ARE213 Problem Set #2A

Peter Alstone & Frank Proulx

November 5, 2013

1 Problem #1

1.1 Part A

Here we will consider the within estimator, as suggested. This suggests that we want to find $\widehat{\beta_{FE}}$ by running the following regression:

$$\ddot{Y_{it}} = \ddot{X}_{it}' \widehat{\beta_{FE}} + \ddot{\epsilon_{it}} \tag{1}$$

Where $\ddot{Y_{it}} = Y_{it} - \bar{Y}_i$, $\ddot{X_{it}} = X_{it} - \bar{X}_i$, $\bar{Y}_i = \frac{1}{T}\sum_{t=1}^{T} Y_{it}$, and $\bar{X}_i = \frac{1}{T}\sum_{t=1}^{T} X_{it}$. Our fixed effects estimator is therefore

$$\widehat{\beta_{FE}} = (\ddot{X}_{it}' \ddot{X_{it}})^{-1} \ddot{X}_{it}' \ddot{Y_{it}} \tag{2}$$

Because we have T=2, this can be rewritten as

$$\widehat{\beta_{FE}} = ((X_{it} - \frac{1}{2}X_{i1} - \frac{1}{2}X_{i2})'(X_{it} - \frac{1}{2}X_{i1} - \frac{1}{2}X_{i2}))^{-1}(X_{it} - \frac{1}{2}X_{i1} - \frac{1}{2}X_{i2})'(Y_{it} - \frac{1}{2}Y_{i1} - \frac{1}{2}Y_{i2}) \tag{3}$$

And looking at the first differences estimator, where we get $\widehat{\beta_{FD}}$ by running the regression

$$\Delta Y_{it} = \Delta X_{it}' \widehat{\beta_{FD}} + \Delta \epsilon_{it} \tag{4}$$

where $\Delta Y_{it} = Y_{it} - Y_{it-1}$, $\Delta X_{it} = X_{it} - X_{it-1}$, and $\Delta \epsilon_{it} = \epsilon_{it} - \epsilon_{it-1}$.

Because we only have T=2, this differencing estimator can only be estimated for t=2, so we regress

$$\Delta Y_{i2} = \Delta X_{i2}' \widehat{\beta_{FD}} + \Delta \epsilon_{i2} \tag{5}$$

$$Y_{i2} - Y_{i1} = (X_{i2} - X_{i1})' \widehat{\beta_{FD}} + \epsilon_{i2} - \epsilon_{i1} \tag{6}$$

thus, our first difference estimator is

$$\widehat{\beta_{FD}} = ((X_{i2} - X_{i1})'(X_{i2} - X_{i1}))^{-1}(X_{i2} - X_{i1})'(Y_{i2} - Y_{i1}) \tag{7}$$

2 Problem #3

2.1 Part A

Running pooled bivariate OLS, adding a quadratic time trend, and adding the covariates that we expect to belong produces the models shown in Table .

Pooled analysis assumes that there are no unobserved unique time constant attributes of individuals and no universal effects across time.

The bivariate case suggests that the log of per capita traffic fatalities decreases by 0.144 with the introduction of primary seatbelt laws, across both place and time. The quadratic time term is significant at a high level, but has a considerably smaller

2.2 Part B

2.3 Part C

The between estimator will give an unbiased estimate of the effect of primary seat belt laws insofar as variation within states (across time) is uncorrelated with the observables.

We don't think that this criterion is met here. For example, within a given state, the total vmt per year probably tracks very closely with fatalities, as the higher VMT within a given year, the more likely there are to be fatal crashes (ceteris paribus).

2.4 Part D

The RE estimator will give an unbiased estimate so long as the within states variation is uncorrelated with observables. Again, this assumption is probably not met here.

