

# The University of Illinois at Chicago

## Economics 346: Econometrics

### Review for Final Examination

1. Why are the errors in cross-section studies unlikely to be serially correlated?
2. A somewhat naive researcher attempts to estimate an aggregate consumption function for the U.S. by regressing a consumption variable  $C$  on disposable income  $Y$  and savings  $S$ . The model is  $C = \beta_1 + \beta_2 Y + \beta_3 S + \varepsilon$ . How good a fit is this researcher likely to get when this equation is run? (Notice that  $C=Y-S$  identically for all observations.)
3. Your regression of agricultural production had a high  $R^2$  (.98), an  $F$  statistic that was significant at the 1% level, but the coefficient on machinery was negative and insignificant. What is the problem with this model?
4. True or false and explain: The problems from including irrelevant variables in a regression model are more serious than the problems from omitting variables that should be included.
6. True or false and explain: The  $R^2$  values of two models, one involving a regression in the first-difference form and the other in the level form, are not directly comparable.
7. Given a sample of 50 observations and 4 explanatory variables, what can you say about autocorrelation if
  - a.  $d = .105$ ?
  - b.  $d = 3.97$ ?
8. You are soon to be interviewed for the position of political analyst for Channel 2. Since this is an election year, all TV stations will devote a great deal of time to election coverage, and your worldly knowledge will not be enough to land you this position. Consequently, you should use a regression model to back your opinions regarding the 1994 Congressional elections. To fill up air time, you will need 3 models. All the models will try to explain differences across states in the percentage of votes received by Democratic candidates for Congressional seats among all votes cast for House of Representatives candidates in each state. Therefore, the dependent variable has 50 observations, one for each state. You have 4 types of explanatory variables:
  1. The unemployment rate in each state
  2. Regional dummy variables for whether the state is in the Northeast, South, Midwest, or West
  3. A dummy variable for whether Bill Clinton appeared in that state to campaign for Congressional candidates.

4. Interaction terms between the regional dummies and the Clinton dummy.

Your 3 models differ only in the explanatory variables they contain:

Model I contains variables (1) and (2).

Model II contains variables (1), (2), and (3).

Model III contains variables (1) and (4).

- a. Write out each model in a regression equation. This can be done in more than one way; pick any formulation you like. Be sure to define all notation, and describe how the explanatory variables will show the effects of interest, by interpreting the effects to be estimated.
- b. Using the variables in these models, how would you test the following hypotheses? (If you propose an F-test, give the restricted and unrestricted regression equations.)
  - i. Clinton's appearance doesn't matter.
  - ii. The Northeast and Midwest (the "Frostbelt") voted uniformly.
  - iii. Regionalism is insignificant: the entire country voted uniformly, with no regional differences.
  - iv. Clinton's appearance had the same effect in all regions.