Meteorological and behavioral correlates of COVID-19 transmissibility across the United States

Venkatsai Bellala

July 21, 2022

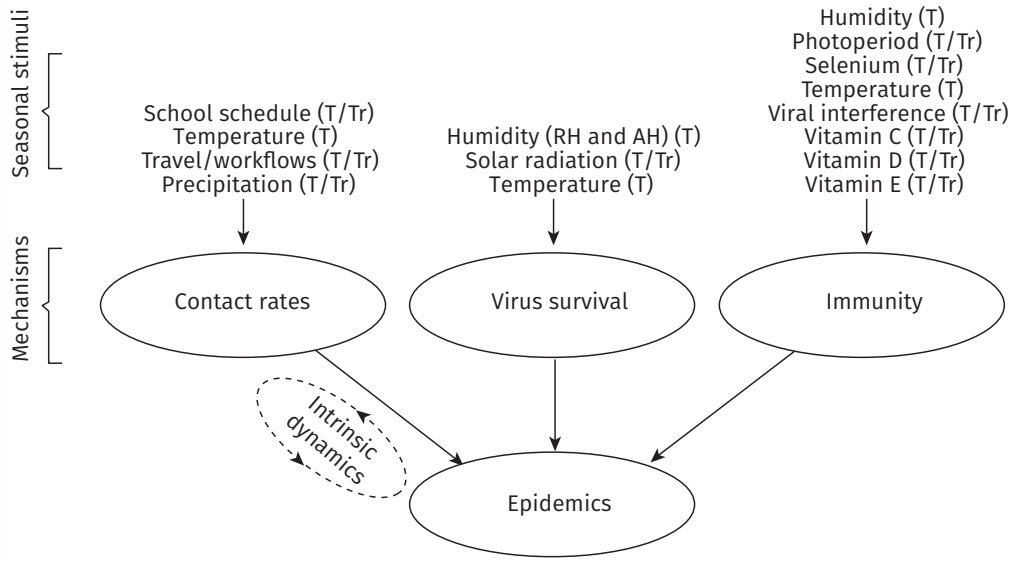
Applied Interdisciplinary Research in Air (AIR^2)

Background i

Virus stability and transmissibility driven by environmental factors and human behaviors

- Temperature (T)
- Relative Humidity (RH)
- Absolute Humidity (AH)
- Gathering indoors / being in close proximity

Figure 1: Factors mediating influenza epidemics [8]

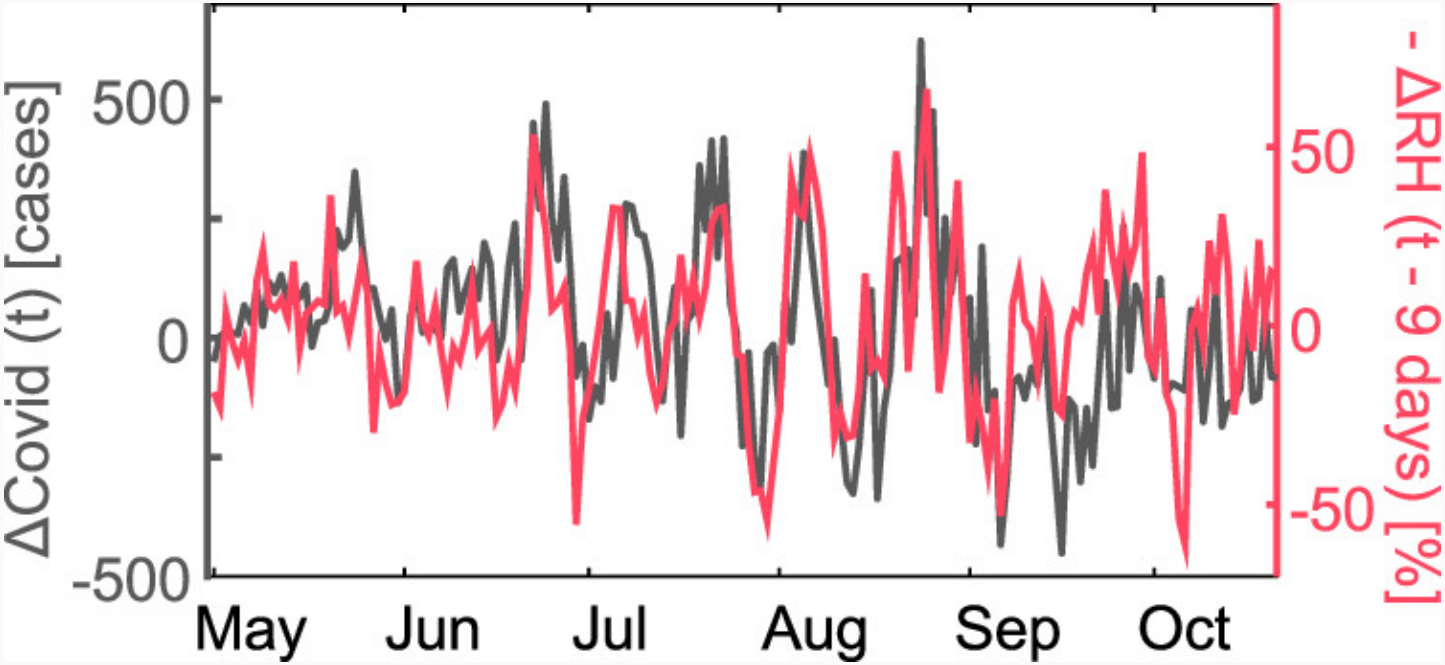


Instantaneous reproduction number (R_t)

