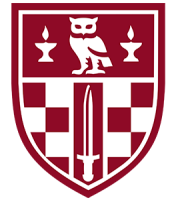
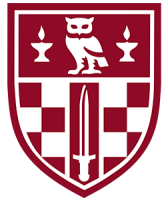
# Advanced Topics in Quantitative Social Research

Welcome and Introduction



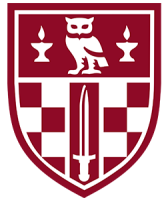
# Plan for today

- **Module overview**
  - Objectives and arrangements
  - Syllabus: Readings and tutorials
  - Assessment and presentations (Week 11)
- **Introduction**: Tell us about yourself
- **Exercise**: A quick recap on OLS (and will be continued)



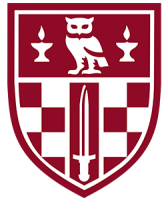
# Objectives and arrangements

- Advanced Quants covers a selection of topics on **model**- and **design**-based statistical inference
- Students are expected to be familiar with **multiple linear regression** and basic `R` **programming** before they join
- The weekly syllabus is **tentative** and can be adjusted depending on our progress
- The in-person tutorials will be used to
  - Review and answer your questions about the readings and recorded lectures (if any)
  - Give you the chance to implement the analysis in `R`
  - Help you prepare for the final project (Weeks 5 and 11)

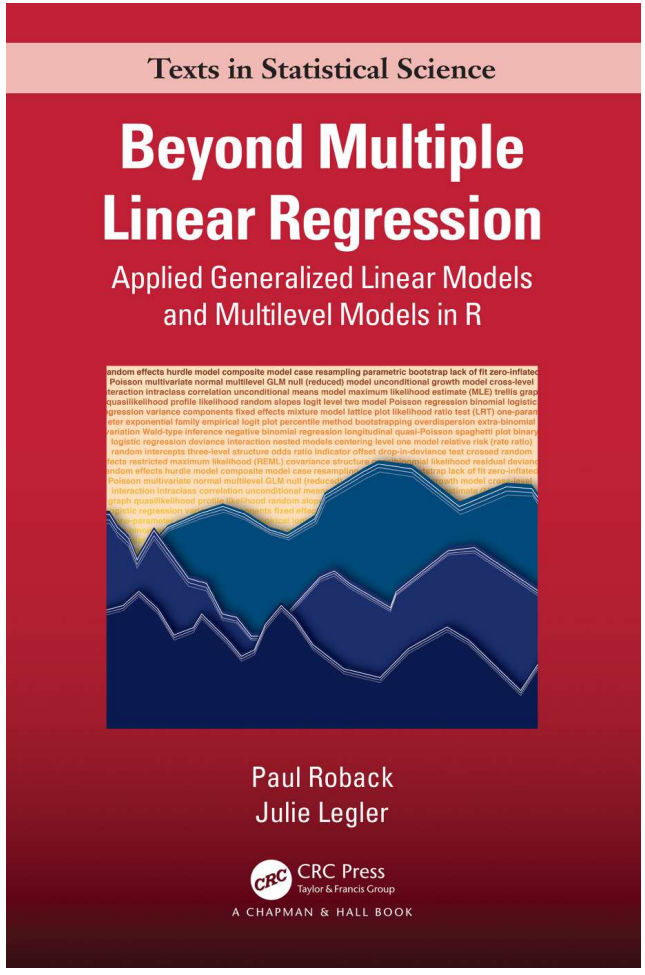


# Weekly syllabus

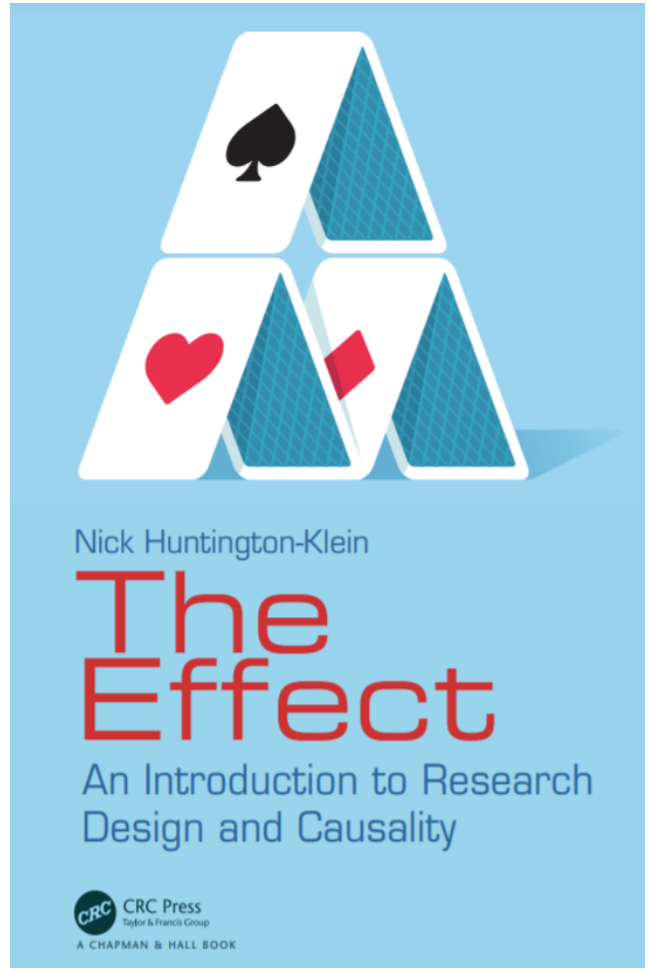
- Part 1: Generalized linear model
  - From OLS to GLM
  - Logit regression
  - Multilevel/hierarichical modeling (maybe)
- Part 2: Applied causal inference (maybe)
  - Selection on the observables (matching)
  - Regression discontinuity design
  - Instrumental variable
  - Difference-in-difference
  - Causal learning



