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

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

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 reftab:3a. The pooled OLS is an ill-informed baseline model but nonetheless tells us that there is a statistically significant negative correlation between the states with primary seatbelt laws and those without. In particular, the mean of the log fatalities per capita is reduced by 0.14 for states with primary seatbelt laws. This does not, however, account for the year to year trends that are captured by including a set of quadratic year terms in the regression as shown in the "quadratic time" regression. This basic trend is that deaths have gone down over time. Correcting for these trends may allow our OLS estimates approach the true ATE, and indeed the apparent effect of primary seatbelt laws is reduced. When further covariates are added that are relevant (see Table reftab:3a) further reductions in the apparent effect occur and the effect is no longer statistically significant.

2.2 Part B

No, the standard errors are most likely not correct (are they ever really correct?). In this case the error comes from

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).

Table 1: Pooled Models of Fatalities Per Capita

| | logfatalpc | | |
| | bivariate | quadratic time | covariates |
	(1)	(2)	(3)
primary	−0.144***	−0.075***	0.015
	(0.026)	(0.026)	(0.027)
year		−1.906**	−3.838***
		(0.968)	(0.760)
sqyears		0.0005*	0.001***
		(0.0002)	(0.0002)
secondary			−0.001
			(0.022)
college			−2.902***
			(0.176)
beer			0.268***
			(0.031)
totalvmt			−0.00000***
			(0.00000)
precip			−0.015**
			(0.006)
snow32			−0.297***
			(0.017)
rural_speed			0.016***
			(0.002)
urban_speed			0.005***
			(0.002)
Constant	−1.703***	−1,863.345*	−3,760.458***
	(0.011)	(952.304)	(747.809)
N	1,127	1,127	1,127
R^2	0.027	0.094	0.596
Adjusted R^2	0.027	0.094	0.590
F Statistic	31.007*** (df = 1; 1125)	39.030*** (df = 3; 1123)	149.510*** (df = 11; 1115)

Notes: 3   ***Significant at the 1 percent level.
**Significant at the 5 percent level.
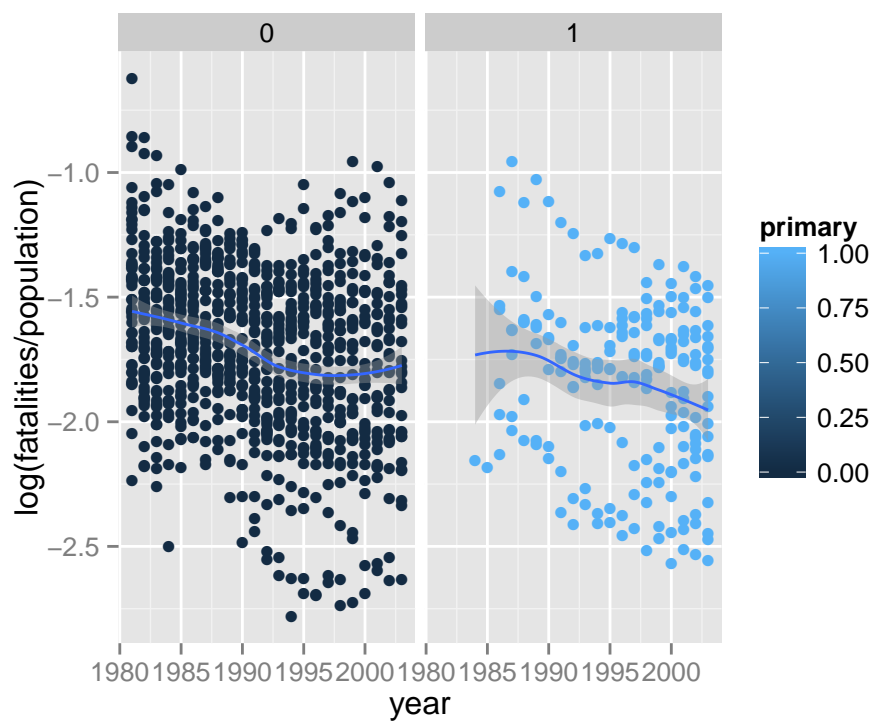*Significant at the 10 percent level.

Figure 1: Year to year trends in the log of traffic fatalities per capita, divided by primary seatbelt law presence. A LOESS fit to each dataset is included for reference but is not necessarily indicative of the true underlying function.

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

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 advantage 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 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

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