Panel Analysis of the impact of WF Smoke on Solar Gen: West US 2013-2020

R Markdown cheat sheet: https://www.rstudio.com/blog/the-r-markdown-cheat-sheet/ (https://www.rstudio.com/blog/the-r-markdown-cheat-sheet/)

Use “?rmarkdown::_____” to look up formatting options for the YAML header. For example, “?rmarkdown::html_document” will provide potential options for html documents. More info on Rmarkdown anatomy can be found at https://bookdown.org/yihui/rmarkdown-cookbook/rmarkdown-anatomy.html (https://bookdown.org/yihui/rmarkdown-cookbook/rmarkdown-anatomy.html)

Short guide to chunk options: https://bookdown.org/yihui/rmarkdown/r-code.html (https://bookdown.org/yihui/rmarkdown/r-code.html)

Panel data regression analysis

```
##Unbalanced Panel
unbalanced_panel <- read.csv('data/processed/unbalanced_panel_data.csv',
                        header = TRUE,
                        sep = ",")

#Ordinary Least Squares
ols_unbal <- lm(total_generation ~ pm_measure, data = unbalanced_panel)

summary(ols_unbal)
```

```
##
## Call:
## lm(formula = total_generation ~ pm_measure, data = unbalanced_panel)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -132940  -32554  -28964  -13719  837756
##
## Coefficients:
##             Estimate Std. Error t value             Pr(>|t|)
## (Intercept)  26178.2     2006.6  13.046 < 0.0000000000000002 ***
## pm_measure     987.6      192.5   5.129          0.000000301 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 81230 on 5307 degrees of freedom
##   (4398 observations deleted due to missingness)
## Multiple R-squared:  0.004933,   Adjusted R-squared:  0.004745
## F-statistic: 26.31 on 1 and 5307 DF,  p-value: 0.0000003013
```

```r
#Fixed Effects
fe_model_unbal <-
  plm(
    total_generation ~ pm_measure,
    data = unbalanced_panel,
    index = c("fips_codes", "date"),
    effect = "individual",
    model = "within"
  )

summary(fe_model_unbal)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = total_generation ~ pm_measure, data = unbalanced_panel,
##     effect = "individual", model = "within", index = c("fips_codes",
##         "date"))
##
## Unbalanced Panel: n = 85, T = 1-96, N = 5309
##
## Residuals:
##       Min.    1st Qu.     Median    3rd Qu.       Max.
## -336468.07   -3332.82    -666.39    2419.69  534040.62
##
## Coefficients:
##            Estimate Std. Error t-value     Pr(>|t|)
## pm_measure  -666.59     125.45 -5.3134 0.0000001121 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    11412000000000
## Residual Sum of Squares: 11351000000000
## R-Squared:      0.0053762
## Adj. R-Squared: -0.01081
## F-statistic: 28.2317 on 1 and 5223 DF, p-value: 0.00000011209
```

```r
#Year and Month Fixed Effects
fe.time_model_unbal <- plm(
  total_generation ~ pm_measure +
    factor(year) + factor(month),
  data = unbalanced_panel,
  index = c("fips_codes", "date"),
  effect = "individual",
  model = "within"
)

summary(fe.time_model_unbal)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = total_generation ~ pm_measure + factor(year) +
