Chris Kalra

Econ 610

Homework 9

1a) 

The estimate on the 1990 dummy variable, which is extremely statistically significant, tells us that when all else is equal, rent increased by 26% over the 10-year period. The coefficient on *pctstu*, which is statistically significant, tells us that a 1 percent increase in *pctstu* increases rent by half a percent (.5%)

1b) The standard errors are not valid since we are ignoring the ai term. If the ai term is nonzero, the errors across the time periods are correlated, which invalidates the OLS standard errors

1c)



The estimated coefficient on *pctstu* (which is statistically significant) is greater than twice what it was before. Now, a one percent increase in *pctstu* results in an estimated increase in rental rates of 1.1%. However, differencing causes us to obtain a less precise estimate, even though we have eliminated ai from our model, and there may still be other variables in uit that are correlated with 

1d)



n = 64, T = 2

There is no intercept, as it is removed due to time demeaning. Note that the estimated coefficient on *y90t* is the same as the estimated intercept coefficient from part (c). Furthermore, all the other variables have the same estimates and standard errors here as they did in part (c)

2a) If past executions serve as a deterrent, β1 would be negative, as more executions would lead to a decrease in murder rate due to executions being a deterrent. Intuitively, β2 would be positive; as unemployment increases, more murders are committed (possibly due to financial disputes, or because in general, a higher unemployment rate tends to be correlated with a higher crime rate), though I do not believe this assertion to be as strong as β1 being negative

2b) The coefficient on *exec* is not statistically significant, meaning execution does not appear to have a deterrent effect

2c) 

Now there is evidence of a deterrent effect, as not only is the coefficient on *exec* negative, but it has a p-value of .02073, meaning it is significant at the 5% level. According to the coefficient estimate, 10 additional executions is estimated to reduce the murder rate by 1.04, which is just over one fewer murder per 100,00 people (murders tend to be reported at a rate of per 100,000)

2d) The state with the most executions is Texas with 34, and the state with the second most executions is Virginia with 11. Therefore, Texas had 23 more executions than the state with the next highest number of executions

2e) 

Now, none of the variables are statistically significant at the 5% level (though *d93t comes close, with a p-value of .0567*). Therefore, we no longer have a deterrent effect

3a) 18.65% more fourth graders in 1995 in this district passed the standardized math test than in 1994

3b) Each 1% increase in real expenditures per student in the district leads to 0.005% additional fourth graders passing the standardized math test

3c) For each additional 1% students eligible for the school lunch program. 0.407% fewer fourth graders are expected to pass the standardized math test

3d) It means the errors have a strong positive correlation with the errors at the next point in time (lag of 1). This is not surprising, as there are a lot of factors that may change over time that we are not measuring, such as

3e) No they are not; the errors across the two time periods are positively correlated, which invalidates the usual OLS standard errors

3g) It becomes positive, which is counterintuitive. However, it also becomes statistically insignificant at the 5% significance level (with an approximate t-value calculated to be 1.22)

3h) β4 would likely get closer to 0, as district is likely correlated with income (and whether or not a student is on the lunch program is directly dependent upon their parent’s income)

3i) The estimate is 0.534 – 9.049 = -8.515. To find the standard error, we would need to create a new variable in R that is (*log(rexppit)* + *log(rexppi, t-1*)). Then, in the output, look at the standard error on the coefficient for *log(rexppit).* This is the estimate for the standard error of 

3j) A  value of 0.58 means the model is slightly more of a fixed effects model than a pooled OLS. A quick glance at some of the variables, such as the intercept, *log(rexpp),* and *lunch*, indicate that this may be a realistic value for , as these variable’s coefficients under the random effects model have values that are slightly closer to the coefficients of those in the fixed effects than the OLS model

*Code (italics)* and Output (Monaco)

1a) *rental=data*

*regr1a=lm(lrent ~ y90 + lpop + lavginc + pctstu, data=rental)*

*summary(regr1a)*

Call:

lm(formula = lrent ~ y90 + lpop + lavginc + pctstu, data = rental)

Residuals:

Min 1Q Median 3Q Max

-0.24233 -0.07824 -0.01642 0.04389 0.48082

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -0.568807 0.534881 -1.063 0.2897

y90 0.262227 0.034763 7.543 8.78e-12 \*\*\*

lpop 0.040686 0.022515 1.807 0.0732 .

lavginc 0.571446 0.053098 10.762 < 2e-16 \*\*\*

pctstu 0.005044 0.001019 4.949 2.40e-06 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1259 on 123 degrees of freedom

Multiple R-squared: 0.8613, Adjusted R-squared: 0.8568

F-statistic: 190.9 on 4 and 123 DF, p-value: < 2.2e-16

1c) *regr1c=lm(clrent ~ clpop + clavginc + cpctstu, data=rental)*

*summary(regr1c)*

Call:

lm(formula = clrent ~ clpop + clavginc + cpctstu, data = rental)

Residuals:

