Department of Sociology
American University
Comprehensive Exam in Social Research Methods
December 1, 2000

This exam has four parts. You must answer one question in each part for a total of four questions. Your answers should: (1) demonstrate the full range and depth of your understanding of the relevant research techniques, (2) refer to appropriate academic studies and or research traditions, and (3) minimize the degree of repetition in your responses. Be sure to budget your time; please *double-space* the document or write on every *other* line of the paper.

Part I. Answer one question

1. You are an urban sociologist and have been doing participant observation for several months studying the inner-city youth subculture of "graffiti crews". Among other things, you are exploring how the production of graffiti operates in the continual attempt by often-times excluded urban groups – young minority men, for example – to gain a sense of identity and community. While not glorifying graffiti, you have begun to see it as a social practice for resisting marginalization and the denial of human dignity. However, the city has recently adopted a "zero tolerance" policy regarding graffiti; graffiti writers have, thus, been criminalized.

Consider the following excerpts from the Ethical Standards of the American Sociological Association. Emphasis (italics) has been added.

> Part (b) of section **11.01 Maintaining Confidentiality** states: Confidential information provided by research participants . . . is treated as such by sociologists *even if there is no legal protection or privilege* to do so.
> Part (c) of section **11.02 Limits of Confidentiality** states: Confidentiality is not required with respect to observations in public places, *activities conducted in public*, or other settings where no rules of privacy are provided by law or custom.

Explain the contradiction embodied in these two statements. If confronted by the authorities, which will you follow? Why? How might you (re)design your research such that both you and the graffiti crews are protected?

or

2. You are about to embark on a cross-national study of attitudes toward the elderly. How would you use Denzin's four types of triangulation -- Data triangulation, Investigator triangulation, Theory triangulation, Methodological triangulation – in your research design? Why is triangulation particularly important for cross-national analysis?

Part II Answer one question.

1. Consider the folk-saying "Absence makes the heart grow fonder." State a sociological research question based on this folk wisdom. How would you operationlize the main concepts in this statement in order to gather empirical evidence concerning its basis in social reality? How would you assess both the reliability and the validity of the measures you developed?

or

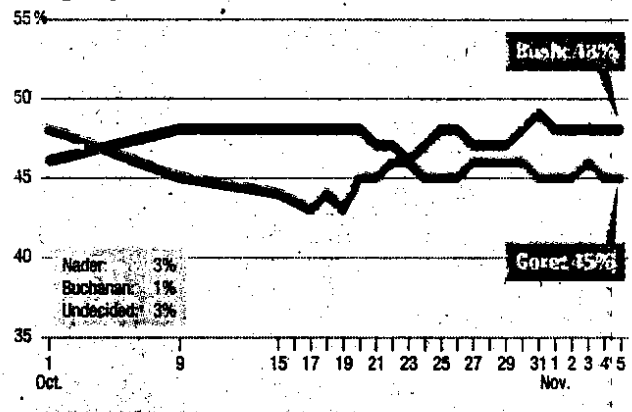
2. The box below presents findings from four tracking polls prior to the recent presidential election. The story that reported these results was entitled "Race Still Too Close to Call." The fine print at the bottom of the box reads as follows:

The Washington Post tracking poll was conducted daily until Election Day. The latest tracking result represents the four-night average of data collected Nov 2-5 among 2,429 self- identified likely voters nationwide. The Post and ABC news collect data jointly but use somewhat different models to identify likely voters. This may produce slightly different estimates of candidate support. The margin of sampling error for overall results is plus or minus 2 percentage points and slightly larger for subsamples. Sampling error is only one of many potential sources of error in this or any other public opinion poll. Interviewing was conducted by TNS Intersearch of Horsham, PA.

Washington Post Daily Tracking Poll

Q: If the election were being held today, whom would you vote for?

Among likely voters



Results from other polls yesterday:	Gore	Bush
MSNBC/Reuters	46%	47%
CNN/USA Today/Gallup	45%	47%
ABC News	45%	48%

Please explain why the conclusion is that the race is too close to call when all the results show Bush in the lead. Give an example of a possible subsample; why would estimates for the opinions in subsamples be less accurate than the overall sample? Explain how different models for identifying likely voters might have impact on the results. The reader is warned about other possible sources of error– what might they be? In the body of the story, results are given from a fifth national poll conducted by Zogby International showing Gore in the lead, 48% to 46%. How do you account for this reversal?

Part III Answer one question.

1. You have been asked to be the keynote speaker at the 2001 ASA conference. The topic you have decided to address is the following: "Qualitative versus Quantitative: Is it necessary to choose?" Write the speech, being sure to include the underlying assumptions and theoretical foundations of each methodological orientation, a description of each and its derivative methods, and provide examples of how each has been/can be used. Of course, be sure to also take a position on the issue, and argue your position.

or

2. Fully discuss the implications for the research process that are raised by the following statement:

Theory, Albert Einstein once said, determines what we can observe. Louis Althusser, a contemporary French Marxist philosopher, slightly transformed this proposition by saying that theory also determines what we cannot observe.