The Random Effects estimator has the advantage over pooled OLS that it allows for (and assumes) unobserved heterogeneity. OLS has the advan-

Table 1: Pooled Models of Fatalities Per Capita

	Dependent variable:		
	logfatalpc		
	bivariate	quadratic time	covariates
	(1)	(2)	(3)
primary	-0.144^{***}	-0.082^{***}	-0.064^{**}
	(0.026)	(0.026)	(0.025)
sqnetyears		-0.001^{***}	-0.0004^{***}
		(0.0001)	(0.0001)
secondary			-0.076^{***}
			(0.020)
college			-3.019^{***}
			(0.180)
beer			0.287^{***}
			(0.032)
totalvmt			-0.00000^{***}
			(0.00000)
precip			-0.021^{***}
			(0.006)
snow32			-0.312^{***}
			(0.018)
rural_speed			0.008^{***}
			(0.002)
urban_speed			0.008^{***}
			(0.002)
Constant	-1.703^{***}	-1.626^{***}	-2.051^{***}
	(0.011)	(0.014)	(0.146)
Observations	1,127	1,127	1,127
R^2	0.027	0.079	0.573
Adjusted R^2	0.027	0.079	0.568
F Statistic	31.007^{***} (df = 1; 1125)	48.494^{***} (df = 2; 1124)	150.050^{***} (df = 10; 1116)

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 2: Between models of effects of primary seatbelt use laws

	Dependent variable:	
	logfatalpc	
	(1)	(2)
primary	−0.071	0.114
	(0.155)	(0.165)
secondary		−0.025
		(0.159)
college		-2.603^{***}
		(0.636)
totalvmt		-0.00000^{**}
		(0.00000)
snow32		-0.266^{***}
		(0.083)
rural_speed		0.063^{***}
		(0.012)
Constant	-1.716^{***}	-4.978^{***}
	(0.052)	(0.789)
Observations	49	49
R^2	0.004	0.758
Adjusted R^2	0.004	0.650
F Statistic	0.212 (df = 1; 47)	21.898^{***} (df = 6; 42)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 3: Random Effects Models

	Dependent variable:	
	logfatalpc	
	(1)	(2)
primary	-0.230^{***}	-0.138^{***}
	(0.016)	(0.015)
secondary		-0.065^{***}
		(0.010)
college		-1.420^{***}
		(0.169)
unemploy		-0.024^{***}
		(0.002)
beer		0.757^{***}
		(0.038)
totalvmt		-0.00000^{***}
		(0.00000)
precip		-0.024^{***}
		(0.006)
snow32		-0.018
		(0.014)
rural_speed		-0.006^{***}
		(0.001)
urban_speed		0.003^{***}
		(0.001)
Constant	-1.688^{***}	-1.902^{***}
	(0.044)	(0.092)
Observations	1,127	1,127
R^2	0.153	0.604
Adjusted R^2	0.153	0.598
F Statistic	203.173^{***} (df = 1; 1125)	170.214^{***} (df = 10; 1116)
Note:		*p<0.1; **p<0.05; ***p<0.01

tage that it is more efficient than the Random Effects estimator when said heterogeneity does not exist.

2.5 Part E

2.6 Part F

2.7 Part G

Table 4: Fixed Effects Models

	Dependent variable:	
	logfatalpc	
	(1)	(2)
primary	-0.098^{***}	-0.099^{***}
	(0.015)	(0.015)
sqnetyears	-0.001^{***}	-0.0003^{***}
	(0.00003)	(0.0001)
secondary		-0.030^{***}
		(0.010)
college		0.176
		(0.239)
beer		0.859^{***}
		(0.040)
totalvmt		-0.00000^{*}
		(0.00000)
precip		-0.033^{***}
		(0.006)
snow32		-0.003
		(0.015)
rural_speed		0.0003
		(0.001)
urban_speed		0.003^{***}
		(0.001)
Observations	1,127	1,127
R^2	0.384	0.609
Adjusted R^2	0.367	0.578
F Statistic	336.044^{***} (df = 2; 1076)	166.694^{***} (df = 10; 1068)
Note:		*p<0.1; **p<0.05; ***p<0.01