PSC-202--Intro to Political Analysis
Week 12 Section Practice

Part 1: Controlled Comparison Exercise

1.1. You suspect there is a state-level difference regarding support for a nation-wide ban on abortion. Specifically, you have reason to believe that southern states are more likely to support a ban on abortion than northern states. To test your theory, you surveyed 200 respondents' opinion on a ban of abortion in Texas and New York separately (200 respondents in each state). Your classmate theorizes that religion is a confounding variable that may affect your results. As a result, you decide to use data from a survey that controls for religion (divided here into two categories: atheist and Christian).

Scenario 1:

Suppose you run another round of the survey and find an association shown in the following table.

Ban on Abortion?	Support Ban on Abortion?					
	Atheist			Christian		
	Texas	New York	Total	Texas	New York	Total
Yes	44 (44%)	42 (42%)	86	38 (38%)	36 (36%)	74
No	56 (56%)	58 (58%)	114	62 (62%)	64 (64%)	126
Total	100	100	200	100	100	200

Q1. Calculate the controlled effect of living in a southern state among atheists.

Q2. Calculate the controlled effect of living in a southern state among Christians.

Q3. Looking at the partial effects, what is the nature of the relationship (additive, spurious, interactive)?

Part 2: Multiple Regression

A researcher is interested in whether higher levels of education cause higher crime rates. She theorizes that higher levels of education should *lower* crime rates, but an examination of data from 67 counties in the state of Florida reveals a *positive* relationship between **crime rate (annual number of crimes per 1000 population)** and **education (percentage of adult residents having at least a high school education)**, with a bivariate relationship that can be represented as:

$$\text{Crime rate} = -51.3 + 1.5(\text{education rate})$$

In other words, as the percentage of county residents having at least a high school education increases, so does the crime rate.

The researcher then wonders if there is a confounding variable that she has not taken into account. Looking at the data, she sees strong positive associations between crime rate and urbanization ($r = 0.68$) and education and urbanization ($r = 0.79$). She wonders if perhaps the association between crime rate and education is spurious. Perhaps urbanization is a common causal factor. Urbanization is operationalized as the percentage living in an urban environment. As urbanization increases, both crime rate and education increase, resulting in a positive correlation between crime rate and education.

When the researcher includes urbanization in her analysis, she finds the following:

Crime rate = 58.9 – 0.6(education rate) + 0.7(urbanization rate)

Table of regression coefficients:

Variable	Coefficient	Standard Error
Intercept	58.9	2.57
Education Rate	-0.6	0.18
Urbanization Rate	0.7	0.25

1) According to these estimates, what is the effect of education rate on the crime rate in a county? Explain this is in a way that would be understood by someone unfamiliar with regression.

2) Suppose a county had 50% of its adult residents with at least a high school diploma and an urbanization rate of 78%. What would be the predicted crime rate in that county (measured in annual number of crimes per 1000 population) (Using the number before the % sign to calculate the result, no need to turn the percentage into a decimal)?

3) Franklin County and Gadsden County both have education rates (percentage of adult residents with at least a high school diploma) of 23%. Franklin County has an urbanization rate of 0% whereas Gadsden County's urbanization rate is 37%. What is the predicted difference in the crime rate (Using the number before the % sign to calculate the result, no need to turn the percentage into a decimal)?

4) Which of these coefficients are statistically significant at the 0.05 level?