log graphs for each county

Cheyenne Ehman

4/7/2021

```
library(readxl)
library(tidyverse)
library(ggformula)
# read in OHIO_CASES_DATA
cases <- read_excel("COVID_CASES_OH_CNTY_20210223_pop.xlsx")
cases$DATE <- as.Date(cases$DATE, "%m/%d/%Y") # convert dates
cases <- cases%>%
  filter( (COUNTY != 'UNASSIGNED') & (COUNTY !='OUT OF OH'))%>% # remove UNASSIGNED and OUT OF OH data
  mutate(FIPS = str_sub(UID,start = 4,end = 8),
         NEWDEATHS = ifelse(is.na(NEWDEATHS),0,NEWDEATHS))%>%
  select(COUNTY,FIPS,DATE,CNTY_LAT,CNTY_LONG,POPULATION,CUMCONFIRMED,CUMDEATHS,NEWDEATHS,NEWCONFIRMED) %
  arrange(DATE)
```

Fix negative values in NEW DEATHS

```
cases[cases$NEWDEATHS<0,]
```

```
## # A tibble: 56 x 10
##    COUNTY FIPS  DATE       CNTY_LAT CNTY_LONG POPULATION CUMCONFIRMED CUMDEATHS
##    <chr>  <chr> <date>        <dbl>     <dbl>      <dbl>        <dbl>     <dbl>
##  1 MIAMI  39109 2020-03-22     40.1     -84.2     106439           13         0
##  2 FRANK~ 39049 2020-03-30     40.0     -83.0    1283688          281         2
##  3 ADAMS  39001 2021-02-18     38.8     -83.5      27960         2118        29
##  4 ALLEN  39003 2021-02-18     40.8     -84.1     106160        10655       228
##  5 ASHLA~ 39005 2021-02-18     40.8     -82.3      53973         3709        91
##  6 ASHTA~ 39007 2021-02-18     41.7     -80.7      98637         5564       137
##  7 BELMO~ 39013 2021-02-18     40.0     -81.0      69738         4828        84
##  8 BUTLER 39017 2021-02-18     39.4     -84.6     383683        34548       452
##  9 CHAMP~ 39021 2021-02-18     40.1     -83.8      39686         2791        44
## 10 CLARK  39023 2021-02-18     39.9     -83.8     136118        12018       320
## # ... with 46 more rows, and 2 more variables: NEWDEATHS <dbl>,
## #   NEWCONFIRMED <dbl>
```

```
#test case
newdeaths <- cases[cases$COUNTY=='MORROW',]$NEWDEATHS
distri_neg <- function(newdeaths){
  for (i in 1:length(newdeaths)) {
    if(newdeaths[i] < 0){
      if(i == 1){
        stop("problem")
      }
      j = i-1
      while (newdeaths[i]<0 && j>=1) {
```

```
        if(newdeaths[j]>0){
          newdeaths[j] = newdeaths[j] - 1
          newdeaths[i] = newdeaths[i] + 1
        }
        j = j - 1
      }
    }
    if(newdeaths[i]<0){print("Still negative, need to double check")}
    }
  return(newdeaths)
}
# test out
#distri_neg(newdeaths)
## apply to each county
cases <- cases%>%
  group_by(COUNTY)%>%
  mutate(rev_NEWDEATHS = distri_neg(NEWDEATHS))
# double check
cases[cases$rev_NEWDEATHS<0,]
```

```
## # A tibble: 0 x 11
## # Groups:   COUNTY [0]
## # ... with 11 variables: COUNTY <chr>, FIPS <chr>, DATE <date>, CNTY_LAT <dbl>,
## #   CNTY_LONG <dbl>, POPULATION <dbl>, CUMCONFIRMED <dbl>, CUMDEATHS <dbl>,
## #   NEWDEATHS <dbl>, NEWCONFIRMED <dbl>, rev_NEWDEATHS <dbl>
```

```
# two counties have negative new deaths at the beginning
```

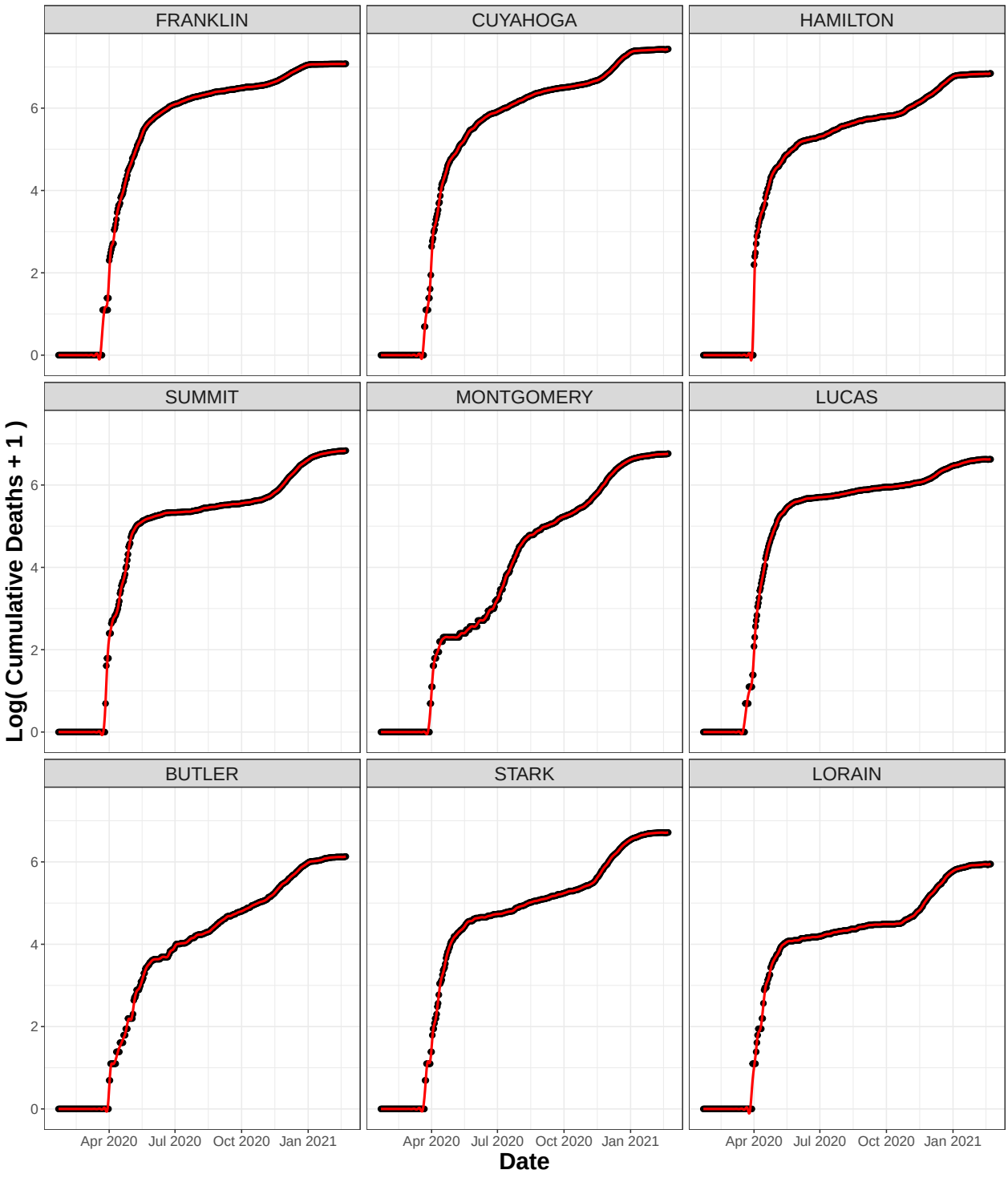
Fit Splines to Log(New Deaths + 1)

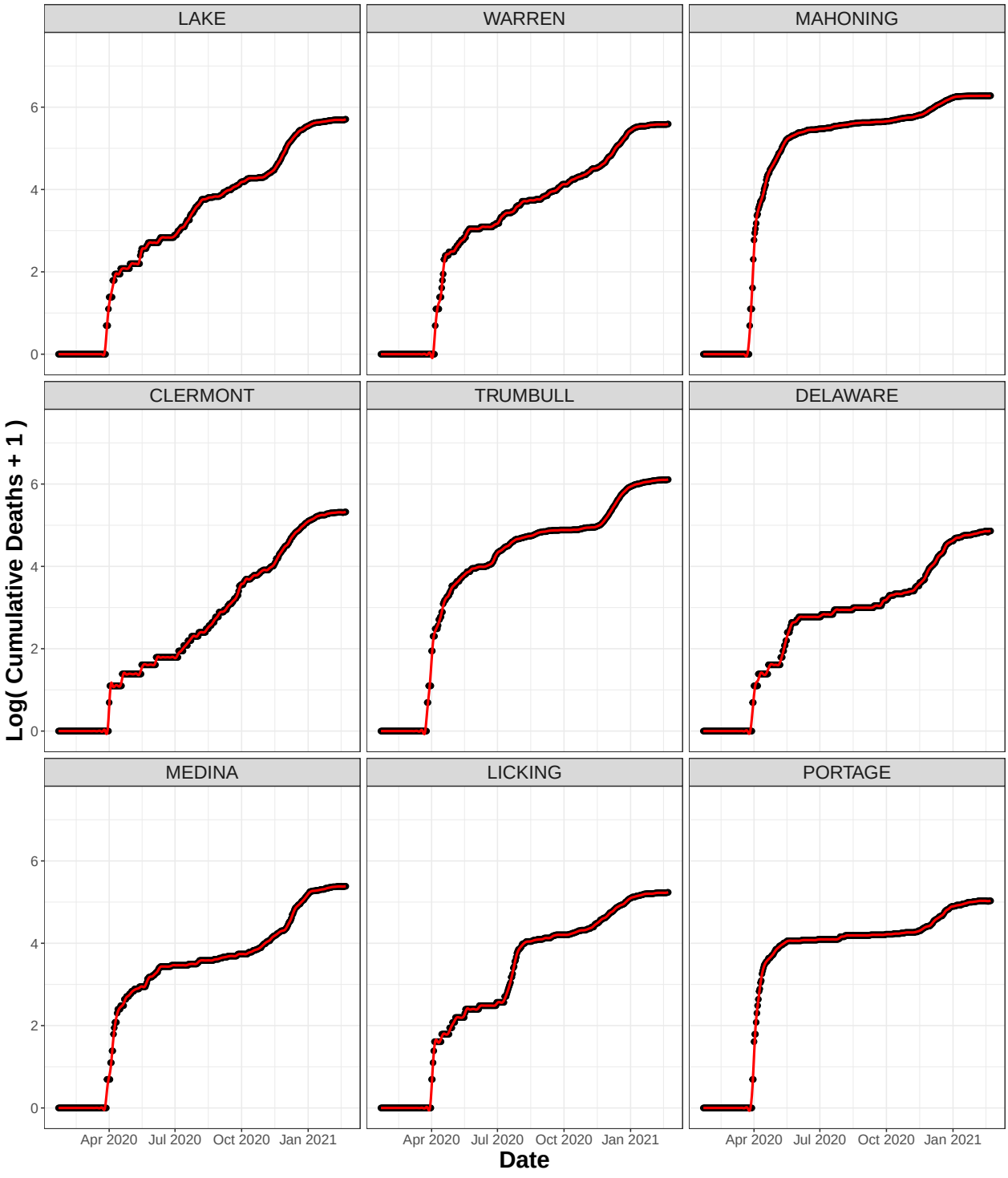
```
log_deaths_county_df <- cases %>%
  group_by(COUNTY) %>%
  mutate(log_tot_deaths = log(CUMDEATHS + 1),
         log_new_deaths = log(rev_NEWDEATHS + 1),
         tot.smoothed.spline = smooth.spline(DATE,log_tot_deaths, cv = TRUE)$y,
         new.smoothed.spline = smooth.spline(DATE,log_new_deaths, cv = TRUE)$y,
         tot.slope = predict(smooth.spline(DATE,log_tot_deaths, cv = TRUE),deriv = 1)$y,
         new.slope = predict(smooth.spline(DATE,log_new_deaths, cv = TRUE),deriv = 1)$y)
#Want to order the counties by most Populous
log_deaths_county_df$COUNTY <- factor(log_deaths_county_df$COUNTY,
                                      levels = levels(fct_reorder(log_deaths_county_df$COUNTY,
```

