

Homework #6

not collected (final exam prep)

Problem 1 The following is an abstract of an economics research paper:

This paper compares corruption in China over the past 15 years with corruption in the U.S. between 1870 and 1930, periods that are roughly comparable in terms of real income per capita. Corruption indicators for both countries and both periods are constructed by tracking corruption news in prominent U.S. newspapers. The comparison indicates that corruption in the U.S. in the early 1870s when its real income per capita was about \$2,800 (in 2005 dollars) was 7 to 9 times higher than Chinas corruption level in 1996, the corresponding year in terms of income per capita. By the time the U.S. reached \$7,500 in 1928 — approximately equivalent to Chinas real income per capita in 2009 — corruption was similar in both countries. The findings imply that, while corruption in China is an issue that merits attention, it is not at alarmingly high levels, compared to the U.S. historical experience.

What objections or concerns is the author most likely to run into when he presents his paper to other economists?

The main issue I see is one of reporting. Are contemporary Chinese newspapers as likely as were turn-of-the-century US newspapers to report corruption?

Problem 2 A doctor's family practice clinic treated 100 patients with flu-like symptoms in 2012. 50% of those who were advised only to rest at home got better within 2 days. 40% of those who were given Fluaway, an experimental new drug, got better within two days.

a. You are hired as a consultant to the clinic. You disaggregate the data so that you can look at children and adults separately. Describe, using the table below, data that would lead you to conclude that Fluaway is a *more* effective treatment than rest alone (there are many correct answers!).

Children only	Improved	Not improved	% improved
Fluaway			
Rest			

Adults only	Improved	Not improved	% improved
Fluaway			
Rest			

Many answers will work. Suppose 50 kids and 50 adults are treated. Suppose further that 10 kids get fluaway, of whom 6 improve, while 23 of the 40 who rest improve. Suppose also that 14/40 adults who receive Fluaway improve, versus 2 of 10 adults who rest. You can verify that these disaggregated data are consistent with Fluaway being more effective.

b. Describe, using the table below, data that are consistent with the conclusion that Fluaway is a *less* effective treatment than rest alone (again, there are many correct answers).

Many answers will work. Suppose 12/30 kids and 12/30 adults who get Fluaway improve, versus 10/20 kids and 10/20 adults who rest. Verify that these numbers are consistent with rest being more effective.

Children only	Improved	Not improved	% improved
Fluaway			
Rest			

Adults only	Improved	Not improved	% improved
Fluaway			
Rest			

Hint: One of parts a-b is hard, one is easy.

Problem 3 A researcher is interested in the effects of smoking and body mass on cardiovascular health. He finds a random sample of 2,300 30 year old smokers, and estimates the following regression equation:

$$CV_HEALTH = \beta_0 + \beta_1 * CIGS + \beta_2 * CIGS^2 + \beta_3 * MALE + \beta_4 * BMI + \beta_5 * BMI * MALE + \epsilon$$

CV_HEALTH is a measure of cardiovascular health determined by a series of tests such as running on a treadmill, and measured on a scale of 0 to 100, with 0 being the lowest possible score, and 100 the highest. $CIGS$ measures the number of cigarettes a person smokes on a typical day. $MALE$ is a dummy variable equal to 1 if the person is male, and 0 if female. BMI equals mass divided by height squared, with higher values meaning someone weighs more.

He obtains the following results:

$$R^2 = .452 \tag{1}$$

variable	coefficients	standard error	t-stat	P-value
<i>Intercept</i>	80	16	5	5.733 E-7
<i>CIGS</i>	.02	.008	2.5	.0124
<i>CIGS</i> ²	-.02	.005	-4	6.3E-5
<i>MALE</i>	2	2.5	.8	.424
<i>BMI</i>	-.8	.267	-3	.0027
<i>BMI * MALE</i>	-.2	.1	-2	.0455

a. According to the regression results, which variables are significant explainers of CV_HEALTH at the 5% level?

CIGS, *CIGS*², *BMI*, and *BMI * MALE*.

b. According to the regression results, what is the average CV_HEALTH score for a non-smoking male with a BMI of 24?

58

c. How much do the results predict that smoking 10 cigarettes per day decreases CV_HEALTH ? How much does smoking 20 cigarettes per day decrease CV_HEALTH ?

10 cigarettes decrease CV_Health by 1.8, while 20 decrease CV_Health by 7.6.

d. Interpret the estimate of the coefficient β_5 .

β_4 is the change in CV_Health given a one-unit change in BMI for women. $\beta_4 + \beta_5$ is the corresponding change for men. So, β_5 is the differential effect of BMI on CV_Health for men.

Problem 4 Lana Sociologist is interested in the effect of breast feeding on the cognitive development of infants. She randomly samples 10,000 US 40-year old adults, finds that 6,000 of them were breast fed as babies, while 4,000 of them were not. She finds that the average annual income of the first group is \$64,000, while the average annual income of the second group is \$52,000. The standard deviations are low enough to reject a hypothesis test with the null hypothesis that $\mu_1 = \mu_2$, so she concludes that breast feeding does indeed increase cognitive development, as reflected in the labor market earnings of adults (a wide variety of prior studies have found a positive link between intelligence and labor market earnings).

Give a detailed analysis of the additional information you would need to evaluate Lana's claim. Your answer should include information on what concerns you have about her study, and what additional data you would need to be shown to address those concerns.

There may be more than one plausible answer, but the main issue I see is one of whether or not we can accurately compare the breast-fed group to the non-breast-fed group. Perhaps breast feeding is correlated with parental involvement, and this is a more important determinant of eventual labor market success.

Problem 5 A recent article in *The Economist* ("Robot recruiters," 4/6/2013) makes several statistical claims. Answer the following.

a. The article states that "...for customer-support calls, people with a criminal background actually perform a bit better." Do you take this result as evidence that having a criminal background or not is irrelevant to performance as an employee?

The main issue seems to be one of comparability of the two groups, convicts and non-convicts who apply for call center jobs. Perhaps the convicts who apply are the best of their group, while the non-convicts who apply to call centers are the least employable of their group. It's hard to imagine that a more selective job (e.g. fireman or plumber) would generate similar results.

b. The article also states that "...people who fill out online job applications using browsers that did not come with the computer (such as Microsofts Internet Explorer on a Windows PC) but had to be deliberately installed (like Firefox or Googles Chrome) perform better and change jobs less often." What do you make of this strange association?

Browser choice by itself is surely irrelevant to employability, but perhaps something else is correlated with browser choice, such as general comfort with computers.

Problem 6 Discuss the main contribution of the Tabarrok and Klick paper from class, "Using terror alert levels to estimate the effect of police on crime". Why is it necessary to resort to such unusual data to

estimate the effect of police on crime? What is the point of including the variable measuring midday Metro ridership? How do you interpret the coefficients for High Alert in Table 2 and for High Alert X District 1 and High Alert X Other Districts in Table 4? Explain how the authors calculate the elasticity of crime with respect to police from their results.