Harvard University Department of Economics  
Spring 2018  
Ec 970: Sophomore Tutorial

Instructor: David Kane  
dave.kane@gmail.com

Time: 4:00 PM -- 5:30 PM; Monday/Wednesday

Office Hours: By appointment, generally either before or after class

Location: CGIS Knafel Building - K108

Elite Education: Rhetoric and Empirics

Course Description

This course reviews the economic literature on elite colleges in the United States while studying the theory and practice of persuasion: using words, statistics and graphics to convince others, and yourself, of some claim. You will learn how to write persuasive essays, how to pick out the flaws in your opponent's argument, how to shift the terms of a debate to your advantage and how to marshal statistics and graphics to your side. We will study topics like: Which students choose to apply to colleges like Harvard? Which are admitted? What influence do athletics, legacy, race and wealth play in admissions? How do students perform while at college? How much of a problem is grade inflation and what might be done about it? Do elite colleges increase or decrease economic mobility? How generous are alumni after graduation? Yet beyond the empirical question concerning how the world works today, we are also interested in discussing the normative questions surrounding how elite colleges ought to function tomorrow. Natural Philosophers, the classical name for Economists, have wrestled with questions of Rhetoric since Plato confronted the Sophists two thousand years ago. Our class will continue that conversation in the context of contemporary debates about elite education in the United States.

Prerequisites  Courses in introductory statistics and in Microeconomics (Ec 1010a or Ec 1011a) are required. Econometrics, generally either Ec 1123 or Ec 1126, is required but may be taken concurrently.
Course Policies

*Stata:* Enrollment in Ec 970 requires completion of four statistical software problem sets to help develop your skill in handling and analyzing data. The Economics Department currently predominantly uses, encourages, and formally supports Stata. However, Ec 970 students are permitted to use the R programming language, with the understanding that there is no Departmental support for R. The Economics Department provides a Stata mini-course, which comprises one “live” session on Tuesday of 2/6, 8:30-10am (location TBD), and several online modules and resources. In addition, we offer several Stata office hours each week, all semester. Given the significant Stata resources we make available, we encourage you to do the problem sets in Stata, and to do all final paper data-related analysis using Stata. However, if you are extremely comfortable with R and would prefer to do your problem sets and final paper analysis using R--without any Departmental support--you may do so. More details on the Stata mini-course and Ec Department Stata resources available online: https://economics.harvard.edu/pages/stata.

*Correspondence:* I aim to respond to emails within 24 hours. Keep this in mind around assignment due dates and note that I may be slower on weekends. Please include Ec 970 in the subject line for all emails so they don’t get lost in my inbox.

*Laptops/tablets:* You may bring laptops/tablets to class. In some sessions, this is actually required as we work through some data analysis together. But, in most sessions, the discussion will be so active that you won’t even have time to take notes. You must have a way of referring to the readings during class, either via a hard copy or electronically. For example, we will often closely examine a specific table from a reading assignment, discussing it in detail. We will all need to look at the table during that discussion.

*Extensions:* Please discuss extensions with me in advance. Be aware that you cannot get an extension on the final paper. There will be no extensions available for the back-and-forth essays so plan ahead to ensure you can complete all of these. Extensions may be available on items like the rough draft of the research paper.

*Plagiarism:* If you plagiarize, you will fail the course. See the Harvard College Handbook for Students for details. Discussing ideas and work-in-progress with others is an important and desirable part of the research process, but in the end, a student’s assignment must be their own effort, written by the student, and ultimately based on their own thinking. All written assignments must use appropriate citation practices. For questions about Harvard’s stance on academic honesty, consult the Academic Information section in the Harvard College Handbook for Students.

*Working with Others:* Students are free (and encouraged) to discuss problem sets and their papers with one another. However, you must hand in your own (unique) code and written work in all cases.

*Expectations for Readings:* Most of the readings assigned in this course were published in economics journals. You are not expected to be able to reproduce the mathematical arguments on your own. When reading, focus on the empirical strategy and the intuition behind the economic arguments. Try to figure out where the paper fits into the literature.

*R:* You may use either Stata or R. Note, however, that R is a much more sophisticated tool for doing high quality data analysis. I will use R for all in-class examples. I recommend, but do not require, that you use R for your research paper. I suspect that R users will get much more from taking this tutorial than Stata users will. Useful resources for R include *R for Data Science, Data Visualization for Social Science* and *R bookdownplus Textbook*.

