# Introduction to data science for social scientists with an application to comparative politics

Teaching:
*"Professorship with a focus on Democratic Political Decision-Making"*
University of Hamburg

Daniel Bischof

November 29, 2024

University of Münster
Aarhus Universitet

Universität
Münster

SOLIDUM PETIT IN PROFUNDIS · UNIVERSITAS ARHUSIENSIS ·

## Context:

Introduction to data science for social scientists with an application to comparative politics

- ▶ BA lecture

- ▶ Last week: Survey Experiments
  *Application: designing a simple experiment using javascript*

Introduction to data science for social scientists with an application to comparative politics

▶ BA lecture

▶ Last week: Survey Experiments
*Application: designing a simple experiment using javascript*

▶ Today: Quasi-Experimental application: Regression Discontinuity Design
*Application: Turning existing research into quasi-experiments*

▶ First part of session: Advanced organizer, overview of different quasi-experimental designs

**More:**
*You can find a draft of a full syllabus with more information about my teaching profile on the handouts!*

▶ Institutions have both **mechanical** but also **psychological** effects

▶ Institutional access allows political elites to have much stronger influence on policies, polity but also citizens

# Contents & literature

▶ Institutions have both **mechanical** but also **psychological** effects

▶ Institutional access allows political elites to have much stronger influence on policies, polity but also citizens

Literature:

▶ CHAPTER 3: Norm Change *in* Bicchieri, Christina (2016). Norms in the wild: How to diagnose, measure, and change social norms. Oxford: Oxford University Press.

▶ CHAPTER 6: Regression Discontinuity *in* Cunningham, Scott (2021). Causal Inference: The Mixtape. New Haven: Yale University Press. `https://mixtape.scunning.com/06-regression_discontinuity`

▶ Pearl, Judea, & Mackenzie, Dana. (2018). The book of why: the new science of cause and effect. Basic books.

▶ Bischof, Daniel, & Wagner, Markus. (2019). Do voters polarize when radical parties enter parliament? *American Journal of Political Science*, 63(4), 888-904.

# Start

MARIA KONNIKOVA

# HOW NORMS CHANGE

**By Maria Konnikova** October 11, 2017



*I don't believe in the rules of Donald Trump's behavior rise of ... as usual"?*

As you have learned throughout the course:

▶ The last years have maximized **data availability**:
getting data is frequently a no-brainer; it is all about using the correct techniques to
access it

As you have learned throughout the course:

▶ The last years have maximized **data availability**:
getting data is frequently a no-brainer; it is all about using the correct techniques to access it

⚡ **The key challenge is to draw the correct conclusions from the excessive data at our disposal**

⚡ **AI cannot understand cause and effect**
(see also: Pearl and Mackenzie 2018)

▶ We have full access to electoral records of parties around the globe

▶ We have an enormous amount of public opinion polls
  with which we can approximate social norms

⚡ But this doesn't allow us to know the causal relationship between political parties and citizens

▶ There are obvious **causes** why parties end up in parliament; parties do not end up in parliament by 'random chance':

▶ There are obvious **causes** why parties end up in parliament; parties do not end up in parliament by 'random chance':

    1. **Parties' behavior:** e.g. positions, personal, campaigning …

▶ There are obvious **causes** why parties end up in parliament; parties do not end up in parliament by 'random chance':

1. **Parties' behavior:** e.g. positions, personal, campaigning …

2. **Citizens' behavior:** e.g. voting, protesting, opinions as such

▶ There are obvious **causes** why parties end up in parliament; parties do not end up in parliament by 'random chance':

1. **Parties' behavior:** e.g. positions, personal, campaigning …

2. **Citizens' behavior:** e.g. voting, protesting, opinions as such

3. *Parties and citizens act in an equilibrium*

1. How can we as data scientists overcome such issues of causal inference?

2. Can we use the "real world" (as opposed to the lab setting) to address these issues in our example?

# *Theory:* how parties drive public opinion

Institutional access: …

▶ increases media presence for parties (visibility)
  (Abou-Chadi and Krause 2020)

▶ can be a shock-like event for citizens (shock)
  (Bischof and Wagner 2019)

▶ legitimizes extreme opinions, might lead to polarization (legitimization)
  (Bischof and Wagner 2019; Valentim 2021)

# *Theory:* how parties drive public opinion

Institutional access: ...

