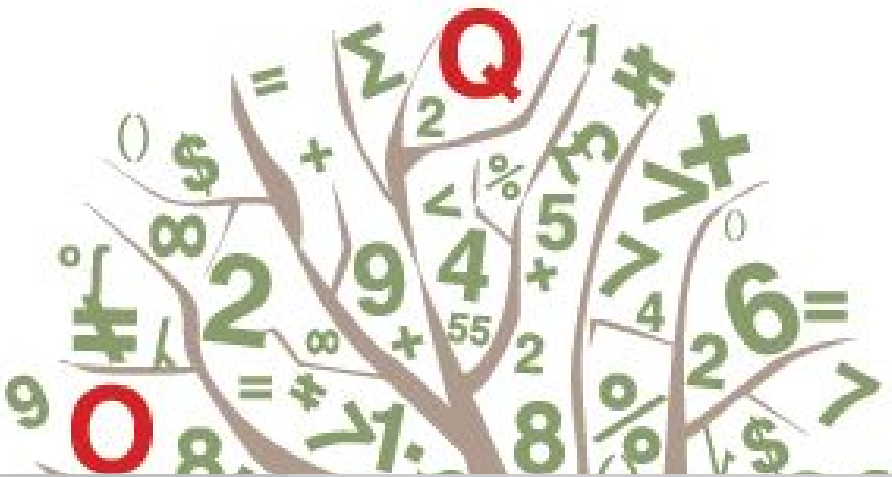
Q-Step: Week 8 Lecture

Synthesis and Review

Spyros Kosmidis

Oxford

March 16, 2022



Roadmap

Previously

- Research Design
- Concepts and Measurement
- Descriptive stastistics
- Case Selection
- Inference
- Correlation & Regression

Today

- Review
- Synthesis
- Q-Step Essay

Research Questions

- Descriptive questions:

How often do revolutions occur?

- Explanatory questions:

Why do (or don't) revolutions occur?

- Causal questions:

What is the effect of military defeat on revolutions?

How should we select our cases?

- Your case selection will affect the answers you get for your research questions.
- Case selection vital for observational studies
- Think about type of question, if explanatory (why X leads to Y?) or causal (what is the effect of X on Y?): DON'T SELECT ON DV
- Key problems due to biased case selection:
- **Guilt by Association**: Erroneous inference that any characteristic that cases with similar outcome share is a cause
- **Cherry picking**: Erroneous inference based on picking cases that support hypothesis whilst ignoring other cases
- **Overgeneralisation**: Erroneous inference that a relationship between variables in selected set of cases reflects the relationship in all cases

Measurement: Why should we care about it?

Example:

Can we draw a sharp distinction between regimes that are democratic and those that are not? If so, what are the criteria? If not, why not?

- Not only concerned with distinction in the abstract, but also whether this is possible empirically
- Requires engagement with empirical work on democracy
- Again, may want to think about research question (and theoretical argument) under examination
- e.g. Harding & Stasavage (2014), "What Democracy Does (and Doesn't Do) for Basic Services"
- EIEC (from Database of Political Institutions), e.g. Lindberg (2006), "Democracy and Elections in Africa"
- Freedom House
- If interested: look at Varieties of Democracy project (www.v-dem.net)

Measurement: Consequences

Binary

A regime is classified as a democracy if all of the following conditions apply. Otherwise, it is classified as a dictatorship.

- The Chief Executive must be elected.
- The Legislature must be elected.
- must be more than one party.
- There must have been at least some alternation of power under existing institutional arrangement.

Measurement: Consequences II

Continuous

Polity IV: Regimes coded on indices of democracy and autocracy.

