Political Methodology Comprehensive Examination, May 2022
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 (but note the game theory question counts for two questions). 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. Suppose each person has a true cognitive skill level CS_i. A researcher gives each person a test that produces a measure IQ_i of their cognitive skill with some error: $IQ = CS + \varepsilon$, where ε is a random draw from some distribution. Additionally, suppose that each person can accurately estimate their own cognitive skill with no error. If the researcher regresses each person's self-report of their cognitive skill on the measured IQ, what are they likely to find besides an overall positive correlation? What kinds of faulty conclusions might one draw from this?

2. Your colleague fit a multivariate linear regression model estimating the county-level relationship between GOP Vote Share in the 2020 election and a series of socio-economic indicators. When examining the results, he found that all but three variables have statistically significant p-values. He decides to drop those three insignificant variables and keep all the remaining predictors. He says this is to minimize the bias-variance trade-off because irrelevant variables just add noise and unnecessary complexity to the model. What are the possible problems with your colleague's method?

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
					F(6, 5274) =	9.86
Model	2054.73167	6	342.455279		Prob > F =	0.0000
Residual	183132.797	5274	34.7237006		R-squared =	0.0111
					Adj R-squared =	0.0100
Total	185187.529	5280	35.0733956		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 **education** is less than 0.141?

b. According to the model, what is the expected difference in growth between a country experiencing civil war in 2000 and a country at peace in 1970?

c. 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?

d. Advisor 1 worries that including **education** in the model might lead you to under-estimate the effect of **democracy**. Under what specific circumstances would this objection have merit?

e. Advisor 2 disagrees and says that it's always better to include too many controls rather than too few. Is this objection correct? If not, when does the principle go wrong?

f. 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. Advisor 2 responds that you can simply run an F test on those two variables to see if there's any joint prediction. Who's wrong and who's right here and why? If you disagree with both, how would you test your theory?

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?

i. Stratify the model by running the regression only in a single decade for each decade in your sample. You find that **democracy** is now significant in two specific decades. Can you conclude that democracy's effect varies over time? What could you do next to investigate further?

Part II

1. Describe what the SUTVA assumption is and give an example of how it can be violated. What are some of the ways that SUTVA violations can be minimized? What are some of the common approaches to dealing with SUTVA violations and what are their strengths and weaknesses?

2. Several years from now, you want to test the effect of state-level abortion bans/limitations in the United States on ensuing support for the two major parties. Let's ignore any potential legal challenges and just think about testing the bills' effects. What are the inferential problems with the simple test where you regress party support on the previous passage of a ban/limitation? Describe several alternative testing strategies. What are the strengths and weaknesses of these alternatives?

3. Suppose we are trying to assess how well an LDA model predicts "fake" news articles from real ones on Facebook. For the purpose of this test, we consider a positive result to be a "fake" article and a negative result to be a "real" news article. After fitting the model in R, we compare predicted values from the actual values, as shown below:

```
> lda.fit = lda(fakenews ~ x1 + x2, data=train)
> lda.pred = predict(lda.fit, train)
> table(lda.pred, train$health)
            train
lda.pred  fake  real
fake       16    23
real       35    89
```

a) What is the misclassification rate of the LDA model?
b) How could we decrease the rate of false positives to false negatives and how is this likely to affect the misclassification rate?
c) Suppose you also ran a logit model for comparison and found the logit model had a smaller misclassification rate than the LDA on the same data. Why might logit perform better than LDA?

4. You are reviewing a paper examining covid rates across Democratic and Republican voters inside of the United States. The author has the following table in their paper:

Table 1: Effect of Vaccinations on Covid Among Different U.S. Voters

	Democrats		Republicans	
	Cases	Total	Cases	Total
Vaccinated	12	600	8	400
Not Vaccinated	20	400	20	600

The author reasons that "The average rate of vaccination among the Democrats is 60% while it is only 40% among Republicans. However, the rate of covid is higher among Democrats (3.2%) than among Republicans (2.8%). Examining the difference in means across the two groups, this suggests vaccinations are largely ineffective."

a) How would you respond to the author's reasoning and with what evidence?
b) What is this error in inference an example of?
c) Moving forward, what would be a better modeling approach to estimating this relationship and why?

5. A major challenge when estimating non-linear models such as logit, ordered probit, multinomial logit, Weibull models, etc., is deciding what marginal effects to present to your audience. What are some of the key ways that marginal effects can vary? What principles should you apply when presenting and interpreting marginal effects of non-linear models? Are different types of marginal effects more appropriate given different research questions?

6. Researchers are increasingly interested in estimating causal mediation models. Why do you believe this is? Give an example of a causal relationship that would be theoretically enriched by estimating the role of a mediator. What substantive insights might be gained from this? When doing causal mediation work, what are the key principles that you think lead to credible and valuable findings?

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?