



Carnegie Mellon University

PHIGHT COVID

Statistical Practice

Capstone Project

April 27, 2022

*Olivia Wang, Daniel Xie, Julia Keating, Wei-Yu Tseng,
Daniel Nason*

COVID and Teaching Posture

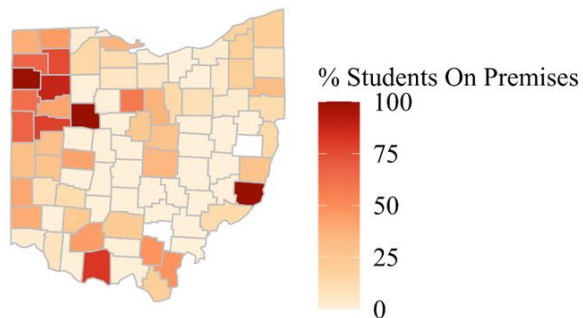
- **Background:** Children are known to act as a vector for influenza virus transmission. Maybe this is also the case for COVID.
- **Goal:** Understand effect of teaching posture on COVID-19 transmission by looking at K-12 public school data in Ohio.

Why Ohio?

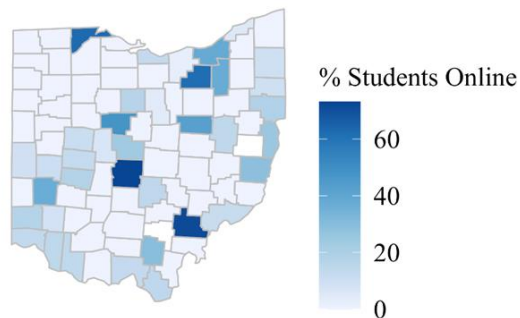
- Small size with uniform geography and climate
- Statewide implementation of interventions
- Multiple school districts
- School districts had diverse teaching posture

Teaching Posture %

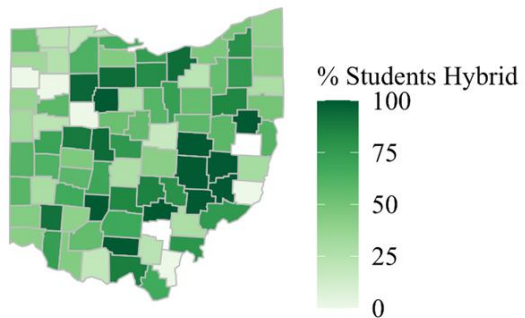
A.



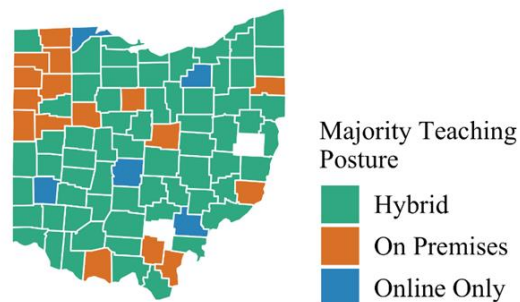
B.



C.

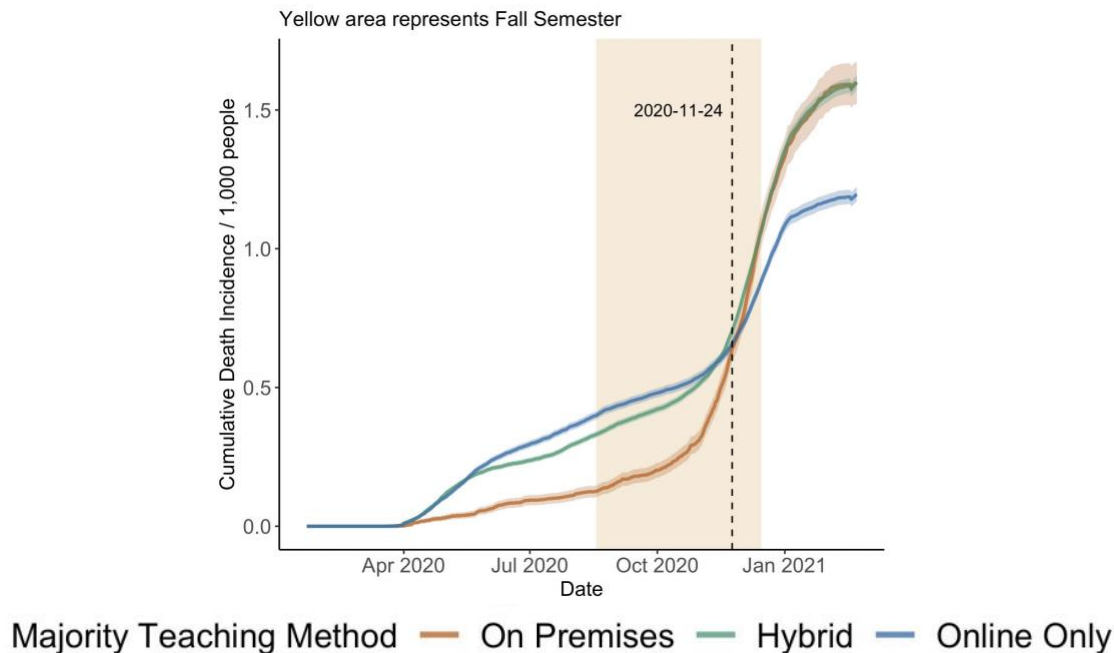


D.



Code cited from Ehman et al. (2021) - School Teaching Posture Correlates with COVID-19 Disease Outcomes in Ohio.

On-premises teaching posture seems associated with increased COVID-19 deaths at the end of the semester



Plot cited from Ehman et al. (2021) - School Teaching Posture Correlates with COVID-19 Disease Outcomes in Ohio.

What is R_t ?

- R_t = average number of secondary cases of disease caused by a single infected individual at time t
- Instantaneous measure of disease spread commonly used to characterize pathogen transmissibility during an epidemic
- Can be compared across counties and teaching postures

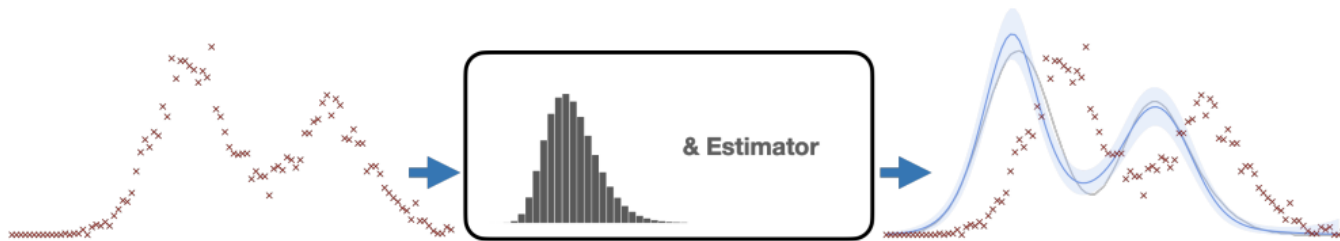
R_t can be estimated if we know the number of new infections I_t for all times t

- We don't observe infections
- Need to infer infections from the events we can observe (e.g. cases or deaths)

I_t , and thus R_t , can be estimated from the time series of deaths if we know the distribution of time from infection to death



(a) Assumed data generating process.



(b) Estimation procedure

Miller et al. (2020)

How we calculated R_t

Using the EpiEstim package in R, with inputs:

1. an incidence time series (e.g. cases or deaths)
1. the corresponding delay distribution of time from infection to incidence event

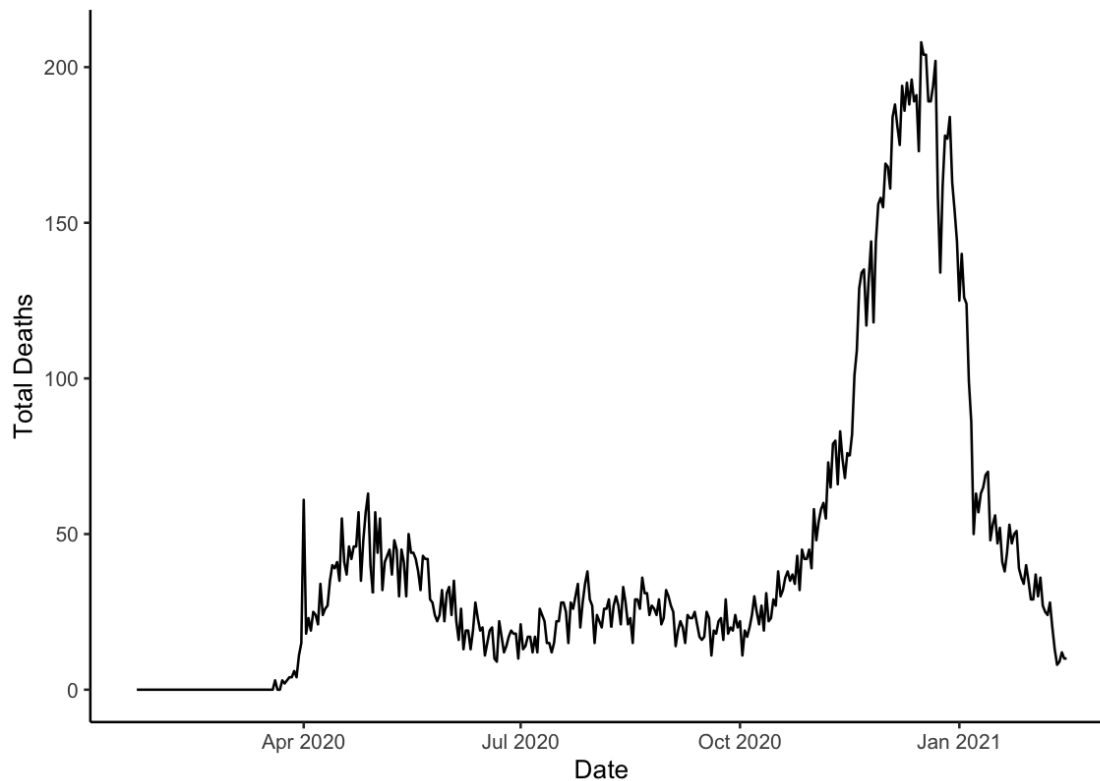
Cases v.s. Deaths

- Benefits of using deaths
 - Deaths reported can be easily validated
- Benefits of using cases
 - Cases happen sooner after infection than deaths
 - Cases has larger sample size, the variance of calculated R_t is smaller

