

*Last Updated January 22, 2020*

# Applied Research Methods II

## (Research Practicum, Semester 2)

**Instructor Information**Name: Mike DenlyEmail: [mdenly@utexas.edu](mailto:mdenly@utexas.edu)Office: Mezes 3.232Office Hours (OH): Thurs. 1:45-4:45pmOH Booking: [mikedenly.youcanbook.me](http://mikedenly.youcanbook.me)Website: [www.mikedenly.com](http://www.mikedenly.com)**Course Information**Abbreviation: GOV355DUnique Number: 38010Time: Friday 1-4pmRoom: RLP 0.122Flags: [Writing](#); [Indep. Inquiry](#); [Quant. Reasoning](#)Website: [canvas.utexas.edu](http://canvas.utexas.edu)**Supervising Professor Information**Name: Dr. Mike FindleyEmail: [mikefindley@utexas.edu](mailto:mikefindley@utexas.edu)Office: BAT 3.102Website: [www.michael-findley.com](http://www.michael-findley.com)**Research Lab Information**Innovations for Peace and DevelopmentLab Location: BEL (Stadium) 214Lab Access Code: Ask the instructorWebsite: [www.ipdutexas.org](http://www.ipdutexas.org)

## 1. Course Description

This course is the second semester of a two-semester program that attempts to provide undergraduate students with a fairly comprehensive introduction to the research process in the social sciences. As part of this program, students will attend regular classes, write their own first-rate quantitative research paper, and gain internship experience with [Innovations for Peace and Development \(IPD\)](#). Students are required to take both semesters of this two-semester, interdisciplinary research program.

During the first semester, the classroom part of the course covered the essential elements of applied social science research, including arguments, concepts, measures, causality, and basic statistics. Given that knowledge of statistical software, text editors, reference management software, and mapping software is increasingly helpful for success in the social sciences, the course also provided training in [Stata](#), [R](#), [Python](#), [L<sup>A</sup>T<sub>E</sub>X](#), [Mendeley](#), and [ArcGIS](#). At the end of the first semester, students completed their own well-developed Research Proposals in lieu of a final exam.

During the second semester, classroom instruction will cover experiments, data structures, data cleaning, hypothesis testing, measurement challenges, linear regression, as well as the basics of panel data, regression discontinuity designs, difference-in-differences, synthetic controls, logistic regression, and network analysis. Training in the above software programs

will continue during the second semester as well. At the end of the second semester, students will complete their own research projects, write-up their results in a formal paper, and present their findings to the class.

## 2. Course Requirements

### 2.1. Prerequisite Coursework

Students need to have already taken Applied Research Methods I (Research Practicum, Semester 1) to enroll in this course. The instructor will only grant exceptions to students who studied abroad during the previous semester and already completed a research proposal that they will then transform into a full-fledged research paper by the end of the semester. Also note that this is an upper-division undergraduate course. Students with previous coursework in political science, economics, sociology, and/or statistics will likely find the course easier.

### 2.2. Required Software and Resources

This course makes use of **Stata**, **R**, **L<sup>A</sup>T<sub>E</sub>X**, **Mendeley**, and **ArcGIS**. Prior knowledge of any of these software programs is not required. I will teach you the basics of all of these programs during the course.

- **R**. For instructions on how to freely download R and its companion program, **R Studio**, consult [here](#).
- **L<sup>A</sup>T<sub>E</sub>X**. Windows users can freely download MiKTeX [here](#). Mac users can freely download MacTeX [here](#). Advanced users may want to consider downloading [SublimeText](#), to be used in combination with [Sumatra PDF](#) (instructions [here](#)). In class, we will be using [Overleaf](#), a program that allows users to use L<sup>A</sup>T<sub>E</sub>X online—that is, without the need to have it installed on one’s computer.
- **Mendeley**. This reference management software program is freely available [here](#). After learning how to use **Mendeley**, or one of its competitors, such as **Zotero** or **EndNote**, it will never be necessary to construct your own bibliography manually ever again.
- **Stata**. Although **Stata** is a relatively expensive commercial software, as a UT student you have free access to **Stata** through the [UT Austin Stats Apps Server](#) as well as numerous computer labs on campus. The IPD lab (BEL [Stadium] 214), which you have access to as part of this class, has **Stata** as well.
- **ArcGIS**. At first, we will be using computers from the computer lab with **ArcGIS** already installed on them. Providing that there is student interest, we will consider obtaining a one-year license for interested students. As with **Stata**, students can access **ArcGIS** in numerous computer labs on campus, including the IPD lab.

To get help with these programs and others, there are two resources that we will utilize:

- *Data Camp*. I have signed up our class for *free* courses from [Data Camp](#). It is an online platform that provides hundreds of courses to learn new skills. The courses are interactive and fun. We will be using some of these courses from Data Camp as required homework. Overall, students will be able to use 900 free hours of coursework.
- *Lynda*. You can also access free courses through [UT-Austin's Lynda Portal](#).

## 2.3. Readings

Students must purchase (or borrow from a library) the course's primary textbooks:

- Gerring, John, and Dino Christenson. 2017. *Applied Social Science Methodology: An Introductory Guide*. Cambridge: Cambridge University Press.
- Kelstedt, Paul, and Guy Whitten. 2018. *The Fundamentals of Political Science Research*. Third Edition. Cambridge: Cambridge University Press.
- Imai, Kosuke. 2017. *Quantitative Social Science: An Introduction*. Princeton, NJ: Princeton University Press. <http://qss.princeton.press/student-resources-for-quantitative-social-science>
- King, Gary, Robert Keohane and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
- Li, Quan. 2018. *Using R for Data Analysis in the Social Sciences: A Research Project-Oriented Approach*. Oxford: Oxford University Press.
- James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2014. *An Introduction to Statistical Learning: With Applications in R* New York: Springer. [This book is freely available via <http://www-bcf.usc.edu/~gareth/ISL>]
- Wickham, Hadley and Grolemund, Garrett. 2017. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. Sebastol, California: O'Reilly Media. [This book is freely available via <https://r4ds.had.co.nz>]

For some weeks, I supplement the textbook readings with other required and optional readings. When these articles can be easily found on the UT Austin Library webpage, I will ask students to download the article(s) themselves—to ensure students know how to use the library website; otherwise, I will post the article(s) on the class website, Canvas. For more information on the specific reading assignments for each week, refer to the Class Schedule (below). Optional readings are not required for each class period, and reading them will not enable students to receive extra credit. However, I may use these readings to supplement the textbook in case it is necessary to facilitate comprehension of important topics.

## 2.4. Attendance, Quizzes, and Participation

All students must come to class prepared, having completed the readings before class. At the beginning of each class, I will give everyone a five-question, multiple-choice quiz.

The quiz serves three purposes. First, the quiz will help keep track of attendance and serve as a commitment device for students to attend class and on-time. Even if students miss both questions on the quiz but are present for class, they will receive full credit toward attendance for the respective class. Overall, attendance will account for 5% of students' final grades.

Second, since the quiz will only cover the most basic ideas from the required reading, the quiz will serve as a reward: you should receive 100% every time if you read. To give students some cushion for bad days or extenuating circumstances, I will drop your lowest 2 scores. I will make no other accommodations should you miss class for any reason or arrive late and miss the quiz. In total, students' average quiz score will comprise 10% of your final grade.