*Contacting Authors:* You must successfully contact at least one of the authors we read. For example,

---

\(^1\) I used very similar (sometimes identical) wording, in this and other sections, to the syllabi used in previous years in Ec 970.
Carrell and West (2010) report (p. 419) that more detailed results are available on request. One student will email them and ask for those results. Please coordinate among yourselves to ensure that only one student contacts a given author. Other options include: McPherson and Jewell (2007), pp. 872 and 877; Hoxby and Avery (2013), p. 29; McPherson et al (2009), p. 49.

Course Requirements

Class Participation (10%) Attendance is mandatory since class discussion is an essential component of the class. Students should read all the readings before class and be ready to actively participate in the class discussion. I will also ask you to email me, before the start of class, results of simple programming assignments, mostly as a commitment mechanism to ensure that you working on your technical skills. These are ungraded. Students are allowed two “free passes” for the semester, days on which I will not expect them to have done the reading or completed the programming assignment, assuming that they inform me ahead of class.

Stata Problem Sets (10%) Students should complete all Stata problem sets. You may use either Stata or R for the problem sets, but note that the department only provides help in Stata. You have one free day for turning in a problem set late. In other words, for one problem set, you may turn in your answers within 24 hours of the due date and not be penalized. Further lateness leads to a bottom rank for that assignment.

Essays (34%) There are four 1,000 word essays and two 500 word essays. The former each count for 7% and the latter for 3%. See the end of the syllabus for details on all assignments. Turning in an essay leads to a bottom rank for that assignment.

Empirical Replications (6%) There are two 500 word empirical replications, each worth 3% of the final grade. These each involve attempting to replicate one of the results from one of the assigned readings.

Research Paper (40%) There is a single 5,000 word research paper. This paper will be permanently archived by Harvard and publicly available. I have found that student work which is publicly displayed is of much higher quality than student work which is not. Although you may not turn in the final paper late, you have three free days, to be used as you like, on the interim due dates for the paper, items like the literature review or rough draft.

Grading

Writing assignments, other than the final paper, will be graded anonymously. Papers will be ranked from 1 to N, from best to worst. One problem with letter grades is that they do not allow me to clearly differentiate between the very best work and the merely good. What point is there in grades that only range from A to B+? Ranking also makes it clear to students who are struggling that, in fact, they are struggling. If you write the worst essay in the class, then I want you (and me) to be clear on that fact. Writing the worst essay does not make you (or me) a bad person, but it is important to see where you stand and to know what you should do to improve.

These rank grades are only for use in understanding where you stand relative to your peers. Your final grade will, of course, be on the traditional scale. I will calculate this final grade by first determining where you stand in the class and then by mapping this ranking to the distribution for Ec 970 as a whole. Class participation and problem sets are not ranked directly. (It is too hard, and probably counterproductive, to determine if a student is #3 or #4 in class participation.) Students receive either full credit, half credit or zero credit. Full credit corresponds to a rank of 1, half credit to a rank of N/2 and zero credit to a rank of N. The overall score is the weighted sum of the ranks across the 11 categories: class participation, Stata problem sets, 6 essays, 2 replications and 1 research paper.
I use the overall score to create a rank in the tutorial. One student will have done best. One will have done worst. (Ties are possible but unlikely, especially at the extremes of the distribution.) Again, these ranks are purely for your information. It is important to know where you stand, in life and in this tutorial.

What does a rank of 5 or 8 or whatever in the tutorial imply for the letter grade which actually appears on your transcript? I do not want to penalize students in the tutorial by grading them too low or reward them by grading too high. So, I will follow the “standard” practice of the college in general, and of the department in particular. In other words, if the lowest grade typically awarded in a tutorial is a B, then the bottom ranked student will receive a B. If the median student usually gets an A-, then the median-ranked student will receive an A-.

There is some chance that a group of especially bright and hard-working students will take this tutorial and that, therefore, the average grade should be higher than that in an “average” tutorial. Perhaps. Since I do not have enough recent teaching experience to make that judgment, I will rely on students’ performance on the final project as a guide. I will show these projects to some colleagues and ask if they think that the work is superior. If they think that it is, I will raise the tutorial average accordingly.
Schedule

Week 1: Jan 29

Session 1: Introduction

Review the syllabus.

No required readings. Please bring your laptop to class. Please install R and RStudio before class. *R for Data Science* is the best single introduction to both. This page provides many useful resources. If you are having trouble setting things up, try this free basic orientation.