- Expected infections by one individual at a given time

Background ii

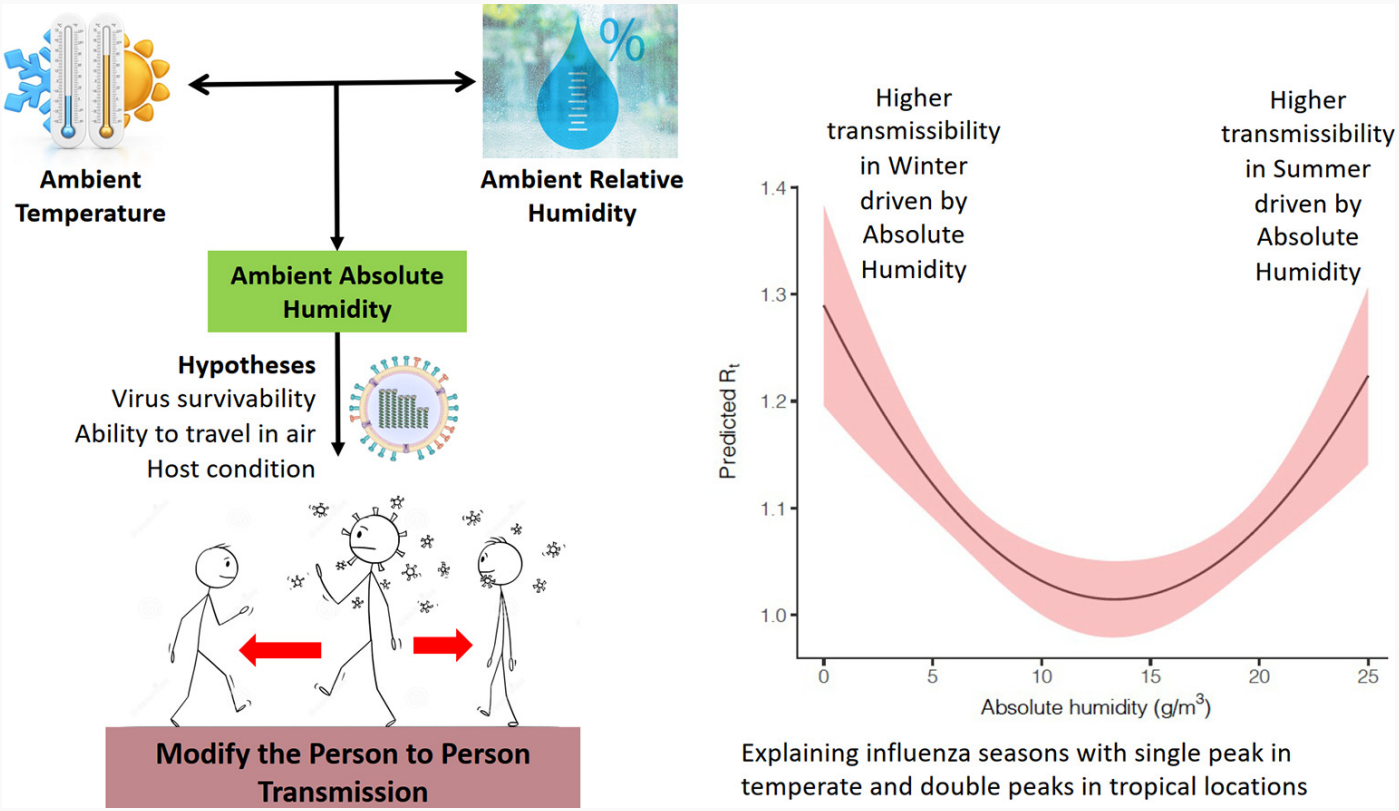
Figure 2: Detail of the co-evolution of ΔCovid(t) (gray; left axis) and negative ΔRH(t – 9 days) (red; right axis) during the months with a high number of COVID-19 confirmed cases. [4]



(Pineda Rojas et al., 2021)

Background iii

Figure 3: AH as a driver of influenza transmissibility [1]



(Ali et al., 2022)

Background iv

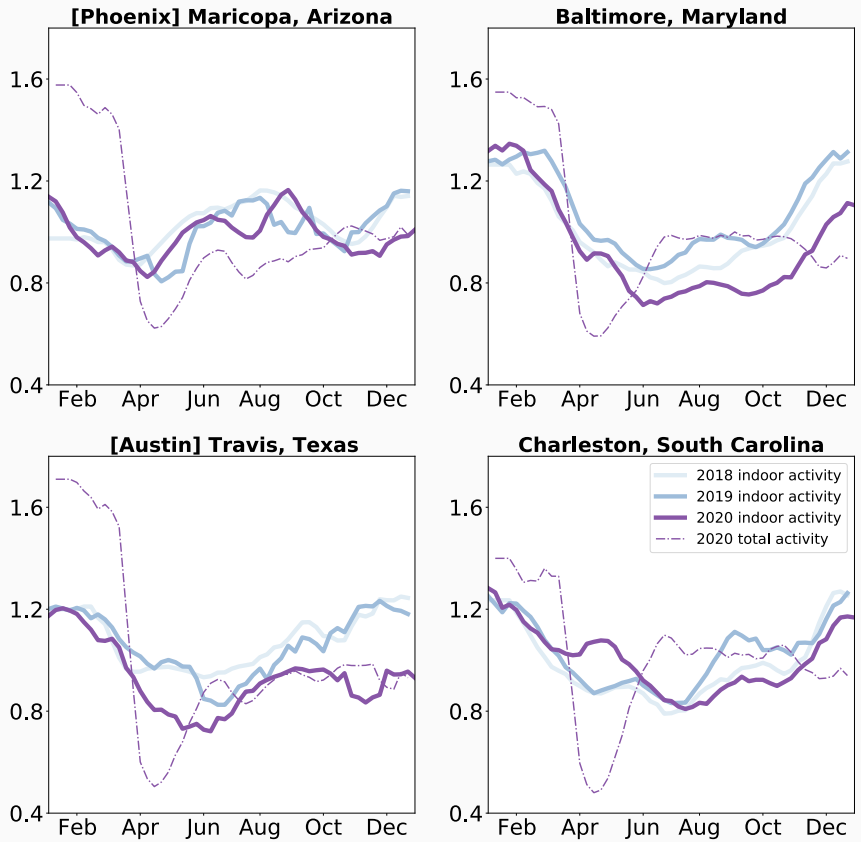
Indoor activity as a metric of "host social behavior"

- Cell phone data from SafeGraph Weekly Patterns
- 4.6 million POIs (excl. home locations)

1.2 $\longrightarrow$ 20% **more** indoor than average

0.8 $\longrightarrow$ 20% **less** indoor than average

Figure 4: Indoor activity trends [7]



(Susswein et al., 2022)

Objective

Quantify the role of

- dry bulb temperature (DBT)

on the instantaneous reproduction number of COVID-19.

