Lab2_RM

Nathan Colbert

10/30/2017

A

For this lab I am going to examine ones confidence in the scientific community as a function of their confidence in organized religion with their panel wave as a factor.

```
panel <- read.csv("Downloads/panel-for-R.csv")
vars <- c("consci", "conclerg", "panelwave", "paeduc", "maeduc", "idnum")
lab2 <- panel[, vars]
lab2 <- ddply(lab2, "idnum", mutate, d.consci = firstD(consci), d.conclerg = firstD(conclerg))
lab2$consci <- 4 - lab2$consci
lab2$conclerg <- 4 - lab2$conclerg

table(lab2$conclerg)
```

```
##
##    1    2    3
##  701 1751  667
```

```
table(lab2$consci)
```

```
##
##    1    2    3
##  232 1627 1188
```

```
lm1 <- lm(consci ~ conclerg + as.factor(panelwave), lab2)
summary(lm1)
```

```
##
## Call:
## lm(formula = consci ~ conclerg + as.factor(panelwave), data = lab2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3991 -0.3424 -0.2751  0.6575  0.7816
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)              2.22907    0.03804  58.593  < 2e-16 ***
## conclerg                 0.05669    0.01677   3.381 0.000731 ***
## as.factor(panelwave)2 -0.02863    0.02602  -1.100 0.271353
## as.factor(panelwave)3 -0.06732    0.02752  -2.446 0.014502 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.605 on 2992 degrees of freedom
##   (3004 observations deleted due to missingness)
## Multiple R-squared:  0.006114,   Adjusted R-squared:  0.005118
## F-statistic: 6.135 on 3 and 2992 DF,  p-value: 0.000372
```

Examining this linear model, we see that over time net of all other factors, confidence in the scientific community is decreasing by approximately .03 per panel wave, however these results are not statistically significant (or maybe barely so) and shouldnt be given very much clout. Interestingly enough, this model suggests that for every one unit increase in confidence in organinized religion, their is a corresponding increase of .06*** in their confidence in the scientific community which is highly statiscally significant.

B

```
con_fe <- plm(consci ~ conclerg + as.factor(panelwave),
              index = c("idnum", "panelwave"),
              model = "within",
              data = lab2)

summary(con_fe)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = consci ~ conclerg + as.factor(panelwave), data = lab2,
##     model = "within", index = c("idnum", "panelwave"))
##
## Unbalanced Panel: n = 1295, T = 1-3, N = 2996
##
## Residuals:
##     Min.  1st Qu.   Median  3rd Qu.     Max.
## -1.37577 -0.11003  0.00000  0.10433  1.37423
##
## Coefficients:
##                          Estimate Std. Error t-value  Pr(>|t|)
## conclerg                 0.101399   0.024060  4.2144 2.636e-05 ***
## as.factor(panelwave)2 -0.043975   0.021991 -1.9997 0.0456956 *
## as.factor(panelwave)3 -0.083327   0.023674 -3.5197 0.0004434 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    386.67
## Residual Sum of Squares: 379.26
## R-Squared:      0.019158
## Adj. R-Squared: -0.73005
## F-statistic: 11.0551 on 3 and 1698 DF, p-value: 3.4555e-07
```

However, when running a fixed effects model, the statisical significance of all three independent variables increases. This suggests that through holding constant things which we wouldnt percieve to change over time, we remove any confounding of these variables. The coefficient on conclerg doubled, simalarily the coefficients on panel wave grew by almost -.02 points.

