Sociology 593
Exam 2
April 4, 1997

I. True-False. (20 points) Indicate whether the following statements are true or false. If false, briefly explain why.

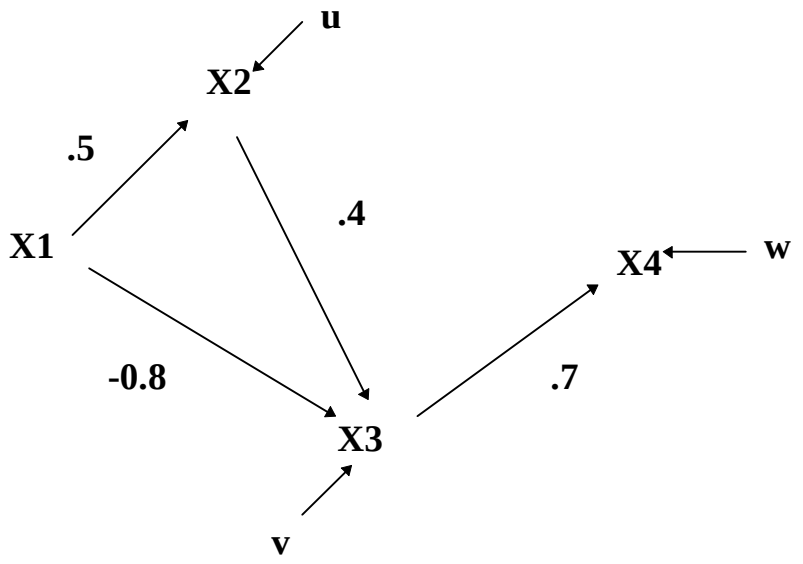
1. A researcher computes SUMX1X2 = X1 + X2. She regresses Y on X1 and SUMX1X2. If the coefficient for X1 is insignificant, she should conclude that X1 does not affect Y.

2. A researcher regresses Y on X1, X2, and X3. If the effect of X1 is 0, she can safely conclude that changes in X1 will have no effect on Y.

3. A researcher has inadvertently omitted an important variable from her model. Fortunately, as the sample size gets larger and larger, the omitted variable bias will diminish and eventually disappear.

4. In a piecewise regression model, the regression line need not be continuous

II. Path Analysis/Model specification. (30 points). A sociologist believes that the following model describes the relationships between X1, X2, X3 and X4. All variables are in standardized form. The hypothesized value of each path is included in the diagram.



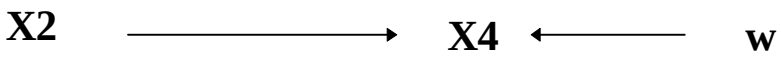
a. Write out the structural equation for each endogenous variable.

b. Determine the complete correlation matrix. (Remember, variables are standardized. You can use either normal equations or Sewell Wright, but you might want to use both as a double-check.)

c. Decompose the correlation between X2 and X3 into

- Correlation due to direct effects
- Correlation due to common causes

d. Suppose the above model is correct, but instead the researcher believed in and estimated the following model:

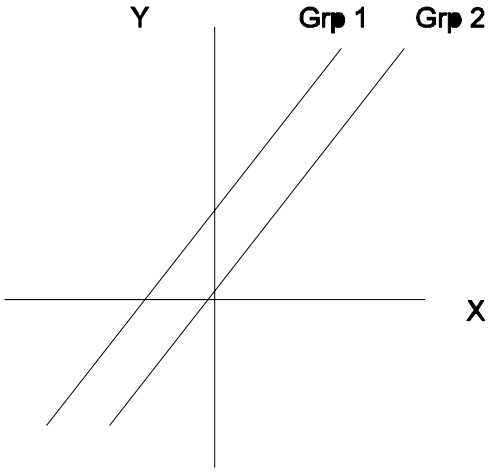


What conclusions would the researcher likely draw? In particular, what would the researcher conclude about the effect of changes in X2 on X4? Why would he make these mistakes? Discuss the consequences of this mis-specification.

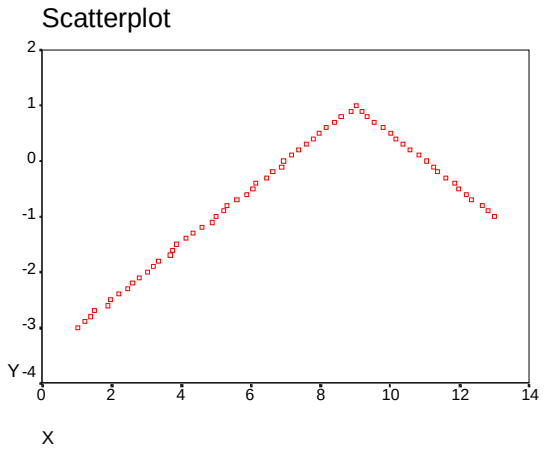
III. *Short answer.* Answer *two* of the following three questions. (25 points each; up to 10 points extra credit if you do all 3).

1. Examine the following plots. What OLS assumptions, if any (e.g., additivity, linearity), would be violated if you simply regressed Y on X? What harm would result, e.g. would values be over-estimated, under-estimated, or what? Indicate the model you think should be estimated, e.g. $E(Y) = \alpha + \beta_1 X + \beta_2 X^2$.