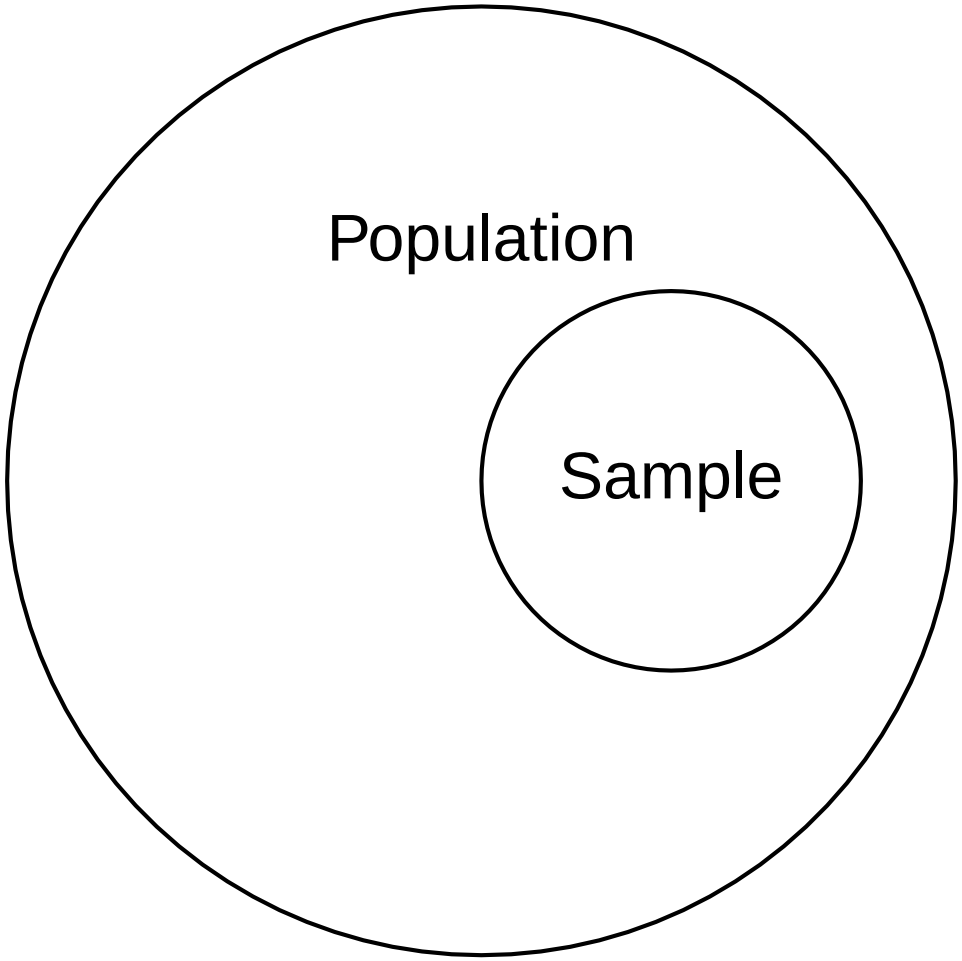
- The competitiveness of political participation (1-3).
- The competitiveness of executive recruitment (1-2),
- The openness of executive recruitment (1), and
- The constraints on the chief executive (1-4).
- Both measures are useful but are based on different assumptions. Examine the latest trends and see for yourself which measure does a better job?

Case Selection

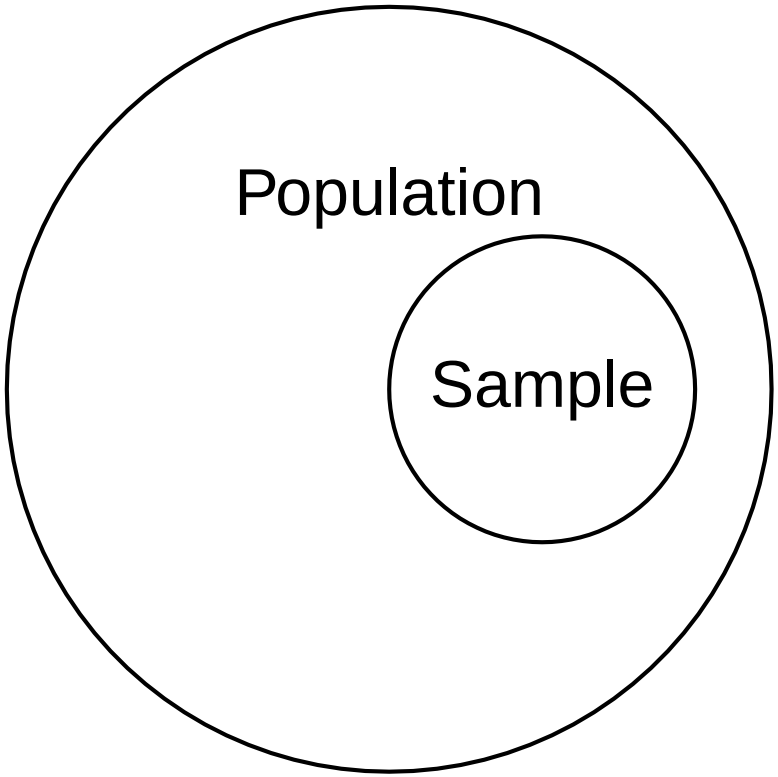
- Population: The entire set of cases that our theory applies to
- Sample: Subset of cases that we analyse

Case Selection

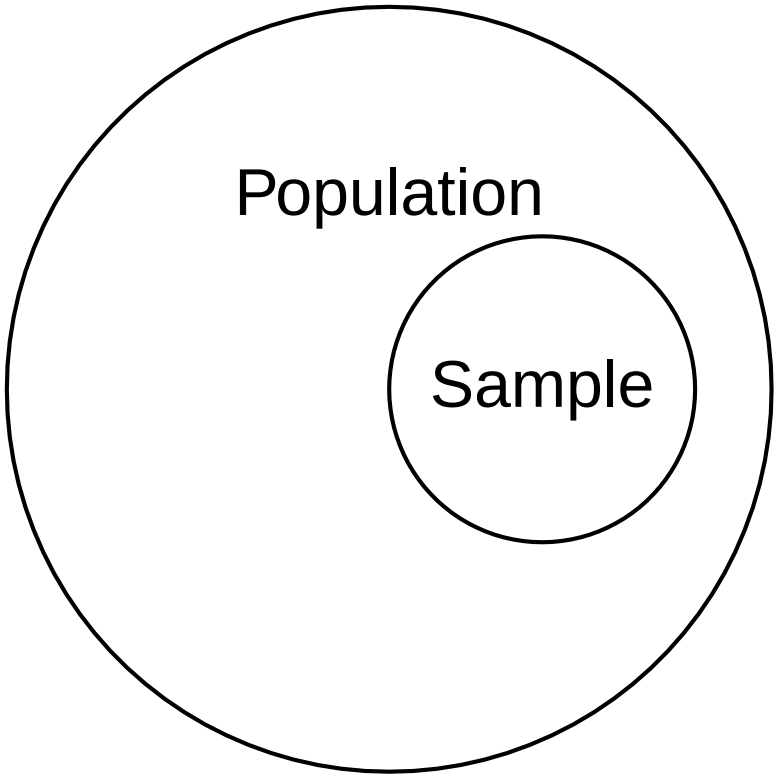
- Population: The entire set of cases that our theory applies to
- Sample: Subset of cases that we analyse