Objective

Quantify the role of

- dry bulb temperature (DBT)
- relative humidity (RH)

on the instantaneous reproduction number of COVID-19.

Objective

Quantify the role of

- dry bulb temperature (DBT)
- relative humidity (RH)
- absolute humidity (AH)

on the instantaneous reproduction number of COVID-19.

Objective

Quantify the role of

- dry bulb temperature (DBT)
- relative humidity (RH)
- absolute humidity (AH)
- indoor activity

on the instantaneous reproduction number of COVID-19.

Methods

Methods i

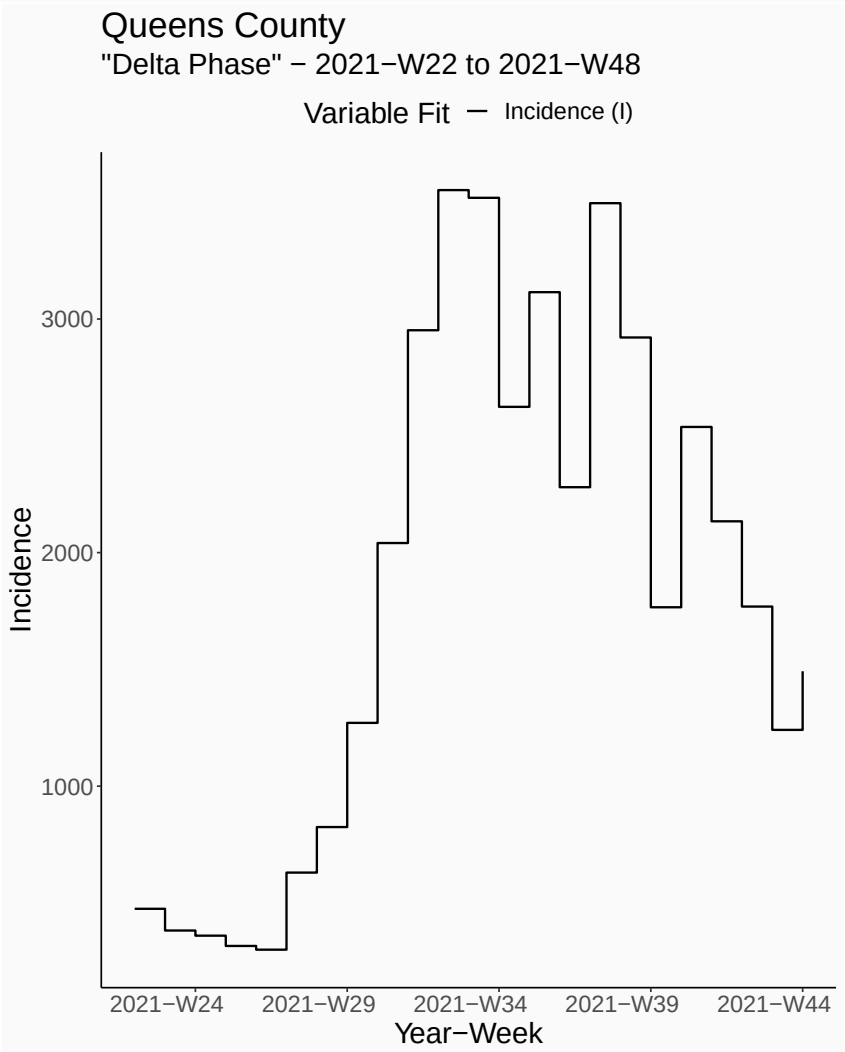
Data sources:

- Daily COVID-19 surveillance data – JHU CSSE
- Hourly meteorological measurements – NOAA
  - AH calculated with *humidity* R package
- Weekly indoor activity – Susswein et al. [7]

Organization:

- Weekly temporal scale
  - 2021-W22 to 2021-W48
- County-level analysis with R version 4.2.0 [5]

