Technical Appendix

Ziyan Zhu, Cheyenne Ehman, Yixuan Luo, Zi Yang

Appendix 1: Motivations

```
## Set up aesthetic theme for all graphs generated in the report
Sys.setlocale("LC_TIME", "English")
```

```
## [1] "English_United States.1252"
```

```r
library(ggrepel)
library(tidyverse)
library(lubridate)
require(scales)
library(readxl)
library(ggpubr)
library(PMCMRplus)
require(DescTools)
library(cowplot)
library(sp)
source("step2_new.R")
# color blind friendly Palette
library(ggthemes)
col_theme <- c("Hybrid"="#009E73","On Premises"="#D55E00","Online Only"="#0072B2")
## plot theme
grid_theme <- theme(axis.line = element_line(colour = "black"),
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    panel.border = element_blank(),
    legend.key = element_blank(),
    panel.background = element_blank(),
    legend.box="vertical", legend.margin=margin())
team_theme <- grid_theme+
  theme(legend.text = element_text(size=12),
        legend.title = element_text(size=12),
        axis.text = element_text(size=13),
        title=element_text(size=13),
        strip.text.x = element_text(size = 10, face = "bold.italic"))
```

1.1 Time series plot of daily cases of children under 19 years old

This plot serves as an motivation example to investigate on the transmissibility of children under 18 years old; then we are interested in whether children act as an vector of transmission after school reopening in the 2020 Fall semester.

```r
## process Age data
cases_by_age <- read_excel("OhiobyAge.xlsx")
```

```r
rolling_age_cases <- cases_by_age %>%
  mutate(youth_prop_roll = zoo::rollmean(`00_19/total(%)`, k = 7, fill = NA),
         all_roll = zoo::rollmean(`00_80+`, k = 7, fill = NA))
colors <- c("Total Daily Cases" = "black",
            "0-19 Age / Total Cases (%)" = "gray")
coeff <- 200
cases_by_age_long <- cases_by_age %>%
  gather(age_group, percent_cases,
         `00_19/total(%)`:`80+/total(%)`,
         factor_key=TRUE) %>%
  group_by(age_group) %>%
  mutate(roll_percent_cases= zoo::rollmean(percent_cases, k = 7, fill = NA))

colors <- c("Total Daily Cases" = "black",
            "0-19 Age / Total Cases (%)" = "gray")
coeff <- 200
ggplot(rolling_age_cases, aes(x=Date)) +
  geom_line( aes(y=youth_prop_roll,
                 color = "0-19 Age / Total Cases (%)"),
             na.rm = T)+
  geom_line( aes(y=all_roll/coeff,
                 color = "Total Daily Cases"),
             na.rm = T) +
  scale_y_continuous(
    # Features of the first axis
    name = "% Daily Cases for 0-19 Year Age Group",
    labels = function(x){paste0(x, "%")},
    # Add a second axis and specify its features
    sec.axis = sec_axis(~.*coeff, name="New Daily Cases for Entire Population",
                        label=comma)
  ) +
  geom_rect(data=rolling_age_cases[1,],
            aes(xmin=as.POSIXct ("2020/08/18"), xmax=as.POSIXct ("2020/12/12"),
                ymin=-Inf,ymax=Inf),
            color = NA,alpha=0.2, show.legend = F, fill = "orange") +
  geom_vline(xintercept = as.POSIXct ("2020/08/18") + days(7),lty = 2)+
  xlim(c(as.POSIXct ("2020/08/01"),as.POSIXct ("2021/01/01"))) +
  labs(title = "Total Cases and Percent of Cases in 0-19 Year Age Group",
       subtitle = "Yellow Area represents the fall semester",
       caption = "Smoothed using a 7 day moving average",
       color = "")+
  scale_color_manual(values = colors)+
  team_theme +
  theme(legend.position='bottom')
```

Total Cases and Percent of Cases in 0−19 Year Age Group

Yellow Area represents the fall semester

— 0−19 Age / Total Cases (%) — Total Daily Cases

Smoothed using a 7 day moving average
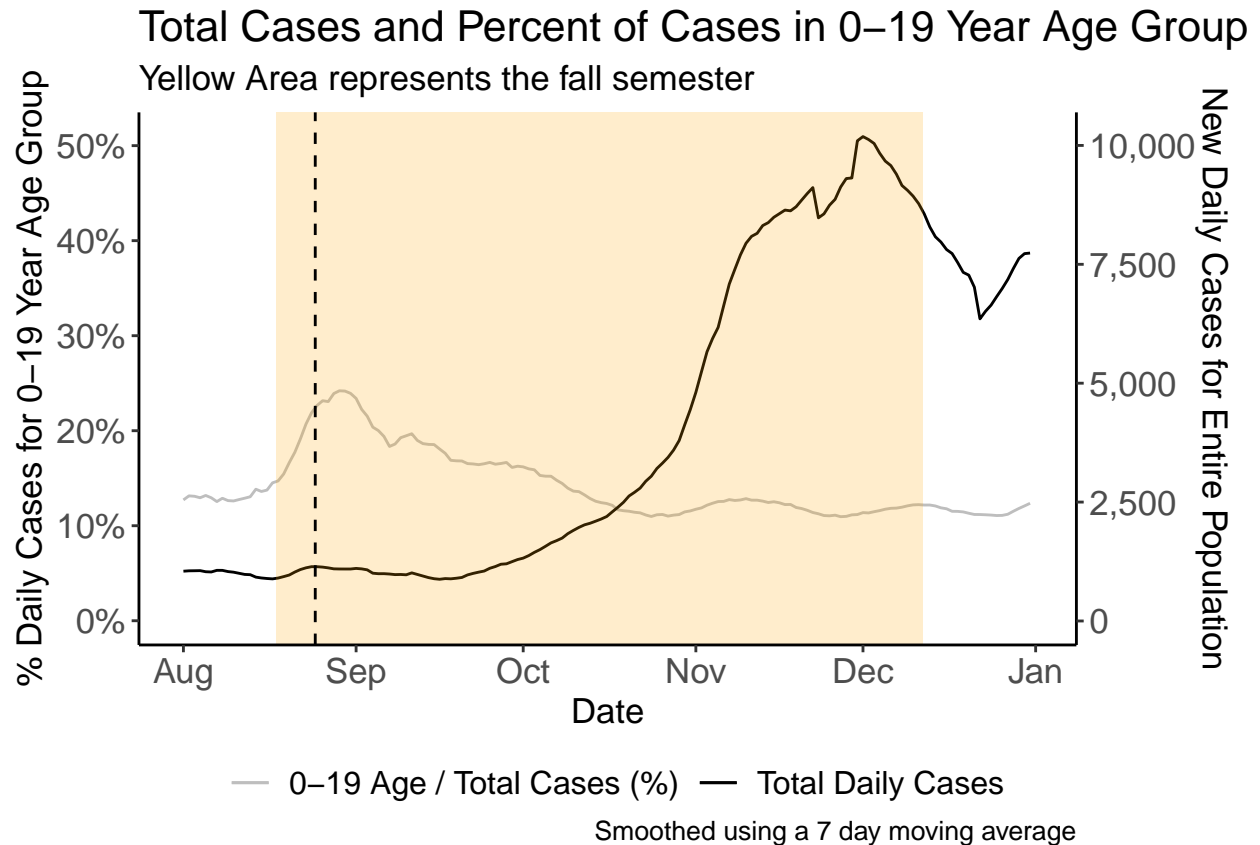
Appendix 2: Maps of Ohio

2.1 Geographical distribution of the teaching posture proportions, population and student enrollmen tat county-level

```r
ohio_map <- map_data("county") %>%subset(region=="ohio")%>%
  mutate(county=toupper(subregion))%>%dplyr::select(long,lat,county,group)

# Map of proportion of students taking online-only classes
wide_teaching_enroll%>%
  left_join(ohio_map,by='county')%>%
  mutate(Online_Only= Online_Only*100)%>%
  ggplot() +
  geom_polygon(aes(x = long, y = lat, group = group, fill = Online_Only), color = "gray") +
  coord_fixed(1.3) + theme_map() +
  scale_fill_distiller(palette = "OrRd",direction = 1)+
  labs(fill='% Online Only')+
  theme(legend.position = "bottom",
        legend.text = element_text(size=),
        legend.title = element_text(size=20))
```
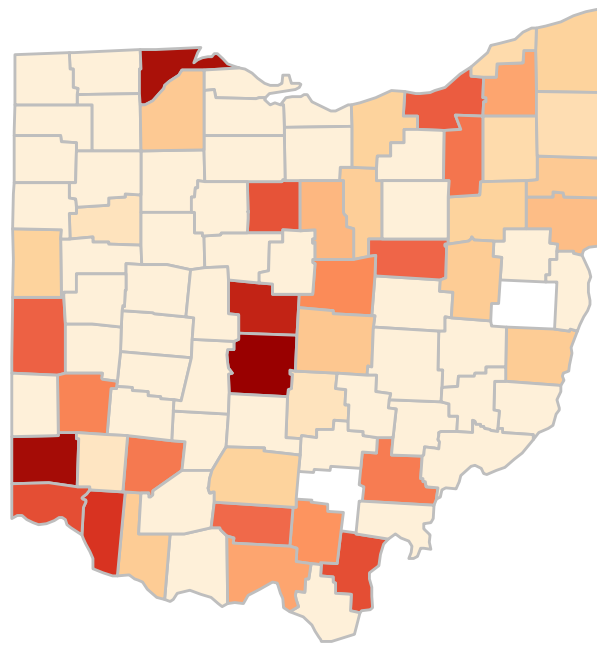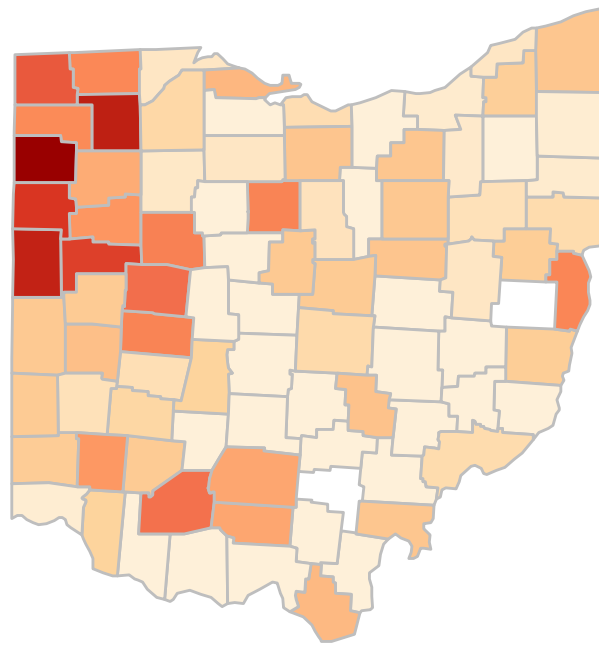
% Online Only

```
# Map of proportion of students taking on-premises classes
wide_teaching_enroll%>%
  left_join(ohio_map,by='county')%>%
  mutate(On_Premises= On_Premises*100)%>%
  ggplot() +
  geom_polygon(aes(x = long, y = lat, group = group, fill = On_Premises), color = "gray") +
  coord_fixed(1.3) + theme_map() +
  scale_fill_distiller(palette = "OrRd",direction = 1)+
  labs(fill='% On Premises')+
  theme(legend.position = "bottom",
        legend.text = element_text(size=),
        legend.title = element_text(size=20))
```
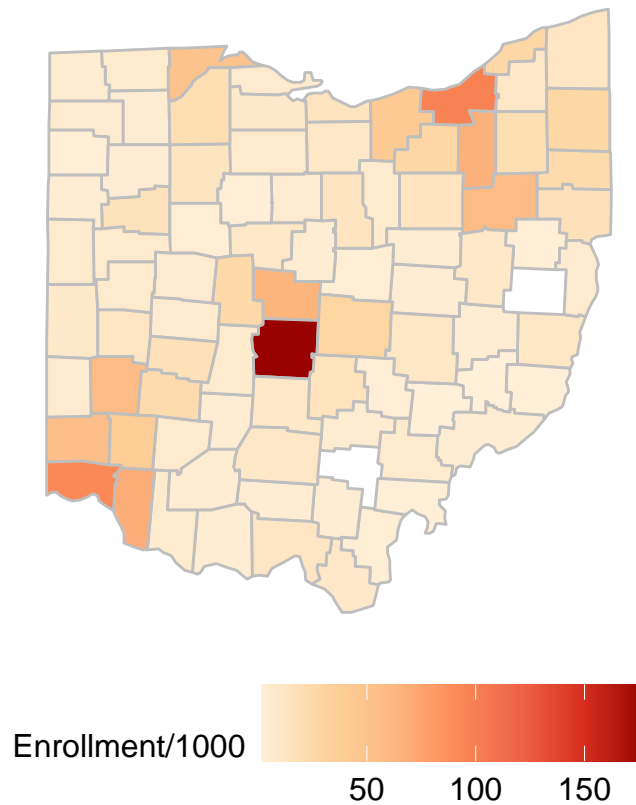
% On Premises

0   25   50   75   100

```r
# Map of proportion of students taking hybrid classes
wide_teaching_enroll%>%
  left_join(ohio_map,by='county')%>%
  mutate(Hybrid= Hybrid*100)%>%
  ggplot() +
  geom_polygon(aes(x = long, y = lat, group = group, fill = Hybrid), color = "gray") +
  coord_fixed(1.3) +
  theme_map() +
  scale_fill_distiller(palette = "OrRd",direction = 1)+
  labs(fill='% Hybrid')+
  theme(legend.position = "bottom",
        legend.text = element_text(size=),
        legend.title = element_text(size=20))
```

% Hybrid
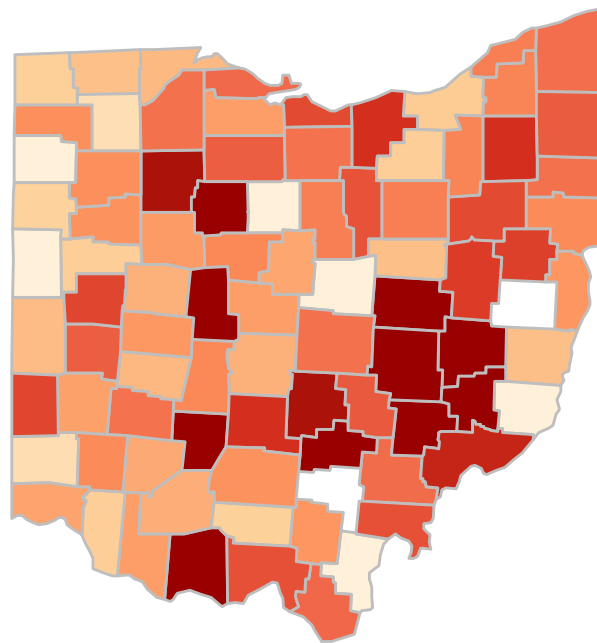
```r
# Map of population size
cases%>%
  distinct(COUNTY,POPULATION)%>%
  left_join(ohio_map,by=c('COUNTY'='county'))%>%
  mutate(population = POPULATION/1000)%>%
  ggplot() +
  geom_polygon(aes(x = long, y = lat, group = group, fill = population), color = "gray") +
  coord_fixed(1.3) + theme_map() +
  scale_fill_distiller(palette = "OrRd",direction = 1)+
  labs(fill='Population/1000')+
  theme(legend.text = element_text(size=12),
        legend.title = element_text(size=12),
        legend.position = "bottom",
        legend.key.size = unit(2,"lines"))
```

Population/1000
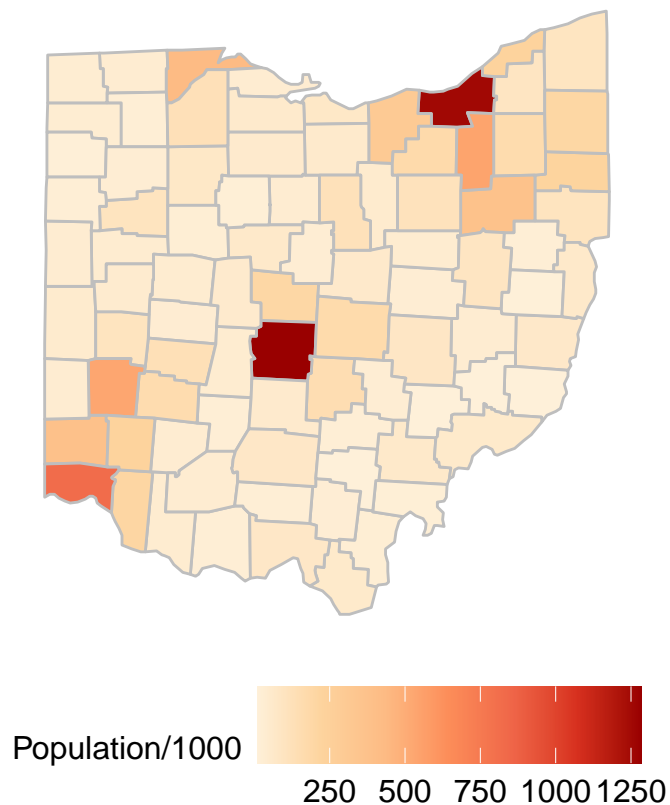
```r
# Map of student enrollments
teachingmethod_enroll%>%
  distinct(county,county_enroll)%>%
  left_join(ohio_map,by=c('county'))%>%
  mutate(county_enroll = county_enroll/1000)%>%
  ggplot() +
  geom_polygon(aes(x = long, y = lat, group = group, fill = county_enroll), color = "gray") +
  coord_fixed(1.3) + theme_map() +
  scale_fill_distiller(palette = "OrRd",direction = 1)+
  labs(fill='Enrollment/1000')+
  theme(legend.text = element_text(size=12),
        legend.title = element_text(size=12),
        legend.position = "bottom",
        legend.key.size = unit(2,"lines"))
```

Enrollment/1000

## 2.2 Geographical distribution of the majority teaching posture at county-level

```
wide_teaching_enroll%>%
  left_join(ohio_map,by='county')%>%
  mutate(On_Premises= On_Premises*100)%>%
  ggplot() + geom_polygon(aes(x = long, y = lat, group = group, fill = as.factor(major_teaching)), colo
  coord_fixed(1.3) + theme_map() +
  scale_fill_manual(values=col_theme)+
  labs(fill='Majority teaching posture')+
  theme(legend.position = "bottom",
        legend.text = element_text(size=14),
        legend.title = element_text(size=14))
```
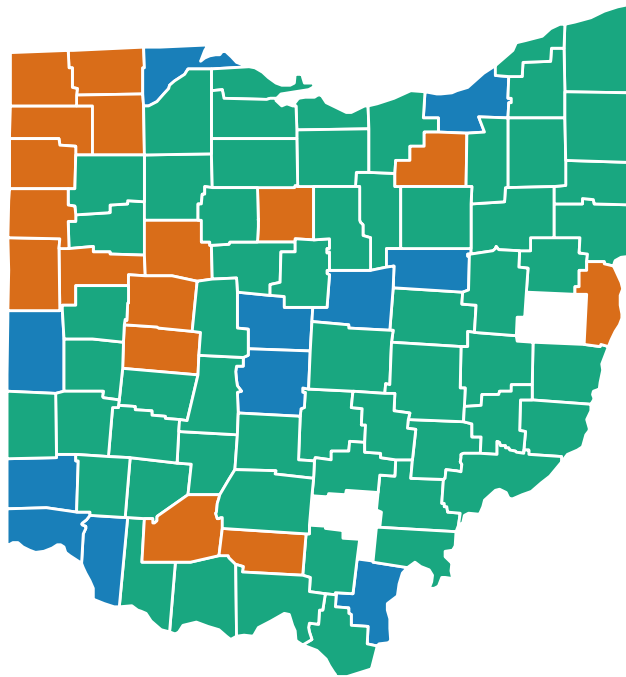
Majority teaching posture ■ Hybrid ■ On Premises ■ Online Only

