590 HW1

Austin Turvy Derek Holste Mason Carhart

4/13/2021

```
# Load packages
library(pacman)
p_load(readr, data.table, dplyr, janitor, haven, here,
       tidyverse, skimr, lfe, stargazer, quantreg, hrbrthemes,
       tinytex, kableExtra, broom)
```

```
# Load data
af = read_dta(here("afghanistan_anonymized_data.dta"))
```

```
# Create nonoutlier column
af <- af %>% mutate(nonoutlier= ifelse((f07_num_ppl_hh_cnt > 20 & f07_observed == 1) |
                                    (f07_jeribs_cnt > 10 & f07_observed == 1)|
                                    (f07_num_sheep_cnt > 50 & f07_observed == 1)|
                                      (s08_num_ppl_hh_cnt > 20 & s08_observed == 1)|
                                      (s08_jeribs_cnt > 10 & s08_observed == 1) |
                                      (s08_num_sheep_cnt > 50 & s08_observed == 1)
                                      ,0,1))


# Tabulating nonoutliers
non_outlier_tab <- af %>% group_by(nonoutlier) %>%
    summarise(count = n(), percentage = n()/nrow(.))

non_outlier_tab
```

```
## # A tibble: 2 x 3
##   nonoutlier count percentage
##        <dbl> <int>      <dbl>
## 1          0    76     0.0421
## 2          1  1728     0.958
```

There are 76 non-outliers in the data given the specifications prior.

```
# Run girls regression table 4 column 1 including outliers
reg3 = felm(f07_formal_school ~ treatment + chagcharan | 0 | 0 | clustercode,
            data = af %>% filter(f07_girl_cnt == 1 &
                                 f07_observed == 1))

# Run boys regression table 4 column 1including outliers
reg4 = felm(f07_formal_school ~ treatment + chagcharan | 0 | 0 | clustercode,
            data = af %>% filter(f07_girl_cnt == 0 &
                                 f07_observed == 1))
```

```
# Run girls regression table 4 column 1 with controls
reg5 = felm(f07_formal_school ~ treatment + chagcharan + f07_heads_child_cnt +
              f07_age_cnt + f07_duration_village_cnt +
              f07_farsi_cnt + f07_tajik_cnt + f07_farmer_cnt +
              f07_age_head_cnt + f07_yrs_ed_head_cnt + f07_num_ppl_hh_cnt +
              f07_jeribs_cnt + f07_num_sheep_cnt +
              f07_nearest_scl| 0 | 0 | clustercode,
            data = af %>% filter(f07_girl_cnt == 1 &
                                 nonoutlier == 1 &
                                 f07_observed == 1))

# Run girls regression table 4 column 1 with controls
reg6 = felm(f07_formal_school ~ treatment + chagcharan + f07_heads_child_cnt +
              f07_age_cnt + f07_duration_village_cnt +
              f07_farsi_cnt + f07_tajik_cnt + f07_farmer_cnt +
              f07_age_head_cnt + f07_yrs_ed_head_cnt + f07_num_ppl_hh_cnt +
              f07_jeribs_cnt + f07_num_sheep_cnt +
              f07_nearest_scl| 0 | 0 | clustercode,
            data = af %>% filter(f07_girl_cnt == 0 &
                                 nonoutlier == 1 &
                                 f07_observed == 1))
```

Question 3:

```
# girls regression for column 3
reg7 = felm(f07_both_norma_total ~ treatment + chagcharan | 0 | 0 | clustercode,
            data = af %>% filter(f07_girl_cnt == 1 &
                                 nonoutlier == 1 &
                                 f07_test_observed == 1))

# boys regression for column 3
reg8 = felm(f07_both_norma_total ~ treatment + chagcharan | 0 | 0 | clustercode,
            data = af %>% filter(f07_girl_cnt == 0 &
                                 nonoutlier == 1 &
                                 f07_test_observed == 1))
```

Question 4:

```
# girls regression column 4
reg9 = felm(f07_both_norma_total ~ treatment + chagcharan + f07_heads_child_cnt +