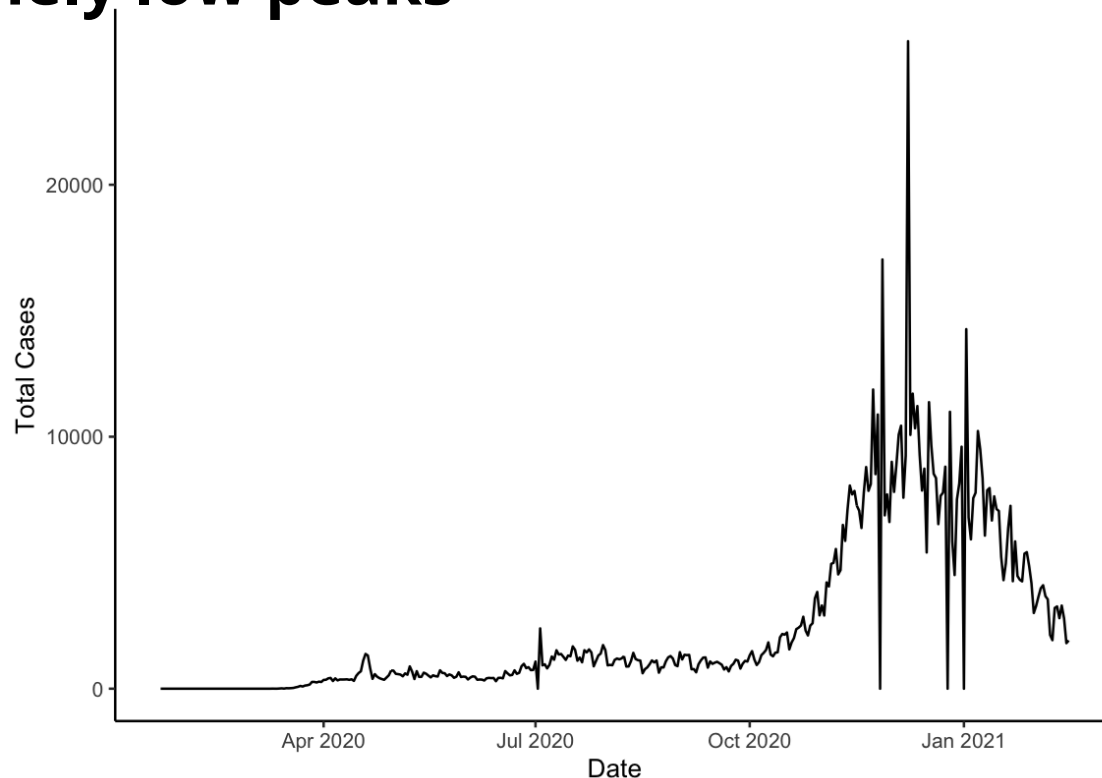
Delay distributions

- Time from infection to death follows a Gamma distribution with:
 - Mean = 23.9 days, coefficient of variation = 0.40
 - Unwin et al. (2020)
- Time from infection to case follows a Gamma distribution with:
 - Mean = 15.2 days, coefficient of variation = 0.60
 - Guessed from a related distribution

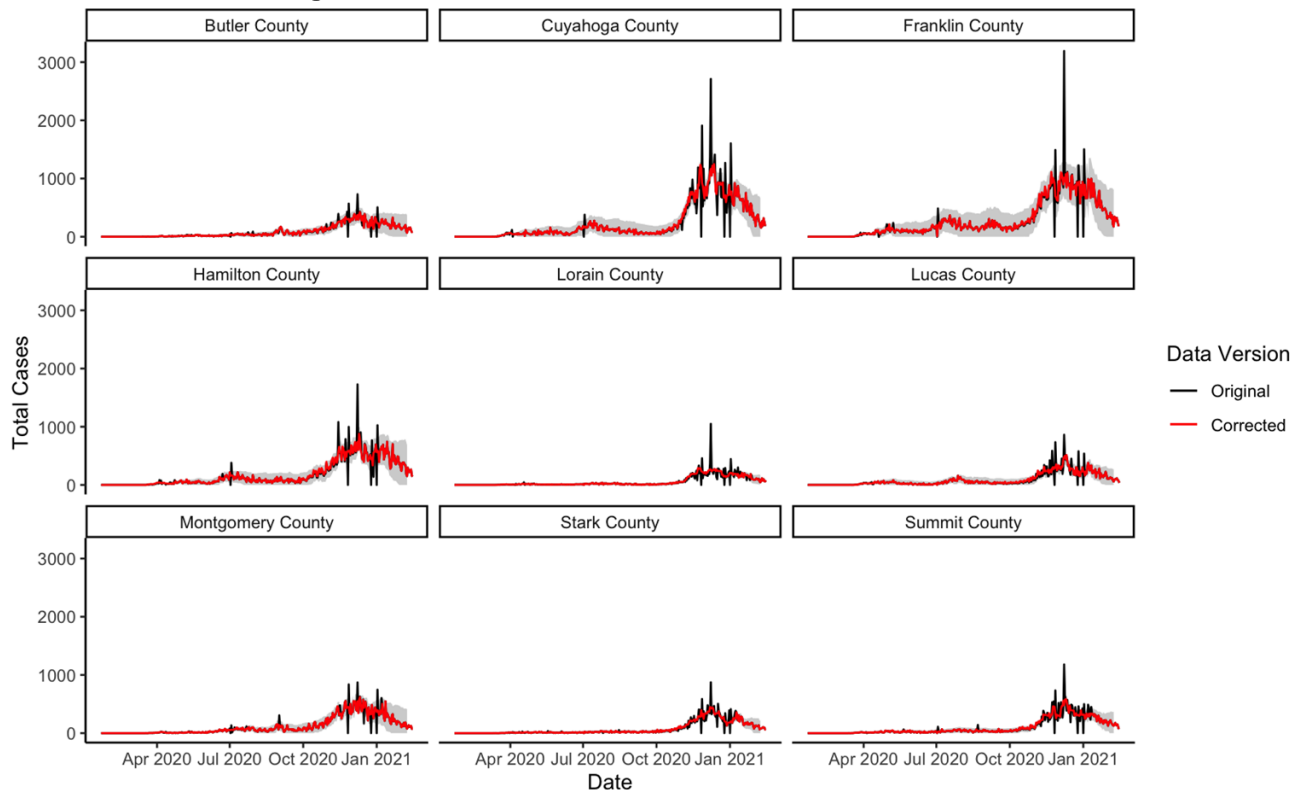
Deaths over time at state level



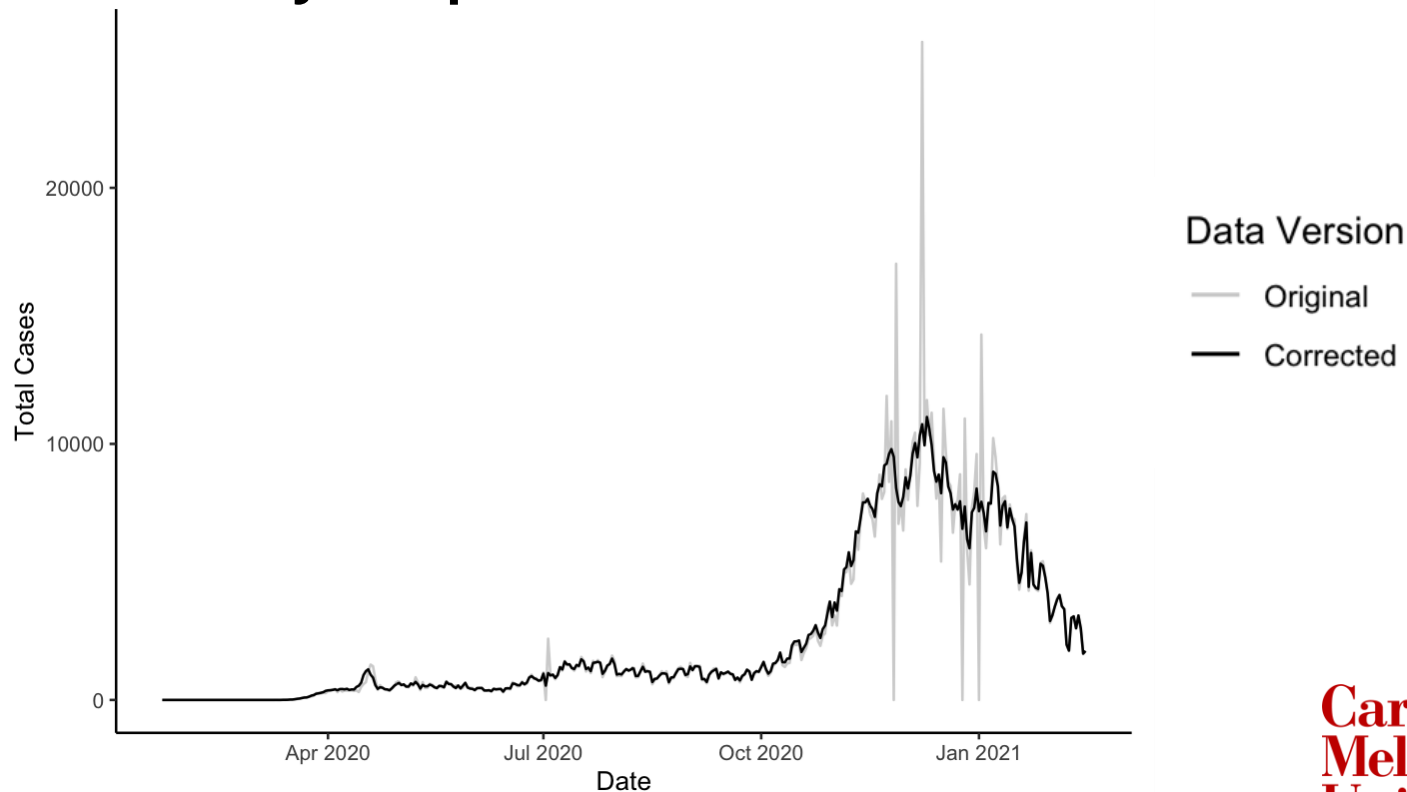
The cases time series has extremely high and extremely low peaks