Part IV Answer one question

1. According to a story published in the Washington Post (A37 Thursday, November 9, 2000), exit polls – polls of people leaving the voting place – conducted on election day showed the following differences among Democrat and Republican voters:

Seventy two percent of Republican voters are married, compared with 59 percent married among Democrats.
A solid majority of Republican voters have guns at home, 58 percent, while an even larger majority of Democrats, 62 percent, does not.
A quarter of Republicans consider themselves members of the religious right, while 92 percent of Democrats do not.
An overwhelming majority of Democrats, 71 percent, believe abortions should be legal all or most of the time, while 56 percent of Republicans think abortions should always or most often be illegal.
Similar proportions of Democrats and Republicans live in suburbs. But the ratio of Democratic city dwellers to rural and small town residents is 37-21, while for Republicans the ratio is reversed, 21-34.

Answer the following questions:

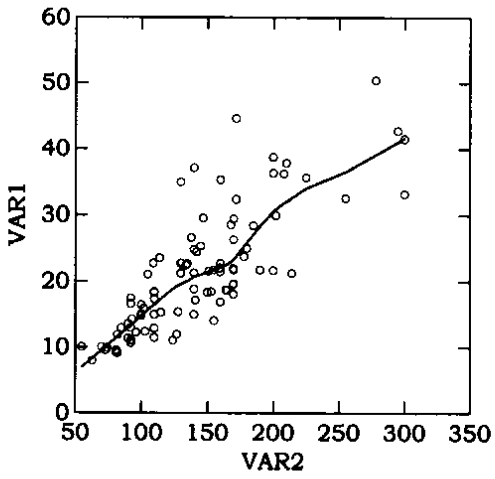
a. What do these bivariate relations tell you about the influences on voting behavior?
b. What further information and what statistical test(s) would you want performed on these findings in order to feel confident in your conclusions?

c. How might one of the factual variables reported on here (marital status, place of residence) influence one of the attitude variables (identification with religious right, circumstances under which abortion should be legal)? Make a hypothetical table showing that relationship.
d. Identify a third variable (not mentioned here) that might be important in the relationship between opinions about legal abortions and voting Democrat or Republican. Make a hypothetical table showing that multivariate relationship.

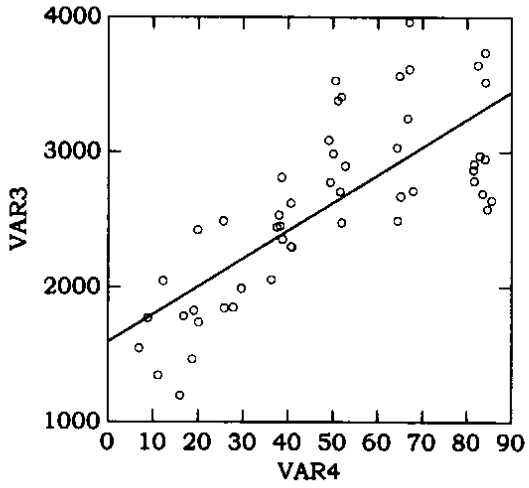
or

2. Answer the following questions regarding regression:

a. Modern data analysis emphasizes examination of residuals from regression models. Explain why.
b. The plot below contains a pathology. Identify the pathology, explain why it is bad, and describe how you would fix it. The line is a lowess smoother.



c. The plot below also contains a pathology. Identify the pathology, explain why it is bad, and describe how you would fix it. The line is a linear smoother.



d. From a dataset about *automobiles* we wish to establish a relationship between the weight of each car (units: lbs.), the suggested dealer price (units: $1000s), turning radius (units: meters), and the luggage capacity (units: liters). Interpret the regression output below.

```
Dep Var: WEIGHT   N: 93   Multiple R: 0.743   Squared multiple R: 0.553

Adjusted squared multiple R: 0.543   Standard error of estimate: 398.910

Effect          Coefficient     Std Error      Std Coef Tolerance      t    P(2 Tail)

CONSTANT           1880.401       121.526         0.000         .     15.473     0.000
PRICE                20.616         5.647         0.124     0.946      3.651     0.011
TURN                 -0.217         0.139         0.222     0.897     -1.591     0.099
LUGGAGE               7.284         1.189         0.647     0.940      6.125     0.000

                              Analysis of Variance

Source             Sum-of-Squares   df  Mean-Square     F-ratio       P

Regression          1.76923E+07      3  8846159.471      55.591       0.000
Residual            1.43216E+07     89   159129.413
-------------------------------------------------------------------------------

Durbin-Watson D Statistic      1.940
First Order Autocorrelation    0.027
```

e. Below is the residuals plot from the regression above. Interpret the plot.

Plot of Residuals against Predicted Values