590 HW1

Austin Turvy Derek Holste Mason Carhart

4/13/2021

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

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

1. An outlier observation is someone who had more than 50 sheep or goats owned by the household in 2007 and in the fall of 2007 or had more than 10 jeribs of land owned by the household in the fall of 2007 or had more than 20 poeple in the houdcount in fall of 2007 or a person that has number of people in a household greater than 20 and are observed in summer of 2008. Or if number of jerribs of land counted is more than 10 and are observed in summer of 2008. Lastly if the number of sheep and goats is above 50 and they are observed in summer of 2008.

```r
# 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.

```r
# 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))
```

```r
# 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:

```r
# 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 +
```

```r
               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))
```

```r
# 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_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_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_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_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
```

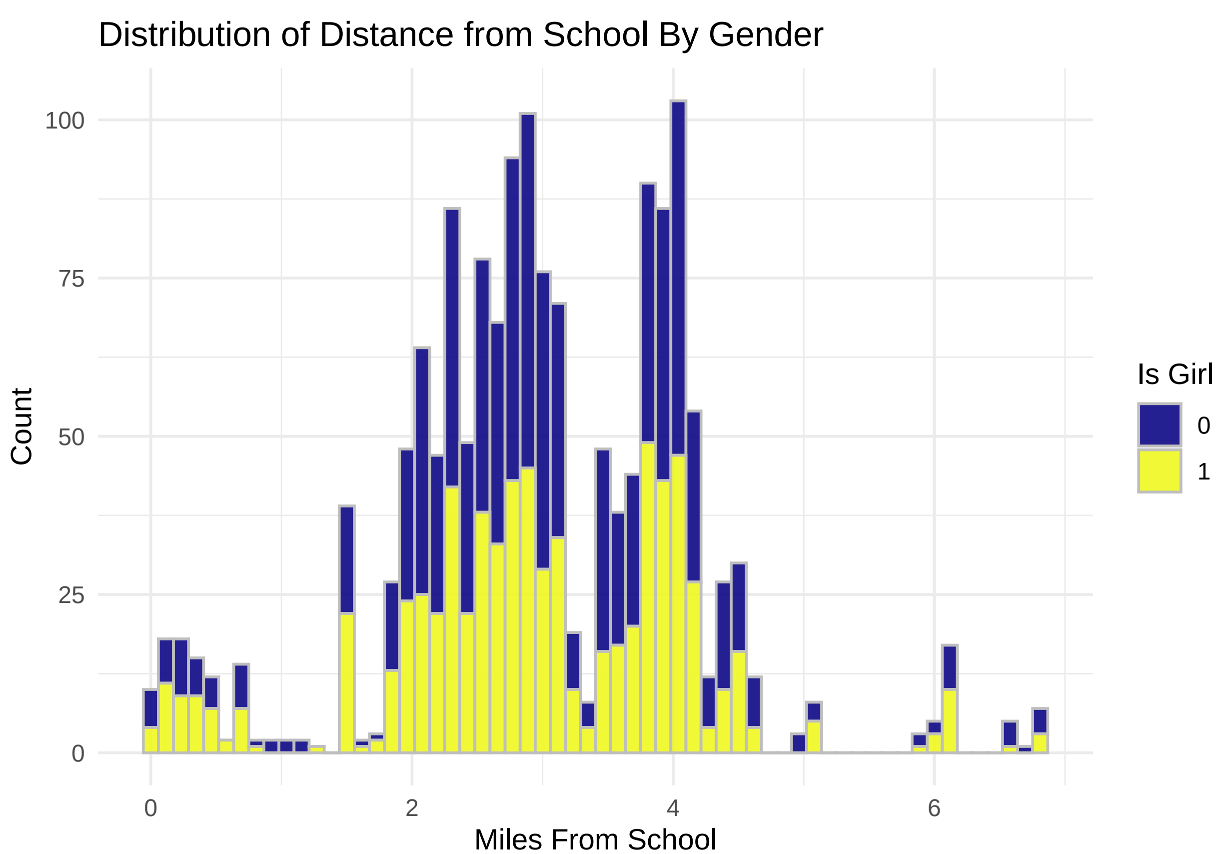
```r
reg_function_se = function(x){
  reg = felm(x ~ treatment | 0 | 0 | clustercode,
             data = af %>% filter(nonoutlier == 1 & f07_observed == 1))
  se = sqrt(diag(vcov(reg)))[2]
  return(se)
}



# 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_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_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_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_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))
```

```
ggplot(af, aes(x = f07_nearest_scl, fill = as.factor(f07_girl_cnt))) +
  geom_histogram(color = "grey", bins = 60, alpha = .9) +
  scale_fill_viridis_d(option = "C") +
  theme_minimal() +
  labs(x = "Miles From School", y = "Count", fill = "Is Girl",
       title = "Distribution of Distance from School By Gender")
```

```
## Warning: Removed 233 rows containing non-finite values (stat_bin).
```



```r
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
```

```r
                              ))

# 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
                              ))
```

```r
# 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
```

```r
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(
```

```r
                              f07_observed == 1 &
                              nonoutlier == 1 &
                              f07_girl_cnt == 0 &
                              dist_far == 1
                              ))
```

```r
# 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",
                               "Jeribs of land","Number of sheep",
                               "Distance to nearest formal school",
                               "Nearest 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)
```

```r
# 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)
```

```r
# Table 2
stargazer(col_1, col_2, col_3_coefs, col_3_se, col_7, col_8,
          keep.stat = c('n', 'rsq'),
          title = "Table 2: Treatment Effects by Gender",
          covariate.labels = c("Household head's child", "Girl",
                               "Age","Years family in village","Farsi",
                               "Tajik","Farmers","Age of household head",
                               "Years of education of household head",
                               "Jeribs of land","Number of sheep",
```