Case Selection



Case Selection



- From a Sample to a Population
- Uncertainty
- Statistical Inference

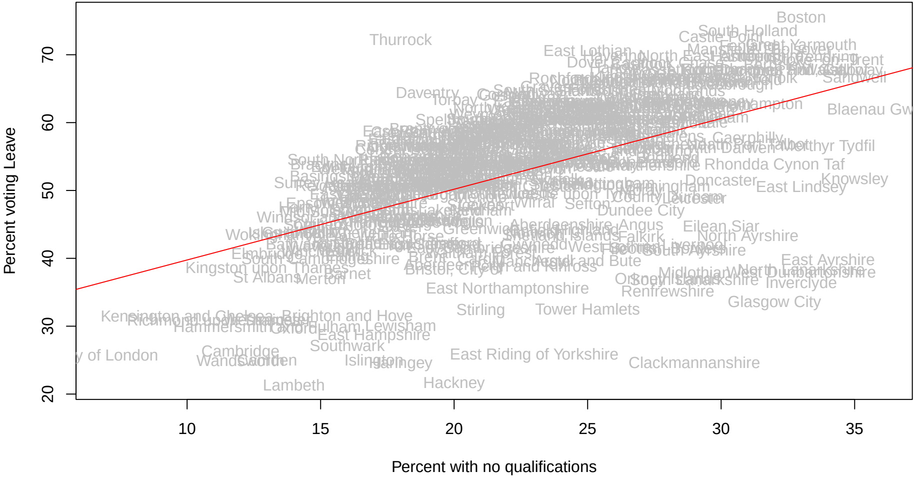
How should we measure the association between variables?

- Simple Contigency Tables (Conditional Means)
- Correlation
  - It shows direction and strength
  - But bad for predictions
  - Only Bivariate
- Regression
  - Good for Prediction and Explanation
  - Both Bivariate and Multivariate
  - The most widely used technique

What you need to remember

- What is a regression coefficient (slope)?
  - How do we interpret it?
- What is the constant?
  - How do we interpret it?
- What is the standard error of a coefficient?
- What is a p-value?
- What is the substantive (i.e. not statistical) significance of the effects?

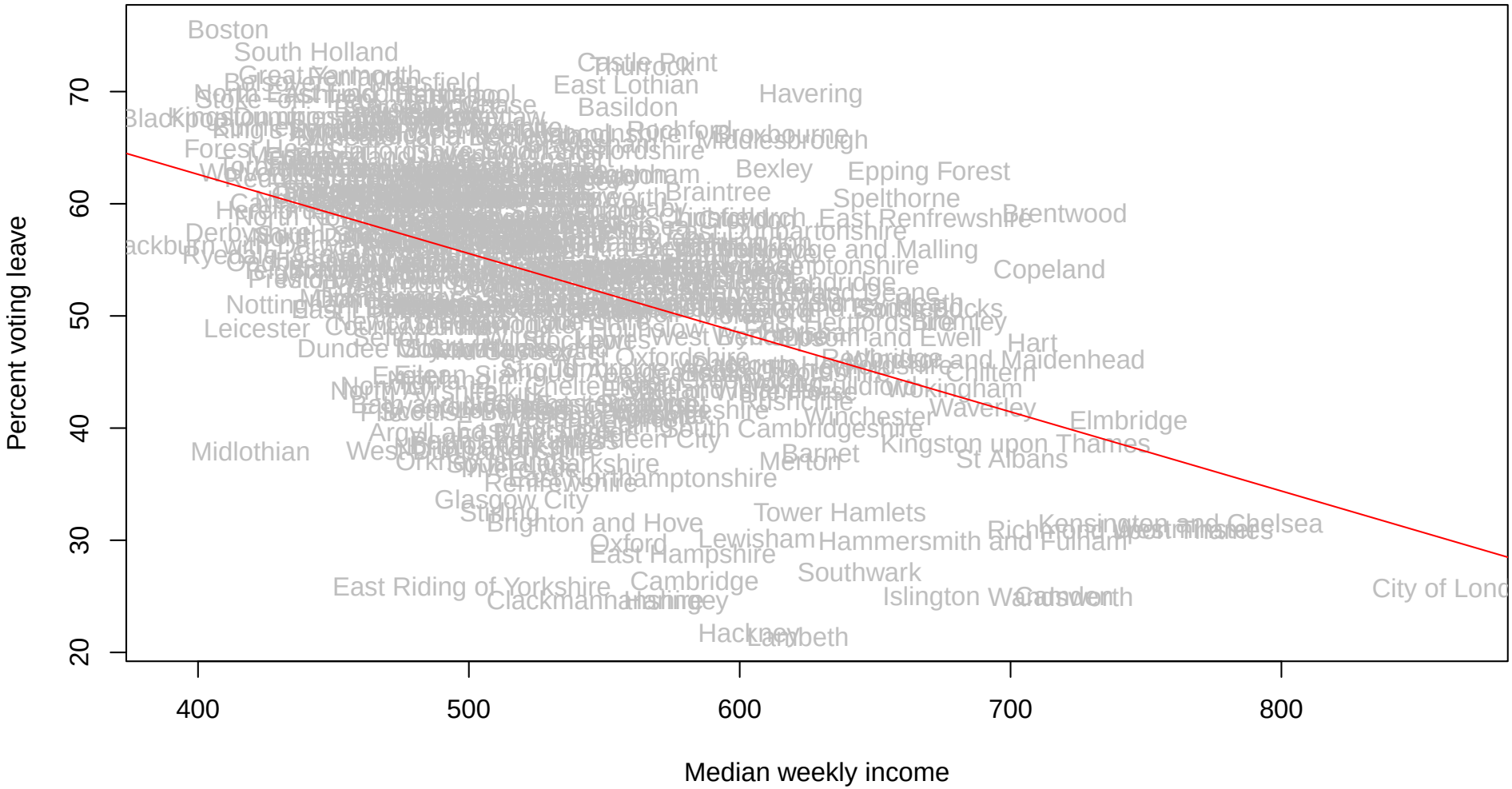
An Example?


