

Assignment #10Due Wednesday 6/30/10 by 6 p.m. in the Econ 300/QAC201 slot in the Economics Alcove

Please show the calculations used to arrive at your answers. Round final answers to the second decimal place if necessary.

- A. A multiple regression was computed using a sample of 60 months, relating the average air quality index in Middletown ( $Y$ ), measured on a hundred-point index, to the average number of cars driven down Main Street ( $X_1$ ), measured in thousands, and the number of days in the month with high winds ( $X_2$ ):

$$Y = 25.95 + 0.46X_1 - 2.13X_2$$

$$(5.96) \quad (0.14) \quad (0.78)$$

The standard errors for the coefficients are in parentheses below the coefficients.

- 1) Calculate the 95% confidence intervals, t ratios, and two-tailed p-values for each coefficient.
- 2) Other things being equal, what would you estimate is the effect of 25,000 cars per day being driven down Main Street?
- 3) Other things being equal, what would you estimate is the effect of having one more windy day in a month?
- 4) Recall from Assignment #9 that the simple regression of  $Y$  on  $X_1$  yielded:

$$Y = 14.85 + 0.43X_1$$

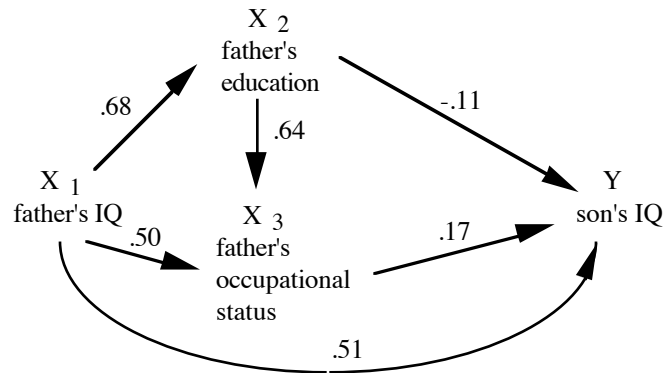
$$(4.58) \quad (0.15)$$

Discuss the pros and cons of using this simpler formulation to predict air quality in place of the multiple regression equation presented above.

- B. Suppose your significant other is a bright person, but he/she only read through Ch. 12 of Wonnacott and Wonnacott when he/she took probability and statistics, and therefore has no notion of the concept of multiple regression. He/she is now writing a paper attempting to explain the U.S. violent crime rate in which he/she has run 100 simple regressions of the crime rate in each county in the U.S. on 100 other variables, taken each in turn. These variables were chosen simply because it was possible to get data measured on a county level for these variables.

- 1) How many significant variables do you think he/she will uncover in this analysis?
- 2) Assuming he/she uncovers 10 statistically discernable variables, what would you now suggest that he/she do next to improve upon this analysis?
- 3) Suppose he/she finds that population density is positively related to the crime rate and decides to draw the policy conclusion that it is better to maintain a low population density in each county. What might you (politely) say in response to that prescription?

C. The path diagram below is from a classic work on intergenerational transmission of inequality (C. Jencks, Inequality, 1972, p. 340):

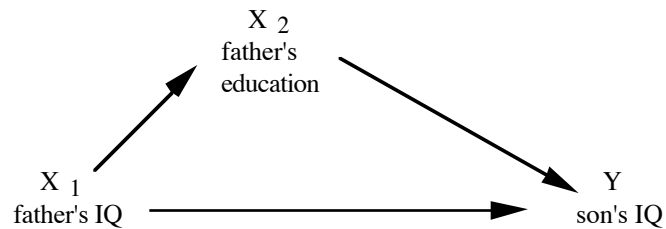


1) Calculate the total effect on Y of each of the following:

- The total effect of  $X_3$
- The total effect of  $X_2$
- The total effect of  $X_1$

2) Compare two people, A and B. How much of an effect on A's IQ would you expect if A and B have identical backgrounds, except A's father has one more year of education?

3) Suppose the researchers doing this study had not measured  $X_3$ , father's occupational status. Then the path diagram would look like this:



Suppose that otherwise the data were exactly the same. What would the path coefficients be?