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)
Jeribs of land			0.007	−0.001			0.007	−0.001
			(0.006)	(0.009)			(0.007)	(0.014)
Number of sheep			−0.009	0.016			0.016	0.018
			(0.013)	(0.011)			(0.030)	(0.032)
Distance to nearest formal school			0.006	0.004			0.008	0.013^{**}
			(0.004)	(0.003)			(0.008)	(0.005)
Nearest 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

Table 2: Table 4b: Treatment Effects by Gender

	Dependent variable:			
	s08_both_norma_total			
	(1)	(2)	(3)	(4)
Treatment	0.735***	0.380***	0.661***	0.413***
	(0.093)	(0.129)	(0.090)	(0.099)
chagcharan	−0.458***	0.281***	−3.052***	−3.144***
	(0.066)	(0.089)	(0.460)	(0.302)
Observations	689	712	687	709
R^2	0.165	0.042	0.378	0.410
Note:			S.E Clustered by village	

```
                          "Distance to nearest formal school"),
          #dep.var.labels = c(""),
          type = 'latex', header = F, float = TRUE,
          notes =c("S.E Clustered by village"), flip = TRUE)
```

Table 3: Table 2: Treatment Effects by Gender

f07_heads_child_cnt	0.935
f07_girl_cnt	0.474
f07_age_cnt	8.321
f07_duration_village_cnt	30.302
f07_farsi_cnt	0.208
f07_tajik_cnt	0.243
f07_farmer_cnt	0.717
f07_age_head_cnt	40.142
f07_yrs_ed_head_cnt	3.315
f07_jeribs_cnt	1.345
f07_num_sheep_cnt	7.552
f07_nearest_scl	2.910

S.E Clustered by village

```
stargazer(reg_near_g,reg_close_g,reg_nclose_g, reg_far_g,
          reg_near_b,reg_close_b,reg_nclose_b, reg_far_b,
          keep.stat = c('n', 'rsq'),
          title = "Further Analysis: Distance Effect on Enrollment",
          covariate.labels = c('Treatment', "chagcharan"),
          column.labels = c("Short", "Med-Short", "Med-Long",
                            "Long", "Short", "Med-Short", "Med-Long", "Long"),
          dep.var.labels = c("Whether or not the Individual is in a Formal School"),
          type = 'latex', header = F, float = TRUE,
          column.sep.width = "-10pt",
          notes =c("S.E Clustered by village 1-4 (Female), 5-8 (Male).",
          "Both groups have increasing distance as the number increases."))
```

