Technical Appendix

# Notes for appendix

The Appendix contains all the code and plots in the IDMRD paper and some other important supplementary plots.

Contents

[Notes for appendix 1](#_Toc72092894)

[Appendix 1: Map 2](#_Toc72092895)

[Teaching method, Population and Enrollment 3](#_Toc72092896)

[Covid deaths during fall semester and death proportion during fall semester 8](#_Toc72092897)

[Micropolitan Map 10](#_Toc72092898)

[Appendix 2: Death Incidence 12](#_Toc72092899)

[Data Process 12](#_Toc72092900)

[Daily Cases for 0-19 Year Age Group 14](#_Toc72092901)

[Death Prop Over Time by the Majority Teaching Method 15](#_Toc72092902)

[Pairwise of Death Incidence 16](#_Toc72092903)

[Appendix 3: Confounding Variables 17](#_Toc72092904)

[Uninsured Population 17](#_Toc72092905)

[Percentage of Senior 19](#_Toc72092906)

[Log Population Density 20](#_Toc72092907)

[Appendix 4: Exponential growth model 22](#_Toc72092908)

[Data process 22](#_Toc72092909)

[Select Max B 22](#_Toc72092910)

[Aggregate counties by teaching method 23](#_Toc72092911)

[Difference in B’s 24](#_Toc72092912)

[Distribution of Maximum Growth Coefficient for All Counties by Majority Teaching posture 27](#_Toc72092913)

[Max B vs. Mobility and Population Density for All Counties 28](#_Toc72092914)

[Appendix 5: Micropolitan Counties 30](#_Toc72092915)

[Distribution of Log Population Density by Rural-Urban Status colored by Majority Teaching posture 30](#_Toc72092916)

[Dotplots of Log Population Density for Micropolitan counties 31](#_Toc72092917)

[Distribution of Maximum Growth Coefficient Micropolitan Counties 32](#_Toc72092918)

[Max B vs. Mobility and Population Density in Micropolitan Counties 33](#_Toc72092919)

[Change in Growth vs Mobility and Log Population Density 35](#_Toc72092920)

[Sensitive Analysis 39](#_Toc72092921)

[Appendix 6: Math Plots 47](#_Toc72092922)

[Gamma distribution for the time lengths from infections to deaths 47](#_Toc72092923)

# Appendix 1: Map

This section contains all the types of map in our main paper.

Sys.setlocale("LC\_TIME", "English")

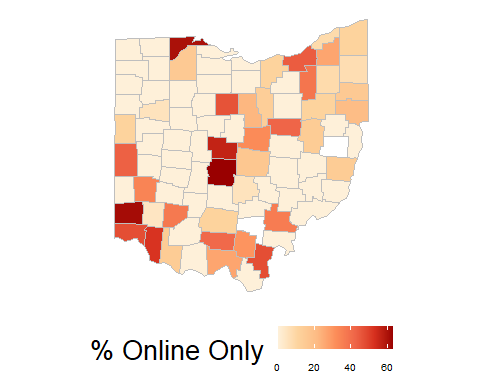
## [1] "English\_United States.1252"

library(ggrepel)  
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"))

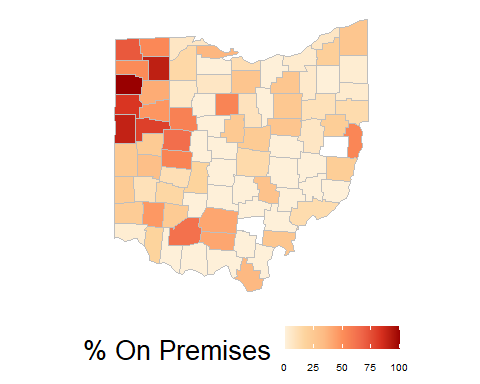
## Teaching method, Population and Enrollment

This section contains the percentage of different teaching methods, the number of enrolled students and population profile in counties.

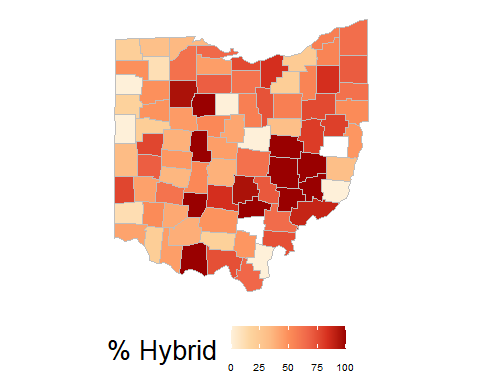
ohio\_map <- map\_data("county") %>%subset(region=="ohio")%>%  
 mutate(county=toupper(subregion))%>%select(long,lat,county,group)  
# create map plots  
wide\_teaching\_enroll%>%  
 left\_join(ohio\_map,by='county')%>%  
 mutate(Online\_Only= Online\_Only\*100)%>%  
 ggplot() +   
 geom\_polygon(aes(x = 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))



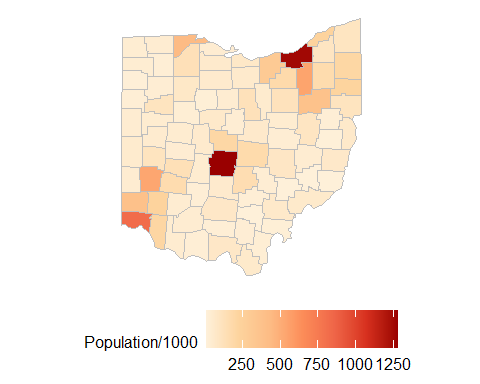
# create map plots  
wide\_teaching\_enroll%>%  
 left\_join(ohio\_map,by='county')%>%  
 mutate(On\_Premises= On\_Premises\*100)%>%  
 ggplot() +   
 geom\_polygon(aes(x = long, y = lat, group = group, fill = On\_Premises), color = "gray") +   
 coord\_fixed(1.3) + theme\_map() +  
 scale\_fill\_distiller(palette = "OrRd",direction = 1)+  
 labs(fill='% On Premises')+  
 theme(legend.position = "bottom",legend.text = element\_text(size=),legend.title = element\_text(size=20))