Pre-processing smooths the extremely high and extremely low peaks at the county level



Pre-processing the cases time series smooths the extremely high and extremely low peaks



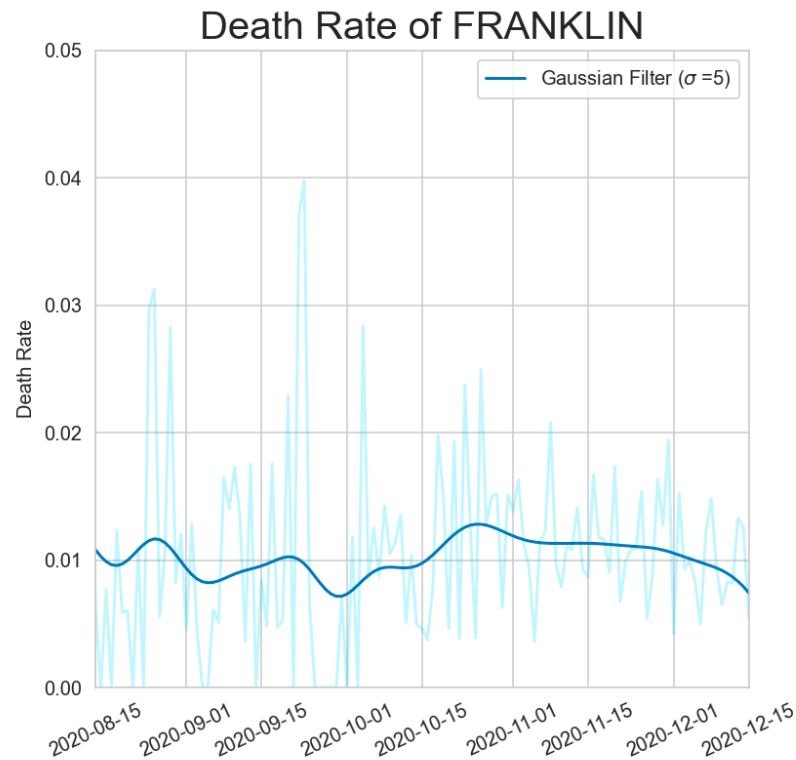
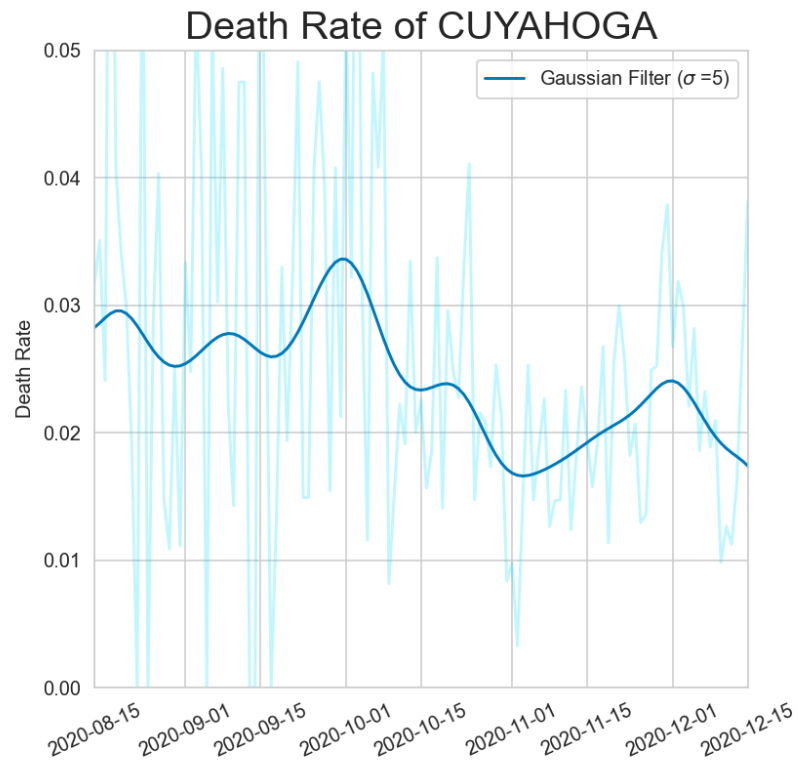
Cases v.s. Deaths

- Benefits of using Deaths
 - Deaths reported can be easily validated
- Benefits of using Cases
 - Cases happen sooner than deaths after infection
 - Cases has larger sample size, the variance of calculated R_t is smaller
- Epidemiologist prefer to use cases
- But can we trust cases?

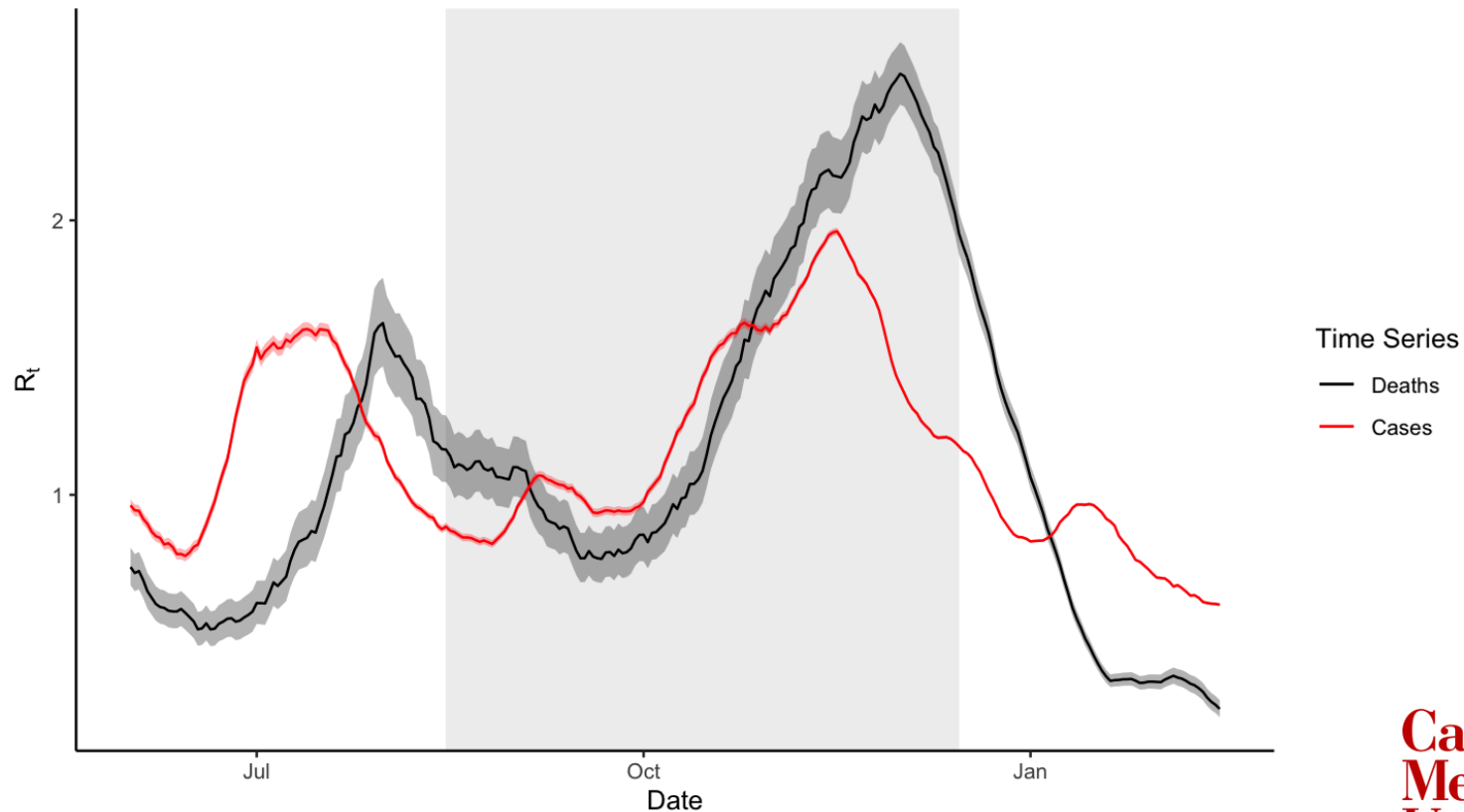
Limitation of cases calculated R_t

- Testing capacity maybe lacking, resulting in a low number of cases
 - If not, Death rate(deaths - cases ratio) should be approximately constant over time, and similar across counties with similar demographics.
 - R_t calculated from cases should be similar as R_t from deaths.

