17.871
Spring 2000
Problem Set # 1: Using STATA

Handed out: February 12, 2002
Due back: February 21, 2002

1. Using a text editor such as EMACS or the STATA do-file editor, type the text from Exhibit 1 in the document "How to Use the *STATA* **infile** and **infix** Commands" into Athena and save it in a file named scores.dat on your home directory. Write a "do" file that will create a STATA data set from this raw data and save it as a file called "scores.dta". Turn in a "log" file that documents the STATA commands you used to read in the data and save it.

2. Find two tables that interest you in the *Statistical Abstract of the United States* which meet the following criteria: (1) they have between 25 and 52 observations and (2) they have the same units of analysis (e.g., states, years, nations).

   a. Call these two tables Table A and Table B. Create a STATA data set that contains one variable, plus the identifying variable (like state or year), from Table A. Save it. Create a STATA dat set that consists of one variable, plus the identifying variable (like state or year), from Table B. Save it. Merge the two data sets. Save the merged data set.

   b. Turn in the following:
      (1) A "do file" that shows how you created the data sets and merged them.
      (2) A printout of the data.
      (3) A short (one paragraph, 2 or 3 sentences) description of the tables you got your data from.

3. Do the following Review Exercises in Freedman, at the end of Chapter 6, beginning on p. 104. When you write your explanations, use complete, well-constructed sentences and paragraphs.

   Review Exercises: 1, 2, 3 (p. 104)
   Special Review Exercises: 13 (p. 107) and 15 (p. 108).

4. MIT wants to improve how it delivers mental health services to its students. It can choose between mutually exclusive (for this exercise) policy choices: expanding the size of the mental health staff or making mental health services more accessible by expanding walk-in hours. MIT wants to make the choice based on sound social science research. It identifies 10 "peer" universities that have recently expanded the size of its mental health staff and 10 other "peer" universities that have recently expanded walk-in hours. MIT researchers go to these universities and administer surveys to students to assess how satisfies they are with campus mental health services. The results of the survey are these: Students on the campuses that

expanded the size of the staff rate their satisfaction an average of 8.2 on a scale of 1–10. Students on the campuses that expanded walk-in hours rate their satisfaction an average of 9.0. The MIT researchers conclude that expanding walk-in hours is more effective in achieving student satisfaction than expanding staff.

Discuss the validity of this research design.

5. In the 2000 presidential election, George W. (Quincy) Bush received 47.9% of the popular vote case in the United States. The following reports the final "horse race" results from the various polls taken right before the election: (All of these results were reported the day before the election)

CBS: 44%
CNN/Gallup: 48%
IBD: 48%
Reuters: 46%
Voter.com: 50%

What accounts for these differences?