Percent voting Leave
Percent with no qualifications
20
30
40
50
60
70
10
15
20
25
30
35
Boston
South Holland
Thurrock
East Lothian
Daventry
Blaenau Gw
Merthyr Tydfil
Rhondda Cynon Taf
Knowsley
East Lindsey
Doncaster
Dundee City
Sefton
Wirral
Eilean Siar
North Ayrshire
East Ayrshire
Inverclyde
Glasgow City
Renfrewshire
East Northamptonshire
Stirling
Tower Hamlets
Kingston upon Thames
St Albans
Merton
Barnet
Brighton and Hove
East Hampshire
Southwark
Cambridge
Islington
East Riding of Yorkshire
Clackmannanshire
Lambeth
Hackney
Falkirk
Angus
Aberdeenshire
Shetland Islands
Argyll and Bute
Midlothian
Caerphilly
y of London

An Example

```
##
## Call:
## lm(formula = brexit$leave ~ brexit$noqual)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -34.855  -3.593   1.971   5.958  24.182
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   29.33773    2.08661   14.06   <2e-16 ***
## brexit$noqual  1.04234    0.08911   11.70   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 8.945 on 377 degrees of freedom
## Multiple R-squared:  0.2663, Adjusted R-squared:  0.2643
## F-statistic: 136.8 on 1 and 377 DF,  p-value: < 2.2e-16
```

An Example


Boston
South Holland
Castle Point
Thurrock
East Lothian
Havering
Basildon
Middlesbrough
Bexley
Epping Forest
Braintree
Spelthorne
East Renfrewshire
Brentwood
Copeland
Leicester
Hart
Chiltern
Waverley
Elmbridge
Midlothian
Kingston upon Thames
St Albans
Barnet
Merton
Glasgow City
Tower Hamlets
Brighton and Hove
Lewisham
Hammersmith and Fulham
Kensington and Chelsea
Oxford
East Hampshire
Southwark
East Riding of Yorkshire
Cambridge
Islington
Camden
City of Lond
Hackney
Lambeth
Percent voting leave
20
30
40
50
60
70
400
500
600
700
800
Median weekly income

An Example

```
##
## Call:
## lm(formula = brexit$leave ~ brexit$income)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -29.804  -5.471   1.452   5.837  23.010
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   90.853597   3.644753   24.93   <2e-16 ***
## brexit$income -0.070581   0.006767  -10.43   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.289 on 360 degrees of freedom
##   (17 observations deleted due to missingness)
## Multiple R-squared:  0.232,  Adjusted R-squared:  0.2299
## F-statistic: 108.8 on 1 and 360 DF,  p-value: < 2.2e-16
```