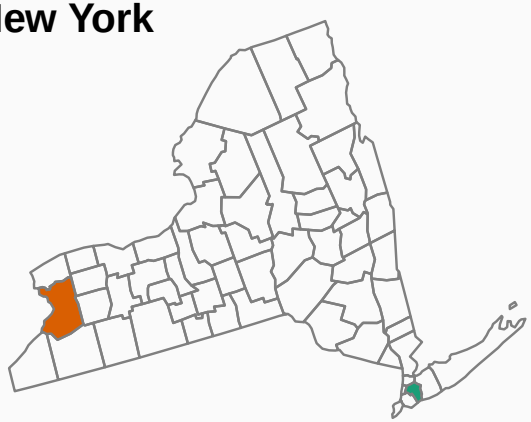
Figure 5: Incidence v. Time



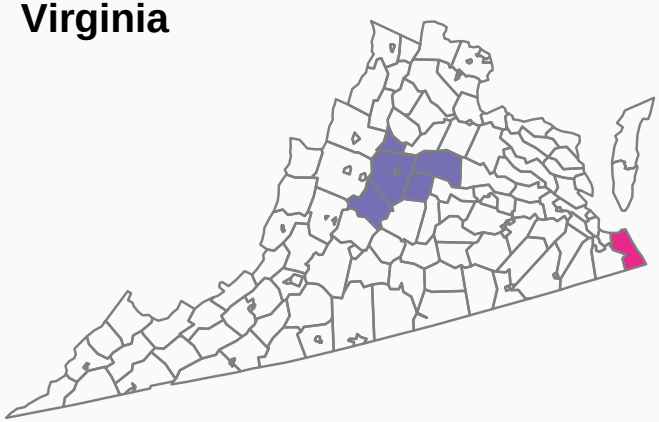
Methods ii

Figure 6: Map of Regions Studied

New York



Virginia



Texas

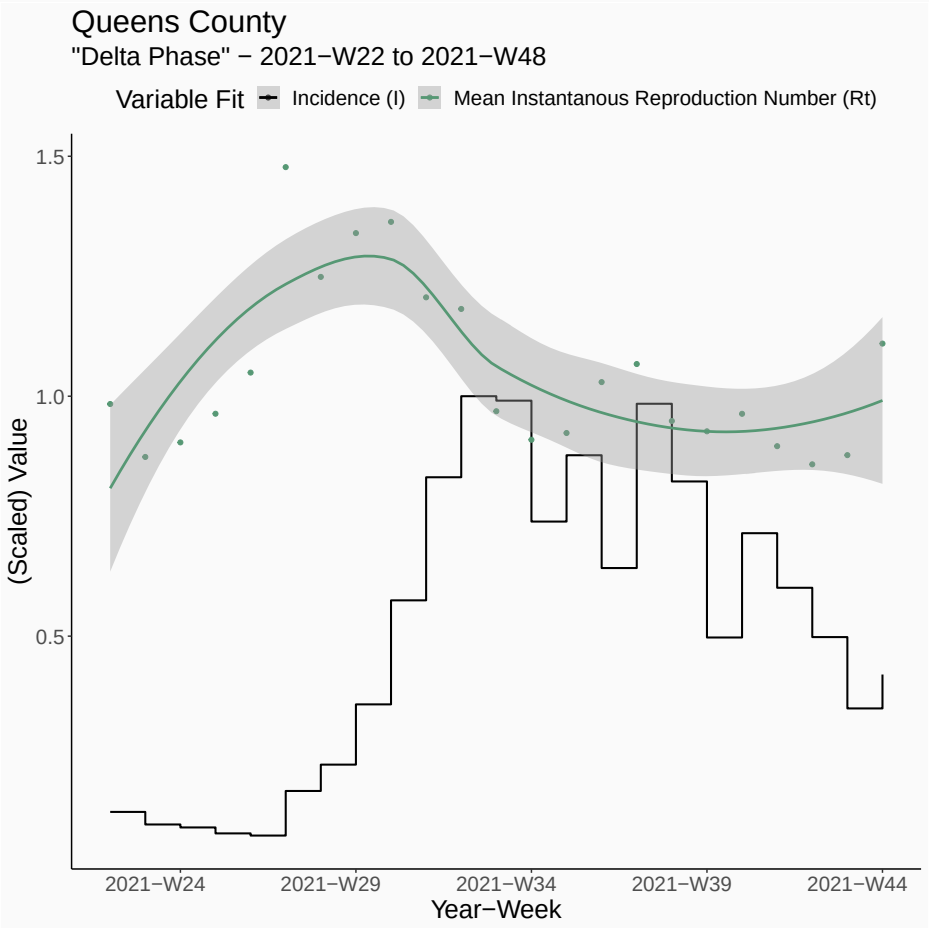
Region	Population
Queens County	2,253,858
Erie County	918,702
Thomas Jefferson Health District	256,206
Virginia Beach Health District	449,974
Cameron County	423,163
Harris County	4,713,325

Methods iii

R_t for SARS-CoV-2 computed with the Cori et al. method [2]

- *EpiEstim* R package
- Incidence $\approx$ 7 day rolling average of confirmed cases
- Sliding window of **7** days
- Parametric serial interval (SI) method
  - $\mu = 4.8, \sigma = 2.3$ [3]

