

Sociology 593
Exam 2
March 26, 1999

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

1. A researcher regresses Income on Education. She does NOT include any dummy variables or interaction terms. One implication of this model is that, if it is true, the mean income for blacks will be the same as the mean income for whites.

2. A researcher believes that job experience has a stronger effect on earnings for men than it does women. If she is correct, this means that the effect of job experience on earnings is non-linear.

3. A researcher has included several extraneous variables in her model. The larger her sample, the less problematic this will be.

4. A researcher hypothesizes that IQ positively affects the political liberalism of women but negatively affects the political liberalism of men. She gets

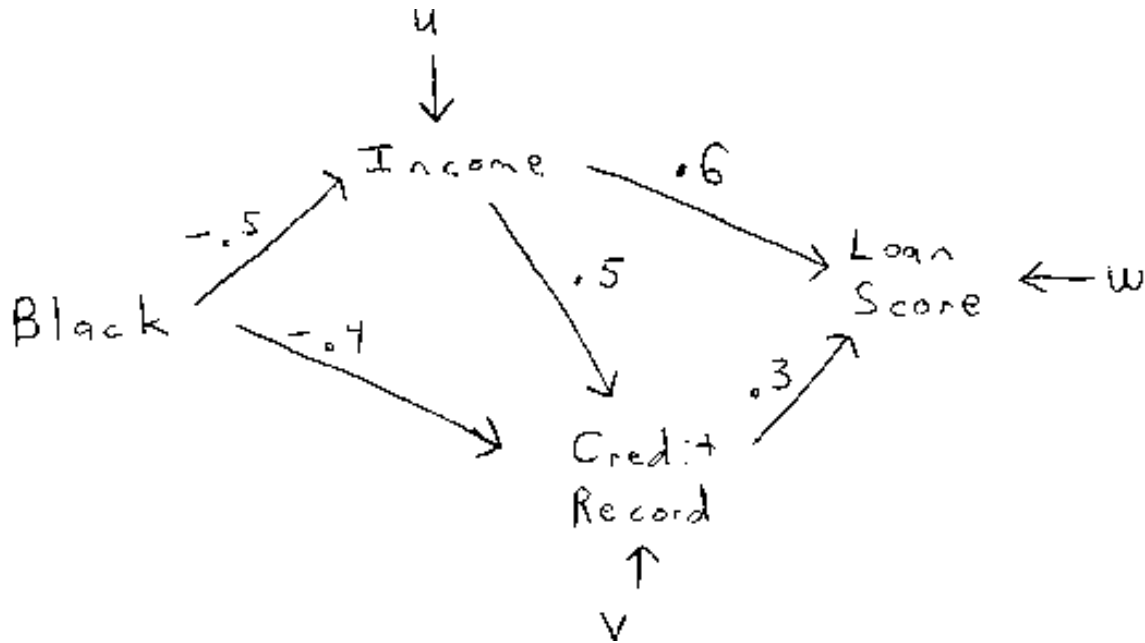
$$\beta_{IQ} = 6$$

$$\beta_{Male} = 0$$

$$\beta_{IQ*Male} = -4$$

Male = 1 if male, 0 if female. The standard error for the interaction term is 1.0. N = 2,000. The evidence supports the researcher's hypothesis.

II. *Path Analysis/Model specification.* (30 points). A researcher is interested in home mortgage lending. She has measures of loan applicant's race, income, credit record (where the higher the score, the better the person's record is) and the Loan Score assigned by the lender (where the higher the score assigned, the stronger the lender thought the application was). All her 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 Income and Loan Score into
 - Correlation due to direct effects
 - Correlation due to indirect effects
 - Correlation due to common causes
- d. Many studies have shown that blacks fair less well in the home mortgage market than would seem justified based on their income alone. That is, even when a black and a white have identical incomes, the black is less likely than the white to get the loan. According to the above model, why does this occur? Would you interpret this model as saying there is racial discrimination against blacks? Why or why not?

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. Draw a scatterplot that illustrates each of the following. Offer a hypothetical example of when you might observe such a relationship between variables. 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. The effect of X on Y is positive for small values of X but then switches to being negative.
- b. The effect of X on Y is the same for both whites and blacks, but the intercepts are different in the two populations.
- c. Each one unit increase in X produces a 5% increase in Y.