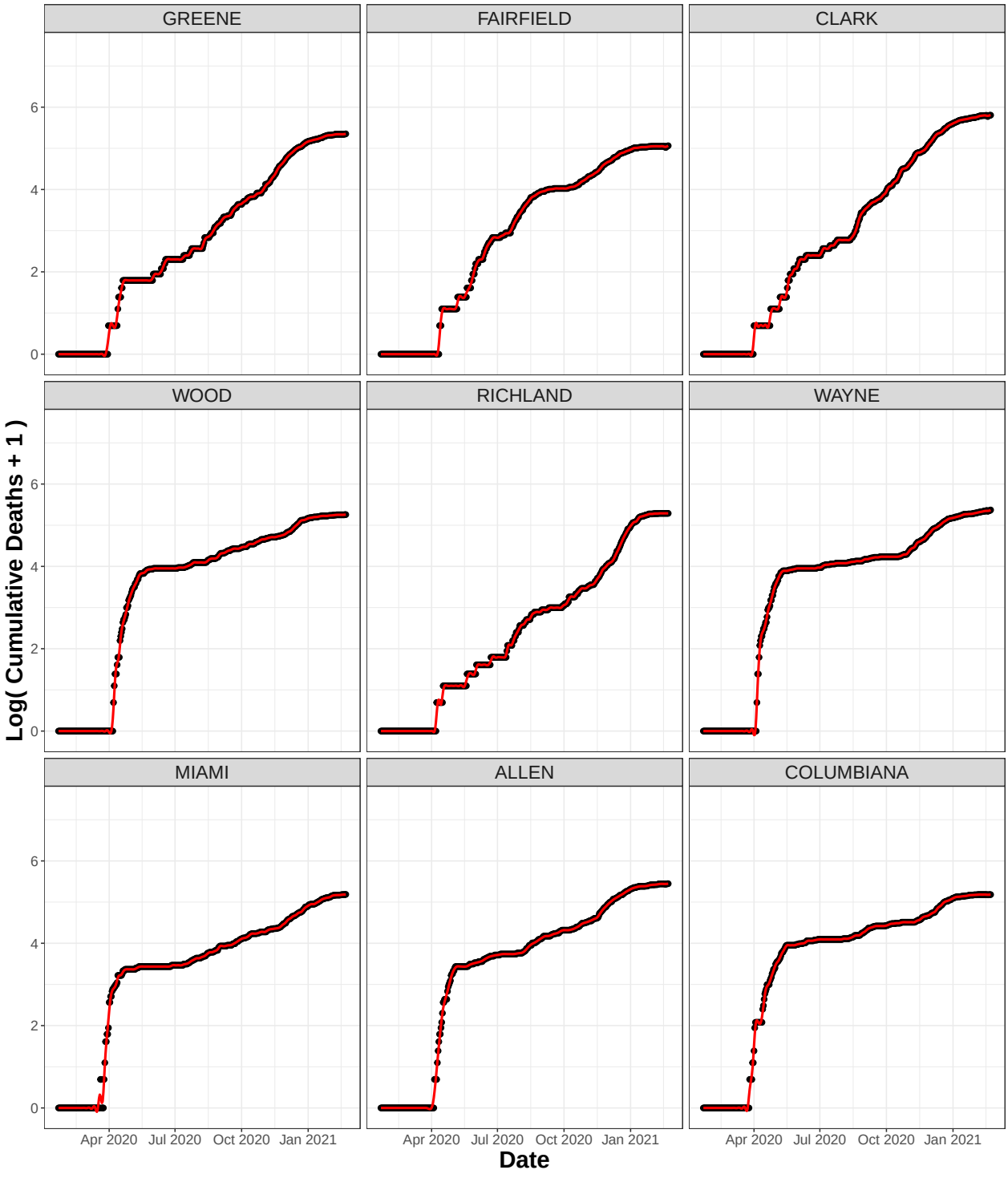
Cumulative Deaths

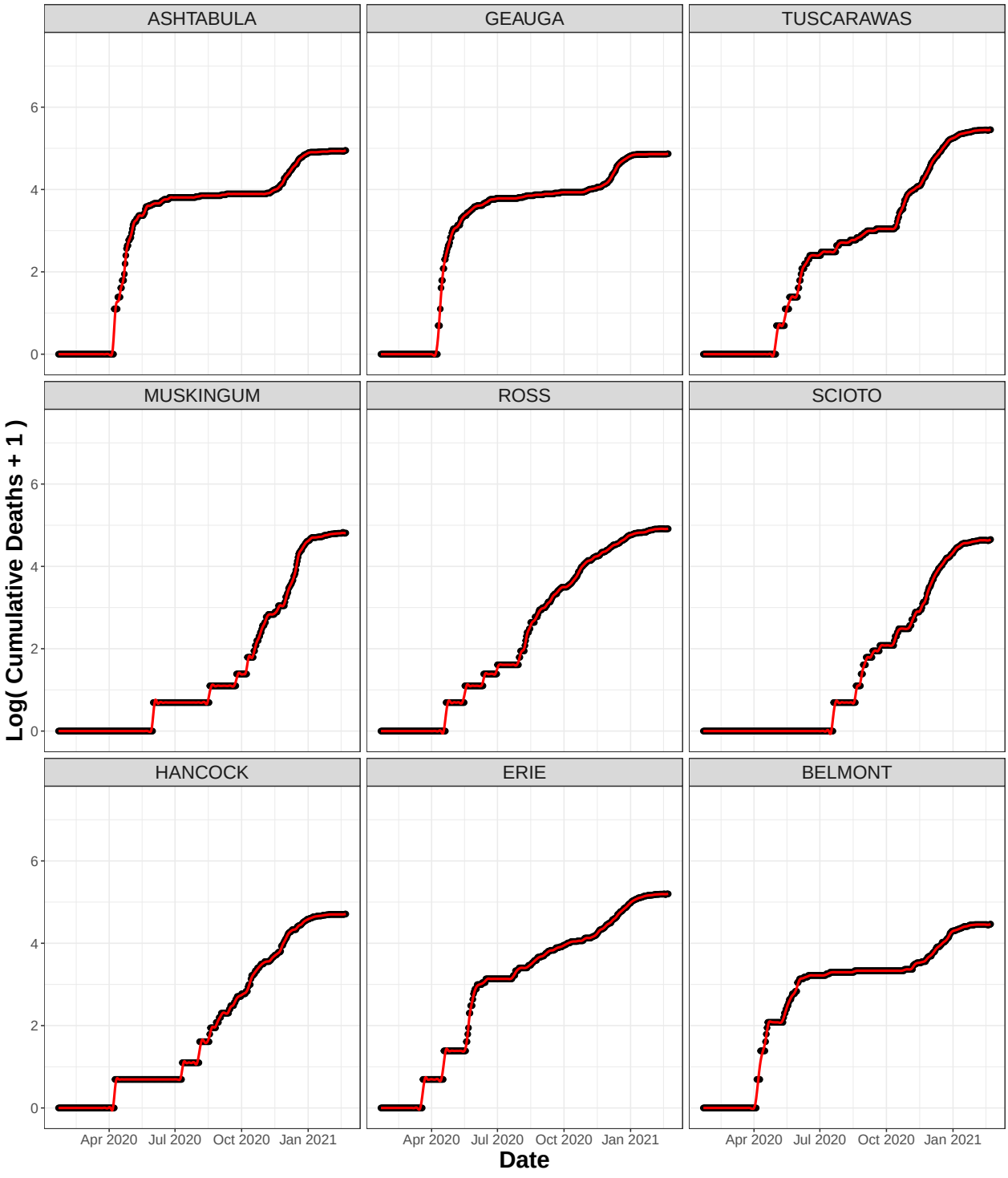
```
library(ggforce)
for(i in 1:10){
  p <- ggplot(log_deaths_county_df, aes(x = DATE, y = log_tot_deaths)) +
      geom_point(size = 2)+
      geom_line(aes(x = DATE,y = tot.smoothed.spline), color = "red",size = 1)+
      facet_wrap_paginate(~COUNTY, ncol = 3, nrow = 3, page = i)+
      theme_bw() +
      labs(x = "Date", y = "Log( Cumulative Deaths + 1 )")+
```