##     factor(month), data = unbalanced_panel, effect = "individual",
##     model = "within", index = c("fips_codes", "date"))
##
## Unbalanced Panel: n = 85, T = 1-96, N = 5309
##
## Residuals:
##      Min.   1st Qu.    Median   3rd Qu.      Max.
## -305739.1  -14041.2   -2677.9   12035.6  504249.4
##
## Coefficients:
##                   Estimate Std. Error t-value              Pr(>|t|)
## pm_measure         -357.95     125.03 -2.8630             0.0042132 **
## factor(year)2014 14662.77    2629.40  5.5765   0.00000002578216895 ***
## factor(year)2015 23491.09    2586.06  9.0837 < 0.00000000000000022 ***
## factor(year)2016 33623.43    2565.67 13.1051 < 0.00000000000000022 ***
## factor(year)2017 46404.01    2519.36 18.4190 < 0.00000000000000022 ***
## factor(year)2018 50671.25    2490.43 20.3464 < 0.00000000000000022 ***
## factor(year)2019 51974.07    2497.14 20.8134 < 0.00000000000000022 ***
## factor(year)2020 32470.12    3098.33 10.4799 < 0.00000000000000022 ***
## factor(month)2    3621.92    2955.12  1.2256             0.2203896
## factor(month)3   11544.25    2980.91  3.8727             0.0001089 ***
## factor(month)4   16719.92    2978.17  5.6142   0.00000002077421912 ***
## factor(month)5   22196.67    2972.09  7.4684   0.00000000000009468 ***
## factor(month)6   25670.27    2949.06  8.7046 < 0.00000000000000022 ***
## factor(month)7   25032.46    2934.49  8.5304 < 0.00000000000000022 ***
## factor(month)8   25119.12    2912.92  8.6233 < 0.00000000000000022 ***
## factor(month)9   20299.32    2920.34  6.9510   0.00000000000407324 ***
## factor(month)10  15923.67    2927.14  5.4400   0.00000005571128297 ***
## factor(month)11   8264.89    2898.35  2.8516             0.0043673 **
## factor(month)12   4543.60    2879.32  1.5780             0.1146235
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    11412000000000
## Residual Sum of Squares: 9621500000000
## R-Squared:      0.15691
## Adj. R-Squared: 0.14023
## F-statistic: 50.9859 on 19 and 5205 DF, p-value: < 0.00000000000000222
```

```r
#Random Effects
random_model_unbal <-
  plm(
    total_generation ~ pm_measure,
    data = unbalanced_panel,
    index = c("fips_codes", "date"),
    effect = "individual",
    model = "random"
  )

summary(random_model_unbal)
```

```
## Oneway (individual) effect Random Effect Model
##    (Swamy-Arora's transformation)
##
## Call:
## plm(formula = total_generation ~ pm_measure, data = unbalanced_panel,
##     effect = "individual", model = "random", index = c("fips_codes",
##         "date"))
##
## Unbalanced Panel: n = 85, T = 1-96, N = 5309
##
## Effects:
##                       var    std.dev share
## idiosyncratic 2173253165      46618 0.339
## individual    4231308457      65049 0.661
## theta:
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.4175  0.9153  0.9220  0.9145  0.9271  0.9271
##
## Residuals:
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -313457   -5103   -2657     458    1022  557042
##
## Coefficients:
##             Estimate Std. Error z-value    Pr(>|z|)
## (Intercept) 31106.97    7198.94  4.3210 0.000015529 ***
## pm_measure   -653.27     125.07 -5.2231 0.000000176 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    11542000000000
## Residual Sum of Squares: 11492000000000
## R-Squared:      0.0044735
## Adj. R-Squared: 0.0042859
## Chisq: 27.2805 on 1 DF, p-value: 0.00000017598
```

```r
#Comparing Models
pFtest(fe_model_unbal, ols_unbal)    #F test for individual effects, null: OLS preferred to FE
```

```
##
##  F test for individual effects
##
## data:  total_generation ~ pm_measure
## F = 129.65, df1 = 84, df2 = 5223, p-value < 0.00000000000000022
## alternative hypothesis: significant effects
```

```
phtest(fe_model_unbal, random_model_unbal)    #Hausman test, null: random effects preferred to F
E
```

```
##
##  Hausman Test
##
## data:  total_generation ~ pm_measure
## chisq = 1.8548, df = 1, p-value = 0.1732
## alternative hypothesis: one model is inconsistent
```

```
##Balanced Panel
balanced_panel <- read.csv('data/processed/balanced_panel_data.csv',
                        header = TRUE,
                        sep = ",")

#Ordinary Least Squares
ols_bal <- lm(total_generation ~ pm_measure, data = balanced_panel)

summary(ols_bal)
```

```
##
## Call:
## lm(formula = total_generation ~ pm_measure, data = balanced_panel)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -184912  -93954  -71321   67099  768051
##
## Coefficients:
##             Estimate Std. Error t value            Pr(>|t|)
## (Intercept)  85107.6     8786.7   9.686 <0.0000000000000002 ***
## pm_measure    2081.4      853.4   2.439              0.0149 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 141200 on 1138 degrees of freedom
## Multiple R-squared:  0.0052, Adjusted R-squared:  0.004326
## F-statistic: 5.948 on 1 and 1138 DF,  p-value: 0.01488
```

```r
#Fixed Effects
fe_model_bal <-
  plm(
    total_generation ~ pm_measure,
    data = balanced_panel,
    index = c("fips_codes", "date"),
    effect = "individual",
    model = "within"
  )

summary(fe_model_bal)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = total_generation ~ pm_measure, data = balanced_panel,
##     effect = "individual", model = "within", index = c("fips_codes",
##         "date"))
##
## Balanced Panel: n = 19, T = 60, N = 1140
##
## Residuals:
##      Min.   1st Qu.    Median   3rd Qu.      Max.
## -326052.3  -13287.7   -1644.1   12198.1  404639.0
##
## Coefficients:
##            Estimate Std. Error t-value  Pr(>|t|)
## pm_measure -1521.00     438.15 -3.4714 0.0005375 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    4762300000000
## Residual Sum of Squares: 4711600000000
## R-Squared:      0.010645
## Adj. R-Squared: -0.0061389
## F-statistic: 12.0505 on 1 and 1120 DF, p-value: 0.00053753
```

```r
#Year and Month Fixed Effects
fe.time_model_bal <- plm(
  total_generation ~ pm_measure +
    factor(year) + factor(month),
  data = balanced_panel,
  index = c("fips_codes", "date"),
  effect = "individual",
  model = "within"
)

summary(fe.time_model_bal)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = total_generation ~ pm_measure + factor(year) +