2.3 Geographical distribution of cumulative COVID-19 deaths and deaths incidence until 02/22/2021.

```
death_prop%>%
  left_join(ohio_map,by=c("COUNTY"='county'))%>%
  ggplot() +
  geom_polygon(aes(x = long, y = lat, group=group,fill = CUMDEATHS), color = "gray")+
  coord_fixed(1.3) + theme_map() +
  scale_fill_distiller(palette = "OrRd",direction = 1)+
  labs(fill='Cumulative deaths \nuntil 2021-02-22')+
  theme(legend.text = element_text(size=12),
        legend.title = element_text(size=12),
        legend.position = "bottom",
        legend.key.size = unit(2,"lines"))
```
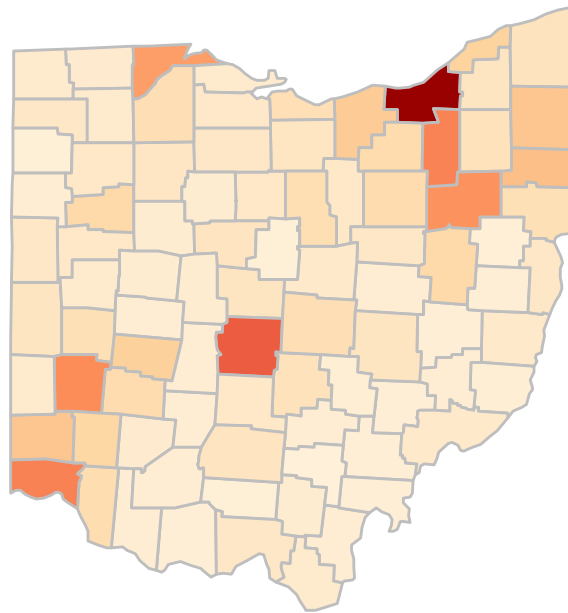
Cumulative deaths
until 2021−02−22
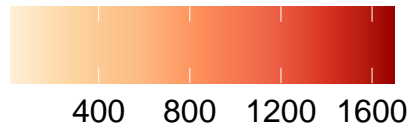
400   800   1200   1600

```
death_prop%>%
  left_join(ohio_map,by=c("COUNTY"='county'))%>%
  ggplot() +
  geom_polygon(aes(x = long, y = lat, group=group,fill = death_per_1000),
               color = "gray") +
  coord_fixed(1.3) + theme_map() +
  scale_fill_distiller(palette = "OrRd",direction = 1)+
  labs(fill='Cumulative deaths \nper 1000 people \nuntil 2021-02-22')+
  theme(legend.text = element_text(size=12),
        legend.title = element_text(size=12),
        legend.position = "bottom",
        legend.key.size = unit(2,"lines"))
```

Cumulative deaths
per 1000 people
until 2021−02−22   0.5   1.0   1.5   2.0   2.5
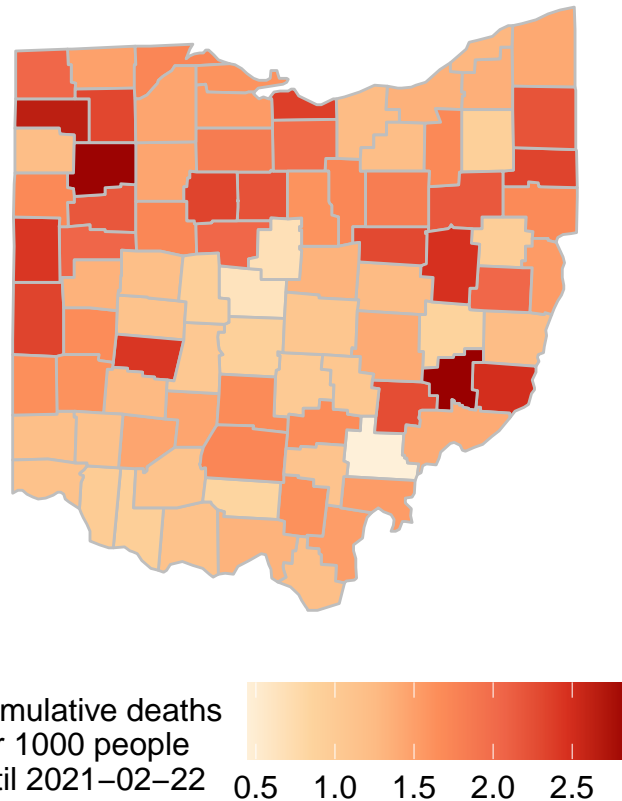
Appendix 3: Difference in death incidence during the Fall semester

3.1 One-way anova test on death incidence by teaching posture

```
county_open_teaching_enroll <- OH_K12%>%
  distinct(county,leaid,teachingmethod,county_enroll,district_enroll,date)%>%
  group_by(county,date,teachingmethod)%>%
  summarise(open_county_enroll = sum(district_enroll),
            opendate_prop = sum(district_enroll)/county_enroll)%>%
  rename(opendate = date)

major_reopening <- county_open_teaching_enroll%>%
  group_by(county)%>%
  slice(which.max(opendate_prop))%>%
  rename(COUNTY=county,major_opendate=opendate)%>%
  distinct(COUNTY,major_opendate,opendate_prop)


# see when the intesection happens
date.intercept <- as.Date("2020-11-24")
# add 95% confidence bans
confidence_level <- .95
z_cl <- qnorm(confidence_level)
# case_policy_wide
case_policy_wide <- cases %>%
```

```r
  left_join(county_policy_wide[,c("county","major_teaching",
                                  "Online_Only","Hybrid","On_Premises")],
            by = c("COUNTY" = "county")) %>%
  mutate(death_prop = CUMDEATHS/POPULATION)
opendate_cases <- case_policy_wide%>%
  inner_join(major_reopening%>%dplyr::select(COUNTY,major_opendate),by=c('COUNTY'))

# filter and summarize deaths incidence in the Fall
fall_cases <- opendate_cases %>%
  filter(DATE >= major_opendate & DATE <= as.Date("2020/12/15")) %>%
  group_by(COUNTY) %>%
  arrange(DATE) %>%
  filter(row_number()==1 | row_number()==n()) %>%
  mutate(death_incidence = diff(CUMDEATHS),
         death_incidence_per_1000 = death_incidence*1000/POPULATION) %>%
  distinct(COUNTY,POPULATION,major_teaching,
           death_incidence,death_incidence_per_1000)

# one-way ANOVA
fall_major_teaching.aov <- aov(death_incidence_per_1000 ~ major_teaching,
                               data = fall_cases)

summary(fall_major_teaching.aov) # p-value of .012
```

```
##                Df Sum Sq Mean Sq F value  Pr(>F)
## major_teaching  2  1.653  0.8264   5.205 0.00761 **
## Residuals      76 12.067  0.1588
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The one-way ANOVA test shows that there is at least one pair of counties grouped by majority teaching posture that have significantly different death incidences during Fall.

3.2 Pairwise test of death incidences by the majority teaching posture

Following the significant result in the one-way ANOVA test, we conduct a Duncan posthoc test to figure out which pair of teaching postures have significantly different death incidences during Fall.

```r
## Duncan test after significant ANOVA test
stat.test <- PostHocTest(fall_major_teaching.aov, method = "duncan")$major_teaching%>%
  as.data.frame()%>%
  rownames_to_column("group") %>%
  separate(group,"-", into = c("group1","group2")) %>%
  mutate(pval = round(pval,3),
         p = case_when(pval <= .01~ "**",
                       pval <= .05 ~ "*",
                       TRUE ~ "NS"))%>%
  dplyr::select(group1, group2, pval, p)

# Box Plots with test statistics
ggplot(fall_cases,aes(y = death_incidence_per_1000, x = major_teaching)) +
  geom_boxplot(aes(fill = major_teaching))+
  stat_compare_means(method = "anova")+
  stat_pvalue_manual(stat.test, label = "p",y.position = 2.5, step.increase = 0.15)+
  ylim(c(0,3.5))+
```

```
theme_bw()+
labs(y = "Death Incidence / 1,000 people",x = "",
     fill = "Majority teaching posture",
     title = "Death Incidence in the Fall Semester",
     caption = "Pairwise p-values come from Duncan pairwise comparison test") +
theme(legend.position = "bottom",
      axis.text.x=element_blank())+team_theme+
scale_colour_manual(values=col_theme)+scale_fill_manual(values=col_theme)
```

Death Incidence in the Fall Semester

Pairwise p–values come from Duncan pairwise comparison test
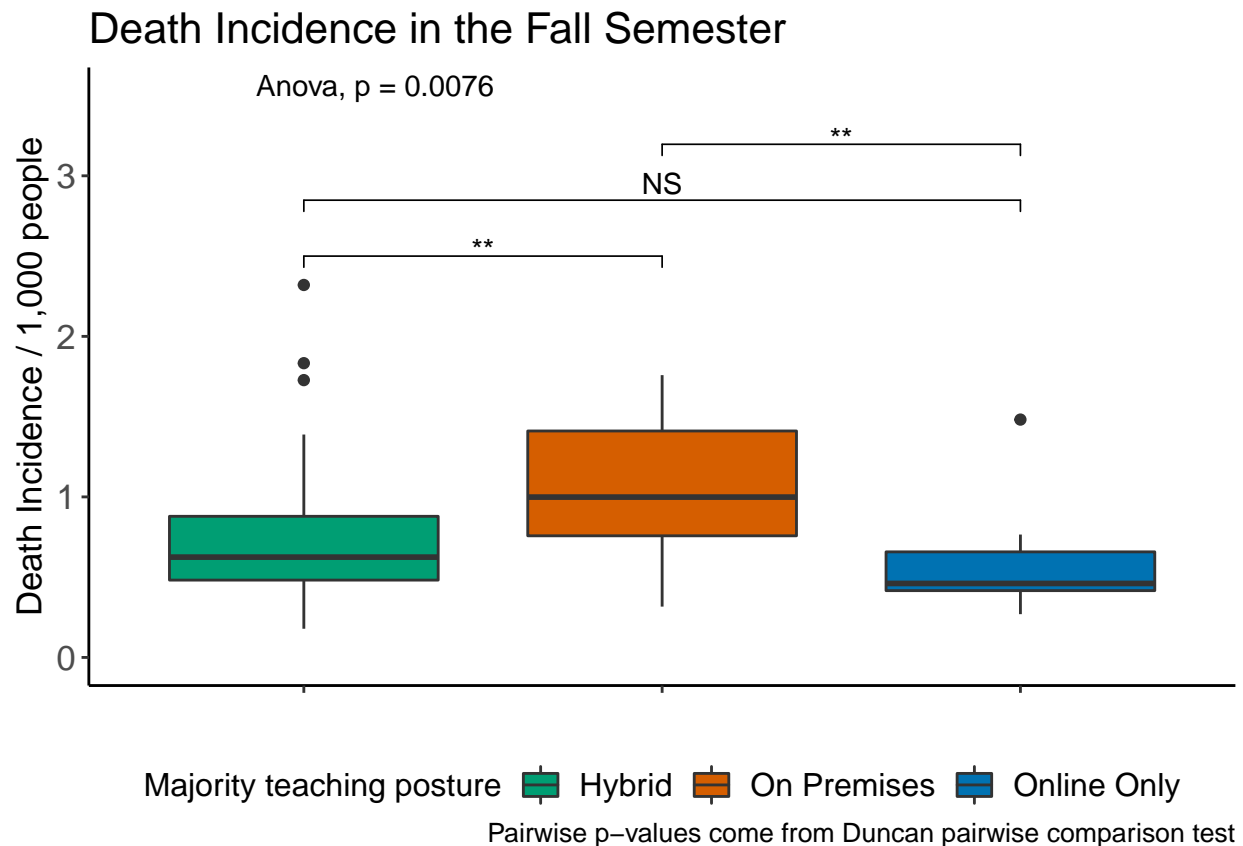
3.3 Time series of death incidences by the majority teaching posture

```
case_policy_wide%>%
  group_by(DATE, major_teaching) %>%
  drop_na(major_teaching)%>%
  summarise(total_deaths = sum(CUMDEATHS),
            total_pop = sum(POPULATION),
            death_prop = total_deaths/total_pop,
            death_prop_upper = death_prop + z_cl*sqrt(death_prop*(1 - death_prop)/total_pop),
            death_prop_lower = death_prop - z_cl*sqrt(death_prop*(1 - death_prop)/total_pop),
            .groups = "drop") %>%
  ggplot(aes(x = DATE, y = death_prop*1000, group = major_teaching))+
  geom_rect(data=case_policy_wide[1,],
            aes(xmin=as.Date("2020/08/18"), xmax=as.Date("2020/12/15"),
                ymin=-Inf,ymax=Inf),
```

```
           color = NA,alpha=0.2, show.legend = F, fill = "#E69F00") +
geom_line(aes(color = major_teaching),size = 1, alpha = .8) +
geom_ribbon(aes(ymin = 1000*death_prop_lower, ymax = 1000*death_prop_upper,
             fill= major_teaching),
           alpha = .3, show.legend = F)+
geom_vline(xintercept = date.intercept, linetype = "dashed") +
annotate("text",x = date.intercept,y = 1.5,
         label = date.intercept,
         hjust = 1.1) +
team_theme + theme(legend.position = "bottom")+
ggtitle("Death Incidences Increase Faster for Red Counties ")+
labs(x = "Date", y = "Cumulative Death Incidence / 1,000 people",
    subtitle = "Yellow area represents Fall Semester",
    color = "Majority teaching posture") +team_theme+
scale_colour_manual(values=col_theme)+scale_fill_manual(values=col_theme)
```

Death Incidences Increase Faster for Red Counties
Yellow area represents Fall Semester
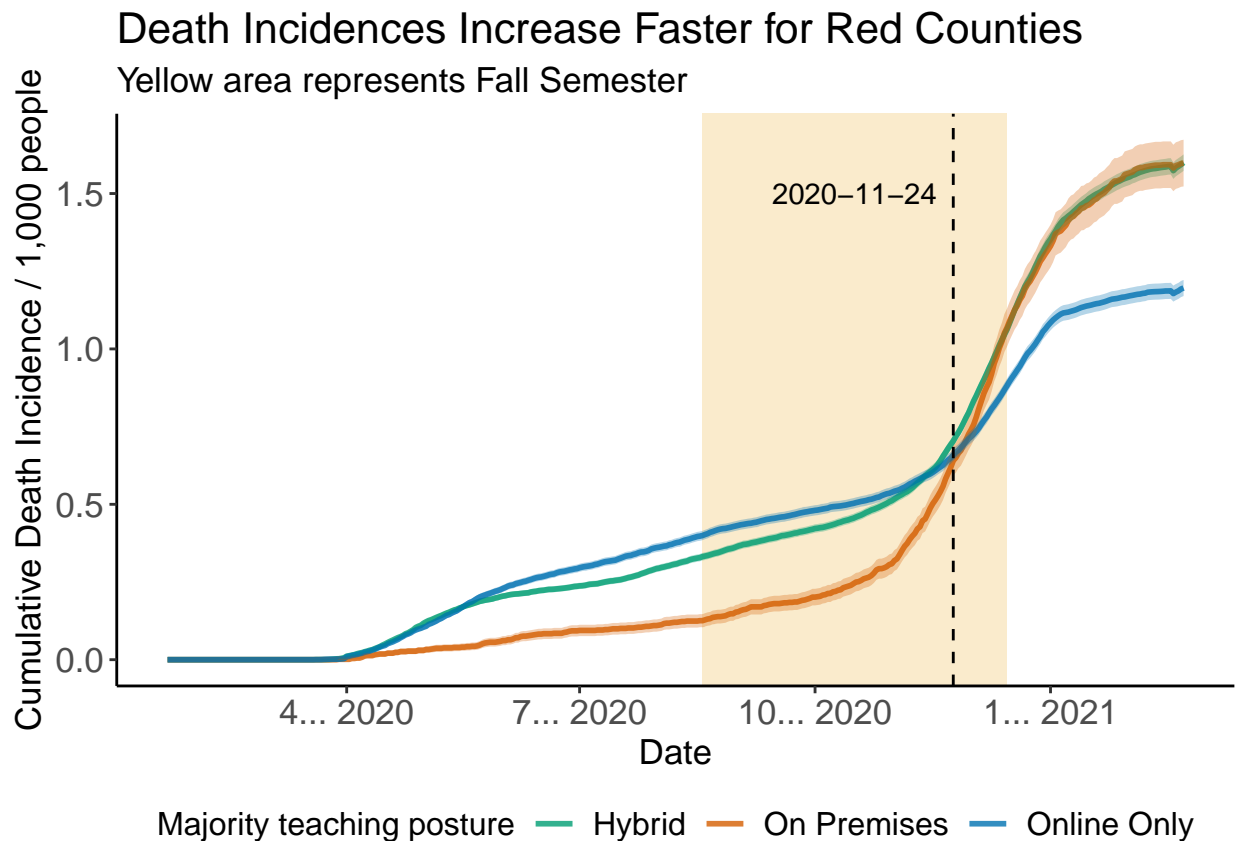
Appendix 4: Confounding Variables

Since what we show above are results aggregated by majority teaching posture. We should notice that there could be other essential factors that are significantly different between counties taking different act in school operations.

4.1 Distribution of percent of uninsured population by majority teaching posture

```r
library(ggpubr)
library(PMCMRplus)
require(DescTools)

ohio_profile <- read.csv("county_level_latest_data_for_ohio.csv")
ohio_profile <- ohio_profile[,c(1,14:20,38:50)]
names(ohio_profile)[1]<-"County"
ohio_profile$County <- toupper(ohio_profile$County)

# set up data
teaching_profile <- ohio_profile%>%
  inner_join(wide_teaching_enroll,by=c("County"="county"))
teaching_profile$major_teaching <- factor(teaching_profile$major_teaching,
                                           levels = c("On Premises","Hybrid","Online Only"))