An Example, Full Model

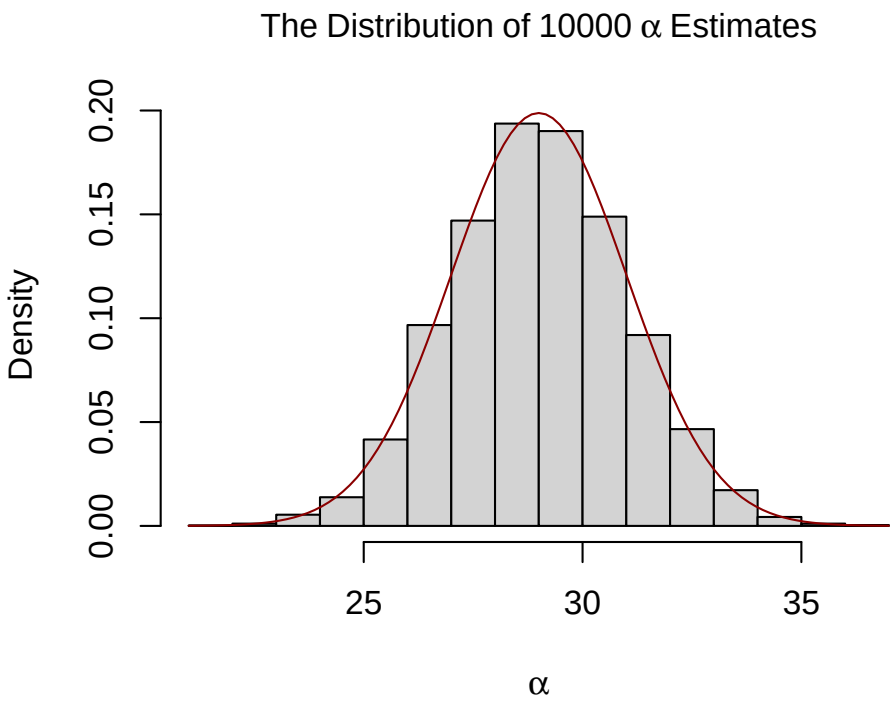
Table 1: Aggregate Models of Leave

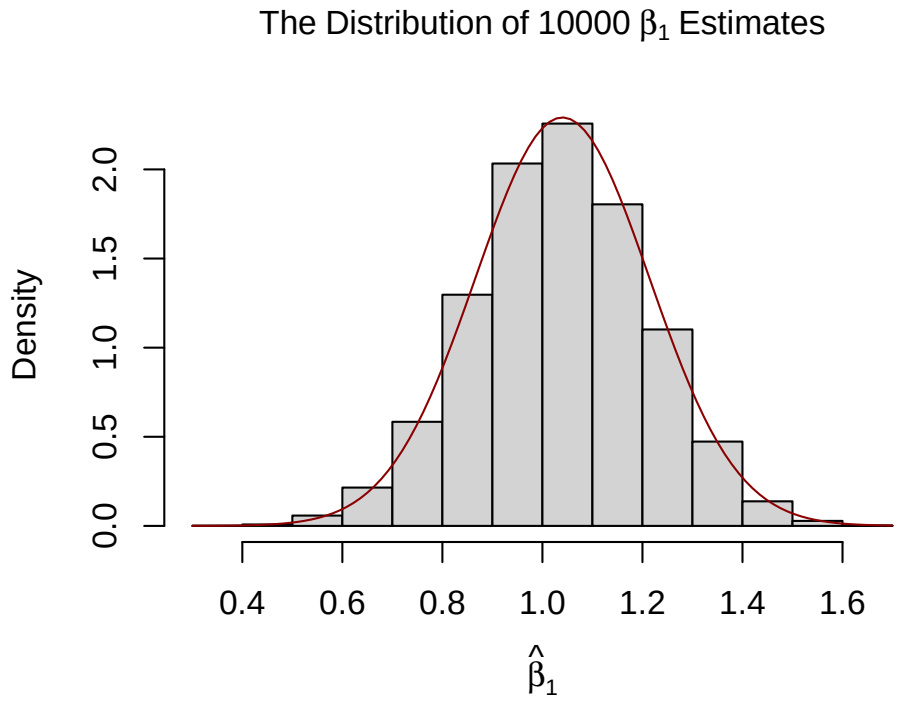
	Dependent variable:		
	Per cent Voting Leave		
	(1)	(2)	(3)
No Qualifictions	1.04^{***} (0.09)		0.72^{***} (0.12)
Weekly Income		-0.07^{***} (0.01)	-0.04^{***} (0.01)
Constant	29.34^{***} (2.09)	90.85^{***} (3.64)	56.05^{***} (6.87)

