Replication materials for Gov 2001 Final Paper: Votes on Fire

Maria Ballesteros, Aleksandra Conevska, Ethan Miles

December 14, 2020

This document replicates the analysis of "Votes on Fire".

Data Preparation

```
## Warning: package 'spDataLarge' is not available (for R version 4.0.2)

## Warning in p_install(package, character.only = TRUE, ...):

## Warning in library(package, lib.loc = lib.loc, character.only = TRUE,
## logical.return = TRUE, : there is no package called 'spDataLarge'

## Warning in pacman::p_load(sf, data.table, lwgeom, RSQLite, MASS, car, tidyverse, : Failed to install/
## spDataLarge
```

Fixing Data from OK county

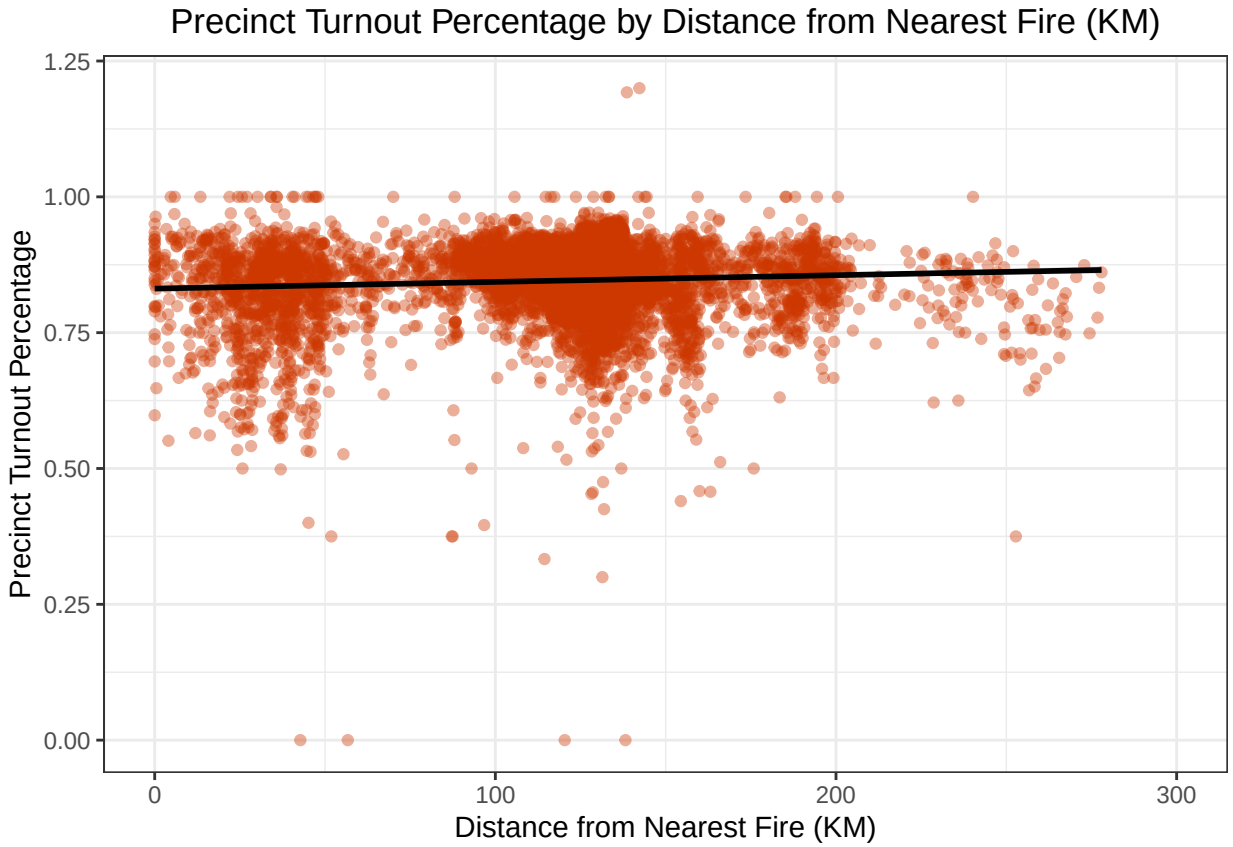
```
################### THIS CHUNK COULD GO IN THE DATA MERGE COMPLETE SCRIPT #########
#### getting OK county data from individual data as there appears to be a data entry error.

reg_v$to2020<- reg_v$voted2020/reg_v$totalreg
## replacing values for OK county with values from Indv level data
  ## first, droping FullPrc.y and renaming FullPrc.x to FullPrc
votefire_1620<- votefire_1620 %>%
                    dplyr::select(-c("FullPrc.y", "countycode")) %>%
                    rename(FullPrc = FullPrc.x)
reg_v<- reg_v %>%
                dplyr::select(c("FullPrc", "totalreg", "voted2020", "to2020"))

## generating dataset with OK replaced
votefire_2<- left_join(votefire_1620, reg_v, by = "FullPrc", all.x = T)
attach(votefire_2)

for( i in 1:nrow(votefire_2)){
if(CountyName[i] == "OK"){
votefire_2$ballotscast2020[i]<-  votefire_2$voted2020[i]
votefire_2$regvoters2020[i]<-    votefire_2$totalreg[i]
votefire_2$turnout2020[i]<-      votefire_2$to2020[i]
}
}



ggplot(votefire_2)+
  geom_point(aes(x = PrecDist5000Fire_km, y = turnout2020), color = "orangered3", alpha = .4) +
```

```
  xlim(0,300) +
  ggtitle("Precinct Turnout Percentage by Distance from Nearest Fire (KM)") +
  xlab("Distance from Nearest Fire (KM)") +
  ylab("Precinct Turnout Percentage") +
  geom_smooth(method = "lm", aes(x = PrecDist5000Fire_km, y = turnout2020), color = "black") +
  theme_bw() +
  theme(plot.title = element_text(hjust = 0.5))
```

```
## `geom_smooth()` using formula 'y ~ x'
```

```
## Warning: Removed 258 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 258 rows containing missing values (geom_point).
```




```
#ggsave("TurnoutByFireDistance.jpeg", path = "/Users/ethan/Dropbox (Harvard University)/Replication/out

## dropping 2 rogue OK precincts and extra variables

votefire_2<- votefire_2 %>%
                        filter(turnout2020<= 1) %>%
                        dplyr::select(-c("totalreg", "voted2020", "to2020"))


### ## creating proportion demographic variables
votefire_2$prop_fem<- votefire_2$fem/votefire_2$regvoters2020
votefire_2$prop_under30<- votefire_2$age18_30/votefire_2$regvoters2020
```

```r
## creating turnout variables by demographic category
votefire_2$fem_turnout <- votefire_2$fem_voted/votefire_2$fem
votefire_2$male_turnout <- votefire_2$male_voted/votefire_2$male
votefire_2$young_turnout <- votefire_2$young_voted/votefire_2$age18_30

## selecting complete cases
delvars <- delvars <- c('turnout2016', 'turnout2020', 'BAWithin0to5km',
             'WHPMean', 'crit_work', 'remindex', 'proximity',
             'exposure', 'urban', 'FIPS')
votefire2_complete <- votefire_2[complete.cases(votefire_2[ , delvars]),]
FullPrc<-votefire2_complete$FullPrc
```