              f07_age_cnt + f07_duration_village_cnt +
              f07_farsi_cnt + f07_tajik_cnt + f07_farmer_cnt +
              f07_age_head_cnt + f07_yrs_ed_head_cnt + f07_num_ppl_hh_cnt +
              f07_jeribs_cnt + f07_num_sheep_cnt +
              f07_nearest_scl| 0 | 0 | clustercode,
            data = af %>% filter(f07_girl_cnt == 1 &
                                 nonoutlier == 1 &
                                 f07_test_observed == 1))
```

```
# boys regression column 4
reg10 = felm(f07_both_norma_total ~ treatment + chagcharan + f07_heads_child_cnt +
               f07_age_cnt + f07_duration_village_cnt +
               f07_farsi_cnt + f07_tajik_cnt + f07_farmer_cnt +
               f07_age_head_cnt + f07_yrs_ed_head_cnt + f07_num_ppl_hh_cnt +
               f07_jeribs_cnt + f07_num_sheep_cnt +
               f07_nearest_scl| 0 | 0 | clustercode,
             data = af %>% filter(f07_girl_cnt == 0 &
                                  nonoutlier == 1 &
                                  f07_test_observed == 1))
```

Question 4b:

```
# girls regression for column 6
reg11 = felm(s08_both_norma_total ~ treatment + chagcharan | 0 | 0 | clustercode,
             data = af %>% filter(s08_girls_cnt == 1 &
                                  nonoutlier == 1 &
                                  s08_test_observed == 1))

# boys regression for column 6
reg12 = felm(s08_both_norma_total ~ treatment + chagcharan | 0 | 0 | clustercode,
             data = af %>% filter(s08_girls_cnt == 0 &
                                  nonoutlier == 1 &
                                  s08_test_observed == 1))
# girls regression for column 7
reg13 = felm(s08_both_norma_total ~ treatment + chagcharan +
                s08_heads_child_cnt + s08_age_cnt +
                s08_duration_village_cnt + s08_farsi_cnt + s08_tajik_cnt +
                s08_farmer_cnt + s08_age_head_cnt + s08_yrs_ed_head_cnt +
                s08_num_ppl_hh_cnt + s08_jeribs_cnt + s08_num_sheep_cnt +
                s08_nearest_scl | 0 | 0 | clustercode,
             data = af %>% filter(s08_girls_cnt == 1 &
                                  nonoutlier == 1 &
                                  s08_test_observed == 1))

# boys regression for column 7
reg14 = felm(s08_both_norma_total ~ treatment + chagcharan +
                s08_heads_child_cnt  + s08_age_cnt +
                s08_duration_village_cnt + s08_farsi_cnt + s08_tajik_cnt +
                s08_farmer_cnt + s08_age_head_cnt + s08_yrs_ed_head_cnt +
                s08_num_ppl_hh_cnt + s08_jeribs_cnt + s08_num_sheep_cnt +
                s08_nearest_scl | 0 | 0 | clustercode,
             data = af %>% filter(s08_girls_cnt == 0 &
                                  nonoutlier == 1 &
                                  s08_test_observed == 1))
```

```
# Getting mean values for column 1
col_1 = af %>%
  filter(treatment == 1 & nonoutlier == 1 & f07_observed == 1) %>%
  select(
```

```
    c(f07_heads_child_cnt, f07_girl_cnt, f07_age_cnt, f07_duration_village_cnt,
                  f07_farsi_cnt,  f07_tajik_cnt,  f07_farmer_cnt,
                  f07_age_head_cnt, f07_yrs_ed_head_cnt, f07_num_ppl_hh_cnt, f07_jeribs_cnt,
                  f07_num_sheep_cnt, f07_nearest_scl)
  ) %>%
  mutate(
  across(.cols = c(f07_heads_child_cnt, f07_girl_cnt, f07_age_cnt, f07_duration_village_cnt,
                  f07_farsi_cnt,  f07_tajik_cnt,  f07_farmer_cnt,
                  f07_age_head_cnt, f07_yrs_ed_head_cnt, f07_num_ppl_hh_cnt, f07_jeribs_cnt,
                  f07_num_sheep_cnt, f07_nearest_scl), mean, na.rm = TRUE)
  ) %>% head(., 1) %>% unlist(.)