–>

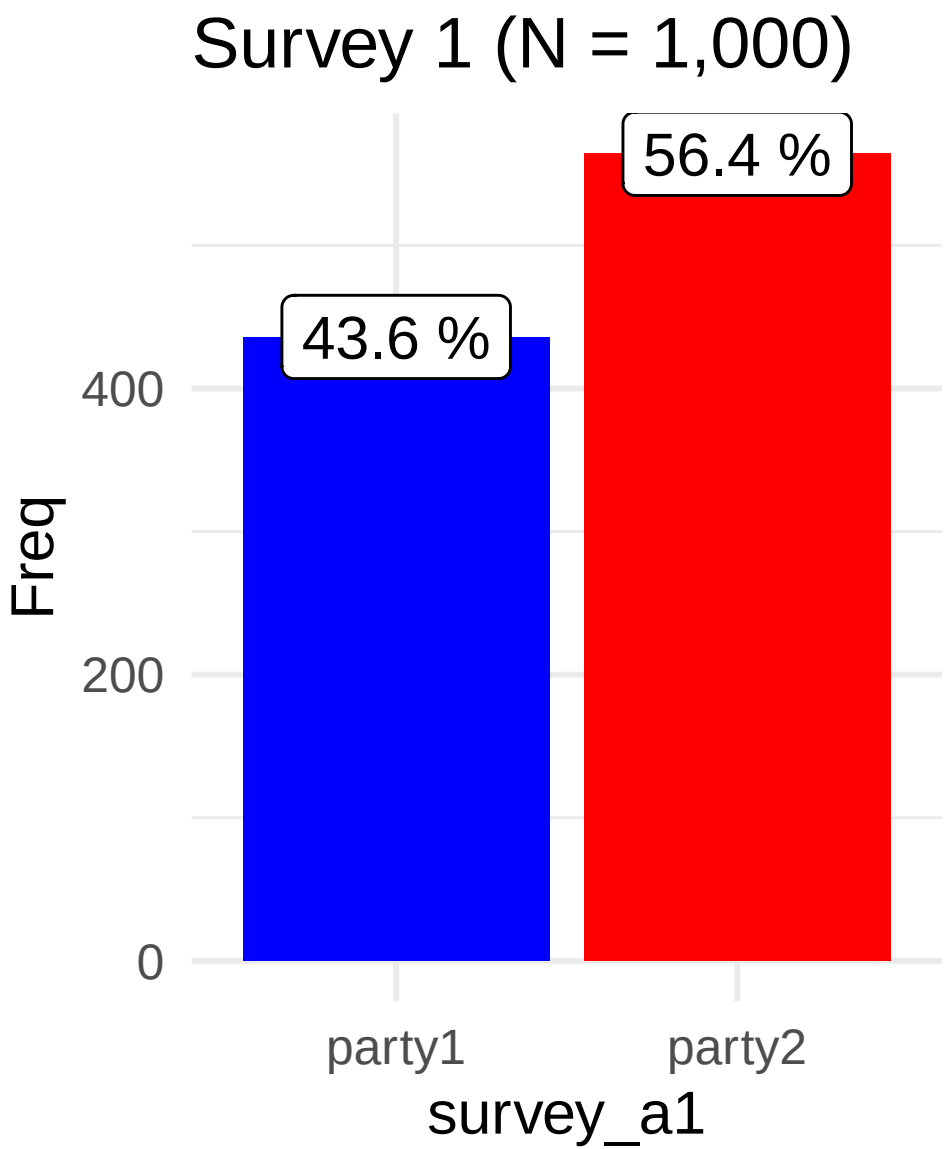
–> –>

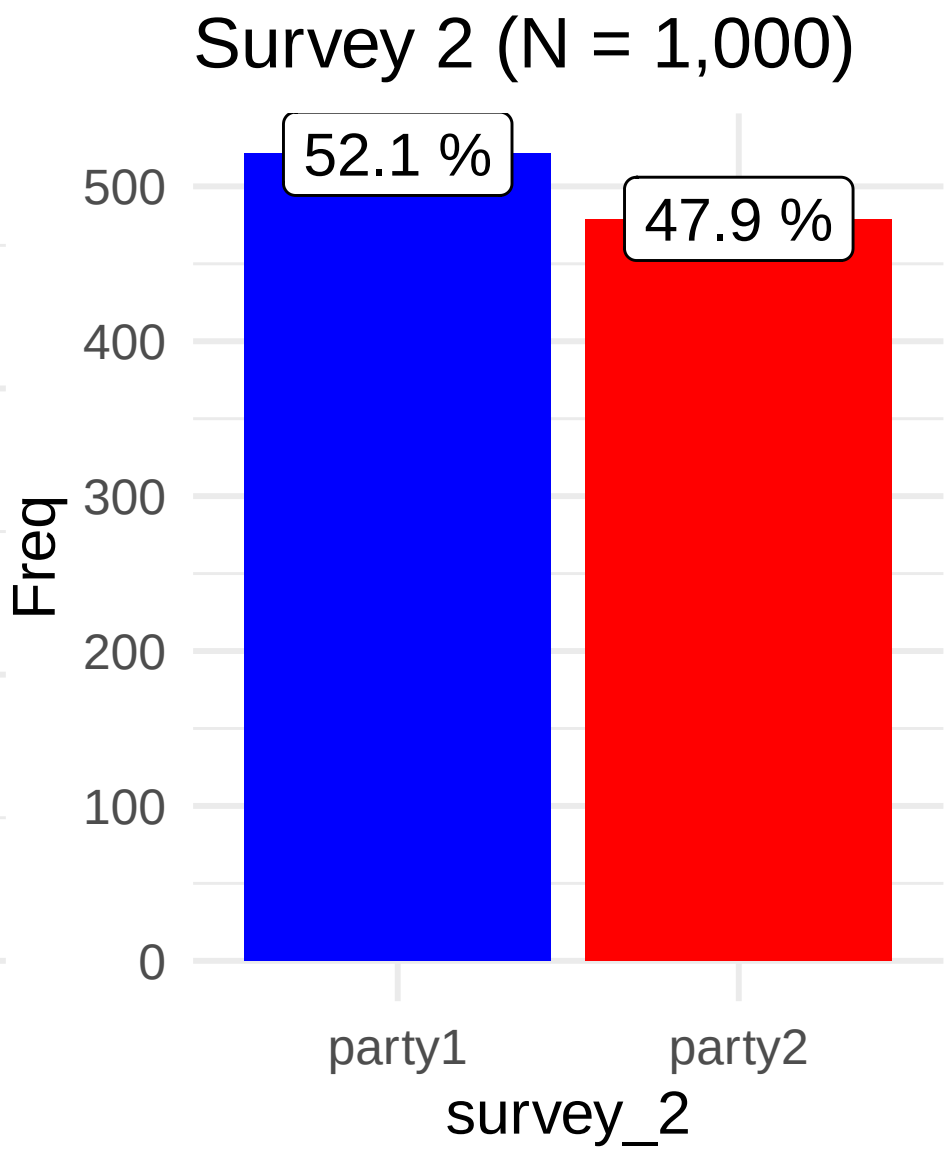
Sampling distribution of the coefficients