Session 2: Overview


Read the preface and chapter 2 from *Data Visualization*. (Feel free to skim chapter 1.) Submit your version of the figure from section 2.6.

Email the NLSF to ask for permission to use the data.


*Gorgias* by Plato.

Week 2: Feb 5

Essay 1 due Sunday February 4 at midnight.

Session 3: View From High School


Session 4: Early Admissions


Read chapter 3 from *Data Visualization*. Submit your own figure using the gapminder data, as discussed in section 3.8.

*Starred (*) readings are optional.
Essay 2 due Thursday February 8 at midnight.

**Week 3: Feb 12**

Essay 3 due Sunday February 11 at midnight.

**Session 5: The Bell Curve**

[Selections from The Bell Curve: Intelligence and Class Structure in American Life by Richard J. Herrnstein and Charles Murray, 1994.](#)


**Session 6: Low Income**


Read chapter 4 from [Data Visualization](#). Submit your own figure using the gapminder data, as discussed in section 4.8.

---

**Week 4: Feb 19**

Essay 4 due Sunday February 18 at midnight.

No class on Monday February 19 because of Presidents’ Day.

**Session 7: Affirmative Action I**


Read chapter 5 from [Data Visualization](#). Submit both your .Rmd file and the knitr pdf, doing something interesting with the elections_historic data, as discussed in section 5.7.

Essay 5 due Thursday February 22 at midnight.
**Week 5: Feb 26**

Essay 6 due Sunday February 25 at midnight.

**Session 8: Affirmative Action II**


**Session 9: Data Analysis**


Gelman, Andrew, and Eric Loken. “The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time,” 2013, working paper, Columbia University. [Link](#)


**Week 6: Mar 5**

Empirical Replication 1 due Sunday March 4 at midnight.

**Session 10: Admissions Preferences I**


**Session 11: Admissions Preferences II**


Empirical Replication 2 due Thursday March 8 at midnight.

**Week 7: Mar 19**
Session 12: Matriculation


Session 13: Peer Effects


Week 8: March 26

March 25: Rough summary (2,000 words) of the argument in your chosen book.

Session 14: Grade Inflation I


Session 15: Grade Inflation II


Week 9: April 2

April 1: Replication of at least two separate tables from your chosen book using NLSF data.

Session 16: Grade Inflation III


Session 17

Read chapters 6, 7 and 8 in Data Visualization.

*Skim Fundamentals of Data Visualization.

Please bring your laptops to class. We will, as a group, work on creating graphics using NLSF data.

Week 10: April 9: Professor Quality and Evaluation

Beautiful graphic using NLSF data for use in research paper due Sunday April 8 at midnight.

Session 18


Session 19


Week 11: April 16

Rough draft of research papers due Tuesday April 17 at midnight.

Session 20: Alumni Giving


Session 21: Research Paper Draft Review

In this class session, we will review the drafts of your research papers. I provide the below citations for your reference. You do not need to read the book you are not reviewing. If you are interested in doing a replication exercise featuring a different paper or book, please let me know. *If you can get access to the data*, we can work something out.


Week 12: Apr 23

Session 22


Session 23

Writing Assignments

Essays

Essays will be graded anonymously. Much of the complexity here is due to the desire for “blind” grading. Fairness has its costs. Students will use “call signs.” These should be single, not-too-stupid-sounding, names which serve to preserve anonymity. “Rabbit” or “Crimson” are reasonable call signs. “IamTheVeryModelofaModernHarvardStudent” is not. There are 6 essays, organized into two groups. Each group is centered around two “instigation” articles from the academic literature.

Group One: Starting January 30


Group Two: Starting February 18


Within each group, students act as the “critic” of one paper and the “defender” of the other. There will be 10 “pairings:” A through J. Each student is part of two pairings because each student acts as both a critic and a defender. The schedule of roles will be assigned at the start of the semester. Students should keep this and similar information to themselves. This schedule will specify both what role you will fulfill each week and who you will be assigned to respond to on those weeks during which you are a “defender.” Your “opponent” in this dialog will be another member of the class whose identity is not known to you (or me).

In writing your essays, keep in mind the following:

- Essay 1 is a 1,000 word essay which attempts to refute the instigation article. The goal is to demonstrate that the original author is wrong, confused, mistaken, dishonest, ignorant or otherwise not worthy of consideration. Five students will be the critic for one of the instigation articles (pairs A through E) and five students will be the critic of the other (pairs F through J).