Third, the quiz will help ensure that students are ready to discuss the material and do not rely entirely on my lecture to learn the materials. The material is challenging, and passive learning will generally not suffice for students to perform well in the course. Since participation comprises 5% of the final grade, I will post a 1-5 cumulative score for the semester on Canvas for each student after the third class and gradually update it during the semester, as appropriate. This way, the final participation grade will not come as a surprise to students at the end of the semester. As stipulated in the Policies section of this syllabus, I will make every possible effort to ensure that students feel comfortable participating. To ensure that you receive a good grade for participation, please try to make at least one meaningful contribution to discussion each class.

## 2.5. IPD Internship and Workshops (and Extra Credit)

Each week, the Task Team Leader from the respective [Innovations for Peace and Development \(IPD\)](#) teams will send students/team members assignments. On average, these assignments will take students around 5 hours to complete. By Friday at 12 noon each week, students will update an individual time-tracking Google Sheet where they will keep track of their hours and the tasks that they complete for their IPD internship. The Task Team Leaders from the respective research team will sign-off on each student's hours. Based on these Google Sheets and feedback from the respective Task Team Leaders, students will receive a grade for their IPD internship participation. That grade will comprise 15% of students' final grade for the course.

Students who consistently and satisfactorily complete their assignments and update their Time-Tracking Google Sheets on-time each week have been promoted from Research Apprentice to Research Affiliate at the end of the first semester. If you have any questions, please ask your IPD Team Leader.

Most weeks, IPD will also offer workshops in programs such as **Stata** and **R**. On many occasions, I will offer these workshops, and integrate them with our course. For each workshop that students attend outside of our normal class time, students will receive 0.33 points toward their final grade. To receive this credit, students will need to sign-in each time with the IPD Task Team Leader offering the workshop. I will then obtain a list of students who attended this workshop directly from the respective IPD Task Team Leader.

## 2.6. Office Hours and Meetings with the Instructor

Students must meet with me at least two times by April 9. During the first meeting, we will discuss necessary steps to make your Research Proposal from the first semester into a full-fledged Research Paper by the end of this second semester of the course. The second meeting will entail discussion of students' progress on their Research Papers. In case students have a work or class commitment, I will be happy to meet with them outside my designated office hours. To obviate having to wait in long lines, students may [book an office hours slot here](#). Students who do not meet with me at least twice before April 9 will lose one point toward their final grade. This requirement is in place because last-minute efforts will generally not suffice to do well in the course.

## 2.7. Homework Assignments (and Extra Credit)

Homework assignments, including **Stata** and **R** assignments, will comprise 10% of students' final grades. For each written assignment that students complete using  $\text{\LaTeX}$ , they will receive two extra credit points for the particular assignment.

## 2.8. Research Week Presentations

All students must participate in UT Research Week. Participation entails making poster presentations for: (1) the UT Government Department on Tuesday, April 14, 2020 in Batts 5.108; and (2) the whole university on Thursday, April 16, 2020 in Union Ballroom 3.202. To qualify for research week, you will need to [submit an abstract to the university](#) by Thursday, February 27, 2020. To ensure that the university can have your posters ready by April 17, you will need to [submit them your poster to the Undergraduate Research Office](#) by April 5, 2020 at 11:59pm. You may obtain more information on the whole process on the [UT Research Week website](#). Since the university will not accept late submissions, the course late submission policy not apply; all students who do not fulfill these requirements will receive zero for the respective assignments.

## 2.9. Research Paper, Sub-Assignments, Referee Reports, and Presentation (and Extra Credit)

Having students write a first-rate Research Paper is a primary goal for the course. In the past, students have produced research papers relating to foreign aid, governance, political economy, political violence, peace processes, international development, and many other topics. Many students have been accepted (and funded) to present their research at the Midwest Political Science Association meeting in Chicago, IL. On that score, two students in the class will be presenting at Midwest this year.

During the first semester, students produced a Research Proposal with the following elements:

1. An introduction to a puzzle in an academic literature of the student's choosing
2. A clear description of the dependent variable
3. A clear critique of an existing literature related to the student's topic, covering how different independent variables have explained their dependent variable
4. A theory/argument that explains the puzzle and mechanistically traces why it causes the dependent variable
5. A research design, articulating how the student plans to test the theory/argument
6. A section describing what else students will need to do during the second semester to complete the paper.

During this second semester of the course, students will make the appropriate updates to their Research Proposals from the first semester. After some hard work, students should be able to convert their Research Proposal into a complete Research Paper by the end of this second semester of the course. To ensure students are on track to complete their Research Papers by the end of the semester (deadline: May 17) and receive adequate feedback along the way, the course will contain the following sub-assignments:

1. Revise & Resubmit Assignment [due January 24]
2. Another paper revision, including an Appendix describing the variables (and description of qualitative cases/interview questions, if mixed method study) [due February 14]
3. Research Week abstract [due February 21 to instructor; February 27 to university]
4. Clean dataset [due March 8]
5. Research Week poster material submission [due April 5]
6. A draft paper with all sections, including an analysis section with estimated regression results, but excluding the conclusion [due April 17]
7. A Referee Report of another student's progress on his/her paper to date [2.5-3 double-spaced pages, due May 1]

## 8. A Presentation of students' final Research Papers [May 8]

For each of these sub-assignments (except the Poster and Presentation), students will receive an extra two points if they submit their assignments using  $\text{\LaTeX}$ . All of these sub-assignments must include a bibliography—generated automatically using **Mendeley** or another program such as **Zotero** or **EndNote**—as learned during the first semester of the course. Students should not be wasting time by manually generating a bibliography.

I will provide feedback on relevant sub-assignments within one week of submission or earlier. The Referee Report will also provide useful feedback for students. Additionally, the Referee Reports serve as a medium for students to learn how to critique others' quantitative studies in a respectful way, thereby enabling the students to become better scholars. The Referee Report will account for 5% of students' overall grades.

After completing all of the sub-assignments and the Referee Reports, students will submit a final Research Paper. The final Research Proposal should incorporate feedback from all the previous assignments and the student Referee Reports as well as include a bibliography made with **Mendeley** or another program. The final Research Paper will comprise 15% of students' final grades and will be due on May 17 at 12 noon.

As with the sub-assignments, the Research Proposal also presents an opportunity for extra credit: students who write their Research Paper in  $\text{\LaTeX}$  will receive an two extra points toward their final grade on the Research Proposal. During office hours, I would be more than happy to help students who are having issues with formatting anything in  $\text{\LaTeX}$ . I cannot provide the same support for anything written in Microsoft Word, Libre Office, Google Docs, etc.

On the final day of class, students will give 5-minute presentations of their work. The presentation will comprise 5% of students' grades. Since we will not have time to cover  $\text{\LaTeX}$  Beamer during class, students will not be able to receive extra credit for submitting their presentations in  $\text{\LaTeX}$  Beamer.

## 3. Policies

### 3.1. Grading Rubric

- Attendance: 5%
- Class Participation: 5%
- Quizzes: 10%
- Homework, including Stata/R Assignments: 10%
- Revise & Resubmit Assignment: 5%
- Paper Submission with Codebook Appendix: 5%

- Clean Dataset: 5%
  - Research Week Poster: 5%
  - Empirical Section/Results Sub-Assignment: 10%
  - Referee Report: 5%
  - Research Paper: 15%
  - Final Presentation: 5%
  - IPD Internship Evaluation: 15%
- 
- Extra Credit: IPD Workshops: 0.33 potential points toward students' final grades for each workshop attended outside of class time.
  - Potential Penalty: Instructor Meetings: 1 potential point for not meeting with the instructor twice before April 9.

