BT2101 Tutorial

Mar 30, 2022

1. Suppose you are researching the value of attending graduate school using a data set collected from 2015 to 2022 from the NUS alumni office. Consider the following estimation equation for the panel data set:

ln (wage$_{it}$) $= \beta_0 + \beta_1 * Education_{it} + D_i + \alpha * X + u_{it}$ ------ (1),

where $Education_{it}$ is the individual's number of years of education, D_i is the individual's unobserved characteristics, X includes control variables that can affect the wage such as experience, characteristics, occupation, tenure, etc. Your explanatory variable is $Education_{it}$.

a. Depending on the assumption of testing the fixed effects for D_i, state the two key assumptions.

b. Suppose you realize that $Education_{it}$ does not necessarily vary over time for individuals once they finish their studies. Explain if model (1) needs to be estimated using fixed effects. Explain how you can determine if applying the fixed effects is necessary.

2. Consider the binary variable version of the fixed effects model in the following equation, except with an additional regressor, $D1_i$; that is, let

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \gamma_1 D1_i + \gamma_2 D2_i + \cdots + \gamma_n Dn_i + \mu_{it}$$

What will happen if you try to estimate the coefficient of the regression by OLS?

3. Alcohol taxes are only one way to discourage drinking and driving. States differ in their punishments for drunk driving, and a state that cracks down on drunk driving could do so by toughening driving laws as well as raising taxes. If so, omitting these laws could produce omitted variable bias in the OLS estimator of the effect of real beer taxes on traffic fatalities, even in regressions with state and time fixed effects. In addition, because vehicle use depends in part on whether drivers have jobs and because tax changes can reflect economic conditions (a state budget deficit can lead to tax hikes), omitting state economic conditions also could result in omitted variable bias. We therefore extend the preceding analysis of traffic fatalities to include other driving laws and economic conditions. The results are summarized in Table 1.

The next four regressions in Table 10.1 include additional potential determinants of fatality rates along with state and time effects. The base specification, reported in Model 1, includes

variables related to drunk driving laws plus variables that control for the amount of driving and overall state economic conditions. The first legal variables are the minimum legal drinking age, represented by three binary variables for a minimum legal drinking age of 18, 19, and 20 (so the omitted group is a minimum legal drinking age of 21 or older).

The other legal variable is the punishment associated with the first conviction for driving under the influence of alcohol, either mandatory jail time or mandatory community service (the omitted group is less severe punishment). The three measures of driving and economic conditions are average vehicle miles per driver, the unemployment rate, and the logarithm of real (1988 dollars) personal income per capita (using the logarithm of income permits the coefficient to be interpreted in terms of percentage changes of income).

Table 1. Results of the econometric model

DV: Traffic fatality rate (death per 10,000)	Model 1	Model 2
Beer tax	-0.45 (0.030)	-0.69* (0.35)
Drinking age 18	0.028 (0.070)	-0.010 (0.070)
Drinking age 19	-0.018 (0.050)	-0.076 (0.068)
Drinking age 20	0.032 (0.051)	-0.100** (0.050)
Mandatory jail or community service	0.038 (0.103)	0.085 (0.112)
Average vehicle miles per driver	0.008 (0.007)	0.017 (0.011)
Unemployment rate	-0.063** (0.013)	-
Real income per capita (logarithm)	1.82** (0.64)	-
Constant	2.049** (0.868)	0.721 (1.019)
Years	1982-88	1982-88
State effects	Yes	Yes
Time effects	Yes	Yes
Category Fixed Effects	Yes	Yes
Clustered standard errors	Yes	Yes

Note: *: $p < 0.1$, **: $p < 0.05$, ***: $p < 0.01$

a. New Jersey has a population of 8.1 million people. Suppose that New Jersey increased the tax on a case of beer by $1 (in 1988 dollars). Use the results in Model 1 to predict the number of lives that would be saved over the next year. Construct a 95% confidence interval for your answer.

b. The drinking age in New Jersey is 21. Suppose that New Jersey lowered its drinking age to 18. Use the results in Model 1 to predict the change in the number of traffic fatalities in the next year. Construct a 95% confidence interval for your answer.

c. Suppose that real income per capita in New Jersey increases by 1% in the next year. Use the results in Model 1 to predict the change in the number of traffic fatalities in the next year. Construct a 90% confidence interval for your answer.

d. Should time effects be included in the regression? Why or why not?

e. A researcher conjectures that the unemployment rate has a different effect on traffic fatalities in the western states than in the other states. How would you test this hypothesis? (Be specific about the specification of the regression and the statistical test you would use.)