# Getting mean values for column 2
col_2 = af %>%
  filter(treatment == 0 & nonoutlier == 1 & f07_observed == 1) %>%
  select(
    c(f07_heads_child_cnt, f07_girl_cnt, f07_age_cnt, f07_duration_village_cnt,
                  f07_farsi_cnt,  f07_tajik_cnt,  f07_farmer_cnt,
                  f07_age_head_cnt, f07_yrs_ed_head_cnt, f07_num_ppl_hh_cnt, f07_jeribs_cnt,
                  f07_num_sheep_cnt, f07_nearest_scl)
  ) %>%
  mutate(
  across(.cols = c(f07_heads_child_cnt, f07_girl_cnt, f07_age_cnt, f07_duration_village_cnt,
                  f07_farsi_cnt,  f07_tajik_cnt,  f07_farmer_cnt,
                  f07_age_head_cnt, f07_yrs_ed_head_cnt, f07_num_ppl_hh_cnt, f07_jeribs_cnt,
                  f07_num_sheep_cnt, f07_nearest_scl), mean, na.rm = TRUE)
  ) %>% head(., 1) %>% unlist(.)



# Create regression functions for column 3
reg_function_coef = function(x){
    reg = felm(x ~ treatment | 0 | 0 | clustercode,
               data = af %>% filter(nonoutlier == 1 & f07_observed == 1))
    coef = coef(reg)[2]
    return(coef)
}


# Regression function for SE
reg_function_se = function(x){
  reg = felm(x ~ treatment | 0 | 0 | clustercode,
             data = af %>% filter(nonoutlier == 1 & f07_observed == 1))
  se = sqrt(diag(vcov(reg)))[2]
  return(se)
}



# Obtain column 3 coefficients
col_3_coefs = af %>%
  filter(nonoutlier == 1 & f07_observed == 1) %>%
  select(
```

```
    c(f07_heads_child_cnt, f07_girl_cnt, f07_age_cnt, f07_duration_village_cnt,
                  f07_farsi_cnt,  f07_tajik_cnt,  f07_farmer_cnt,
                  f07_age_head_cnt, f07_yrs_ed_head_cnt, f07_num_ppl_hh_cnt, f07_jeribs_cnt,
                  f07_num_sheep_cnt, f07_nearest_scl)
  ) %>%
  mutate(
  across(.cols = c(f07_heads_child_cnt, f07_girl_cnt, f07_age_cnt, f07_duration_village_cnt,
                  f07_farsi_cnt,  f07_tajik_cnt,  f07_farmer_cnt,
                  f07_age_head_cnt, f07_yrs_ed_head_cnt, f07_num_ppl_hh_cnt, f07_jeribs_cnt,
                  f07_num_sheep_cnt, f07_nearest_scl), reg_function_coef)
  ) %>% head(., 1) %>% unlist(.)

# Obtain column 3 SE
col_3_se = af %>%
  filter(nonoutlier == 1 & f07_observed == 1) %>%
  select(
    c(f07_heads_child_cnt, f07_girl_cnt, f07_age_cnt, f07_duration_village_cnt,
                  f07_farsi_cnt,  f07_tajik_cnt,  f07_farmer_cnt,
                  f07_age_head_cnt, f07_yrs_ed_head_cnt, f07_num_ppl_hh_cnt, f07_jeribs_cnt,
                  f07_num_sheep_cnt, f07_nearest_scl)
  ) %>%
  mutate(
  across(.cols = c(f07_heads_child_cnt, f07_girl_cnt, f07_age_cnt, f07_duration_village_cnt,
                  f07_farsi_cnt,  f07_tajik_cnt,  f07_farmer_cnt,
                  f07_age_head_cnt, f07_yrs_ed_head_cnt, f07_num_ppl_hh_cnt, f07_jeribs_cnt,
                  f07_num_sheep_cnt, f07_nearest_scl), reg_function_se)
  ) %>% head(., 1) %>% unlist(.)