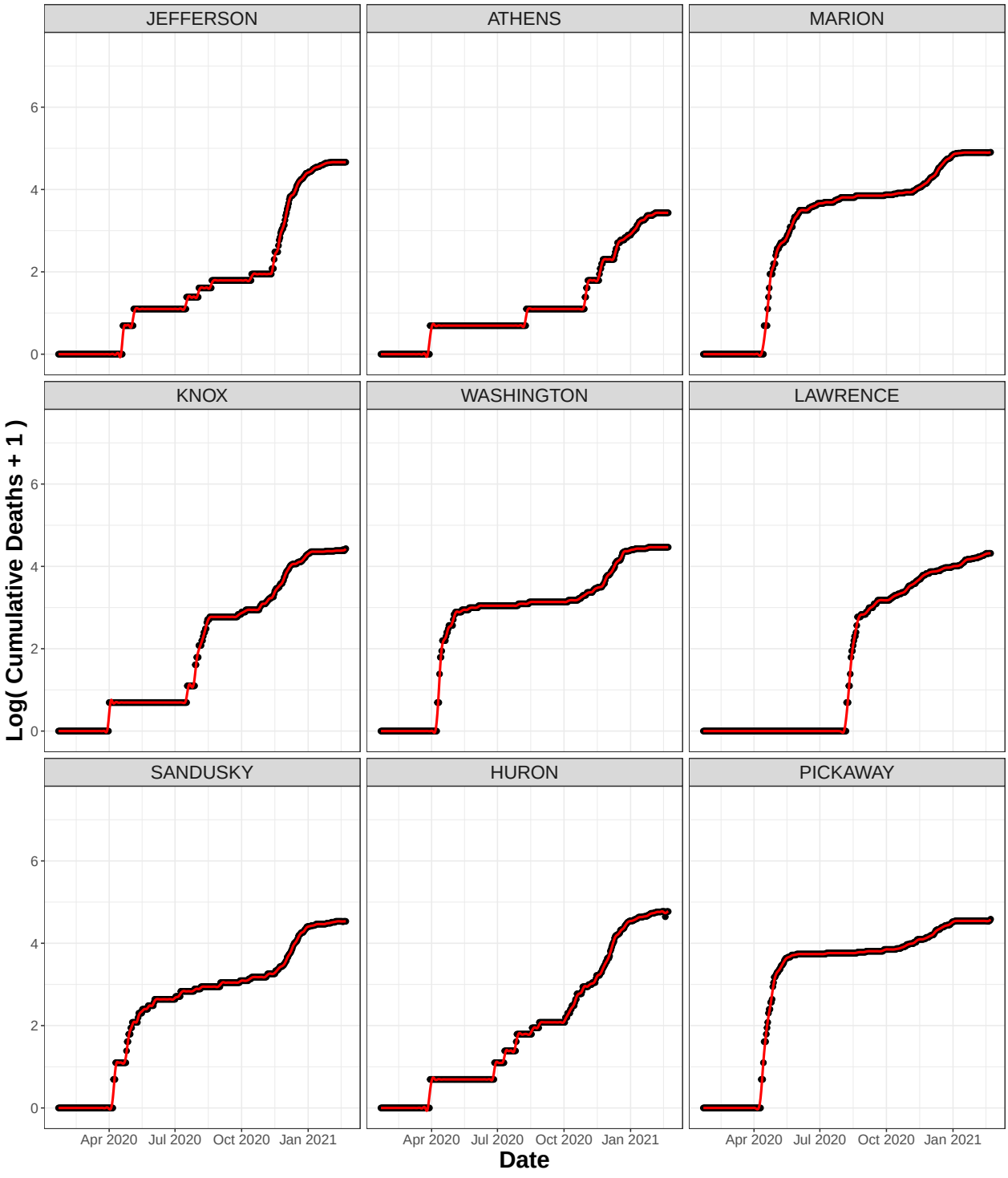
```
      theme(axis.text=element_text(size=12),
        axis.title=element_text(size=20,face="bold"),
        strip.text.x = element_text(size = 16))
  print(p)
  cat("\n\n\\newpage\n")
}
```

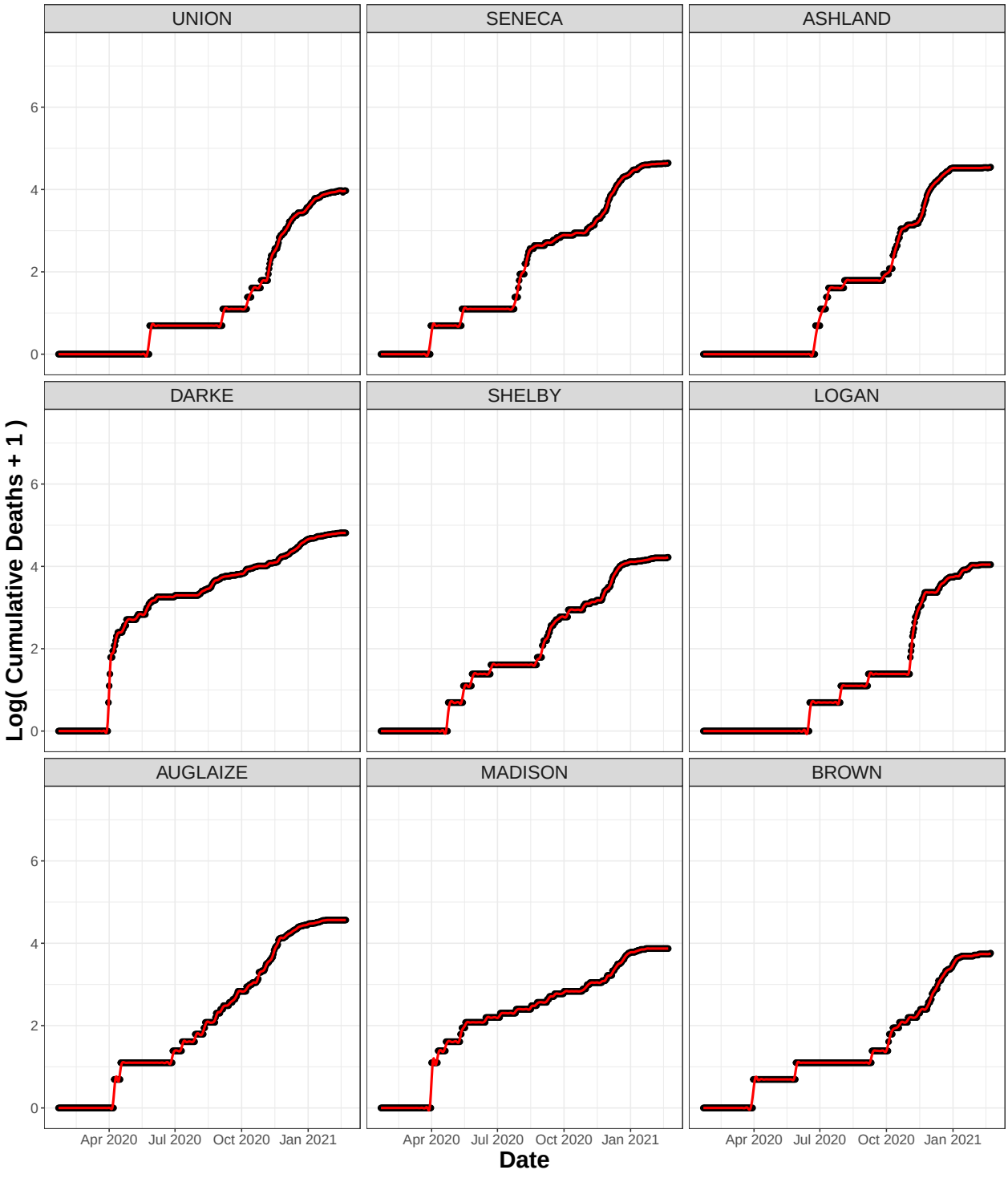


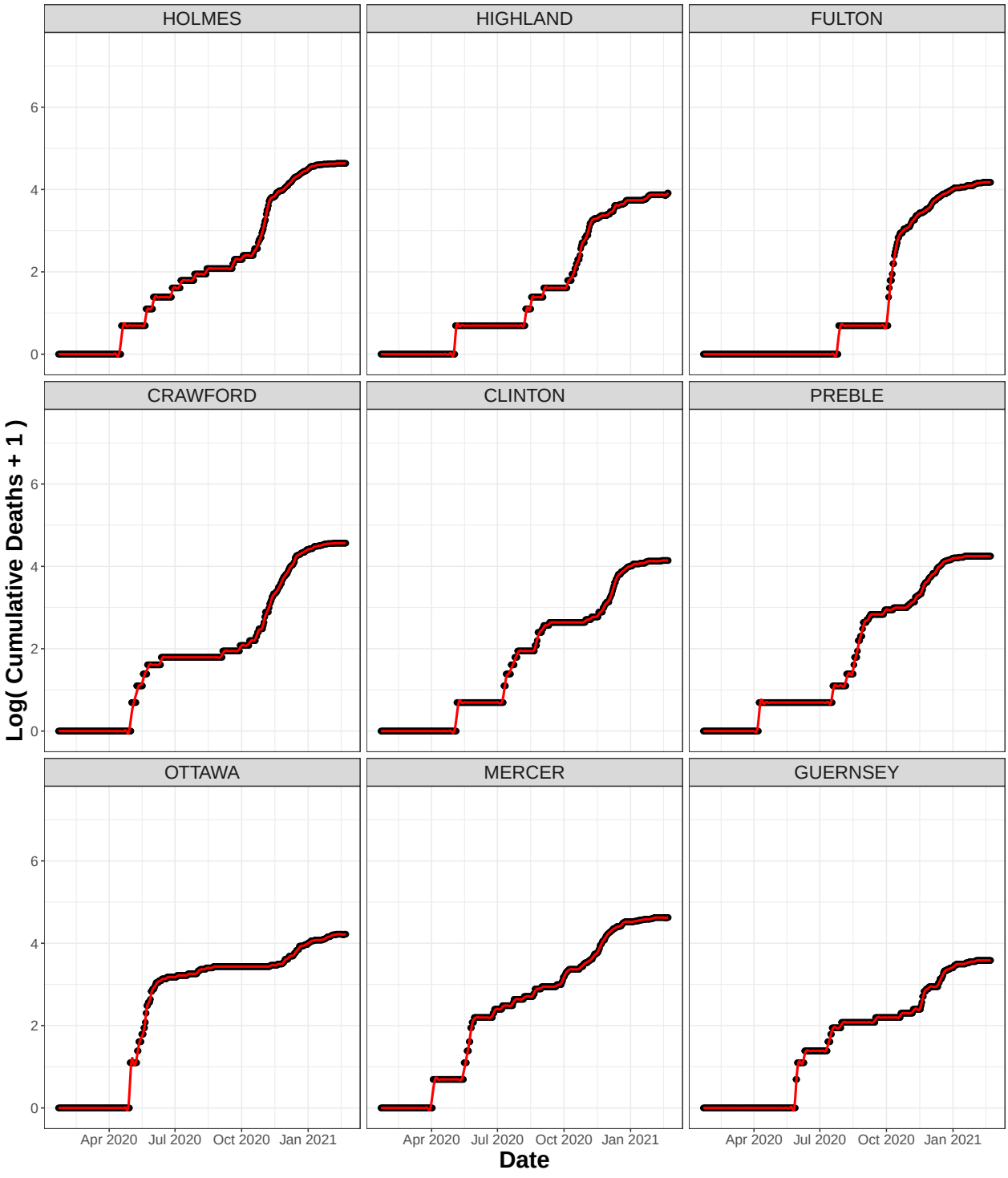

LAKE
WARREN
MAHONING
CLERMONT
TRUMBULL
DELAWARE
MEDINA
LICKING
PORTAGE
Log(Cumulative Deaths + 1)
0
2
4
6
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date