# one-way ANOVA test
profile_major_teaching.aov <- aov(Percent.uninsured ~ major_teaching,data = teaching_profile)
summary(profile_major_teaching.aov)
```
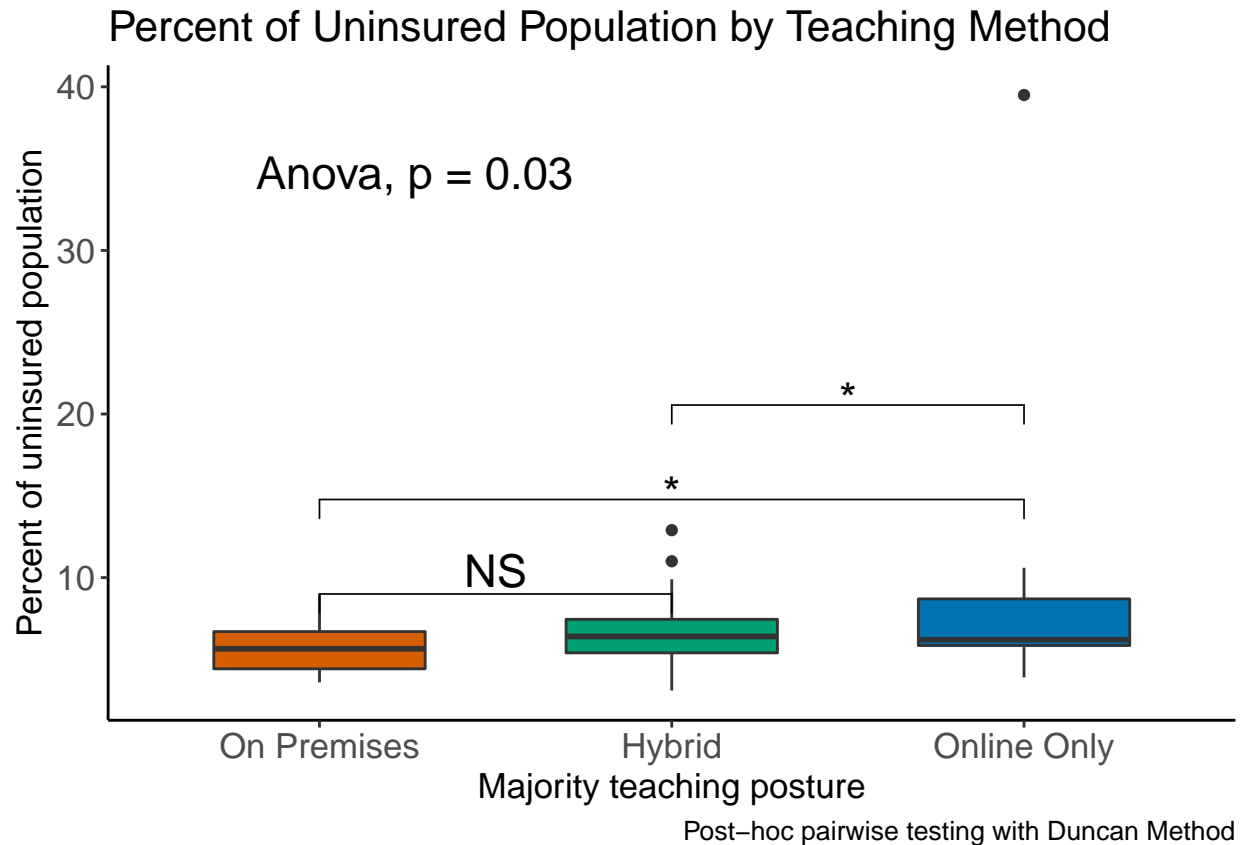
```
##               Df Sum Sq Mean Sq F value Pr(>F)
## major_teaching  2  108.8    54.38   3.645 0.0304 *
## Residuals      83 1238.2    14.92
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Duncan test, p-value of .012
stat.test <- PostHocTest(profile_major_teaching.aov, method = "duncan")$major_teaching %>%
  as.data.frame()%>%
  rownames_to_column("group") %>%
  separate(group,"-", into = c("group1","group2")) %>%
  mutate(pval = round(pval,3),
         p = case_when(pval <= .01~ "**",
                       pval <= .05 ~ "*",
                       TRUE ~ "NS"))%>%
  dplyr::select(group1, group2, pval, p)

teaching_profile%>%
  ggplot(aes(x=major_teaching,y=Percent.uninsured))+
  geom_boxplot(aes(fill=major_teaching),width=0.6)+
  stat_compare_means(method = "anova",size=6,label.y.npc=0.85)+
  stat_pvalue_manual(stat.test, label = "p",y.position = 1,
                     step.increase = 0.15,size = 6,bracket.nudge.y = 8)+
  labs(title="Percent of Uninsured Population by Teaching Method",
       x="Majority teaching posture",
       y="Percent of uninsured population",
       caption = "Post-hoc pairwise testing with Duncan Method")+
  team_theme+theme(legend.position = "")+scale_fill_manual(values=col_theme)
```

Percent of Uninsured Population by Teaching Method

Anova, p = 0.03

Post–hoc pairwise testing with Duncan Method

4.2 Distribution of percent of senior population by majority teaching posture

```r
# one-way ANOVA
senior_major_teaching.aov <- aov(Percent.Population.65..yrs~ major_teaching,data = teaching_profile)

summary(senior_major_teaching.aov)

##               Df Sum Sq Mean Sq F value  Pr(>F)
## major_teaching  2   61.8   30.88   5.297 0.00684 **
## Residuals      83  483.9    5.83
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
# Duncan test: p-value of .012
stat.test <- PostHocTest(senior_major_teaching.aov, method = "duncan")$major_teaching %>%
  as.data.frame()%>%
  rownames_to_column("group") %>%
  separate(group,"-", into = c("group1","group2")) %>%
  mutate(pval = round(pval,3),
         p = case_when(pval <= .01~ "**",
                       pval <= .05 ~ "*",
                       TRUE ~ "NS"))%>%
  dplyr::select(group1, group2, pval, p)
# boxplot
teaching_profile%>%
  ggplot(aes(x=major_teaching,y=Percent.Population.65..yrs))+
```

```
geom_boxplot(aes(fill=major_teaching),width=0.6)+
stat_compare_means(method = "anova",size=6,label.y.npc=0.95)+
stat_pvalue_manual(stat.test, label = "p",y.position = 1,
                   step.increase = 0.15,size = 5,bracket.nudge.y = 2)+
labs(title="Percent of 65years+ Population by Teaching Method",
    x="Major Teaching Method",y="Percent of 65+yrs population",
    fill="Majority teaching posture",
    caption = "Post-hoc pairwise testing with Duncan Method")+team_theme+
theme(legend.position = "")+
scale_fill_manual(values=col_theme)
```
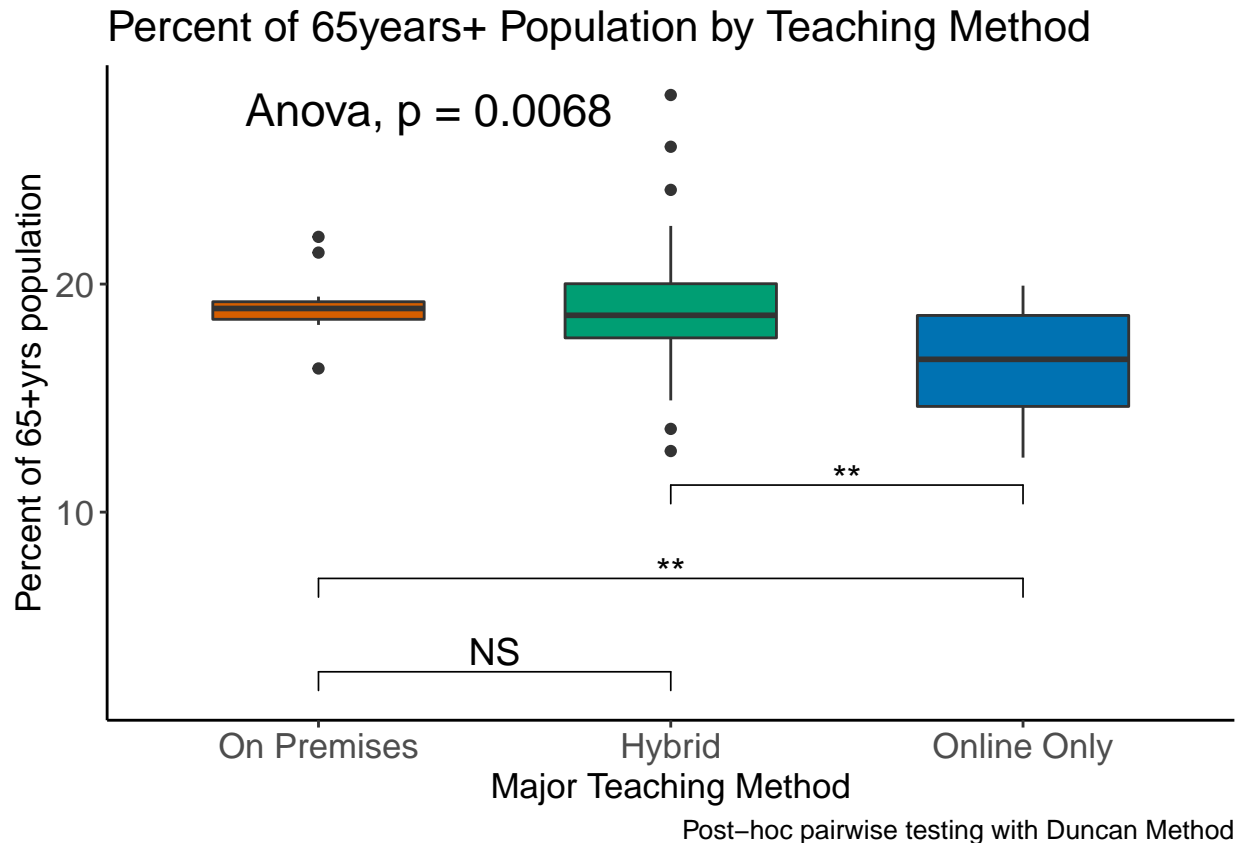

4.3 Distribution of log population density by majority teaching posture

We rescale the population density by log transformation because we want to better show the ylab.

```
# one-way ANOVA
pop_den_major_teaching.aov <- aov(log(Population.density)~ major_teaching,data = teaching_profile)

summary(pop_den_major_teaching.aov)

##               Df Sum Sq Mean Sq F value   Pr(>F)
## major_teaching  2  17.04   8.520   10.69 7.41e-05 ***
## Residuals      83  66.16   0.797
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Duncan test: p-value of .012
stat.test <- PostHocTest(pop_den_major_teaching.aov, method = "duncan")$major_teaching %>%
  as.data.frame()%>%
  rownames_to_column("group") %>%
  separate(group,"-", into = c("group1","group2")) %>%
  mutate(pval = round(pval,3),
         p = case_when(pval <= .01~ "**",
                       pval <= .05 ~ "*",
                       TRUE ~ "NS"))%>%
  dplyr::select(group1, group2, pval, p)

# boxplot
teaching_profile%>%
  ggplot(aes(x=major_teaching,y=log(Population.density)))+
  geom_boxplot(aes(fill=major_teaching),width=0.6)+
  stat_compare_means(method = "anova",size=6,label.y.npc=0.95)+
  stat_pvalue_manual(stat.test, label = "p",y.position = 1, step.increase = 0.15,
                     size = 6,bracket.nudge.y = -0.5)+
  labs(title="Log Population Density by Teaching Method",
       x="Majority teaching posture",
       y="Log Population Density",
       caption = "Post-hoc pairwise testing with Duncan Method")+
  team_theme+theme(legend.position = "")+scale_fill_manual(values=col_theme)
```


Log Population Density by Teaching Method

Post–hoc pairwise testing with Duncan Method
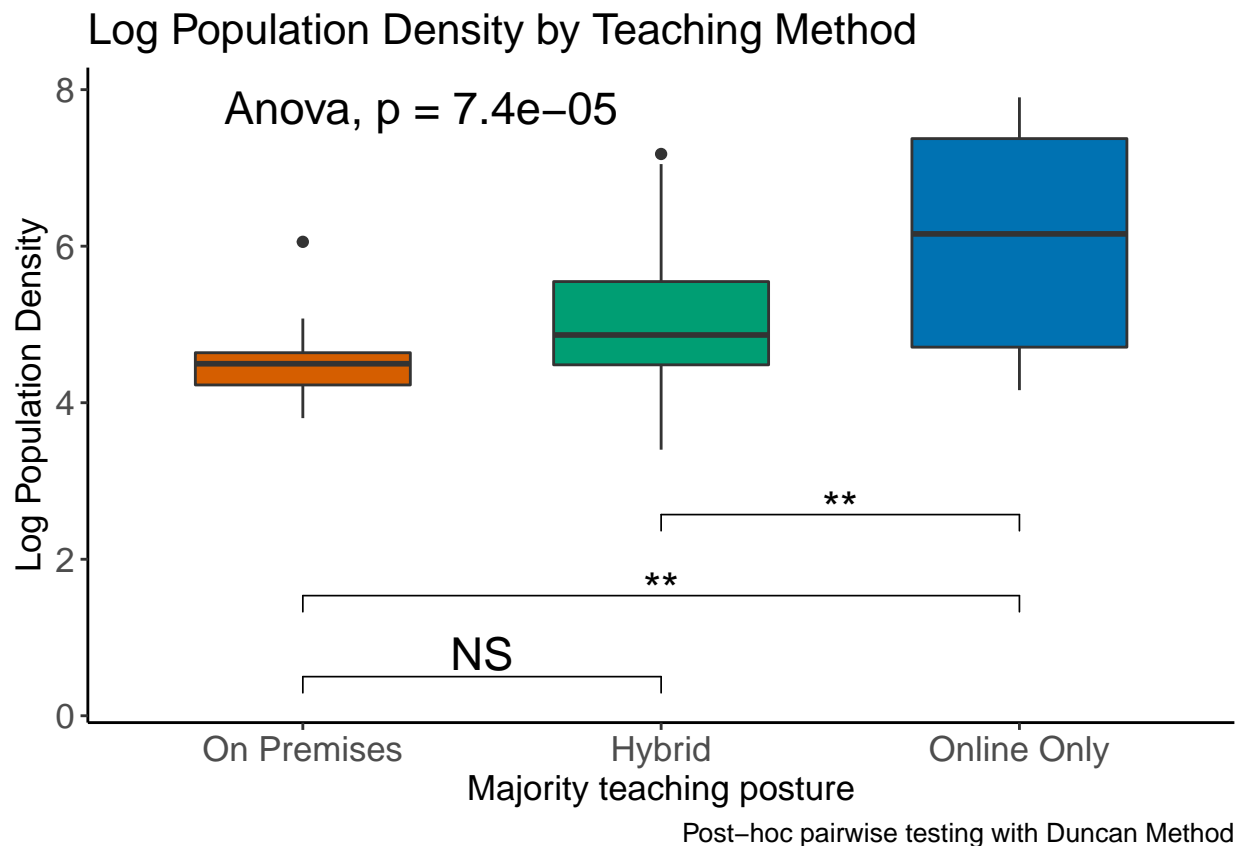
Appendix 5: Exponential growth model

We construct the exponential growth model to measure the state of pandemic. Please refer to the details in our Methods section.

5.1 Process growth coefficent

```r
# read in the estimated coefficients
cases_slope <- read.csv("county_splines.csv", header = T)%>%
  dplyr::select(COUNTY,DATE,POPULATION,CUMDEATHS,
                log_tot_deaths,tot.slope,NEWDEATHS,rev_NEWDEATHS,
                log_new_deaths,new.slope)

# SHIFT THE DATE 24 days forward
cases_slope$DATE <- as.Date(cases_slope$DATE)-24

# get Majority teaching posture wide_teaching_enroll
cases_slope_teach <-death_teaching%>%
  dplyr::select(-DATE,-POPULATION,-CUMDEATHS,-NEWDEATHS)%>%
  distinct()%>%
  right_join(cases_slope,by=c("COUNTY"))%>%
  filter(DATE>as.Date("2020-01-23"))
write.csv(cases_slope_teach,"cases_slope_teach.csv",row.names = F)


## ordering the teaching method factor to ensure the color order
cases_slope_teach$major_teaching <- factor(cases_slope_teach$major_teaching,
                                           levels = c("On Premises","Hybrid","Online Only"))
cases_slope_teach$DATE <- as.Date(cases_slope_teach$DATE)
```

5.2 Compute the maximum B values for each county during the Fall semester

Maximum B values represent the severity of the disease for the county during the Fall semester.

```r
maxB1 <- cases_slope_teach%>%
  group_by(COUNTY)%>%
  filter(DATE >= as.Date("2020-08-18") & DATE<=as.Date("2020-12-15"))%>%
  summarise(max_B1 = max(new.slope), .groups = 'drop')
avgB1 <- cases_slope_teach%>%
  group_by(COUNTY)%>%
  filter(DATE >= as.Date("2020-08-18") & DATE<=as.Date("2020-12-15"))%>%
  summarise(avg_B1 = mean(new.slope), .groups = 'drop')
## avg3w_B0
## average B0 of the first 3 weeks of school reopening
## avg1w_2w_B0
## OR average B0s between  2020-08-18 -7days and +14days
##[before the rate bounce back around the dashed line]
## avg3w_bf_B0 ## OR average B0s between  2020-08-18 -21days and 2020-08-18
##[before the rate bounce back around the dashed line]
avgB0 <- cases_slope_teach%>%
  group_by(COUNTY)%>%
  filter(DATE > as.Date("2020-08-18") & DATE<as.Date(major_opendate)+21)%>%
  summarise(avg3w_B0 = mean(new.slope), .groups = 'drop')%>%
  left_join(cases_slope_teach%>%
```

```r
  group_by(COUNTY)%>%
  filter(DATE > as.Date("2020-08-18")-7 & DATE<as.Date("2020-08-18")+14)%>%
  summarise(avg1w_2w_B0 = mean(new.slope)),by="COUNTY", .groups = 'drop')%>%
  left_join(cases_slope_teach%>%
  group_by(COUNTY)%>%
  filter(DATE < as.Date("2020-08-18") & DATE>=as.Date("2020-08-18")-21)%>%
  summarise(avg3w_bf_B0 = mean(new.slope)),by="COUNTY", .groups = 'drop')
#  B0 and B1
B0B1 <- death_teaching%>%
  distinct(COUNTY,POPULATION,NCHS.Urban.Rural.Status,Population.density)%>%
  left_join(maxB1,by="COUNTY")%>%
  left_join(wide_teaching_enroll, by = c("COUNTY" = "county"))%>%
  left_join(avgB1,by="COUNTY")%>%
  left_join(avgB0,by="COUNTY") %>%
  left_join(avg_mobility,by="COUNTY")
## ordering the teaching method factor to ensure the color order
B0B1$major_teaching <- factor(B0B1$major_teaching,levels = c("On Premises","Hybrid","Online Only"))
```

5.3 Time series of growth coefficient aggregation by counties majority teaching posture

```r
cases_slope_teach_agg <- cases_slope_teach %>%
  drop_na(major_teaching)%>%
  group_by(DATE, major_teaching) %>%
  summarise(total_new_deaths = sum(rev_NEWDEATHS), .groups = "drop") %>%
  mutate(log_new_deaths = log(total_new_deaths + 1)) %>%
  group_by(major_teaching) %>%
  mutate(smooth.spline = smooth.spline(DATE,log_new_deaths,df = 398/28)$y,
         B = predict(smooth.spline(DATE,log_new_deaths,df = 398/28),deriv = 1)$y,
         B2 = predict(smooth.spline(DATE,log_new_deaths,df = 398/28),deriv = 2)$y)
week3_after_start <- as.Date("2020/08/18") + 21