- Because this is a course in the rhetoric of economics, the critic should focus on the numbers. Throughout, you may assume that a reader of your essay has read the original article. You may use any other sources in your essay. Although you do not have time for a full scale research project each week, a well-chosen fact or number can be very persuasive. Do not plagiarize. Cite your sources appropriately.

- Essays should not have title pages. Instead, the top of the first page will consist of four items: Your title (preceded by the essay number and pair ID), your call sign, the call sign of the person you are paired with and a word count. (In the first essay, you won’t know the call sign of the person you are paired with, so you should just use the names of the authors of the article in its place.) Example:
Essay 2 and Pair C: “Mistaking Context for Contact”
Call sign: Crimson
Responding to: Rabbit
Word count: 987

● The word counts are only guidelines but you should try to follow them. A failure to fully support the arguments or to adequately edit and focus your prose will be counted against your ranking. I will reward neither excessive succinctness nor verbosity.

● Note that the purpose of the essay is not to tell us what you think. You get to do that during class discussion. The purpose of the essay is to convince us that the author of the original article is wrong. One of the overriding goals of the tutorial is to discover and review all the techniques for accomplishing this task.

● Essay 2 will involve each student as a “defender,” writing a 1,000 word rebuttal to the critic. The defender will act as if she is the author of the original piece. She will write sentences like, “Rabbit misunderstands my point when I claim ...” The purpose of the defense is to demonstrate that the thesis of the original article stands unfuted. The critic should be ignored. His argument has no merit. You refer to the critic using her call sign. Again, each student is, simultaneously, engaged in two debates, one as a critic of one of the articles and the other as the defender of the other article. Each of these debates is a pairing, so a student might be the critic in Pair B and the defender in Pair G.

● Essays 3 and 4 will involve another back-and-forth round, this time using just 500 words. Remember, these essay are not what you, personally, think. You are an advocate, making the best possible case. Critics write essay 3, responding to the arguments that the defenders have made in essay 2. In essay 1, critics referred to the authors of the instigation article by name. They do the same in essay 3. Defenders, again writing as if they are the original authors, get the final say in essay 4, referring to the authors of essays 1 and 3 using the call sign of those authors.

● Essay 5 and 6 will restart the cycle with new instigation essays. In these cases, there are only two rounds, a 1,000 word initial criticism and a 1,000 word response, just as with Essays 1 and 2.

● All essays will be submitted by e-mail in PDF format to me using an anonymous email service like Guerilla Mail.

● Make sure that your 1,000 word essays fit on no more than 4 pages and your 500 word essays on two pages during the first cycle This makes printing easier. This requirement does not apply if you are using R Markdown (or a similar set up) since it can be difficult to control page numbers.

● Although overrated, The Elements of Style, is worth reading at least once. More ambitious students should consult Artful Sentences: Syntax as Style by Virginia Tufte.

The goal of these essays is to persuade the reader. You should make the best case you possibly can even if you disagree with the arguments that you are making.

Keep in mind that you should make use of the published literature in economics. Students who just offer their personal opinions will not be graded highly. (The use of the academic literature in related fields is also welcome, but never forget that you are in an economics tutorial.) Although the syllabus provides a list of relevant articles from the literature, you do not need to use those articles, nor should you feel restricted
to only using them. Get good at finding relevant articles on your own. At the same time, you do not have the space (or time) for a full scale literature review.

**Example Assignments**

<table>
<thead>
<tr>
<th>Person</th>
<th>Pair 1</th>
<th>Pair 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>F</td>
</tr>
<tr>
<td>2</td>
<td>A</td>
<td>G</td>
</tr>
<tr>
<td>3</td>
<td>B</td>
<td>H</td>
</tr>
<tr>
<td>4</td>
<td>B</td>
<td>I</td>
</tr>
<tr>
<td>5</td>
<td>C</td>
<td>J</td>
</tr>
<tr>
<td>6</td>
<td>C</td>
<td>F</td>
</tr>
<tr>
<td>7</td>
<td>D</td>
<td>G</td>
</tr>
<tr>
<td>8</td>
<td>D</td>
<td>H</td>
</tr>
<tr>
<td>9</td>
<td>E</td>
<td>I</td>
</tr>
<tr>
<td>10</td>
<td>E</td>
<td>J</td>
</tr>
</tbody>
</table>

Persons 1,3,5,7,9 are critics of Espenshade and defenders of Dale. Persons 2,4,6,8,10 are critics of Dale and Defenders of Espenshade. Let’s walk through an example. Person 8 writes her first essay as a critic of Dale in Pair H. Person 3 responds to her as Dale. Person 8 then writes (Essay 2) as a defender of Espenshare as part of Pair D. She is responding to Person 7.