# Regression for column 7
col_7 = felm(f07_formal_school ~ f07_heads_child_cnt +
               f07_girl_cnt + f07_age_cnt + f07_duration_village_cnt +
               f07_farsi_cnt + f07_tajik_cnt + f07_farmer_cnt +
               f07_age_head_cnt + f07_yrs_ed_head_cnt + f07_num_ppl_hh_cnt +
               f07_jeribs_cnt + f07_num_sheep_cnt + f07_nearest_scl
             | 0 | 0 | clustercode,
             data = af %>%
               filter(nonoutlier == 1 & f07_observed == 1 & treatment == 0))

# Regression for column 8
col_8 = felm(f07_both_norma_total ~ f07_heads_child_cnt +
               f07_girl_cnt + f07_age_cnt + f07_duration_village_cnt +
               f07_farsi_cnt + f07_tajik_cnt + f07_farmer_cnt +
               f07_age_head_cnt + f07_yrs_ed_head_cnt + f07_num_ppl_hh_cnt +
               f07_jeribs_cnt + f07_num_sheep_cnt + f07_nearest_scl
             | 0 | 0 | clustercode,
             data = af %>%
               filter(nonoutlier == 1 & f07_observed == 1 &
                        f07_test_observed == 1 & treatment == 0))
```

```
# further analysis regression
reg_fa = felm(f07_formal_school ~ f07_girl_cnt +
         f07_nearest_scl + f07_girl_cnt*f07_nearest_scl +
         I(f07_nearest_scl^2) | 0 | 0 | clustercode,
```

```
        data = af %>% filter(treatment==0 &
                             f07_observed == 1 &
                             nonoutlier == 1))
```

```
af = af %>% mutate(dist_near = if_else(f07_nearest_scl < 2, 1, 0),
               dist_close = if_else(f07_nearest_scl >= 2 & f07_nearest_scl < 3, 1, 0),
               dist_nclose = if_else(f07_nearest_scl >= 3 & f07_nearest_scl < 4, 1, 0),
               dist_far = if_else(f07_nearest_scl >= 4, 1, 0))

# Regression for near individuals
reg_near_g = felm(f07_formal_school ~ treatment +
                    chagcharan | 0 | 0 | clustercode,
         data = af %>% filter(
                             f07_observed == 1 &
                             nonoutlier == 1 &
                             f07_girl_cnt == 1 &
                             dist_near == 1
                             ))

# Regression for close individuals
reg_close_g = felm(f07_formal_school ~ treatment +
                    chagcharan | 0 | 0 | clustercode,
         data = af %>% filter(
                             f07_observed == 1 &
                             nonoutlier == 1 &
                             f07_girl_cnt == 1 &
                             dist_close == 1
                             ))

# Regression for not close individuals
reg_nclose_g = felm(f07_formal_school ~ treatment +
                    chagcharan | 0 | 0 | clustercode,
         data = af %>% filter(
                             f07_observed == 1 &
                             nonoutlier == 1 &
                             f07_girl_cnt == 1 &
                             dist_nclose == 1
                             ))

# Regression for far individuals
reg_far_g = felm(f07_formal_school ~ treatment +
                    chagcharan | 0 | 0 | clustercode,
         data = af %>% filter(
                             f07_observed == 1 &
                             nonoutlier == 1 &
                             f07_girl_cnt == 1 &
                             dist_far == 1
                             ))

# Regression for near individuals
reg_near_b= felm(f07_formal_school ~ treatment +
                    chagcharan | 0 | 0 | clustercode,
         data = af %>% filter(
```

```
                              f07_observed == 1 &
                              nonoutlier == 1 &
                              f07_girl_cnt == 0 &
                              dist_near == 1
                              ))

# Regression for close individuals
reg_close_b = felm(f07_formal_school ~ treatment +
                   chagcharan | 0 | 0 | clustercode,
         data = af %>% filter(
                              f07_observed == 1 &
                              nonoutlier == 1 &
                              f07_girl_cnt == 0 &
                              dist_close == 1