####
ggplot(cases_slope_teach_agg, aes(x = DATE, color = major_teaching)) +
  geom_line(aes(y = B), size = 1) +
  geom_rect(data = cases_slope_teach_agg[1,],
            aes(xmin=as.Date("2020/08/18"), xmax=as.Date("2020/12/15"),
                ymin=-Inf,ymax=Inf),
            color = NA,alpha=0.2, show.legend = F, fill = "#E69F00") +
  geom_vline(xintercept = week3_after_start, lty = 2) +
  annotate("text",label = "3 Weeks After",
           x = week3_after_start, y = .05, hjust = 1.1)+
  geom_vline(xintercept = as.Date("2020/08/18")+42, lty = 2) +
  annotate("text",label = "6 Weeks After",
           x = as.Date("2020/08/18")+130, y = .06, hjust = 1.3)+
  labs(x = "Date", y = "Exponential Growth Coefficient",
       color = "Majority teaching posture",
       caption = "Smoothing window set to every 4 weeks",
       subtitle = "Yellow area represents the fall semester (08/18 - 12/15)") +
  theme(legend.position = "bottom")+
  team_theme+scale_color_manual(values=col_theme)
```

Yellow area represents the fall semester (08/18 – 12/15)

3 Weeks After

6 Weeks After

Majority teaching posture — On Premises — Hybrid — Online Only

Smoothing window set to every 4 weeks
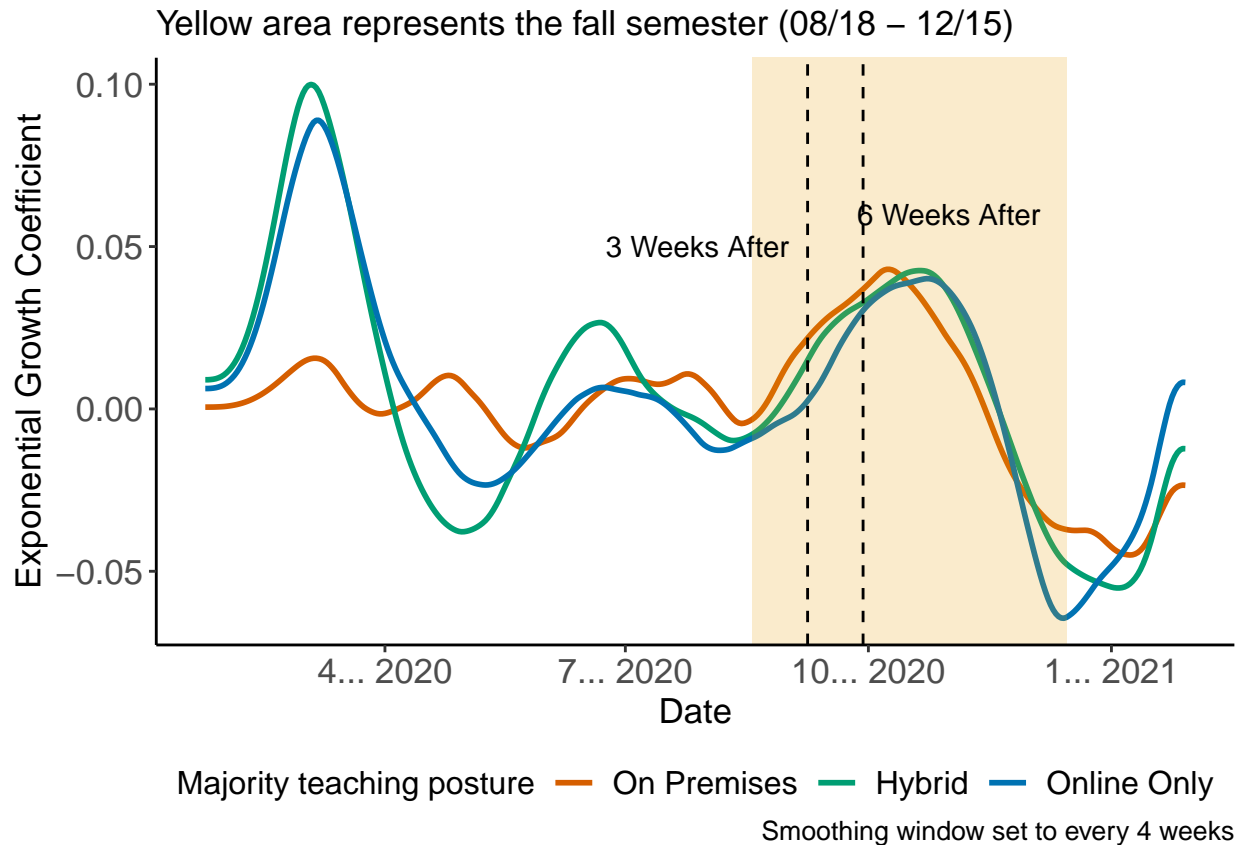
5.4 Compute the change in growth B values and corresponding average mobility for each county during the Fall semester

```r
B0w <- cases_slope_teach%>%
  filter(DATE==as.Date("2020-08-18"))%>%
  drop_na(major_teaching)%>%
  rename(new.slope0w=new.slope)
B1w <- cases_slope_teach%>%
  filter(DATE==as.Date("2020-08-18")+7)%>%
  drop_na(major_teaching)%>%
  rename(new.slope1w=new.slope)
B2w <- cases_slope_teach%>%
  filter(DATE==as.Date("2020-08-18")+14)%>%
  drop_na(major_teaching)%>%
  rename(new.slope2w=new.slope)
B3w <- cases_slope_teach%>%
  filter(DATE==as.Date("2020-08-18")+21)%>%
  drop_na(major_teaching)%>%
  rename(new.slope3w=new.slope)
B4w <- cases_slope_teach%>%
  filter(DATE==as.Date("2020-08-18")+28)%>%
  drop_na(major_teaching)%>%
  rename(new.slope4w=new.slope)
B5w <- cases_slope_teach%>%
  filter(DATE==as.Date("2020-08-18")+35)%>%
```

```r
  drop_na(major_teaching)%>%
  rename(new.slope5w=new.slope)
B6w <- cases_slope_teach%>%
  filter(DATE==as.Date("2020-08-18")+42)%>%
  drop_na(major_teaching)%>%
  rename(new.slope6w=new.slope)
B7w <- cases_slope_teach%>%
  filter(DATE==as.Date("2020-08-18")+49)%>%
  drop_na(major_teaching)%>%
  rename(new.slope7w=new.slope)

Bm1w <- cases_slope_teach%>%
  filter(DATE==as.Date("2020-08-18")-7)%>%
  drop_na(major_teaching)%>%
  rename(new.slopem1w=new.slope)
Bm2w <- cases_slope_teach%>%
  filter(DATE==as.Date("2020-08-18")-14)%>%
  drop_na(major_teaching)%>%
  rename(new.slopem2w=new.slope)
Bm3w <- cases_slope_teach%>%
  filter(DATE==as.Date("2020-08-18")-21)%>%
  drop_na(major_teaching)%>%
  rename(new.slopem3w=new.slope)

avg_mobi_0w3w <- case_mobility%>%
  left_join(major_reopening,by=c("COUNTY"))%>%
  group_by(COUNTY)%>%
  filter(DATE >= as.Date("2020-08-18")& DATE <as.Date("2020-08-18") + 21)%>%
  summarise(avg_full_work_prob = mean(full_work_prop_7d))

avg_mobi_3w6w <- case_mobility%>%
  left_join(major_reopening,by=c("COUNTY"))%>%
  group_by(COUNTY)%>%
  filter(DATE >= as.Date("2020-08-18")+ 21 & DATE <=as.Date("2020-08-18") + 42)%>%
  summarise(avg2_full_work_prob = mean(full_work_prop_7d))

# Before slope mobility
avg_mobi_m1w2w <- case_mobility%>%
  left_join(major_reopening,by=c("COUNTY"))%>%
  group_by(COUNTY)%>%
  filter(DATE >= as.Date("2020-08-18")-7 & DATE <=as.Date("2020-08-18") + 14)%>%
  summarise(avg_full_work_prob_m1w2w = mean(full_work_prop_7d))

avg_mobi_m2w1w <- case_mobility%>%
  left_join(major_reopening,by=c("COUNTY"))%>%
  group_by(COUNTY)%>%
  filter(DATE >= as.Date("2020-08-18")-14 & DATE <=as.Date("2020-08-18") + 7)%>%
  summarise(avg_full_work_prob_m2w1w = mean(full_work_prop_7d))

avg_mobi_m3w0w <- case_mobility%>%
  left_join(major_reopening,by=c("COUNTY"))%>%
  group_by(COUNTY)%>%
  filter(DATE >= as.Date("2020-08-18")-21 & DATE <=as.Date("2020-08-18"))%>%
```

```r
  summarise(avg_full_work_prob_m3w0w = mean(full_work_prop_7d))

# After slope mobility
avg_mobi_1w4w <- case_mobility%>%
  left_join(major_reopening,by=c("COUNTY"))%>%
  group_by(COUNTY)%>%
  filter(DATE >= as.Date("2020-08-18")+7 & DATE <=as.Date("2020-08-18")+28)%>%
  summarise(avg_full_work_prob_1w4w = mean(full_work_prop_7d))

avg_mobi_2w5w <- case_mobility%>%
  left_join(major_reopening,by=c("COUNTY"))%>%
  group_by(COUNTY)%>%
  filter(DATE >= as.Date("2020-08-18")+14 & DATE <=as.Date("2020-08-18")+35)%>%
  summarise(avg_full_work_prob_2w5w = mean(full_work_prop_7d))

avg_mobi_4w7w <- case_mobility%>%
  left_join(major_reopening,by=c("COUNTY"))%>%
  group_by(COUNTY)%>%
  filter(DATE >= as.Date("2020-08-18")+28 & DATE <=as.Date("2020-08-18")+49)%>%
  summarise(avg_full_work_prob_4w7w = mean(full_work_prop_7d))

# Construct B_diff
B_diff <- B6w[,c(1:9,13,20)]%>%
  left_join(B3w%>%dplyr::select(COUNTY,new.slope3w),by="COUNTY")%>%
  left_join(B0w%>%dplyr::select(COUNTY,new.slope0w),by="COUNTY")%>%
  left_join(B1w%>%dplyr::select(COUNTY,new.slope1w),by="COUNTY")%>%
  left_join(B2w%>%dplyr::select(COUNTY,new.slope2w),by="COUNTY")%>%
  left_join(B4w%>%dplyr::select(COUNTY,new.slope4w),by="COUNTY")%>%
  left_join(B5w%>%dplyr::select(COUNTY,new.slope5w),by="COUNTY")%>%
  left_join(B7w%>%dplyr::select(COUNTY,new.slope7w),by="COUNTY")%>%
  left_join(Bm1w%>%dplyr::select(COUNTY,new.slopem1w),by="COUNTY")%>%
  left_join(Bm2w%>%dplyr::select(COUNTY,new.slopem2w),by="COUNTY")%>%
  left_join(Bm3w%>%dplyr::select(COUNTY,new.slopem3w),by="COUNTY")%>%
  mutate(new.slope.diff = new.slope3w-new.slope0w,
         new.slope.diff2 = new.slope6w-new.slope3w,
         new.slope.diff2m1 = new.slope2w-new.slopem1w,
         new.slope.diff1m2 = new.slope1w-new.slopem2w,
         new.slope.diff0m3 = new.slope0w-new.slopem3w,
         new.slope.diff52 = new.slope5w-new.slope2w,
         new.slope.diff41 = new.slope4w-new.slope1w,
         new.slope.diff74 = new.slope7w-new.slope4w)%>%
  left_join(avg_mobi_0w3w,by="COUNTY")%>%
  left_join(avg_mobi_3w6w,by="COUNTY")%>%
  left_join(avg_mobi_m1w2w,by="COUNTY")%>%
  left_join(avg_mobi_m2w1w,by="COUNTY")%>%
  left_join(avg_mobi_m3w0w,by="COUNTY")%>%
  left_join(avg_mobi_1w4w,by="COUNTY")%>%
  left_join(avg_mobi_2w5w,by="COUNTY")%>%
  left_join(avg_mobi_4w7w,by="COUNTY")
B_diff$major_teaching <- factor(B_diff$major_teaching,
                                levels = c("On Premises","Hybrid","Online Only"))
```

5.5 Distribution of maximum growth coefficient for all counties by majority teaching posture

Online-only counties have highest max B value, which is not what we expected.

```
na.omit(B0B1)%>%
  ggplot(aes(x=major_teaching,y=max_B1))+geom_boxplot(aes(fill=major_teaching),width=0.6)+
  labs(title="",x="Majority teaching posture",y="Max Growth",fill="Majority teaching posture")+
  team_theme+theme(legend.position = "bottom")+
  scale_fill_manual(values=col_theme)
```
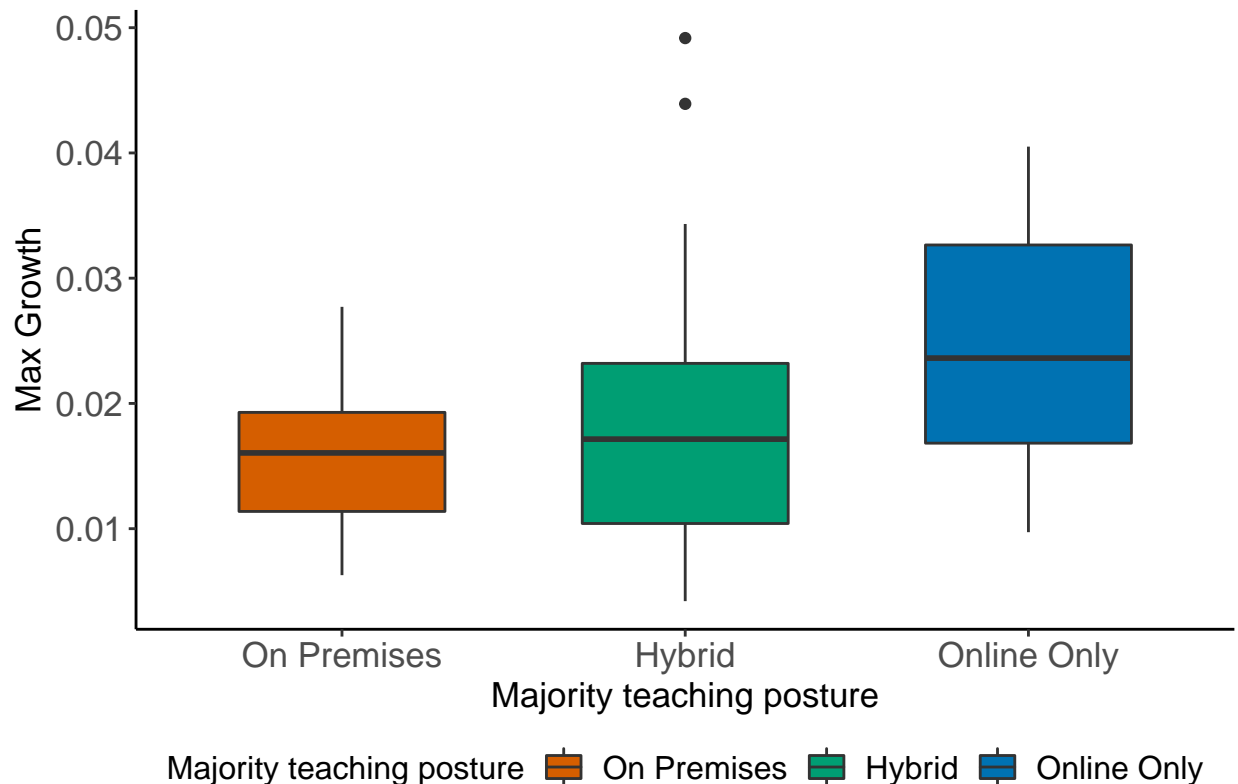

5.6 Maximum growth coefficient vs. mobility and population density for all counties

```
na.omit(B0B1)%>%
  drop_na(major_teaching)%>%
  ggplot(aes(x=avg_full_work_prob,y=max_B1,group=major_teaching,color=major_teaching))+
  geom_point(aes(size=log(Population.density),alpha=log(Population.density)))+
  geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+
  labs(y="Max B",
       x="Averaged \n%6hr+ Away Home in Fall semester",
       title="Max B in Fall v.s. Average Mobility \nAll Counties",
       color="Majority teaching posture",
       size = "Log of Population Density",
       alpha= "Log of Population Density" )+
  team_theme+theme(legend.position = "bottom")+
```

```
scale_color_manual(values=col_theme)
```

Max B in Fall v.s. Average Mobility
All Counties
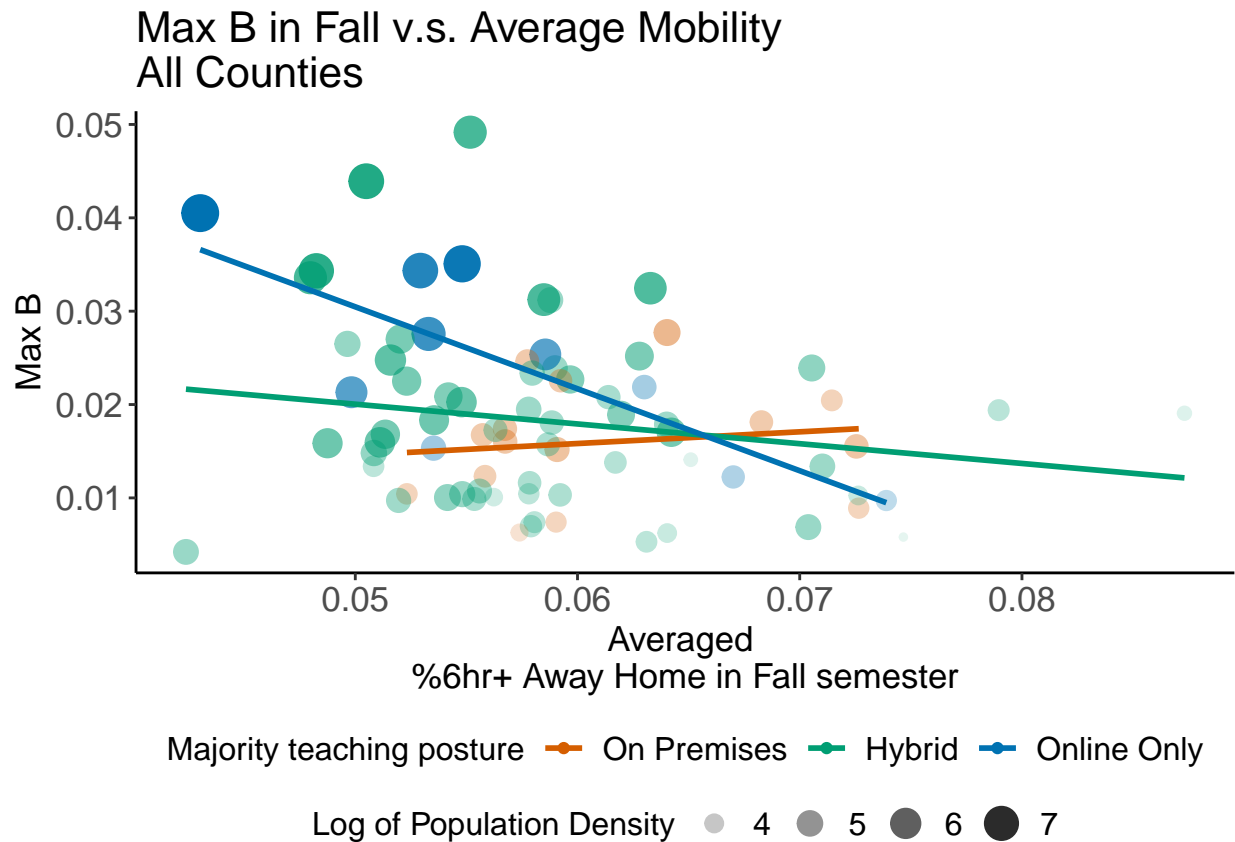

```
na.omit(B0B1)%>%
  drop_na(major_teaching)%>%
  ggplot(aes(x=log(Population.density),y=max_B1,group=major_teaching,color=major_teaching))+
  geom_point(aes(size=avg_full_work_prob,alpha=avg_full_work_prob))+
  geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+
  labs(y="Max B",x="Log of Population Density",
       title="Max B in Fall v.s. Log of Population Density \nAll Counties",
       color="Majority teaching posture",
       size = "Averaged \n%6hr+ Away Home in Fall semester",
       alpha= "Averaged \n%6hr+ Away Home in Fall semester" )+
  team_theme+theme(legend.position = "bottom")+
  scale_color_manual(values=col_theme)
```

Max B in Fall v.s. Log of Population Density
All Counties
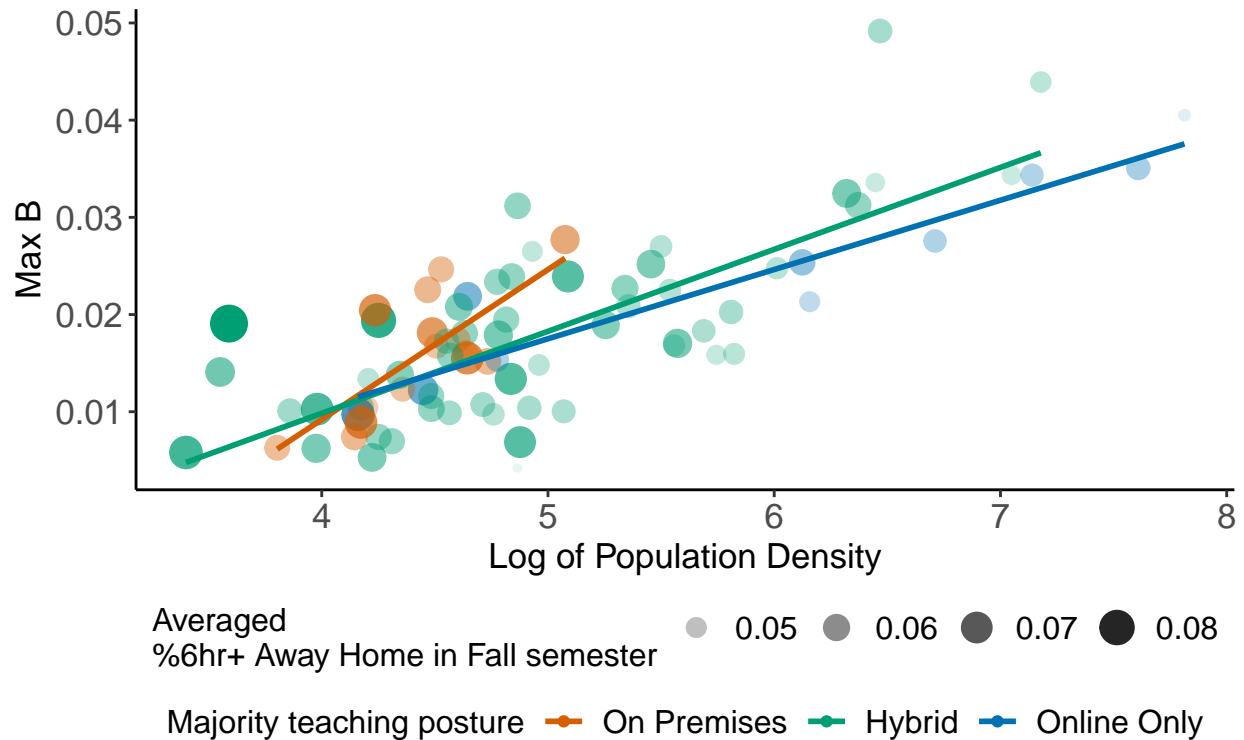
5.7 Change in growth vs. mobility and population density for all counties