We will draw numbers at random for the initial assignments. But, once you are Person 7, you no longer care about that number. You only need to remember which Pairs you are in. And you need to keep track of your two call signs, one for each pair.

**Empirical Replications**

The 7th and 8th assignments are brief replications of a result from one of the assigned papers. The easiest to do will involve the NLSF data which are the most important source for the second set of instigation articles. But you may use a different paper from the syllabus, if you like, or even an unassigned paper, but please check with me first. This is inspired by Harvard Professor Gary King’s guide on how to replicate published research as a class assignment. These are meant to be simpler versions of the work that you will complete for your research paper. Much of the below advice for the research paper applies for these replications as well. These will be graded anonymously. Notes:

- Use R markdown to write the paper, and then knit it into a PDF for submission to me.
- These are not full scale papers! You should not try to replicate all the results from the papers, much less try to extend those results. Instead, pick a single table, or even a portion of a table.
- Your replication will describe what how the authors (claim to have) calculated their result, what their result was (perhaps using a screenshot) and your attempt to replicate it.
- This is not as essay with a thesis statement, supporting arguments, and so on.
- Your own replication should be replicable. If you give your .Rmd file to a member of the class, they should be able to, on their own computer, knit it into the same PDF you submitted to me.
- If you can’t replicate the result, that is OK! Perhaps they made a mistake.
Research Paper

Students will write one research paper of approximately 5,000 words. The format of the paper is that of a review article about one of these two books:


The style you should aim for is that of a review essay from the *Journal of Economic Literature*. Good examples of such reviews include:


Think about this report as an exercise in replication. In writing this paper, please keep in mind all the lessons you have learned from your essays. In addition:

- Include an abstract. These are the most important 150 or so words of your paper. Make them count.

- Include an introduction. Each sentence in the abstract is expanded to a paragraph in the introduction. Remember that a research paper is not a mystery novel. Reveal the most important things that you have discovered in the introduction.

- Include a conclusion. The conclusion restates every thing in the introduction. Many readers (sadly!) only read either the introduction or the conclusion, so anything that matters needs to be included in both sections. At the end of the conclusion, you are free to suggest productive avenues for future research and/or speculate about the implications of your findings. These sorts of observations belong in neither the abstract nor the introduction.

- You will produce portions of this paper over the second half of the course. See the schedule for details. The goal is to avoid a mad rush to complete it on April 30.

- Use the data associated with the *National Longitudinal Survey of Freshmen*. Please request permission to use this data using the contact information at their homepage.

- You must produce at least one beautiful graphic as part of your review. Read *Data Visualization for Social Science* (free online) for excellent advice. Note that all figures and tables need thorough captions, ones which explain all the associated details involved in their creation. Much
of that material will also be included in the text of the paper itself. It should be possible for a reader to look at just the figure/tables/captions or at just the text of the paper and still understand your point.

- You must also do some original quantitative research. The best way to begin this is by replicating some of the results from the book. Once you are able to replicate some of the work, you can extend/challenge the results by modifying the statistical models which the authors use.

- Because you are creating graphics and doing some original research, your work is actually more ambitious than the typical review in the *JEL*.

### Graded Assignment Schedule

All assignments are due at midnight on the specified day, unless otherwise specified.

- February 4: Essay 1
- February 8: Essay 2
- February 11: Essay 3
- February 16: Stata Problem Set #1
- February 18: Essay 4
- February 22: Essay 5
- February 25: Essay 6
- March 2: Stata Problem Set #2
- March 4: Empirical Replication 1
- March 8: Empirical Replication 2
- March 23: Stata Problem Set #3
- March 25: Rough summary (2,000 words) of the argument in your chosen book
- April 1: Replication of at least one table from your chosen book using NLSF data
- April 6: Stata Problem Set #4
- April 8: Beautiful graphic using NLSF data
- April 17: Rough draft of paper
- April 30: Final research paper due by 5:00 PM. No extensions.

### Conclusion

If you had tried to write a similar book review before taking this tutorial, you would have done X well. Now that you have taken the tutorial -- now that you have learned how to analyze an argument, evaluate statistical data, create beautiful graphics and write persuasively -- you will do Y well with your review. The success (or failure) of the tutorial can be measured by comparing Y with X.