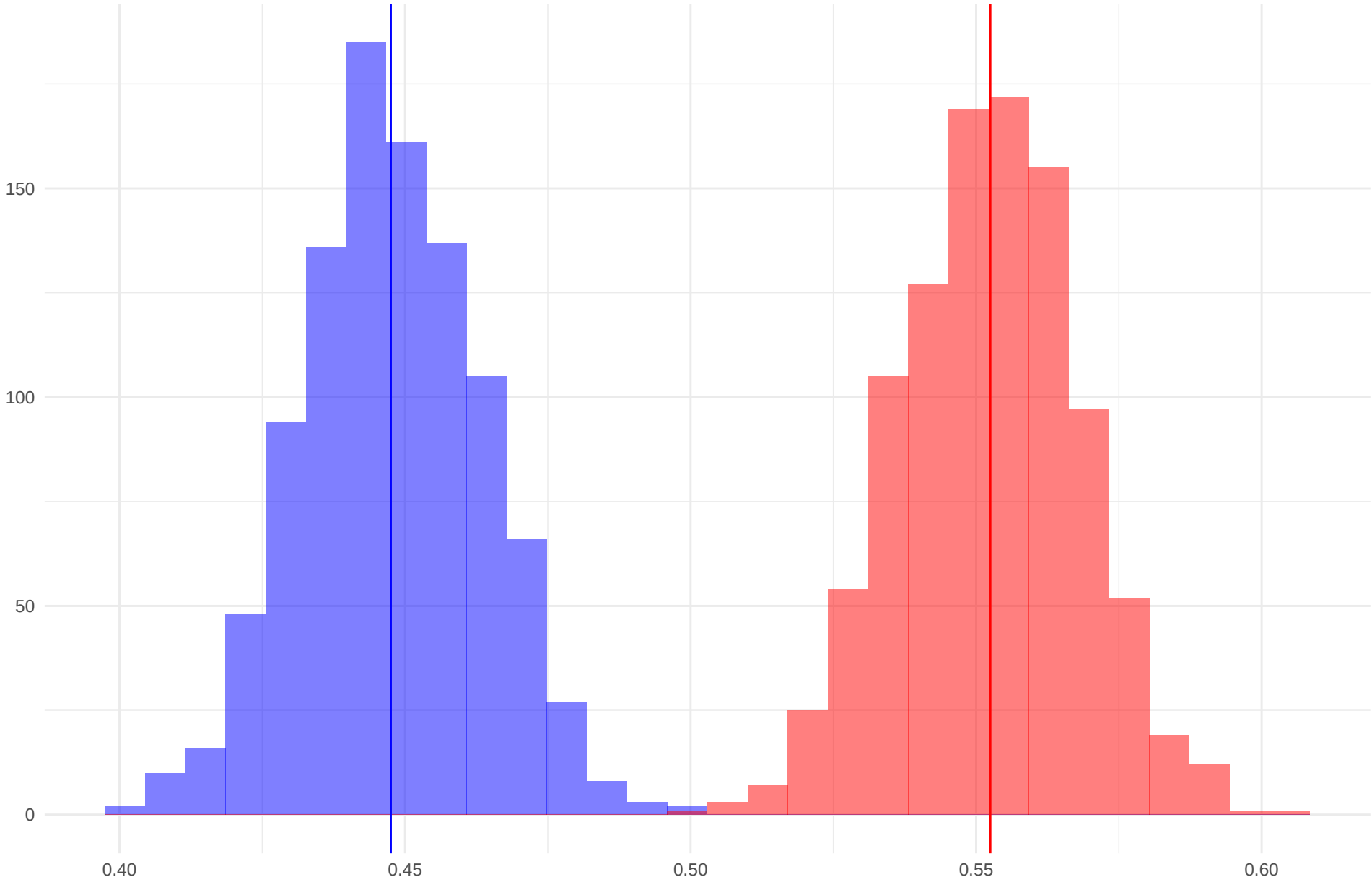
Statistical Inference





What if we had 1000 surveys of 1000 respondents?