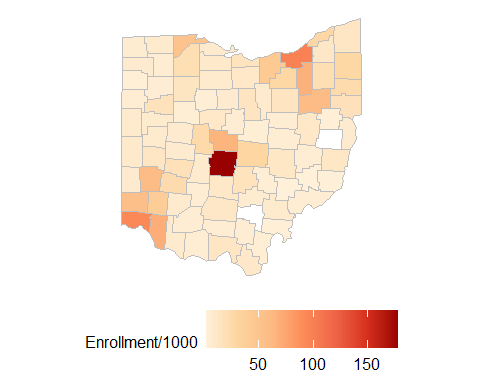
# create map plots for population  
wide\_teaching\_enroll%>%  
 left\_join(ohio\_map,by='county')%>%  
 mutate(Hybrid= Hybrid\*100)%>%  
 ggplot() +   
 geom\_polygon(aes(x = long, y = lat, group = group, fill = Hybrid), color = "gray") +   
 coord\_fixed(1.3) +   
 theme\_map() +  
 scale\_fill\_distiller(palette = "OrRd",direction = 1)+  
 labs(fill='% Hybrid')+  
 theme(legend.position = "bottom",legend.text = element\_text(size=),legend.title = element\_text(size=20))



# create map plots  
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"))

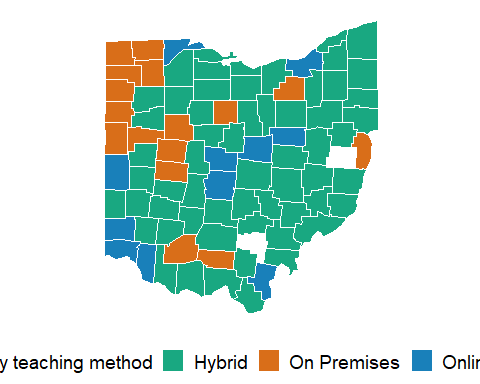


# create map plots  
teachingmethod\_enroll%>%  
 distinct(county,county\_enroll)%>%  
 left\_join(ohio\_map,by=c('county'))%>%  
 mutate(county\_enroll = county\_enroll/1000)%>%  
 ggplot() +   
 geom\_polygon(aes(x = long, y = lat, group = group, fill = county\_enroll), color = "gray") +   
 coord\_fixed(1.3) + 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"))



The following map shows the major teaching method in a certain county.

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)), color = "white",alpha=0.9) +   
 coord\_fixed(1.3) + theme\_map() +   
 scale\_fill\_manual(values=col\_theme)+   
 labs(fill='Majority teaching method')+  
 theme(legend.position = "bottom",legend.text = element\_text(size=14),legend.title = element\_text(size=14))

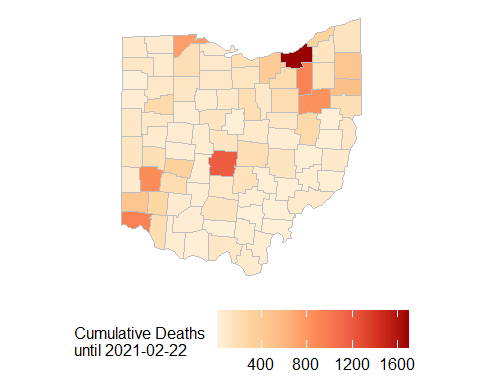


## Covid deaths during fall semester and death proportion during fall semester

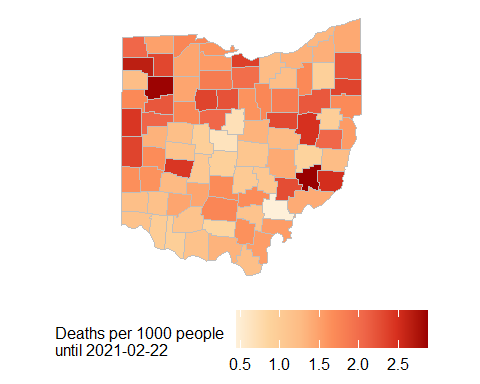
This section shows the death profile in the county(both cumulative death numbers and death per 1000 people).

getLabelPoint <- # Returns a county-named list of label points  
function(county) {Polygon(county[c('long', 'lat')])@labpt}  
centroids = by(ohio\_map, ohio\_map$county, getLabelPoint)# Returns list  
centroids2 <- do.call("rbind.data.frame", centroids)# Convert to Data Frame  
centroids2$county = str\_to\_title(rownames(centroids))  
names(centroids2) <- c('clong', 'clat', "county") # Appropriate Header

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"))



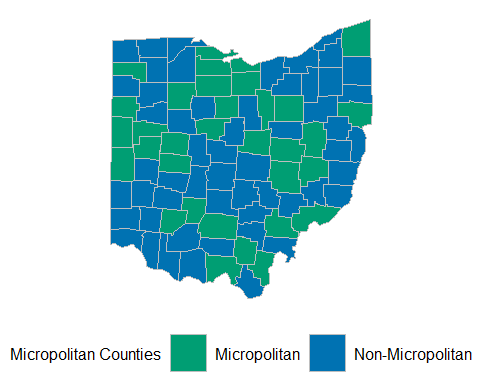
death\_prop%>%  
 left\_join(ohio\_map,by=c("COUNTY"='county'))%>%  
 ggplot() +   
 geom\_polygon(aes(x = long, y = lat, group=group,fill = death\_per\_1000),   
 color = "gray") +   
 coord\_fixed(1.3) + theme\_map() +  
 scale\_fill\_distiller(palette = "OrRd",direction = 1)+  
 labs(fill='Deaths per 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"))