Table 4: Table 2: Treatment Effects by Gender

f07_heads_child_cnt	0.911
f07_girl_cnt	0.455
f07_age_cnt	8.312
f07_duration_village_cnt	27.594
f07_farsi_cnt	0.209
f07_tajik_cnt	0.208
f07_farmer_cnt	0.727
f07_age_head_cnt	39.970
f07_yrs_ed_head_cnt	3.076
f07_jeribs_cnt	1.274
f07_num_sheep_cnt	5.631
f07_nearest_scl	3.163

S.E Clustered by village

Table 5: Table 2: Treatment Effects by Gender

f07_heads_child_cnt.treatment	0.024
f07_girl_cnt.treatment	0.020
f07_age_cnt.treatment	0.009
f07_duration_village_cnt.treatment	2.709
f07_farsi_cnt.treatment	-0.001
f07_tajik_cnt.treatment	0.035
f07_farmer_cnt.treatment	-0.010
f07_age_head_cnt.treatment	0.172
f07_yrs_ed_head_cnt.treatment	0.239
f07_jeribs_cnt.treatment	0.071
f07_num_sheep_cnt.treatment	1.921
f07_nearest_scl.treatment	-0.253

S.E Clustered by village

Table 6: Table 2: Treatment Effects by Gender

f07_heads_child_cnt.treatment	0.015
f07_girl_cnt.treatment	0.020
f07_age_cnt.treatment	0.040
f07_duration_village_cnt.treatment	1.605
f07_farsi_cnt.treatment	0.054
f07_tajik_cnt.treatment	0.049
f07_farmer_cnt.treatment	0.034
f07_age_head_cnt.treatment	1.101
f07_yrs_ed_head_cnt.treatment	0.442
f07_jeribs_cnt.treatment	0.107
f07_num_sheep_cnt.treatment	1.504
f07_nearest_scl.treatment	0.349

S.E Clustered by village

Table 7: Table 2: Treatment Effects by Gender

	Dependent variable:	
	f07_formal_school	f07_both_norma_total
	(1)	(2)
Household head's child	0.038	−0.090
	(0.061)	(0.113)
Girl	-0.208^{***}	-0.682^{***}
	(0.080)	(0.102)
Age	0.046^{***}	0.287^{***}
	(0.017)	(0.018)
Years family in village	−0.001	-0.005^{**}
	(0.001)	(0.002)
Farsi	−0.039	0.074
	(0.074)	(0.119)
Tajik	−0.006	0.042
	(0.078)	(0.065)
Farmers	−0.050	−0.020
	(0.081)	(0.120)
Age of household head	−0.004	-0.004^{**}
	(0.003)	(0.002)
Years of education of household head	0.001	0.036^{***}
	(0.006)	(0.007)
Jeribs of land	0.004	−0.004
	(0.006)	(0.011)
Number of sheep	0.022^{***}	0.056^{***}
	(0.008)	(0.011)
Distance to nearest formal school	0.010^{***}	0.016^{**}
	(0.002)	(0.007)
f07_nearest_scl	−0.060	-0.089^{**}
	(0.043)	(0.037)
Constant	0.246	-1.649^{***}
	(0.324)	(0.369)
Observations	708	653
R^2	0.155	0.401
Note:		*p<0.1; **p<0.05; ***p<0.01
		S.E Clustered by village

Table 8: Further Analysis: Distance Effect on Enrollment

	Dependent variable:							
	Whether or not the Individual is in a Formal School							
	Short	Med-Short	Med-Long	Long	Short	Med-Short	Med-Long	Long
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.448***	0.528***	0.696***	0.481***	0.319*	0.447***	0.235**	0.504***
	(0.122)	(0.140)	(0.073)	(0.066)	(0.190)	(0.160)	(0.093)	(0.092)
chagcharan	0.404***	0.192	−0.156	0.368***	0.377**	0.028	0.058	0.153
	(0.069)	(0.117)	(0.126)	(0.064)	(0.150)	(0.152)	(0.111)	(0.094)
Constant	0.128	0.022	0.072	0.049	0.357*	0.332***	0.414***	0.094
	(0.124)	(0.026)	(0.067)	(0.046)	(0.203)	(0.063)	(0.060)	(0.084)
Observations	94	272	204	123	88	328	232	149
R^2	0.173	0.400	0.403	0.420	0.225	0.213	0.068	0.310