Analysis

Table 1

```r
# Creating lengths for cluster robust se
G <- length(unique(votefire2_complete$CountyCode))
N <- length(votefire2_complete$CountyCode)

# model 1 - binary iv, no turnout trend, controls, no fixed effects
m1 <-  lm(
        turnout2020 ~ BAWithin0to5km  + WHPMean + crit_work +
        remindex + proximity + exposure + urban + pop,
        data = votefire2_complete)

dfa1 <- (G/(G - 1)) * (N - 1)/m1$df.residual#computing df adjustment for clustered SE
firm_c_vcov1 <- dfa1 * vcovHC(m1, type = "HC0", cluster = "group", adjust = T) # display with cluster V
m1_r <- coeftest(m1, vcov = firm_c_vcov1)# run models with clustered SE



# model 2 - binary iv, turnout trend, controls, no fixed effects
m2 <-  lm(
       turnout2020 ~ BAWithin0to5km  + WHPMean + turnout2016 +
       crit_work + remindex + proximity + exposure + urban + pop,
       data = votefire2_complete)

dfa2 <- (G/(G - 1)) * (N - 1)/m2$df.residual#computing df adjustment for clustered SE
firm_c_vcov2 <- dfa2 * vcovHC(m2, type = "HC0", cluster = "group", adjust = T) # display with cluster V
m2_r <- coeftest(m2, vcov = firm_c_vcov2)# run models with clustered SE


# model 3 - binary iv, turnout trend, no controls,  fixed effects

# m3_rob = lm_robust(
#  turnout2020 ~ BAWithin0to5km  + WHPMean + turnout2016,
 # fixed_effects = ~ CountyCode,
#  cluster = CountyCode,
#  se_type = "CR0",
#  data = votefire2_complete
#)
```

```r
m3 <- lm(turnout2020 ~ BAWithin0to5km  + WHPMean + turnout2016 + factor(CountyCode),
          data = votefire2_complete)
dfa3 <- (G/(G - 1)) * (N - 1)/m3$df.residual #computing df adjustment for clustered SE
firm_c_vcov3 <- dfa3 * vcovHC(m3, type = "HC0", cluster = "group", adjust = T) # display with cluster V
m3_r <- coeftest(m3, vcov = firm_c_vcov3)# run models with clustered SE


## PLM m3 <- plm(turnout2020 ~ BAWithin0to5km  + WHPMean + turnout2016,
#           data = votefire2_complete,
#           model = 'within',
#           index = c('CountyCode'))


dfa3 <- (G/(G - 1)) * (N - 1)/m3$df.residual #computing df adjustment for clustered SE
firm_c_vcov3 <- dfa3 * vcovHC(m3, type = "HC0", cluster = "group", adjust = T) # display with cluster V
m3_r <- coeftest(m3, vcov = firm_c_vcov3)# run models with clustered SE


# model 4 -  continuous iv, no turnout trend, covid controls,  no fixed effects
m4 <-  lm(
  turnout2020 ~ PrecDist5000Fire_km  + WHPMean + crit_work +
    remindex + proximity + exposure + urban + pop,
  data = votefire2_complete
)

dfa4 <- (G/(G - 1)) * (N - 1)/m4$df.residual#computing df adjustment for clustered SE
firm_c_vcov4 <- dfa4 * vcovHC(m4, type = "HC0", cluster = "group", adjust = T) # display with cluster V
m4_r <- coeftest(m4, vcov = firm_c_vcov4)# run models with clustered SE




# model 5 - continuous iv, turnout trend, covid controls, no fixed effects
m5 <-  lm(
  turnout2020 ~ PrecDist5000Fire_km  + WHPMean + turnout2016 +
    crit_work + remindex + proximity + exposure + urban + pop,
  data = votefire2_complete
)

dfa5 <- (G/(G - 1)) * (N - 1)/m5$df.residual#computing df adjustment for clustered SE
firm_c_vcov5 <- dfa5 * vcovHC(m5, type = "HC0", cluster = "group", adjust = T) # display with cluster V
m5_r <- coeftest(m5, vcov = firm_c_vcov5)# run models with clustered SE




# model 6 -  continuous iv, turnout trend, no covid controls,  fixed effects


#m6_rob = lm_robust(
 # turnout2020 ~ PrecDist5000Fire_km  + WHPMean + turnout2016,
#  fixed_effects = ~ CountyCode,
```

```
#  cluster = CountyCode,
#  se_type = "CR0",
#  data = votefire2_complete
#)

# THIS WORKS INSTEAD OF LM_ROBUST. PRODUCES THE SAME SE AND COEFFS ALMOST EXACTLY. MIGHT WANT TO JUST U
# IN STARGAZER. IF EVERYONE AGREES WE CAN JUST QUICKLY CHANGE THE CODE

## PLM m6 <- plm(turnout2020 ~ PrecDist5000Fire_km  + WHPMean + turnout2016, data = votefire2_complete,

m6 <- lm(turnout2020 ~ PrecDist5000Fire_km  + WHPMean + turnout2016 + factor(CountyCode), data = votefi

dfa6 <- (G/(G - 1)) * (N - 1)/m6$df.residual #computing df adjustment for clustered SE
firm_c_vcov6 <- dfa6 * vcovHC(m6, type = "HC0", cluster = "group", adjust = T) # display with cluster V
m6_r <- coeftest(m6, vcov = firm_c_vcov6)# run models with clustered SE


#### on how to add lm_robust objects to stargazer: https://declaredesign.org/r/estimatr/articles/regres
```

Stargazer tables

```
##

stargazer(m1, m2, m3, m4, m5, m6,
           se = starprep(m1, m2, m3, m4, m5, m6,
                         clusters = votefire2_complete$FullPrc, se_type = "stata"),
          header = FALSE,
          df=FALSE,
          style = "apsr",
          title = "Main result with different measure of fire exposure",
          omit = c("WHPMean", "crit_work", "remindex", "proximity",
                   "exposure", "urban", "pop", "turnout2020", "Constant", "CountyCode"),
          add.lines = list(c("County FE", "", "", "X", "", "", "X"),
                           c("Controls", "X", "X", "", "X", "X", "")),
          covariate.labels = c("Fire within 0-5km", "Distance to nearest fire"),
          dep.var.caption = "Turnout 2020 (mean = ; sd = 0.",
          keep.stat = c("rsq", "n"),
          digits = 3)
```

Testing turnout change

```
#Creating vote change variable
votefire_2$turnout_change <- votefire_2$turnout2020 - votefire_2$turnout2016
summary(votefire_2$turnout_change)
```

```
Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
```

-0.81028 0.02809 0.05321 0.05735 0.07949 1.00000

```
#running without fixed effects
m7 <-  lm(
  turnout_change ~ PrecDist5000Fire_km  + WHPMean + crit_work +
    remindex + proximity + exposure + urban + pop,
  data = votefire_2
```

Table 1: Main result with different measure of fire exposure

	turnout2020					
	(1)	(2)	(3)	(4)	(5)	(6)
Fire within 0-5km	−0.002	0.001	0.010			
	(0.012)	(0.010)	(0.010)			
Distance to nearest fire				0.0002***	0.0001***	−0.0003***
				(0.00003)	(0.00002)	(0.0001)
turnout2016		0.415***	0.404***		0.414***	0.404***
		(0.022)	(0.022)		(0.022)	(0.022)
County FE			X			X
Controls	X	X		X	X	
N	6,768	6,768	6,768	6,768	6,768	6,768
R^2	0.046	0.407	0.440	0.051	0.410	0.443

