Second Mid-Term Exam
Multiple Regression

Question A:

Public policy analysts are interested in understanding how and why individuals come to develop the opinions they do of various public policy issues. In Colorado, the policy known as Amendment 2 has recently been the subject of much attention and controversy. Amendment 2 is a voter initiative which prohibits anti-discrimination legislation protecting homosexuals. One of the groups lobbying against Amendment 2 has asked a public policy analyst to help them determine which factors are related to an individual's stance on Amendment 2. To answer the lobby group's questions, the analyst collected data on the following variables from a random sample of 200 registered voters in Denver and Boulder:

STANCE: An individual's policy stance on Amendment 2 (on a 1-5 scale, where 1 = strongly support, 3 = neutral, 5 = strongly oppose)

Political Variables:

PARTY: An individual's political affiliation (0 = Republican, l = Democrat)

LIBERAL: Self-ratings of an individual's liberalness or conservativeness (higher scores indicate more liberalness, lower scores indicate more conservativeness)

Sociodemographic Variables:

AGE: (in years)

GENDER: Individual's gender (0 = Male, l = Female)

INCOME: The average annual household income of the voter (in $1000/per year)

The public policy analyst would like to use the data collected to answer the following questions, and has asked you to help. Please use the above variables to provide Models C and A, n-PA, and PA-PC for each question. You may need to create some new variables for your Models C and A. If you do, please make sure you define these new variables somewhere in your answer.

1) Does the average stance towards Amendment 2 differ from neutrality? Please provide the simplest test which answers this question.

2) Is how liberal one rates oneself related to one's stance on Amendment 2?

3) Sociologists would suggest that one's sociodemographic background predicts one's opinion of Amendment 2. Is this true given the current data?

4) After statistically equating individuals on the sociodemographic variables, does one's liberalness predict one's stance on Amendment 2?

5) The lobby organization provides the analyst with a statistic from a past study which says that men are more likely to support Amendment 2 than women. Furthermore, given a 5 point scale of policy stance like the one used in the current study, men's stances are, on average, 1 point higher than women's. The analyst would like you to test this statistic given the current data.

6) Do the political variables as a set allow us to improve our prediction of one's opinion of Amendment 2, over and above the sociodemographic variables?

7) Assuming that all the sociodemographic variables are useful predictors, does the average individual oppose Amendment 2? Please provide the most powerful test which answers this question.

Question B:

Researchers conduct a survey that is to be presented in court on behalf of the plaintiff in a civil suit. Among the many survey questions are 14 individual attitude items that the researchers believed to be related. To simplify the analysis, the researchers construct a scale score for each respondent by summing the 14 responses to the individual attitude items. (We will call this summed score SUMATT.) The researchers regress the key dependent variable KEY on SUMATT and obtain a large $R^2 = .67$, which is statistically significant. The defendant, a large industrial company, hires a statistician to critique the analysis. This statistician regresses KEY on all 14 individual items. In this analysis, none of the individual coefficients are significantly different from zero, even though it seems obvious to a lay person, including the judge, that many of the individual items ought to be related to the dependent variable KEY. The defendant argues that

failure to find these expected relationships in the more detailed analysis casts doubt on the plaintiff's survey. How would you advise the judge?

Question C:

In this question we will continue our examination of the Faculty Course Questionnaire (FCQ) dataset. Our focus will continue to be on the variable R_INSTR, the average rating received by the instructor on a 0-to-4 rating scale. We will consider three types of variables:

Influenced or Controlled by the Instructor:

A_GRADE: The average grade (0-to-4 scale) assigned in that section.

Alternative Evaluation Measure:

P_RETURN: The proportion of the students enrolled in the class who returned an FCQ rating sheet. Possibly, some students in classes in which they did not like the instructor had already voted with their feet.