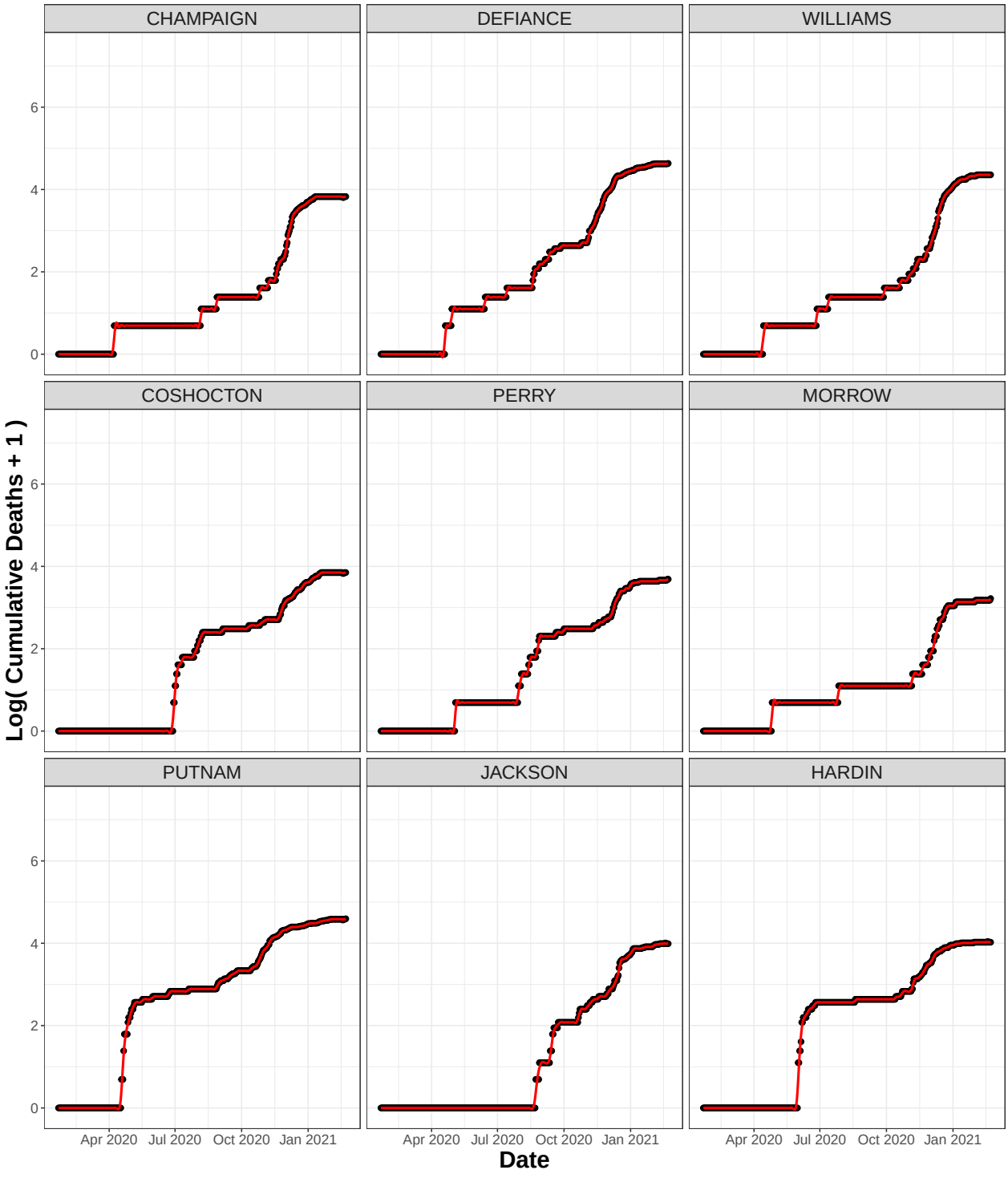

GREENE
FAIRFIELD
CLARK
WOOD
RICHLAND
WAYNE
MIAMI
ALLEN
COLUMBIANA
Log(Cumulative Deaths + 1)
0
2
4
6
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date

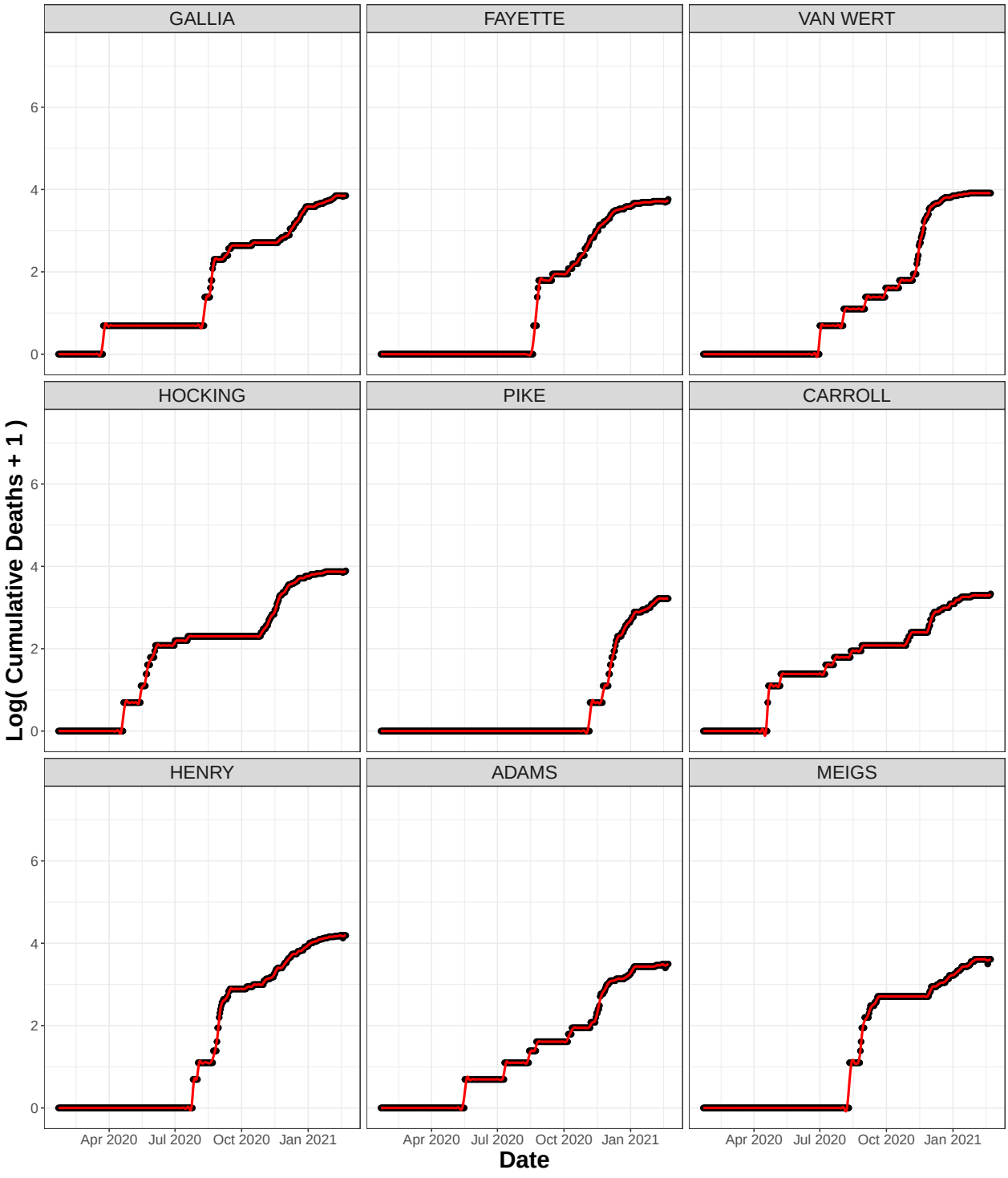

ASHTABULA
GEAUGA
TUSCARAWAS
MUSKINGUM
ROSS
SCIOTO
HANCOCK
ERIE
BELMONT
Log(Cumulative Deaths + 1)
0
2
4
6
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date

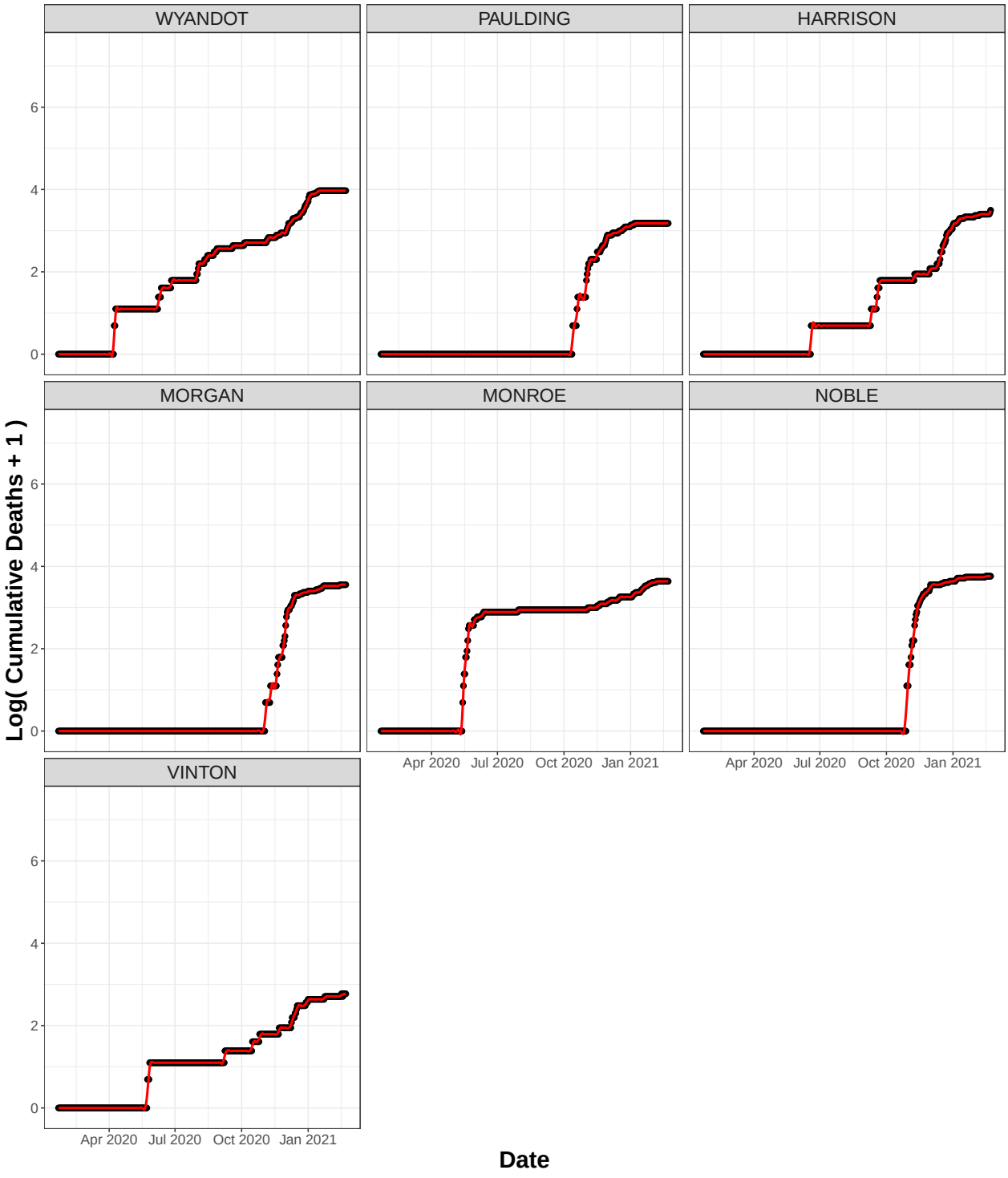

JEFFERSON
ATHENS
MARION
KNOX
WASHINGTON
LAWRENCE
SANDUSKY
HURON
PICKAWAY
Log(Cumulative Deaths + 1)
0
2
4
6
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date