                              ))

# Regression for not close individuals
reg_nclose_b = felm(f07_formal_school ~ treatment +
                   chagcharan | 0 | 0 | clustercode,
         data = af %>% filter(
                              f07_observed == 1 &
                              nonoutlier == 1 &
                              f07_girl_cnt == 0 &
                              dist_nclose == 1
                              ))

# Regression for far individuals
reg_far_b = felm(f07_formal_school ~ treatment +
                   chagcharan | 0 | 0 | clustercode,
         data = af %>% filter(
                              f07_observed == 1 &
                              nonoutlier == 1 &
                              f07_girl_cnt == 0 &
                              dist_far == 1
                              ))
```

```
# Regression for near individuals
reg_test_near_g = felm(f07_both_norma_total ~ treatment +
                   chagcharan | 0 | 0 | clustercode,
         data = af %>% filter(
                              f07_observed == 1 &
                              nonoutlier == 1 &
                              f07_girl_cnt == 1 &
                              dist_near == 1
                              ))

# Regression for close individuals
reg_test_close_g = felm(f07_both_norma_total ~ treatment +
                   chagcharan | 0 | 0 | clustercode,
         data = af %>% filter(
                              f07_observed == 1 &
                              nonoutlier == 1 &
                              f07_girl_cnt == 1 &
```

```
                              dist_close == 1
                              ))

# Regression for not close individuals
reg_test_nclose_g = felm(f07_both_norma_total ~ treatment +
                    chagcharan | 0 | 0 | clustercode,
          data = af %>% filter(
                              f07_observed == 1 &
                              nonoutlier == 1 &
                              f07_girl_cnt == 1 &
                              dist_nclose == 1
                              ))

# Regression for far individuals
reg_test_far_g = felm(f07_both_norma_total ~ treatment +
                    chagcharan | 0 | 0 | clustercode,
          data = af %>% filter(
                              f07_observed == 1 &
                              nonoutlier == 1 &
                              f07_girl_cnt == 1 &
                              dist_far == 1
                              ))

# Regression for near individuals
reg_test_near_b= felm(f07_both_norma_total ~ treatment +
                    chagcharan | 0 | 0 | clustercode,
          data = af %>% filter(
                              f07_observed == 1 &
                              nonoutlier == 1 &
                              f07_girl_cnt == 0 &
                              dist_near == 1
                              ))

# Regression for close individuals
reg_test_close_b = felm(f07_both_norma_total ~ treatment +
                    chagcharan | 0 | 0 | clustercode,
          data = af %>% filter(
                              f07_observed == 1 &
                              nonoutlier == 1 &
                              f07_girl_cnt == 0 &
                              dist_close == 1
                              ))

# Regression for not close individuals
reg_test_nclose_b = felm(f07_both_norma_total ~ treatment +
                    chagcharan | 0 | 0 | clustercode,
          data = af %>% filter(
                              f07_observed == 1 &
                              nonoutlier == 1 &
                              f07_girl_cnt == 0 &
                              dist_nclose == 1
                              ))
```

```
# Regression for far individuals
reg_test_far_b = felm(f07_both_norma_total ~ treatment +
                    chagcharan | 0 | 0 | clustercode,
          data = af %>% filter(
                              f07_observed == 1 &
                              nonoutlier == 1 &
                              f07_girl_cnt == 0 &