Death Rate should be approx constant over time



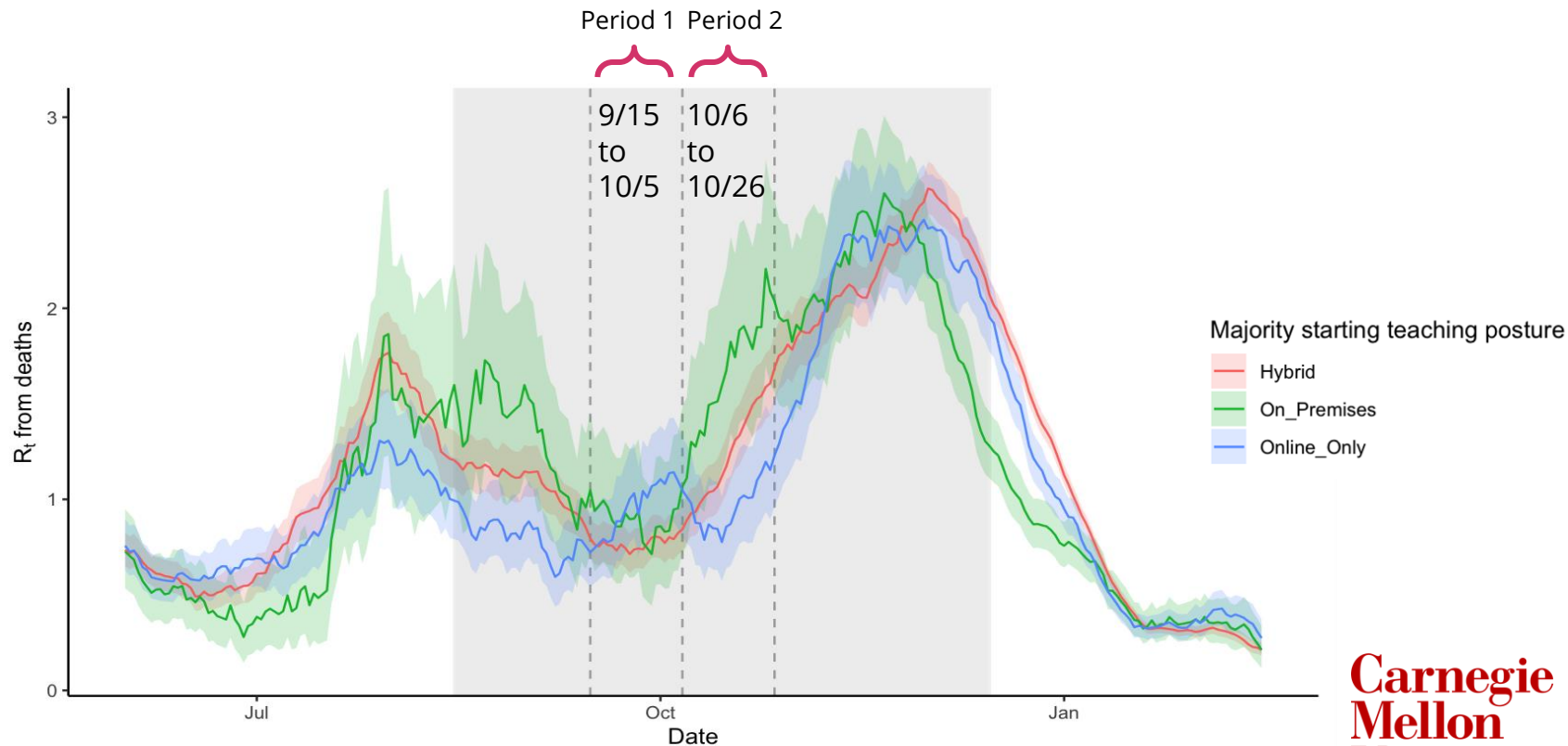
Calculated R_t 's should overlap



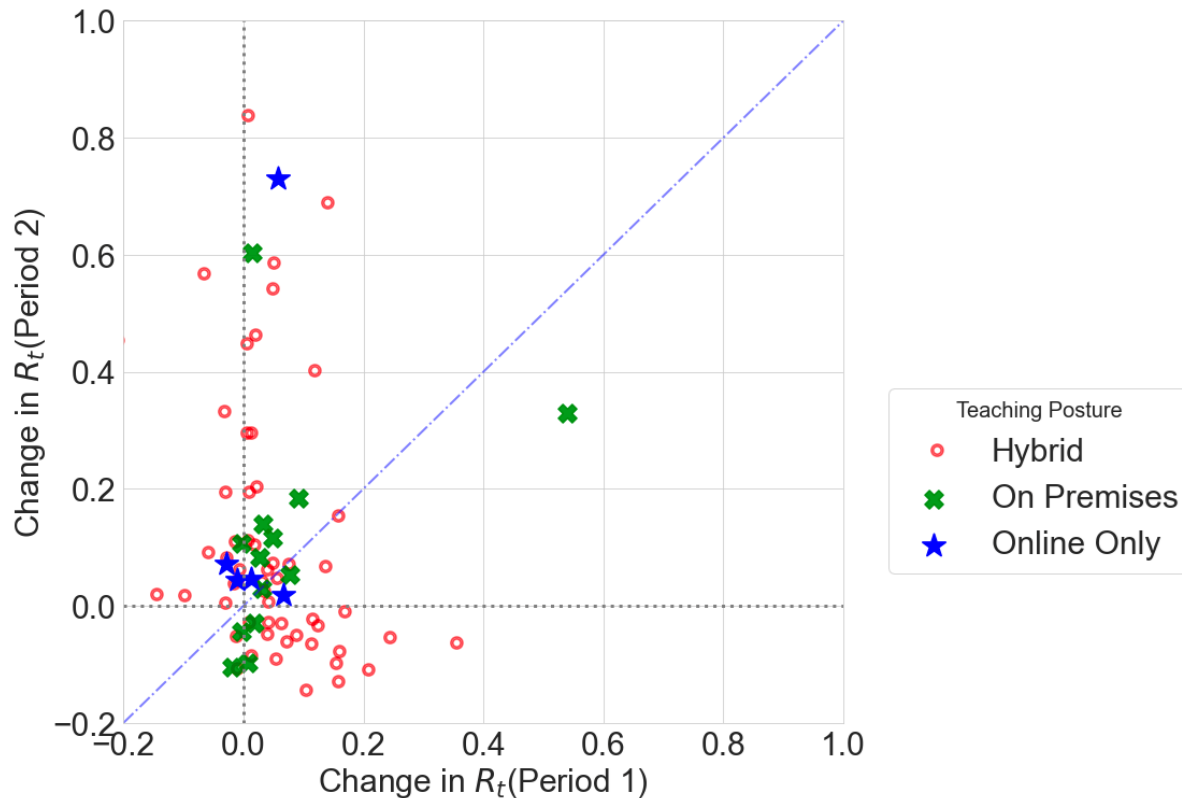
Recap on R_t

- We don't trust number of new cases
 - The numbers weren't correctly reported
- We trust the number of new deaths
 - Deaths reported can be easily validated
- In the rest of the talk, we use R_t calculated from Deaths

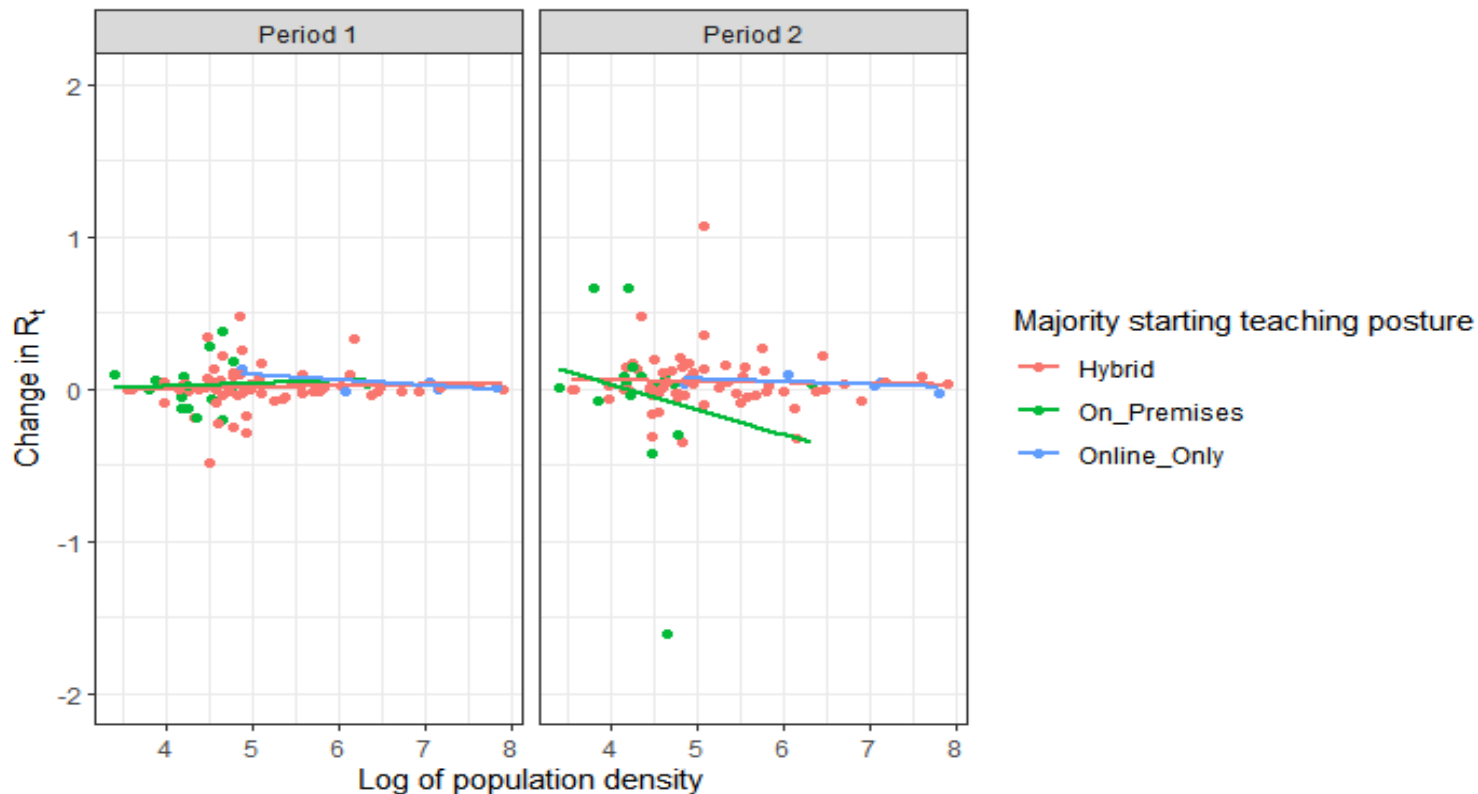
On premise counties have a relatively higher R_t at the start of the semester