C

In order to attempt to explain this relationship further, I am going to include mother and fathers education as variables.

```
con_fe2 <- plm(consci ~ conclerg + as.factor(panelwave) + maeduc + paeduc,
              index = c("idnum", "panelwave"),
```

```
              model = "within",
              data = lab2)

summary(con_fe2)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = consci ~ conclerg + as.factor(panelwave) + maeduc +
##     paeduc, data = lab2, model = "within", index = c("idnum",
##     "panelwave"))
##
## Unbalanced Panel: n = 1034, T = 1-3, N = 2183
##
## Residuals:
##      Min.   1st Qu.    Median   3rd Qu.      Max.
## -1.367864 -0.069826  0.000000  0.067147  1.312729
##
## Coefficients:
##                          Estimate Std. Error t-value  Pr(>|t|)
## conclerg                0.0552459  0.0295377  1.8704 0.0616896 .
## as.factor(panelwave)2 -0.0570899  0.0261085 -2.1866 0.0289709 *
## as.factor(panelwave)3 -0.1017476  0.0278688 -3.6509 0.0002731 ***
## maeduc                -0.0078304  0.0112585 -0.6955 0.4868792
## paeduc                 0.0208886  0.0117802  1.7732 0.0764622 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    243.17
## Residual Sum of Squares: 238.33
## R-Squared:      0.019874
## Adj. R-Squared: -0.86944
## F-statistic: 4.63929 on 5 and 1144 DF, p-value: 0.00033839
```

Inclusion of mother and fathers education completely decimates the statistcal significance in confidence in organized religion as a predictor for confidence in the scientific community. This is an interesting result. Not only is conclerg no longer statistically significant, but maeduc and paeduc are not statistically significant either. Looking at this fixed effect model with these added variables, it appears neither mother or fathers education, or respondents confidence in organized religion are statistically significant predictors of confidence in the scienctific community in either direction. However, net of all other factors, panel wave remains a statistically significant predictor of a decline in confidence in the scientific community across each panel by roughly -.06 points.

D

```
con_re <-plm(consci ~ conclerg + as.factor(panelwave) + maeduc + paeduc,
              index = c("idnum", "panelwave"),
              model = "random",
              data = lab2)
summary(con_re)
```

```
## Oneway (individual) effect Random Effect Model
##    (Swamy-Arora's transformation)
```

```
##
## Call:
## plm(formula = consci ~ conclerg + as.factor(panelwave) + maeduc +
##     paeduc, data = lab2, model = "random", index = c("idnum",
##     "panelwave"))
##
## Unbalanced Panel: n = 1034, T = 1-3, N = 2183
##
## Effects:
##                  var std.dev share
## idiosyncratic 0.2083  0.4564 0.604
## individual    0.1367  0.3697 0.396
## theta:
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.2229  0.3424  0.4196  0.3726  0.4196  0.4196
##
## Residuals:
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## -1.43011 -0.29854 -0.13238  0.00331  0.43172  1.04254
##
## Coefficients:
##                          Estimate Std. Error t-value  Pr(>|t|)
## (Intercept)            2.0191280  0.0642248 31.4384 < 2.2e-16 ***
## conclerg               0.0594126  0.0200606  2.9617 0.0030929 **
## as.factor(panelwave)2 -0.0513367  0.0243034 -2.1123 0.0347723 *
## as.factor(panelwave)3 -0.0966636  0.0258252 -3.7430 0.0001865 ***
## maeduc                 0.0019851  0.0052892  0.3753 0.7074692
## paeduc                 0.0190303  0.0048090  3.9572 7.826e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    511.05
## Residual Sum of Squares: 456.61
## R-Squared:      0.10947
## Adj. R-Squared: 0.10742
## F-statistic: 51.908 on 5 and 2177 DF, p-value: < 2.22e-16
```

Looking at the random effects model, statistical significance returns to conclerg, and fathers education becomes highly statistically significant. Under the random effects model, a one category growth in a fathers education level leads to a .019*** increase in their confidence in the scientific community. To determine which model is more accurate, I will use a hausman test.

E

```
phtest(con_fe2, con_re)
```

```
##
##  Hausman Test
##
## data:  consci ~ conclerg + as.factor(panelwave) + maeduc + paeduc
## chisq = 1.4019, df = 5, p-value = 0.9241
## alternative hypothesis: one model is inconsistent
```

The p-value of .9241 is greater then .05, suggesting that we should use the random effects model. Our test

suggests that our unique errors are not correlated with our predictors, and thus, a random effects model will be more appropriate. This also suggests that our fixed effects model was too stringent to provide and accurate understanding of the relationship.