is written as follows:

95% confidence interval for outcome y = [y - [2 * se(y)] , y + [2 * se(y)]]

Normal distribution example—Magpie swooping

How to calculate the margin of error?

Sampling error at 5% significance level= $1.96\sqrt{Var/\mathrm{n}}$

With variance (var)=p(1–p) where is the number of respondents, and p is the proportion favouring an outcome.

LECTURE PART 3—Surveys

Why conduct a survey?

Surveys are a popular means of gathering information about:
Events or behaviours
Evaluations or attitudes

Events or behaviours

Example from the readings of an Ontario university cafeteria

Stages of a survey

Conceptualising
Survey design
Instrumentation
Planning
Sampling
Training/briefing
Pretesting
Surveying
Monitoring
Verifying
Processing
Analysing
Reporting

Survey ethics approval is a must

Question wording challenges

Excessive length
Ambiguous wording
Two questions in one

Bias—questions encourage one type of answer
Response set (agree/disagree) bias—people want to agree
Argumentativeness
Social desirability bias
Forcing a response when people do not have one—"don't know"

Hint: These are potential ways of critiquing or building on existing surveys.

Questionnaire challenges

Questionnaire structure
Explanation
Warm-up questions
Substantive questions
Demographic items

Questionnaire format
Questions per page
Number of questions
Length of time

Types of surveys

Personal interviews
Pros—allows for a variety of questions, can hold attention for longer, higher response rate
Cons—expensive, potential for bias, principle-agent issues, time-consuming

Mail survey
Pros—cheaper, reduces response bias, more time to respond, less people to conduct
Cons—Not all cases have clear mailing addresses, less questions are feasible, low response rates
Telephone survey
Pros—speed, reduced bias compared to in-person, cheap, can be automated
Cons—hard to reach unbiased sample, fewer home phones

Internet survey
Pros—cheap, reduced personnel, format easy to adapt, respondent convenience, live monitoring, global reach, easy software
Cons—not everyone has a computer, hard to reach representative sample, self-selection bias, lack of personalisation, technology variation, attention wanes

Unrepresentative samples

Example of how a teenager on TikTok disrupted thousands of online studies

Another survey question type: List experiments

Example from Shy Trump voters

Survey database examples

Australian Election Survey
American National Election Studies
The European Union
The UK Data Archive
The Pew Research Centre
Gallup
Roper Centre for Public Opinion Research
World Values Survey
Afrobarometer
Asia Barometer
ICPSR
National Opinion Research Centre
Latinobarometer
Harvard Dataverse
World Bank country surveys
World Bank enterprise surveys
Populism public opinion surveys (Stanford)
YouGov

Non-random or stratified focus groups

The focus of focus groups is not on the individual but the group as a unit.

Key elements
1. Clear objective/problem
2. Group characteristics
3. Atmosphere
4. A listening facilitator
5. Structure and direction but restrained facilitator
6. Recording discussion (text or video)
7. Analysis

Not random or representative: Expert surveys

Pros—Potentially deep knowledge of issue, group, or time; repeated panel
Con—finding experts, proximity biases, high non-response

Examples: V-Dem, PEI, TRIP surveys

Important terms this week

Population
Sample
Representative sample
Random sample
Systematic random sample
Cluster or multistage sampling
Stratified random sample

Selection bias
Nonprobability sample
Convenience sample
Volunteer sample
Purposive sample
Snowball sample
Sampling error
Confidence interval
Variance
List experiment

III. WEEK 5 WORKSHOP

The focus of today's workshop is on applying the readings and lecture material. Given that this is the last week where we are not going to explicitly focus on quantitative data description and analysis, I want to do things a bit differently. Instead of starting with individual activities and then moving into small group work, we are going to focus on discussing the readings, lecture, and assignments as student groups. Therefore, please wait until all students have completed the quiz before assembling into groups of three to four students.

Part 1: Page (1997)

First let us focus on the Page (1997) article.

1. **Schaeffer and Presser say that there are two types of survey questions. What type is being used by Page 1997?**

2. **What is the hypothesis being tested?**

3. **Now think about the types of surveys discussed in lecture. Do you think we could address the hypothesis from question 2 (or the underlying causal mechanisms) using different forms of surveys? If so, how? If not, why not?**

4. **What questions or concerns about the sample used in this article do you have?**

5. **What population do you think Page (1997) is trying to generalise to?**

Submit your questions to Wattle/POLS2044/Week 5/Item 5.1 before you move on. Be sure to have one student submit for the group and include the other student's names & uniid.

Part 2: Levitt and Venkatesh (2000)

Next, let us dig into the Levitt and Venkatesh (2000) article. This article's focus is clearly on the economics of an urban gang. The article is vague about the location and name of the street gang, but in later work Professor Venkatesh identifies the location as Chicago. In later work he also admits that he started his gang research by approaching gang members with an in-person written survey. However, he quickly realised that this was not the most appropriate method to

get the data he wanted, and Professor Venkatesh shifted to a qualitative sociological approach. Yet in the end he gathered the quantitative data reported in this article.

6. **Do you think it would be possible to try a different type of survey to learn more about this type of often illegal business (hint, look at the survey types I talked about in lecture)? If so, what type? What actors would you survey? If not, why not?**

7. **How might this one gang represent (or fail to represent) other gangs either across space or across time?**

Submit your questions to Wattle/POLS2044/Week 5/Item 5.2 before you move on. Be sure to have one student submit for the group and include the other student's names & uniid.

Part 3: Sampling & surveys

As a last effort at discussing issues related to sampling let us look at real-world surveys. It connects to this week's discussion of sampling and heterogeneity of larger populations.

I have included three ranked lists below. As a group chose one list to explore (using whatever selection mechanism you choose).

1. The 20 most popular travel destinations in the world (https://www.insider.com/the-20-most-popular-travel-destinations-in-the-world-2016-9)

2. The top 10 social media platforms (https://www.searchenginejournal.com/social-media/biggest-social-media-sites/)

3. The top 21 most popular fruits (https://www.worldatlas.com/articles/the-most-popular-fruit-in-the-world.html)

Then one student will go to random.org (https://www.random.org/) and chose a random number between 1 and the top end (max) of your selected ranked list (e.g., 20, 10, or 21).

Look at the selection chosen and the rest of the list. Then think about the consumers of the selected travel destination, social media website, and selected fruit.

8. **What random selection did your group make?**

9. **What type of person is <u>most likely</u> to consume this selection?**

10. **What types of people are <u>less likely</u> to consume this selection?**

11. **If you were going to study this population of consumers, what <u>types of survey</u> would you chose and what questions would you ask?**

12. **Would it matter <u>when</u> or <u>where</u> you conduct your survey?**

13. **Could you <u>generalise</u> from your imaginary results to the consumers of the other options on your chosen list?**