```r
B_diff_micro <- B_diff%>%
  drop_na(major_teaching)%>%
  filter(NCHS.Urban.Rural.Status=="Micropolitan") %>%
  mutate(acc = new.slope.diff2 - new.slope.diff)
#At start of reopen 3w-0w
B_diff_micro%>%
  ggplot(aes(x=log(Population.density),
             y=new.slope.diff,group=major_teaching,
             color=major_teaching))+
  geom_point(aes(size=avg_full_work_prob,alpha=avg_full_work_prob))+
  geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+theme_minimal()+team_theme+
  labs(y="Change in Growth",x="Log of Population Density",
       title="Change in Growth Right After School Reopen\nOnly Micropolitan Counties",
       color="Majority teaching posture",
       size = "0w~3w\nAveraged %6hr+ Away Home",
       alpha= "0w~3w\nAveraged %6hr+ Away Home" ,fill="Majority teaching posture")+
  scale_color_manual(values=col_theme)+
  theme(legend.position = "bottom")+
  geom_text(data =B_diff_micro%>%
              filter(major_teaching=="On Premises"),
            aes(label=COUNTY),color='black',size=3,hjust=0.8, vjust=-0.2)
```
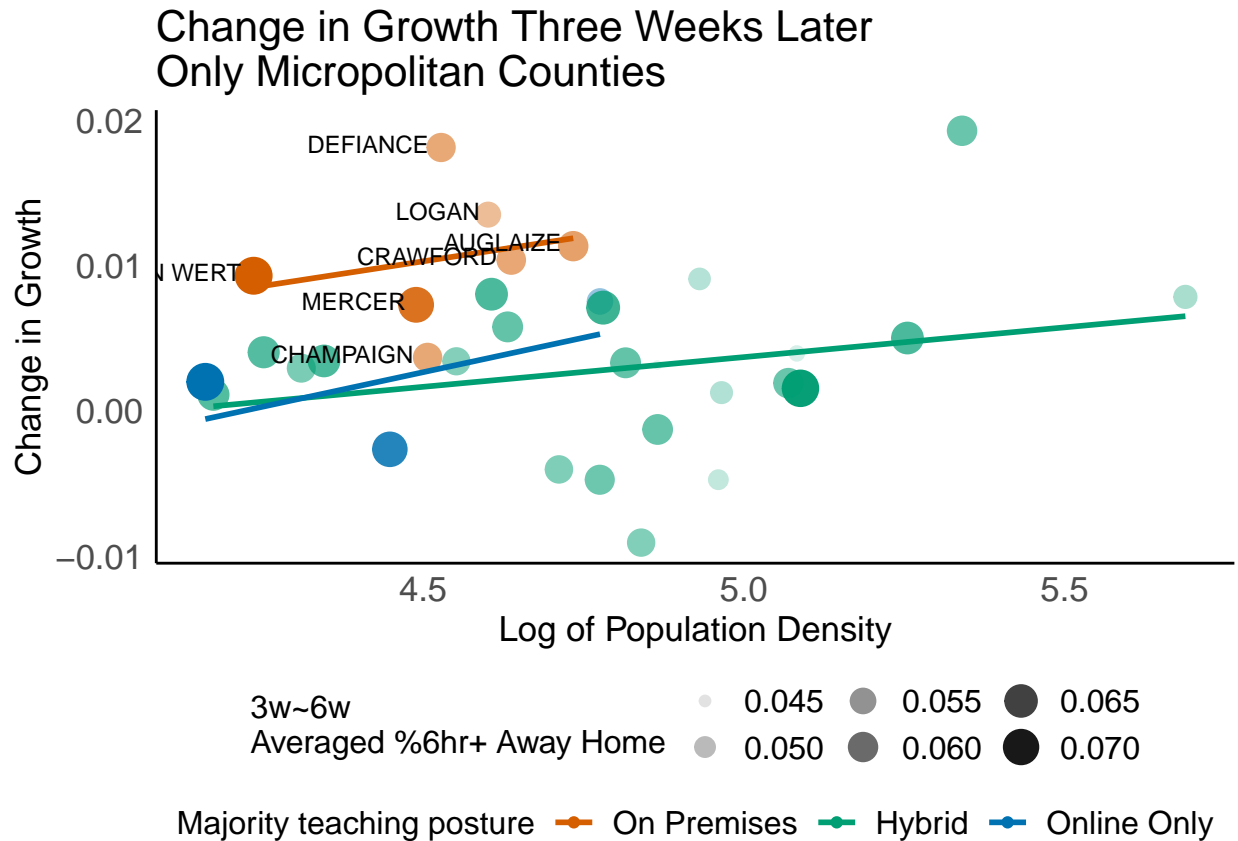
Change in Growth Right After School Reopen
Only Micropolitan Counties


```r
#After reopen for 6w-3w
B_diff_micro%>%
  ggplot(aes(x=log(Population.density),
             y=new.slope.diff2,
             group=major_teaching,
             color=major_teaching))+
  geom_point(aes(size=avg2_full_work_prob,alpha=avg2_full_work_prob))+
  geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+
  theme_minimal()+team_theme+
  labs(y="Change in Growth",x="Log of Population Density",
       title="Change in Growth Three Weeks Later\nOnly Micropolitan Counties",
       color="Majority teaching posture",
       size = "3w~6w\nAveraged %6hr+ Away Home",
       alpha= "3w~6w\nAveraged %6hr+ Away Home" ,fill="Majority teaching posture")+
  scale_color_manual(values=col_theme)+theme(legend.position = "bottom")+
  geom_text(data =B_diff_micro%>%
              filter(major_teaching=="On Premises"),
            aes(label=COUNTY),color='black',size=3,hjust=1.1, vjust=0.3)
```

Change in Growth Three Weeks Later
Only Micropolitan Counties

Appendix 6: Micropolitan Counties

Here we found that the rural-urban status is differentitating three groups of counties. Thus we block on the Micropolitan counties where the three groups are most comparable to reduce the confounding effect of urban-rural status.

6.1 Distribution of log population density by rural-Urban status and majority teaching posture

```r
# Pop density vs RURAL-Urban status
ohio_profile%>%
  left_join(wide_teaching_enroll[,c("county","major_teaching")],
            by = c("County" = "county"))%>%
  drop_na(major_teaching)%>%
  ggplot(aes(y=NCHS.Urban.Rural.Status,x=log(Population.density),
             fill=major_teaching))+
  facet_grid(~major_teaching)+
  geom_boxplot()+
  labs(fill="Majority teaching posture",
      size="Averaged \n%6hr+ Away Home",
      title="Distribution of Log Population Density \nby Rural-urban Status",
      x="Log Population Density", y= "Rural-urban Status")+
  team_theme+
  scale_fill_manual(values=col_theme)+
  theme(legend.position = "bottom")
```

Distribution of Log Population Density by Rural–urban Status
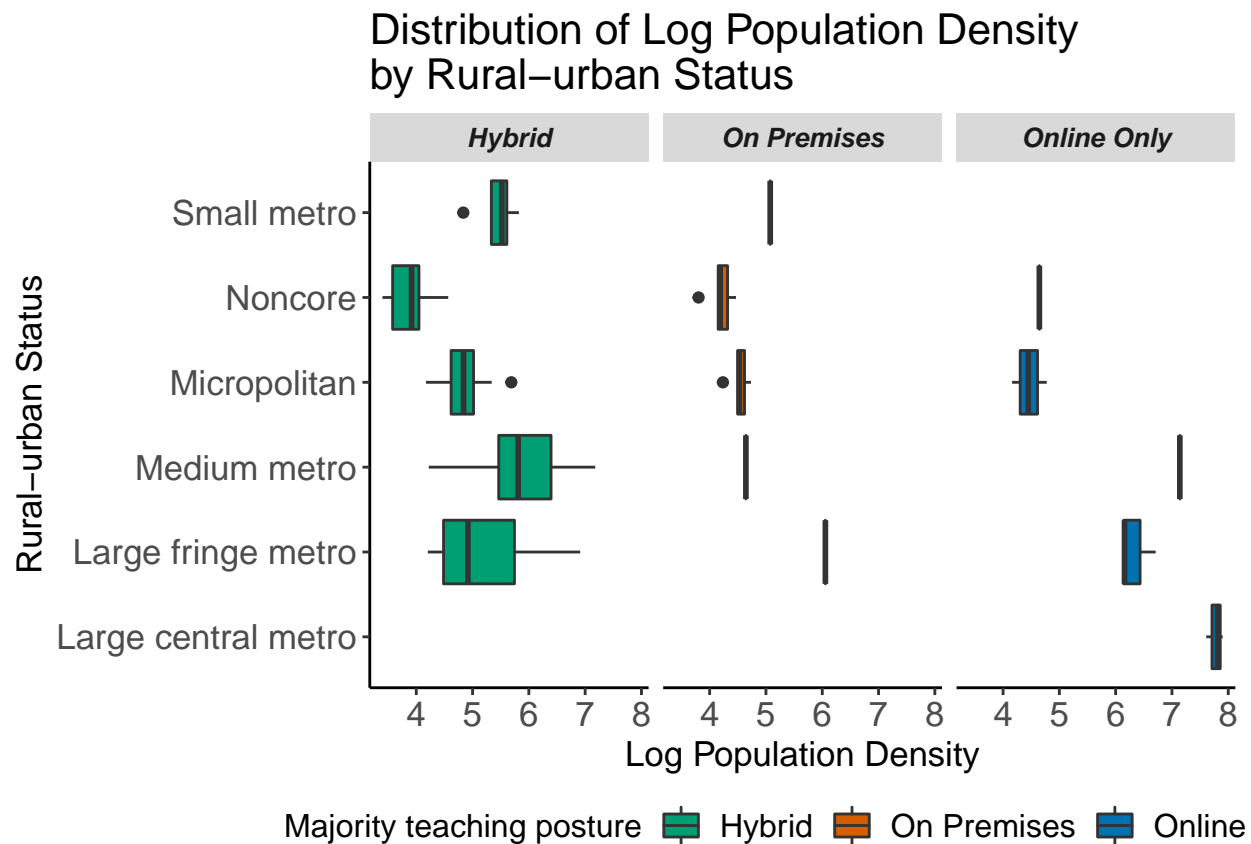
## 6.2 Distribution of log population density majority teaching postue in Micropolitan counties

Here we can see that in the micropolitan counties, online-only and on-premises ones have similar population density.

```
ohio_profile%>%
  filter(NCHS.Urban.Rural.Status=="Micropolitan")%>%
 left_join(BOB1%>%dplyr::select(COUNTY,major_teaching),by=c("County"="COUNTY"))%>%
  ggplot(aes(x=major_teaching))+
  geom_dotplot(aes(y=log(Population.density),fill=major_teaching),
               binaxis='y', stackdir='center')+
  team_theme+guides(fill=FALSE)+
  labs(y="Log Population Density",x="Majority teaching posture",
       title="Log of population density vs teaching posture\nfor Micropolitan Counties")+
  scale_fill_manual(values=col_theme)
```

Log of population density vs teaching posture for Micropolitan Counties
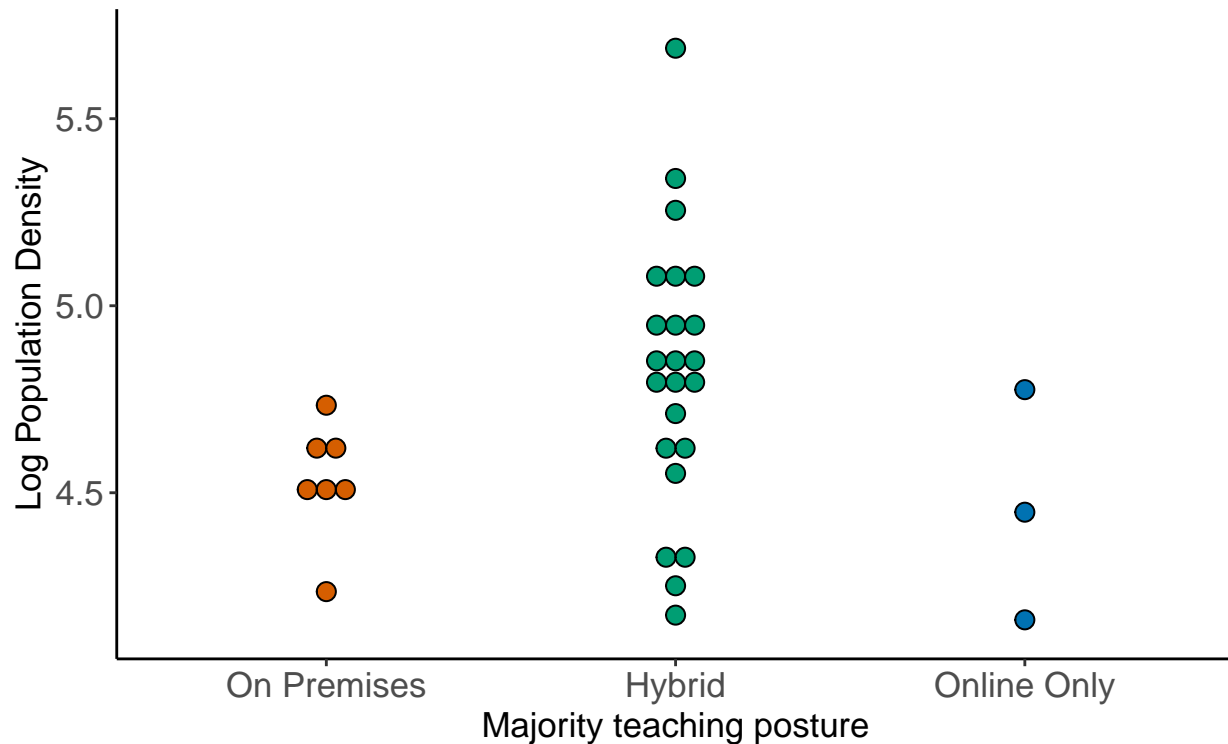
6.3 Geographical distribution of counties' Micropolitan status

We can see that micropolitan counties are well-spread across the state.

```r
# Map of Micropolitan status
ohio_profile%>%
  distinct(County,NCHS.Urban.Rural.Status) %>%
  mutate(is_micro = factor(ifelse(NCHS.Urban.Rural.Status == "Micropolitan",
                                  "Micropolitan","Non-Micropolitan")))%>%
  left_join(ohio_map,by=c('County'='county'))%>%
  ggplot() +
  geom_polygon(aes(x = long, y = lat, group = group, fill = is_micro),
               color = "gray") +
  coord_fixed(1.3) + theme_map() +
  scale_fill_manual(values = c("Non-Micropolitan" ="#0072B2",
                    "Micropolitan" ="#009E73"))+
  labs(fill='Micropolitan Counties')+
  theme(legend.text = element_text(size=12),
        legend.title = element_text(size=12),
        legend.position = "bottom",
        legend.key.size = unit(2,"lines"))
```
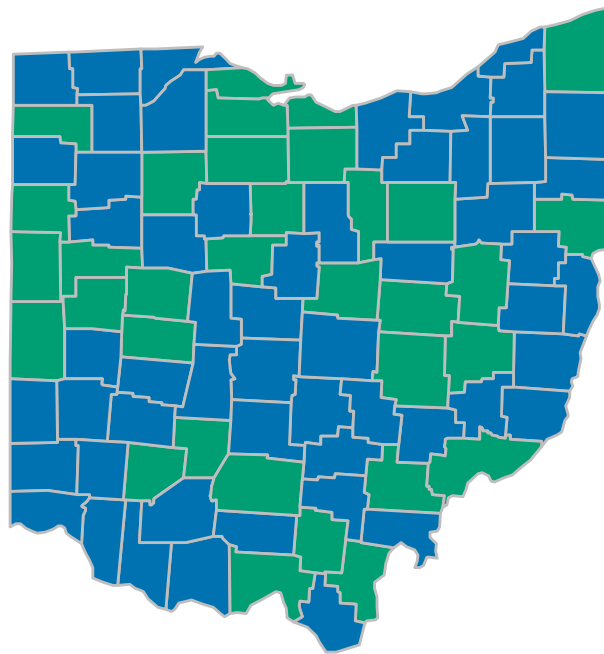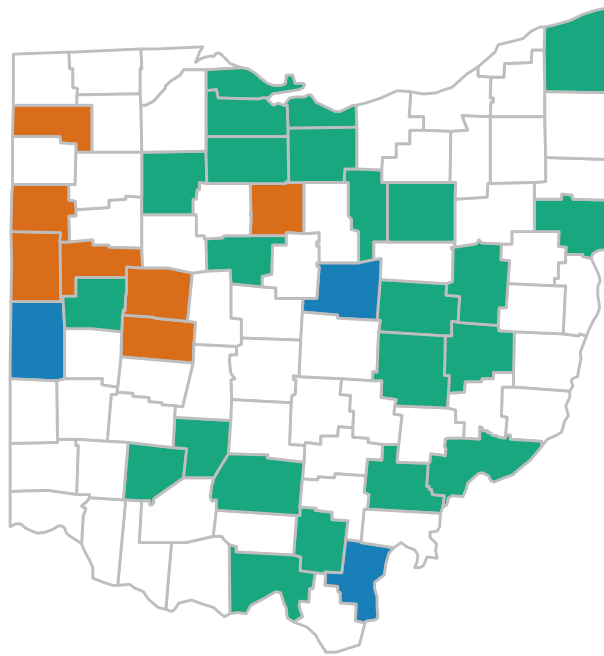
Micropolitan Counties 🟩 **Micropolitan** 🟦 **Non−Micropolitan**

```r
# Map of majority teaching posture in Micropolitan counties
ohio_profile%>%
  distinct(County,NCHS.Urban.Rural.Status) %>%
  left_join(wide_teaching_enroll[,c("county","major_teaching")],
            by = c("County" = "county"))%>%
  mutate(is_micro = factor(ifelse(NCHS.Urban.Rural.Status == "Micropolitan",1,0)),
         micro_teach = factor(ifelse(is_micro == 1, major_teaching, "Not Micro")))%>%
  left_join(ohio_map,by=c('County'='county'))%>%
  ggplot() +
  geom_polygon(aes(x = long, y = lat, group = group, fill = micro_teach),
               color = "gray",alpha=0.9) +
  coord_fixed(1.3) + theme_map() +
  scale_fill_manual(values = c(col_theme, "Not Micro" = "white"))+
  labs(fill='Micropolitan \nCounties')+
  theme(legend.text = element_text(size=12),
        legend.title = element_text(size=12),
        legend.position = "bottom",
        legend.key.size = unit(2,"lines"))
```

| Micropolitan Counties | Hybrid | Not Micro | On Premises | Online Only |

6.4 Distribution of maximum growth ccoefficient in Micropolitan counties

On Premises counties have significant higher maximum growth coefficient than online only counties.