*p < .1; **p < .05; ***p < .01

```
)

dfa7 <- (G/(G - 1)) * (N - 1)/m7$df.residual#computing df adjustment for clustered SE
firm_c_vcov7 <- dfa7 * vcovHC(m7, type = "HC0", cluster = "group", adjust = T) # display with cluster V
m7_r <- coeftest(m7, vcov = firm_c_vcov7)# run models with clustered SE


#running with fixed effects
m8 = lm_robust(
  turnout_change ~ PrecDist5000Fire_km  + WHPMean,
  fixed_effects = ~ CountyCode,
  cluster = CountyCode,
  se_type = "CR0",
  data = votefire_2
)


## Fixed effects, in lm form (vs lm_robust)
m8 <-  lm(
  turnout_change ~ PrecDist5000Fire_km  + WHPMean + crit_work +
    remindex + proximity + exposure + urban + pop + factor(CountyCode),
  data = votefire_2
)


dfa8 <- (G/(G - 1)) * (N - 1)/m8$df.residual#computing df adjustment for clustered SE
firm_c_vcov8 <- dfa8 * vcovHC(m8, type = "HC0", cluster = "group", adjust = T) # display with cluster V
m8_r <- coeftest(m8, vcov = firm_c_vcov8)# run models with clustered SE


## is this the correct setup for SEs?
stargazer(m7, m8,
           se = starprep(m7, m8,
                         clusters = votefire2_complete$FullPrc, se_type = "stata"),
          header = FALSE,
```

```
          df=FALSE,
          style = "apsr",
          title = "Turnout Change",
          omit = c("WHPMean", "crit_work", "remindex", "proximity",
                   "exposure", "urban", "pop", "turnout2020", "Constant", "CountyCode"),
          add.lines = list(c("County FE", "", "X"),
                           c("Controls", "X", "X")),
          dep.var.caption = "Turnout 2020 (mean = ; sd = 0.",
          keep.stat = c("rsq", "n"),
          digits = 3)
```

Table 2: Turnout Change

	turnout_change	
	(1)	(2)
PrecDist5000Fire_km	0.0001**	−0.0004***
	(0.00003)	(0.0001)
County FE		X
Controls	X	X
N	6,768	6,768
R^2	0.014	0.039

*p < .1; **p < .05; ***p < .01

creating model with demographic controls

```
### first, creating percent female and percent age


m4 <-  lm(
        turnout2020 ~ PrecDist5000Fire_km  + WHPMean + crit_work +
        remindex + proximity + exposure + urban + pop,
        data = votefire2_complete)

c1<-  lm(
  turnout2020 ~ PrecDist5000Fire_km  + WHPMean + crit_work +
    remindex + proximity + exposure + urban + pop + prop_under30,
  data = votefire2_complete
)


c2<-  lm(
  turnout2020 ~ PrecDist5000Fire_km  + WHPMean + crit_work +
    remindex + proximity + exposure + urban + pop + prop_fem,
  data = votefire2_complete
)


c3<-  lm(
  turnout2020 ~ PrecDist5000Fire_km  + WHPMean + crit_work +
    remindex + proximity + exposure + urban + pop + prop_under30*prop_fem,
  data = votefire2_complete
```

```
)


stargazer(
  m4, c1, c2, c3,
  se = starprep(m4, c1, c2, c3,
                clusters = votefire2_complete$FullPrc, se_type = "stata"),
          header = FALSE,
          df=FALSE,
          style = "apsr",
          title = "Main result with different measure of fire exposure",
          omit = c("WHPMean", "crit_work", "remindex", "proximity",
                   "exposure", "urban", "pop", "turnout2020", "Constant"),
          column.labels = c("Baseline", "Baseline w Gender", "Baseline w young", "Baseline w age \ gende
          covariate.labels = c( "Distance to nearest fire"),
          dep.var.caption = "Turnout 2020 (mean = ; sd = 0.",
          keep.stat = c("rsq", "n"),
          digits = 3
)
```

Table 3: Main result with different measure of fire exposure

	turnout2020			
	Baseline	Baseline w Gender	Baseline w young	Baseline w age gender
	(1)	(2)	(3)	(4)
Distance to nearest fire	0.0002***	0.0001***	0.0002***	0.0001***
	(0.00003)	(0.00003)	(0.00003)	(0.00003)
prop_under30		−0.167***		−0.342***
		(0.059)		(0.026)
prop_fem			−0.237***	−0.241***
			(0.088)	(0.069)
prop_under30:prop_fem				0.121***
				(0.012)
N	6,768	6,768	6,768	6,768
R^2	0.051	0.157	0.093	0.248

*p < .1; **p < .05; ***p < .01

```
#### Removing cases where age is missing #only 2#
delvars <- delvars <- c("fem_turnout", "young_turnout", "male_turnout")

votefire_dem <- votefire2_complete[complete.cases(votefire2_complete[ , delvars]),]

M4 <-  lm(
        turnout2020 ~ PrecDist5000Fire_km  + WHPMean + crit_work +
        remindex + proximity + exposure + urban + pop,
        data = votefire_dem)
c4 <-  lm(
        fem_turnout ~ PrecDist5000Fire_km  + WHPMean + crit_work +
        remindex + proximity + exposure + urban + pop + prop_fem,
        data = votefire_dem)
c5<-  lm(
  male_turnout ~ PrecDist5000Fire_km  + WHPMean + crit_work + prop_fem +
```

```
    remindex + proximity + exposure + urban + pop,
  data = votefire_dem
)


c6<-  lm(
  young_turnout ~ PrecDist5000Fire_km  + WHPMean + crit_work +
    remindex + proximity + exposure + urban + pop + prop_under30,
  data = votefire_dem
)

stargazer(
  M4, c4, c5, c6,
  se = starprep(M4, c4, c5, c6,
                clusters = votefire_dem$FullPrc, se_type = "stata"),
          header = FALSE,
          df=FALSE,
          style = "apsr",
          title = "Main result with different measure of fire exposure",
          omit = c("WHPMean", "crit_work", "remindex", "proximity",
                   "exposure", "urban", "pop", "turnout2020", "Constant"),
          column.labels = c("Baseline", "Turnout: Female Voters", "Turnout:Male Voters", "Turnout: age 3
          covariate.labels = c( "Distance to nearest fire"),
          dep.var.caption = "Turnout 2020 (mean = ; sd = 0.",
          keep.stat = c("rsq", "n"),
          digits = 3
)
```

Table 4: Main result with different measure of fire exposure

	turnout2020 Baseline	fem_turnout Turnout: Female Voters	male_turnout Turnout:Male Voters	young_turnout Turnout: age 30 and under
	(1)	(2)	(3)	(4)
Distance to nearest fire	0.0002***	0.0001***	0.0001***	−0.0001***
	(0.00003)	(0.00003)	(0.00004)	(0.00004)
prop_fem		−0.520***	−0.500***	
		(0.129)	(0.133)	
prop_under30				−0.185***
				(0.042)
N	6,733	6,733	6,733	6,733
R^2	0.056	0.202	0.152	0.230