- ▶ increases media presence for parties (visibility)
  (Abou-Chadi and Krause 2020)

- ▶ can be a shock-like event for citizens (shock)
  (Bischof and Wagner 2019)

- ▶ legitimizes extreme opinions, might lead to polarization (legitimization)
  (Bischof and Wagner 2019; Valentim 2021)

➡ The perception of social norms is changing

▶ Use of real world situations (=data) where the happening of an event (=parliamentary access) is determined by a random or external process.

▶ <u>Objective:</u> Imitate randomization in experiments to estimate causal effects in observational data.

▶ *Example:* natural disasters

The electoral threshold is the **minimum share of votes** that a candidate or political party requires before they become **entitled to representation** in a legislature.

# *Electoral thresholds:* an Overview (Valentim, 2021)

**Table 1.** List of Elections Included in the Sample and Their Electoral Thresholds.

| Elections | Electoral threshold |
|---|---|
| Austria 2008, 2013, 2017 | 4 |
| Bulgaria 2014 | 4 |
| Croatia 2007 | 5 |
| Czech Republic 1996, 2002, 2006 | 5 |
| Denmark 1998, 2001, 2007 | 2 |
| Estonia 2011 | 5 |
| Germany 1998, 2002, 2005, 2009, 2013, 2017 | 5 |
| Greece 2009, 2012, 2015 (Jan), 2015 (Sep) | 3 |
| Hungary 1998, 2002, 2018 | 5 |
| Israel 1996, 2003 | 1.5 |
| Israel 2006, 2013 | 2 |
| Italy 2006 | 4 |
| Italy 2018 | 3 |
| Latvia 2010, 2011, 2014 | 5 |
| Netherlands 1998, 2002, 2006, 2010 | 0.67 |
| Norway 1997, 2001, 2005, 2009, 2013 | 4 |
| Poland 1997, 2001, 2005, 2007 | 5 |
| Romania 1996 | 3 |
| Romania 2004 | 5 |
| Serbia 2012 | 5 |
| Slovakia 2010, 2016 | 5 |
| Slovenia 1996 | $\approx 3$ |
| Slovenia 2004, 2008 | 4 |
| Sweden 2006, 2014 | 4 |
| Ukraine 1998 | 4 |

## *Solution:* Regression Discontinuity Design (RDD)

1. Electoral systems come with electoral thresholds

2. Parties must receive more votes than the threshold to enter parliament

1. Electoral systems come with electoral thresholds

2. Parties must receive more votes than the threshold to enter parliament

$\rightarrow$ some parties are just above the threshold

## *Solution:* <u>R</u>egression <u>D</u>iscontinuity <u>D</u>esign (RDD)

1. Electoral systems come with electoral thresholds

2. Parties must receive more votes than the threshold to enter parliament

$\rightarrow$ some parties are <u>just above</u> the threshold

$\leftarrow$ some parties are <u>just below</u> the threshold

## *Solution:* <u>R</u>egression <u>D</u>iscontinuity <u>D</u>esign (RDD)

1. Electoral systems come with electoral thresholds

2. Parties must receive more votes than the threshold to enter parliament

→ some parties are <u>just above</u> the threshold

← some parties are <u>just below</u> the threshold

! If we focus on parties very close to the threshold, e.g., within 1%, the parties are quasi-randomly in parliament or not

# *Solution:* <u>R</u>egression <u>D</u>iscontinuity <u>D</u>esign (RDD)

1. Electoral systems come with electoral thresholds

2. Parties must receive more votes than the threshold to enter parliament

$\rightarrow$ some parties are <u>just above</u> the threshold
  **treatment**

$\leftarrow$ some parties are <u>just below</u> the threshold
  **control**

! If we focus on parties very close to the threshold, e.g., within 1%, the parties are quasi-randomly in parliament or not