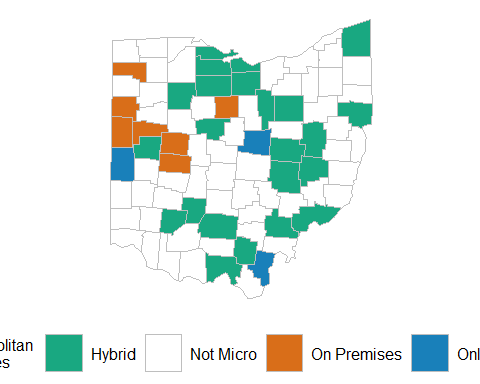
## Micropolitan Map

Since we reduced our county samples to micropolitan counties, we also show how micropolitan county distribute in Ohio and which teaching method they used. We can find that micropolitan counties are well-spread.

# Micropolitan map plots  
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 map plots  
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"))



# Appendix 2: Death Incidence

## Data Process

library(tidyverse)  
library(lubridate)  
require(scales)  
library(readxl)  
cases\_by\_age <- read\_excel("OhiobyAge.xlsx")  
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))  
county\_policy\_wide$major\_teaching <- factor(county\_policy\_wide$major\_teaching,  
 levels = c("On Premises","Hybrid","Online Only"))  
# 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 %>%  
 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%>%select(COUNTY,major\_opendate),by=c('COUNTY'))  
# Box Plots in Fall semester  
library(PMCMRplus)  
require(DescTools)  
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)   
fall\_major\_teaching.aov <- aov(death\_incidence\_per\_1000 ~ major\_teaching,  
 data = fall\_cases)  
summary(fall\_major\_teaching.aov) # p-value of .012

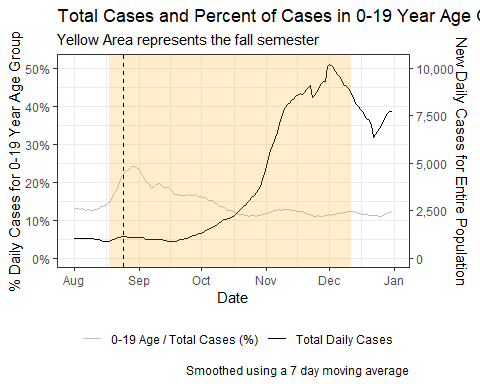
## Df Sum Sq Mean Sq F value Pr(>F)   
## major\_teaching 2 1.653 0.8264 5.205 0.00761 \*\*  
## Residuals 76 12.067 0.1588   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

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"))%>%  
 select(group1, group2, pval, p)  
library(ggpubr)

## Daily Cases for 0-19 Year Age Group

This plot shows total cases and percent of cases in 0-19 Year Age group, this special figure is one reason why we would like investigate schooling effect.

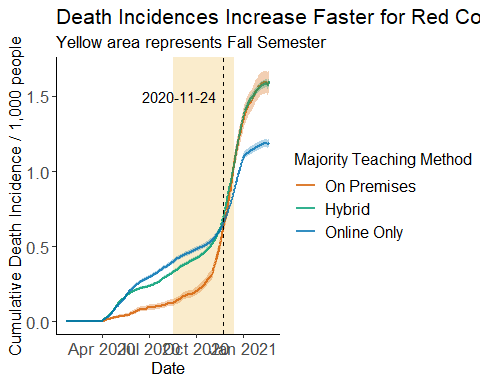
colors <- c("Total Daily Cases" = "black",  
 "0-19 Age / Total Cases (%)" = "gray")  
coeff <- 200  
ggplot(rolling\_age\_cases, aes(x=Date)) +   
 geom\_line( aes(y=youth\_prop\_roll,  
 color = "0-19 Age / Total Cases (%)"),   
 na.rm = T)+  
 geom\_line( aes(y=all\_roll/coeff,  
 color = "Total Daily 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)+   
 theme\_bw() +   
 theme(legend.position='bottom')



## Death Prop Over Time by the Majority Teaching Method

This section shows a special death incidence plot over time, which is the reason we want to investigate the change of death incidence.

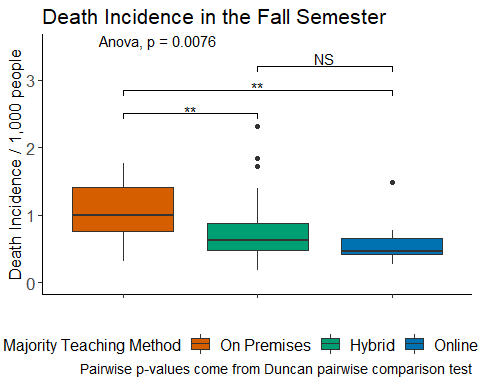
case\_policy\_wide%>%  
 group\_by(DATE, major\_teaching) %>%  
 drop\_na(major\_teaching)%>%  
 summarise(total\_deaths = sum(CUMDEATHS),  
 total\_pop = sum(POPULATION),  
 death\_prop = total\_deaths/total\_pop,  
 death\_prop\_upper = death\_prop + z\_cl\*sqrt(death\_prop\*(1 - death\_prop)/total\_pop),  
 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) +   
 theme\_bw() +   
 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 Method") +team\_theme+ scale\_colour\_manual(values=col\_theme)+scale\_fill\_manual(values=col\_theme)



Pairwise of Death Incidence

We find that death incidence during the fall semester for different teaching methods show obvious significant difference.

