POS 3713: Homework Assignment #6
Spring 2001
Due on Friday, April 20th

Instructions: **Type** your answers to the following questions. You are permitted to do any calculations by hand on paper and attach the work to your typed responses. You should, however, report the final results of your calculations in the typed portion of your assignment.

The purpose of this assignment is to introduce bivariate and multiple regression. We will be using data from the 1996 National Election Study.

Part A. Bivariate Explanations for Presidential Approval

Throughout this class, we have been looking at examples of the factors US citizens use to evaluate the President of the United States. A simple model of Presidential approval might take into account three factors: the citizen's party identification, evaluation of Presidential job performance (such as handling of the economy), and evaluation of Presidential personality (such as leadership skills). In mathematical form, Presidential Evaluation= f (party identification, job performance, personality). This exercise is designed to find out which of these factors has the strongest influence on Presidential approval. We will be examining 4 variables in this exercise: the Clinton feeling thermometer (as a measure of presidential approval), a respondent's evaluation of Clinton's handling of the economy, a respondent's party identification, and a respondent's evaluation of Clinton's leadership abilities. The variables are measured as follows:

Clinther: ranges from 0-100 with 0 representing the lowest evaluation of Clinton, and 100 the highest evaluation
Economy: ranges from 1-4 where 1 = approve strongly of the way Clinton is handling the economy, 2 = approve not strongly, 3 = disapprove not strongly, and 4 = disapprove strongly (note that higher values indicate a lower evaluation of Clinton's handling of the economy)
PartyID: ranges from 0-6 with 0 = strong democrat, 1 = weak democrat, 2 = independent leaning democrat, 3 = independent, 4 = independent leaning republican, 5 = weak republican, and 6 = strong republican
Leader: the question asked is "Does 'provides strong leadership' describe Bill Clinton extremely well, quite well, not too well, or not well at all?" The measure ranges from 1-4 where 1 = extremely well, 2 = quite well, 3 = not too well, and 4 = not too well at all. Note again that higher numbers on the scale indicate a lower evaluation of Clinton's leadership abilities.

Throughout these exercises, we will be using *ClinTher* as the dependent variable measuring Presidential approval, while *economy, partyID,* and *leader* are independent variables.

Question A1: What component of the presidential performance model does each independent variable measure? In words, describe the causal relationship you expect between each independent variable and Presidential approval.

Many times researchers begin a data analysis task by examining the bivariate relationships between the dependent variable and all possible independent variables, just to get a feel for the nature of the relationships. In other words, they estimate the coefficients for the following three bivariate regression models:

I. $ClinTher = a + B_1(Economy)$
II. $ClinTher = a + B_1(PartyID)$
III. $ClinTher = a + B_1(Leader)$

The output for each of these regression models is presented below. Use this information to answer questions A2-A4.

Model I: ClinTher= a + B_1(*Economy*)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.675[a]	.455	.455	21.74

a. Predictors: (Constant), CLINTON ECONOMY STRENGTH

ANOVA [b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	653605.2	1	653605.208	1382.522	.000[a]
	Residual	782895.5	1656	472.763		
	Total	1436501	1657			

a. Predictors: (Constant), CLINTON ECONOMY STRENGTH

b. Dependent Variable: Clinton thermometer

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	99.830	1.205		82.867	.000
	CLINTON ECONOMY STRENGTH	-18.239	.491	-.675	-37.182	.000

a. Dependent Variable: Clinton thermometer

Model II: ClinTher= a + B_1(*PartyID*)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.680[a]	.462	.462	21.74

a. Predictors: (Constant), PARTY ID SUMMARY

ANOVA [b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	685197.4	1	685197.390	1449.560	.000[a]
	Residual	796488.6	1685	472.694		
	Total	1481686	1686			

a. Predictors: (Constant), PARTY ID SUMMARY

b. Dependent Variable: Clinton thermometer

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	85.131	.858		99.191	.000
	PARTY ID SUMMARY	-9.591	.252	-.680	-38.073	.000

a. Dependent Variable: Clinton thermometer

Model III: ClinTher= a + B_1(*Leader*)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.730[a]	.533	.533	20.26

a. Predictors: (Constant), CLINTON STRONG LEADER

ANOVA [b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	792128.2	1	792128.191	1930.223	.000[a]
	Residual	692724.3	1688	410.382		
	Total	1484853	1689			

a. Predictors: (Constant), CLINTON STRONG LEADER

b. Dependent Variable: Clinton thermometer

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	118.521	1.432		82.744	.000
	CLINTON STRONG LEADER	-24.972	.568	-.730	-43.934	.000

a. Dependent Variable: Clinton thermometer

Question A2: In symbolic form, state the null and research hypotheses for the slope coefficients in each of the bivariate regression models.
Question A3: Evaluate the goodness of fit of each regression model by interpreting R^2 and the significance of the F-test.
Question A4: Based on the t-statistics, are the slope coefficients in each model significantly different from zero? In words, interpret the meaning of each slope coefficient. Based on these results, which of these variables influence Presidential approval?

Part B. Multivariate Explanations for Presidential Approval

Now that you have gained some intuition about the relationships between the independent variables and Presidential approval, you can use multiple regression to see how all the variables work together. We will estimate the following multiple regression model:

ClinTher = a + B_1 (*Economy)* + B_2 (*PartyID*)+ B_3(*Leader*)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.835[a]	.697	.697	16.28

a. Predictors: (Constant), CLINTON STRONG LEADER, PARTY ID SUMMARY, CLINTON ECONOMY STRENGTH

ANOVA [b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	996554.1	3	332184.698	1252.585	.000[a]
	Residual	433070.7	1633	265.199		
	Total	1429625	1636			

a. Predictors: (Constant), CLINTON STRONG LEADER, PARTY ID SUMMARY, CLINTON ECONOMY STRENGTH

b. Dependent Variable: Clinton thermometer

Coefficients [a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	120.240	1.199		100.307	.000
	CLINTON ECONOMY STRENGTH	-7.705	.475	-.284	-16.218	.000
	PARTY ID SUMMARY	-4.328	.241	-.308	-17.984	.000
	CLINTON STRONG LEADER	-13.591	.612	-.400	-22.225	.000

a. Dependent Variable: Clinton thermometer

Question B1: The null and research hypotheses for each slope coefficient are the same as above. However, the slope coefficients are now different, reflecting the influence of each independent variable while controlling for the influence of the others. Evaluate the goodness of fit of the multiple regression using the F-test and R^2.

Question B2: Based on the t-statistics, are the partial slope coefficients significantly different from zero? In words, interpret each of the slope coefficients. Do you notice any changes in the magnitude of the slope coefficients going from the bivariate regression models to the multiple regression models? Which variable appears to have the largest influence on Presidential approval? What are some possible theoretical explanations for these results?

Question B3: Regression equations can be used to predict levels of presidential approval for different types of citizens. Using the unstandardized partial slope coefficients estimated in the three variable model above (in the column labeled B), make two predictions about presidential approval. First, predict the level of approval for a Strong Democrat when *Economy* and *Leader* are at their mean values in the sample. Second, predict the level of approval for Strong Republican when *Economy* and *Leader* are at their mean values in the sample. For an example of

how to use regression to make predictions, read section 18.6 (starting p.456) in Healey. Discuss whether or not the predictions corroborate you theoretical expectations.

The means for the independent variables are presented below.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
CLINTON ECONOMY STRENGTH	1663	1	4	2.20	1.09
PARTY ID SUMMARY	1695	0	6	2.68	2.10
CLINTON STRONG LEADER	1697	1	4	2.37	.87
Valid N (listwise)	1641				