### 3.2. Grading Scale

- |                    |                      |
|--------------------|----------------------|
| • 92.50-100 (A)    | • 72.50-76.49 (C)    |
| • 92.49-89.50 (A-) | • 69.50-72.49 (C-)   |
| • 86.50-89.49 (B+) | • 66.50-69.49 (D+)   |
| • 82.50-86.49 (B)  | • 62.50-66.49 (D)    |
| • 79.50-82.49 (B-) | • 59.50-62.49 (D-)   |
| • 76.50-79.49 (C+) | • 59.49 or below (F) |

### 3.3. Grade Rounding

The above grading scale already incorporates very generous grade rounding, not to mention the multitude of extra credit opportunities. Accordingly, there will be no additional rounding of grades under any circumstance.

### 3.4. Grade Posting on Canvas

I will post all grades to the class website, Canvas. I will also use the option where students may discern the average score of the class. This way, students will know where they stand by the end of semester.



### **3.5. Grade Appeals**

If you would like to appeal your grade on any assignment, you must make the request to me in writing, over email, within 5 days of receiving your grade. In your grade appeal, you must specify the reason(s) why you think I misgraded the paper. Acceptable reasons include those pertaining to the concepts and material covered during the course. I will not consider requests for grade changes that are not germane to the course.

### **3.6. Independent Inquiry Flag**

This course carries the Independent Inquiry flag. Independent Inquiry courses are designed to engage you in the process of inquiry over the course of a semester, providing you with the opportunity for independent investigation of a question, problem, or project related to your major. You should therefore expect a substantial portion of your grade to come from the independent investigation and presentation of your own work.

### **3.7. Quantitative Reasoning Flag**

This course carries the Quantitative Reasoning flag. Quantitative Reasoning courses are designed to equip you with skills that are necessary for understanding the types of quantitative arguments you will regularly encounter in your adult and professional life. You should therefore expect a substantial portion of your grade to come from your use of quantitative skills to analyze real-world problems.

### **3.8. Quantitative Papers Only**

Although this course covers some qualitative research, the focus of the course is quantitative. Accordingly, students must write a quantitatively-oriented or mixed methods paper for their Final Research Papers.

### **3.9. Writing Flag**

This course carries the Writing Flag. Writing Flag courses are designed to give students experience with writing in an academic discipline. In this class, you can expect to write regularly during the semester, complete substantial writing projects, and receive feedback from your instructor to help you improve your writing. You will also have the opportunity to revise one or more assignments, and you may be asked to read and discuss your peers' work. You should therefore expect a substantial portion of your grade to come from your written work. Writing Flag classes meet the Core Communications objectives of Critical Thinking, Communication, Teamwork, and Personal Responsibility, established by the Texas Higher Education Coordinating Board.

### 3.10. Writing Quality of Papers and Assignments

I expect that students will submit their papers and assignments using proper grammar and writing, etc. I will alert students early in the semester if I see that they are having trouble with their writing so that they may seek help from the appropriate source. Since part of research involves being able to communicate in a clear writing style, the quality of exposition will be one element that I will consider when examining students' submissions.

### 3.11. Absences

As described in the Course Requirements section of the syllabus (above), it will be very difficult to perform well in the course if you do not attend regularly. The only absences that I will consider legitimate include those pertaining to religious holidays, illness, extenuating circumstances due to an emergency, and university-excused absences. For illnesses, you will need to either provide me with a doctor's note, or you will need to send me an email before class to inform me that you are sick and won't be attending. If you are sick and do not provide me with a doctor's note or email me before class, your absence will not be excused except under very extenuating circumstances.

### 3.12. Late Work

Unless you receive prior approval from me, I will not accept late final Research Papers or Empirical Analysis/Results sections, and I will discount all other late assignments as follows:

- 1-15 minutes: 0% (grace period for last-minute issues)
- 15 minutes-24 hours late: -10%
- 24-48 hours late: -25%
- more than 2 days late: -50%
- more than one week: -75%
- more than two weeks: no credit offered

### 3.13. Homework Policies, Including for Stata/R Assignments

Students may consult with other members of the class and/or work in groups for the **Stata/R** assignments but not other homework assignments. Regardless of whether students choose to work in groups on the **Stata/R** assignments, students must submit their own copies of their work—i.e., no group submissions. Students are also not allowed to post their homework questions on the Stack Exchange, R help forums, and **Stata** help forums. Additionally, students may not seek help from people outside the class, such as from a friend,

professor, PhD student, etc. Students who received any sort of prohibited outside help will receive a zero for that particular assignment. These policies are in place because the only way to become proficient in these programs is to actually use them and make mistakes until you get it right.

If you need help with a particular question, feel free to [book an office hours slot](#). Provided that you attended the class where I covered the material at hand or missed class due to an excused absence (see above), I am very happy to help! I will not provide additional make-up training during office hours if you missed class for a non-excused absence.

### 3.14. Use of Computers During Class

To start the course, I will let everyone take notes in the manner that them suits them best. However, if students are consistently surfing the internet during class and/or not paying sufficient attention to class discussion, I reserve the right to alter this policy and ban the use of computers.

### 3.15. Students Rights and Responsibilities

- You have a right to a learning environment that supports mental and physical wellness.
- You have a right to respect.
- You have a right to be assessed and graded fairly.
- You have a right to freedom of opinion and expression.
- You have a right to privacy and confidentiality.
- You have a right to meaningful and equal participation, to self-organize groups to improve your learning environment.
- You have a right to learn in an environment that is welcoming to all people. No student shall be isolated, excluded or diminished in any way.

With these rights come these responsibilities:

- You are responsible for taking care of yourself, managing your time, and communicating with the instructor if things start to feel out of control or overwhelming.
- You are responsible for acting in a way that is worthy of respect and always respectful of others.

### 3.16. Personal Pronoun and Name Preferences

Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with differences of race, culture, religion, politics, sexual orientation,

gender, gender variance, and nationalities. Class rosters are provided to the instructor with the student's legal name. I will gladly honor your request to address you by an alternate name or gender pronoun. Please advise me of this preference early in the semester so that I may make appropriate changes to my records.

### **3.17. Academic Integrity**

Each student in the course is expected to abide by the University of Texas Honor Code: "As a student of The University of Texas at Austin, I shall abide by the core values of the University and uphold academic integrity." Plagiarism is taken very seriously at UT. Therefore, if you use words or ideas that are not your own (or that you have used in previous class), you must cite your sources. Otherwise you will be guilty of plagiarism and subject to academic disciplinary action, including failure of the course. You are responsible for understanding UT's Academic Honesty and the University Honor Code, which can be found at the following web address: [http://deanofstudents.utexas.edu/sjs/acint\\_student.php](http://deanofstudents.utexas.edu/sjs/acint_student.php)

### **3.18. Drop Policy**

If you want to drop a class after the 12th class day, you'll need to execute a Q drop before the Q-drop deadline, which typically occurs near the middle of the semester. Under Texas law, you are only allowed six Q drops while you are in college at any public Texas institution. For more information, see: <http://www.utexas.edu/ugs/csacc/academic/adddrop/qdrop>

### **3.19. University Resources for Students**

Your success in this class is important to me. We will all need accommodations because we all learn differently. If there are aspects of this course that prevent you from learning or exclude you, please let me know as soon as possible. Together we'll develop strategies to meet both your needs and the requirements of the course. There are also a range of resources on campus:

#### **3.19.1. Services for Students with Disabilities**

This class respects and welcomes students of all backgrounds, identities, and abilities. If there are circumstances that make our learning environment and activities difficult, if you have medical information that you need to share with me, or if you need specific arrangements in case the building needs to be evacuated, please let me know. I am committed to creating an effective learning environment for all students, but I can only do so if you discuss your needs with me as early as possible. I promise to maintain the confidentiality of these discussions.