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 Method",  
 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)



# Appendix 3: Confounding Variables

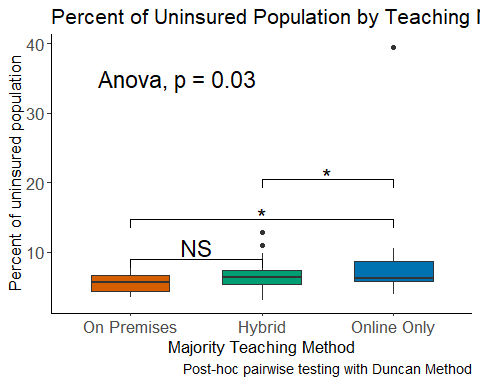
There are also some other variables which may affect death incidence. We do some pairwise analysis based on teaching methods for these variables.

## Uninsured Population

library(ggpubr)  
library(PMCMRplus)  
require(DescTools)  
  
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"))  
  
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

# 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"))%>%  
 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 Method",y="Percent of uninsured population",caption = "Post-hoc pairwise testing with Duncan Method")+  
 team\_theme+theme(legend.position = "")+scale\_fill\_manual(values=col\_theme)

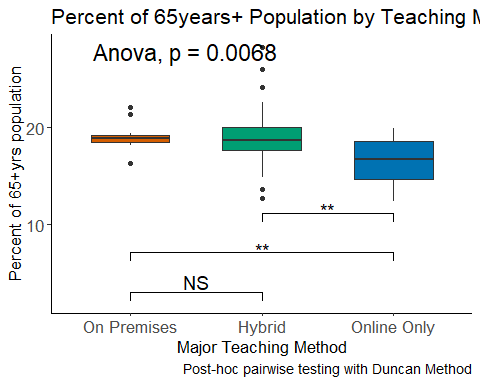


## Percentage of Senior

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

# 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"))%>%  
 select(group1, group2, pval, p)  
  
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 Method",caption = "Post-hoc pairwise testing with Duncan Method")+team\_theme+theme(legend.position = "")+scale\_fill\_manual(values=col\_theme)

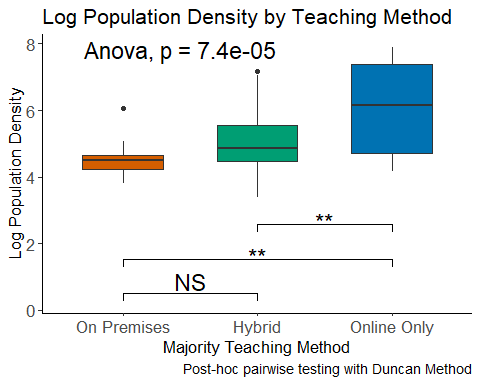


## Log Population Density

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

# 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"))%>%  
 select(group1, group2, pval, p)  
  
  
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 Method",y="Log Population Density",caption = "Post-hoc pairwise testing with Duncan Method")+  
 team\_theme+theme(legend.position = "")+scale\_fill\_manual(values=col\_theme)



These confounding variables show significant difference for different teahcing methods.

# Appendix 4: Exponential growth model

We construct the exponential growth in order to better measure the state of pandemic.

## Data process

cases\_slope <- read.csv("county\_splines.csv", header = T)%>%  
 select(COUNTY,DATE,POPULATION,CUMDEATHS,log\_tot\_deaths,tot.slope,NEWDEATHS,rev\_NEWDEATHS,log\_new\_deaths,new.slope)  
# SHIFT THE DATE!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!  
cases\_slope$DATE <- as.Date(cases\_slope$DATE)-24  
# get majority teaching method wide\_teaching\_enroll  
cases\_slope\_teach <-death\_teaching%>%  
 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)

## Select Max B

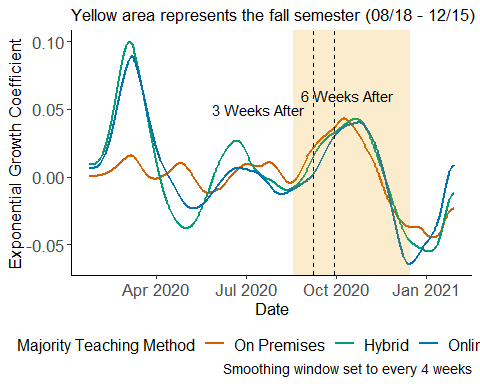
Max B during the fall semester is used to meaure how severe the pandemic is in the county.

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%>%  
 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"))

## Aggregate counties by teaching method

We aggregate the growth coefficient based on the major teaching method used, and draw a time series plot for different teahcing methods.

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 Method",  
 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)



## Difference in B’s

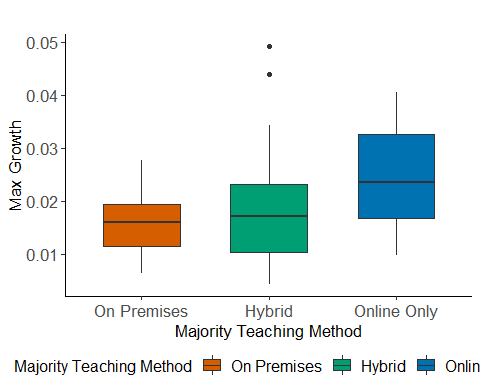
We capture the B values in some specific time point.