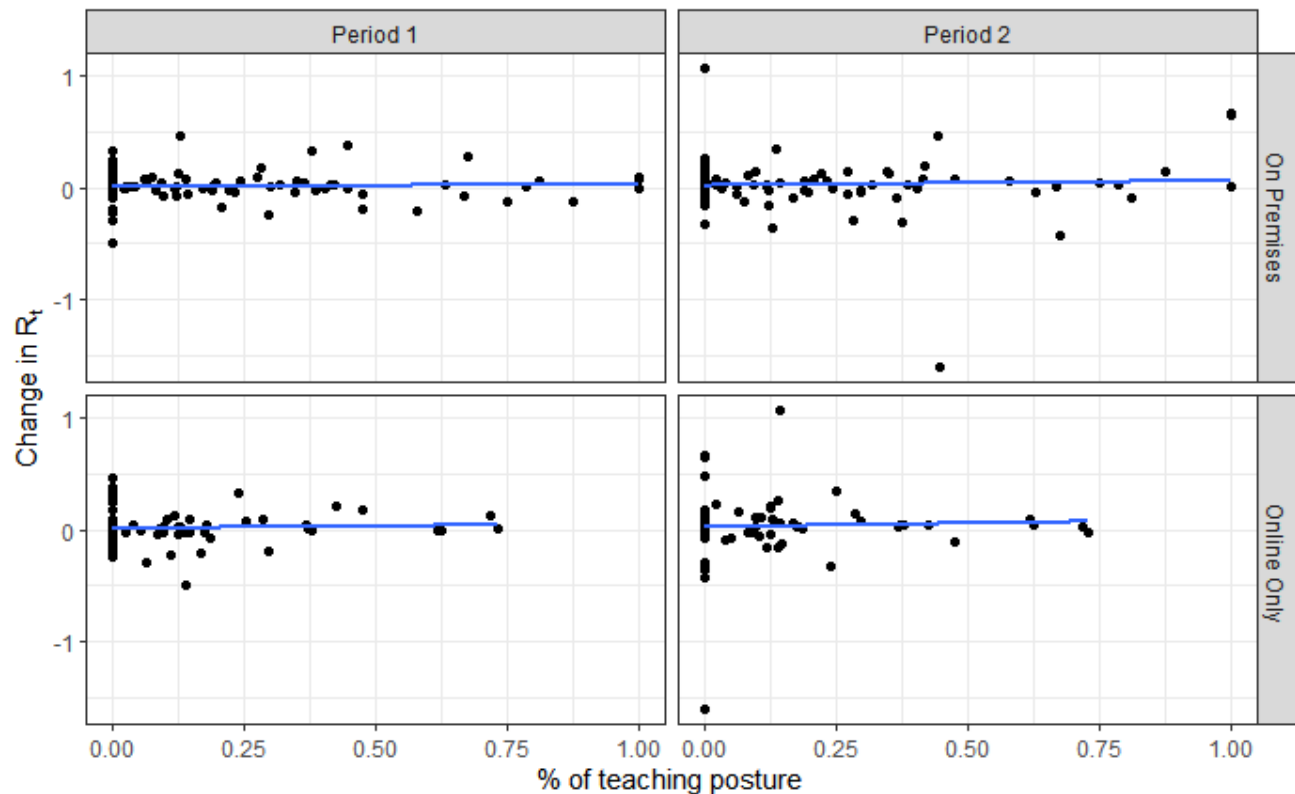
No clear evidence of the change in R_t for counties with an on-premises majority posture compared to other postures at the start of the school year



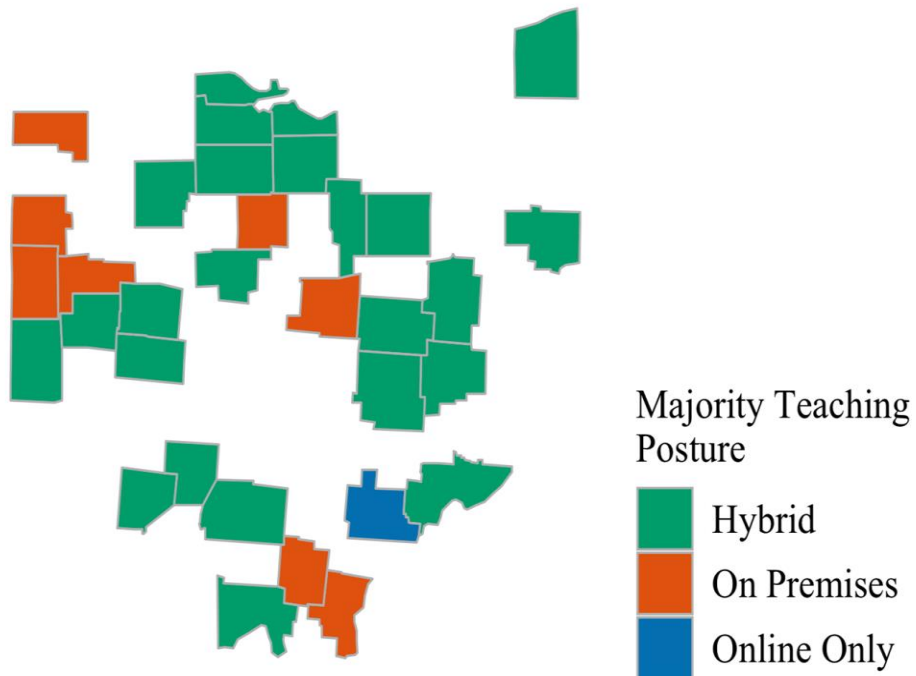
The change in R_t does not increase for majority in-person teaching posture counties at the beginning of the school year accounting for population density



Counties with more students on premises at the start of the school year do not have corresponding increases in change in R_t