```r
# one-way ANOVA
maxB_major_teaching.aov <- aov(max_B1 ~ major_teaching,data = na.omit(BOB1)%>%
  filter(NCHS.Urban.Rural.Status=="Micropolitan"))
summary(maxB_major_teaching.aov)
```

```
##                Df    Sum Sq   Mean Sq F value Pr(>F)
## major_teaching  2 0.0000764 3.821e-05   1.002   0.38
## Residuals      28 0.0010679 3.814e-05
```

```r
# Duncan test
stat.test <- PostHocTest(maxB_major_teaching.aov, method = "duncan")$major_teaching %>%
  as.data.frame()%>%
  rownames_to_column("group") %>%
  separate(group,"-", into = c("group1","group2")) %>%
  mutate(pval = round(pval,3),
         p = case_when(pval <= .01~ "**",
                       pval <= .05 ~ "*",
                       TRUE ~ "NS"))%>%
  dplyr::select(group1, group2, pval, p)


# boxplot
na.omit(BOB1)%>%
  filter(NCHS.Urban.Rural.Status=="Micropolitan")%>%
```
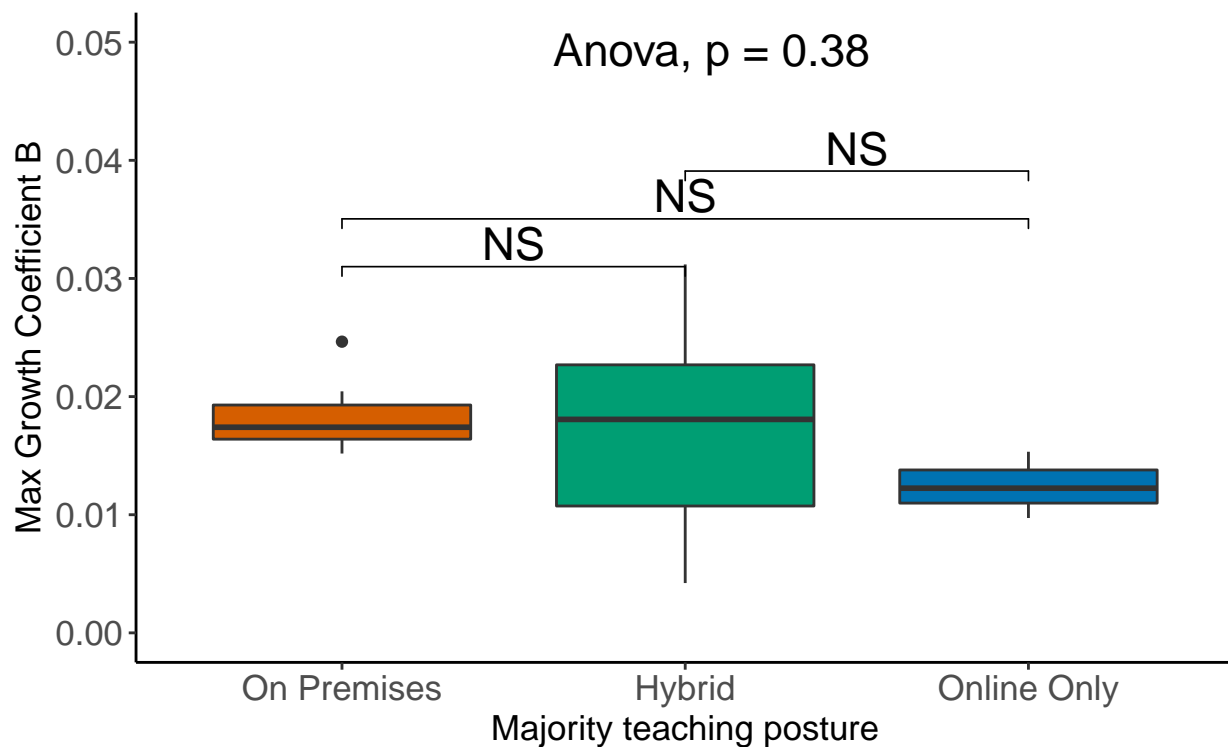
```
drop_na(major_teaching)%>%
ggplot(aes(x=major_teaching,y=max_B1))+
geom_boxplot(aes(fill=major_teaching))+
ylim(c(0,0.05))+
stat_compare_means(method = "anova",size=6,label.y.npc=0.96,label.x.npc = 0.4)+
stat_pvalue_manual(stat.test, label = "p",y.position = 0.03,
                   step.increase = 0.15,size = 6,bracket.nudge.y = 0.001)+
team_theme+
theme(legend.position = " ")+
labs(y="Max Growth Coefficient B",
     x="Majority teaching posture",
     title="Distribution of Maximum Growth Coefficient \nin Micropolitan Counties",
     fill="Majority teaching posture")+
scale_fill_manual(values=col_theme)
```


6.5 Maximum B vs. mobility and population density in Micropolitan counties

We can see the relationship between severity of the disease and mobility or population density become unclear after we block the data on its Micropolitan status and the sample size is limited.
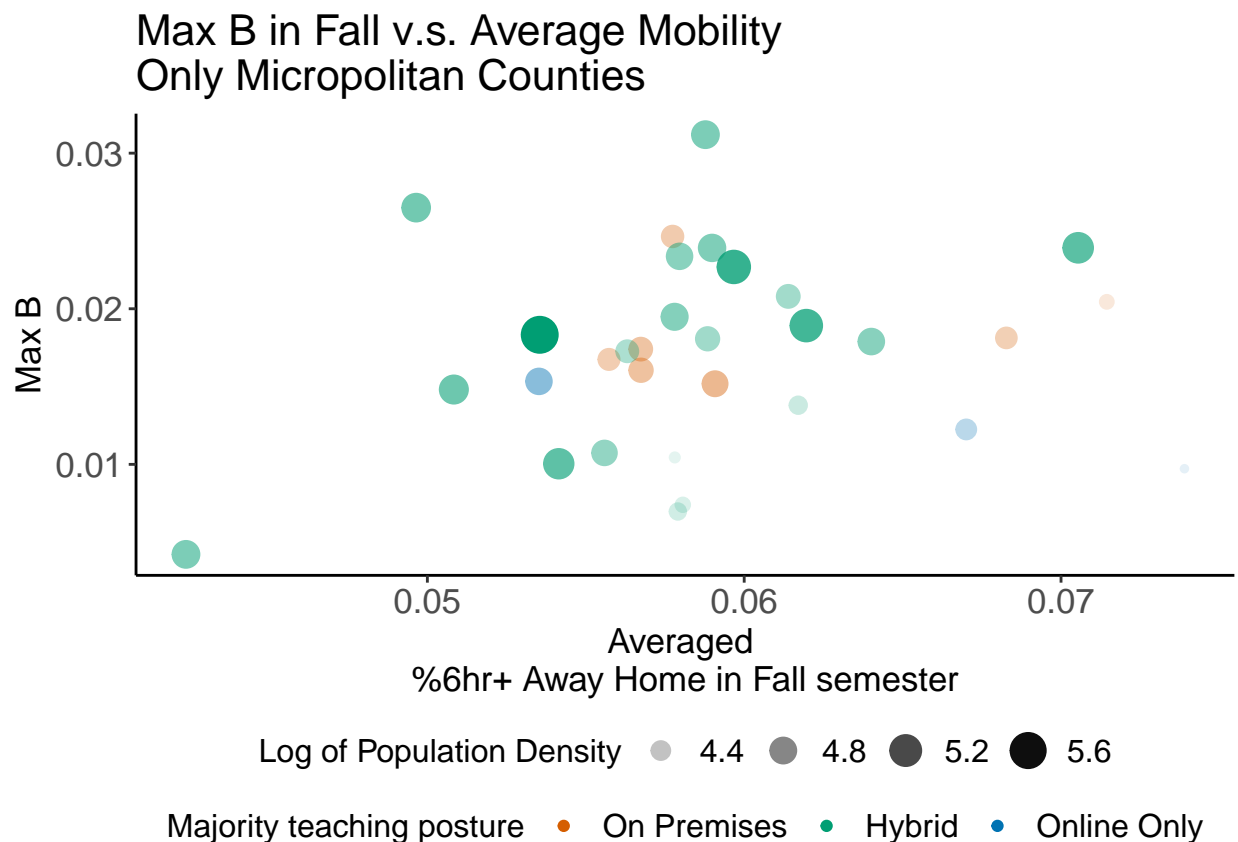
```
na.omit(B0B1)%>%
  filter(NCHS.Urban.Rural.Status=="Micropolitan")%>%
  drop_na(major_teaching)%>%
  ggplot(aes(x=avg_full_work_prob,y=max_B1,group=major_teaching,color=major_teaching))+
  geom_point(aes(size=log(Population.density),alpha=log(Population.density)))+
  #geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+
```

```
labs(y="Max B",x="Averaged \n%6hr+ Away Home in Fall semester",
    title="Max B in Fall v.s. Average Mobility \nOnly Micropolitan Counties",
    color="Majority teaching posture",
    size = "Log of Population Density",
    alpha= "Log of Population Density" )+
team_theme+theme(legend.position = "bottom")+
scale_color_manual(values=col_theme)+
guides(
  size = guide_legend(order = 1),
  alpha = guide_legend(order = 1),
  fill = guide_legend(order = 0)
)
```


Max B in Fall v.s. Average Mobility
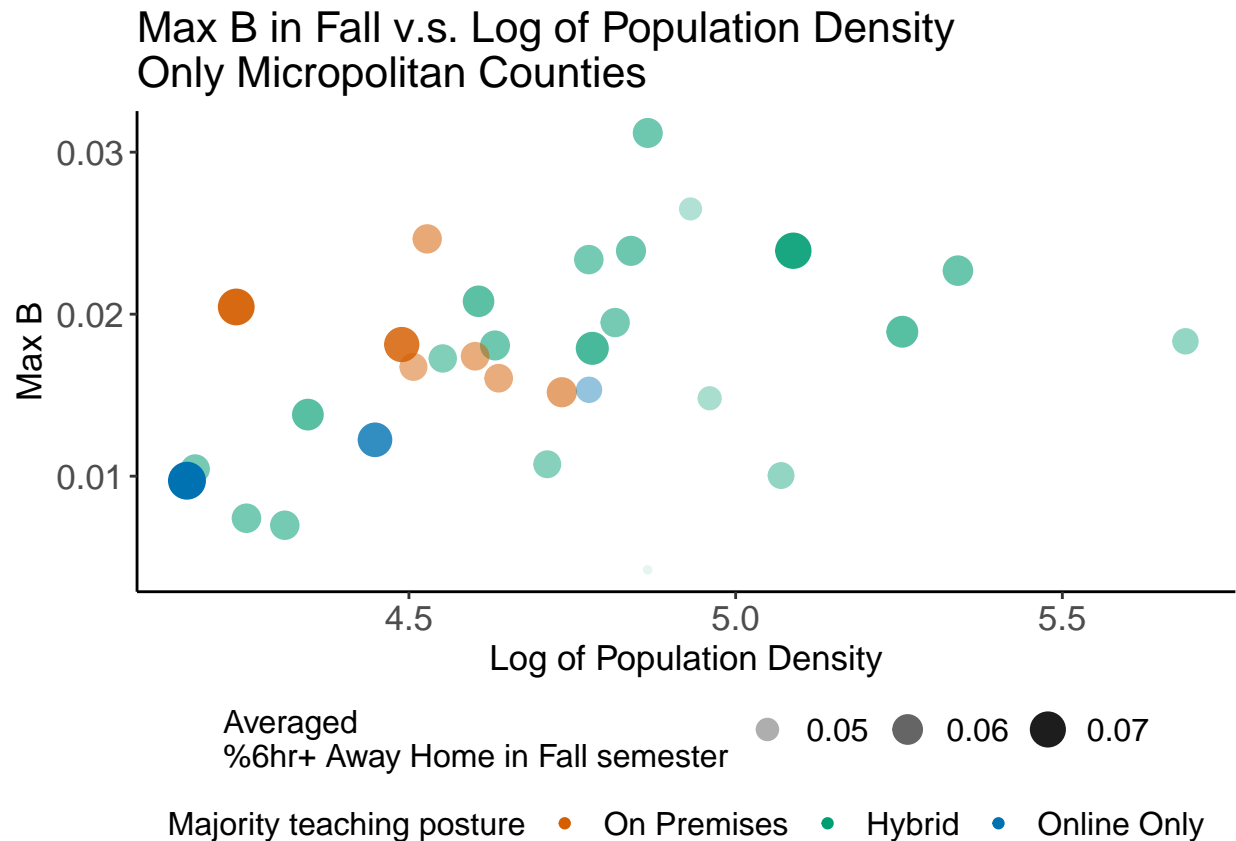Only Micropolitan Counties

```
na.omit(BOB1)%>%
  filter(NCHS.Urban.Rural.Status=="Micropolitan")%>%
  drop_na(major_teaching)%>%
  ggplot(aes(x=log(Population.density),
            y=max_B1,group=major_teaching,
            color=major_teaching))+
  geom_point(aes(size=avg_full_work_prob,alpha=avg_full_work_prob))+
  #geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+
  labs(y="Max B",x="Log of Population Density",
      title="Max B in Fall v.s. Log of Population Density \nOnly Micropolitan Counties",
      color="Majority teaching posture",
      size = "Averaged \n%6hr+ Away Home in Fall semester",
      alpha= "Averaged \n%6hr+ Away Home in Fall semester" )+
```

```
team_theme+theme(legend.position = "bottom")+
scale_color_manual(values=col_theme)
```

Max B in Fall v.s. Log of Population Density
Only Micropolitan Counties

6.6 Change in Growth vMaximum growth coefficient vs. mobility and population density for Micropolitan counties

```
## filter only micro
B_diff_micro <- B_diff%>%
  drop_na(major_teaching)%>%
  filter(NCHS.Urban.Rural.Status=="Micropolitan") %>%
  mutate(acc = new.slope.diff2 - new.slope.diff)
```
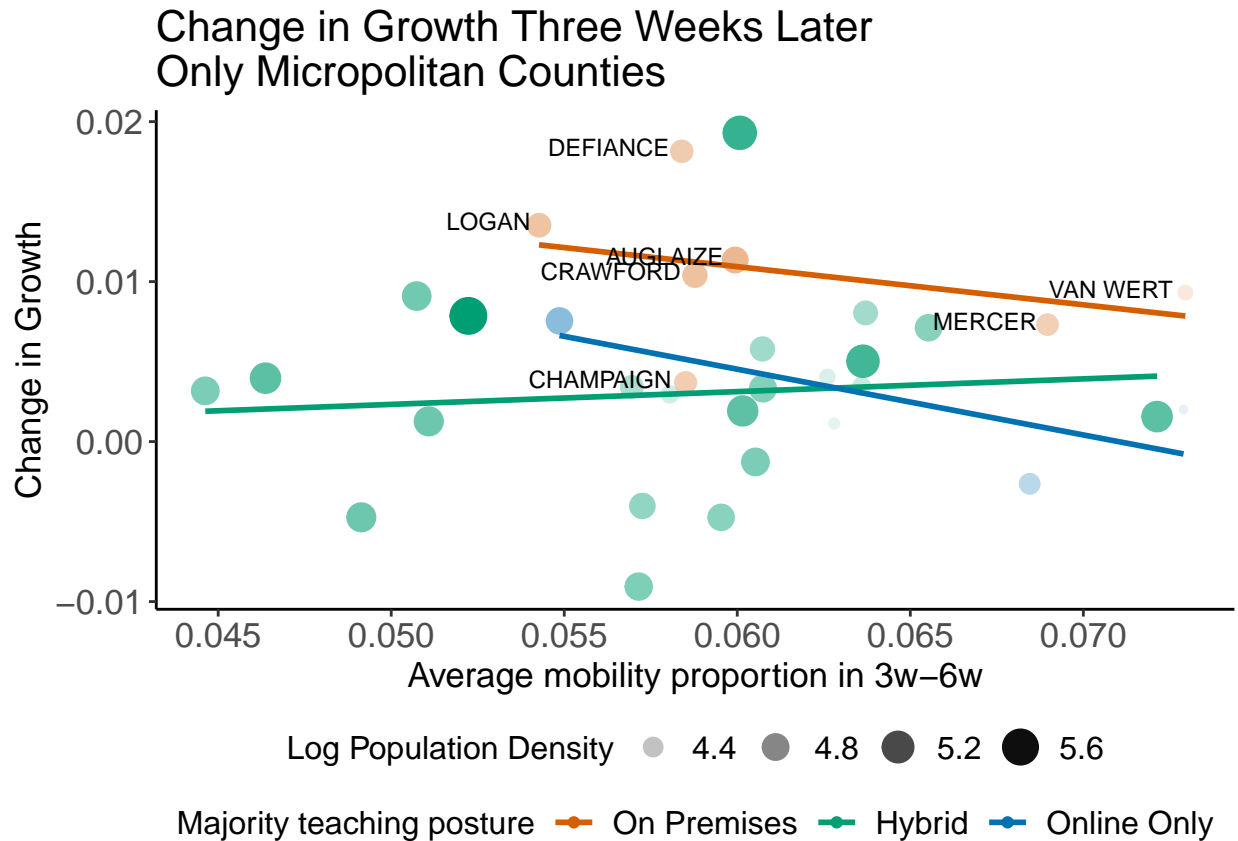
Mobility

```
B_diff_micro%>%
  ggplot(aes(x=avg_full_work_prob,
             y=new.slope.diff,
             group=major_teaching,color=major_teaching))+
  geom_point(aes(size=log(Population.density),alpha=log(Population.density)))+
  geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+
  labs(y="Change in Growth",x="Average mobility proportion in 0w-3w",
       title="Change in Growth Right After School Reopen\nOnly Micropolitan Counties",
       color="Majority teaching posture",
       size = "Log Population Density",
       alpha= "Log Population Density")+
  scale_color_manual(values=col_theme)+theme(legend.position = "bottom")+team_theme+
```