2. The American Cancer Institute is concerned about the recent rise in teenage smoking. It has come up with an anti-smoking video that parents can order for their smoking teenagers. Much to its surprise, its research shows that teen smokers who see the video are less likely to quit than teen smokers who do not. Some critics within the Institute maintain that the video is doing more harm than good and should be abandoned, while others continue to insist that the video is serving its purpose. The Institute 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 what should be done about the video. To be fair, you will want to present one or more models that suggest that the video discourages smoking, one or more models which imply the video encourages smoking, and one or two models which suggest that the video is not achieving what the Institute wants but the problems are correctable (i.e. you don't have to completely scrap the video to solve the problem). When presenting your answer, keep in mind that Institute staff do not know very much about the logic of causal order, so you will have to make things very clear for them.

3. A researcher is interested in how the effect of a message differs by characteristics of the speaker. College student subjects are randomly assigned to read passages whose content ranges from very conservative (MSGLIB = -100) to very liberal (MSGLIB = 100). Half the students are told that the author of the passage is a first year college student (PROF = 0), while the other half are told that the author is a distinguished Nobel Prize winning professor (PROF = 1). Students then complete a questionnaire that measures their own liberalism and conservatism (LIBRLISM = -100 if very conservative, 100 if very liberal). The researcher then runs the following regressions.

Indicate whether there appear to be statistically significant differences in the determinants of liberalism between those who thought the author was a college student and those who thought the author was a professor. If so, tell whether these differences are limited to differences in the intercepts, or whether the effect of the passages differs between the two types of speakers. If differences are found, be specific as to what they are, e.g. how much greater (or weaker) is the effect of messages on beliefs when the author is a professor rather than a student. Briefly discuss the substantive implications of what you think is the best model. Be sure to indicate how the printout supports your arguments.

Compute profmsg = prof * msglib

Regression

Descriptive Statistics

	Mean	Std. Deviation	N
LIBRLISM	23.3895	20.1731	200
MSGLIB	44.8919	28.0179	200
PROF	.5000	.5013	200
PROFMSG	22.4459	29.9809	200

Correlations

		LIBRLISM	MSGLIB	PROF	PROFMSG
Pearson Correlation	LIBRLISM	1.000	.746	.429	.772
	MSGLIB	.746	1.000	.000	.467
	PROF	.429	.000	1.000	.751
	PROFMSG	.772	.467	.751	1.000
Sig. (1-tailed)	LIBRLISM	.	.000	.000	.000
	MSGLIB	.000	.	.500	.000
	PROF	.000	.500	.	.000
	PROFMSG	.000	.000	.000	.
N	LIBRLISM	200	200	200	200
	MSGLIB	200	200	200	200
	PROF	200	200	200	200
	PROFMSG	200	200	200	200

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	MSGLIB ^a	.	Enter
2	PROF ^a	.	Enter
3	PROFMSG ^a	.	Enter

a. All requested variables entered.

b. Dependent Variable: LIBRLISM

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	.746 ^a	.556	.554	13.4766	.556	247.905	1	198	.000
2	.860 ^b	.740	.737	10.3405	.184	139.309	1	197	.000
3	.887 ^c	.787	.784	9.3751	.047	43.664	1	196	.000

- a. Predictors: (Constant), MSGLIB
- b. Predictors: (Constant), MSGLIB, PROF
- c. Predictors: (Constant), MSGLIB, PROF, PROFMSG

ANOVA^d

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	45023.863	1	45023.863	247.905	.000 ^a
	Residual	35960.279	198	181.618		
	Total	80984.141	199			
2	Regression	59919.641	2	29959.821	280.191	.000 ^b
	Residual	21064.500	197	106.926		
	Total	80984.141	199			
3	Regression	63757.366	3	21252.455	241.803	.000 ^c
	Residual	17226.776	196	87.892		
	Total	80984.141	199			

- a. Predictors: (Constant), MSGLIB
- b. Predictors: (Constant), MSGLIB, PROF
- c. Predictors: (Constant), MSGLIB, PROF, PROFMSG
- d. Dependent Variable: LIBRLISM

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.711	1.803		-.394	.694
	MSGLIB	.537	.034	.746	15.745	.000
2	(Constant)	-9.341	1.565		-5.970	.000
	MSGLIB	.537	.026	.746	20.520	.000
	PROF	17.260	1.462	.429	11.803	.000
3	(Constant)	-2.305	1.774		-1.299	.195
	MSGLIB	.380	.034	.528	11.332	.000
	PROF	3.188	2.509	.079	1.271	.205
	PROFMSG	.313	.047	.466	6.608	.000

- a. Dependent Variable: LIBRLISM