

Week 8 and 9 Discussion Section Questions

Questions 7.9, 7.2 and 7.14

Ryan Martin

February 20, 2018

Table of Contents

Question 1

Question 2

Question 3

Question 1

7.9 i

In the STAR experiment (Section 7.5.3), children were randomly assigned within schools into three types of classes: small classes with 13-17 students, regular-sized classes with 22-25 students, and regular-sized classes with a full-time teacher aide to assist the teacher. Student scores on achievement tests were recorded, as was some information about the students, teachers and schools. Data on kindergarten classes is contained in the data file `star.dat`

- a Calculate the average of `TOTALSCORE` for (i) students in regular-sized classrooms with full time teachers, but no aide; (ii) students in regular-sized classrooms with full time teachers and an aide; and (iii) students in small classrooms. What do

7.9 ii

you observe about test scores in these three types of learning environments?

- b Estimate the regression model

$$TOTALSCORE_i = \beta_1 + \beta_2 SMALL_i + \beta_3 AIDE_i + e_i$$

- c Use a t-test to test the hypothesis $H_0 : \beta_2 \geq 1$ against the alternative $H_1 : \beta_2 < 1$
- c To the regression in (b) add the additional explanatory variable TCHEXPER. Is this variable statistically significant? Does its addition to the model affect the estimates of β_2 and β_3 ?

7.9 iii

- d To the regression in (c) add the additional explanatory variables BOY, FREELUNCH and WHITE_ASIAN. Are any of these variables statistically significant? Does their addition to the model affect the estimates of β_2 and β_3 ?
- e To the regression in (d), add the additional explanatory variables TCHWHITE, TCHMASTERS, SCHURBAN and SCHRURAL. Are any of these variables statistically significant? Does their addition to the model affect the estimates of β_2 and β_3 ?
- f Discuss the importance of parts (c), (d) and (e) to our estimation of the “treatment” effects in part(b).

- g Add to the models in (b) through (e) indicator variables for each school.

$$SCHOOL_j = \begin{cases} 1 & \text{if student is in school } j \\ 0 & \text{otherwise} \end{cases}$$

Test the significance of these school “fixed effects.” Does the inclusion of these fixed effect indicator variables substantially alter the estimate of β_2 and β_3

Question 2

In September 1998, a local TV station contacted an econometrician to analyze some data for them. They were going to do a Halloween story on the legend of full moons' affecting behavior in strange ways. They collected data from a local hospital on emergency room cases for the period from January 1, 1998 until mid-August. There were 229 observations (days). During this time, there were eight full moons and seven new moons (a related myth concerns new moons) and three holidays (NYD, Memorial Day and Easter). If there is a full-moon effect, then hospital administrators will adjust numbers of emergency room doctors and nurses and local police may change the number of officers on duty. (This sounds like a very 90s era question.)

7.2 ii

Using the data in the file `fullmoon.dat` we obtain the regression results in the following table: T is a time trend ($T = 1, 2, 3, \dots, 229$) and the rest are indicator variables. $HOLIDAY = 1$ if the day is a holiday; 0 otherwise. $NEWMOON = 1$ if there is a new moon; 0 otherwise.

Emergency Room Cases Regression—Model 1

Variable	Coefficient	Std. Error	<i>t</i> -Statistic	Prob.
<i>C</i>	93.6958	1.5592	60.0938	0.0000
<i>T</i>	0.0338	0.0111	3.0580	0.0025
<i>HOLIDAY</i>	13.8629	6.4452	2.1509	0.0326
<i>FRIDAY</i>	6.9098	2.1113	3.2727	0.0012
<i>SATURDAY</i>	10.5894	2.1184	4.9987	0.0000
<i>FULLMOON</i>	2.4545	3.9809	0.6166	0.5382
<i>NEWMOON</i>	6.4059	4.2569	1.5048	0.1338

$R^2 = 0.1736$

$SSE = 27108.82$

- a Interpret these regression results. When should emergency rooms expect more calls?
- b The model was reestimated omitting the variables FULLMOON and NEWMOON, as shown below. Comment on any changes you observe.
- c Test the joint significance of FULLMOON and NEWMOON. State the null and alternative hypotheses and indicate the test statistic you use. What do you conclude?

Emergency Room Cases Regression—Model 2

Variable	Coefficient	Std. Error	<i>t</i> -Statistic	Prob.
<i>C</i>	94.0215	1.5458	60.8219	0.0000
<i>T</i>	0.0338	0.0111	3.0568	0.0025
<i>HOLIDAY</i>	13.6168	6.4511	2.1108	0.0359
<i>FRIDAY</i>	6.8491	2.1137	3.2404	0.0014
<i>SATURDAY</i>	10.3421	2.1153	4.8891	0.0000

 $R^2 = 0.1640$ $SSE = 27424.19$

Question 3

Professor Ray C. Fair's voting model was introduced in Exercise 2.14. He builds models that explain and predict the U.S. presidential elections. See his website at <http://fairmodel.econ.yale.edu/vote2008/index2.htm>. The basic premise of the model is that the incumbent party's share of the two-party popular vote is affected by a number of factors relating to the economy and variables relating to the politics, such as how long the incumbent party has been in power and whether the president is running for reelection. Fair's data, 33 observations for the election years from 1880 to 2008, are in the file `fair4.dat`. The dependent variable is $VOTE =$ percentage share of the

popular vote won by the incumbent party. The explanatory variables include:

- PARTY: Binary variable that's either 1 or -1. 1 if Democratic incumbent, -1 if Republican.
- PERSON: binary (0-1), 1 if incumbent is running for (re)election.
- DURATION: discrete variable. 0 if the incumbent party has been in power for one term, 1 if the incumbent party has been in power for two consecutive terms, 1.25 if the incumbent has been in power for three consecutive terms, 1.50 for four, etc.
- WAR. binary (0-1). 1 in 1920, 1944, 1948. 0 otherwise

- GROWTH: growth rate of real per capita GDP in the first 3 quarters of the election year (annual rate).
- INFLATION: absolute value of the growth rate of GDP deflator in first 15 quarters of the administration (annual rate) except for 1920, 1944 and 1948 (where values are 0).
- GOODNEWS: The number of quarters in the first 15 quarters of the administration in which the growth rate of real per capita GDP is greater than 3.2% at an annual rate, except for 1920, 1944 and 1948, where the values are 0.

7.14 iv

- a Consider the model

$$VOTE = \beta_1 + \beta_2 GROWTH + \beta_3 INFLATION + \beta_4 GOODNEWS + \beta_5 PE$$

Discuss the anticipated effects of the dummy variables
PERSON and WAR

- b The binary variable PARTY is somewhat different from the dummy variables we have considered. Write out the regression function $E(VOTE)$ for the two values of PARTY. Discuss the effects of this specification.
- c Use the data for the period 1916-2004 to estimate the proposed model. Discuss the estimation results. Are the signs as expected? Are the estimates statistically significant? How well does the model fit the data?

- d Predict the outcome of the 2008 election using the given 2008 data for values of the explanatory variables. Based on the prediction, would you have picked the outcome of the election correctly?
- e Construct a 95% prediction interval for the outcome of the 2008 election.
- f Using data values of your choice (you must explain them), predict the outcome of the 2012 election.