UNION
SENECA
ASHLAND
DARKE
SHELBY
LOGAN
AUGLAIZE
MADISON
BROWN
Log(Cumulative Deaths + 1)
0
2
4
6
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date


HOLMES
HIGHLAND
FULTON
CRAWFORD
CLINTON
PREBLE
OTTAWA
MERCER
GUERNSEY
Log(Cumulative Deaths + 1)
0
2
4
6
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date

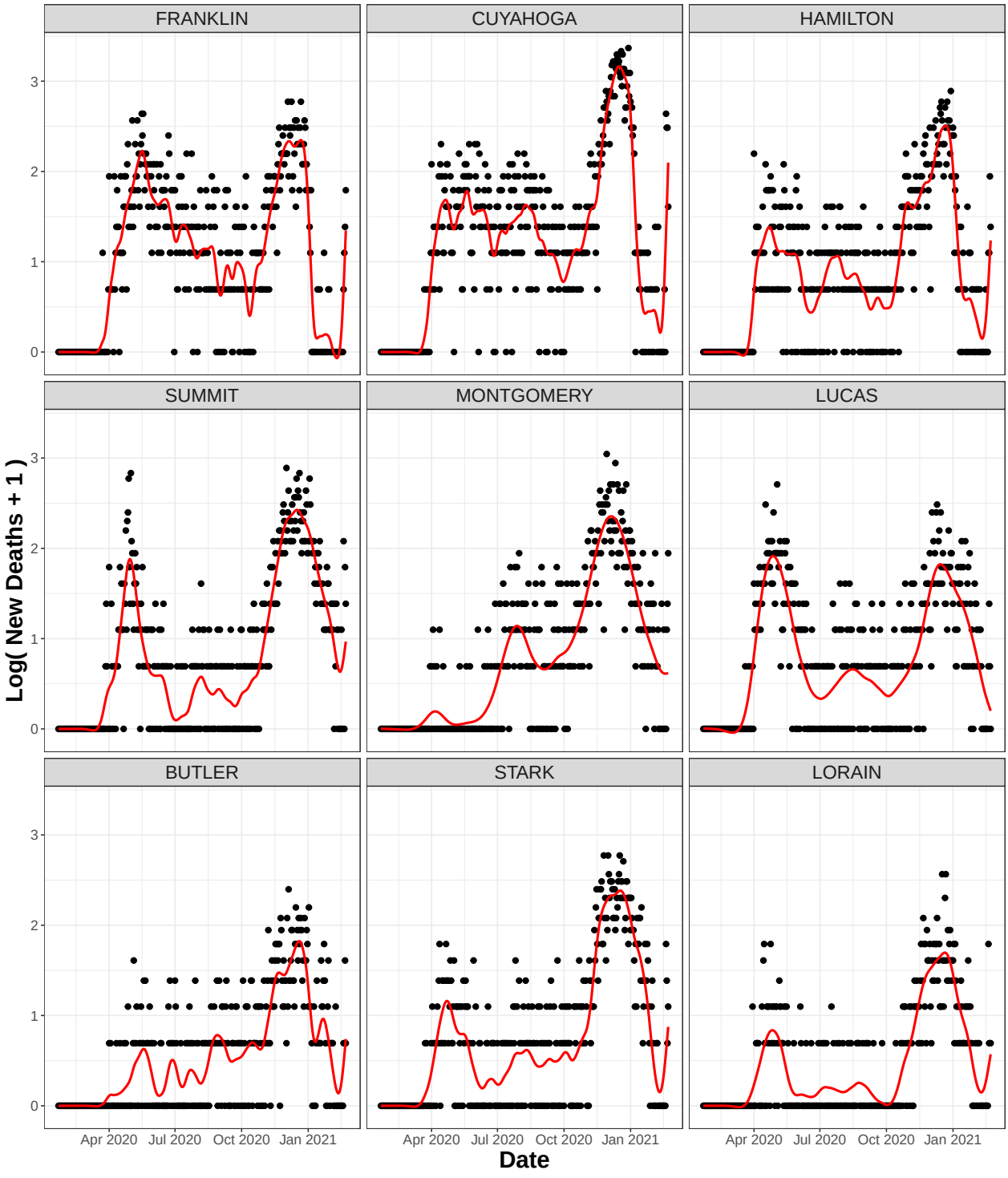

CHAMPAIGN
DEFIANCE
WILLIAMS
COSHOCTON
PERRY
MORROW
PUTNAM
JACKSON
HARDIN
Log(Cumulative Deaths + 1)
0
2
4
6
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date

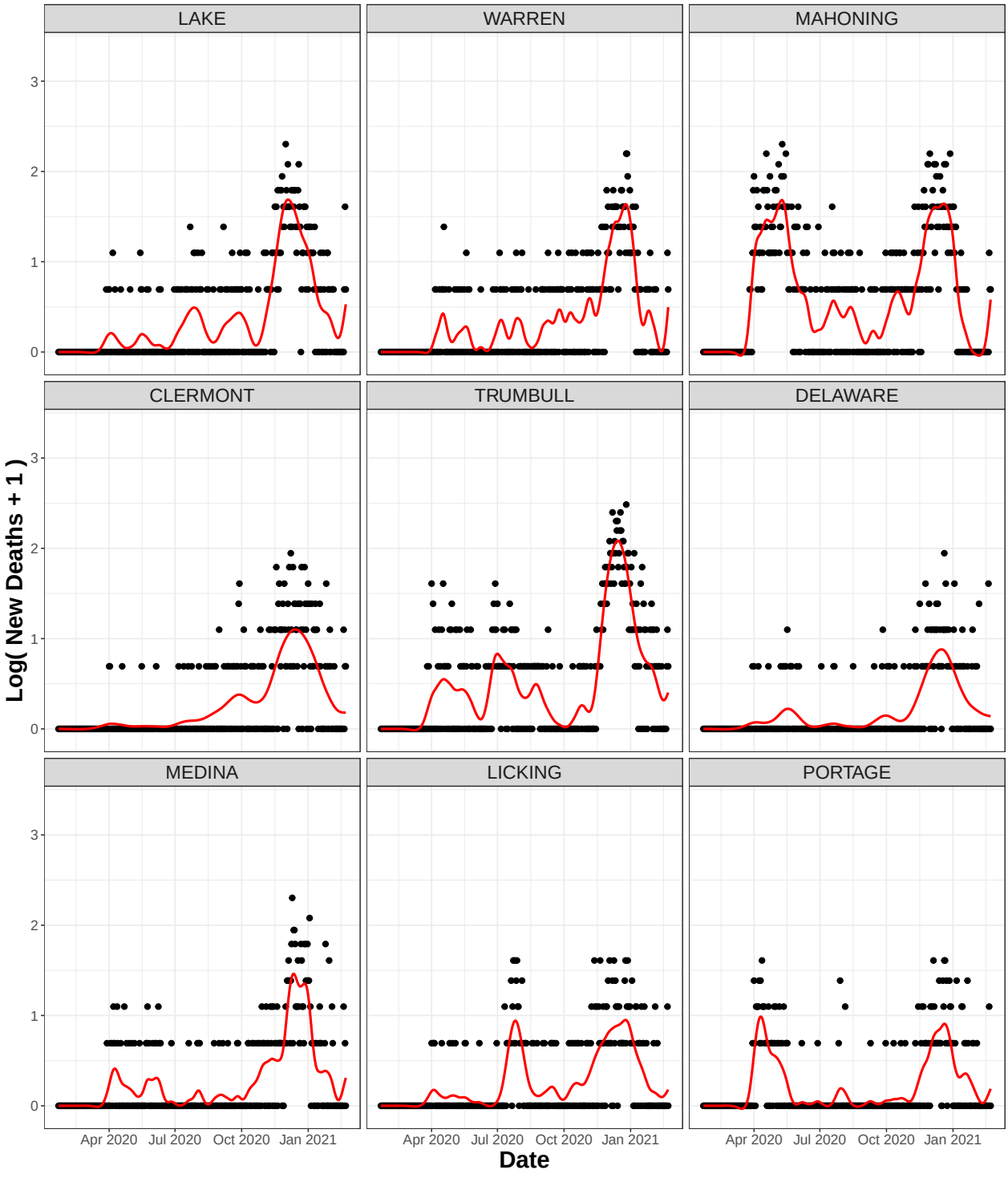

GALLIA
FAYETTE
VAN WERT
HOCKING
PIKE
CARROLL
HENRY
ADAMS
MEIGS
Log(Cumulative Deaths + 1)
0
2
4
6
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date