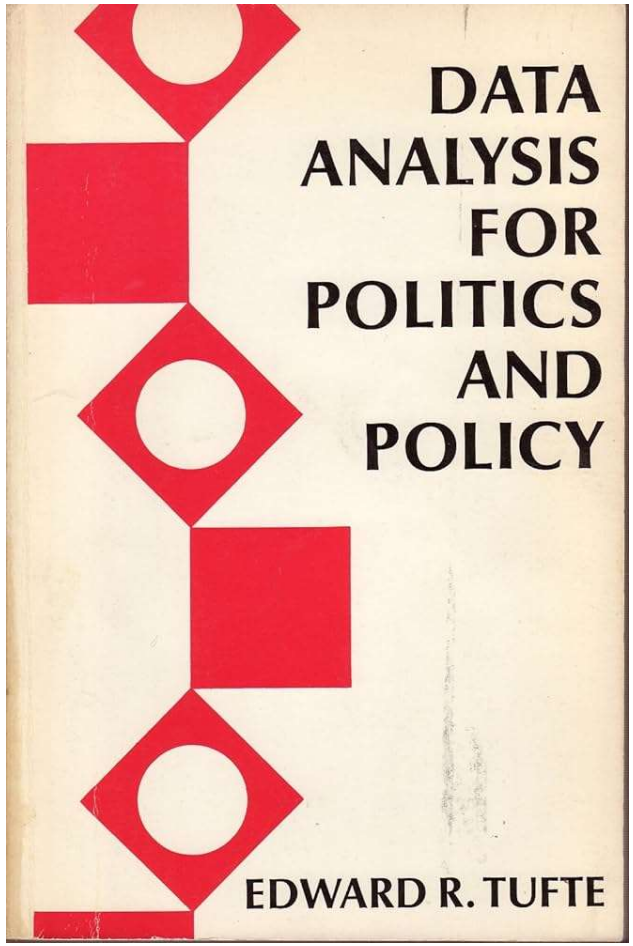
# Textbooks: Clear explanation without much math (under most circumstances)

