

17.871
Spring 2013
Group Projects

Assignment summary

Working with your assigned group, answer the question posed to you *with regression analyses*. You will hand in draft slides of your presentation (see the next page) at the *end of the day* on Monday, April 1 and will give a 15-minute presentation on your work on Wednesday, April 3. Your group will also turn in a 10-page, double-spaced written report on your project on Friday, April 5. (Please e-mail a copy to me and Nick by 5:00 pm. in one of the following formats: doc or pdf. Please also attach your Stata do file and data file.) The 10-page limit includes tables, figures, and bibliography. The report should be in the form of a (mini) term paper meaning, among other things, that it should follow the format described by Kate Turabian, *A Manual for Writers of Term Papers, Theses, and Dissertations*. The paper should cover the same material as the presentation, which is outlined below ("The five slides").

Data requirements

The data set you create must have least 30 cases and at least *three* explanatory variables.

Statement about Collaboration

You are encouraged to seek and extend as much help as you can, both within and between groups. I expect you to be meticulous in citing the written work of others that you use.

Grading

We will assign a letter grade to each group's project. That will be the grade you receive, plus or minus an adjustment that will be determined as follows: We will ask each member of the group to indicate the relative amount of effort each person contributed to the successful completion of the project. If someone in the group stands out as being a conspicuous over-contributor or under-contributor to the group effort, that person's letter grade will be adjusted upward or downward as appropriate.

Grading rubric

Writing/Organization - 30 points
Literature Review - 15 points
Data Collection - 15 points
Data Analysis - 20 points
Data Presentation - 20 points

Presentations: The five slides

Your presentations should consist of the *five* slides.

1. Research question
 - Pose your question
 - What is the current state of academic thinking about your question?
 - Develop a hypothesis about one key explanatory variable
 - Explain why it's important to democracy, public policy, etc.
 2. Research design (methodology: usually regression)
 - Describe dependent variable, key explanatory variable, control variables (justify the controls)
 - Descriptive statistics (recode all to 0-1, present means and sds)
 3. Bivariate relationship for key variable
 - Scatter plot for continuous variables, box plots or crosstabs for nominal variables, label axes, label data points
 - Present and *interpret* bivariate coeff. (include a sentence on the slide), put regression line on scatter plot
 4. Present and *interpret* multivariate regression coefficients
 - Raw Stata output is okay, but only this time!
 - Does adding the control variables to your regression change the results? Why? Interpret the coefficient with a sentence on the slide.
 - Interpret SER (include a sentence on the slide!)
 5. Conclusion
 - What's the most important inference?
 - What are the policy implications or implications for democracy?
 - How strong of an inference can we draw? Problems? Alternative explanations?
 - Internal validity? (Nonrandom selection and reverse causation)
 - External validity?
-
- Please practice your presentations and ensure that you are under 15 minutes!
 - Please bring copies of your slides for everyone in class.

Some final notes

For variables that have been rescaled to vary between 0 and 1, interpret the coefficient as a shift from the minimum to the maximum leads to a (?) change in (dependent variable). In the past, I noticed a few groups instead using the phrase "one unit change" with rescaled variables. Although this is technically correct, the minimum to the maximum language is more informative for readers.

1. When interpreting your results, be very careful about whether your variables are percentages or not. Sometimes groups will interpret their coefficients and their SER in terms of percent change even though the variables were not percentages. The interpretation of the coefficients and the SER are always in the unit of the variables (with the exception of log variables). So, if your

variables aren't percentages (or proportions), don't use percentage language in your interpretations.

2. Don't forget to include a brief literature review. The main point of the literature review is to discuss the most important works relevant to your research question.
3. For this paper, you can paste Stata output straight into the paper. However, make sure you convert the typeface to Courier (or another fixed-width font), so that it will be legible. Please paste in the output of the summary command for the variables used in your regressions (so we can see your Ns, means, standard deviations, and ranges).

Finishing group projects can be difficult interpersonally. A surprisingly useful trick is to be kind in all your interactions with group members.

Project 1: Women in State Legislatures

Names: Cai, Charon, Han

Question: What explains variation in the proportion of U.S. state legislators who are women?