B0w <- cases\_slope\_teach%>%  
 filter(DATE==as.Date("2020-08-18"))%>%  
 drop\_na(major\_teaching)%>%  
 rename(new.slope0w=new.slope)  
B3w <- cases\_slope\_teach%>%  
 filter(DATE==as.Date("2020-08-18")+21)%>%  
 drop\_na(major\_teaching)%>%  
 rename(new.slope3w=new.slope)  
B6w <- cases\_slope\_teach%>%  
 filter(DATE==as.Date("2020-08-18")+42)%>%  
 drop\_na(major\_teaching)%>%  
 rename(new.slope6w=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)  
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)%>%  
 drop\_na(major\_teaching)%>%  
 rename(new.slope5w=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"))%>%  
 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%>%select(COUNTY,new.slope3w),by="COUNTY")%>%  
 left\_join(B0w%>%select(COUNTY,new.slope0w),by="COUNTY")%>%  
 left\_join(B1w%>%select(COUNTY,new.slope1w),by="COUNTY")%>%  
 left\_join(B2w%>%select(COUNTY,new.slope2w),by="COUNTY")%>%  
 left\_join(B4w%>%select(COUNTY,new.slope4w),by="COUNTY")%>%  
 left\_join(B5w%>%select(COUNTY,new.slope5w),by="COUNTY")%>%  
 left\_join(B7w%>%select(COUNTY,new.slope7w),by="COUNTY")%>%  
 left\_join(Bm1w%>%select(COUNTY,new.slopem1w),by="COUNTY")%>%  
 left\_join(Bm2w%>%select(COUNTY,new.slopem2w),by="COUNTY")%>%  
 left\_join(Bm3w%>%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"))

## 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.

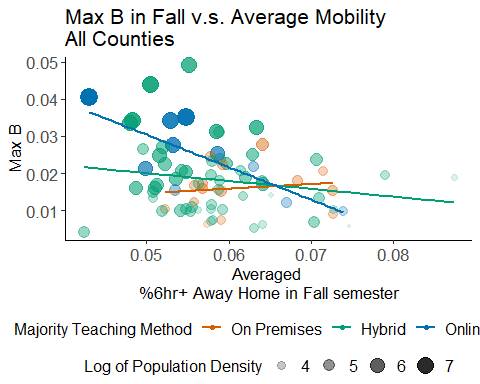
na.omit(B0B1)%>%  
 ggplot(aes(x=major\_teaching,y=max\_B1))+geom\_boxplot(aes(fill=major\_teaching),width=0.6)+  
 labs(title="",x="Majority Teaching Method",y="Max Growth",fill="Majority Teaching Method")+  
 team\_theme+theme(legend.position = "bottom")+scale\_fill\_manual(values=col\_theme)



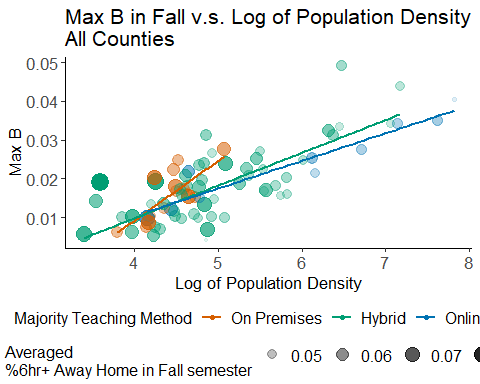
Max B vs. Mobility and Population Density for All Counties

The following plots show how Max Growth Coefficient correlated with Mobility and Log Population Density.

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 Method",size = "Log of Population Density",alpha= "Log of Population Density" )+  
 team\_theme+theme(legend.position = "bottom")+  
 scale\_color\_manual(values=col\_theme)



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 Method",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)



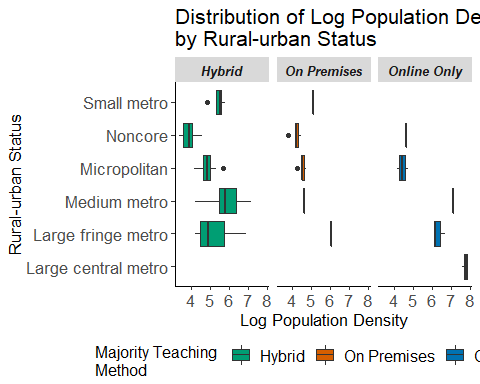
# Appendix 5: Micropolitan Counties

We reduce our sample to micropolitan counties in order to reduce the effect from other confounders. This section gives plots related to micropolitan coungties.

## Distribution of Log Population Density by Rural-Urban Status colored by Majority Teaching posture

Micropolitan counties show similar log population density and have data for all three teaching methods.

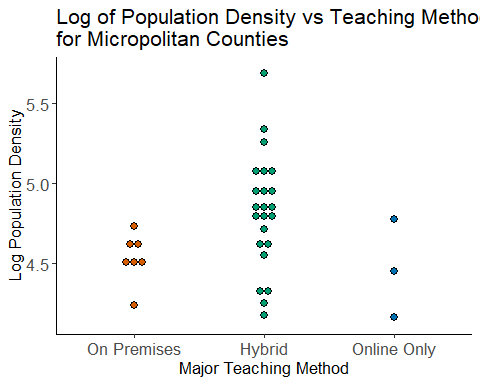
# Pop density vs RURAL  
ohio\_profile%>%  
 left\_join(wide\_teaching\_enroll[,c("county","major\_teaching")],  
 by = c("County" = "county"))%>%  
 drop\_na(major\_teaching)%>%  
 ggplot(aes(y=NCHS.Urban.Rural.Status,x=log(Population.density),  
 fill=major\_teaching))+  
 facet\_grid(~major\_teaching)+  
 geom\_boxplot()+  
 labs(fill="Majority Teaching \nMethod",  
 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")



## Dotplots of Log Population Density for Micropolitan counties

The dot plots show that the distributions of population density for micropolitan counties taking On Premises and Online Only method are quite similar, which means that it is reasonable to compare these two type of counties.