► Schleswig-Holstein (2022): 4,4%

► Schleswig-Holstein (2022): 4,4%       ► Hamburg (2020): 5,3%
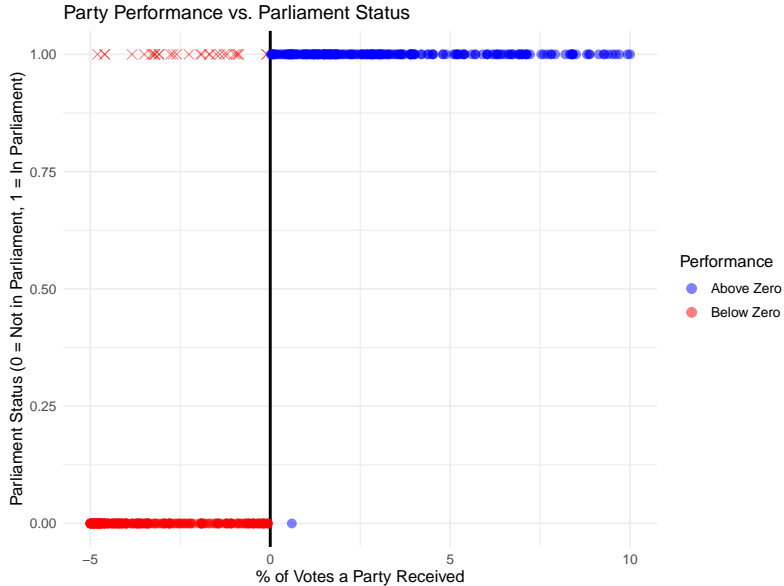
- ▶ **Discuss** in a group of 5 during the break:
- ▶ Repeat the intuition behind the RDD.
- ▶ What does the RDD allow us to do as data scientists?
- ▶ Come up with examples for similar thresholds in the real world *(e.g. driving under the influence)*

End

1. Repetition of intuition of RDD

2. Mathematical background of RDD:
   2.1 Potential outcomes framework
   2.2 how to calculate RDDs

All R code available on my GitHub: `https://github.com/danbischof/example`

## Me as a teacher:

- ▶ My potential teaching portfolio for Hamburg is laid out on the handout

- ▶ I'm able to teach both **substantive courses in CP (and beyond)** as well as **methods, research design** or combinations thereof

- ▶ In my current teaching I already seek to combine these two components; your position would allow me to do that even more extensively

- ▶ I would love to teach on the graduate level: collective decision-making + WISO

- ▶ Evaluations show a **very high student satisfaction, also for methods teaching** *(e.g. Causal Inference, BA, Münster)*

# Thank you very much for your attention!
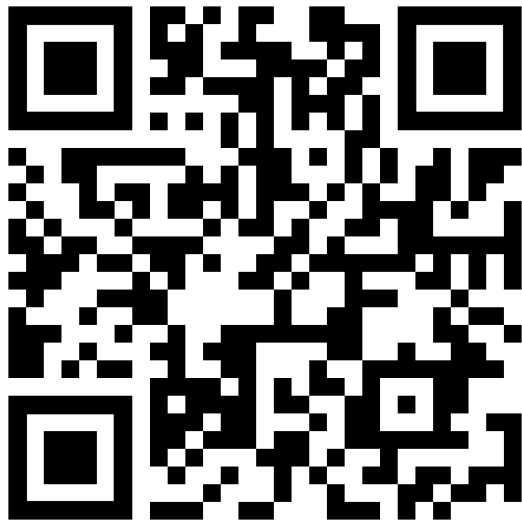
# References

Abou-Chadi, Tarik, and Werner Krause. 2020. "The causal effect of radical right success on mainstream parties' policy positions: A regression discontinuity approach." *British Journal of Political Science* 50 (3): 829–847.