Note: *p<0.1; **p<0.05; ***p<0.01
S.E Clustered by village 1-4 (Female), 5-8 (Male).
Both groups have increasing distance as the number increases.

```
stargazer(reg_test_near_g, reg_test_close_g, reg_test_nclose_g,
          reg_test_far_g, reg_test_near_b, reg_test_close_b,
          reg_test_nclose_b,reg_test_far_b ,
          title = "Further Analysis: Distance Effect on Test Scores",
          keep.stat = c('n', 'rsq'),
          covariate.labels = c('Treatment', "chagcharan"),
          column.labels = c("Short", "Med-Short", "Med-Long", "Long",
                            "Short", "Med-Short", "Med-Long", "Long"),
          dep.var.labels = c("Total Normalized Test Score, Fall 2007"),
          type = 'latex', header = F, float = TRUE,
          column.sep.width = "-10pt",
          notes =c("S.E Clustered by village 1-4 (Female), 5-8 (Male).",
          "Both groups have increasing distance as the number increases."))
```

#**Written Questions:**

##Table 4 Question 5: Both of our coefficient estimates were very similar as the treatment for girls was off by 0.004 and the treatment for men is 0.001. The standard error for women was exactly the same while the standard error for men is unknown from table four.

##Table 4 Question 6: The coefficients on all estimates are minimuly different with the largest change being the difference between the treated groups in column 2 and 4, (the difference is .018). The standard errors are also negligibly different.

##Table 4 Column 2 Question 6: Like in table 4, both treatments are statically significant. For girls, the only statistically significant control is the child's age in the fall of 2007 (5%). For The males there were more statistically significant controls. The age of the child, (1% level) if the child's family speaks Tajik, (5% level) if the head of household is a farmer or not, (10% level) and lastly the nearest schools distance, (1% level) . For both genders the child's age was significant and positive which means that as the child grows older, they are more likely to get a better test score.

Table 9: Further Analysis: Distance Effect on Test Scores

	Dependent variable:							
	Total Normalized Test Score, Fall 2007							
	Short	Med-Short	Med-Long	Long	Short	Med-Short	Med-Long	Long
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	1.126***	0.789***	0.853***	0.563***	0.103*	0.512***	0.202*	0.895***
	(0.262)	(0.233)	(0.198)	(0.065)	(0.061)	(0.159)	(0.116)	(0.152)
chagcharan	0.773***	0.360*	−0.277	0.445***	−0.027	0.146	−0.111	0.371***
	(0.248)	(0.184)	(0.197)	(0.066)	(0.053)	(0.114)	(0.194)	(0.121)
Constant	−1.169***	−0.496***	−0.399**	−0.559***	0.396***	0.281***	0.497***	−0.351***
	(0.295)	(0.134)	(0.177)	(0.070)	(0.048)	(0.040)	(0.121)	(0.080)
Observations	91	259	196	121	81	290	198	138
R^2	0.217	0.212	0.174	0.193	0.002	0.069	0.006	0.221

Note: *p<0.1; **p<0.05; ***p<0.01
S.E Clustered by village 1-4 (Female), 5-8 (Male).
Both groups have increasing distance as the number increases.

##Further Analysis: To start looking at the relationship at the between school we regressed formal enrolled on the distance to the nearest school, whether the child was a girl, an interaction between the two and a variables for distance squared to control for any quadratic patterns. Surprisingly the only statisally significant variable in that regression was whether or not the child was a girl. This means that being farther away does not affect formal enrollment and distance does not affect girls more than boys. The regression output does implicate that women do attend school less than their boy counterparts. Now when re-looking at formal enrollment based on four distance types, short medium short, medium long, and long, we can see that for girl, the amount of distance does not change the fact that treatment is significantly increasing enrollment or test scores. More specifically as distance increases, the effect of the treatment on tests for girls decreases while it increases for men. Meaning that the treatment is most effective for girls who live close and boys who live farther away. Interestingly the treatment for both boys and girls are mostly similar except when the schools were a medium long distance away meaning that distance does not play a significant role in how the treatment will affect the sexes. Overall it seems as though distance is a more important variable when looking at the test scores for boys and girls and not as much for their attendance rates.