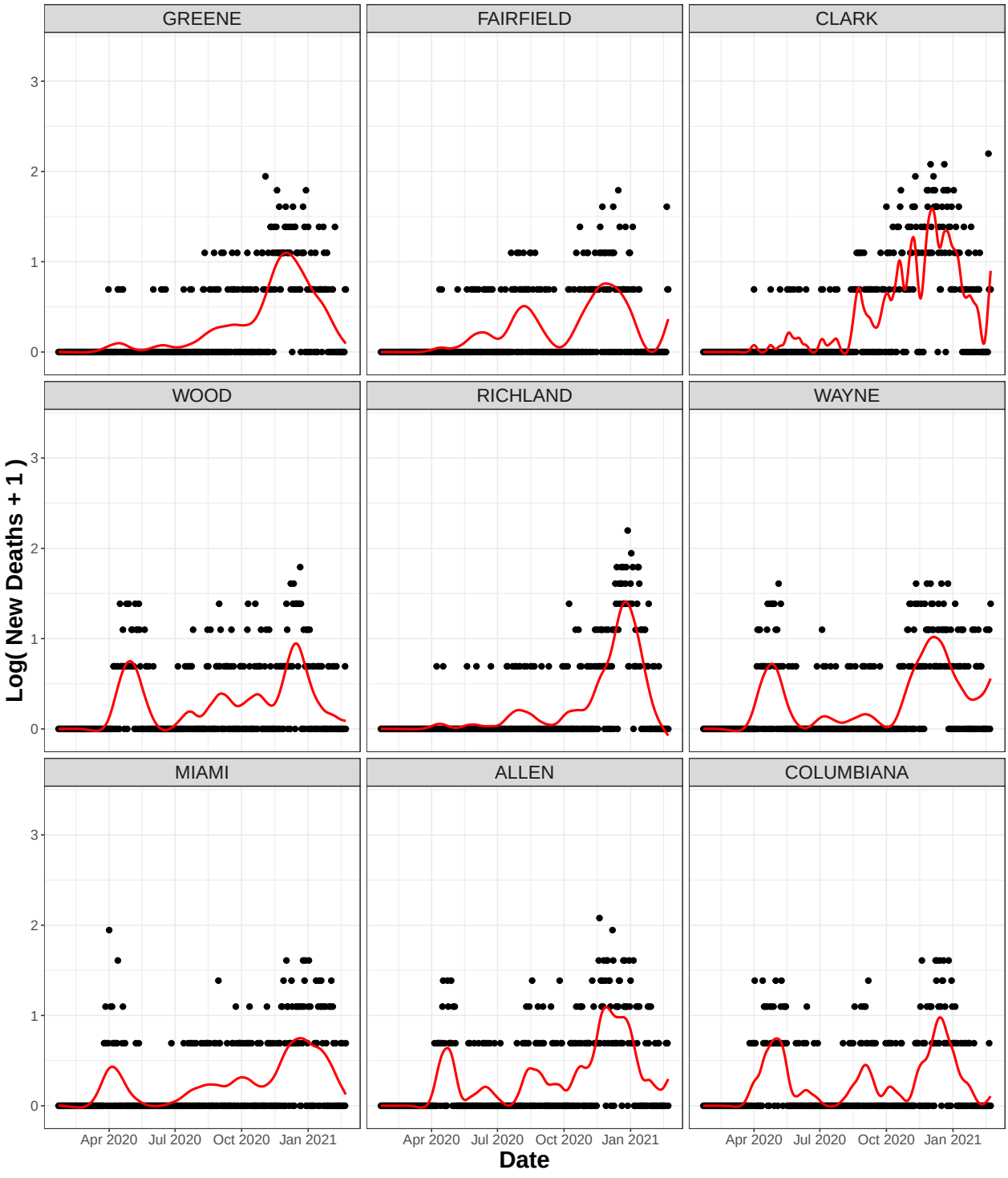

WYANDOT
PAULDING
HARRISON
MORGAN
MONROE
NOBLE
VINTON
Log(Cumulative Deaths + 1)
0
2
4
6
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date

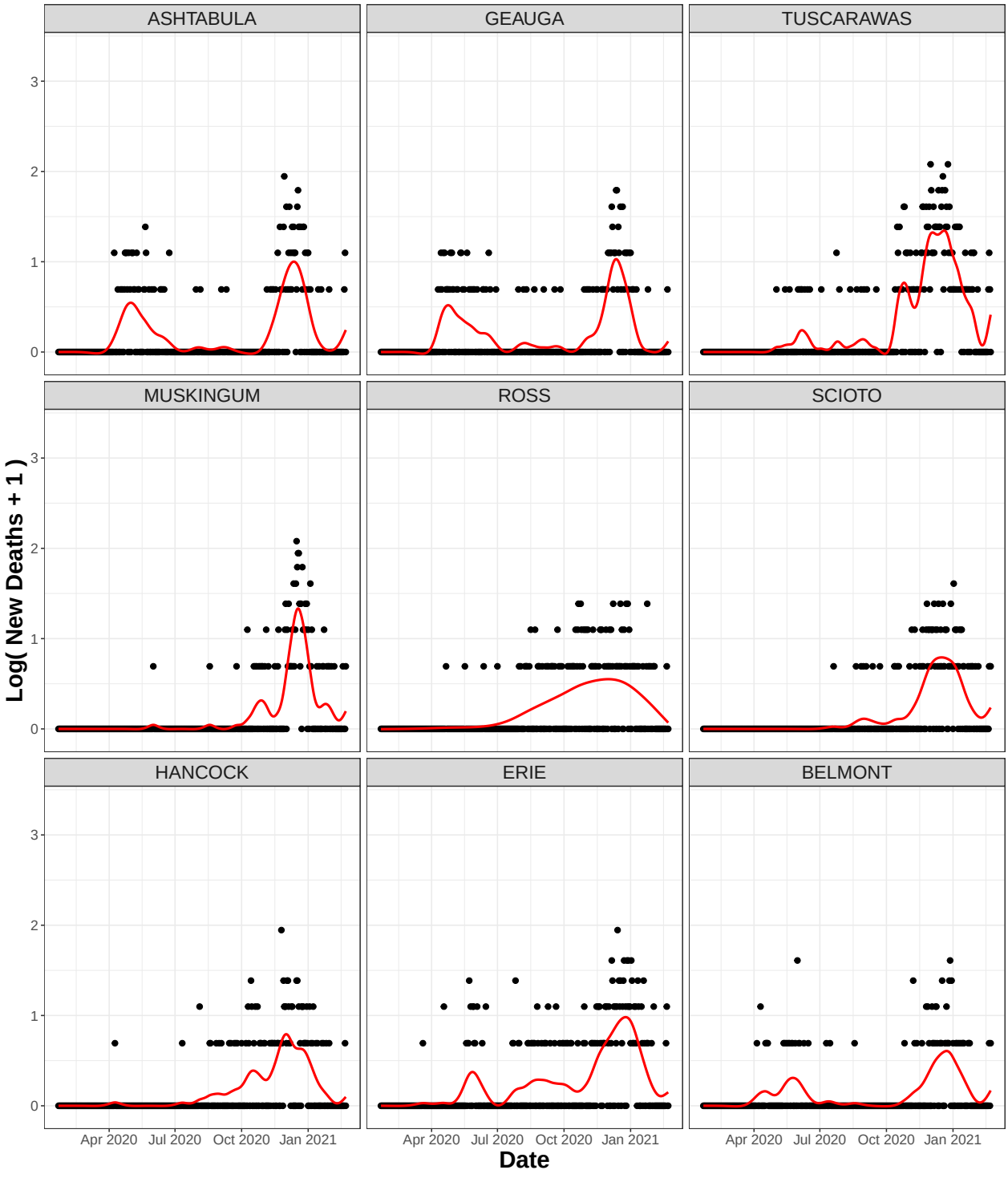
New Deaths

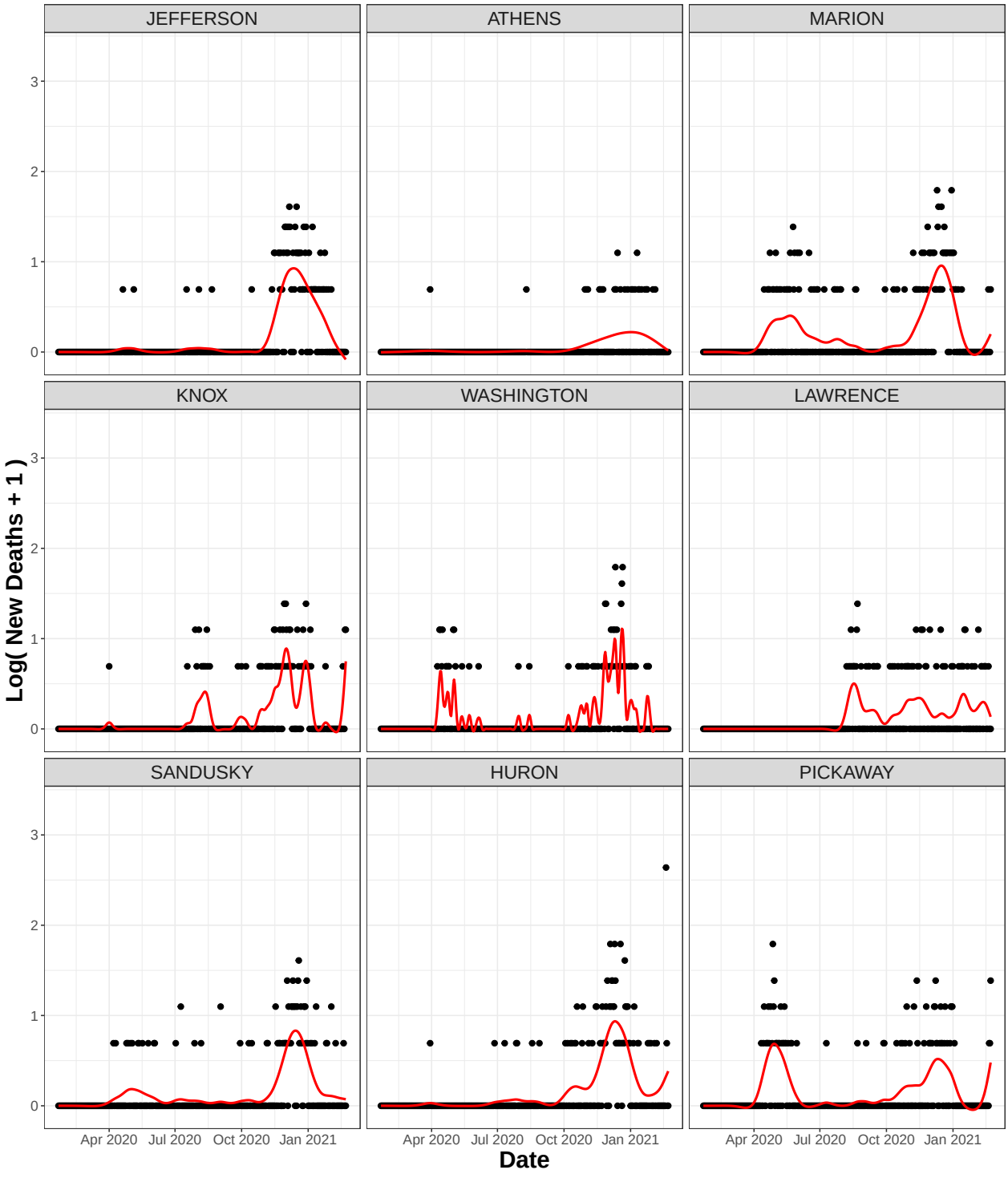
```
for(i in 1:10){
  p <- ggplot(log_deaths_county_df, aes(x = DATE, y = log_new_deaths)) +
      geom_point(size = 2) +
      geom_line(aes(x = DATE,y = new.smoothed.spline), color = "red",size = 1)+
      facet_wrap_paginate(~COUNTY, ncol = 3, nrow = 3, page = i)+
      theme_bw() +
      labs(x = "Date", y = "Log( New Deaths + 1 )")+
      theme(axis.text=element_text(size=12),
        axis.title=element_text(size=20,face="bold"),
        strip.text.x = element_text(size = 16))
  print(p)
  cat("\n\n\\newpage\n")
}
```