If appropriate, also contact Services for Students with Disabilities, 512-471-6259 (voice) or 1-866-329- 3986 (video phone). <http://ddce.utexas.edu/disability/about/>

### 3.19.2. Counseling and Mental Health Center

Do your best to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress.

All of us benefit from support during times of struggle. You are not alone. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. Asking for support sooner rather than later is often helpful.

If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, we strongly encourage you to seek support: <http://www.cmhc.utexas.edu/individualcounseling.html>

### 3.19.3. The Sanger Learning Center

Did you know that more than one-third of UT undergraduate students use the Sanger Learning Center each year to improve their academic performance? All students are welcome to take advantage of Sanger Center's classes and workshops, private learning specialist appointments, peer academic coaching, and tutoring for more than 70 courses in 15 different subject areas. For more information, please visit <http://www.utexas.edu/ugs/slc> or call 512-471-3614 (JES A332).

Undergraduate Writing Center: <http://uwc.utexas.edu/>

Libraries: <http://www.lib.utexas.edu/>

ITS: <http://www.utexas.edu/its/>

Student Emergency Services: <http://deanofstudents.utexas.edu/emergency/>

### 3.19.4. Important Safety Information

If you have concerns about the safety or behavior of fellow students, TAs or Professors, call BCAL (the Behavior Concerns Advice Line): 512-232-5050. Your call can be anonymous. If something doesn't feel right – it probably isn't. Trust your instincts and share your concerns.

The following recommendations regarding emergency evacuation from the Office of Campus Safety and Security (512-471-5767, <http://www.utexas.edu/safety/>):

- Occupants of buildings on The University of Texas at Austin campus are required to evacuate buildings when a fire alarm is activated. Alarm activation or announcement requires exiting and assembling outside.

- Familiarize yourself with all exit doors of each classroom and building you may occupy. Remember that the nearest exit door may not be the one you used when entering the building.
- Students requiring assistance in evacuation shall inform their instructor in writing during the first week of class.
- In the event of an evacuation, follow the instruction of faculty or class instructors. Do not re-enter a building unless given instructions by the following: Austin Fire Department, The University of Texas at Austin Police Department, or Fire Prevention Services office.
- Link to information regarding emergency evacuation routes and emergency procedures can be found at: [www.utexas.edu/emergency](http://www.utexas.edu/emergency)

## 4. Class Schedule, Readings, and Homework

### Week 1: Stata Training (January 24)

#### Class:

- Stata training
  - Loading and saving data
  - Inspecting data (browsing, tabulating, etc.)
  - Creating new variables
  - Labeling variables
  - Correlation
  - Cross tabs
  - Difference in means
  - Merging data
  - Finding and removing duplicates
  - Getting World Bank World Development Indicators data directly from Stata

#### Required Assignments:

- Carefully read the new course syllabus for this semester
- Please watch this 3-minute [Tour of the Stata 14 Interface](#)
  - The video will be covered on the quiz

- Please ensure the [UT Austin Stats Apps Server](#) is setup for your computer so that you can use **Stata**
  - This will comprise part of your quiz grade
- Revise and Resubmit (R&R) Assignment: due Friday, January 24 at 12 noon
  - Congratulations! The editor at a prestigious journal is impressed by the paper you submitted for publication last semester, and so is the anonymous referee who read your paper at the editor's request. Consequently, the editor has offered you a chance to revise and resubmit (R&R) your paper for further review. Since R&R opportunities at prestigious journals are exceptionally difficult to obtain, you definitely want to take advantage of the opportunity and submit your best possible work. In doing so, please revise your paper based on the latest round of comments you received last semester from the instructor (i.e. the editor) and the referee report you received from your classmate (i.e. anonymous reviewer). In addition to addressing the feedback you received, please write a letter to the journal editor that explains exactly how you addressed his feedback as well as that of your anonymous reviewer. You can find example letters to the editor under "Files" on Canvas. When submitting your R&R assignment on Canvas by Friday, January 24 at 12 noon, please include as attachments on Canvas: (a) a revised version of your paper that incorporates the feedback you received from the editor and your anonymous reviewer; (b) a letter to the editor that explains how you addressed the specific points of feedback that you received from the editor and reviewer; and (c) the anonymous review (referee report) that you received last semester.

## Week 2: The Basics of R (January 31)

### Class:

- R Training I
  - Setting the working directory
  - Objects, vectors, entering in data manually, and creating data frames
  - Classes (numeric, character, factors)
  - Dealing with missing values
  - Installing packages and loading libraries
  - Descriptive statistics (mean, median, mode, quantiles)
  - Tables with **stargazer**
  - Getting help
  - Histograms, bar plots, scatter plots in **ggplot2**

- Correlations
- Exporting graphs
- Saving data

### Required Reading:

- Li, Quan. 2018. *Using R for Data Analysis in the Social Sciences: A Research Project-Oriented Approach*. Oxford: Oxford University Press. Chapter 1.
  - Read: pages 1-23, 32-42
  - Optional: pages 23-32
- Imai, Kosuke. 2017. *Quantitative Social Science: An Introduction*. Princeton, NJ: Princeton University Press. Chapter 1.
  - Skip or briefly skim: Sections 1.1-1.2
  - Read: Section 1.3

### Optional Videos:

- Watch [this Getting Started with R and R Studio video](#) (from last semester)
- Watch [this Introduction to R video](#).
- Watch the videos on [UT-Austin's Lynda Portal](#)

### Required Assignments

- Continue revising your paper after receiving my corrections.
- Please update your IPD time-tracking Google Sheet by Friday, January 31 at 12 noon
- **Stata** assignment: due Friday, January 31 at 12 noon. The assignment mirrors your R assignment from last semester:
  - Dichotomize the V-Dem political corruption score variable (`v2x_corr`) into new numeric variables called `corrupt` and `not_corrupt` based on the mean value of `v2x_corr`. (Note: “dichotomize” means separate into binary—i.e., 0 or 1)
  - Create a new string variable called `regime_corrupt`, in which one value is “corrupt regime” if `corrupt== 1`; and the other value is “not corrupt regime” if `not_corrupt== 1`
  - Pick one variable of your choosing from the World Bank’s World Development Indicators (WDI) that you think is correlated with corruption, and explain in words why you think that variable is correlated with corruption.
  - Merge the WDI variable into the V-Dem dataset
  - Subset the data to only keep the data from the year 2016. (hint: the command starts with “keep if”)



- Test if the WDI variable that you chose is correlated with the the V-Dem political corruption score variable (`v2x_corr`), by showing (a) a pairwise correlation table; and (b) a labeled scatter plot. Explain in words what your table and scatterplot suggest.
- Create a crosstab with your WDI variable and the `regime_corrupt` variable that you created above. Explain in words what your crosstab suggests in terms of how your variable varies in corrupt and non-corrupt regimes.
- Using your WDI variable, test if there is a statistically significant difference between corrupt regimes and non corrupt regimes (hint: use difference in means test)
- Please submit both your commented `.do` file and PDF showing your results. To be clear, I don't just want your code; I would like to see the graphs and tables that your code generates as well as some accompanying text. I provided you the code to produce L<sup>A</sup>T<sub>E</sub>X tables, and you also still have the [L<sup>A</sup>T<sub>E</sub>X Training Overleaf](#) from last semester. However, if you are short on time, you can also just paste in screen shots for this homework assignment—though please don't paste in screen shots in your final paper.

#### Extra Credit Assignment:

- [UT Undergraduate Research Fellowship](#) application: due on Canvas Friday, January 31 at 12 noon. Please contact me well before the deadline if you are planning on applying. I will provide feedback on any applications right away so that students can meet the university application deadline of February 3. A complete application entails a 4-page proposal with budget and recommendation letter (see above link). There are quite a few awards: last year, circa 50% of applicants received an award. These awards can be up to \$1000.