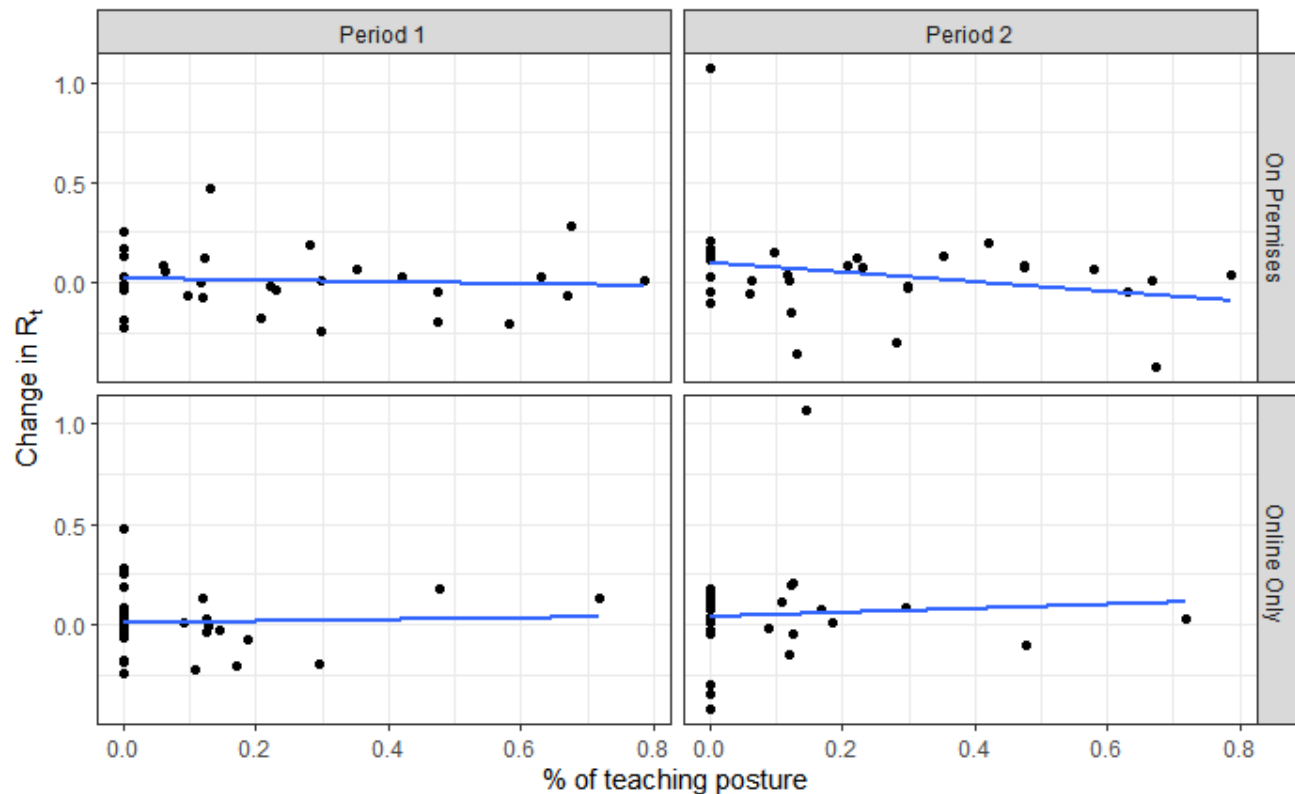


Overview of Micropolitan Counties



- Urban areas with population between 10,000-50,000
- Benefit: comparison across similar counties
- Starting teaching posture for majority of students:
 - Online Only: 1
 - Hybrid: 24
 - On Premises: 8

In Micropolitan counties, there is no increase in the change in R_t when students are on premises compared to being online



Conclusions

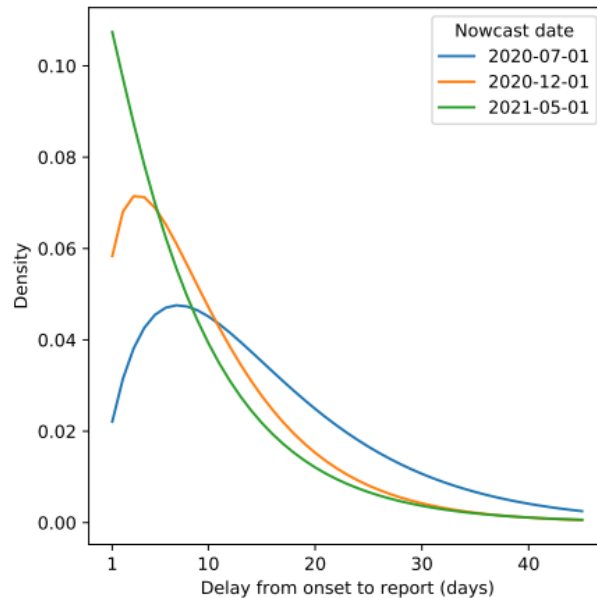
- R_t calculated from deaths is more reliable than that from cases
 - R_t from deaths and cases do not overlap well
 - Non-constant death rate illustrates issues with cases time series
 - Reported deaths are easier to validate
- No significant evidence shows the teaching posture is related to R_t
 - Change in R_t stays relatively constant at the start of the school year
 - This is also the case for micropolitan counties
- Results suggest that it safe to send kids back to in-person teaching

Limitations

- Case: Testing capacity is inconsistent across counties
 - Counties with similar demographics do not share similar death rate pattern
 - Death Rate is not constant over time in many counties
- Death: There are fewer observations
 - Higher variance on calculated R_t
- Delay from infection to report: We assumed constant delay distribution
 - Recent paper shows that the distribution varies over time

Next Steps

- Utilize different parameters for the delay distribution of cases over time
 - Jahja et al. (2022), Figure 3
- Include teaching posture changes by county as a time series and see how R_t varies based on teaching posture
- Extend the analysis to other states, such as Nebraska
- Regression analysis to estimate relationships between posture and R_t





Carnegie Mellon University

Thank you!



Carnegie Mellon University

Technical Appendix

References

Ehman, C., Luo, Y., Yang, Z., Zhu, Z., Donovan, S., Avery, A. J., ... & Lakdawala, S. S. (2021). K-12 School Teaching Posture Correlates with COVID-19 Disease Outcomes in Ohio. *medRxiv*.

Miller, A. C., Hannah, L., Futoma, J., Foti, N. J., Fox, E. B., D'Amour, A., Sandler, M., Saurous, R. A., & Lewnard, J. A. (2020). Statistical deconvolution for inference of Infection Time Series. <https://doi.org/10.1101/2020.10.16.20212753>

Unwin, H. J. T., Mishra, S., Bradley, V. C., Gandy, A., Mellan, T. A., Coupland, H., Ish-Horowicz, J., Vollmer, M. A., Whittaker, C., Filippi, S. L. et al. (2020). State-level tracking of COVID-19 in the United States. *Nature communications* 11 1–9.

Maria Jahja, Andrew Chin, and Ryan J. Tibshirani. (2022). Real-Time Estimation of COVID-19 Infections: Deconvolution and Sensor Fusion. <https://arxiv.org/pdf/2112.06697.pdf>

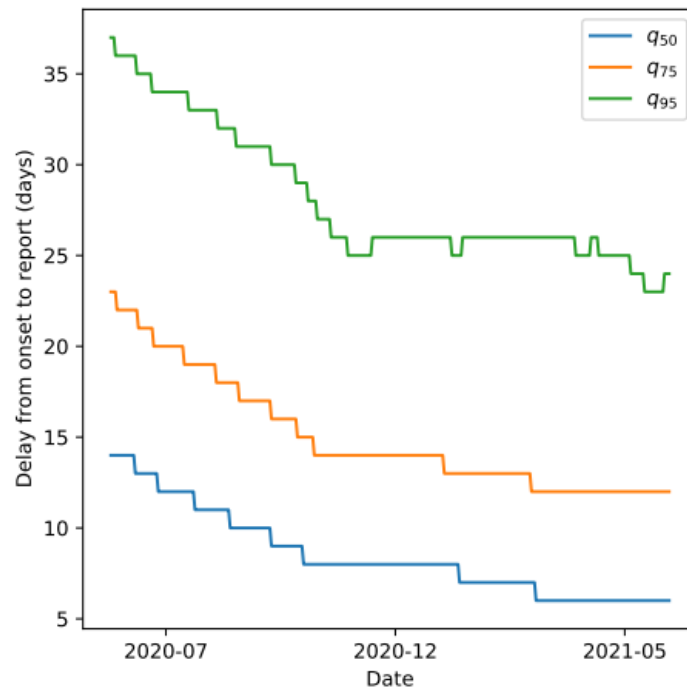


Carnegie Mellon University

Delay Distribution Parameters

Changing parameters for Delay Distribution

- Figure 3 of Jahja et al. (2022) shows that the delay distribution parameters change over time
 - Include (or approximate) these changing parameters when calculating the estimates for R_t
- Could be used to improve reliability of cases time series for analysis