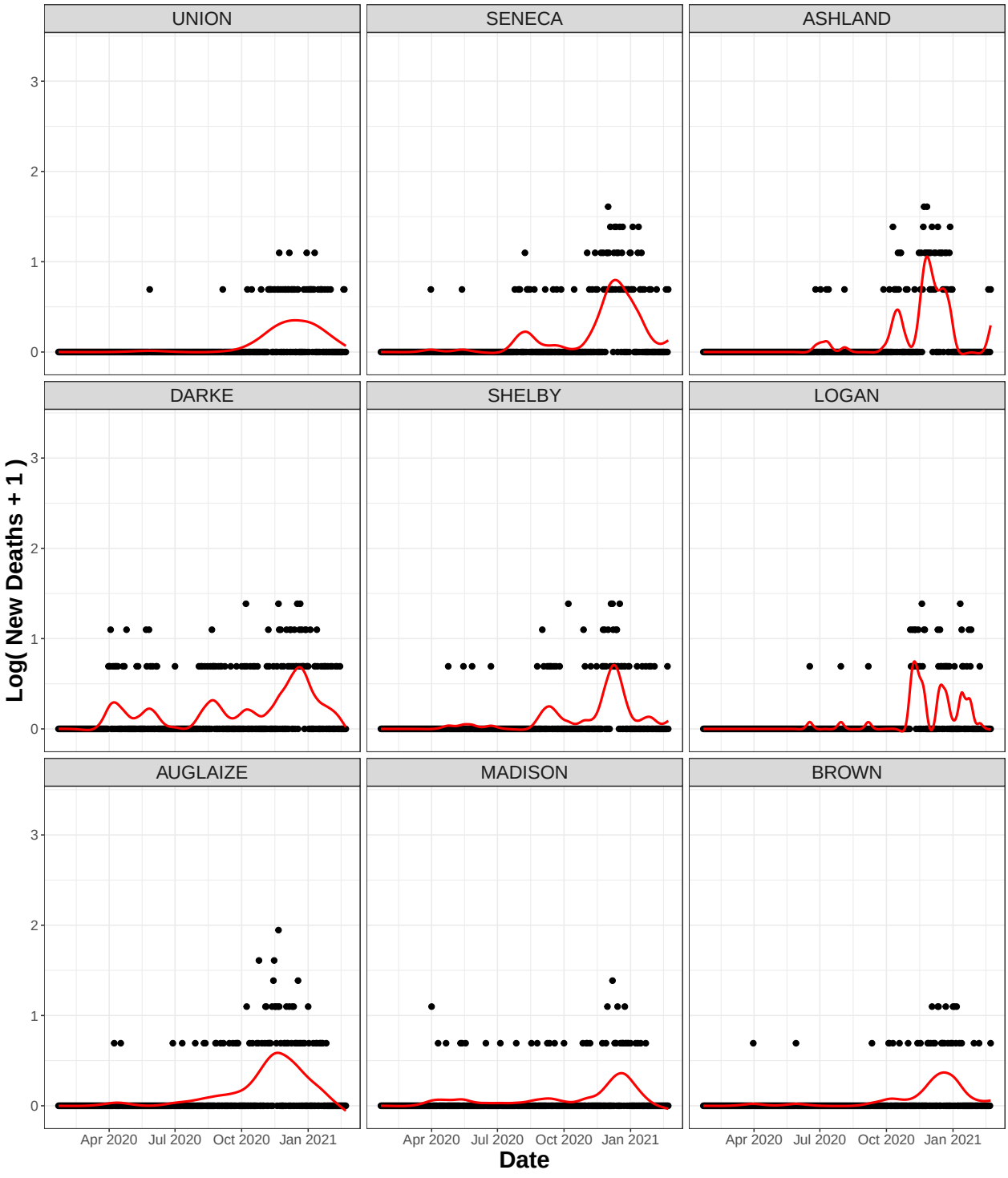

FRANKLIN
CUYAHOGA
HAMILTON
SUMMIT
MONTGOMERY
LUCAS
BUTLER
STARK
LORAIN
Log(New Deaths + 1)
0
1
2
3
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date

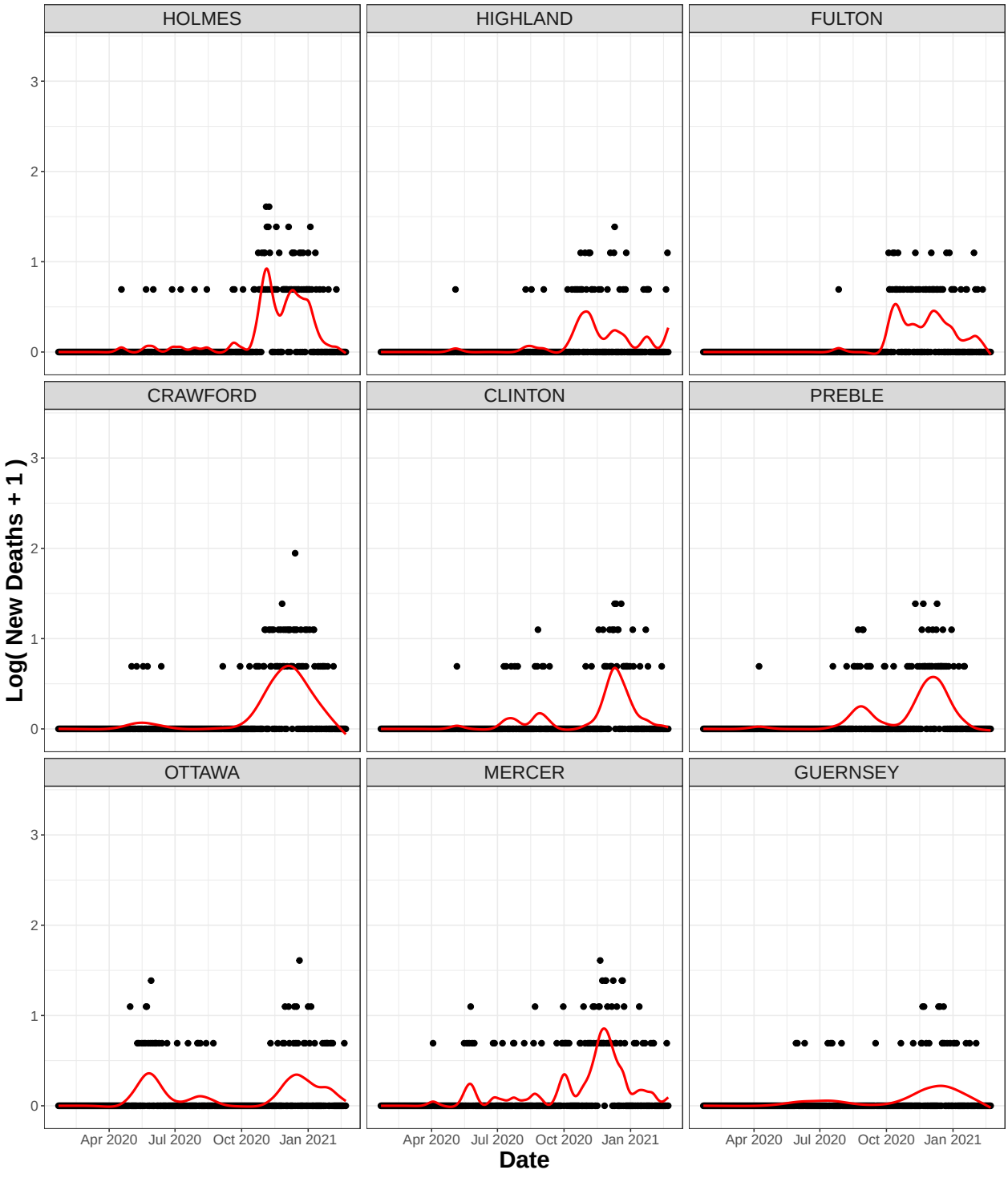

LAKE
WARREN
MAHONING
CLERMONT
TRUMBULL
DELAWARE
MEDINA
LICKING
PORTAGE
Log(New Deaths + 1)
0
1
2
3
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date