Demographic Variables about the Course and Instructor:

SIZE: Number of students enrolled in class

CRSLEV_N: The level of the course (from 1 to 4 representing 1000 to 4000 level courses)

TTRACK: Whether the teacher is tenured or on the track towards tenure (1) or whether the teacher is a limited-term instructor (0).

This time we are examining the sections (over several) years of a single, large department; hence, the number of sections is much lower than when we examined these data before. Use the attached printouts to answer the following questions. Some of the required statistics will be available directly on the printout, while others you will need to calculate from available information.

1. For this department, is the average grade assigned in a class related to the average rating the instructor receives? Using the simplest comparison that answers this question, provide PRE, F*, a statistical conclusion, and interpretation of the coefficient(s) related to this question.

2. Controlling for the average grade assigned in each section, does the alternative evaluation measure P_RETURN indeed predict instructor ratings? Provide PRE, F*, a statistical conclusion, and interpretation of the coefficient(s) related to this question. Be precise in your interpretation of the coefficient(s).

3. Is the average grade students receive in a class related to the proportion of students attending class on the day the FCQ was administered? Provide PRE, F*, and a statistical conclusion. You will not be able to get the coefficient from available information.

4. Why is the coefficient for A_GRADE different in models M1 and M2?

5. Do the demographic variables as a set predict instructor ratings? Provide PRE, F*, a statistical conclusion, and a brief substantive conclusion.

6. If we statistically equate the sections on the demographic variables, or if we control for the demographic variables, do average grades and proportion of forms returned, as a set, predict instructor ratings? Provide PRE, F*, a statistical conclusion, and a brief substantive conclusion.

7. If we first control for average grade and the proportion of forms returned, do the demographic variables as a set predict instructor ratings? Provide PRE, F*, a statistical conclusion, and a brief substantive conclusion. Also, indicate why you reach a different conclusion about the demographic variables in this question versus question #5 above.

8. Once we have controlled for the proportion of students returning forms, do the demographic variables as a set predict the average grades assigned in courses? Provide PRE, F*, a statistical conclusion, and a brief substantive conclusion. [This question is easy, involving just a few simple calculations, if you have a complete conceptual understanding of what multiple regression does and what all the numbers on the printout mean. However, I expect this question to be sufficiently difficult at this stage in the course, that only a couple of students will answer it correctly. So, don't waste too much time if you don't see how to do it.]

9. The department is deciding whether or not to toughen grading by returning the average grade to a C (2.0). In the discussion, one professor, based on the most complete model in this analysis, argues that doing so would cause the instructor ratings in the department to decrease by __. Fill

in the blank and, as a good data analyst, what would be your response to this statement at the faculty meeting?

10. The chair of a smaller department wants to try to replicate the results for the relationship between A_GRADE and R_INSTR after controlling for P_RETURN and the demographic variables. Over this same period, this department had 55 sections. Assuming the relationship is indeed about the same, approximately what are the chances that there will be a significant relationship between A_GRADE and R_INSTR in the full model for this smaller department?

Variable	Label	N	Minimum	Maximum
A_GRADE	Average grade assigned	547	1.8623047	4.0000000
P_RETURN	Percent of forms returned	547	0.3309353	1.0000000
SIZE	Class size (forms requested)	547	5.0000000	508.0000000
TTRACK	Tenure/Tenure-track	547	0	1.0000000
CRSLEV_N	Course level 1-4 (numeric)	547	1.0000000	4.0000000

Variable	Label	Mean	Variance
A_GRADE	Average grade assigned	2.9860675	0.1687751
P_RETURN	Percent of forms returned	0.7532189	0.0206493
SIZE	Class size (forms requested)	50.5429616	4451.09
TTRACK	Tenure/Tenure-track	0.5118830	0.2503164
CRSLEV_N	Course level 1-4 (numeric)	2.5100548	1.4481588

Model: M1
Dependent Variable: R_INSTR Instructor rating, item 12

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	1	38.13722	38.13722	122.256	0.0001
Error	545	170.01049	0.31195		
C Total	546	208.14771			