## **Week 3: Data Structures, Getting Data Ready, and Codebooks (February 7)**

### Class:

- R Training II
  - Importing Data (`.csv`, `.xls`, `.xlsx`, `.dta`)
  - Subsetting (i.e. creating new data frames)
  - Inspecting the data (`head`, `View`, `dim`, `summary`)
  - Creating new variables and indexing
  - Conditional statements (`ifelse`)
  - Importing World Bank World Development Indicators data directly from R

- Merging data
- Appending data
- Reshaping data
- Finding and removing duplicates
- Converting characters/string variables to numeric variables
- Recoding data
- Sorting data
- Creating lag and lead variables
- Taking logs
- Labeling variables

#### Required Reading:

- Imai, Kosuke. 2017. *Quantitative Social Science: An Introduction*. Princeton, NJ: Princeton University Press. Chapter 2.
  - Read: Sections 2.1-2.4
- Li, Quan. 2018. *Using R for Data Analysis in the Social Sciences: A Research Project-Oriented Approach*. Oxford: Oxford University Press. Chapter 2.
  - Read: pages 43-78, 91
  - Optional: pages 78-85
  - Skip: pages 86-90, 92-93
- Denly, Michael, Michael Findley, Joelean Hall, Andrew Stravers, and James Igoe Walsh. 2019. “Natural Resources and Civil Conflict: Evidence from a New, Geo-referenced Dataset.” Revise & Resubmit at *Journal of Conflict Resolution*.
  - Skim: The Codebook (see paper’s Appendix), devoting 5-10 minutes to your skim.
  - Optional: Rest of the paper.

#### Required Assignments:

- Please update your IPD time-tracking Google Sheet by Friday, February 7 at 12 noon
- Complete all chapters from [Data Camp’s Free Introduction to R course](#) and upload your personalized certificate of completion to Canvas by Friday, February 7 at 12 noon. You may, however, skip the chapter on Matrices. If you do, please upload a screen shot for each of the other chapters that you completed by the same deadline.

#### Further Reading (Not Required):

- Coppedge, Michael, John Gerring, *et al.* 2018. “Varieties of Democracy (V-Dem) Codebook, Version 8” Gothenburg, Sweden: University of Gothenburg, V-Dem Institute.

## Week 4: Hypothesis Testing (February 14)

### Class:

- Confidence intervals
- Null hypothesis significance testing and  $p$ -values
- R tidbits
  - Difference in means
  - Testing for equal variance
  - Power analysis (if time allows)
- Introduction to Bayesian statistics and hypothesis testing (if time allows)

### Required Reading:

- Li, Quan. 2018. *Using R for Data Analysis in the Social Sciences: A Research Project-Oriented Approach*. Oxford: Oxford University Press. Chapters 3 and 4.
  - Chapter 3
    - \* Read: pages 94-128, 138-142
    - \* Optional: pages 128-137
  - Chapter 4
    - \* Read: pages 144-160
    - \* Optional: pages 160-168
- Kellstedt, Paul, and Guy Whitten. 2018. *The Fundamental of Political Science Research*. Third Edition. Cambridge: Cambridge University Press.
  - Read: Chapter 8.3: The Logic of  $P$ -Values (pages 163-166).
    - \* Note: If you are using the 2013 edition, it's Chapter 7.3 (pages 147-150).

### Required Assignment:

- Please update your IPD time-tracking Google Sheet by Friday, February 14 at 12 noon
- From [Data Camp's Intermediate R Course](#), complete only the freely-available Chapter 1, Conditionals and Control Flow. Once you are done, take a screenshot to prove that you have completed this chapter 1, and upload your screenshot to Canvas by Friday, February 14 at 12 noon

- Please re-submit your full paper by Friday, February 14 at 12 noon, including the following changes: (a) all corrections implemented from the previous round of feedback; (b) very specific details in the Research Design section on the exact variables (dependent, independent, and control) that you will be using for your analysis; (c) an Appendix to the paper with the following, more detailed table filled out for each variable in your dataset (L<sup>A</sup>T<sub>E</sub>X code [here](#) if you need help):

Table 1: Description of Variables

Variable	Description	Source
Democracy	This variable corresponds to the polyarchy measure of the Varieties of Democracy (V-Dem) index. Per the codebook, the variable corresponds to an “average of, on the one hand, the weighted average of the indices measuring freedom of association thick, clean elections, freedom of expression, elected officials, and suffrage and, on the other, the five-way multiplicative interaction between those indices.”	Coppedge et al. (2018)
Infant Mortality	“Infant mortality rate is the number of infants dying before reaching one year of age, per 1,000 live births in a given year.”	World Bank (2019)

- Mixed Methods Studies with Qualitative Elements: For qualitative studies involving human subjects (i.e. interviews, focus groups, and/or an ethnography), please also submit an Institutional Review Board (IRB) application and a complete list of interview/focus group questions attached as the second Appendix to your paper; for case studies, please include a description and justification of the cases selected, specifying the exact sources (e.g. newspaper articles, books, articles).

#### Extra Credit Assignment:

- Submit a YouTube link for a video regarding your project to the [UT Austin Student Research Showdown](#) to me by Friday, February 14 at 12 noon. See the link for the info on the huge cash prizes available, etc.
- Note: if you submit a video, I will provide you with feedback immediately so that you can make appropriate changes before submitting your final version to the university by February 18 at 11:59pm.

#### Further Reading (Not Required):

- Imai, Kosuke. 2017. *Quantitative Social Science: An Introduction*. Princeton, NJ: Princeton University Press. Chapter 3.
- Gerring, John, and Dino Christenson. 2017. *Applied Social Science Methodology: An Introductory Guide*. Cambridge: Cambridge University Press. Chapter 20.

## Week 5: Experiments, Ethics, and Transparency (February 21)

### Class:

- Experiments
  - The logic of the randomized control trial (RCT)
  - Survey experiments
  - Lab experiments
  - Field experiments
  - Natural experiments
  - Noncompliance
  - Attrition
  - Spillover
  - Social desirability bias and Hawthorne effects
- Ethics
  - Why (Milgrom, Stanford Prison, etc.)
  - Belmont Report
  - Institutional Review Board (IRB)
- Transparency
  - *P*-hacking and false positives
  - The replication crisis
  - Pre-analysis plans/pre-registration

### Required Assignments:

- Continue revising your paper after receiving my corrections.
- Submit an [abstract for your Research Week poster](#) on Canvas by Thursday, February 21 at 12 noon and through this link to the university by February 27, 2020—after receiving my feedback. Your abstract should include:
  - 1-2 sentence(s) on why your topic is important, ensuring that you show without “telling” the reader, and definitely do not use the word “important”
  - 1-2 sentence(s) on your argument
  - 1-2 sentence(s) on how you test your argument, including justification of your case/data, if necessary
  - 1 sentence stating the expected policy or theoretical contribution of your paper.

- Please update your IPD time-tracking Google Sheet by Friday, February 21 at 12 noon.
- From [Data Camp's Cleaning Data in R Course](#), complete the "Introduction and Exploring Raw Data", "Tidying Data", and "Preparing Data for Analysis" chapters, and upload screen shots to prove completion on Canvas by Friday, February 21 at 12 noon.

