590 HW1

Austin Turvy Derek Holste Mason Carhart

4/13/2021

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.

```
mytable <- table(af$nonoutlier,af$nonoutlier)
mytable</pre>
```

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

af = read_dta(here("afghanistan_anonymized_data.dta"))

Table 1: Regression Results

	Dependent variable:				
	Girl Enrol	lled in Formal School Fall 2007			
	(1)	(2)			
Treatment	0.517***	0.372^{***}			
	(0.091)	(0.094)			
chagcharan	0.178**	0.086			
	(0.085)	(0.090)			
Constant	0.087*	0.328***			
	(0.051)	(0.057)			
Observations	667	707			
\mathbb{R}^2	0.335	0.167			
Note:		*p<0.1; **p<0.05; ***p<0.01			
		S.E Clustered by village			

f07_test_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_test_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_test_observed == 1))
```

Table 2: Regression Results

	Dependent variable:				
		Girls			
	(1)	(2)	(3)		
Treatment	0.517***	0.372***	0.511***		
	(0.091)	(0.094)	(0.083)		
chagcharan	0.178**	0.086	0.156*		
	(0.085)	(0.090)	(0.082)		
Constant	0.087^{*}	0.328***	-0.159		
	(0.051)	(0.057)	(0.256)		
Observations	667	707	667		
\mathbb{R}^2	0.335	0.167	0.366		
Note:	*p<0.1; **p<0.05; ***p<0.01 S.E Clustered by village				

Question 3:

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,
```

```
# 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
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) %>%
    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,
```

Table 3: Regression Results

$f07_heads_child_cnt$	$f07_girl_cnt$	$f07_age_cnt$	$f07_duration_village_cnt$	$f07_farsi_cnt$	$f07_tajik_cnt$
0.935	0.474	8.321	30.302	0.208	0.243

S.E Clustered by village

Table 4: Regression Results

f07_heads_child_cnt	f07_girl_cnt	f07_age_cnt	f07_duration_village_cnt	f07_farsi_cnt	f07_tajik_cnt
0.911	0.455	8.312	27.594	0.209	0.208

S.E Clustered by village

Table 5: Regression Results

f07_heads_child_cnt.treatment	f07_girl_cnt.treatment	$f07_age_cnt.treatment$	f07_duration_village_cnt.treatm
0.024	0.020	0.009	2.709

S.E Clustered by village

Table 6: Regression Results

$f07_heads_child_cnt.treatment$	$f07_girl_cnt.treatment$	$f07_age_cnt.treatment$	$f07_duration_village_cnt.treatreatreatreatreatreatreatreatreatrea$
0.015	0.020	0.040	1.605

S.E Clustered by village

```
#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_modern_rc(ticks = TRUE, plot_title_size = 13) +
```

Table 7: Regression Results

	Depend	lent variable:
	f07_formal_school	f07_both_norma_total
	(1)	(2)
f07_heads_child_cnt	0.038	-0.090
	(0.061)	(0.113)
f07_girl_cnt	-0.208***	-0.682^{***}
	(0.080)	(0.102)
f07_age_cnt	0.046***	0.287***
	(0.017)	(0.018)
f07_duration_village_cnt	-0.001	-0.005^{**}
	(0.001)	(0.002)
f07_farsi_cnt	-0.039	0.074
	(0.074)	(0.119)
f07_tajik_cnt	-0.006	0.042
_ v _	(0.078)	(0.065)
f07_farmer_cnt	-0.050	-0.020
	(0.081)	(0.120)
f07_age_head_cnt	-0.004	-0.004**
	(0.003)	(0.002)
f07_yrs_ed_head_cnt	0.001	0.036***
	(0.006)	(0.007)
f07_num_ppl_hh_cnt	0.004	-0.004
	(0.006)	(0.011)
f07_jeribs_cnt	0.022***	0.056***
~ –	(0.008)	(0.011)
f07_num_sheep_cnt	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
\mathbb{R}^2	0.155	0.401

Note:

*p<0.1; **p<0.05; ***p<0.01 S.E Clustered by village

```
title = "Distribution of Distance from School By Gender")
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),</pre>
               dist_nclose = if_else(f07_nearest_scl >= 3 & f07_nearest_scl < 4, 1, 0),</pre>
               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 &
```

labs(x = "Miles From School", y = "Count", fill = "Is Girl",

```
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
```

stargazer(reg_near_g,reg_close_g,reg_nclose_g, reg_far_g, reg_near_b,reg_close_b,reg_nclose_b, reg_far_covariate.labels = c('Treatment', "chagcharan"), column.labels = c("Short", "Med-Short", "Med dep.var.labels = c("Whether or not the Individual is in a Formal School"),
 type = 'latex', header = F, float = TRUE, notes = c("S.E Clustered by village 1-4 (Female), "Both groups have increasing distance as the number increases."))

Table 8: Regression Results

	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
\mathbb{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_covariate.labels = c('Treatment', "chagcharan"), column.labels = c("Short", "Med-Short", "
```

Table 9: Regression Results

	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
$\frac{R^2}{R^2}$	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.