Solutions to HW3

Guanxi Li

June 1, 2024

1

The model can be written as

$$Y_{it} = \sum_{i=1}^{n} \alpha_i D_i + u_{it}$$

1.1

The objective function is

$$Q = \min_{\alpha_1, \alpha_2, \ldots, \alpha_n} \sum_{t=1}^{T} \left(Y_{it} - \sum_{i=1}^{n} \alpha_i D_i \right)^2$$

F.O.C. is

$$\frac{\partial}{\partial \alpha_i} \sum_{t=1}^{T} \left(Y_{it} - \sum_{i=1}^{n} \alpha_i D_i \right)^2 = -2 \sum_{t=1}^{T} (Y_{it} - \alpha_i) = 0$$

Thus

$$\hat{\alpha}_i = \frac{1}{T} \sum_{t=1}^{T} Y_{it}$$

1.2

$$E(\hat{\alpha}_i) = E\left(\frac{1}{T}\sum_{t=1}^{T} Y_{it}\right) = \alpha_i + E\left(\frac{1}{T}\sum_{t=1}^{T} u_{it}\right) = \alpha_i$$

$$\mathrm{Var}(\hat{\alpha}_i) = \mathrm{Var}\left(\frac{1}{T}\sum_{t=1}^{T} u_{it}\right) = \frac{1}{T^2}\sum_{t=1}^{T} \mathrm{Var}(u_{it}) = \frac{1}{T^2}\sum_{t=1}^{T} \sigma_u^2 = \frac{\sigma_u^2}{T}$$

The variance won't converge to zero as N goes to infinity. Thus $\hat{\alpha}_i$ is inconsistent.

2 P292, 2

Table 1: Results

	(1) fatalityrate	(2) fatalityrate	(3) fatalityrate
sb_useage	0.004 07***	−0.005 77***	−0.003 72***
	(0.00123)	(0.00116)	(0.00113)
speed65	0.000 148	−0.000 425	−0.000 783*
	(0.000408)	(0.000334)	(0.000424)
speed70	0.002 40***	0.001 23***	0.000 804**
	(0.000472)	(0.000329)	(0.000340)
ba08	−0.001 92***	−0.001 38***	−0.000 822**
	(0.000361)	(0.000373)	(0.000352)
drinkage21	0.000 079 9	0.000 745	−0.001 13**
	(0.000987)	(0.000507)	(0.000535)
lnincome	−0.0181***	−0.0135***	0.006 26
	(0.00109)	(0.00142)	(0.00387)
age	−0.000 007 22	0.000 979**	0.001 32***
	(0.000164)	(0.000382)	(0.000383)
Constant	0.197***	0.121***	−0.0850**
	(0.00925)	(0.00977)	(0.0403)
Observations	556	556	556
R-squared	0.549	0.887	0.910
State F.E.	No	Yes	Yes
Year F.E.	No	No	Yes

* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

2.1

The coefficient on seat belt useage is positive and significant at the 1 percent level, suggesting that using seat belt will lead to an increase in fatality rate.

2.2

The coefficient on seat belt useage is significantly negative now. It suggests that states with more dangerous driving conditions have more people wearing seat belts.

2.3

The coefficient on seat belt useage turns bigger and is -0.00372 now.

2.4

$F = 8.85$ for time effects, so (3) is better.

2.5

A 38% increase in seat belt useage could lower the fatality rate by $0.38 \times -0.00372 = -0.0014$. The number of fatalities prevented is $0.0014 \times 41447.73 = 58$. (Note: Number of fatalities $=$ fatalityrate $\times$ vmt)

2.6

The coefficients on primary and secondary are 0.206 and 0.109 respectively, and are both significant at 1 percent level.

2.7

This results in an estimated increase in seatbelt useage of $0.206 - 0.109 = 0.097$. The fatality rate is estimated to be reduced by $0.00372 \times 0.097 = 0.00036$ fatalities per million traffic miles. $0.00036 \times 63008 = 22$ lives were saved.

3 P354, 6

$F = \frac{(R_{un}^2 - R_r^2)/q}{(1-R_{un}^2)/(n-k_{un}-1)} = \frac{0.05/1}{(1-0/05)/98} = 5.16$ with 100 observations, and $F = \frac{(R_{un}^2 - R_r^2)/q}{(1-R_{un}^2)/(n-k_{un}-1)} = \frac{0.05/1}{(1-0/05)/498} = 26.2$ with 500 observations.

4

$$\hat{\beta}_{IV} = \frac{Cov(Z_i, Y_i)}{Cov(Z_i, X_i)} = \beta_1 + \frac{\beta_2 Cov(Z_i, W_i)}{Cov(Z_i, X_i)}$$

4.1

If $Cov(Z_i, W_i) = 0$, $\hat{\beta}_{IV}$ is consistent.

4.2

If $Cov(Z_i, W_i) \neq 0$, $\hat{\beta}_{IV}$ is not consistent.

5

Table 2: Results

	(1) OLS	(2) IV	(3) IV
morekids	−5.387***	−6.314***	−5.821***
	(0.0871)	(1.275)	(1.246)
agem1			0.832***
			(0.0229)
black			11.62***
			(0.229)
hispan			0.404
			(0.260)
othrace			2.131***
			(0.206)
Constant	21.07***	21.42***	−4.792***
	(0.0561)	(0.487)	(0.407)
Observations	254,654	254,654	254,654
R-squared	0.014	0.014	0.044
Controls	No	No	Yes

* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

5.1

The coefficient is -5.387, indicating that women with more than two kids work 5.387 fewer weeks per year than women with two or fewer kids.

5.2

Fertility and weeks worked are both choice variable. Women with more weeks worked could lead to lower fertility, implicating that the dependent variable is correlated with the error term.

5.3

When $samesex = 1$, couples are 6.7% more likely to have an additional child, which is significant at 1 percent level.

5.4

Whether the first two children of a couple is relatively exogenous, implying that it is uncorrelated with the error term. Also, it can only affect the weeks worked through whether the couple want to have more kids. This is because in the United States, families who have had two children of the same sex are more likely to continue having children.

5.5

The F-statistic in the first stage is 1237, implicating that *samesex* is a valid IV.

5.6

The coefficient is -6.314.

5.7

The results do not change much. It is because $agem1$, $black$, $hispan$, $othrace$ are uncorrelated with $samesex$. Adding more exogenous control variables can make the effects we are interested in clearer.