

17.871  
 Spring 2006  
 Take Home Assignment # 2  
 Handed out March 9  
 Due at the beginning of class March 23

- The following table reports the total level of federal spending from FY 1980 to FY 2003, in nominal (i.e., non-inflation-adjusted) dollars. Letting Y denote spending and X the year, run the following two regressions:

$$Y_t = \beta_0 + \beta_1 X_t + u_t \quad (\text{Eq. 1})$$

and

$$\ln Y_t = \beta_0 + \beta_1 X_t + u_t \quad (\text{Eq. 2})$$

Estimate each regression equation and report how you would substantively interpret  $\beta_1$  in each case.

Fiscal Year	Expenditures (\$millions)	Fiscal Year	Expenditures (\$millions)	Fiscal Year	Expenditures (\$millions)
1980	590,941	1990	1,253,165	2000	1,788,773
1981	678,241	1991	1,324,369	2001	1,863,770
1982	745,743	1992	1,381,655	2002	2,010,970
1983	808,364	1993	1,409,489	2003	2,157,637
1984	851,853	1994	1,461,877		
1985	946,396	1995	1,515,802		
1986	990,430	1996	1,560,535		
1987	1,004,082	1997	1,601,250		
1988	1,064,455	1998	1,652,585		
1989	1,143,646	1999	1,701,891		

Turn in a “do file” that would do any of the variable transformations needed to produce the answer you provide, plus the “log file” that shows the results of the regressions you ran.

2. African Americans vote at a lower rates than whites, even though it's now a generation since the passage of the Voting Rights Act. Some argue that this is because old habits die hard in the South, where the most stringent forms of Jim Crow were imposed. Others argue that this is purely due to the economic disadvantages that African Americans continue to labor under.

Address this question using a regression. On the class web site you will find a Stata dataset named `voting2004.dta`. This data set was taken from the Voting and Registration Supplement to the November 2004 Current Population Survey. It contains information about whether people voted in the 2004 election, plus some selective demographics.

The codebook is in the file `voting2004.txt`. Note that the variables of interest are rarely coded in an immediately useful form. So, an initial issue is how to recode the variables you find in order to do the analysis you want to do.

I expect you to do a multiple regression to address this question. Choose at least two independent variables, and possibly more. You may find it useful to run two regressions to address this question.

The one hint I will give you is about the dependent variable. The dependent variable should be a binary, or dummy variable, equal to 1 if the person voted, 0 otherwise. You will notice that the variable in the data set that carries this information, `PES1`, is equal to 1 if the answer was yes, 2 if the answer was no, and three different negative numbers for missing values. You are going to have to do something about this.

After settling on a regression that helps you answer the question the way you feel comfortable with, write two pages, or so, double-spaced, that do the following:

- (1) Describe the independent variables and why you chose them. Describe how you recoded the original variables in the data set to produce the final variables you used in your analysis.
- (2) Describe the regression you settled on, what the coefficients were, and what the coefficients mean.
- (3) Discuss how the results from your multiple regression differ from running a series of bivariate regressions and why they differ. (If you do more than one multiple regression, you may answer this question with respect to only one of the multiple regressions.)

Finally, turn in a printout of the “do files” that would allow us to reproduce any regressions or other analyses (like graphs or tables) that you report in your write-up. Also include a copy of the “log file” that is produced when you run these “do files.”

**File locations:**

data: /mit/17.871/www/2006/voting2004.dta

codebook: /mit/17.871/www/2006/voting2004.txt