ohio\_profile%>%  
 filter(NCHS.Urban.Rural.Status=="Micropolitan")%>%  
 left\_join(B0B1%>%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="Major Teaching Method",title="Log of Population Density vs Teaching Method \nfor Micropolitan Counties")+scale\_fill\_manual(values=col\_theme)



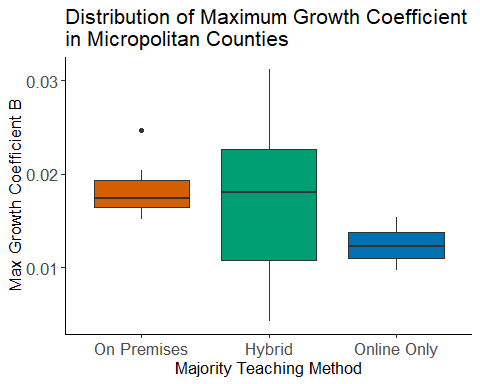
## Distribution of Maximum Growth Coefficient Micropolitan Counties

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

##### Micropolitan counties  
maxB\_major\_teaching.aov <- aov(max\_B1 ~ major\_teaching,data = na.omit(B0B1)%>%  
 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

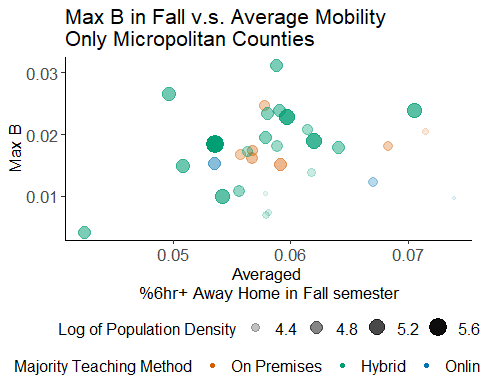
# p-value of .012  
stat.test <- PostHocTest(maxB\_major\_teaching.aov, method = "duncan")$major\_teaching %>%  
 as.data.frame()%>%  
 rownames\_to\_column("group") %>%  
 separate(group,"-", into = c("group1","group2")) %>%  
 mutate(pval = round(pval,3),  
 p = case\_when(pval <= .01~ "\*\*",  
 pval <= .05 ~ "\*",  
 TRUE ~ "NS"))%>%  
 select(group1, group2, pval, p)  
na.omit(B0B1)%>%  
 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 Method",title="Distribution of Maximum Growth Coefficient \nin Micropolitan Counties",  
 fill="Majority Teaching Method")+  
 scale\_fill\_manual(values=col\_theme)



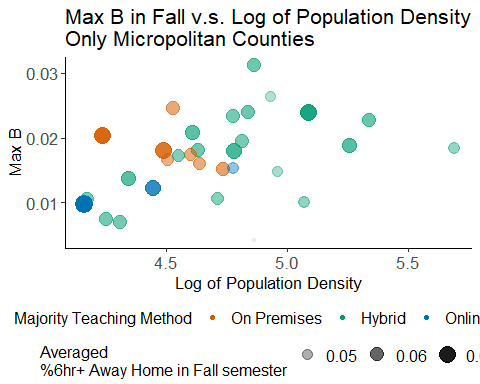
## Max B vs. Mobility and Population Density in Micropolitan Counties

These plots show that after blocking data to micropolitan counties, the relationship between growth coefficient and mobility & log population density is not so obvious.

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 Method",  
 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)  
 )



######  
na.omit(B0B1)%>%  
 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 Method",  
 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)



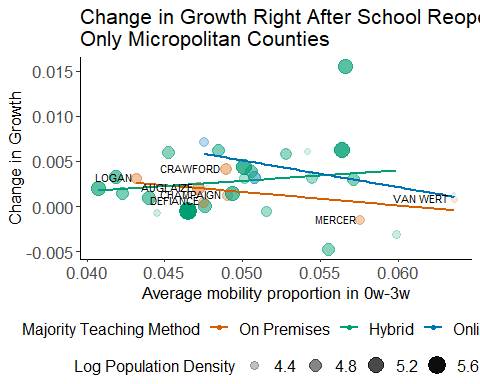
## Change in Growth vs Mobility and Log Population Density

The follwing plots show the change in growth right after school and three weeks later and we observed obvious shifting of red line.

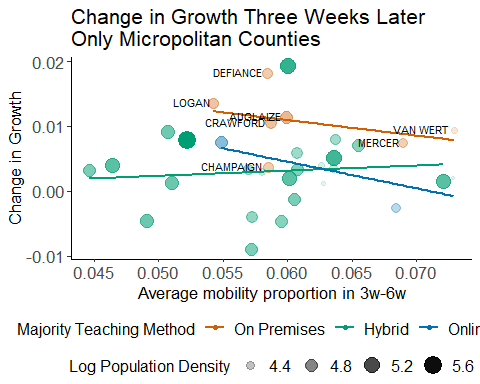
B\_diff\_micro <- B\_diff%>%  
 drop\_na(major\_teaching)%>%  
 filter(NCHS.Urban.Rural.Status=="Micropolitan") %>%  
 mutate(acc = new.slope.diff2 - new.slope.diff)

**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 Method",  
 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)

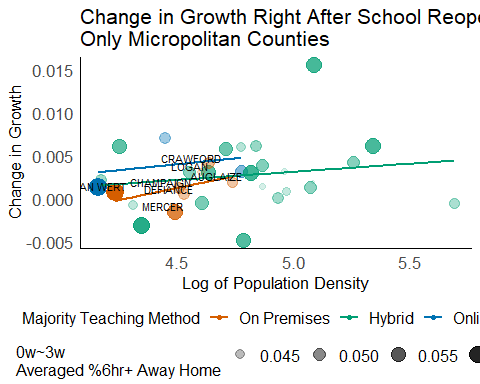


B\_diff\_micro%>%  
 ggplot(aes(x=avg2\_full\_work\_prob,y=new.slope.diff2,group=major\_teaching,color=major\_teaching))+geom\_point(aes(size=log(Population.density),alpha=log(Population.density)))+geom\_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+  
 labs(y="Change in Growth",x="Average mobility proportion in 3w-6w",  
 title="Change in Growth Three Weeks Later\nOnly Micropolitan Counties",  
 color="Majority Teaching Method",  
 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)

