Some Examples of Regression

Juanling Liu (15320171151901)*

March 29, 2019

1 Example In Stata

To build a railway, If the railway is a line, so there must be a city that passes through and a city that has not been crossed. If city i is crossed, remember D_i=1; if city i is not crossed, remember D_i=0. Now we want to know, after the railway is repaired, is the city through the railway growing faster? in other words,we want to know $\beta_3$in the following model.

	$Di = 0$	$Di = 1$
$Ti = 0$	0	0
$Ti = 1$	0	1

Model:

$$Y_{it} = \beta_0 + \beta_1 D_i + \beta_2 T + \beta_3 (D_i * T) + \beta_4 X_{it} + \varepsilon_{it} \tag{1}$$

Method: DID

Data sources: http://dss.princeton.edu/training/Panel101.dta

Assume the policy execution time is 1994 and set it as a dummy variable, the place where the policy is executed is also a dummy variable,where the country is greater than 4

*Email: 854230829l@qq.com

```
. use "E:\微观计量\双重差分\Panel101.dta", clear

. gen time = (year>=1994) & !missing(year)

. gen treated = (country>4) & !missing(country)

. gen did = time*treated

. reg y time treated did, r

Linear regression                               Number of obs     =         70
                                                F(3, 66)          =       2.17
                                                Prob > F          =     0.0998
                                                R-squared         =     0.0827
                                                Root MSE          =     3.0e+09
```

y	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
time	2.29e+09	9.00e+08	2.54	0.013	4.92e+08	4.09e+09
treated	1.78e+09	1.05e+09	1.70	0.094	-3.11e+08	3.86e+09
did	-2.52e+09	1.45e+09	-1.73	0.088	-5.42e+09	3.81e+08
_cons	3.58e+08	7.61e+08	0.47	0.640	-1.16e+09	1.88e+09

Figure 1: The effect of railway construction on economic growth

Obviously, at the 10% level, policy implementation has a significant negative effect.That is, the construction of the railway will reduce the economy of the cities along the route.

2 Example In R

In the multiple regression analysis, we would like to explore the relationship between crime rates and factors in a state, including population, illiteracy rate, average income and frosting days (average number of days when the temperature is below freezing)

Model:

$$Murder = \beta_0 + \beta_1 population + \beta_2 illiteracy + \beta_3 income + \beta_4 frost + \varepsilon_i \tag{2}$$

Method: OLS

Data sources: datasets stata.x77 in the underlying package

```
> states <- as.data.frame(state.x77[,c("Murder", "Population", "Illiteracy", "Income", "Frost")])
> fit <- lm(Murder ~ Population + Illiteracy + Income + Frost, data=states)
> summary(fit)

Call:
lm(formula = Murder ~ Population + Illiteracy + Income + Frost,
    data = states)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.7960 -1.6495 -0.0811  1.4815  7.6210 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 1.235e+00  3.866e+00   0.319   0.7510    
Population  2.237e-04  9.052e-05   2.471   0.0173 *  
Illiteracy  4.143e+00  8.744e-01   4.738 2.19e-05 ***
Income      6.442e-05  6.837e-04   0.094   0.9253    
Frost       5.813e-04  1.005e-02   0.058   0.9541    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.535 on 45 degrees of freedom
Multiple R-squared:  0.567,     Adjusted R-squared:  0.5285 
F-statistic: 14.73 on 4 and 45 DF,  p-value: 9.133e-08
```

Figure 2: The relationship between crime rate and other factors

The regression coefficient of the illiteracy rate is 4.14, indicating that when the population is controlled, income and temperature remain unchanged, the illiteracy rate increases by 1%, the murder rate will rise by 4.14%, and its coefficient is significantly not equal to 0 at the p<0.001 level. On the contrary, the p-value=0.954 of the Frost indicates that Frost and Murder are not linearly correlated when the control of other variables remains unchanged.