#### Required Reading:

- Gerring, John, and Dino Christenson. 2017. *Applied Social Science Methodology: An Introductory Guide*. Cambridge: Cambridge University Press. Chapter 7.
  - Read the whole chapter and learn/know the key terms, listed at the end of the chapter.
- Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94(4): 991-1013
  - Read pages 991-994, paying particular attention to the excellent literature critique.
  - Quickly skim the rest of the article, devoting about 10 minutes to your skim
- Miguel, Edward *et al.* 2014. "Promoting Transparency in Social Science Research." *Science* 343(6166): 30-31.
  - Read the whole article (only 2 pages)
- The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research 1979. "[The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research](#)" US Department of Health, Education, and Welfare.
  - Quickly skim the report, devoting about 10 minutes to your skim.

#### Further Reading (Not Required):

- Dunning, Thad. 2016. "Transparency, Replication, and Cumulative Learning: What Experiments Alone Cannot Achieve." *Annual Review of Political Science* 19: 541-563
- Olken, Benjamin. 2015. "The Promises and Perils of Pre-Analysis Plans." *Journal of Economic Perspectives* 29(3): 61-80.
- Blair, Graeme, Jasper Cooper, Alexander Coppock, and Macartan Humphreys. 2019. "Declaring and Diagnosing Research Designs." *American Political Science Review* 113(3): 838-859.
- Kellstedt, Paul, and Guy Whitten. 2018. *The Fundamental of Political Science Research*. Third Edition. Cambridge: Cambridge University Press. Chapter 4.2: "Experimental Research Designs."
- Nosek, Brian *et al.* 2014. "Estimating the Reproducibility of Psychological Science." *Science* 349(6251): aac4716-1 - aac4716-8.

- Ioannidis, John. 2005. "Why Most Published Findings Are False." *PLoS Medicine* 2(8): 0696-0701.
- Gerber, Alan, and Neil Malhotra. 2008. "Do Statistical Reporting Standards Affect What Is Published? Publication Bias in Two Leading Social Science Journals." *Quarterly Journal of Political Science* 3(3): 313-326.
- Gerber, Alan, and Neil Malhotra. 2008. "Publication Bias in Empirical Sociological Research." *Sociological Methods and Research* 37(1): 3-30.

## Week 6: Measurement Challenges (February 28)

### Class:

- Geddes (2003) and King, Keohane, and Verba (1994) readings
- Challenges from qualitative researchers

### Required Reading:

- Geddes, Barbara. 2003. *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics*. Ann Arbor, MI: University of Michigan Press.
  - Read Chapter 3, pages 89-129. [See Canvas]
- King, Gary, Robert Keohane and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
  - Section 4.1., Indeterminate Research Designs
    - \* Read only pages 118-119.
  - Section 4.2., Limits of Random Selection
    - \* Read only the bottom of page 124.
  - Section 4.3., Selection Bias and Selecting on the Dependent Variable
    - \* Read only page 128 until "avoid them!" on page 130.
  - Section 4.3.1., Selecting on an Explanatory Variable
    - \* Read only the first paragraph of page 137.
  - Section 5.1.1., Systematic Measurement Error
    - \* Read only page 156.
  - Section 5.1.2.1., Nonsystematic Measurement Error in the Dependent Variable
    - \* Read only the first paragraph of pages 158-159.
  - Section 5.2., Excluding Relevant Variables: Bias
    - \* Read only pages 168-169.

- Section 5.3., Including Irrelevant Variables: Inefficiency
  - \* Read only the middle of page 182-183.
- Section 5.4., Endogeneity
  - \* Read only the bottom page of 185-186.

#### Required Assignment:

- Please start getting your data ready for analysis, so that your clean dataset can be ready by the due date of March 6 (next week). If you have questions, feel free to book office hours. I am happy to help. Do not wait till the last minute here. You won't be able to finish on time if you wait.
- Please update your IPD time-tracking Google Sheet by Friday, February 28 at 12 noon

#### Further Reading (Not Required):

- Collier, David, and James Mahoney. 1996 "Insights and Pitfalls: Selection Bias in Qualitative Research." *World Politics* 49(1): 56-91.
- Gerring, John. 2017. "Qualitative Methods." *Annual Review of Political Science* 20(1): 15-36.
- Mahoney, James "After KKV: The New Methodology of Qualitative Research." *World Politics* 62(1): 120-147.

## **Week 7: Linear Regression I (March 6)**

#### Class:

- R Training III
  - Bivariate linear regression with continuous variable
  - Bivariate linear regression with a dummy variable
  - Interpreting effects,  $R^2$ ,  $F$ -test
  - Substantive vs statistical significance
  - Multivariate linear regression
  - Producing beautiful regression tables with **stargazer**
  - Producing beautiful coefficient plots with **ggcoef**

#### Required Reading:

- Li, Quan. 2018. *Using R for Data Analysis in the Social Sciences: A Research Project-Oriented Approach*. Oxford: Oxford University Press. Chapter 5.
  - Read: 170-194



- Optional: 195-205
- Kellstedt, Paul, and Guy Whitten. 2018. *The Fundamental of Political Science Research*. Third Edition. Cambridge: Cambridge University Press.
  - Read: pages 221-225 (Section 10.4), 227 (Section 10.6)
    - \* Note: If you are using the 2013 edition, it's pages 202-206 (Section 9.4), 207-209 (Section 9.6)
  - Optional: Chapter 9, rest of Chapter 10
    - \* Note: If you are using the 2013 edition, it's the rest of Chapters 8 and 9

#### Required Assignment:

- Please update your IPD time-tracking Google Sheet by Friday, March 6 at 12 noon
- Submit your clean dataset for your paper by Friday, March 6 at 12 noon. Your submission should include two attachments: (1) your clean dataset (in one file) with labeled variables (exported as a Stata `.dta` file); and (2) an accompanying R script or Stata `.do` file to show your work and how you labeled your variables, etc.

#### Further Reading (Not Required):

- James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2014. *An Introduction to Statistical Learning: With Applications in R* New York: Springer. Chapter 3.
- Gerring, John, and Dino Christenson. 2017. *Applied Social Science Methodology: An Introductory Guide*. Cambridge: Cambridge University Press. Chapter 22.
- Imai, Kosuke. 2017. *Quantitative Social Science: An Introduction*. Princeton, NJ: Princeton University Press. Chapter 4.2-4.3.
- King, Gary. 1986. "How to Not Lie with Statistics: Avoiding Common Mistakes in Quantitative Political Science." *American Journal of Political Science* 30(3): 666-687
- Luskin, Robert. 1991. "Abusus Non Tollit Usus: Standardized Coefficients, Correlations, and  $R^2$ s." *American Journal of Political Science* 35(4): 1032-1046.

## **Week 8: Linear Regression II: Assumptions & Diagnostics (March 13)**

#### Class:

- R Training IV
  - Generating and plotting the residuals and fitted values
  - Testing for collinearity (VIF)

- Testing for heteroskedasticity
- Clustering standard errors
- Interaction terms
- Testing for outliers and removing them
- Normality tests (qqplot)

#### Required Reading:

- Li, Quan. 2018. *Using R for Data Analysis in the Social Sciences: A Research Project-Oriented Approach*. Oxford: Oxford University Press. Chapter 6.
  - Read: 206-250, 261
  - Optional: 251-260

#### Required Assignment:

- Please update your IPD time-tracking Google Sheet by Friday, March 13 at 12 noon.
- From Data Camp's [Correlation and Regression Course](#), complete only the chapters on Simple Linear Regression, Interpreting Regression Models, and Model Fit. Once you are done, please upload screen shots to Canvas to prove that you completed these chapter by Friday, March 13 at 12 noon.

#### Optional Reading:

- Kellstedt, Paul, and Guy Whitten. 2018. *The Fundamental of Political Science Research*. Third Edition. Cambridge: Cambridge University Press. Chapter 10
  - Note: if you are using the 2013 edition, it's Chapter 9
- Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis* 27: 163-192.