*p < .1; **p < .05; ***p < .01

Testing if WHP predicts fires

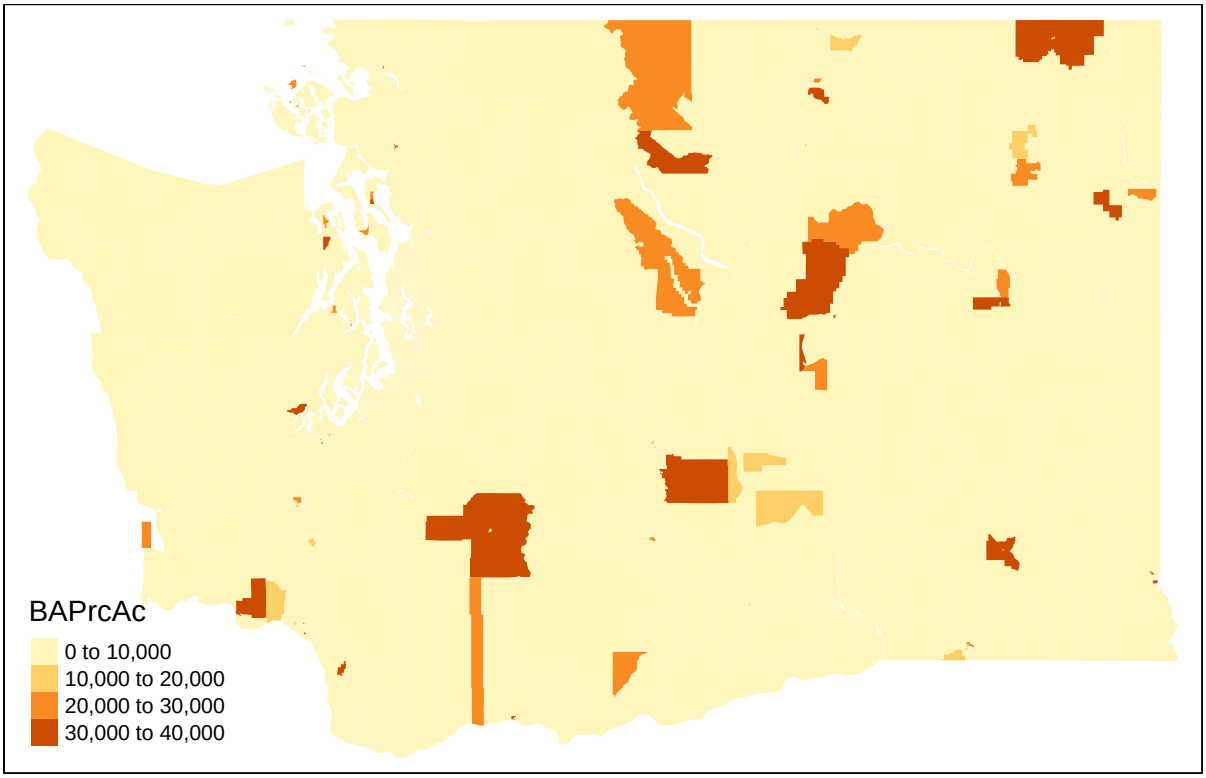
```
WHP_test <- plm(BASqMSum ~ WHPMean, data = votefire2_complete, model = 'within', index = c('CountyCode'

dfatest <- (G/(G - 1)) * (N - 1)/WHP_test$df.residual#computing df adjustment for clustered SE
firm_c_vcovtest <- dfatest * vcovHC(WHP_test, type = "HC0", cluster = "group", adjust = T) # display wi
WHPtest_r <- coeftest(WHP_test, vcov = firm_c_vcovtest)# run models with clustered SE
```

```
stargazer(
  WHPtest_r,
          header = FALSE,
          df=FALSE,
          style = "apsr",
          title = "Testing relationship between fire vulnerability and burned area",
          omit = c("Constant"),
          column.labels = c("Burned Area (Sq m)"),
          covariate.labels = c("Wildfire Hazard Potential"),
          dep.var.caption = "",
          keep.stat = c("rsq", "n"),
          digits = 3
)
```

```
##
## \begin{table}[!htbp] \centering
##   \caption{Testing relationship between fire vulnerability and burned area}
##   \label{}
## \begin{tabular}{@{\extracolsep{5pt}}lc}
## \\[-1.8ex]\hline \\[-1.8ex]
## \\[-1.8ex] &   \\
##  & Burned Area (Sq m) \\
## \hline \\[-1.8ex]
##  Wildfire Hazard Potential & 1,212.155 \\
##   & (1,330.625) \\
##  \hline \\[-1.8ex]
## \multicolumn{2}{l}{$^{*}$p $<$ .1; $^{**}$p $<$ .05; $^{***}$p $<$ .01} \\
## \end{tabular}
## \end{table}
```

Figure X: Map of Wildfire

```
## Warning in sp::proj4string(obj): CRS object has comment, which is lost in output
```




```
## Warning in sp::proj4string(obj): CRS object has comment, which is lost in output
```

```
## Map saved to /Users/mariaballesteros/Dropbox/Harvard/G1/GOV 2001/Replication/output/Figures/washfire_
```

```
## Resolution: 3200 by 2540 pixels
```

```
## Size: 9.142857 by 7.257143 inches (350 dpi)
```

Figure X: Main results with different measures of fire exposure

```
### Ethan  ### this plot needs confidence intervals
distlist.granular = seq(from=5, to=40, by=5)

f.reltomedian = as.formula("turnout2020~BAWithin0to5km + BAWithin5to10km +
        BAWithin10to15km + BAWithin15to20km + BAWithin20to25km +
        BAWithin25to30km + BAWithin30to35km + BAWithinOver40km +
        turnout2020 +  WHPMean ")


lmr.reltomedian = lm(
                  formula = f.reltomedian,
                  data=votefire_2)
```

```
## Warning in model.matrix.default(mt, mf, contrasts): the response appeared on the
## right-hand side and was dropped
```

