

Problem Set 3

1. Table 1 presents the estimated coefficients and related statistics for a number of alternative models on poverty rates. The dependent variable is the poverty rate.

a. Explain why each of the variables might belong in the model. What signs would you expect a priori for each of the variables in Model A? Do the observed signs agree with your intuition? If not, do you have a rationalization for the unexpected sign?

b. In each model, apply the Wald F test to test the joint significance of the variables in the model. Be sure to state the null and alternative hypotheses in each case and the degrees of freedom for the F-statistic. What do you conclude about the models?

c. Apply the t-test to each of the regression coefficients in each of the models (be sure to indicate the degrees of freedom) and test whether the regression coefficients are significant. Be sure to state the null and alternative hypotheses.

d. Consider Model A as the unrestricted model and Model C as the restricted model. Use these two models to perform a relevant test. State the null and alternative hypotheses, the appropriate test statistic, its distribution under the null hypothesis and its degrees of freedom. Based on your test, would you recommend that certain variables be dropped from Model A? If yes, what are they and why? If not, why not?

e. Which of the four models do you consider the best? Explain your reasons.

f. In Model D, suppose you want to test the hypothesis that the sum of the regression coefficients for EDUC2 and EDUC3 is zero. For each of the three methods described in the chapter, show how you will carry out a test for this hypothesis. More specifically, state what regressions, if any, you would run; what statistic you would compute; its distribution under the null hypothesis, including the degrees of freedom; and the criteria for accepting/rejecting the hypothesis.

g. Are there any missing variables? Justify your answer.

2. In a study of early retirements, the following equations were estimated using 1980 census data for 44 States (values in parentheses are t-statistics):

$$\begin{aligned}
 \widehat{RETRD} &= -3,930 + 1.627HLTH - 0.0005MSSEC + 0.0005 MPUBAS \\
 &\quad (-0.41) \quad (5.4) \quad (-0.26) \quad (0.28) \\
 &\quad + 0.549 + UNEMP + 0.153 DEP + 0.077 RACE \\
 &\quad (2.2) \quad (1.6) \quad (2.3)
 \end{aligned}$$

$$\bar{R}^2 = 0.654 \quad \hat{\sigma} = 2.175 \quad ESS = 175.088$$

$$\begin{aligned}
 \widehat{RETRD} &= 5.093 + 1.596 HLTH + 0.557 UNEMP + 0.153 DEP \\
 &\quad (-1.6) \quad (6.5) \quad (2.3) \quad (1.8) \\
 &\quad + 0.083 RACE \\
 &\quad (3.5)
 \end{aligned}$$

$$\bar{R}^2 = 0.671 \quad \hat{\sigma} = 2.121 \quad ESS = 175.524$$

where

RETRD = returned men who are between the ages of 16 and 65
 HLTH = percent of people between 16 and 64 who are prevented from working due to disability
 MSSEC = mean social security income (\$)
 MPUBAS = mean public assistance income (\$)
 UNEMP = unemployment rate (in percent)
 DEP = percent of households that represent married couples with children under 18
 RACE = percent of men who are nonwhite

Test both models for overall significance (you have enough information to answer the question). Also test each regression coefficient (except the constant term) for significance at the 5 percent and 10 percent levels. Next test the hypothesis that the coefficients for MSSEC and MPUBAS are both jointly insignificant. State your null and alternative hypotheses, test statistic, distribution and d.f., and the criterion to accept or reject the null. What do you conclude? State the signs you would expect for the coefficients in Model B and whether the actual signs agree with your intuition. Finally, comment on the adequacy of the model.

3. For each of the following models, derive the elasticities of Y with respect to X:

$$\ln Y = \alpha + \beta X + \text{error}$$

$$Y = \alpha + \beta X + \gamma \ln X + \text{error}$$

$$Y = \alpha + \beta X + \gamma X^2 + \text{error}$$

$$Y = \alpha + \beta X + \gamma XZ + \text{error}$$

$$Y = \alpha + \beta/X + \text{error}$$

4. Assume that: $Y_t = a + bX_{1t} + cX_{2t} + u_t$ is your regression model of interest, where X_{1t} and X_{2t} are explanatory variables, and u_t is the error term. If I know that $b = c$, how can I obtain estimates of a , b , and c ? If I know that $b = 1$, how can I obtain estimates of a and c ?

5. Define the growth rate of some economic time series, say Y_t . How do I construct the growth rate of this variable, in practice? List 3 reasons why growth rates are often used for examining time series data.

6. Discuss briefly the issue of nonstationarity in time series analysis.

POVRATE = % of families with income below poverty line
 URB = % of urban population
 FAMSIZE = average number of family members
 EDUC1 = % of population 25 years old + that completed only 8 years of education
 EDUC2 = % of population 25 years old + that graduated from high school
 EDUC3 = % of population 25 years old + that completed 4 years of college
 UNEMP = unemployment rate among persons 16 years old and older
 MEDINC = median family income in thousands of dollars

TABLE 1 Estimated Models for Poverty Rate

| Variable | Model A | Model B | Model C | Model D |
|-----------------------------|--------------------|--------------------|--------------------|--------------------|
| CONST | 27.955 (1.83) | 27.851 (3.21) | 25.924 (9.07) | 27.646 (10.28) |
| URB | 0.0238 (2.09) | 0.0238 (2.19) | 0.0235 (2.20) | 0.0206 (1.93) |
| FAMSIZE | -0.4576 (-0.20) | -0.4667 (-0.24) | | |
| EDUC1 | -0.0018 (-0.01) | | | |
| EDUC2 | -0.1632 (-0.82) | -0.1617 (-2.58) | -0.1517 (-3.29) | -0.1594 (-3.43) |
| EDUC3 | 0.1458 (1.72) | 0.1453 (2.23) | 0.1425 (2.25) | 0.1298 (2.03) |
| UNEMP | 0.1112 (1.55) | 0.1113 (1.58) | 0.1118 (1.61) | |
| MEDINC | -0.5392 (-4.00) | -0.5393 (-4.06) | -0.5496 (-4.42) | -0.5364 (-4.26) |
| R^2 | 0.570 | 0.578 | 0.586 | 0.574 |
| F | 11.795 | 14.036 | 17.144 | 20.183 |
| ESS = $\sum \hat{u}_t^2$ | 133.952 | 133.953 | 134.098 | 140.755 |
| SGMASQ = $\hat{\sigma}_u^2$ | 2.679 | 2.627 | 2.579 | 2.656 |
| AIC | 3.043 | 2.940 | 2.843 | 2.883 |
| SCHWARZ | 4.043 | 3.770 | 3.519 | 3.444 |

Values in parentheses are the t-statistics.