## **Week 9: No Class due to Spring Break (March 20)**

## **Week 10: Panel Data Linear Regression (March 27)**

#### Class:

- Least-Square Dummy Variable (LSDV) Model
- Fixed effects
- Testing for serial correlation
- Testing for unit roots/stationarity

- Testing for heteroskedasticity
- Robust standard errors
- Clustered standard errors
- Applications in R and Stata

#### Required Reading:

- Torres-Reyna, Oscar. 2010. “[Getting Started in Fixed/Random Effects Models Using R](#).” Manuscript. Princeton University.
  - Read: Slides 2, 8, 9, 11, 12, 14, 18, 22, 23
- Blattman, Chris. 2017. “Clusterjerk” [On Clustered Standard Errors]. Blog Post. Accessible at: <https://chrisblattman.com/2015/12/08/clusterjerk/>.

#### Required Assignment:

- Please update your IPD time-tracking Google Sheet by Friday, March 27 at 12 noon.
- If you have not already, start estimating your models and writing the empirical section of your paper. You will need at least some preliminary results for your poster sessions.

#### Reminder:

- If you have not already met with me twice, it would be advisable to [book another 15-minute office hours slot](#) before April 9. You do not want to lose 1 point toward your final grade for failing to fulfill this requirement.

#### Further Reading (Not Required)

- Torres-Reyna, Oscar. 2007. “[Panel Data Analysis: Fixed and Random Effects Using Stata](#).” Manuscript. Princeton University.
  - Read: Slides 2-3, 5, 9-12, 15-19, 23, 25-27, 29
  - For those doing a panel data model in their papers: 30-39
- Linden, Ariel. 2015. “Conducting Interrupted Time-Series Analysis for Single- and Multiple-Group Comparisons.” *Stata Journal* 15(2):480-500.

## **Week 11: Logistic Regression (April 3)**

#### Class:

- The basics of logistic regression
- Panel logistic regression with fixed effects
- Applications in R and Stata

#### Required Reading:

- Kellstedt, Paul, and Guy Whitten. 2018. *The Fundamental of Political Science Research*. Third Edition. Cambridge: Cambridge University Press.
  - Read: pages 273-281 (Section 12.1-12.2)
    - \* Note: If you are using the 2013 edition, it's pages 247-256 (Section 11.1-11.2)
- Li, Quan. 2018. *Using R for Data Analysis in the Social Sciences: A Research Project-Oriented Approach*. Oxford: Oxford University Press. Appendix.
  - Read: pages 313-322
- Rodriguez, German. 2019. “[Longitudinal Logits](#).”
  - Take a brief look, and keep this in mind for later if you are using logistic regression for your project.

#### Required Assignment:

- From [Data Camp's course on Multiple and Logistic Regression](#), complete the “Logistic Regression” chapter, and upload a screenshot that proves you completed the chapter by Friday, April 3 at 12 noon.
- Please ensure that you submit your completed poster on Canvas to me on Canvas by Friday, April 3 at 12 noon. Since you will need to submit the poster to the University printing shop by April 5 at 11:59pm, I will provide feedback right away. More [details here](#). You will need this poster for the Government Department Poster Session (April 14) and the UT Research Week Poster Session (April 16).

#### Optional Reading

- Karaca-Mindic, Pinar, Edward C. Norton, and Bryan Dowd. 2012. “Interaction Terms in Nonlinear Models.” *Health Services Research* 47(1): 255-274.

## **Week 12: Difference-in-Differences & Synthetic Control (April 10)**

#### Class

- The logic of dif-in-dif
- Card and Krueger (1994)
- Abadie, Diamond, and Hainmueller (2010)
- How to estimate dif-in-dif and synthetic control in R and Stata

#### Required Reading and Video:

- Imai, Kosuke. 2017. *Quantitative Social Science: An Introduction*. Princeton, NJ: Princeton University Press.
  - Read: Section 2.5 (pages 54-63)

- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. “Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program.” *Journal of the American Statistical Association* 105(490): 493-505.
  - Read: Introduction (pages 493-494); Section 3.2-Conclusion (pages 498-503)
  - Watch: [Video that shows you how to estimate the results from the article in Stata](#). The video also helps with the understanding of the concept of the synthetic control.
- Xu, Yiqing. 2017. “Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models.” *Political Analysis* 25: 57-76
  - Read: pages 68 (starting at Empirical Example)-69 (first two lines) [yes, one paragraph only]
  - Optional: How to run [gsynth in R](#)

#### Required Assignments:

- Please continue working on your Empirical Analysis section, which is due next week. Given that two poster sessions the following week, you really do not want to leave everything till the last minute.
- Please update your IPD time-tracking Google Sheet by Friday, April 10 at 12 noon.

#### Further Reading (Not Required)

- Card, David, and Alan Kruger. 1994. “Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania.” *American Economic Review* 90(5): 1397-1420.
- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2015. “Comparative Politics and the Synthetic Control Method.” *American Journal of Political Science* 59(2): 495-510.

## **Week 13: Regression Discontinuity Designs (RD/RDD) (April 17)**

#### Class:

- The logic of the sharp RD
- How to estimate RD models in R and Stata with `rdrobust`

#### Required Reading:

- Gerring, John, and Dino Christenson. 2017. *Applied Social Science Methodology: An Introductory Guide*. Cambridge: Cambridge University Press. Chapter 8.
  - Review (from last semester): Regression Discontinuity (RD) Designs (pages 132-133)

- Angrist, Joshua, and Pischke, Jörn-Steffen. 2014. *Mastering 'Metrics: The Path from Cause to Effect*. Princeton, New Jersey: Princeton University Press.
  - Read: pages 147-165.
- Imai, Kosuke. 2017. *Quantitative Social Science: An Introduction*. Princeton, NJ: Princeton University Press.
  - Read: Section 4.3.4: Regression Discontinuity Designs (pages 176-181)
- Eggers, Andrew, and Jens Hainmueller. 2009. “MPs for Sale? Returns to Office in Postwar British Politics.” *American Political Science Review* 103(4): 513-533.
  - Read: Introduction (pages 513-514)
  - Optional: Rest of the article.

#### Required Assignment:

- Please update your IPD time-tracking Google Sheet by Friday, April 17 at 12 noon.
- Please attend the UT Research Week Poster Session and the Government Department Poster Session. Each session will take about an hour or two of time. You do not need to stay forever.
- By Friday, April 17 at 12 noon, please submit the latest version of your entire paper, which will now include an Empirical Analysis/Results section as well as a discussion of those results. I will be looking for the following elements:
  - The codebook that you wrote for the respective earlier assignment, included in an Appendix or as a Table
  - Your R script or **Stata** .do file as well your labeled data attached as separate items
  - A description of the data and how it was measured
  - Some summary statistics (see the R and **Stata** trainings and the Quan Li book for how to make summary stats and put them in LaTeX)
  - Some description of the method you are using as well as why it is appropriate for you problem
  - Some tables and, preferably, some coefficient plots
  - An interpretation (in words) of your results
  - How certain are you about your findings? Recall how this was one of the major things Gerring and Christenson (2017, Chapter 1) and King, Koehane, and Verba (1994, Chapter 1) discussed. Hint: look at things such as standard errors, confidence intervals, p-values, R squared, whether your sample is representative of some broader population, etc.

- Can your findings be interpreted causally (internal validity)? (Hint: it has to do with the method that you use.)
- What is the external validity of your study? In other words, how well do your results generalize to other places/contexts, people, time periods? If you think the results generalize, say where/when and why. If the results do not generalize, why not?
- Consult Appendix [A](#) of this syllabus as well as the Greene guide, Gerring and Christenson (2017, Chapter 14), and the other guides on Canvas for further ideas.