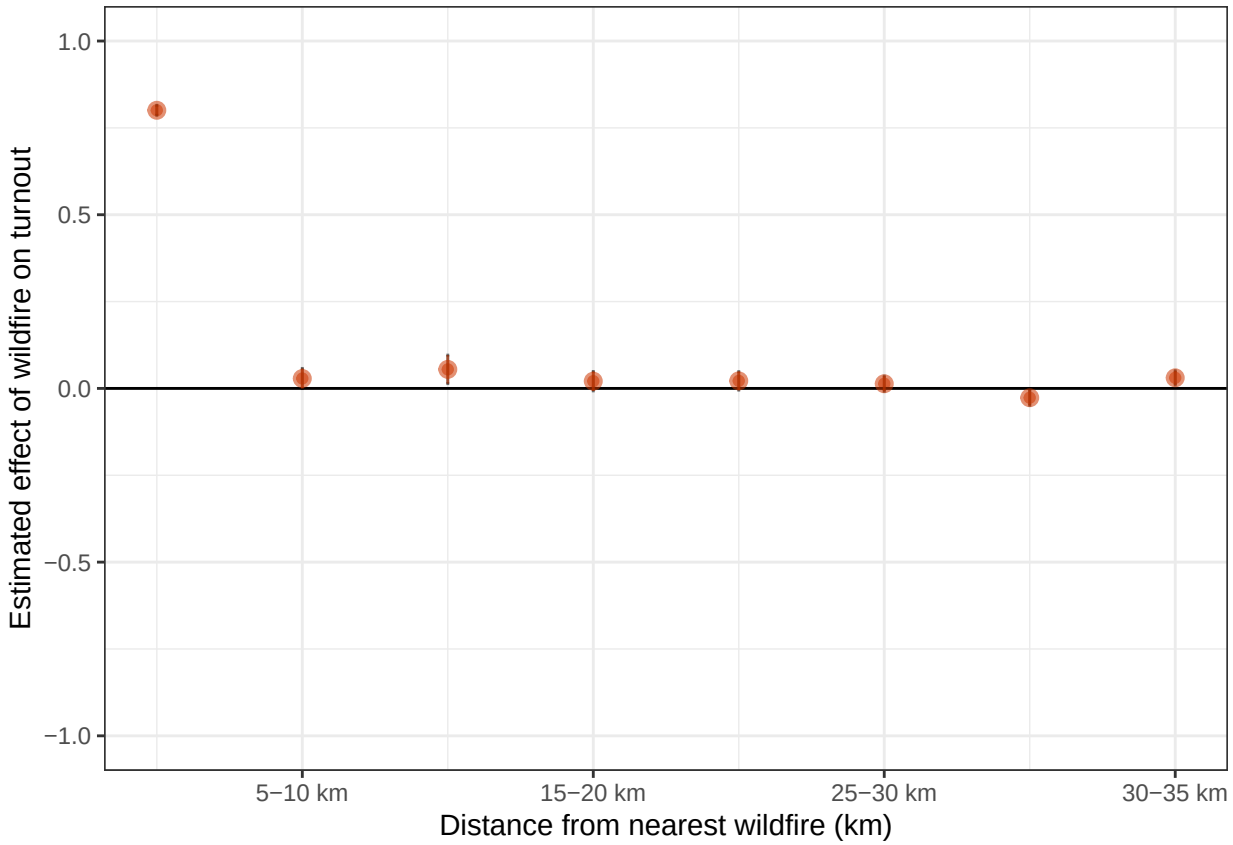
```
## Warning in model.matrix.default(mt, mf, contrasts): problem with term 9 in
## model.matrix: no columns are assigned
```

```
coef.reltomedian = summary(lmr.reltomedian)$coef
```

```r
# Regresion table to accompany figure - could go in Appendix
rownames(coef.reltomedian)=
  c("Fire within 0-5km","Fire within 5-10km", "Fire within 10-15km",
    "Fire within 15-20km","Fire within 20-25km","Fire within 25-30km",
    "Fire within 30-35km", "Fire over 40km away",
    "Voter turnout 2016", "Fire Vulnerability")

xtable::xtable(coef.reltomedian[1:10,1:4], digits = 3)
```

```
## % latex table generated in R 4.0.2 by xtable 1.8-4 package
## % Mon Dec 14 11:34:10 2020
## \begin{table}[ht]
## \centering
## \begin{tabular}{rrrrr}
##   \hline
##  & Estimate & Std. Error & t value & Pr($>$$|$t$|$) \\
##   \hline
## Fire within 0-5km & 0.801 & 0.006 & 143.943 & 0.000 \\
##   Fire within 5-10km & 0.029 & 0.011 & 2.533 & 0.011 \\
##   Fire within 10-15km & 0.055 & 0.016 & 3.437 & 0.001 \\
##   Fire within 15-20km & 0.021 & 0.011 & 1.925 & 0.054 \\
##   Fire within 20-25km & 0.022 & 0.010 & 2.104 & 0.035 \\
##   Fire within 25-30km & 0.013 & 0.008 & 1.587 & 0.113 \\
##   Fire within 30-35km & -0.027 & 0.008 & -3.155 & 0.002 \\
##   Fire over 40km away & 0.031 & 0.008 & 3.862 & 0.000 \\
##   Voter turnout 2016 & 0.047 & 0.006 & 8.423 & 0.000 \\
##   Fire Vulnerability & 0.000 & 0.000 & 5.528 & 0.000 \\
##    \hline
## \end{tabular}
## \end{table}
```

```r
# Prepare for plotting
whichdistances = distlist.granular[1:8]
numdistances = length(whichdistances)
coefs.intervals = coef.reltomedian[1:numdistances,1]
ses.intervals = coef.reltomedian[1:numdistances,2]

plot.dat = data.frame(coefs.intervals, ses.intervals, ub=coefs.intervals+2.58*ses.intervals,
                      lb = coefs.intervals-2.58*ses.intervals, dist=whichdistances)

g = ggplot(plot.dat, aes(y=coefs.intervals, x=dist, color = I("orangered3"), alpha = .4)) +
    geom_point() +
    theme_bw() +
    geom_errorbar(aes(ymin=lb,ymax=ub), color = "black", width=0.1) +
    labs(x="Distance from nearest wildfire (km)") +
    scale_y_continuous(name="Estimated effect of wildfire on turnout", limits=c(-1, 1)) +
    geom_hline(yintercept = 0) +
    geom_pointrange(aes(ymin=lb, ymax=ub)) +
    theme(legend.position = "none")



g + scale_x_continuous(breaks=seq(10,40,10),
                       labels=c("5-10 km","15-20 km","25-30 km","30-35 km"))
```




```
#ggsave("Effect_Of_Fire_On_Turnout.jpeg", path = "/Users/ethan/Dropbox (Harvard University)/Replication


#dev.off()
#ggsave(filename = "fig_dist_reltomedian.jpg", plot = last_plot(),
       #device = "jpeg", path = "figures",
      # width = 6, height = 4, units = "in")
```

##Running additional models and robustness checks

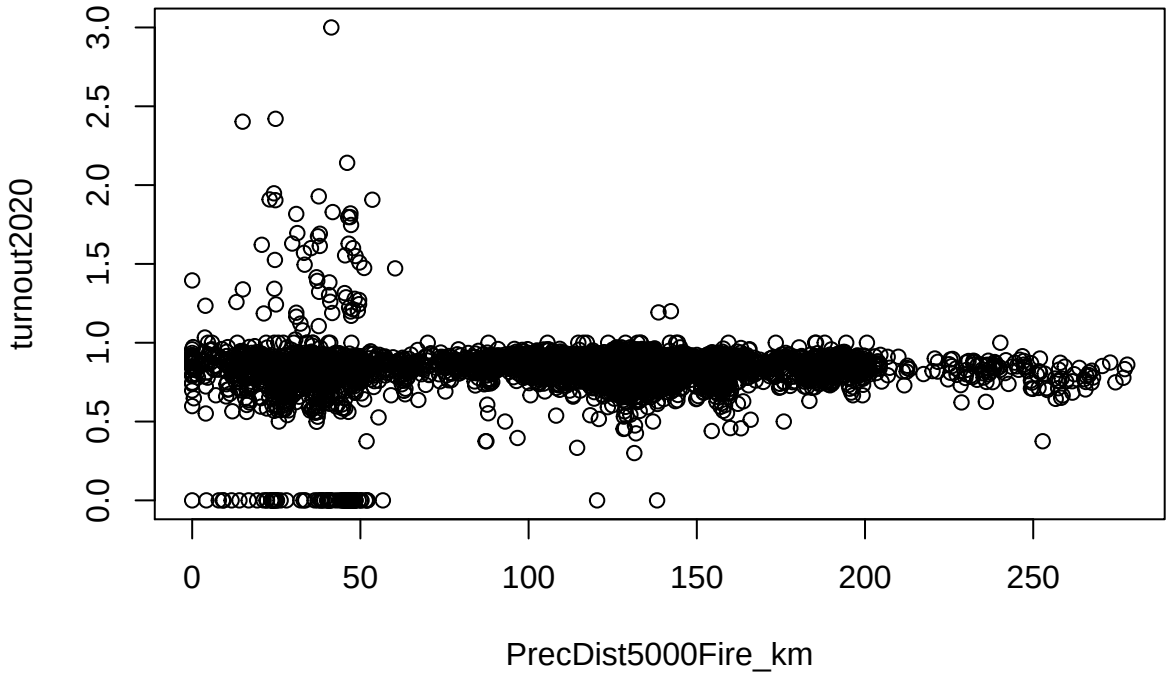
```
## Set working directory
#Aleksandra
#setwd( '/Users/aleksandraconevska/Dropbox/Harvard/2020/Fall2020/Gov-2001/Replication-Paper/Replication

#María

##### seeing how result change OK original vs Ok replaced
attach(votefire_1620)
```

```
## The following object is masked _by_ .GlobalEnv:
##
##     FullPrc
```

```
## The following objects are masked from votefire_2:
##
##     AFCount, age18_30, age30_45, age45_65, age65plus, BAAcresSum,
##     ballotscast, ballotscast2016, ballotscast2020, BAPrecPct, BASqMSum,
##     BAWithin0to5km, BAWithin10to15km, BAWithin15to20km,
```

```
##     BAWithin20to25km, BAWithin25to30km, BAWithin30to35km,
##     BAWithin35to40km, BAWithin5to10km, BAWithinOver40km, cocountyname,
##     CountyCd, CountyCode, CountyName, crit_work, exposure, fem,
##     fem_voted, fid, FIPS, FullPrc, gender_unknown, male, male_voted,
##     pop, PrcCode, PrecAreaSqM, preccodesource, PrecDist5000Fire_km,
##     PrecDistTo5000Fire, precinctcode, PrecinctCode, precinctname,
##     PrecinctPart, proximity, regvoters, regvoters2016, regvoters2020,
##     remindex, State, turnout2016, turnout2020, unique, urban, WHPMean,
##     X, young_voted
```