                              dist_far == 1
                              ))
```

```
# Table 4.1
stargazer(reg1, reg2, reg5, reg6, reg7, reg8, reg9, reg10,
          keep.stat = c('n', 'rsq'),
          column.labels = c("      ", "      ", "      ", "      ",
                            "      ", "      ", "      ", "      "),
          title = "Table 4: Treatment Effects by Gender",
          covariate.labels = c('Treatment', "chagcharan",
                               "Household head's child",  "Age",
                               "Years family in village","Farsi",
                               "Tajik","Farmers","Age of household head",
                               "Years of education of household head","Number of people in Household",
                               "Jeribs of land","Number of sheep",
                               "Distance to nearest formal school"),
          dep.var.labels = c("Formaly Enroled", "Fall Test Scores"),
          #dep.var.labels = c(""),
          type = 'latex', header = F, float = TRUE,
          font.size = "small",
          column.sep.width = "-15pt",
          omit.stat=c("f", "ser"),
          notes = c("S.E Clustered by village"), notes.append = FALSE)
```

```
# Table 4b
stargazer(reg11, reg12, reg13, reg14,
          keep.stat = c('n', 'rsq'),
          title = "Table 4b: Treatment Effects by Gender",
          covariate.labels = c('Treatment', "chagcharan"),
          #dep.var.labels = c(""),
          omit = c("s08_heads_child_cnt", "s08_girl_cnt", "s08_age_cnt",
                   "s08_duration_village_cnt", "s08_farsi_cnt",
                   "s08_tajik_cnt", "s08_farmer_cnt", "s08_age_head_cnt",
                   "s08_yrs_ed_head_cnt", "s08_num_ppl_hh_cnt",
                   "s08_jeribs_cnt", "s08_num_sheep_cnt", "s08_nearest_scl",
                   "chagcharan"),
          type = 'latex', header = F, float = TRUE,
          notes =c("S.E Clustered by village"), notes.append = FALSE)
```

```
# Table 2
stargazer(col_7, col_8,
          keep.stat = c('n', 'rsq'),
          title = "Table 2: Demographic Characteristics By Research Groups",
          covariate.labels = c("Household head's child", "Girl",
                               "Age","Years family in village","Farsi",
```

Table 1: Table 4: Treatment Effects by Gender

	Dependent variable:							
	Formaly Enroled				Fall Test Scores			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.521**	0.371***	0.515***	0.347***	0.691***	0.424***	0.654***	0.400***
	(0.091)	(0.101)	(0.082)	(0.094)	(0.130)	(0.107)	(0.123)	(0.091)
chagcharan	0.176**	0.081	0.154*	0.086	0.282**	0.113	0.275**	0.118
	(0.085)	(0.096)	(0.082)	(0.088)	(0.123)	(0.104)	(0.117)	(0.075)
Household head's child			−0.043	0.022			−0.156	0.125
			(0.051)	(0.051)			(0.168)	(0.098)
Age			0.037**	0.065***			0.243***	0.367***
			(0.016)	(0.019)			(0.030)	(0.021)
Years family in village			−0.001	−0.0001			−0.003	−0.003
			(0.001)	(0.002)			(0.002)	(0.002)
Farsi			−0.082	0.019			−0.115	0.094
			(0.051)	(0.063)			(0.097)	(0.141)
Tajik			−0.063	0.077***			−0.005	0.173***
			(0.068)	(0.030)			(0.079)	(0.052)
Farmers			−0.017	−0.082**			0.0002	−0.082
			(0.035)	(0.039)			(0.076)	(0.113)
Age of household head			−0.00004	−0.003			−0.001	0.005
			(0.002)	(0.002)			(0.004)	(0.003)
Years of education of household head			0.003	0.002			0.026**	0.048***
			(0.005)	(0.004)			(0.011)	(0.011)
Number of people in Household			0.007	−0.001			0.007	−0.001
			(0.006)	(0.009)			(0.007)	(0.014)
Jeribs of land			−0.009	0.016			0.016	0.018
			(0.013)	(0.011)			(0.030)	(0.032)
Number of sheep			0.006	0.004			0.008	0.013**
			(0.004)	(0.003)			(0.008)	(0.005)
Distance to nearest formal school			−0.007	−0.059**			0.001	−0.070
			(0.022)	(0.024)			(0.050)	(0.049)
Constant	0.084*	0.325***	−0.168	0.082	−0.486**	0.267***	−2.421***	−3.017***
	(0.050)	(0.058)	(0.249)	(0.239)	(0.097)	(0.091)	(0.600)	(0.204)
Observations	693	797	693	797	667	707	667	707
R^2	0.339	0.164	0.371	0.245	0.167	0.045	0.357	0.404

Note: S.E Clustered by village