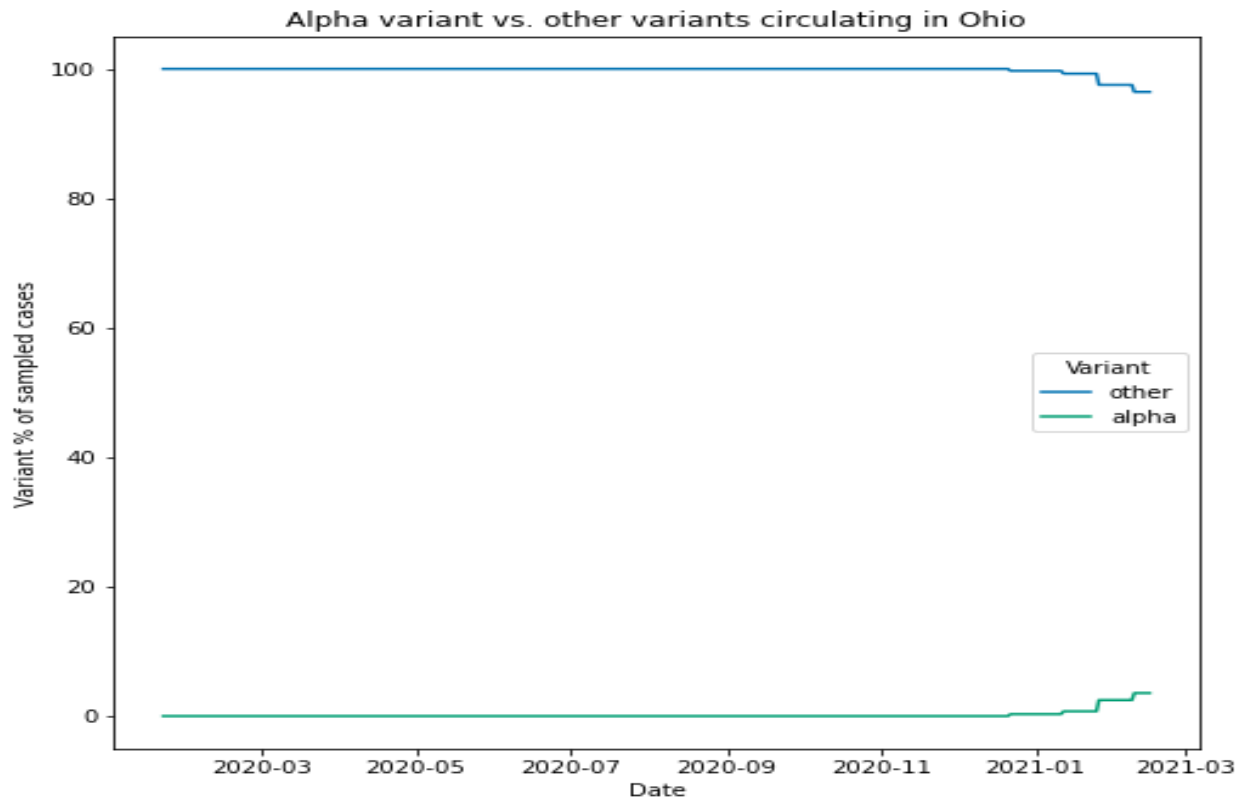




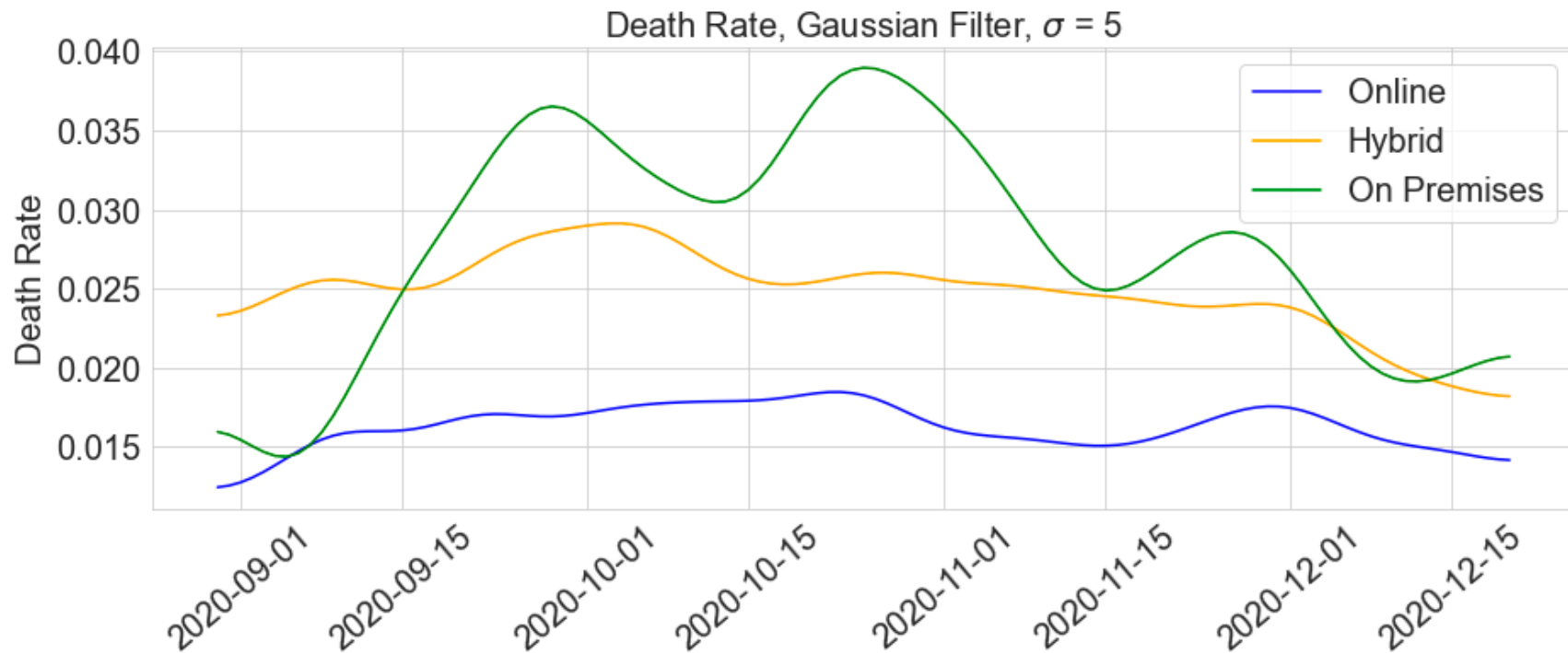
Carnegie Mellon University

Death Rate Analysis: Further Details

Change in dominant variant strain is not responsible for the variation in death rate over time in Ohio