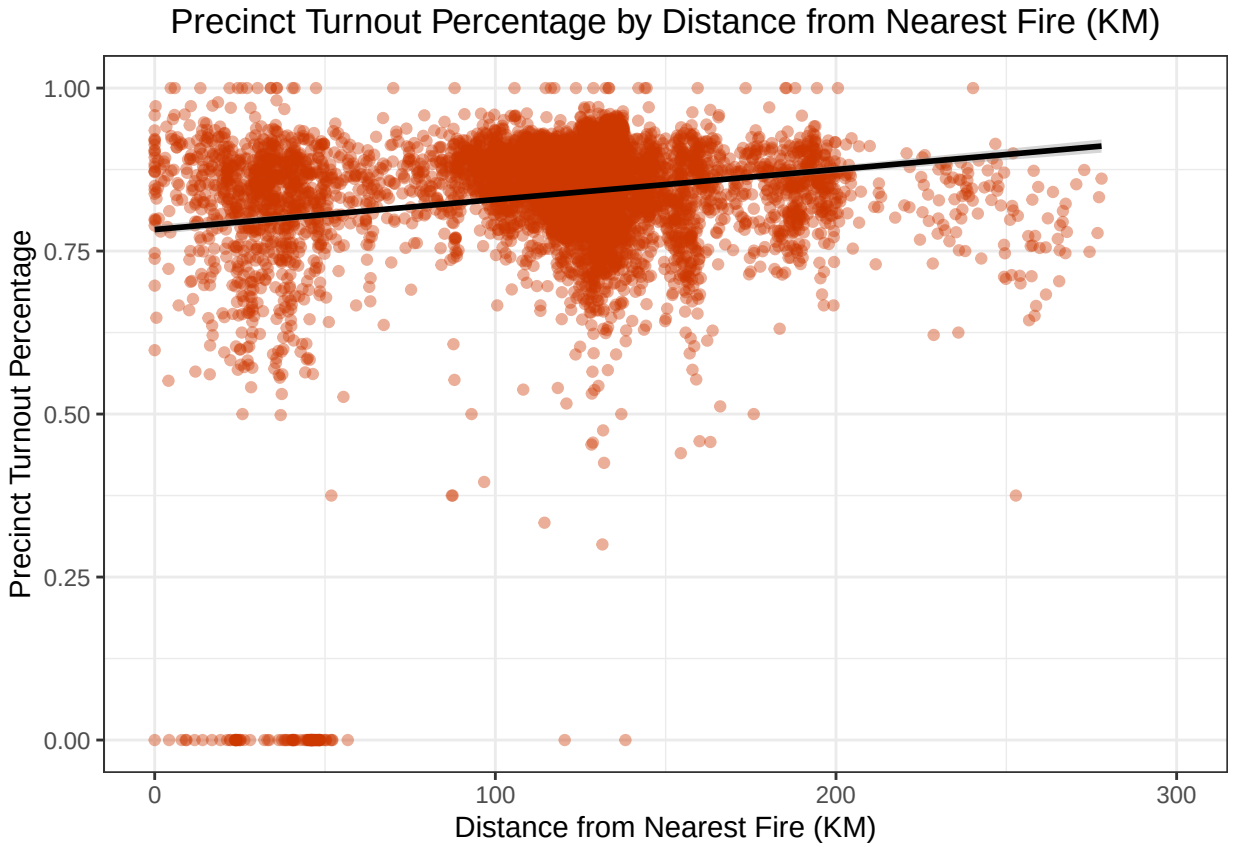
```
## This looks sus? idk why there are so many < 1
plot(x = PrecDist5000Fire_km, y = turnout2020)
```




```
### our standard errors are much larger for the clustered model, so i'll try transformations

############ getting rid of precincts with turnout higher than one to check results

# model t1-  continuous iv, no turnout trend, covid controls,  no fixed effects w/o precincts with turn
test<- votefire_1620 %>% filter(turnout2020 <=1)

ggplot(test)+
  geom_point(aes(x = PrecDist5000Fire_km, y = turnout2020), color = "orangered3", alpha = .4) +
  xlim(0,300) +
  ggtitle("Precinct Turnout Percentage by Distance from Nearest Fire (KM)") +
  xlab("Distance from Nearest Fire (KM)") +
  ylab("Precinct Turnout Percentage") +
  geom_smooth(method = "lm", aes(x = PrecDist5000Fire_km, y = turnout2020), color = "black") +
  theme_bw() +
  theme(plot.title = element_text(hjust = 0.5))
```

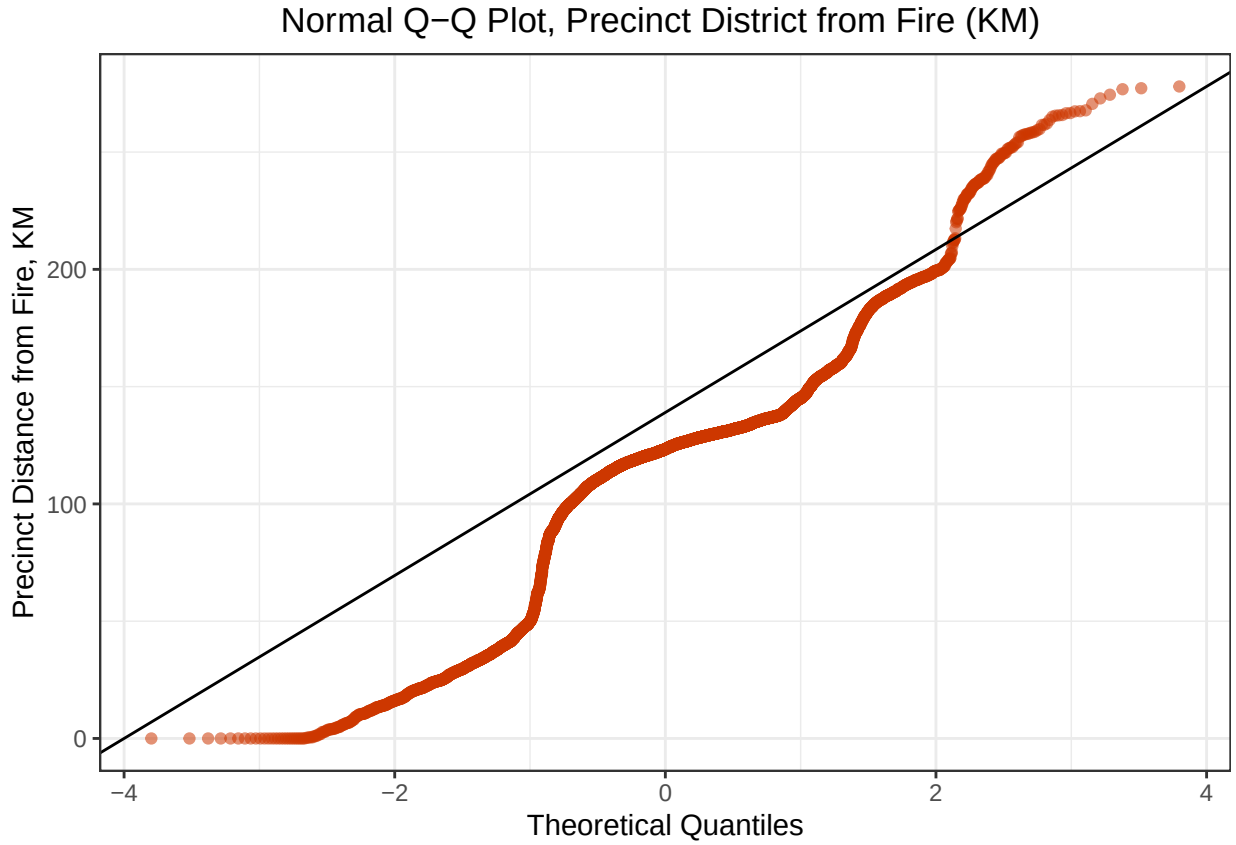
```
## `geom_smooth()` using formula 'y ~ x'
## Warning: Removed 1 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 1 rows containing missing values (geom_point).
```