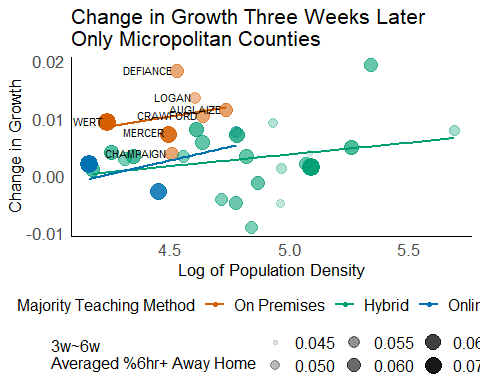


**Log Population Density**

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 Method",  
 size = "0w~3w\nAveraged %6hr+ Away Home",  
 alpha= "0w~3w\nAveraged %6hr+ Away Home" ,fill="Majority Teaching Method")+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)



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 Method",  
 size = "3w~6w\nAveraged %6hr+ Away Home",  
 alpha= "3w~6w\nAveraged %6hr+ Away Home" ,fill="Majority Teaching Method")+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)



## Sensitive Analysis

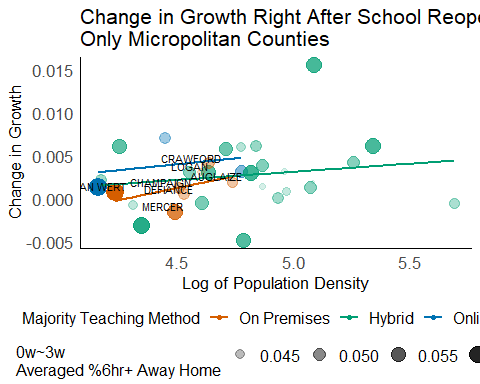
In order to make sure that the shifting of the red line (On Premises counties) is not brought by chance, we will conduct a sensitive analysis to detect how the change in growth varies throughout time. The whole sensitive analysis is based on ‘Change in growth versus Log Population Density.

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:

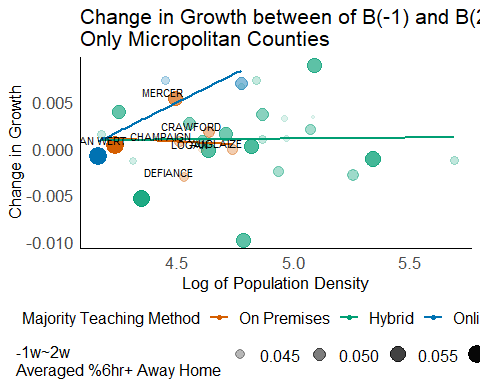
Before school posture taking effect: B(0)-B(-3), B(1)-B(-2), B(2)-B(-1), B(3)-B(0) (also known as change in growth right after the start of school reopen).

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)

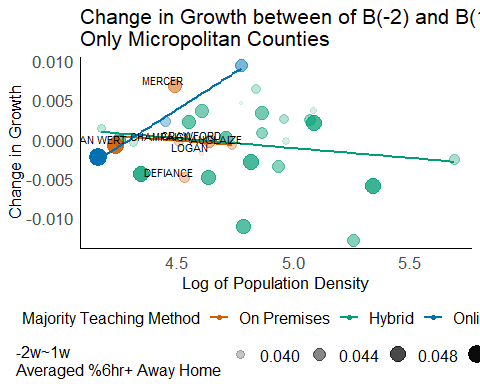
#Before reopen  
##3w-0w  
B\_diff\_micro%>%  
 ggplot(aes(x=log(Population.density),y=new.slope.diff,group=major\_teaching,color=major\_teaching))+geom\_point(aes(size=avg\_full\_work\_prob,alpha=avg\_full\_work\_prob))+geom\_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+theme\_minimal()+team\_theme+  
 labs(y="Change in Growth",x="Log of Population Density",  
 title="Change in Growth Right After School Reopen\nOnly Micropolitan Counties",  
 color="Majority Teaching Method",  
 size = "0w~3w\nAveraged %6hr+ Away Home",  
 alpha= "0w~3w\nAveraged %6hr+ Away Home" ,fill="Majority Teaching Method")+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)



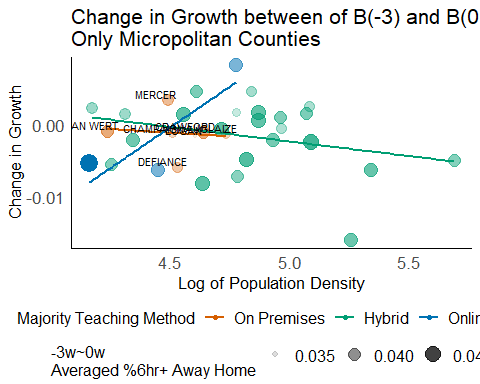
##2w-(-1w)  
B\_diff\_micro%>%  
 ggplot(aes(x=log(Population.density),y=new.slope.diff2m1,group=major\_teaching,color=major\_teaching))+geom\_point(aes(size=avg\_full\_work\_prob\_m1w2w,alpha=avg\_full\_work\_prob\_m1w2w))+geom\_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+theme\_minimal()+team\_theme+  
 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 Method",  
 size = "-1w~2w\nAveraged %6hr+ Away Home",  
 alpha= "-1w~2w\nAveraged %6hr+ Away Home" ,fill="Majority Teaching Method")+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)



