

BOSTON COLLEGE
Department of Economics

EC 228 01
Econometric Methods
Fall 2006
Prof. Baum

Final Exam
12 December 2006

Answer all questions. Total of 126 points. Partial credit given for partial answers. Exam ends at 3:00 PM sharp.

1. (25 pts) Briefly explain each term. Use examples to illustrate your explanation.
 - a. Breusch–Godfrey test
 - b. Newey–West standard errors
 - c. RESET test
 - d. Breusch–Pagan test
 - e. joint hypothesis test

2. (25 pts) True, False, Explain. Indicate clearly whether each of the following statements is **true** or **false**, and **explain** your answer. No credit without explanation.
 - a. Including variables in a regression equation that are not part of the population model yields unreliable, biased and inconsistent estimates of the true parameters.
 - b. Heteroskedasticity in a cross-section regression is a more serious problem than autocorrelation in a time-series regression.
 - c. In the presence of correlation between regressor and error, instrumental variables estimates (unlike OLS estimates) will be unbiased and consistent.
 - d. Dummy variables may be used to evaluate the effects of structural change in a time-series regression model.
 - e. First-order autocorrelation in the errors of a time-series regression with a lagged dependent variable leads to inconsistent parameter estimates.

3. (21 pts) Write an essay on the theory of instrumental variables estimation. Describe the conditions that make a variable an appropriate instrument; the concepts of the order condition for identification and overidentification; and explain why one should use IV estimation only when it is necessary to generate consistent parameter estimates.
4. (15 pts) In the NLSW88 dataset, 2,246 individuals are classified by **race** (white/black/other, coded 1,2,3), **industry** (5 classifications, coded 1,2,...,5) and whether they are a **union** member (0/1).
- Specify how you would test the hypothesis that these workers' **wage** is influenced by their **race** and **union** status. Indicate the transformations you would apply to the data, the regression equation you would use, and the tests you would apply after that equation.
 - The model of part (a) assumes that **race** and **union** status have independent effects on **wage**. Indicate how you would test for non-independent effects of those factors on **wage**.
 - Returning to the model of part (a): how would you include the effects of **industry** on **wage**? Write down the equation you would use, and indicate the test that you would apply to consider the importance of **industry** on **wage**.
5. (16 pts) We have quarterly timeseries data on New England employment for 1990–2000. Write down the models you would use to test each of the following hypotheses:
- The employment series contains a significant trend.
 - Employment have been growing at 1.5 per cent per annum.
 - The employment series contains significant seasonality.
 - New England employment is, on average, 10% higher in the 4th quarter than in the 1st quarter of the year.

6. (24 pts) The following models of $narr86$, the number of times a sample of men were arrested in 1986, yields:

	(1)	(2)	(3)
	$narr86$	$narr86$	$narr86$
$pcnv$	-0.150 (3.67)**	-0.151 (3.69)**	-0.130 (3.22)**
$ptime86$	-0.034 (4.01)**	-0.037 (4.25)**	-0.041 (4.72)**
$qemp86$	-0.104 (10.02)**	-0.103 (9.94)**	-0.095 (9.13)**
$avgsen$		0.007 (1.57)	0.004 (0.80)
$black$			0.338 (7.43)**
$hispan$			0.203 (5.10)**
Constant	0.712 (21.56)**	0.707 (21.32)**	0.585 (16.25)**
Observations	2725	2725	2725
Adjusted R^2	0.04	0.04	0.06
F-test: $B=H=0$			33.59
$Pr > F$			0.00

Absolute value of t statistics in parentheses

* significant at 5%; ** significant at 1%

$pcnv$ measures the proportion of arrests prior to 1986 that led to convictions, $avgsen$ is the average sentence length served for prior convictions, $ptime86$ is months spent in prison in 1986, $qemp86$ is quarters during which the man was employed in 1986, $black$ and $hispan$ are “dummy variables” equal to one if the individual is, respectively, Black or Hispanic.

- a. In the model in column (1), do the coefficients have the expected signs? Explain how you would interpret each of the coefficients, and whether they make sense in this context.
- b. Compare the models in columns (1) and (2). Which do you prefer, and why? How do you interpret the coefficient on *avgsen*?
- c. Does the model of column (3) improve upon the model of column (1)? On what basis do you make that judgment?
- d. How do you interpret the *F*-test for the coefficients of *black* and *hispan*? What conclusion do you draw from these coefficients about the likelihood of being arrested?

Happy Holidays!