Bischof, Daniel, and Markus Wagner. 2019. "Do voters polarize when radical parties enter parliament?" *American Journal of Political Science* 63 (4): 888–904.

Pearl, Judea, and Dana Mackenzie. 2018. *The book of why: the new science of cause and effect.* Basic books.

Valentim, Vicente. 2021. "Parliamentary representation and the normalization of radical right support." *Comparative Political Studies* 54 (14): 2475–2511.

## My teaching philosophy:

- Competency and practice-oriented knowledge transfer
- Critical reflection of literature, methods, arguments, and instructors
- Match of my substantial research interests with teaching content
- Teaching how to deal with data embedded into substantial research questions
- Diversity, flat hierarchy:
  - All courses can be offered in German or English.
  - All courses, including lectures, use diverse exercises (e.g., peer review systems, group discussions, surveys, mini-conferences, expert panels with examples from journalists, writing blog posts, and much more)
  - As my evaluations show, I insist on very flat hierarchies: I want to discuss with students, engage in dialogue.
  - I ensure balanced literature in terms of gender, methods, and fields of study.
  - "Desired content": I inquire in the first session of seminars if students are missing certain topics within the seminar and try to incorporate these.

## Possible BA Offerings at the University of Hamburg:

- **Introductory courses:**
  - Lecture: "Introduction to Comparative Democracy Research"
  - Political institutions in comparative perspective
  - **Introduction to data science for social scientists with an application to comparative politics**
- **Intro to data science:**
  - Basic Course in Empirical Social Research Methods
  - Causal Inference in Social Science
  - Generating and Managing Data
  - Analyzing Text as Data
  - Introduction to R/Stata (with some Python + JavaScript)
- **Advanced courses:**
  - SE Political Culture and Social Norms
  - SE Cleavages: Political Conflict Lines and Their Shifts
  - SE Revolutions and Democracy Research
  - SE Comparative Democracies
  - SE The Rise of the Radical Right in Western Europe
  - SE Political Representation
  - Reading course: "How Democracies Die" by Steven Levitsky & Dan Ziblatt
  - Reading course: "Engineers of Jihad" by Diego Gambetta & Steffen Hertog
- **Colloquia**

- Substantial Content: **Institutions, Democracy Research, and Political Culture**
  - Introduction to political institutions
  - How do political institutions emerge and change? How do they affect politics and societies?
  - Introduction to research on electoral systems
  - Democracies and political culture in comparative perspective
  - Extremism: how it arises, what fosters it, how it impacts societies.

**Possible MA Offerings at the University of Hamburg:**

- **Democracy Research:**
  - Democratic Decision-making in Times of Backsliding
  - Democratic Backsliding in Comparative Perspective
  - The Rise of Democracies: Revolutions
  - How can we measure democratic support?
  - When extreme parties enter the state: causes and consequences
- **Comparative Extremism**
  - Migration and Radical Right Voting
  - The Roots of Political Extremism
  - The Legacy of Political Regimes and Extremism
  - The Role of Social Media for Extremism and Hate
  - Remedies for Extremism
- **Minorities and Democracy:**
  - Political Representation of Minorities
  - Political Institutions and Women's Rights
  - Orchestrating Hate: Reasons for Political Violence and Hate Crimes
  - Antisemitism: Legacies, Continuity, and Conspiracies
- **Data Science & Digitalization:**
  - Research Design
  - Causal Inference
  - Experiments in Social Sciences
  - Data collection and measurement in Social Sciences (e.g. web-scrapping, 'text as data', integrating AI, )
  - Advanced Coding (R, Python, JavaScript)
- **Project Module/Research Seminars**
  - Key challenges to contemporary societies (incl. Inequality, Racism, Global Warming, Poverty, Populism)
  - Advanced Causal Inference (design your own experiment, conduct data collection, write pre-analysis plan)
  - Applying Generative AI as a Data Scientist: promises and pitfalls
- **Colloquia**