Possible explanations

- *Demographics.* States with more women elect more state legislators who are women. (Believe it or not, there's variance here. 50.9% of the US is women, but the standard deviation across states is 0.81%.)
- *Ideology.* More liberal states are more likely to elect more women.
- *Supply.* Because lawyers tend to be over-represented among state legislators, states with a higher proportion of women as lawyers should elect more women.
- *Size of constituencies.* Smaller legislative districts may be more conducive to women candidates, under the theory that women are disadvantaged in things like money, and money is more important as constituencies get larger.

Data sources

- *Book of the States.* In Dewey and Rotch libraries. Gives basic information about the structure of state governments.
- *Current Population Survey,* available online and in Dewey and Rotch libraries. Special reports give professional breakdowns by demographic characteristics.
- *Center for American Women and Politics at Rutgers University.* Lots of reports about women elected officials.

Bibliographic example

- Kira Sanbonmatsu. 2002. "Political parties and the recruitment of women to state legislatures." *Journal of Politics* 64(3): 791–809.

Comments/hints: A major issue is the unit of analysis. Will you study all the states in a particular year, which is called a cross-sectional study? In this case, the unit of analysis is the state. Will you study one particular state over a number of years, which is called a longitudinal study? Here, the unit of analysis is the year. Or, will you combine the two, in a panel study, where the state-year is an observation?

Project 2: Democracy and Foreign Direct Investment

Names: Kapoor, McCrary, Melchor, Powell

Question: Do democracies tend to attract greater foreign direct investment than countries with different types of electoral systems?

Possible explanations

- *Democracy.* Democracies will tend to attract more foreign direct investment because they are usually organized with a set of robust property rights protections.
- *Natural resources.* Nations with lots of extractable national resources will attract greater FDI in order to mine desired the desired good and trade with it.
- *Domestic market economic growth.* Countries with higher degrees of economic growth will attract higher levels of foreign direct investment, because foreign companies will regard investments in those countries as being good bets.

Data sources

- *World Bank Development Indicators.* Consists of a large number of measures of foreign direct investment, along with measures like travel levels, market size, and GDP.
- *Polity IV Project.* This is a long-term database project that measures characteristics of countries.

Bibliographic example

- Jensen, Nathan M. 2003. "Democratic Governance and Multinational Corporations: Political Regimes and Inflows of Foreign Direct Investment." *International Organization* 57 (3): 587-616.

Comments/hints: A major issue is the unit of analysis. Will you study all nations in a particular year, which is called a cross-sectional study? In this case, the unit of analysis is the country. Will you study one particular country over a number of years, which is called a longitudinal study?

Project 3: Support for Medical Use of Marijuana in Massachusetts

Names: Mahajan, Rodarte, Wynter

Question: What explains support for the 2012 proposition in Massachusetts allowing for the use of “medical marijuana” in the State?

Possible explanations

- *Baby boomers.* The baby boomers, who now range in age from 54 to 74, are the age group most likely to vote. The referendum passed because this age group has risen in number, compared to other age groups in the electorate.
- *College students.* Young people, regardless of generation, are more likely to smoke pot than older people. Massachusetts has the largest fraction of its population in college of any state in the country. It was the large number of college students that accounted for support for the referendum.
- *Libertarianism.* Despite its reputation for being a solidly “blue state,” Massachusetts has a strong libertarian streak on social issues that cuts across parties. It was this sense of social liberalism that allowed the referendum to pass.

Data sources

- *State Elections Division.* This is the office in the Secretary of State’s office that is in charge of elections in Massachusetts. While the state has yet to publish town-level election returns on its web site, you *might* be able to get them from the state if you call and ask real nice. For older election results, you have to rely on a printed volume, named “PD 43.”
- *Population statistics.* The best place to get data about the age distribution of towns is from the American FactFinder (<http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml>), which is part of the U.S. Census Bureau.

Bibliographic example

- Matthew E. Kahn and John G. Matsusaka. “Demand for Environmental Goods: Evidence from Voting Patterns on California Initiatives.” *Journal of Law and Economics* 40(April): 137–173.

Comments/hints: This is a project in which getting the data may be more work than you initially think, so don’t hesitate assembling your datasets. You should conduct this analysis at the municipality level, not the county level, since Massachusetts has few counties (14) but a lot of municipalities (351).