```
geom_text(data =B_diff_micro%>%
            filter(major_teaching=="On Premises"),
          aes(label=COUNTY),color='black',
          size=3,hjust=1.1, vjust=0.3,size=6)
```


Change in Growth Right After School Reopen
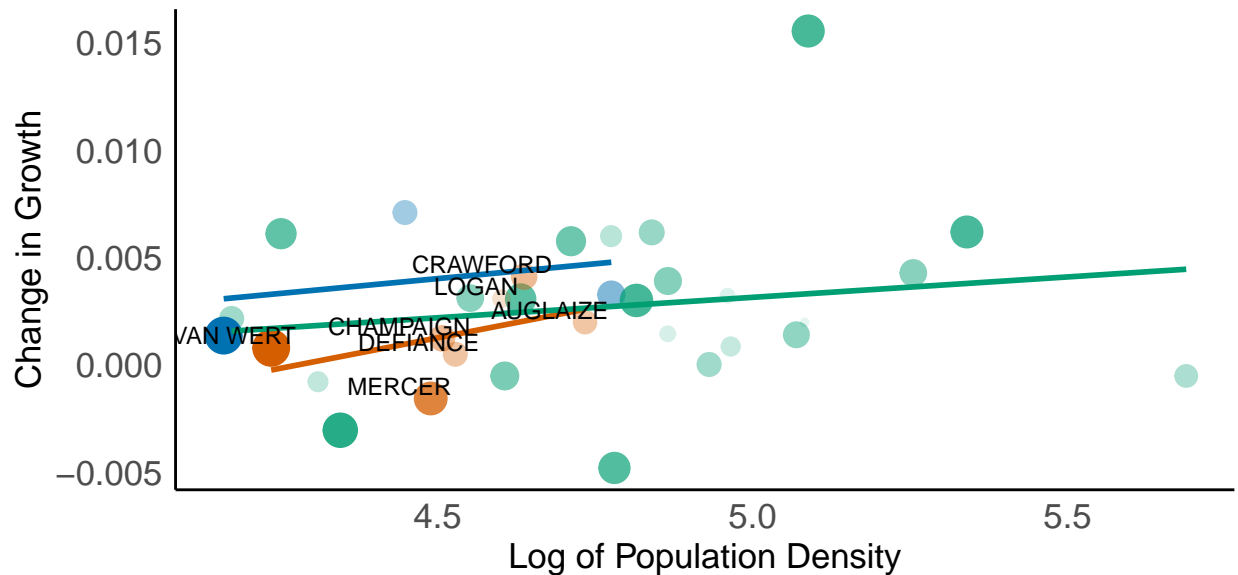Only Micropolitan Counties
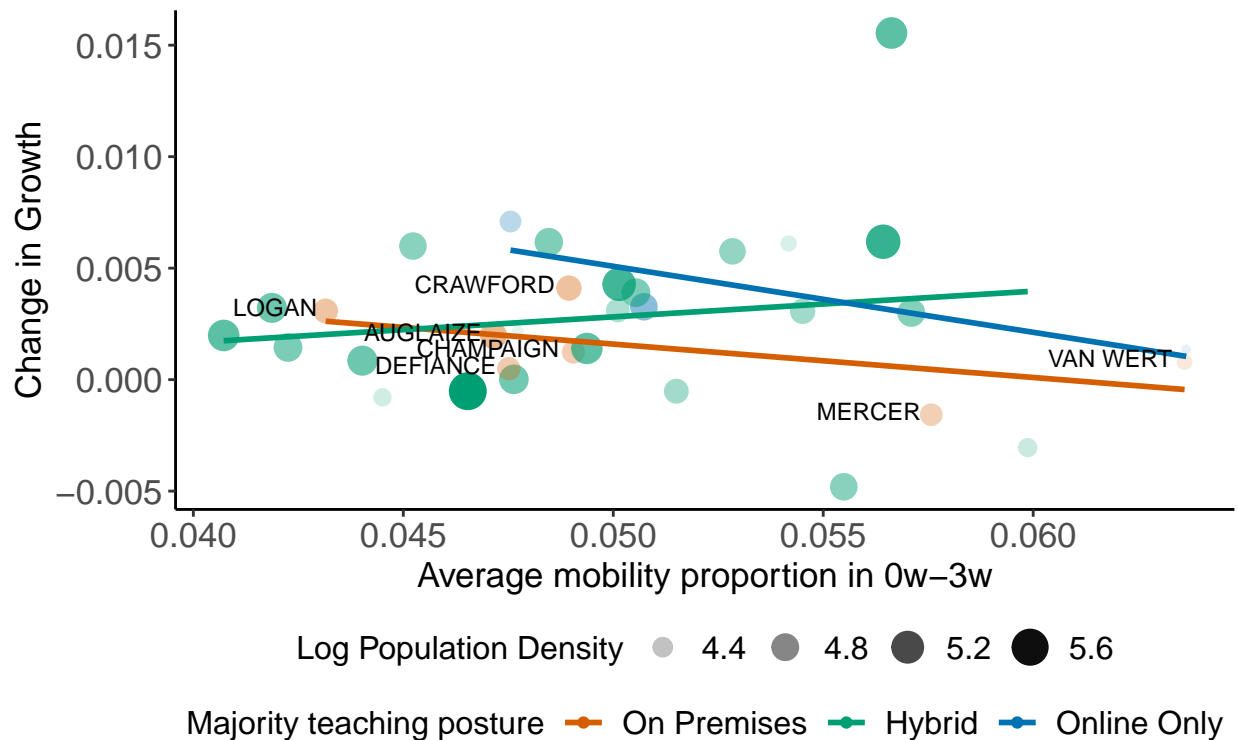
```
B_diff_micro%>%
  ggplot(aes(x=avg2_full_work_prob,
             y=new.slope.diff2,
             group=major_teaching,
             color=major_teaching))+
  geom_point(aes(size=log(Population.density),alpha=log(Population.density)))+
  geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+
  labs(y="Change in Growth",x="Average mobility proportion in 3w-6w",
       title="Change in Growth Three Weeks Later\nOnly Micropolitan Counties",
       color="Majority teaching posture",
       size ="Log Population Density",
       alpha= "Log Population Density")+
  scale_color_manual(values=col_theme)+
  theme(legend.position = "bottom")+
  team_theme+geom_text(data =B_diff_micro%>%
                         filter(major_teaching=="On Premises"),
                       aes(label=COUNTY),color='black',
                       size=3,hjust=1.1, vjust=0.3,size=6)
```

Change in Growth Three Weeks Later
Only Micropolitan Counties

Log Population Density

```
B_diff_micro%>%
  ggplot(aes(x=log(Population.density),
             y=new.slope.diff,
             group=major_teaching,color=major_teaching))+
  geom_point(aes(size=avg_full_work_prob,alpha=avg_full_work_prob))+
  geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+
  theme_minimal()+team_theme+
  labs(y="Change in Growth",x="Log of Population Density",
       title="Change in Growth Right After School Reopen\nOnly Micropolitan Counties",
       color="Majority teaching posture",
       size = "0w~3w\nAveraged %6hr+ Away Home",
       alpha= "0w~3w\nAveraged %6hr+ Away Home" ,fill="Majority teaching posture")+
  scale_color_manual(values=col_theme)+
  theme(legend.position = "bottom")+
  geom_text(data =B_diff_micro%>%
              filter(major_teaching=="On Premises"),
            aes(label=COUNTY),color='black',
            size=3,hjust=0.8, vjust=-0.2)
```
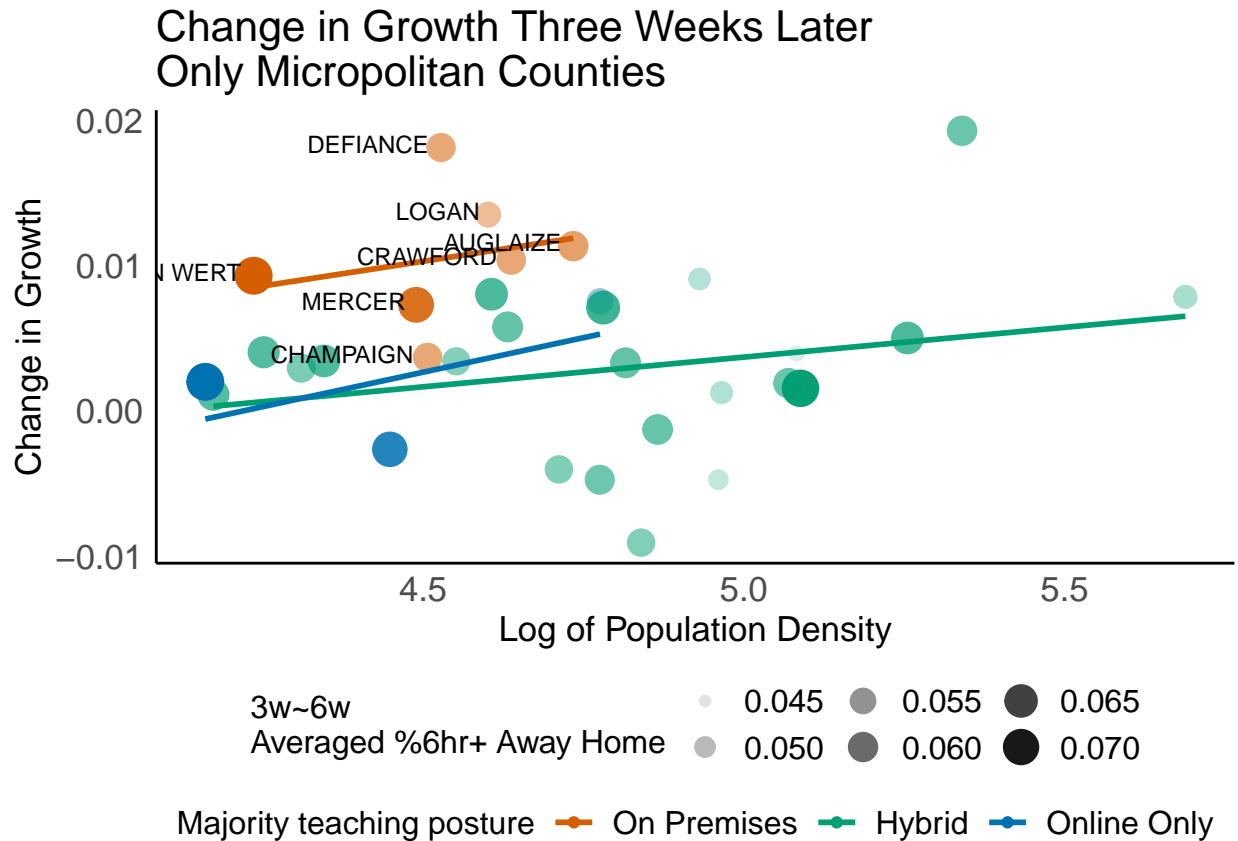
Change in Growth Right After School Reopen
Only Micropolitan Counties


```
B_diff_micro%>%
  ggplot(aes(x=log(Population.density),
             y=new.slope.diff2,
             group=major_teaching,color=major_teaching))+
  geom_point(aes(size=avg2_full_work_prob,
                 alpha=avg2_full_work_prob))+
  geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+theme_minimal()+team_theme+
  labs(y="Change in Growth",x="Log of Population Density",
       title="Change in Growth Three Weeks Later\nOnly Micropolitan Counties",
       color="Majority teaching posture",
       size = "3w~6w\nAveraged %6hr+ Away Home",
       alpha= "3w~6w\nAveraged %6hr+ Away Home" ,
       fill="Majority teaching posture")+
  scale_color_manual(values=col_theme)+theme(legend.position = "bottom")+
  geom_text(data =B_diff_micro%>%
              filter(major_teaching=="On Premises"),
            aes(label=COUNTY),color='black',size=3,hjust=1.1, vjust=0.3)
```

Change in Growth Three Weeks Later
Only Micropolitan Counties

Appendix 7: Sensitive Analysis

In order to make sure that the shifting of the red line (On Premises counties) is not brought by chance, we will conduct a sensitive analysis to detect how the change in growth varies throughout time. The whole sensitive analysis is based on 'Change in growth versus Log Population Density.

Since we assume that the school posture takes three weeks to reflect on the growth coefficient, the growth coefficients before 3 weeks after the start of school are all regarded as not taking effect. So, we use B(3) as a turning point. The changes in growth we want to test are as below:
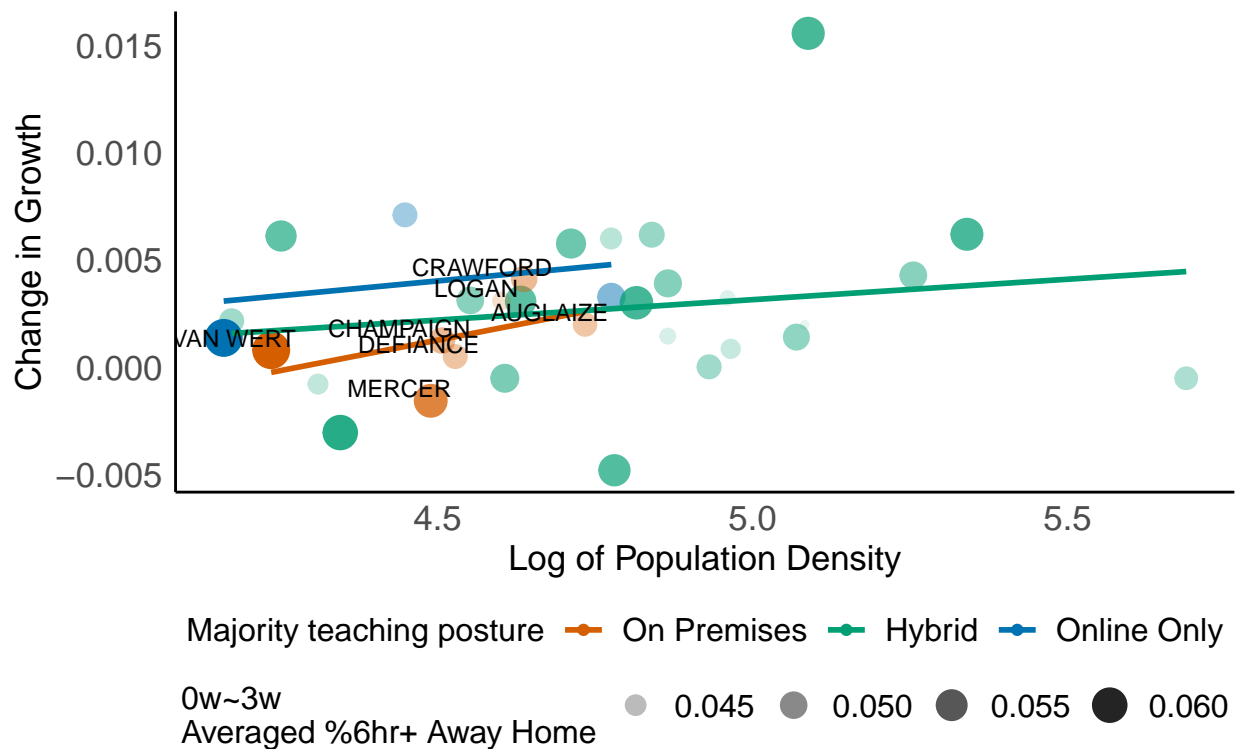
Before school posture taking effect: B(0)-B(-3), B(1)-B(-2), B(2)-B(-1), B(3)-B(0) (also known as change in growth right after the start of school reopen).

After school posture taking effect: B(4)-B(1), B(5)-B(2), B(6)-B(3) (also known as change in growth three weeks later), B(7)-B(4)

```
B_diff_micro <- B_diff%>%
  drop_na(major_teaching)%>%
  filter(NCHS.Urban.Rural.Status=="Micropolitan") %>%
  mutate(acc = new.slope.diff2 - new.slope.diff)
#Before reopen
##3w-0w
B_diff_micro%>%
  ggplot(aes(x=log(Population.density),
             y=new.slope.diff,
             group=major_teaching,
             color=major_teaching))+
```

```
geom_point(aes(size=avg_full_work_prob,alpha=avg_full_work_prob))+
geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+
theme_minimal()+team_theme+
labs(y="Change in Growth",x="Log of Population Density",
     title="Change in Growth Right After School Reopen\nOnly Micropolitan Counties",
     color="Majority teaching posture",
     size = "0w~3w\nAveraged %6hr+ Away Home",
     alpha= "0w~3w\nAveraged %6hr+ Away Home" ,fill="Majority teaching posture")+
scale_color_manual(values=col_theme)+theme(legend.position = "bottom")+
geom_text(data =B_diff_micro%>%filter(major_teaching=="On Premises"),
          aes(label=COUNTY),color='black',size=3,hjust=0.8, vjust=-0.2)
```


Change in Growth Right After School Reopen
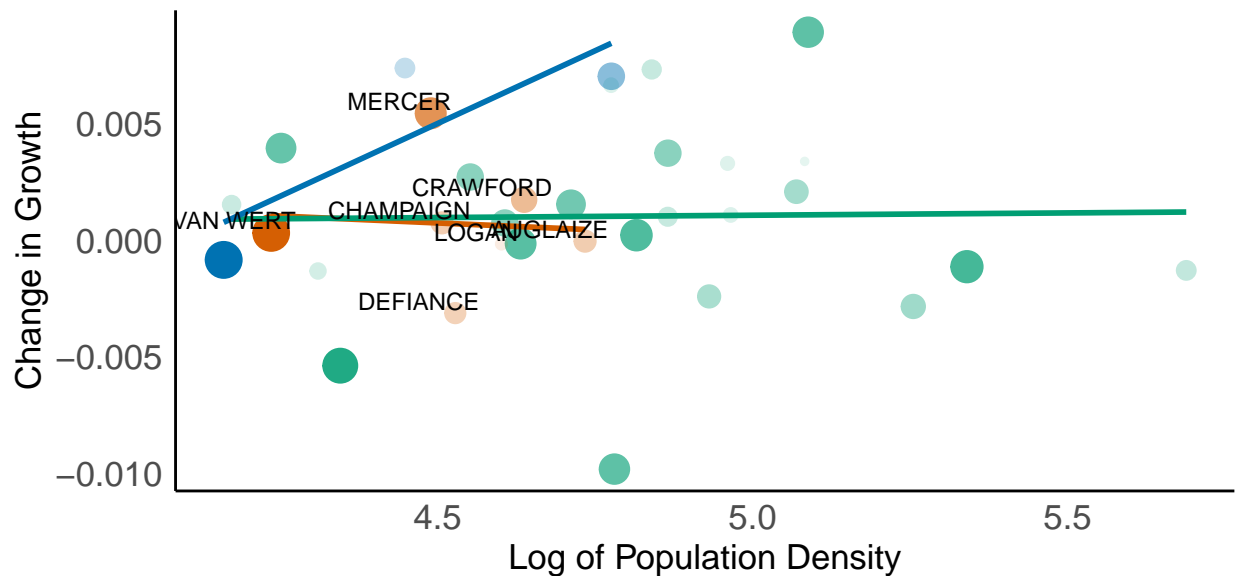Only Micropolitan Counties

```
##2w-(-1w)
B_diff_micro%>%
  ggplot(aes(x=log(Population.density),
             y=new.slope.diff2m1,group=major_teaching,
             color=major_teaching))+
  geom_point(aes(size=avg_full_work_prob_m1w2w,alpha=avg_full_work_prob_m1w2w))+
  geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+theme_minimal()+team_theme+
  labs(y="Change in Growth",x="Log of Population Density",
       title="Change in Growth between of B(-1) and B(2)\nOnly Micropolitan Counties",
       color="Majority teaching posture",
       size = "-1w~2w\nAveraged %6hr+ Away Home",
       alpha= "-1w~2w\nAveraged %6hr+ Away Home" ,
       fill="Majority teaching posture")+
  scale_color_manual(values=col_theme)+
```

```
theme(legend.position = "bottom")+
geom_text(data =B_diff_micro%>%
          filter(major_teaching=="On Premises"),
       aes(label=COUNTY),color='black',size=3,hjust=0.8, vjust=-0.2)
```

Change in Growth between of B(−1) and B(2) Only Micropolitan Counties