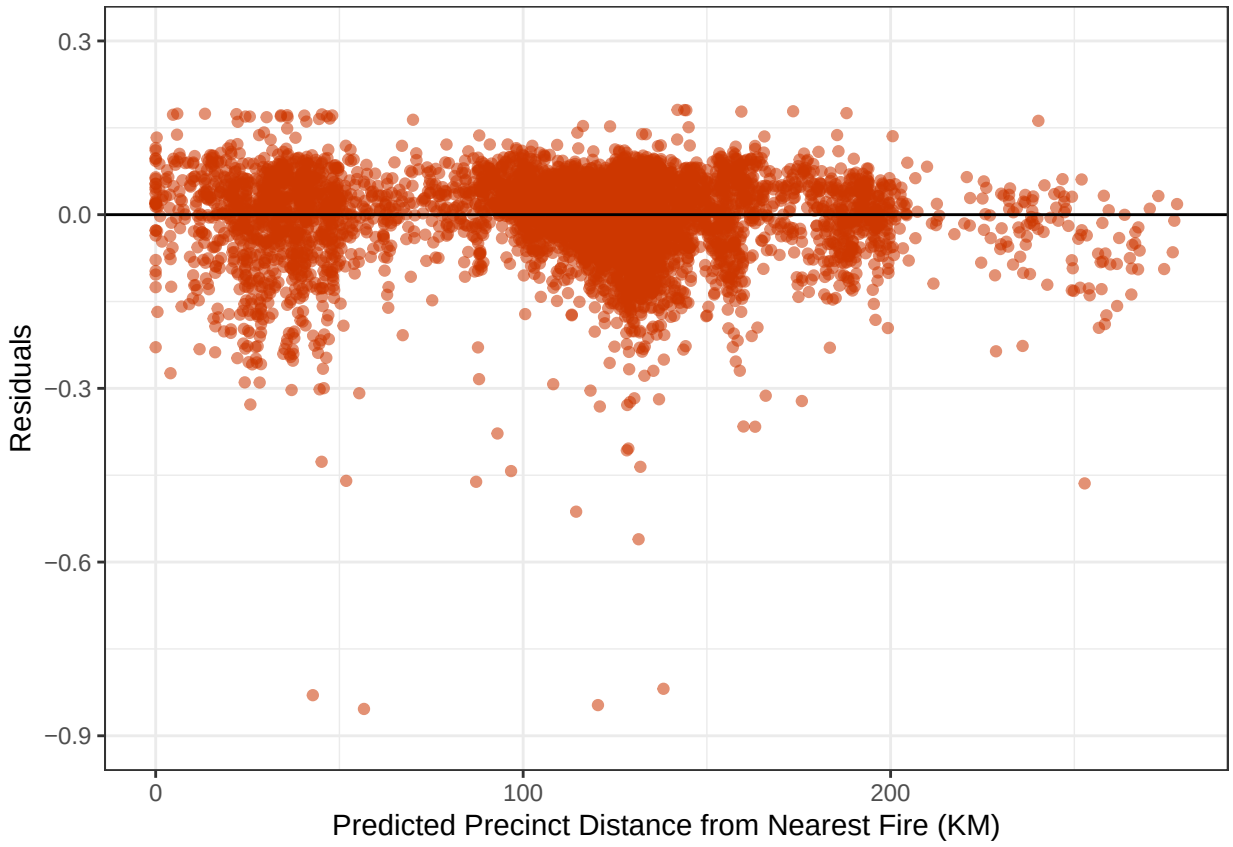

```
#ggsave("TurnoutByFireDistance_Test.jpeg", path = "/Users/ethan/Dropbox (Harvard University)/Replicatio

t1 <-  lm(
  turnout2020 ~ PrecDist5000Fire_km  + WHPMean + crit_work +
    remindex + proximity + exposure + urban + pop,
  data = test
)

dfat1 <- (G/(G - 1)) * (N - 1)/t1$df.residual#computing df adjustment for clustered SE
firm_c_vcovt1 <- dfat1 * vcovHC(t1, type = "HC0", cluster = "group", adjust = T) # display with cluster
t1_r <- coeftest(t1, vcov = firm_c_vcovt1)# run models with clustered SE
#



### trying models with new values of OK
# model t1-  continuous iv, no turnout trend, covid controls,  no fixed effects w/o precincts with turn

t2 <-  lm(
  turnout2020 ~ PrecDist5000Fire_km  + WHPMean + crit_work +
    remindex + proximity + exposure + urban,
  data = votefire_2
)

dfat2 <- (G/(G - 1)) * (N - 1)/t2$df.residual#computing df adjustment for clustered SE
```

```r
firm_c_vcovt2 <- dfat2 * vcovHC(t2, type = "HC0", cluster = "group", adjust = T) # display with cluster
t2_r <- coeftest(t2, vcov = firm_c_vcovt2)# run models with clustered SE


## Turnout



# QQ plot

qplot(sample = PrecDist5000Fire_km, na.rm = TRUE, color = I("orangered3"), alpha = .4) +
      ylab("Precinct Distance from Fire, KM") +
      xlab("Theoretical Quantiles") +
      ggtitle("Normal Q-Q Plot, Precinct District from Fire (KM) ") +
      geom_abline(intercept = max(PrecDist5000Fire_km, na.rm = TRUE)/2, slope = max(PrecDist5000Fire_km
      theme_bw() +
      theme(legend.position = "none") +
      theme(plot.title = element_text(hjust = 0.5))
```




```r
#ggsave("QQ.jpeg", path = "/Users/ethan/Dropbox (Harvard University)/Replication/output/Figures")


## RESIDUALS PLOT

ggplot(m4) +
  aes(m4$model$PrecDist5000Fire_km, m4$residuals, color = I("orangered3"), alpha = .4) +
  geom_point() +
```

```
  geom_hline(aes(yintercept = 0, color ="black")) +
  xlab("Predicted Precinct Distance from Nearest Fire (KM)") +
  ylab("Residuals") +
  scale_y_continuous(limits = c(-.9, .3), breaks = c(-.9, -.6, -.3, 0, .3)) +
  theme_bw() +
  theme(legend.position = "none")
```




```
#ggsave("Residuals.jpg", path = "/Users/ethan/Dropbox (Harvard University)/Replication/output/Figures")



## Trying distance measure transformations
#votefire2_complete$PrecDist5000Fireksq <- sqrt(PrecDist5000Fire_km)

#votefire2_complete$PrecDist5000Fireksq2 <- PrecDist5000Fire_km^2

#attach(votefire2_complete)

#is.na(votefire2_complete$PrecDist5000Fire_km) <- votefire2_complete$PrecDist5000Fire_km == "0"
#is.na(votefire2_complete$PrecDist5000Fire_km) <- votefire2_complete$PrecDist5000Fire_km < "1"


#m <- lm(turnout2020 ~ PrecDist5000Fire_km, data = votefire2_complete)

#MASS::boxcox(PrecDist5000Fire_km, lambda = seq(-0.25, 0.75, by = 0.05), plotit = TRUE)
#MASS::boxcox(turnout2020 ~ PrecDist5000Fire_km)
```

```
#car::boxCox(PrecDist5000Fire_km, lambda = seq(-0.25, 0.5, 1/10), plotit = TRUE)


#distance <- votefire2_complete %>%
                  # dplyr::select(c("PrecDist5000Fire_km"))

#MASS::boxcox(, lambda = "auto")



## RESIDUALS AND QQ PLOT POST TRANSFORM

#m4t <-  lm(
# turnout2020 ~ PrecDist5000Fire_km  + WHPMean + crit_work +
 #   remindex + proximity + exposure + urban + pop,
 # data = votefire2_complete
#)

## (ethan) -- I still need to get these to work

#Res plot
#ggplot(m4t) +
#  aes(m4t$model$PrecDist5000Fireksq, m4t$residuals, color = I("orangered3"), alpha = .4) +
# geom_point() +
#  geom_hline(aes(yintercept = 0, color ="black")) +
#  xlab("Predicted Precinct Distance from Nearest Fire (KM)") +
# ylab("Residuals") +
#  scale_y_continuous(limits = c(-.9, .3), breaks = c(-.9, -.6, -.3, 0, .3)) +
 # theme_bw() +
 # theme(legend.position = "none")

#ggsave("Transformed_Residuals.jpeg", path = "/Users/ethan/Dropbox (Harvard University)/Replication/out


#QQ plot
#qplot(
#  sample = PrecDist5000Fireksq2,
#  na.rm = TRUE,
#  color = I("orangered3"),
##  alpha = .4
#) +
#  ylab("Precinct Distance from Fire, KM") +
#  xlab("Theoretical Quantiles") +
 # ggtitle("Normal Q-Q Plot, Precinct District from Fire (KM) ") +
# geom_abline(
 #   intercept = max(PrecDist5000Fireksq2, na.rm = TRUE) / 2,
 #   slope = max(PrecDist5000Fireksq2, na.rm = TRUE) / 8
 # ) +
#  theme_bw() +
#  theme(legend.position = "none") +
 # theme(plot.title = element_text(hjust = 0.5))

#ggsave("Transformed_QQ.jpeg", path = "/Users/ethan/Dropbox (Harvard University)/Replication/output/Fig
```