Figure 7: Incidence and R_t v. Time



Methods iv

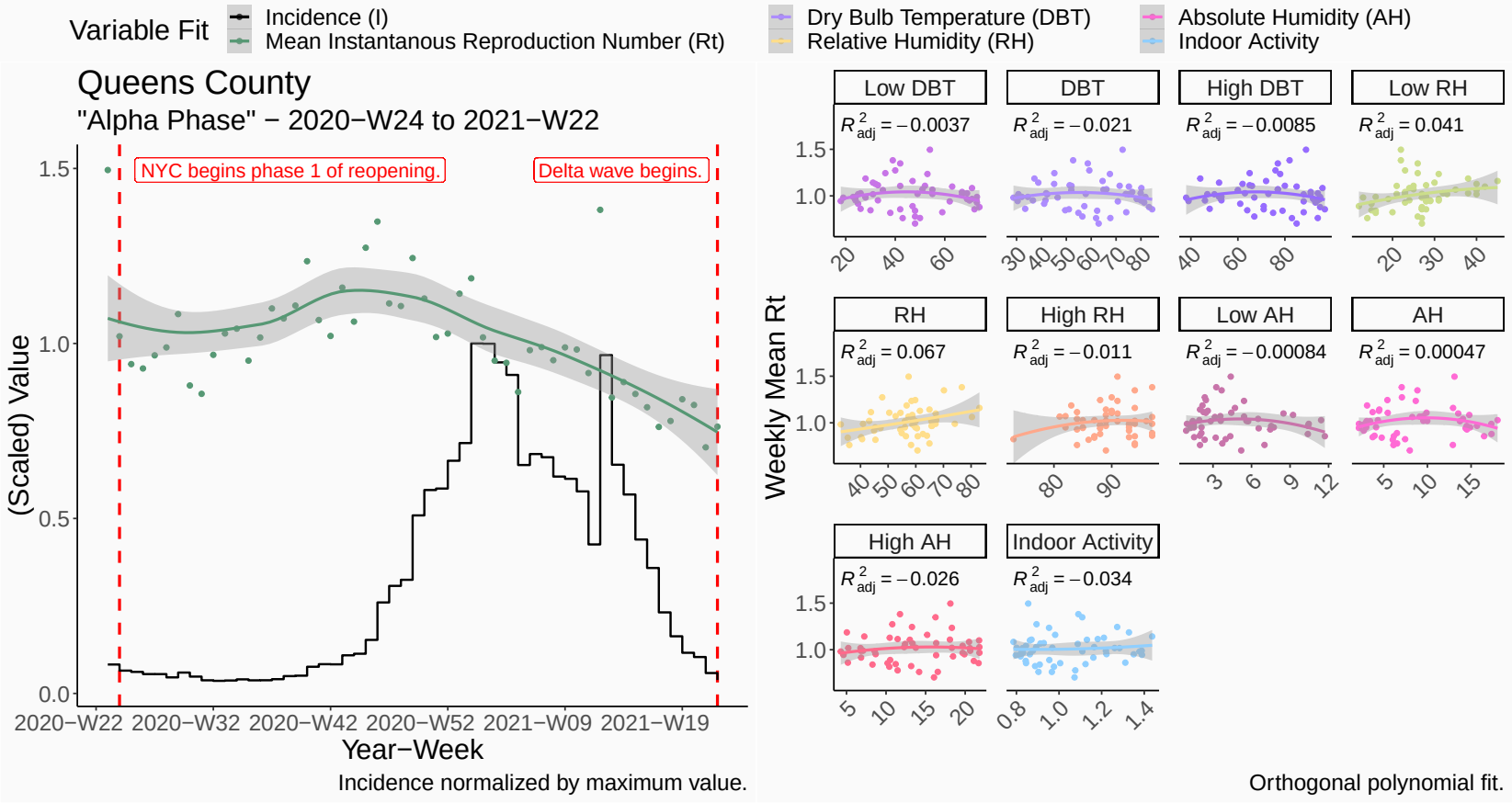
Univariate linear regression models $\longrightarrow$ multivariate regression

Lead-lag analysis with cross-correlation

- *CCF( )* in *feasts* R package

Results

Figure 8: COVID-19 “Alpha Phase” - Queens County



Results ii

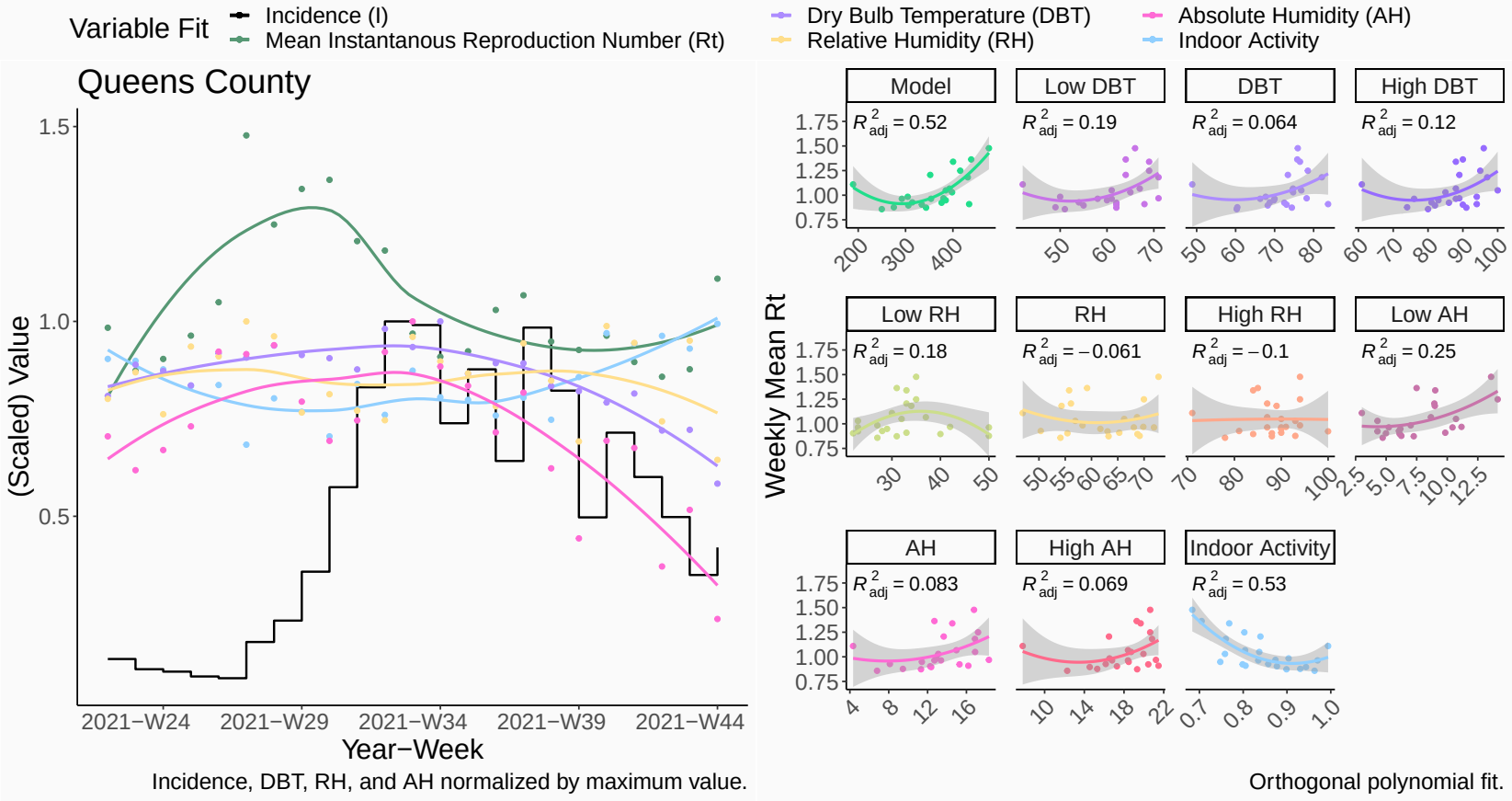
- Poor fits indicate weather patterns and indoor activity are weak correlates
  - Drop in total activity in 2020
  - *Preventative measures* are potentially a stronger correlate
- Solution: Analyze the **Delta Wave**
  - Relaxation of preventative measures
  - Presence of vaccinations and immune population

Results iii

$$\bar{R} \approx \text{poly}\left(\frac{\text{T} \cdot \ln\left(\text{RH}\right)}{\text{Indoor Activity}}\ ,\ 2\right) \qquad \text{(Orthogonal polynomial.)}$$

Results iv

Figure 9: Queens County



Incidence, DBT, RH, and AH normalized by maximum value.