Min 1Q Median 3Q Max

-0.18697 -0.06216 -0.01438 0.05518 0.23783

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.385521 0.036824 10.469 3.66e-15 \*\*\*

clpop 0.072246 0.088343 0.818 0.41671

clavginc 0.309961 0.066477 4.663 1.79e-05 \*\*\*

cpctstu 0.011203 0.004132 2.711 0.00873 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.09013 on 60 degrees of freedom

(64 observations deleted due to missingness)

Multiple R-squared: 0.3223, Adjusted R-squared: 0.2884

F-statistic: 9.51 on 3 and 60 DF, p-value: 3.136e-05

1d) *install.packages("plm")*

*library("plm")*

*rental=pdata.frame(rental, index=c("city", "year"), drop.index=TRUE)*

*regr1d=plm(lrent ~ y90 + lpop + lavginc + pctstu, data=rental)*

*summary(regr1d)*

Oneway (individual) effect Within Model

Call:

plm(formula = lrent ~ y90 + lpop + lavginc + pctstu, data = rental)

Balanced Panel: n = 64, T = 2, N = 128

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-0.118915 -0.029559 0.000000 0.029559 0.118915

Coefficients:

Estimate Std. Error t-value Pr(>|t|)

y90 0.3855214 0.0368245 10.4692 3.661e-15 \*\*\*

lpop 0.0722456 0.0883426 0.8178 0.416714

lavginc 0.3099605 0.0664771 4.6627 1.788e-05 \*\*\*

pctstu 0.0112033 0.0041319 2.7114 0.008726 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares: 10.383

Residual Sum of Squares: 0.24368

R-Squared: 0.97653

Adj. R-Squared: 0.95032

F-statistic: 624.146 on 4 and 60 DF, p-value: < 2.22e-16

2b) *murder=data*

*regr2b = plm(mrdrte~d93 + exec + unem, data=murder, subset=d90+d93==1, model='pooling', index=c("id", "year"))*

*summary(regr2b)*

Pooling Model

Call:

plm(formula = mrdrte ~ d93 + exec + unem, data = murder, subset = d90 +

d93 == 1, model = "pooling", index = c("id", "year"))

Balanced Panel: n = 51, T = 2, N = 102

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-13.0666 -3.3556 -1.6472 1.6071 66.3873

Coefficients:

Estimate Std. Error t-value Pr(>|t|)

(Intercept) -5.27800 4.42781 -1.1920 0.236134

d93 -2.06742 2.14463 -0.9640 0.337421

exec 0.12773 0.26324 0.4852 0.628599

unem 2.52889 0.78172 3.2350 0.001659 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares: 11401

Residual Sum of Squares: 10243

R-Squared: 0.10161

Adj. R-Squared: 0.07411

F-statistic: 3.69475 on 3 and 98 DF, p-value: 0.0144

2c) *regr3c=plm(mrdrte~d93 + exec + unem, data=murder, subset=d90+d93==1, model='within', index=c("id", "year"))*

*summary(regr3c)*

Oneway (individual) effect Within Model

Call:

plm(formula = mrdrte ~ d93 + exec + unem, data = murder, subset = d90 +

d93 == 1, model = "within", index = c("id", "year"))

Balanced Panel: n = 51, T = 2, N = 102

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-1.27948 -0.34897 0.00000 0.34897 1.27948

Coefficients:

Estimate Std. Error t-value Pr(>|t|)

d93 0.413266 0.209385 1.9737 0.05418 .

exec -0.103840 0.043414 -2.3918 0.02073 \*

unem -0.066591 0.158686 -0.4196 0.67662

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares: 33.47

Residual Sum of Squares: 27.936

R-Squared: 0.16533

Adj. R-Squared: -0.75627

F-statistic: 3.16936 on 3 and 48 DF, p-value: 0.032614

2d) *sort(year93$exec)*

[1] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 2 2 2 2 2 2 3 3 6 7 11 34

2e) *murder.2=murder[-c(130,131,132), ]*

*regr3e=plm(mrdrte~d93 + exec + unem, data=murder.2, subset=d90+d93==1, model='within', index=c("id", "year"))*

*summary(regr3e)*

Oneway (individual) effect Within Model

Call:

plm(formula = mrdrte ~ d93 + exec + unem, data = murder.2, subset = d90 +

d93 == 1, model = "within", index = c("id", "year"))

Balanced Panel: n = 50, T = 2, N = 100

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-1.3183 -0.3701 0.0000 0.3701 1.3183

Coefficients:

Estimate Std. Error t-value Pr(>|t|)

d93 0.412523 0.211283 1.9525 0.05685 .

exec -0.067471 0.104913 -0.6431 0.52328

unem -0.070032 0.160371 -0.4367 0.66434

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares: 31.05

Residual Sum of Squares: 27.85

R-Squared: 0.10306

Adj. R-Squared: -0.8893

F-statistic: 1.80012 on 3 and 47 DF, p-value: 0.16008