

a. In words, carefully explain and interpret the coefficient estimate for the treatment effect

in Table 3, column 4. [2 points]

b. Table 3 reports standard errors for the coefficient estimates in brackets. In your own

words, explain what these quantities represent. (Ignore the fact that these are “clustered”

standard errors). What is the difference between a standard error and a standard deviation?

What is the standard deviation of math test scores in this study? [4 points]

2) The following six scenarios describe a naïve conclusion based on the observed association

between two variables. For each scenario, identify: (1) the “treatment” being described; (2) the “outcome” variable of interest; and (3) the relevant counterfactual outcome necessary to assess the causal effect implied. In each case, describe whether the stated conclusion is a sound one, and if not, another plausible explanation (or “data generating process”) for the observed association.

a. When comparing crime rates across U.S. cities, cities that have more police per capita

also have higher crime. High-crime cities are therefore likely to see little to no effect on

crime from additional investment in police protection.

b. A study in the Annals of Improbable Research once reported that counties with large

numbers of mobile-home parks had higher rates of tornadoes than the rest of the

population (your professor grew up in Kansas—this definitely appears to be true). From

this observation, the authors discourage construction of mobile-home parks, as they

raise the likelihood of severe weather events.

c. An analysis of over 50 million anonymous Google search queries including the word

“depression” found comparatively fewer searches in certain locales, dates, and seasons.

For example, North Dakota had many more searches per capita related to depression

than California; mid-August had fewer searches than January; Christmas had fewer

searches than almost all other days; and Mondays had more searches than Sundays.

These findings suggest that depression is strongly predicted by location (i.e. weather) and

time of year/weak.

d. In a recent study of men aged 21-30, the number of hours shooting pool (billiards) per

week was strongly correlated with liver disease later in life. A policy recommendation

resulting from this study would be a high tax on billiards play (or, better yet, banning this

dangerous activity altogether).

e. In a study on the effect of class size on the performance of 6th

graders on a standardized test, it was found that kids who were in small classes frequently performed much better than kids who were in large classes. We can conclude from this finding that smaller class sizes are very important for student success on standardized tests.

f. A researcher found that average test performance of children with divorced parents was

lower than average test performance of children with intact families. This researcher then

concluded that divorce is bad for children’s test outcomes.

3) Imagine you are conducting a study of the effectiveness of adjunct statistics professors relative to tenure-track and tenured professors in the same subject. Your university, Big State U, offers multiple sections of introductory statistics, some of which are taught by adjunct professors and some of which are taught by tenure-track or tenured professors. You have access to two potentially useful measures of effectiveness: professors’ teaching evaluations, and a standard end-of-semester final exam that all Big State U statistics students are required to take. You also have access to other relevant information about students who took these tests: their GPA, high school class rank, gender, and so on.

Write 2-3 sentences assessing each of the following strategies for evaluating the teaching

effectiveness of adjunct professors relative to full-time professors. Explain why each is not an

ideal test of their relative effectiveness.

a. You compare the average teaching evaluation score (on a scale of 1-100) for the two

groups and find that adjunct professors score an average of 20 points higher than regular

professors.

b. You compare the average final exam score (also on a scale of 1-100) for students of the

two groups and find that students of adjunct professors score 15 points lower than

students of regular professors.

c. You estimate a multiple regression for final exam scores, that relates these test scores to

(1) type of professor, and (2) other students characteristics—their GPA, high school

class rank, gender, etc.