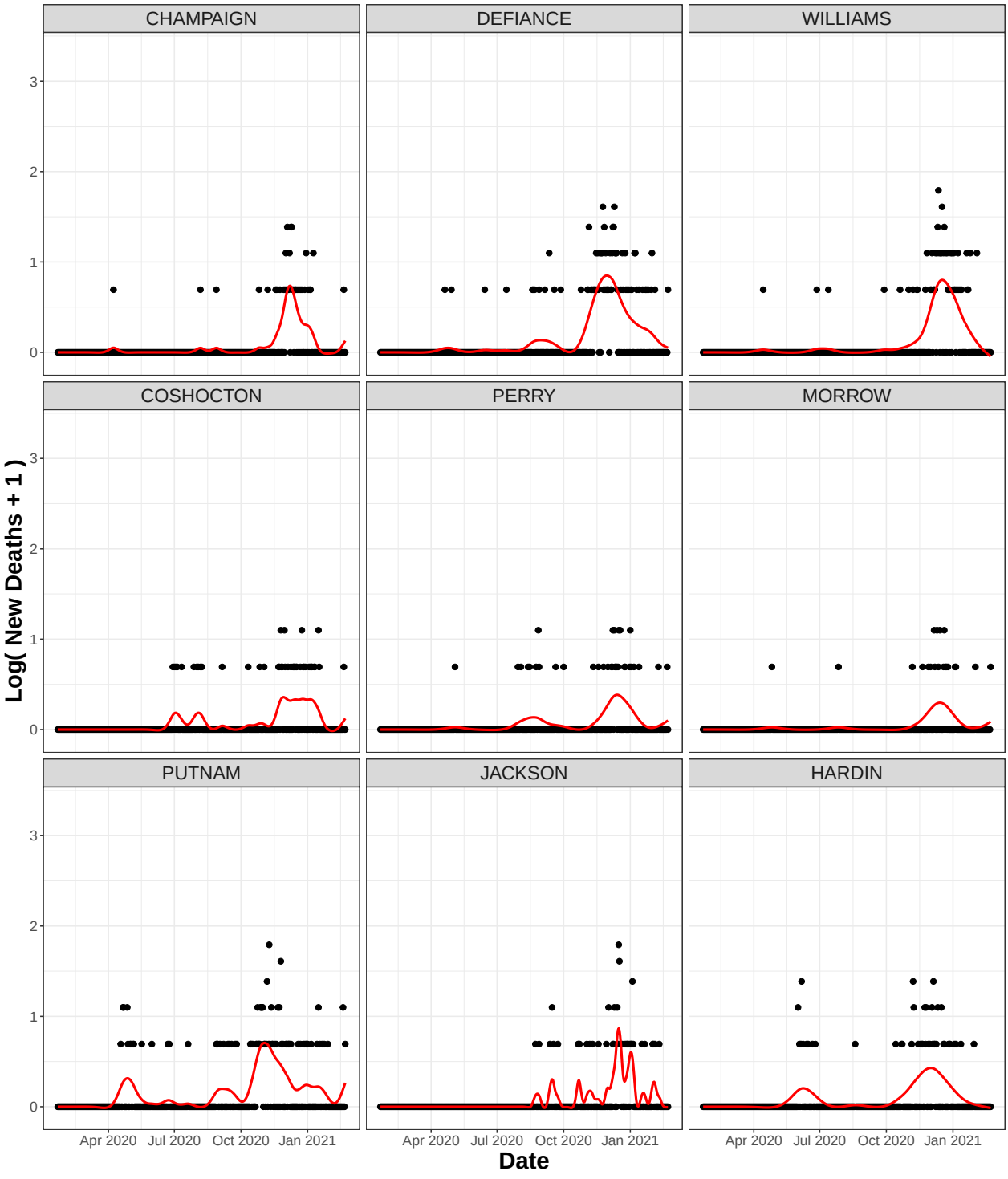

GREENE
FAIRFIELD
CLARK
WOOD
RICHLAND
WAYNE
MIAMI
ALLEN
COLUMBIANA
Log(New Deaths + 1)
0
1
2
3
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date

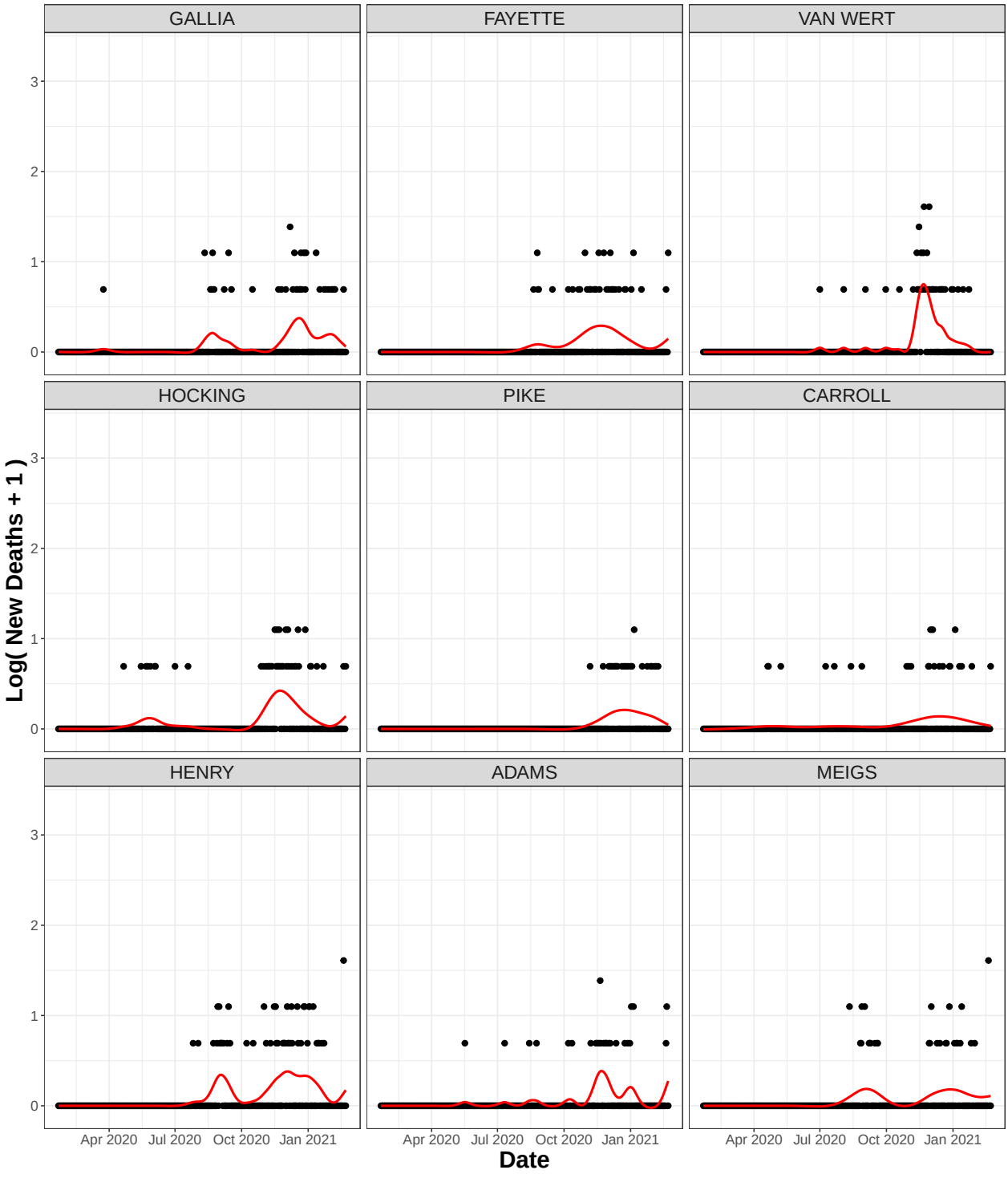

ASHTABULA
GEAUGA
TUSCARAWAS
MUSKINGUM
ROSS
SCIOTO
HANCOCK
ERIE
BELMONT
Log(New Deaths + 1)
0
1
2
3
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date

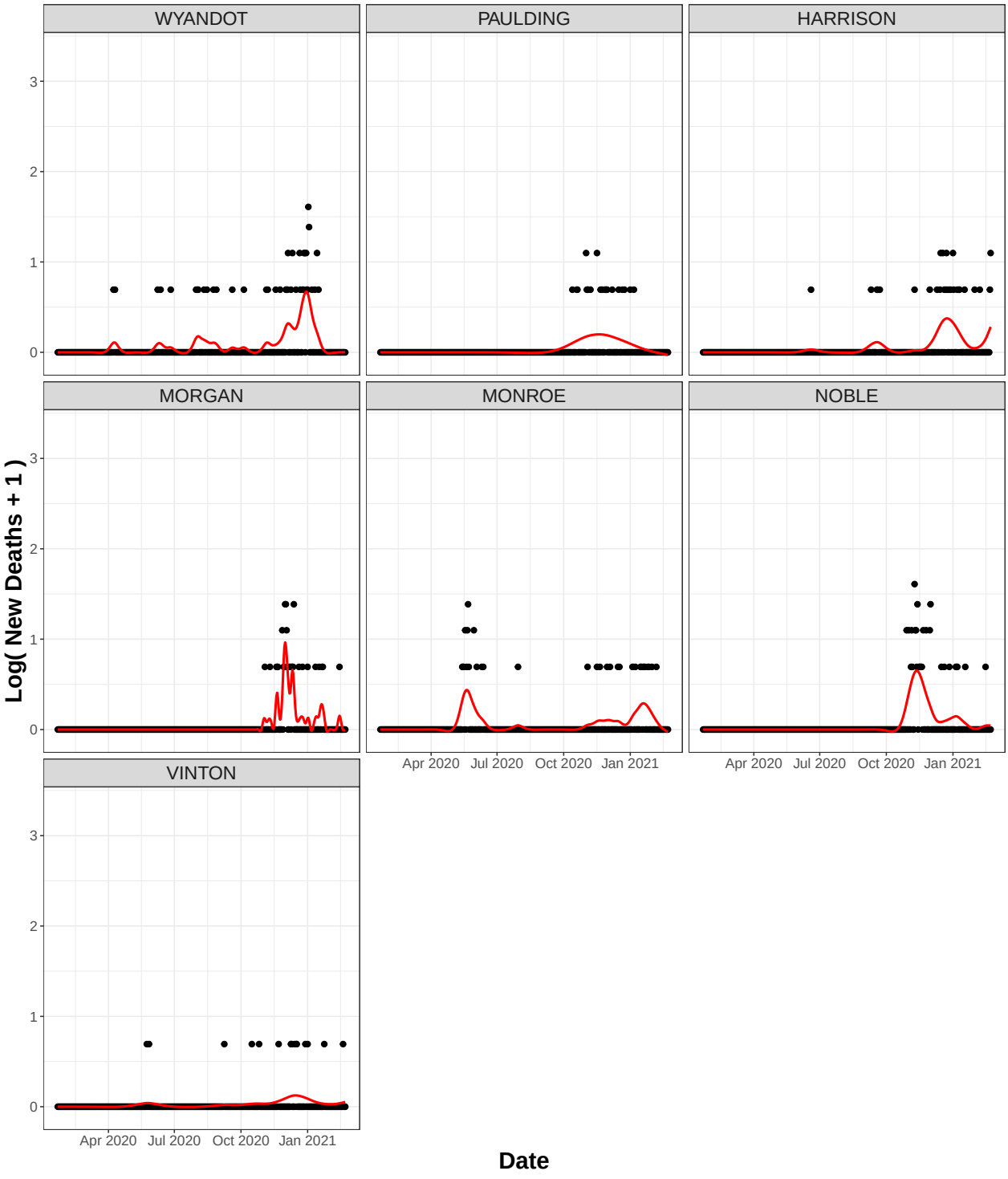

JEFFERSON
ATHENS
MARION
KNOX
WASHINGTON
LAWRENCE
SANDUSKY
HURON
PICKAWAY
Log(New Deaths + 1)
0
1
2
3
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date


UNION
SENECA
ASHLAND
DARKE
SHELBY
LOGAN
AUGLAIZE
MADISON
BROWN
Log(New Deaths + 1)
0
1
2
3
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date


HOLMES
HIGHLAND
FULTON
CRAWFORD
CLINTON
PREBLE
OTTAWA
MERCER
GUERNSEY
Log(New Deaths + 1)
0
1
2
3
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date


CHAMPAIGN
DEFIANCE
WILLIAMS
COSHOCTON
PERRY
MORROW
PUTNAM
JACKSON
HARDIN
Log(New Deaths + 1)
0
1
2
3
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date


GALLIA
FAYETTE
VAN WERT
HOCKING
PIKE
CARROLL
HENRY
ADAMS
MEIGS
Log(New Deaths + 1)
0
1
2
3
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date


WYANDOT
PAULDING
HARRISON
MORGAN
MONROE
NOBLE
VINTON
Log(New Deaths + 1)
0
1
2
3
Apr 2020
Jul 2020
Oct 2020
Jan 2021
Date