Root MSE	0.55852	R-square	0.1832
Dep Mean	3.17175	Adj R-sq	0.1817
C.V.	17.60923		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	1.250768	0.17536914	7.132	0.0001
A_GRADE	1	0.643316	0.05818207	11.057	0.0001

Model: M2
Dependent Variable: R_INSTR Instructor rating, item 12

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	46.35855	23.17928	77.938	0.0001
Error	544	161.78916	0.29741		
C Total	546	208.14771			

Root MSE	0.54535	R-square	0.2227
Dep Mean	3.17175	Adj R-sq	0.2199
C.V.	17.19396		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	0.876153	0.18546591	4.724	0.0001
A_GRADE	1	0.541110	0.06004381	9.012	0.0001
P_RETURN	1	0.902537	0.17166003	5.258	0.0001

Variable	DF	Type II SS	Squared Partial Corr Type II	Tolerance
INTERCEP	1	6.637159	.	.
A_GRADE	1	24.153790	0.12989893	0.89518571
P_RETURN	1	8.221331	0.04835778	0.89518571

Variable	DF	Variable Label
INTERCEP	1	Intercept
A_GRADE	1	Average grade assigned
P_RETURN	1	Percent of forms returned

Model: M3
Dependent Variable: R_INSTR Instructor rating, item 12

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	3	5.12806	1.70935	4.572	0.0036
Error	543	203.01966	0.37389		
C Total	546	208.14771			

Root MSE	0.61146	R-square	0.0246
Dep Mean	3.17175	Adj R-sq	0.0192
C.V.	19.27835		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	3.176516	0.06822799	46.557	0.0001
SIZE	1	-0.001220	0.00040440	-3.016	0.0027
TTRACK	1	-0.050430	0.05877262	-0.858	0.3912
CRSLEV_N	1	0.032944	0.02493290	1.321	0.1870

Variable	DF	Type II SS	Squared Partial Corr Type II	Tolerance
INTERCEP	1	810.429537	.	.
SIZE	1	3.400423	0.01647332	0.94073603
TTRACK	1	0.275275	0.00135407	0.79196524
CRSLEV_N	1	0.652731	0.00320481	0.76064842

Variable	DF	Variable Label
INTERCEP	1	Intercept
SIZE	1	Class size (forms requested)
TTRACK	1	Tenure/Tenure-track
CRSLEV_N	1	Course level 1-4 (numeric)

Model: M4
Dependent Variable: R_INSTR Instructor rating, item 12

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	5	46.75932	9.35186	31.349	0.0001
Error	541	161.38840	0.29831		
C Total	546	208.14771			

Root MSE	0.54618	R-square	0.2246
Dep Mean	3.17175	Adj R-sq	0.2175
C.V.	17.22020		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	0.801565	0.21070031	3.804	0.0002
A_GRADE	1	0.557229	0.06191271	9.000	0.0001
P_RETURN	1	0.962861	0.18537652	5.194	0.0001
SIZE	1	0.000307	0.00039537	0.777	0.4376
TTRACK	1	0.010467	0.05276357	0.198	0.8428
CRSLEV_N	1	-0.015882	0.02274366	-0.698	0.4853

Variable	DF	Type II SS	Squared Partial Corr Type II	Tolerance
INTERCEP	1	4.317394	.	.
A_GRADE	1	24.164809	0.13023116	0.84452878
P_RETURN	1	8.048089	0.04749915	0.76995720
SIZE	1	0.180058	0.00111444	0.78524315
TTRACK	1	0.011740	0.00007274	0.78401567
CRSLEV_N	1	0.145471	0.00090056	0.72936603

Variable	DF	Variable Label
INTERCEP	1	Intercept
A_GRADE	1	Average grade assigned
P_RETURN	1	Percent of forms returned
SIZE	1	Class size (forms requested)
TTRACK	1	Tenure/Tenure-track
CRSLEV_N	1	Course level 1-4 (numeric)