Orthogonal polynomial fit.

Figure 10: Erie County

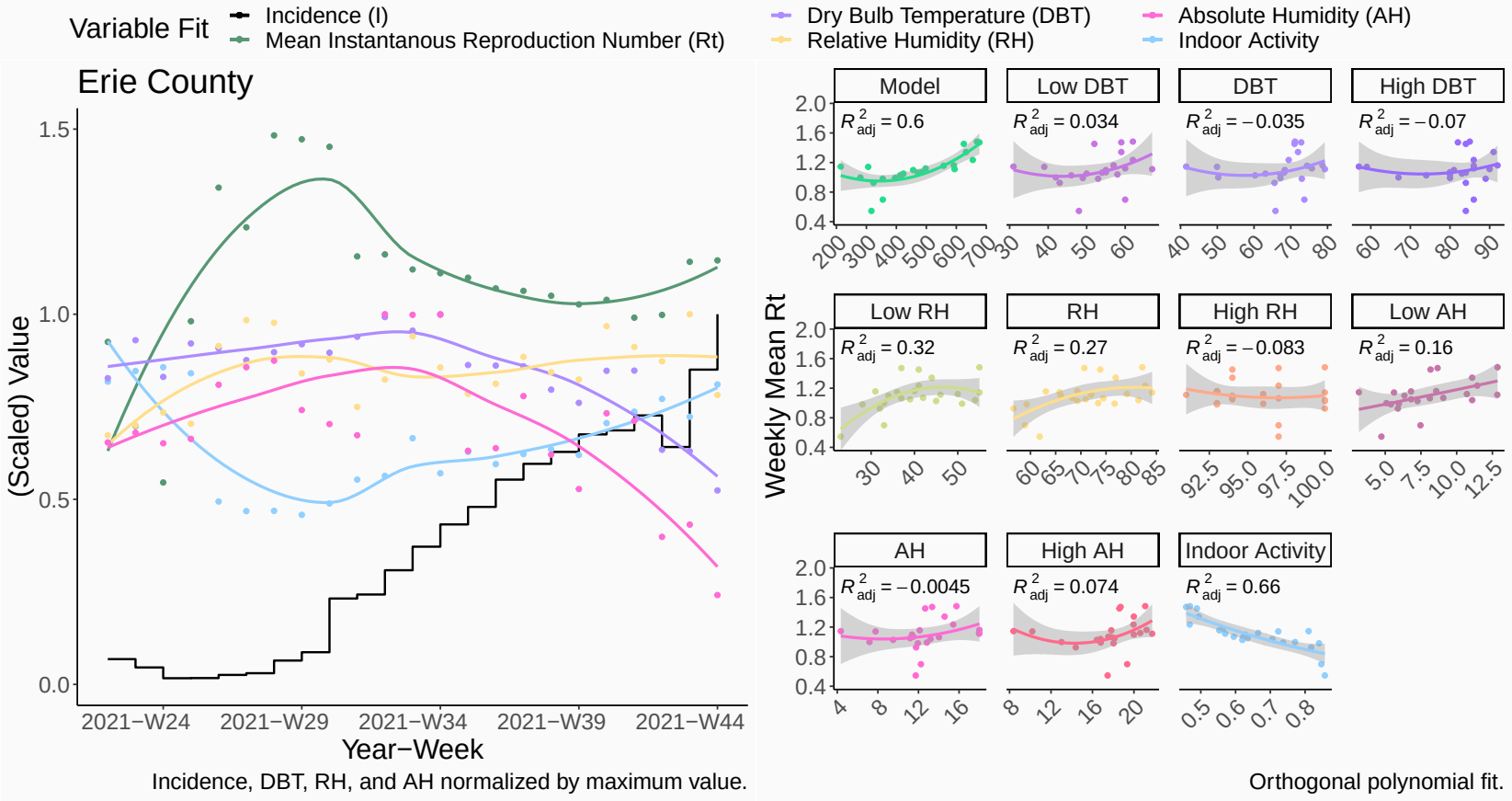
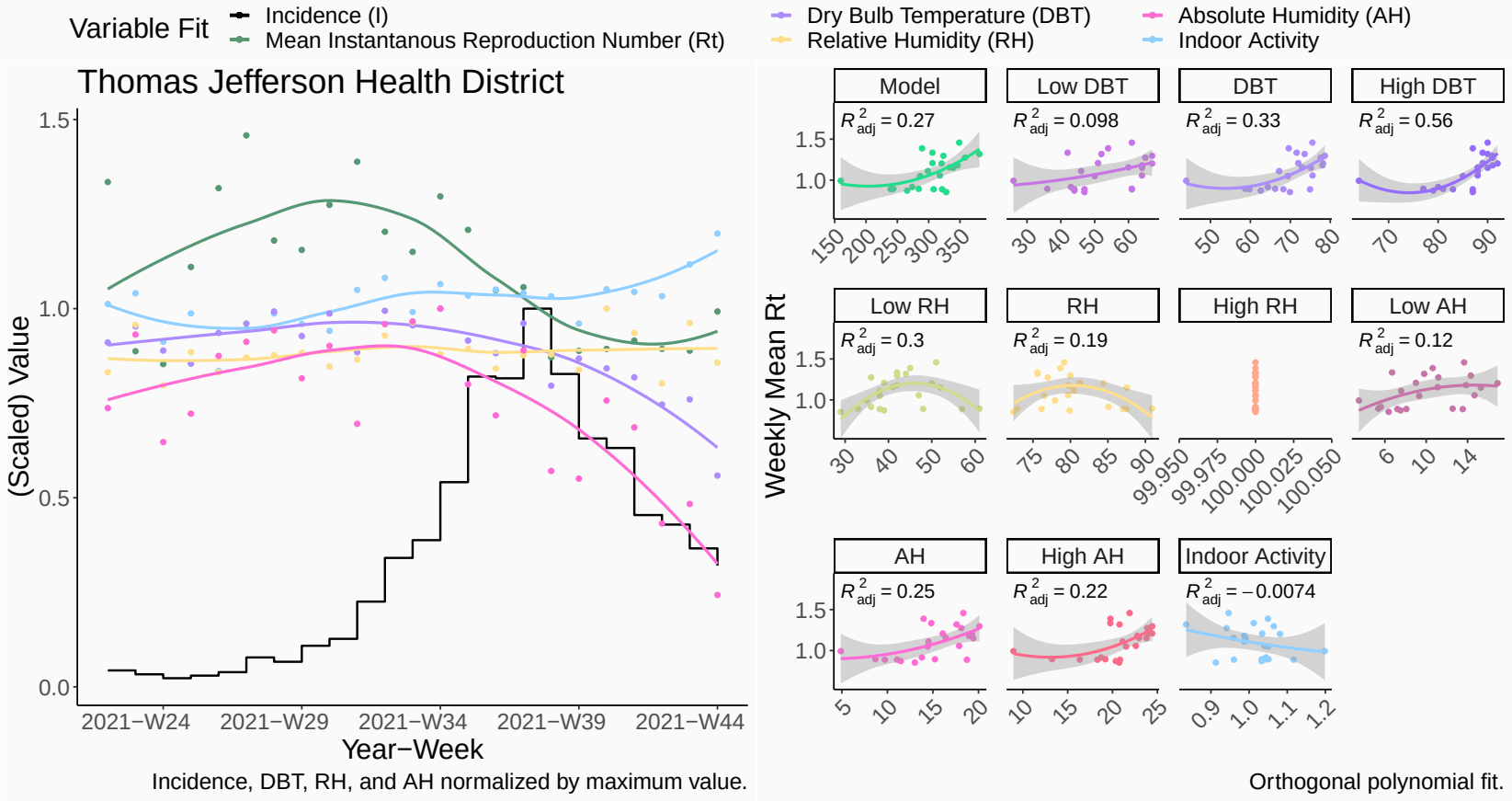


