PECOS4022

Written Examination

November 9th, 2021

Instructions: This exam consists of four overall questions, which each includes some sub-questions. Each of the question counts equally towards the final grade and you must answer all of them.

Question 1: Interpreting a regression

The regression output below is from the article "Insurgency and Ivory: The Territorial Origins of Illicit Resource Extraction in Civil Conflicts" by Felix Haass, published in Comparative Political Studies in 2021. Haass argues "that **insurgents extract resources as a consequence of their armed struggle against a government over territorial authority**. When rebels lack territorial control, they also lack the necessary market access to start extracting resources. Alternatively, without income or civilian support structures generated at least partially through resource extraction it is difficult for rebels to firmly establish territorial control. Instead, it is when rebels actively compete with a government to establish political authority over territory that they intensify resource extraction."

He tests this argument with the *proportion of illegally killed elephants* (PIKE) as the dependent variable in an OLS regression model. The main independent variable is *Territorial Competetion*, counting the number of rebel actions with the intention to gain political control over a given territory. These actions are (1) headquarters or bases being established by rebels; (2) strategic nonviolent activities by rebels, such as recruitment drives or troop movements; (3) rebels taking over territory; and (4) one-sided violence by rebels. Each observation is a monitoring site, which is a specific area. For example, Blindern could be one such site and Sagene another site.

1. How does Haass find support for his claim? How would you interpret the coefficients for the main independent variable?

2. The model includes two terms that are log-transformed, population and GDP per capita What is the rationale for transforming these variables?

3. Do insurgents extract more resources in sites with higher liberal democracy scores?

4. Haass, among other things, include site fixed effects and year fixed effects. What is the motivation for this?

Question 2: Types of regression

For each of the four questions below, identify the independent variable, the dependent variable, how they can be coded, and the preferred type of regression to analyze the question. Give a brief explanation for your answer.

1. What is the effect of earthquakes on the outbreak of civil wars?

2. Do democracies provide higher economic growth than autocracies?

Table 1. Territorial Competition and Poaching Rates.

	Dependent variable				
	PIKE				
	(1)	(2)	(3)	(4)	(5)
Territorial competition	0.023***	0.018***	0.024**	0.020**	0.021**
	(0.007)	(0.006)	(0.010)	(0.008)	(0.010)
Gov't/rebel battles			−0.004	−0.004	−0.003
			(0.003)	(0.003)	(0.003)
Liberal democracy$_{t-1}$					0.466
					(0.696)
Corruption$_{t-1}$					0.384
					(0.457)
Population (log)$_{t-1}$					−0.669
					(1.440)
GDP/PC (log)$_{t-1}$					−0.019
					(0.187)
Site fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	No	Yes	Yes	Yes	Yes
Country-specific trends	No	No	No	Yes	No
Number of countries	15	15	15	15	15
Number of sites	33	33	33	33	33
Observations	429	429	429	429	377
Adjusted R^2	.362	.413	.390	.450	.397

Robust standard errors clustered by country in parentheses: *p < .1. **p < .05. ***p < .01.

Figure 1: Regression output from Haass (2021).

3. Were users of social media more likely to protest following the 2021 Russian legislative election?

4. Does the crime level in a community affect how many times police personnel uses their guns?

Question 3: Research Design

Imagine you are planning to study whether female leaders have done a better job at handling the COVID-19 pandemic than male leaders. You have unlimited resources to conduct the project. Outline how you would design this study. You can, for example, think about the variables would you need to collect or the type of analyses you would conduct. Furthermore, discuss the advantages and disadvantages of your chosen research design.

Question 4: Regression Diagnostics

The image below shows four diagnostic plots of an OLS regression. Using the plots, answer the following questions:

1. Which of the OLS assumptions are violated?

2. How are the violations visible in the `Histogram of Residuals` and the `Residual vs. x` plots, and what would the two plots look like if no assumption was violated?

3. What could be the cause for the assumption violations, and what could you change about the model specification to address the assumption violations?

4. Imagine the data for the model stems from one country. If we collected the same data for multiple other countries and repeated the regression with the additional data from multiple countries, what main problem could this cause for the model's standard errors, and how could this be addressed?