```
##1w-(-2w)
B_diff_micro%>%
  ggplot(aes(x=log(Population.density),
             y=new.slope.diff1m2,group=major_teaching,color=major_teaching))+
  geom_point(aes(size=avg_full_work_prob_m2w1w,alpha=avg_full_work_prob_m2w1w))+
  geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+
  theme_minimal()+team_theme+
  labs(y="Change in Growth",x="Log of Population Density",
      title="Change in Growth between of B(-2) and B(1)\nOnly Micropolitan Counties",
      color="Majority teaching posture",
      size = "-2w~1w\nAveraged %6hr+ Away Home",
      alpha= "-2w~1w\nAveraged %6hr+ Away Home" ,fill="Majority teaching posture")+
  scale_color_manual(values=col_theme)+theme(legend.position = "bottom")+
  geom_text(data =B_diff_micro%>%
             filter(major_teaching=="On Premises"),
          aes(label=COUNTY),color='black',size=3,hjust=0.8, vjust=-0.2)
```

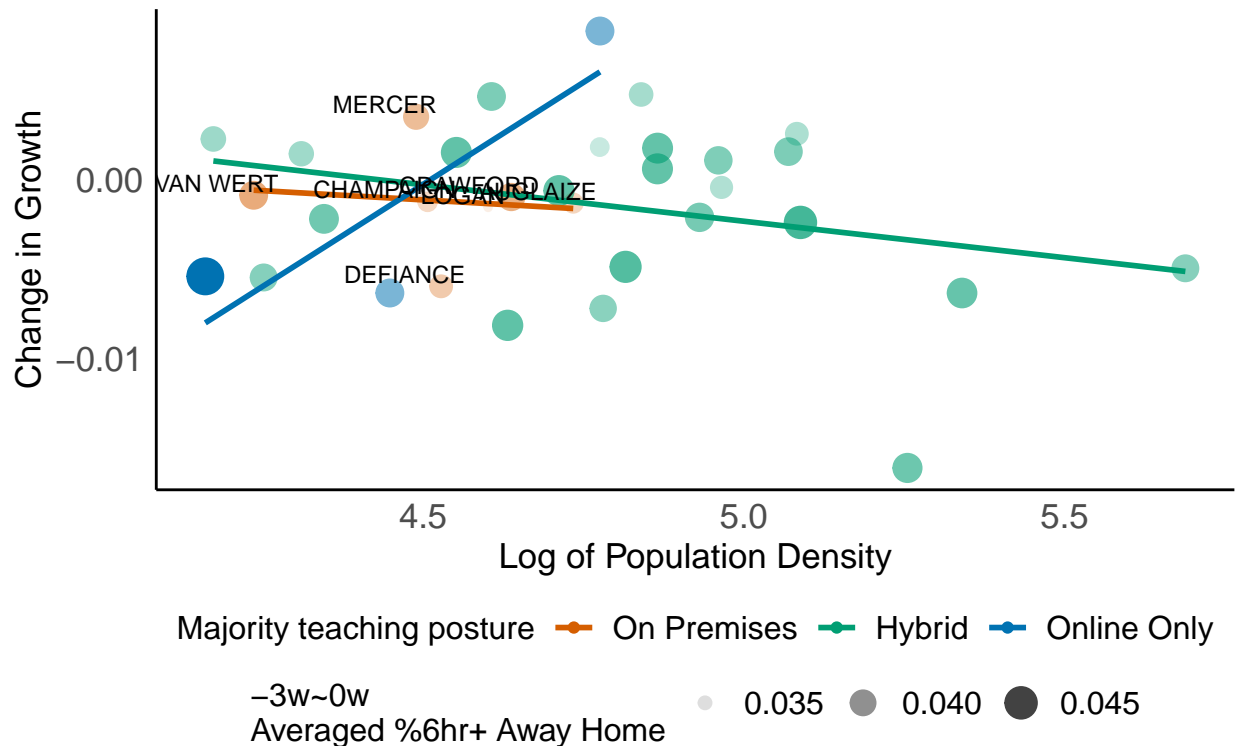Change in Growth between of B(−2) and B(1)
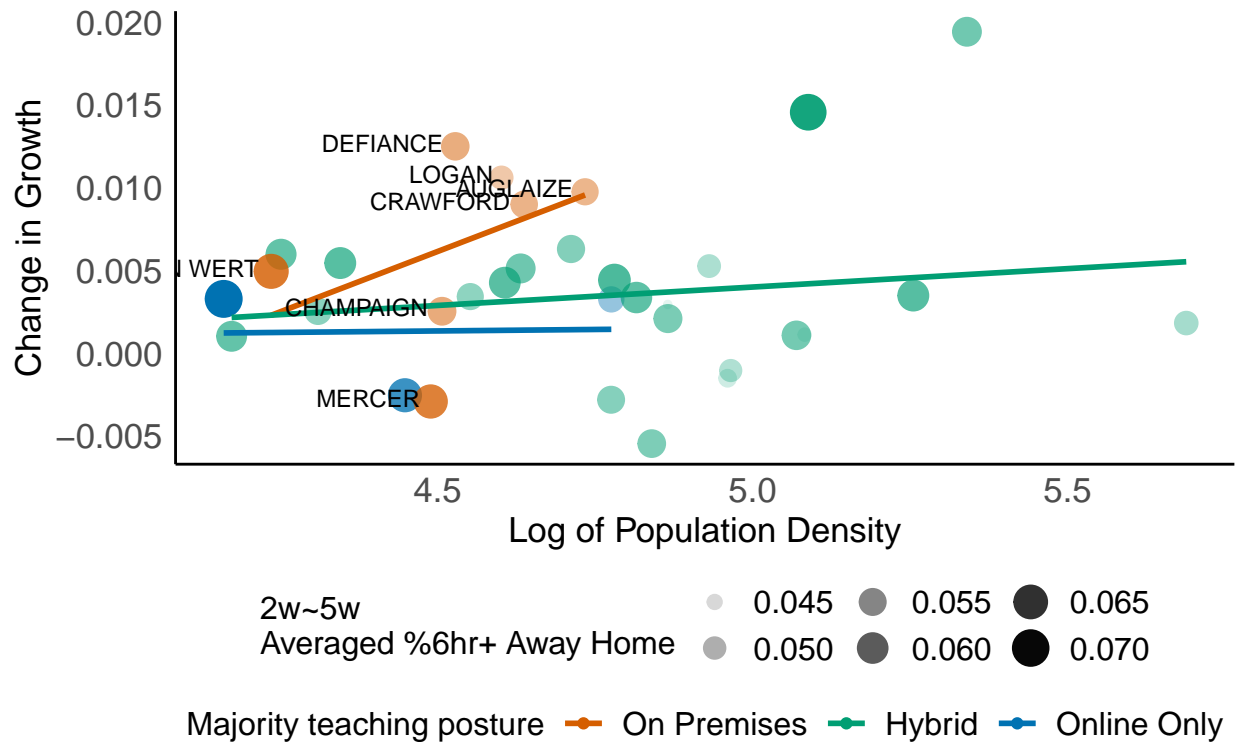Only Micropolitan Counties
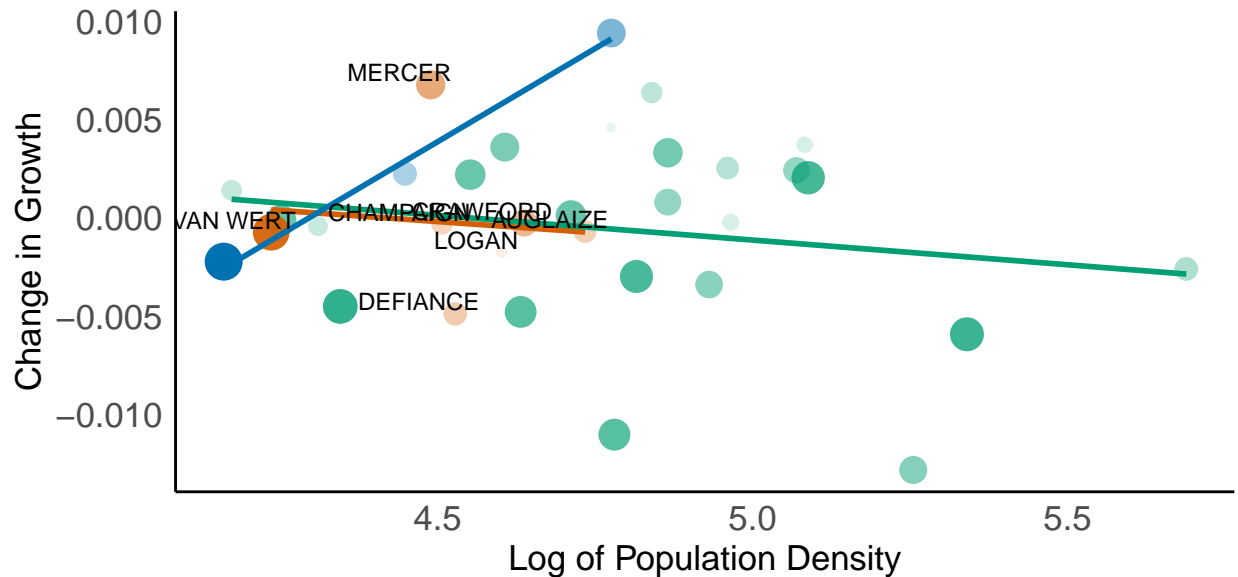
```
##0w-(-3w)
B_diff_micro%>%
  ggplot(aes(x=log(Population.density),
             y=new.slope.diff0m3,group=major_teaching,color=major_teaching))+
  geom_point(aes(size=avg_full_work_prob_m3w0w,alpha=avg_full_work_prob_m3w0w))+
  geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+theme_minimal()+team_theme+
  labs(y="Change in Growth",x="Log of Population Density",
       title="Change in Growth between of B(-3) and B(0)\nOnly Micropolitan Counties",
       color="Majority teaching posture",
       size = "-3w~0w\nAveraged %6hr+ Away Home",
       alpha= "-3w~0w\nAveraged %6hr+ Away Home" ,fill="Majority teaching posture")+
  scale_color_manual(values=col_theme)+theme(legend.position = "bottom")+
  geom_text(data =B_diff_micro%>%filter(major_teaching=="On Premises"),
            aes(label=COUNTY),color='black',
            size=3,hjust=0.8, vjust=-0.2)
```

Change in Growth between of B(−3) and B(0)
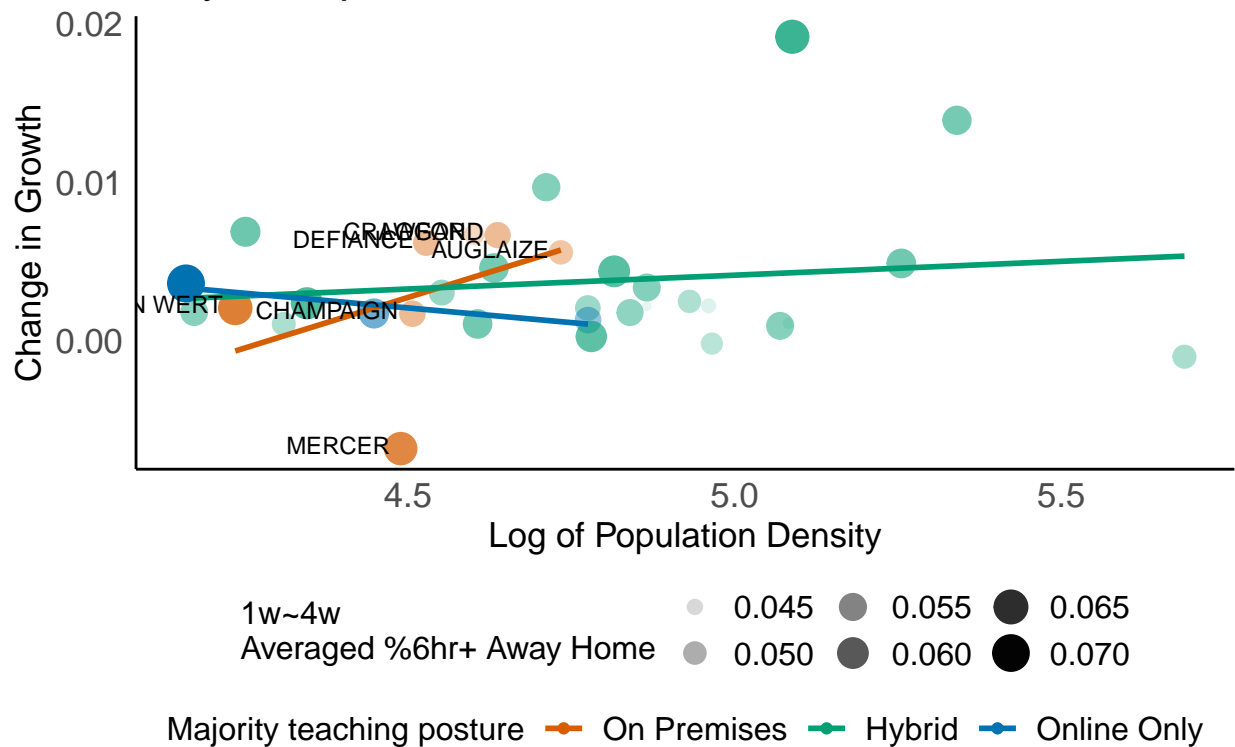## Only Micropolitan Counties


```
##5w-2w
B_diff_micro%>%
  ggplot(aes(x=log(Population.density),y=new.slope.diff52,
             group=major_teaching,color=major_teaching))+
  geom_point(aes(size=avg_full_work_prob_2w5w,alpha=avg_full_work_prob_2w5w))+
  geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+
  theme_minimal()+team_theme+
  labs(y="Change in Growth",x="Log of Population Density",
       title="Change in Growth between B(2) and B(5)\nOnly Micropolitan Counties",
       color="Majority teaching posture",
       size = "2w~5w\nAveraged %6hr+ Away Home",
       alpha= "2w~5w\nAveraged %6hr+ Away Home" ,fill="Majority teaching posture")+
  scale_color_manual(values=col_theme)+theme(legend.position = "bottom")+
  geom_text(data =B_diff_micro%>%filter(major_teaching=="On Premises"),
            aes(label=COUNTY),color='black',
            size=3,hjust=1.1, vjust=0.3)
```

Change in Growth between B(2) and B(5)
## Only Micropolitan Counties


```
##4w-1w
B_diff_micro%>%
  ggplot(aes(x=log(Population.density),
            y=new.slope.diff41,group=major_teaching,color=major_teaching))+
  geom_point(aes(size=avg_full_work_prob_1w4w,alpha=avg_full_work_prob_1w4w))+
  geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+theme_minimal()+team_theme+
  labs(y="Change in Growth",x="Log of Population Density",
      title="Change in Growth between B(1) and B(4)\nOnly Micropolitan Counties",
      color="Majority teaching posture",
      size = "1w~4w\nAveraged %6hr+ Away Home",
      alpha= "1w~4w\nAveraged %6hr+ Away Home" ,fill="Majority teaching posture")+
  scale_color_manual(values=col_theme)+theme(legend.position = "bottom")+
  geom_text(data =B_diff_micro%>%
              filter(major_teaching=="On Premises"),
          aes(label=COUNTY),color='black',
          size=3,hjust=1.1, vjust=0.3)
```

Change in Growth between B(1) and B(4)
Only Micropolitan Counties

```
##7w-4w
B_diff_micro%>%
  ggplot(aes(x=log(Population.density),
             y=new.slope.diff74,group=major_teaching,color=major_teaching))+
  geom_point(aes(size=avg_full_work_prob_4w7w,alpha=avg_full_work_prob_4w7w))+
  geom_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+theme_minimal()+team_theme+
  labs(y="Change in Growth",x="Log of Population Density",
       title="Change in Growth between B(4) and B(7)\nOnly Micropolitan Counties",
       color="Majority teaching posture",
       size = "4w~7w\nAveraged %6hr+ Away Home",
       alpha= "4w~7w\nAveraged %6hr+ Away Home" ,fill="Majority teaching posture")+
  scale_color_manual(values=col_theme)+
  theme(legend.position = "bottom")+
  geom_text(data =B_diff_micro%>%filter(major_teaching=="On Premises"),
            aes(label=COUNTY),
            color='black',size=3,hjust=1.1, vjust=0.3)
```

**Change in Growth between B(4) and B(7)
Only Micropolitan Counties**
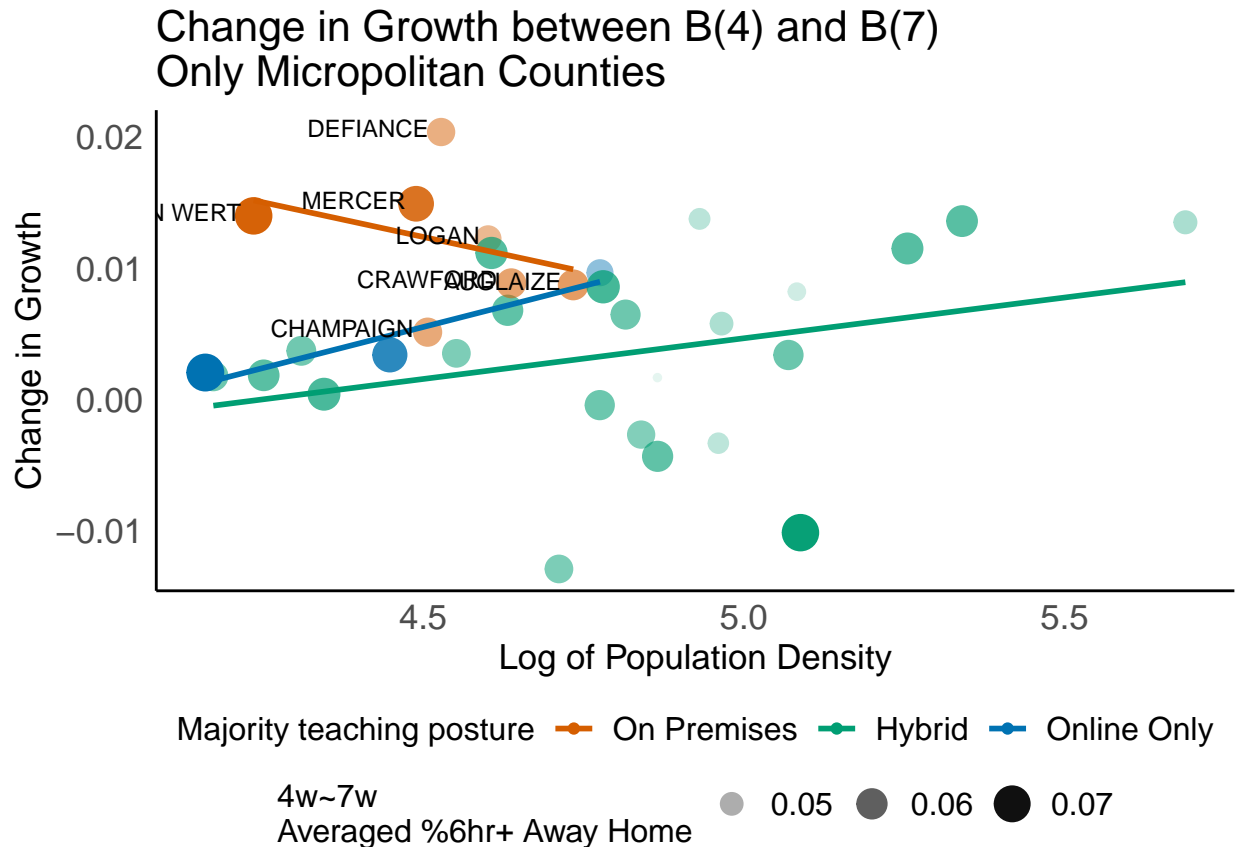
Appendix 8: Math Plots

Gamma distribution for the time lengths from infections to deaths

We know from previous study that the mean for this Gamma distribution is 23.9, with a coefficient of variation being 0.4.

```r
# package for The Gamma Distribution (Alternative Parameterization)
# install.packages("EnvStats")
library(EnvStats)
time_to_deaths <- 1:50
prob_time_to_deaths <- dgammaAlt(x = time_to_deaths,mean = 23.9, cv = 0.4)
## shift x
gamma_plot <- data.frame(prob_time_to_deaths,time_to_deaths,
                         time_to_deaths+5,time_to_deaths+10,time_to_deaths+15)
colnames(gamma_plot) <- c("prob","time1","time2","time3","time4")
ggplot(gamma_plot)+
  geom_line(aes(x=time1,y=prob),colour = "black")+
  geom_vline(xintercept = 5.2,lty=2,colour="darkgreen")+
  geom_vline(xintercept = 15.2,lty=2,colour="darkgreen")+
  geom_vline(xintercept = 25.2,lty=2,colour="darkgreen")+
  labs(x="Time from infections to deaths",
      y="Probability of died after x days")+team_theme+
  theme(legend.position = "bottom")
```