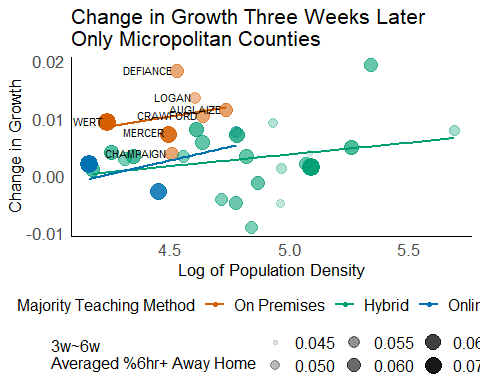
##1w-(-2w)  
B\_diff\_micro%>%  
 ggplot(aes(x=log(Population.density),y=new.slope.diff1m2,group=major\_teaching,color=major\_teaching))+geom\_point(aes(size=avg\_full\_work\_prob\_m2w1w,alpha=avg\_full\_work\_prob\_m2w1w))+geom\_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+theme\_minimal()+team\_theme+  
 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 Method",  
 size = "-2w~1w\nAveraged %6hr+ Away Home",  
 alpha= "-2w~1w\nAveraged %6hr+ Away Home" ,fill="Majority Teaching Method")+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)



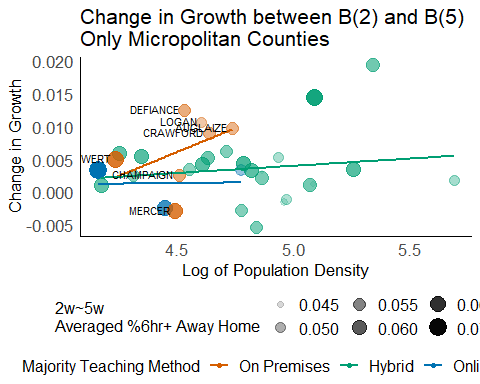
##0w-(-3w)  
B\_diff\_micro%>%  
 ggplot(aes(x=log(Population.density),y=new.slope.diff0m3,group=major\_teaching,color=major\_teaching))+geom\_point(aes(size=avg\_full\_work\_prob\_m3w0w,alpha=avg\_full\_work\_prob\_m3w0w))+geom\_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+theme\_minimal()+team\_theme+  
 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 Method",  
 size = "-3w~0w\nAveraged %6hr+ Away Home",  
 alpha= "-3w~0w\nAveraged %6hr+ Away Home" ,fill="Majority Teaching Method")+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)



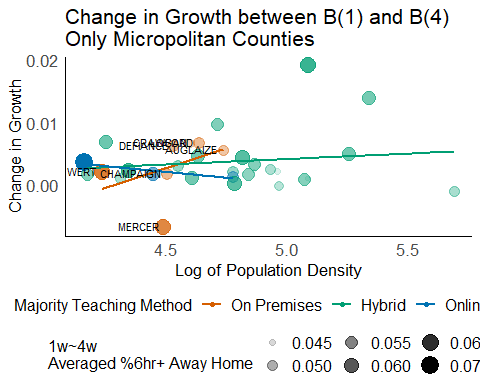
#After reopen  
##6w-3w  
B\_diff\_micro%>%  
 ggplot(aes(x=log(Population.density),y=new.slope.diff2,group=major\_teaching,color=major\_teaching))+geom\_point(aes(size=avg2\_full\_work\_prob,alpha=avg2\_full\_work\_prob))+geom\_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+theme\_minimal()+team\_theme+  
 labs(y="Change in Growth",x="Log of Population Density",  
 title="Change in Growth Three Weeks Later\nOnly Micropolitan Counties",  
 color="Majority Teaching Method",  
 size = "3w~6w\nAveraged %6hr+ Away Home",  
 alpha= "3w~6w\nAveraged %6hr+ Away Home" ,fill="Majority Teaching Method")+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)



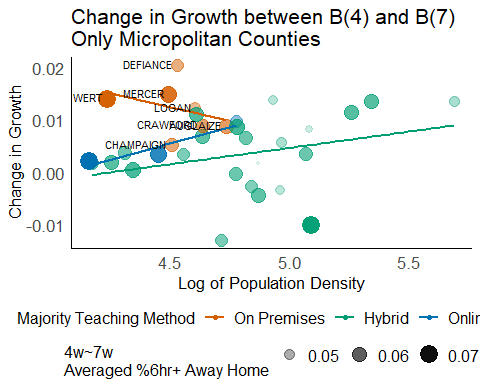
##5w-2w  
B\_diff\_micro%>%  
 ggplot(aes(x=log(Population.density),y=new.slope.diff52,group=major\_teaching,color=major\_teaching))+geom\_point(aes(size=avg\_full\_work\_prob\_2w5w,alpha=avg\_full\_work\_prob\_2w5w))+geom\_smooth(method = "lm", se=F, formula = y ~ x,alpha=0.1)+theme\_minimal()+team\_theme+  
 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 Method",  
 size = "2w~5w\nAveraged %6hr+ Away Home",  
 alpha= "2w~5w\nAveraged %6hr+ Away Home" ,fill="Majority Teaching Method")+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)



##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 Method",  
 size = "1w~4w\nAveraged %6hr+ Away Home",  
 alpha= "1w~4w\nAveraged %6hr+ Away Home" ,fill="Majority Teaching Method")+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)



##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 Method",  
 size = "4w~7w\nAveraged %6hr+ Away Home",  
 alpha= "4w~7w\nAveraged %6hr+ Away Home" ,fill="Majority Teaching Method")+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)



# Appendix 6: 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.

# 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")