Why is our coefficient flipping? Different versions of our baseline model

```
### first version no clusters no controls
mt1 <-  lm( turnout2020 ~ PrecDist5000Fire_km  + WHPMean,
        data = votefire2_complete)

se1 <- coef(summary(mt1))[,2]

### no controls clusters
mt2 <-  lm( turnout2020 ~ PrecDist5000Fire_km  + WHPMean,
        data = votefire2_complete)

se2 <- starprep(mt2, clusters = votefire2_complete$FullPrc, se_type = "stata")

### clusters and covid controls

mt3 <-  lm( turnout2020 ~ PrecDist5000Fire_km  + WHPMean + crit_work +
    remindex + proximity + exposure,
        data = votefire2_complete)

se3 <- starprep(mt3, clusters = votefire2_complete$FullPrc, se_type = "stata")

### clusters and urban controls

mt4 <-  lm( turnout2020 ~ PrecDist5000Fire_km  + WHPMean + pop + urban,
        data = votefire2_complete)

se4 <- starprep(mt4, clusters = votefire2_complete$FullPrc, se_type = "stata")

## clusters, urban controls and covid controls.

mt5 <-  lm( turnout2020 ~ PrecDist5000Fire_km  + WHPMean + pop + urban+ crit_work +
    remindex + proximity + exposure,
        data = votefire2_complete)

se5 <- starprep(mt5, clusters = votefire2_complete$FullPrc, se_type = "stata")

## clusters, trend controls

mt6 <-  lm( turnout2020 ~ PrecDist5000Fire_km  + WHPMean + turnout2016,
        data = votefire2_complete)

se6 <- starprep(mt6, clusters = votefire2_complete$FullPrc, se_type = "stata")

## clusters, urban controls, trend controls,  and covid controls.

mt7 <-  lm( turnout2020 ~ PrecDist5000Fire_km  + WHPMean + pop + urban+ crit_work +
    remindex + proximity + exposure + turnout2016,
        data = votefire2_complete)

se7 <- starprep(mt7, clusters = votefire2_complete$FullPrc, se_type = "stata")

## clusters and fixed effects
```

```r
mt8 <-  lm( turnout2020 ~ PrecDist5000Fire_km  + WHPMean  + factor(CountyCd) + turnout2016,
        data = votefire2_complete)

se8 <- starprep(mt8, clusters = votefire2_complete$FullPrc, se_type = "stata")

## no clusters and fixed effects

mt9 <-  lm( turnout2020 ~ PrecDist5000Fire_km  + WHPMean  +  factor(CountyCd) + turnout2016,
        data = votefire2_complete)

se9 <- coef(summary(mt9))[,2]

stargazer(mt1,mt2, mt3, mt4, mt5, mt6, mt7, mt8, mt9,
           se = c(se1, se2, se3, se4, se5, se6, se7, se8, se9),
          header = FALSE,
          df=FALSE,
          style = "apsr",
          title = "Main result with different measure of fire exposure",
          keep = c("Intercept", "PrecDist5000Fire_km", "urban", "pop", "crit_work",
                   "remindex" , "proximity" , "exposure" ),
          add.lines = list(c("County FE", "", "", "", "", "", "", "", "X", "X"),
                           c("Covid Controls", "", "", "X", "", "X", "", "X", "", ""),
                           c("Trend Controls", "", "", "", "", "", "X", "X", "", ""),
                           c("Urban Controls", "", "", "", "X", "X", "", "X", "", ""),
                           c("Clusters", "", "X", "X", "X", "X", "X", "X", "X", "")),
          covariate.labels = c("Distance to nearest fire"),
          dep.var.caption = "Turnout 2020 (mean = ; sd = 0.",
          keep.stat = c("rsq", "n"),
          digits = 3)
```

Table 5: Main result with different measure of fire exposure

	turnout2020							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Distance to nearest fire	0.0001	0.0001	0.0001	0.0001***	0.0002***	0.0001***	0.0001***	−0.0003***
				(0.00002)	(0.00003)	(0.00002)	(0.00003)	(0.00002)
crit_work			−0.027		0.102***		0.069*	
					(0.039)		(0.038)	
remindex			0.066		0.185***		0.104***	
					(0.015)		(0.021)	
proximity			−0.014		−0.068**		0.024	
					(0.028)		(0.037)	
exposure			−0.015		−0.139***		−0.085**	
					(0.028)		(0.033)	
pop				0.000	−0.00000		−0.000***	
							(0.000)	
urban				−0.009	−0.017		−0.005	
							(0.005)	
County FE								X
Covid Controls			X		X		X	
Trend Controls						X	X	
Urban Controls				X	X		X	
Clusters		X	X	X	X	X	X	X
N	6,768	6,768	6,768	6,768	6,768	6,768	6,768	6,768
R^2	0.010	0.010	0.040	0.031	0.051	0.401	0.410	0.443

*p < .1; **p < .05; ***p < .01