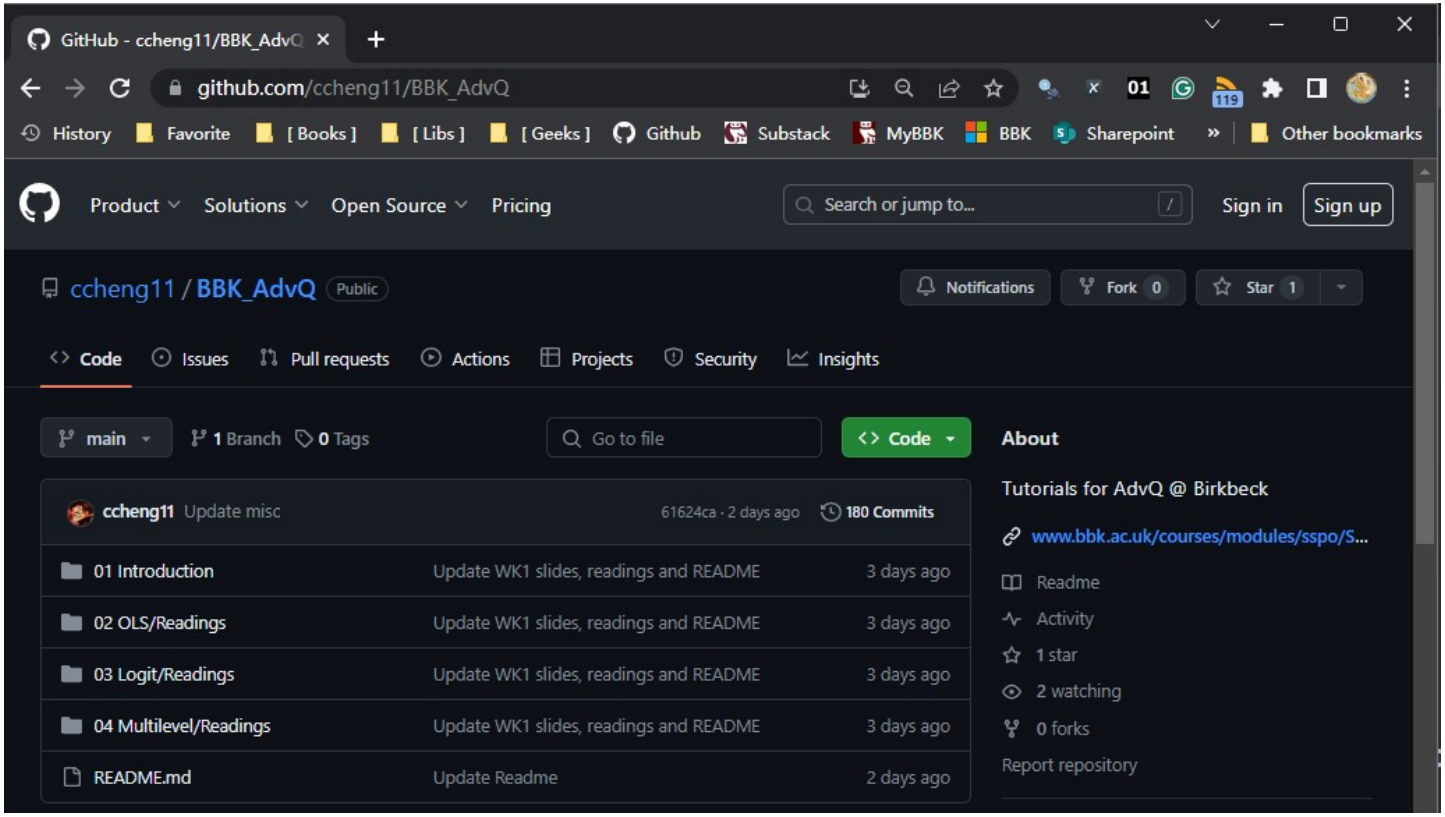



**For Weeks 2-4**




**For Weeks 7-10**

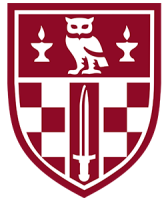

DATA
ANALYSIS
FOR
POLITICS
AND
POLICY
EDWARD R. TUFTE




https://github.com/ccheng11/BBK_AdvQ

# Additional resources on Moodle

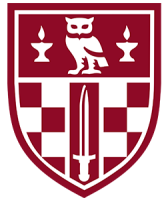
- You can find a variety of additional/supplementary resources – the key principle is to be **selectively thorough**
  - Supplementary textbooks: Use them when you find the main text is not clear; feel free to discuss with me if you are interested in purchasing any of them
  - Mathematics refresher: Review key math concepts and terms, especially **basic algebra**, if necessary; **Khan Academy** is a good start
  - R Programming: Help yourself when you have a hard time; pick one of the books and use the online community (e.g., **R-bloggers** and **Stack Overflow**) to keep yourself updated
  - Podcasts: Get a sense of how quant social researchers present and discuss their work



# Assessment

- **Exercises I and II** (20% each): Students will complete **two short data analytical exercises** based on the weekly tutorials

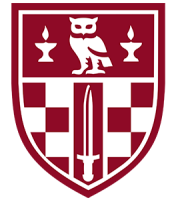
- **Final project** (60%): Students will complete **a quantitative research note** of about 4,000 words, excluding footnotes, bibliography, and appendices (R script and supplementary information); more will be discussed in **Week 5**



# Next week

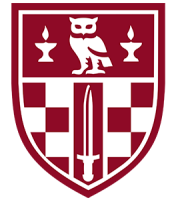
- We will review OLS and discuss why we need GLM (non-linear multiple regression)

- Before you come to class
  - Install *R* and *RStudio* on your laptop and bring it to class
  - Read the assigned chapters (either *Beyond Multiple Linear Regression* OR *The Effect*)
  - Watch the recorded lecture (uploaded on Thursday or Friday)

