

Political Methodology Comprehensive Examination, January 2021
Department of Political Science, George Washington University

Instructions: Read all questions before answering any of them. When you use substantive examples in your answers, we strongly prefer to see examples from political science. Answer all questions in part I. Answer 3 questions in part II (or the game theory question and one other). If you are completing the exam at GW, feel free to hand-write answers or parts of answers in a blue book, but carefully label them and note that you are using the blue book in your typed document. If you are completing the exam at home, you can similarly include photos/scans of hand-written material. Good luck!

Part I

1. After showing a scatterplot exhibiting a positive linear correlation between a pretest X and a posttest Y ($r \approx .65$), a presenter comments that the lowest decile observations on the X variable increased more on average than the highest decile observations on the X variable. He proceeds to give an elaborate explanation as to why this suggests a major endogeneity threat: Perhaps the low-X units were given an intervention? How would you respond to this?
2. When discussing a paper at a conference, you raise questions about the endogeneity of the main independent variable of interest (X) in the author's OLS model. The author dismisses your concerns, stating that he saved the residuals (e) and then computed the correlation of e and X. Since the correlation was only 0.11, he claimed that "if there is endogeneity there, its effects must be very small." How do you respond to that? What follow-up questions do you ask?
3. In your dissertation, one chapter looks at the effect of democracy on economic growth. You have a main outcome of interest **gdp_grow** (measured as % growth), a dichotomous treatment of interest **democracy**, and a set of five other established controls from the literature: **education** (schooling years in the population), **loggdp** (logged GDP/capita), **gini** (inequality), **ucdp_civwar** (a dichotomous civil war indicator), and the **year**. There are no country fixed or random effects. The output from a simple OLS model (using a country-year panel) is shown below.

Source	SS	df	MS	Number of obs =	5281
Model	2054.73167	6	342.455279	F(6, 5274) =	9.86
Residual	183132.797	5274	34.7237006	Prob > F =	0.0000
Total	185187.529	5280	35.0733956	R-squared =	0.0111
				Adj R-squared =	0.0100
				Root MSE =	5.8927

gdp_grow	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
democracy	.1641523	.1910718	0.86	0.390	-.2104275 .5387321
education	.050069	.0463657	1.08	0.280	-.0408269 .140965
loggdp	.1733442	.1160689	1.49	0.135	-.0541989 .4008873
gini	-.029179	.010368	-2.81	0.005	-.0495046 -.0088534
ucdp_civwar	-.5038125	.2525438	-1.99	0.046	-.9989029 -.0087222
year	-.0264924	.0064385	-4.11	0.000	-.0391146 -.0138702
_cons	54.21794	12.89089	4.21	0.000	28.94645 79.48942

- a. According to the model, what is the % likelihood that the effect of **democracy** is greater than -0.21?
- b. According to the model, what is the expected difference in growth between a country experiencing civil war in 1980 and a country at peace in 2000?
- c. Advisor 1 worries that including **education** in the model might lead you to underestimate the effect of **democracy**. Advisor 2 objects and says that it's always better to include too many controls rather than too few. Is this objection correct?
- d. Advisor 1 points to your F statistic of 9.86 and says you're doing a great job explaining what leads to economic growth. Advisor 2 disagrees and points to a different statistic. What statistic should he or she have pointed to? What is the proper way to interpret the F statistic?
- e. You're curious whether **democracy** might have a more positive effect on growth in poor countries. Advisor 1 tells you that can't be since the coefficients on **democracy** and **loggdp** are both non-significant. Is that a valid response? How would you test your theory?
- f. Someone asks you whether your results on the gini coefficient are driven by a specific country-year. Is this more likely for a country-year with an extreme value on the gini coefficient or an average value? How would you check whether your results are dependent on a single case like this?

After running the regression pictured above, you:

- g. Save the fitted values ($\widehat{gdp_grow}$), add the square and cube of the fitted values to your original equation, run that regression, then test for the joint significance of the squared and cubed terms. Why did you do this? What should you do if you reject the null hypothesis?
- h. Save the residuals and lag the residuals. You then regress the residuals on the lagged residuals, and test for the significance of the coefficient on the lagged residuals. Why did you do this? What should you do if you reject the null hypothesis?