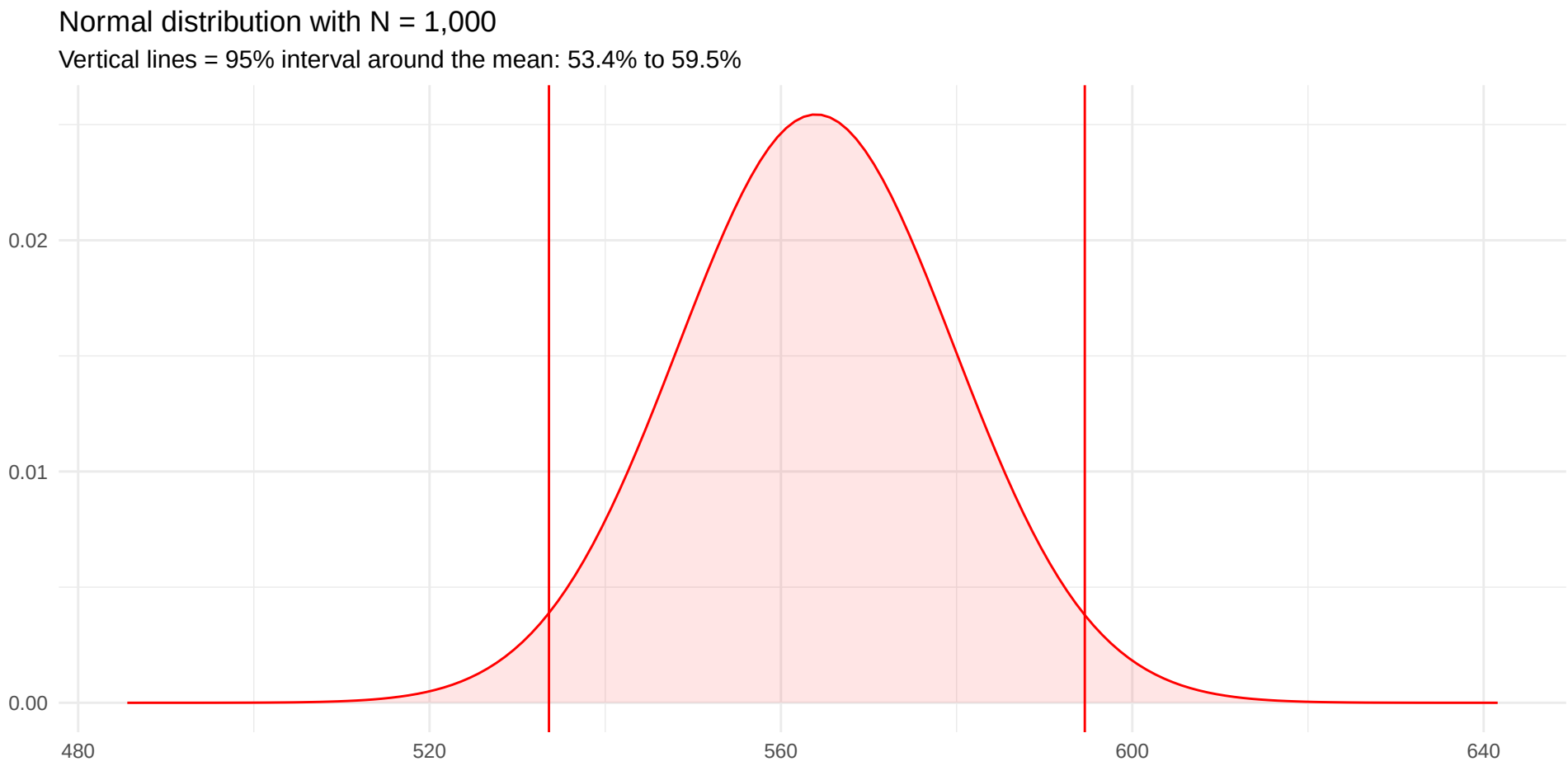
Histograms of simulated values for Party2 (red) and Party1 (blue) percentages



A first view of the confidence interval

Let's assume we are polling 1000 respondents and we find the 56% of voters prefer Party1

```
## [1] 0.44751 0.55249
```



Why bother with Stats?

Why bother with Stats?

- Engage with the evidence critically (critical: involving skilful judgement as to truth, merit, etc.)
- Those choosing, Comparative Government, Political Sociology etc, will HAVE to know how to interpret tables
- Even if your goal in life is to become a political philosopher. You need to be able to understand Stats
- The simplest argument: Read the latest issues of any top journal in political science. If you don't get the numbers you cannot engage with the content.
- You need to be able to assess the empirical evidence you cite
- Outside of Oxford: Knowing your numbers gives you a HUGE advantage!

The fear of R: Lots of people have it, it is just in your head! Get over it by practicing the materials!

Q-Step Essay

- Three topics to choose!

  1. `Does consensus democracy reduce social inequality?`
  2. `Does consensus democracy improve economic outcomes?`
  3. `Does consensus democracy improve the quality of government?`

- Same data as in the lab
- The analysis should be testing your hypothesis
- Requirement: One Plot - One Model, Full interpretation!
- But, I am confident that you can do much more and much better than this!
- This is not a matter of quantity (e.g. 15 models 125 plots), but quality.
- Try to answer the questions better than Lipjhart did!
- If something that has not been addressed by L, just figure it out using the methods you have learned
- Get more/new data
- Enjoy the process!!!

Life after Q-Step

Life after Q-Step

- Q-Step II is here for you!

Life after Q-Step

- Q-Step II is here for you!
- Q-Step Internships

Life after Q-Step

- Q-Step II is here for you!
- Q-Step Internships
- But also:

Life after Q-Step

- Q-Step II is here for you!
- Q-Step Internships
- But also: The Q-Step Advanced Methods Workshop for Second year students every TT [starting TT2023)!

Life after Q-Step

- Q-Step II is here for you!
- Q-Step Internships
- But also: The Q-Step Advanced Methods Workshop for Second year students every TT [starting TT2023)!
  - Causal Inference
  - Experiments
  - Machine Learning
  - Text Analysis

Life after Q-Step

- Q-Step II is here for you!
- Q-Step Internships
- But also: The Q-Step Advanced Methods Workshop for Second year students every TT [starting TT2023)!
  - Causal Inference
  - Experiments
  - Machine Learning
  - Text Analysis
- In the meantime, email me ASAP to express your interest!

Thank you!