#### Further Reading (Not Required):

- Cattaneo, Matías, Nicolás Idrobo, and Rocío Titiunik. 2019. [A Practical Introduction to Regression Discontinuity Designs](#) (Elements in Quantitative and Computational Methods for the Social Sciences). Volume I. Cambridge: Cambridge University Press.
- Querubín, Pablo, and James Snyder. 2013. “The Control of Politicians in Normal Times and Times of Crisis: Wealth Accumulation by U.S. Congressman, 1850-1880.” *Quarterly Journal of Political Science* 8(4): 409-450.

## **Week 14: Network Analysis and Test/Review (April 24)**

#### Class:

- The basics of network analysis
- Some applications of network analysis in R
- A review of what we covered this semester

#### Required Reading:

- Imai, Kosuke. 2017. *Quantitative Social Science: An Introduction*. Princeton, NJ: Princeton University Press. Chapter 5
  - Read: Section 5.2 (pages 205-220)

#### Required Assignments:

- Prepare for a 10 question, multiple choice test that covers some key points from this semester. It will be graded as a quiz but will cover more material.
- Please update your IPD time-tracking Google Sheet by Friday, April 24 at 12 noon.
- From [Data Camp's Network Analysis in R course](#), complete the “Introduction to Networks” chapter, take a screenshot to prove that you completed it, and load that screenshot on Canvas by Friday, April 24 at 12 noon.

## Week 15: Geomapping in R (May 1)

### Class:

- Geomapping in R with the `sf` package
- Loading shape files
- Using `mapview` to view shape files directly from the web
- Data structures (raster, vector)

### Required Assignment:

- From [Data Camp's Visualizing Geospatial Data in R course](#), complete the “Basic mapping with `ggplot2` and `ggmap`” and “Data import and projects” chapters, take relevant screen shots to prove you completed everything, and upload the screen shots to Canvas by Friday, May 1 at 12 noon.
- Please submit a respectful referee report of one of your colleagues' empirical papers by Friday, May 1 at 12 noon. Although I would *strongly* encourage you to look at the example referee reports on Canvas, your referee reports should have the following elements:
  - In the first paragraph, summarize the author's theory and how he/she tested his/her theory. (Referee reports *always* start with this one paragraph summary of the paper).
  - Discuss whether or not you found the theory and hypothesis compelling, and explain why or why not.
  - Discuss whether the research design and empirical section adequately map to the theory. In other words, talk about the operationalization. (Hint: Think about the diagram in Chapter 1 of Kellstedt and Whitten (2018) that links the theory to hypothesis.) Are the data put forth useful to test the hypothesis? Did the author choose the best possible design to test his/her hypothesis? Why or why not? If another design would have been better, explicitly say so. However, please do not just propose another design without a lot of thought. In all likelihood, the author thought carefully about the particular design that he/she chose.
  - Do the statistical models appear to be estimated correctly? If the author does not use an experiment or a design that attempts to mimic an experiment, did the author include relevant covariates? If not, which covariates are missing that would impact the outcome? (Hint: Refer to the slides from Week 6 on Measurement Challenges to ensure any new covariates that you think should be added to the model would actually impact the dependent variable.)
  - Would you recommend that the author perform any robustness tests? If so, which robustness tests would you suggest?



- See Appendix A of this syllabus for additional guiding questions that may be useful for assessing your colleague’s study.

## Week 16: Student Presentations (May 8)

### Class:

- Student presentations

### Required Assignment:

- Please prepare an 10-minute (timed) presentation in which you explain the results from your study. Kindly submit your presentations on Canvas by Friday, May 8 at 12 noon.

## Final Research Paper Due Date: May 17 at 12 noon

- Since the registrar has very tight deadlines for instructor grade submissions, I unfortunately will not be able to accept any late papers—beyond a 15-minute grace period to account for technical issues at submission, etc.

## Appendix A Questions to Consider for Papers

### A.1 All Studies

1. Does the study answer an important question to the world, and does the author justify its importance with a factual argument—as opposed to justifying the topic in overtly normative terms?
2. Does the study contribute to a scholarly literature, and does the author demonstrate sufficient knowledge of that literature to critique it and add to it?
3. Does the study abide by the rules of (descriptive or causal) inference—and contain public procedures, uncertainty estimates, a disinterested posture toward the truth, attention to possible error, and scope conditions?
4. Is the author clear and consistent about the type of relationships, theory, and objectives of the research?
5. Is the writing clear, does the author avoid the passive voice and colloquial language, and are there any grammar issues?

### A.2 Quantitative Studies

1. Is the dependent variable in its concept form clear to the reader?

2. Is the principal independent variable in its concept form clear to the reader?
3. Are there cleanly measured, credible data available that clearly map to the dependent variable and independent variable in their concept forms?
4. Is there a falsifiable theory that is formulated at a high level of abstraction?
5. Is there a clear hypothesis that is formulated at a lower level of abstraction than the theory?
6. Does the hypothesis clearly map onto the theory?
7. Is the argument coherent and credible?
8. Are the research design and data appropriate for the research question, theory, and hypothesis?

### **A.3 Mixed Methods Studies with Qualitative Elements**

### **A.4 All Mixed Method Studies**

:

1. Does the author use integration of the quantitative and qualitative methods appropriately, and avoid triangulating methods for purposes for which they are ill-suited?

#### Case Studies:

1. Is/are the type(s) of case studies appropriate given the research question?
2. Does the author provide a clear and compelling justification for the selection of the case(s) being examined?
3. Is the detective work for the case study complete enough to sufficiently answer the research question?

#### Conceptualization Pieces:

1. Does the author have a clear grasp of the semantic field—that is, does she/he identify like terms and appropriately distinguish them from the concept of interest?
2. Is the author clear about the domain of his/her concept, and appropriately classify it as experience-near (i.e. relative to a particular area) or experience-distant (i.e. universal, positivist)?
3. Does the author choose the appropriate conceptualization strategy—that is, does she/he pay appropriate attention to the abstraction of the concept, avoid conceptual stretching, specifically state the intension and extension, and justify why couching the concept in terms of cumulation, a radial concept, re-definition or family resemblance is appropriate?

Interviews:

1. Does the author select subjects who can speak to the author's research question, and is there a clear, compelling, and ethical selection criteria for subject inclusion/exclusion?
2. Does the author seek out subjects who can provide a diverse range of perspectives on the research question at hand?
3. Does the author provide details about how he/she found the subjects, whether they received compensation, and locations/conditions of the interviews?
4. Does the author ask subjects appropriate questions that map to the research question of interest?
5. Are the answers from the subjects compelling enough to answer the research question definitively?

Focus Groups:

1. All of the above questions for interviews are relevant for focus groups as well.
2. Does the author make a concerted effort to ensure that dominant people in the room do not monopolize the discussion, and that more shy people are able to contribute to the discussion?

Ethnographies:

1. Does the author provide a compelling justification for why the ethnography she/he undertakes is relevant to the research question of interest?
2. Does the author have the training and capacity (e.g. language abilities, skills) to be able to credibly undertake the ethnography?
3. Is the author's role in the ethnography distracting to the extent that it alters the behavior of the actors under study, and does the author provide a compelling justification for why his/her role is not distracting?
4. Does the author make specific reference to the ethics of the ethnography, and are there any ethical concerns regarding the author's presence?
5. Is the author's attention to method (e.g. keeping daily field notes, minimizing distraction) apparent to the reader?