##     factor(month), data = balanced_panel, effect = "individual",
##     model = "within", index = c("fips_codes", "date"))
##
## Balanced Panel: n = 19, T = 60, N = 1140
##
## Residuals:
##       Min.    1st Qu.     Median    3rd Qu.       Max.
## -257993.97  -29739.98    -934.94   26644.65  347896.89
##
## Coefficients:
##                   Estimate Std. Error t-value              Pr(>|t|)
## pm_measure         -334.97     421.23 -0.7952               0.42665
## factor(year)2017 30182.74    5150.48  5.8602 0.0000000060964863587 ***
## factor(year)2018 42139.73    5154.13  8.1759 0.0000000000000007991 ***
## factor(year)2019 50283.88    5168.41  9.7291 < 0.00000000000000022 ***
## factor(year)2020 -4418.50    5183.62 -0.8524               0.39418
## factor(month)2   14017.38    8000.35  1.7521               0.08003 .
## factor(month)3   35550.59    8102.65  4.3875 0.0000125615271663756 ***
## factor(month)4   52153.54    8074.30  6.4592 0.0000000001574220259 ***
## factor(month)5   69745.81    8065.81  8.6471 < 0.00000000000000022 ***
## factor(month)6   78833.98    7999.41  9.8550 < 0.00000000000000022 ***
## factor(month)7   75558.39    7971.33  9.4788 < 0.00000000000000022 ***
## factor(month)8   72988.85    7986.39  9.1392 < 0.00000000000000022 ***
## factor(month)9   58531.85    7987.07  7.3283 0.0000000000004489280 ***
## factor(month)10  45319.64    7964.29  5.6904 0.0000000162285981622 ***
## factor(month)11  18172.30    7976.98  2.2781               0.02291 *
## factor(month)12   5898.46    7995.22  0.7377               0.46082
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    4762300000000
## Residual Sum of Squares: 3329200000000
## R-Squared:      0.30092
## Adj. R-Squared: 0.27941
## F-statistic: 29.7274 on 16 and 1105 DF, p-value: < 0.00000000000000222
```

```r
#Random Effects
random_model_bal <-
  plm(
    total_generation ~ pm_measure,
    data = balanced_panel,
    index = c("fips_codes", "date"),
    effect = "individual",
    model = "random"
  )

summary(random_model_bal)
```

```
## Oneway (individual) effect Random Effect Model
##    (Swamy-Arora's transformation)
##
## Call:
## plm(formula = total_generation ~ pm_measure, data = balanced_panel,
##     effect = "individual", model = "random", index = c("fips_codes",
##         "date"))
##
## Balanced Panel: n = 19, T = 60, N = 1140
##
## Effects:
##                        var     std.dev share
## idiosyncratic  4206775532       64860 0.206
## individual    16175552459      127183 0.794
## theta: 0.9343
##
## Residuals:
##      Min.   1st Qu.    Median   3rd Qu.      Max.
## -301983.6  -16528.1   -7537.2    7892.5  428724.7
##
## Coefficients:
##              Estimate Std. Error z-value   Pr(>|z|)
## (Intercept) 117548.31   29519.22  3.9821 0.00006831 ***
## pm_measure   -1501.58     438.08 -3.4277  0.0006088 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    4840200000000
## Residual Sum of Squares: 4790800000000
## R-Squared:      0.010219
## Adj. R-Squared: 0.009349
## Chisq: 11.749 on 1 DF, p-value: 0.00060876
```

```r
#Two-ways Fixed Effects
twoways_model_bal <-
  plm(
    total_generation ~ pm_measure,
    data = balanced_panel,
    index = c("fips_codes", "date"),
    effect = "twoways",
    model = "within"
  )

summary(twoways_model_bal)
```

```
## Twoways effects Within Model
##
## Call:
## plm(formula = total_generation ~ pm_measure, data = balanced_panel,
##     effect = "twoways", model = "within", index = c("fips_codes",
##         "date"))
##
## Balanced Panel: n = 19, T = 60, N = 1140
##
## Residuals:
##      Min.   1st Qu.    Median   3rd Qu.      Max.
## -264886.1  -31268.9    3140.9   29216.2  330938.7
##
## Coefficients:
##            Estimate Std. Error t-value Pr(>|t|)
## pm_measure  -262.91     500.57 -0.5252   0.5995
##
## Total Sum of Squares:    3191800000000
## Residual Sum of Squares: 3191000000000
## R-Squared:      0.00025994
## Adj. R-Squared: -0.073237
## F-statistic: 0.275866 on 1 and 1061 DF, p-value: 0.59953
```