Figure 11: Thomas Jefferson Health District

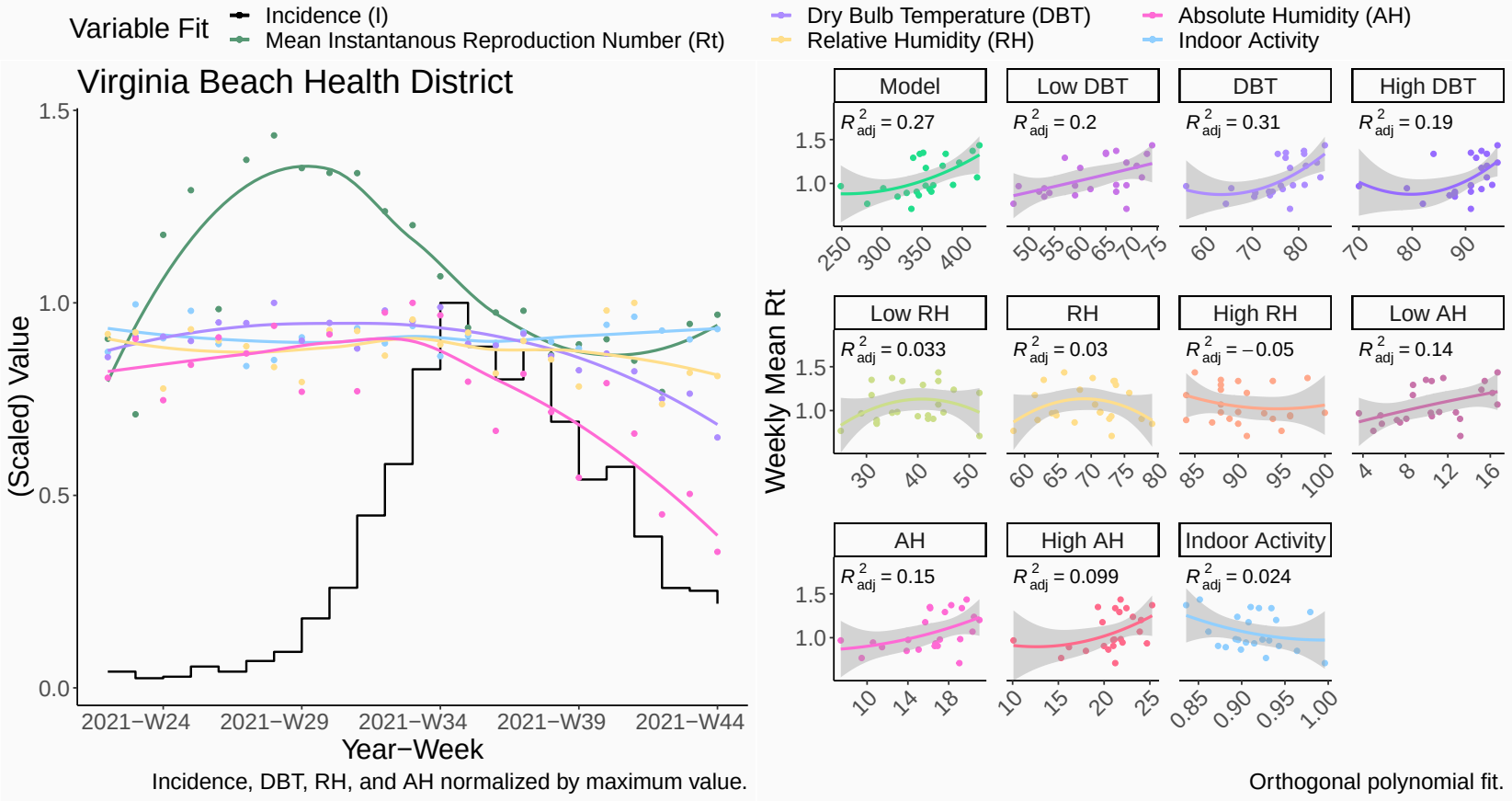


Incidence, DBT, RH, and AH normalized by maximum value.

Orthogonal polynomial fit.