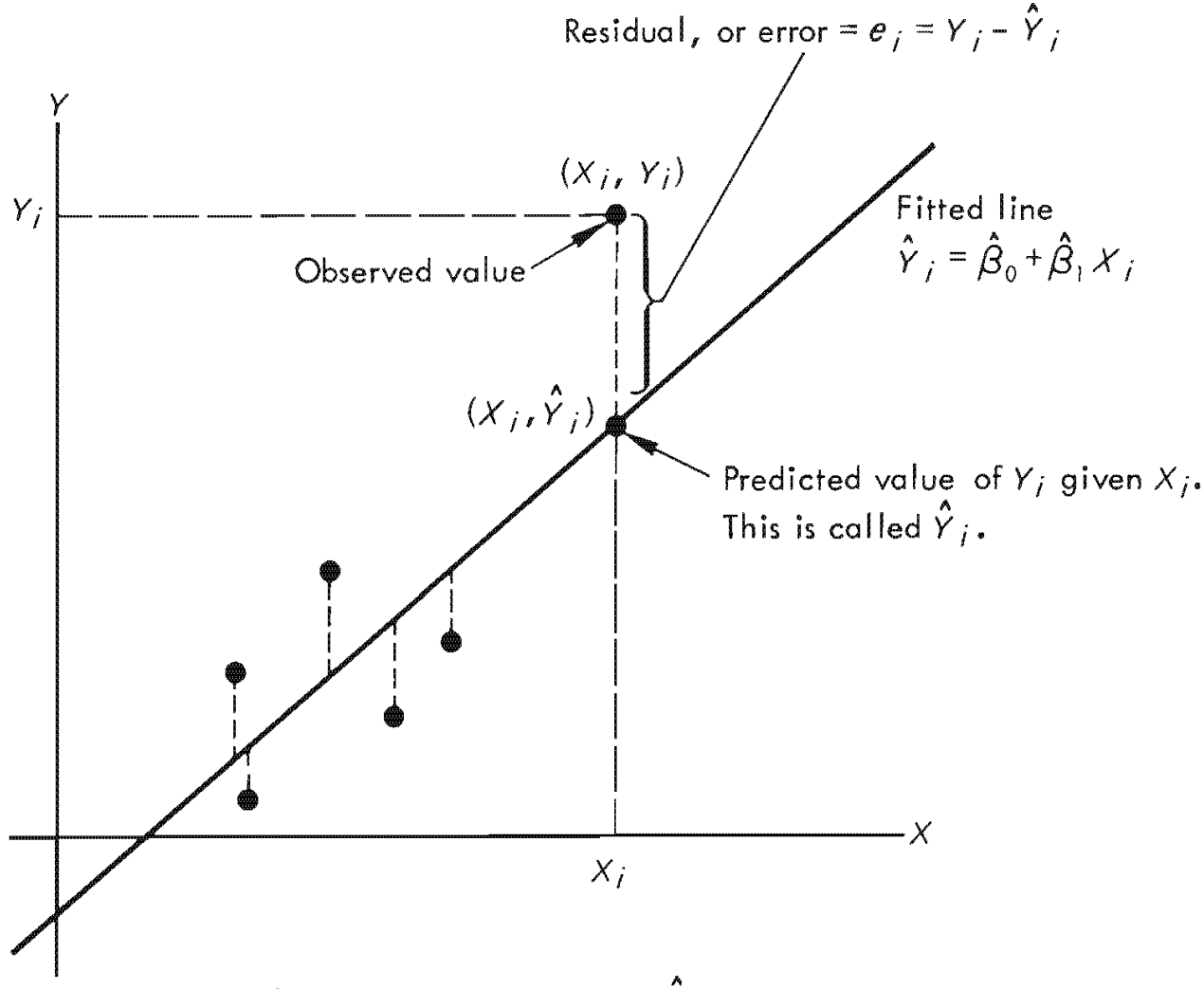


# Introduction

- Please share your name, course, and subject area (e.g., political science) with us
- Please share your experiences with quant social research. What are some of the most rewarding and/or challenging moments?
- Please share your potential research ideas for the final project. Any datasets you plan to study? We will continue the discussion on your research ideas in Week 5

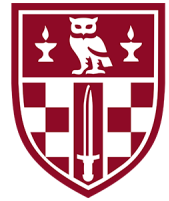


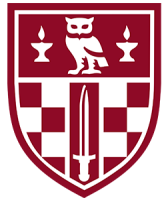
# Recap: OLS




Predicted value of $Y$ for a given $X_i = \hat{Y}_i$

$= \hat{\beta}_0 + \hat{\beta}_1 X_i$



# Recap: OLS

- Go through the **two-dimensional scatterplot**; locate and define/explain the following concepts:
  - Explanatory and dependent variables
  - Data point or individual observation
  - Fitted line (intercept and slope)
  - Residuals and goodness-of-fit
  - Model specification
- What is the idea behind the name of ordinary least squares (OLS)?