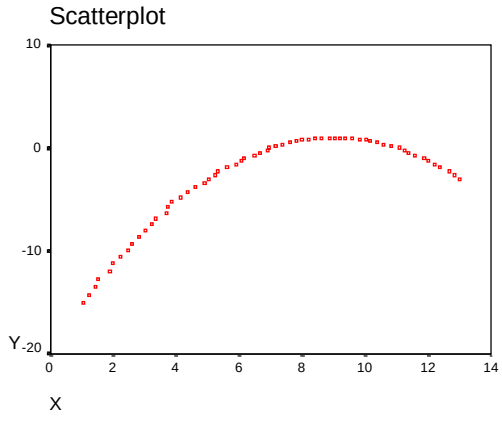
a.



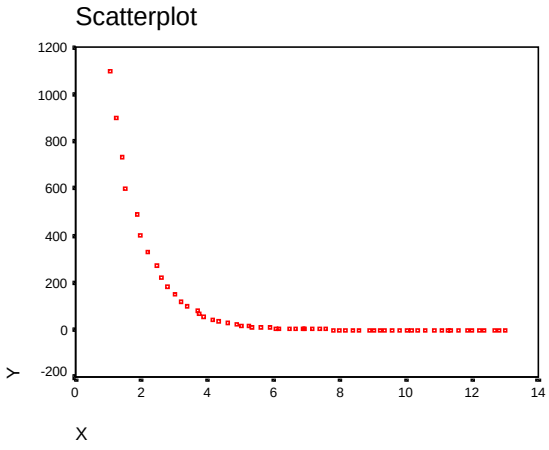
b.



c.



d.



2. A state has adopted a program to help welfare recipients get jobs. Much to its surprise, a recent study has shown that welfare recipients who participate in these programs actually do less well than recipients who do not participate. Participants are less likely to get jobs, when they do get jobs they aren't as good, and are less likely to get off welfare. Some critics are therefore arguing the programs do more harm than good and should be abolished, while others continue to maintain that the programs are beneficial. The state has hired you, a professionally trained social scientist, to give it insight on why these relationships exist. Drawing on your knowledge of the logic of causal order, present different models that could account for the observed relationships. Indicate what implications the different models have for policies that should be adopted. To be fair, you will want to present one or more models that suggest that the job programs are good, one or more models which imply the programs are not good, and one or two models which suggest that the programs are not good but the problems are correctable (i.e. you don't have to abolish the programs to solve the problem). When presenting your answer, keep in mind that the state officials do not know very much about the logic of causal order, so you will have to make things very clear for them.

3. A pollster has constructed a scale which measures support for her candidate (LIKECND), where high values indicate strong support and low values indicate strong opposition. She has also measured years of education (EDUC). She wants to see whether different models apply to men and women. She therefore runs the following regressions. Indicate whether there appear to be statistically significant differences between men and women. If so, tell whether these differences are limited to differences in the intercepts, or whether the effect of education differs between men and women. Briefly discuss the substantive implications of what you think is the best model.

```
Compute FemEd = Female * Educ.
REGRESSION
  /STATISTICS COEFF R ANOVA CHA
  /DEPENDENT LikeDem  /METHOD=ENTER Educ
  /METHOD=ENTER Female  /METHOD=ENTER FemEd  .
```

Regression

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	EDUC[a]	.	Enter
2	FEMALE[a]	.	Enter
3	FEMED[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: LIKECND

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.310[a]	.096	.093	11.2264	.096	30.725	1	288	.000
2	.775[b]	.601	.598	7.4744	.504	362.711	1	287	.000
3	.824[c]	.680	.676	6.7064	.079	70.495	1	286	.000

a. Predictors: (Constant), EDUC

b. Predictors: (Constant), EDUC, FEMALE

c. Predictors: (Constant), EDUC, FEMALE, FEMED

ANOVA[d]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3872.327	1	3872.327	30.725	.000[a]
	Residual	36297.54	288	126.033		
	Total	40169.87	289			
2	Regression	24135.98	2	12067.99	216.012	.000[b]
	Residual	16033.88	287	55.867		
	Total	40169.87	289			
3	Regression	27306.61	3	9102.204	202.377	.000[c]
	Residual	12863.26	286	44.976		
	Total	40169.87	289			

a. Predictors: (Constant), EDUC

b. Predictors: (Constant), EDUC, FEMALE

c. Predictors: (Constant), EDUC, FEMALE, FEMED

d. Dependent Variable: LIKECND

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	26.633	1.123		23.723	.000
	EDUC	-7.688	1.387	-.310	-5.543	.000
2	(Constant)	25.654	.749		34.241	.000
	EDUC	-5.730	.929	-.231	-6.168	.000
	FEMALE	2.447	.128	.715	19.045	.000
3	(Constant)	26.030	.674		38.636	.000
	EDUC	-5.708	.834	-.231	-6.847	.000
	FEMALE	1.506	.161	.440	9.373	.000
	FEMED	1.937	.231	.393	8.396	.000

a. Dependent Variable: LIKECND