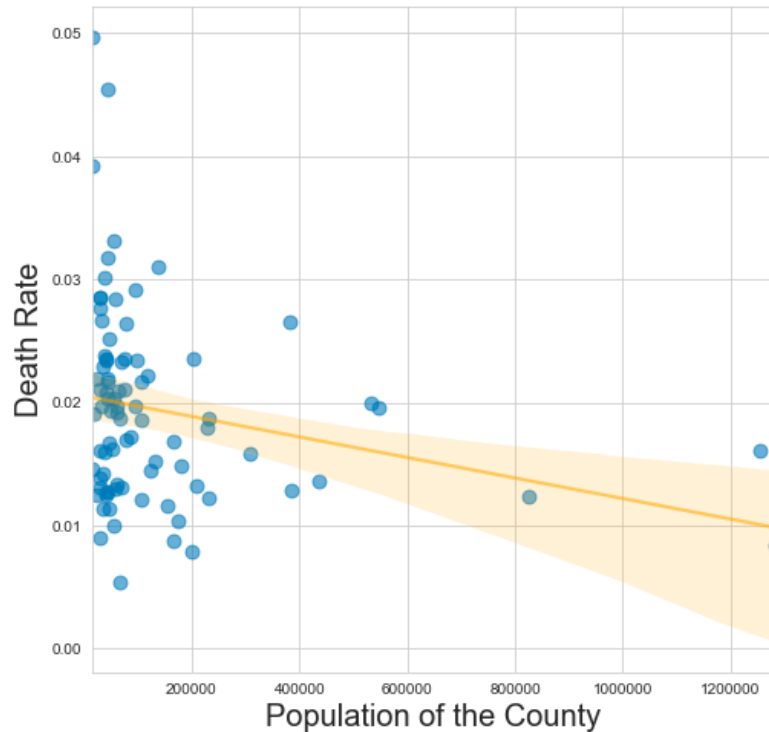


Death Rate breakdown by Teaching Posture

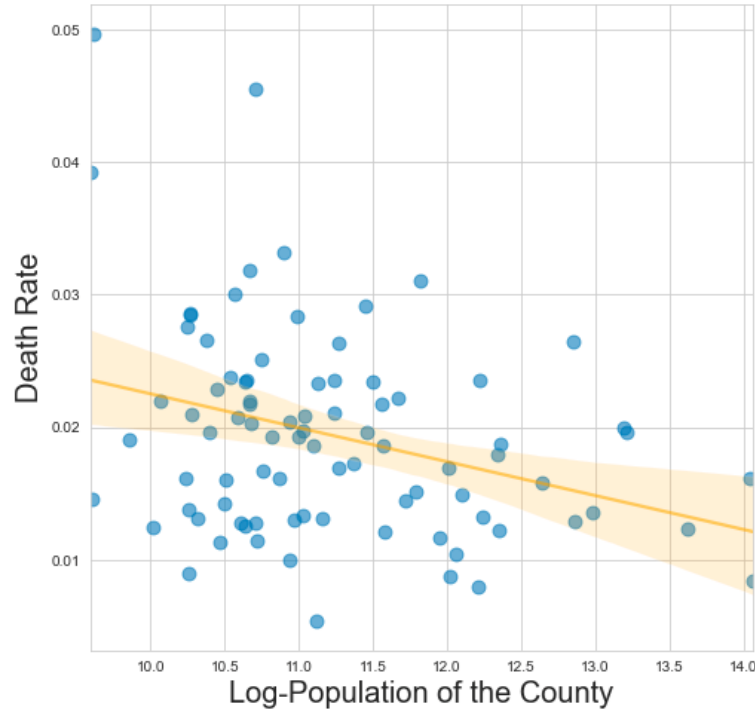


Death Rate v.s. Population

Population v.s. Death Rate(Aug 2020 - Dec 2020)
Correlation = -0.2348

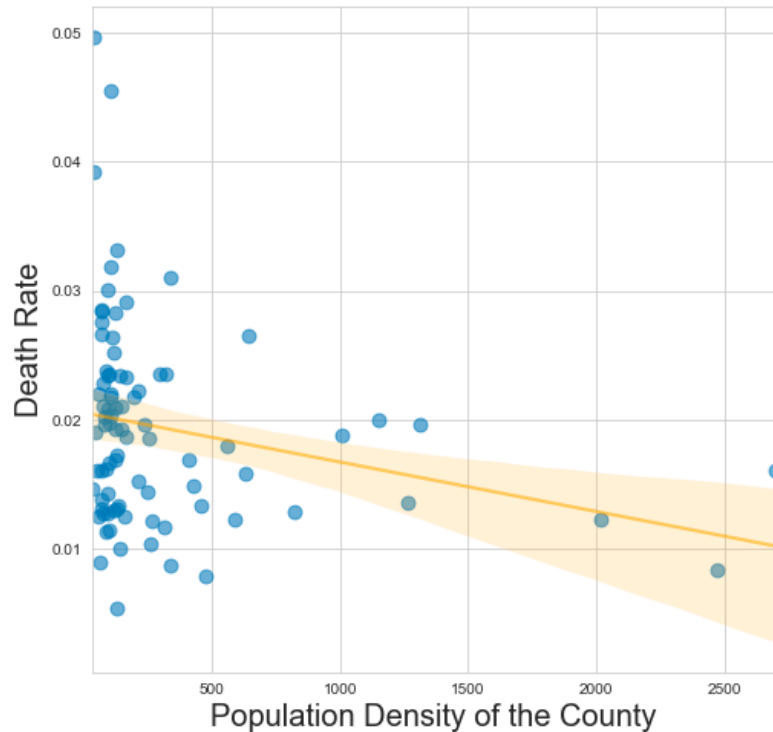


Log-Population v.s. Death Rate(Aug 2020 - Dec 2020)
Correlation = -0.3194

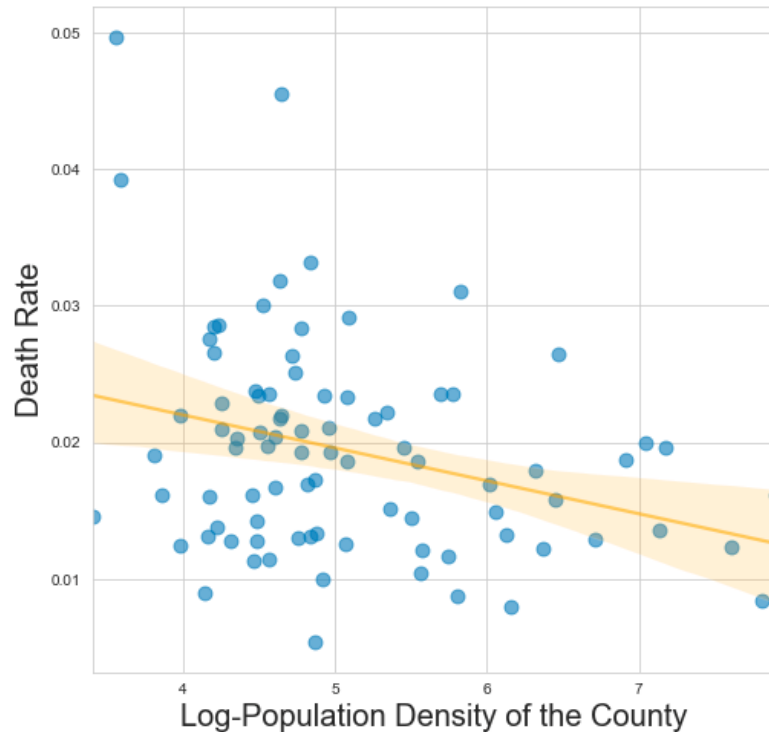


Death Rate v.s. Population Density

Population Density v.s. Death Rate(Aug 2020 - Dec 2020) Correlation = -0.2357



Log-Population Density v.s. Death Rate(Aug 2020 - Dec 2020) Correlation = -0.3042





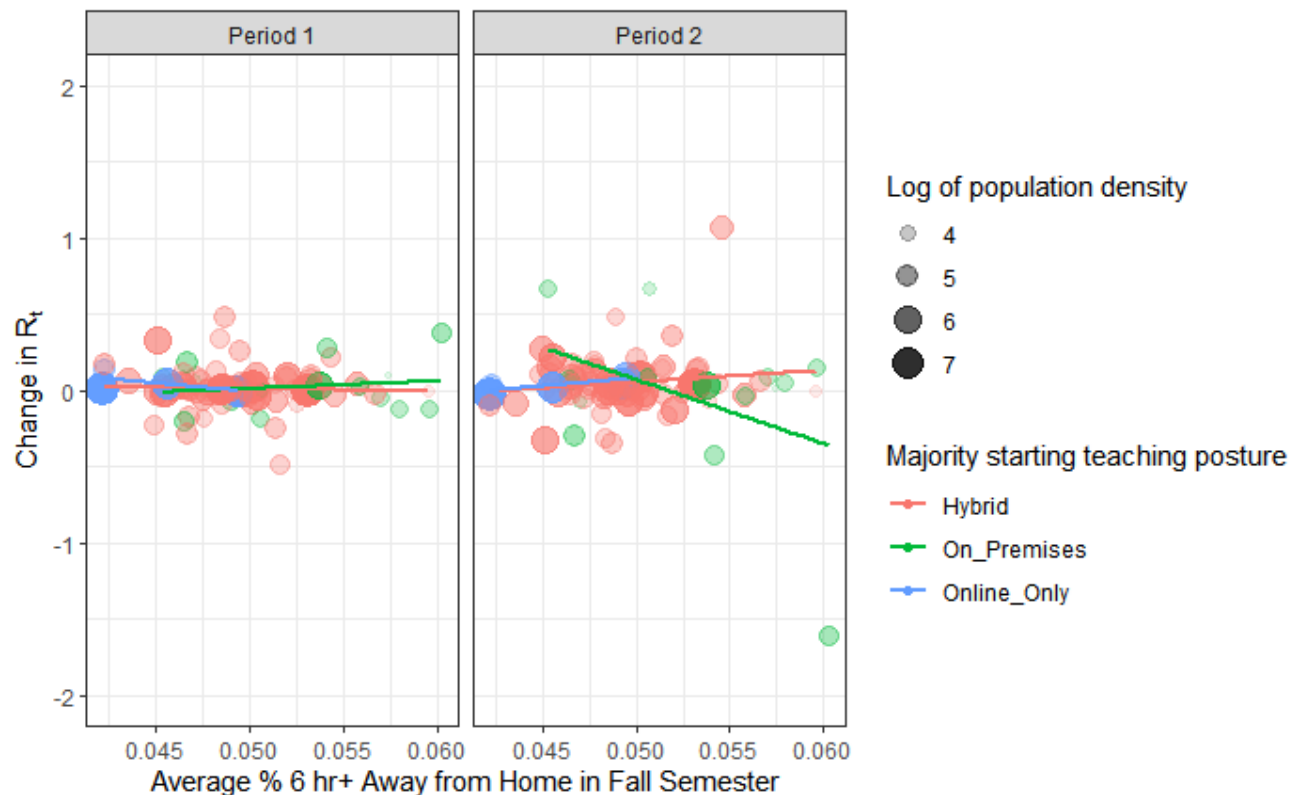
Carnegie Mellon University

Additional R_t Plots

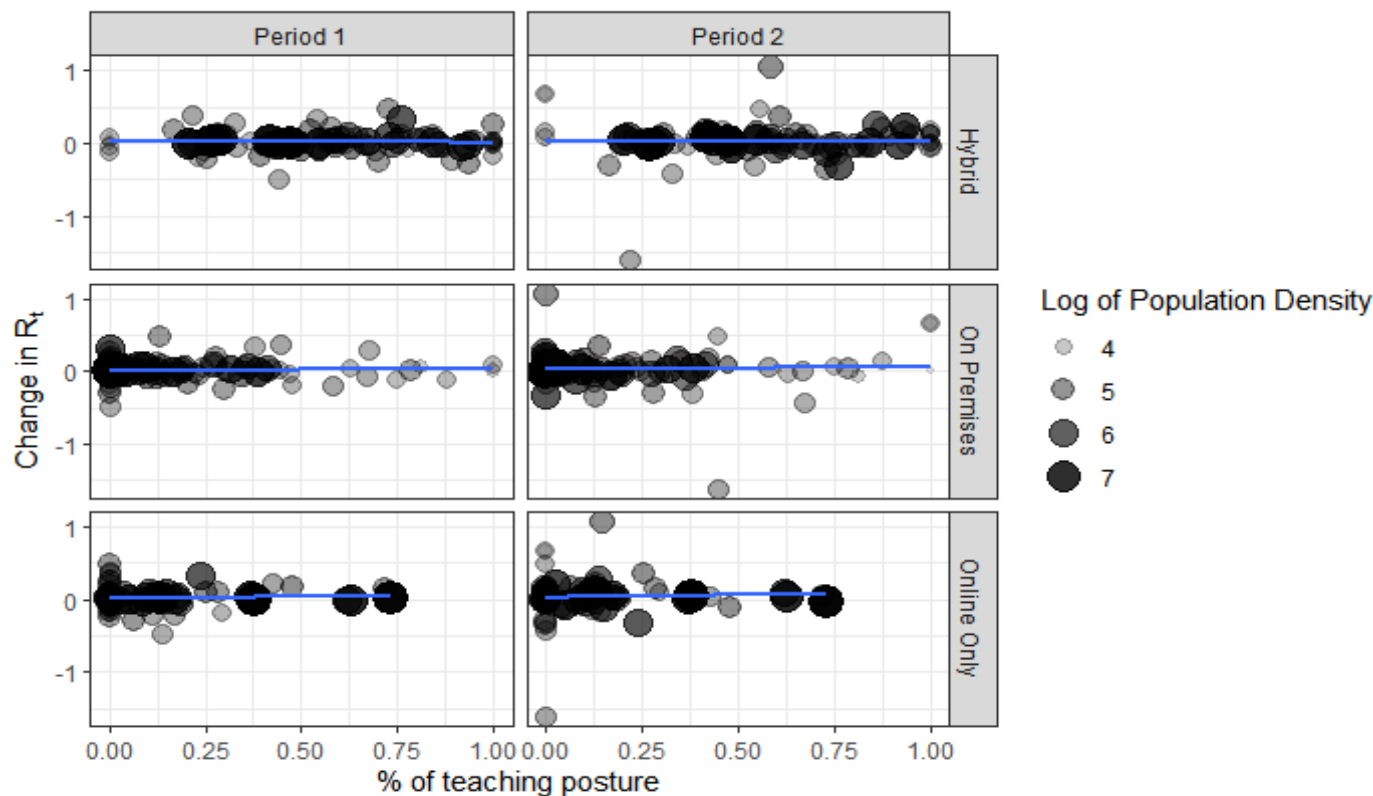
No notable difference in R_t for majority in-person teaching posture counties at the beginning of the school year



Change in R_t does not increase for majority in-person teaching posture counties at the beginning of the school year



Counties with more students on premises do not see increases in the change in R_t at the start of the school year



No increase in change in R_t at the beginning of the school year for Micropolitan counties regardless of teaching posture