Part II

1. Proponents of experimental research argue that a key advantage of experiments is that randomization ensures unbiased estimates of causal effects. What core assumptions are required for this to be the case? In political science research, what are types of research areas or situations when these assumptions might fail to hold? When designing experimental research, what steps can scholars take to make these assumptions more likely to hold?
2. Sampling for population surveys involves a process of moving from a target population, to the sample that is drawn for the survey, to the sample of respondents that actually answers survey questions. What are the different forms of error that can arise in this process? What strategies can survey researchers use to minimize each form of potential error?
3. Some studies suggest that government spending is inefficient when combined with widespread corruption. For several decades, the formula for government transfers to Brazilian municipalities was determined in part by the municipality's population: municipalities with populations below a certain cutoff did not receive additional revenue, while states above the cutoff did. Suppose you want to estimate the causal effect of increased government spending on a number of development

outcomes, such as educational attainment and poverty rates. To do so, you will implement a regression discontinuity design. What are the required assumptions for this design in the context of this specific example? What would be a scenario in which this assumption is violated? How would you estimate how the treatment effect of spending varies with corruption levels?

4. Network analysis includes, among other things, a set of measurement tools. Give examples of two of these tools. Explain the theoretical basis for these tools, what advantages they have over more conventional measurement tools, and their limitations. For each tool, give an example of how one might use it in applied research, including an explanation of why it would be useful in this context and how it can improve our inferences.
5. Because of attrition or survey item non-response, experimental studies often have missing outcomes for some subjects/units. What are the potential consequences of this situation? What might you do to assess the severity of the problem? What strategies can you use to account for this issue in your analysis and what are their strengths and limitations?
6. Political scientists are often interested in testing causal mechanisms. What are some of the ways in which political scientists test hypotheses related to causal mechanisms and what are their strengths and limitations? If you were going to design an experiment that was mainly concerned with testing different potential causal mechanisms, what would you do?
7. Policymakers have implemented a wide range of interventions to fight the spread of COVID-19. Suppose you want to evaluate the effectiveness of lockdown policies on COVID cases by comparing locality A to locality B. The former imposed a strict lockdown policy in mid-March, while the latter imposed a similar policy in early May. You will employ a difference-in-differences research design to estimate causal effects of such policy on covid cases. How would you do it? What are the required assumptions for this design to be valid? Discuss potential threats to the validity of this design in this particular example, as well as the broader challenges of this approach to estimating causal effects of counter-COVID measures.

Game Theory: Counts as two questions

Consider an activist who wants a dictator to implement a political reform. The activist comes in three types: Radical, Moderate, and Quiet. The dictator's prior beliefs over these types are given by q_R , q_M , and $q_Q = 1 - q_R - q_M$. The order of the game is as follows:

1. The activist chooses to protest or not at cost c .
2. The dictator implements the reform or not.
3. The activist chooses to launch a revolution or not, at cost d to both players and with likelihood of success p .

The payoffs are such that Radical types will revolt no matter what. Quiet types will never revolt, but prefer getting the reform. Moderate types will revolt if and only if the reform is *not* granted (i.e., the reform satisfies them). Implementing the reform costs the dictator 1, with $d > 1$. The dictator also gets benefit W from ruling and 0 otherwise. If a revolution is attempted, the activist's payoff does not depend

on whether the reform was granted (since they'll either be in charge or in jail), but assume the dictator still loses 1 by granting the reform.

- (a) What is the total payoff to the dictator if they do not reform and face revolt? What is the total payoff to the dictator if they reform and avoid revolt?
- (b) After seeing step 1, the dictator will update their beliefs on the type they are facing. Call the updated beliefs in step 2 q'_R , q'_M , and q'_Q . For what set of updated beliefs will the dictator implement the reform in step 2?
- (c) What are the conditions for each type of activist to protest in step 1?
- (d) Using (b) and (c), under what conditions is there a separating equilibrium? (This includes cases where two of the three types overlap, but the third does something different.)
- (e) In the separating equilibrium, what is the probability that reform occurs (assuming the dictator's initial beliefs give the correct probabilities of each type)? What is the probability of revolt?
- (f) How does the structure of signaling in step 1 and/or payoffs for the activist types need to change to get an equilibrium that is maximally beneficial for the dictator?