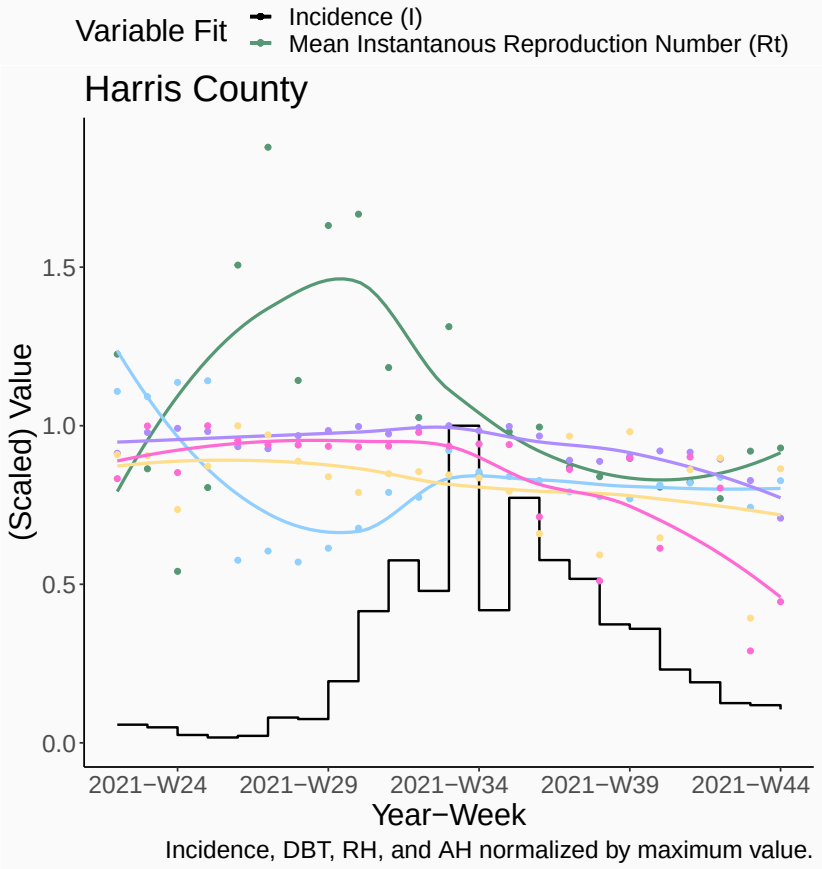
Results vii

Figure 12: Virginia Beach Health District

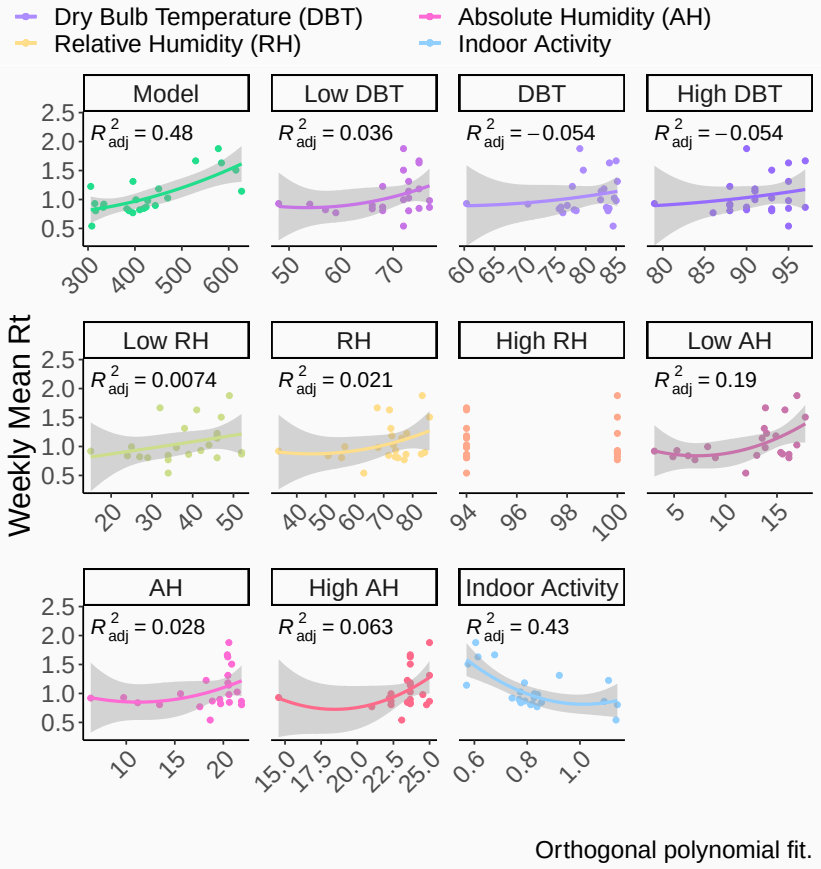


Results viii

Figure 13: Harris County

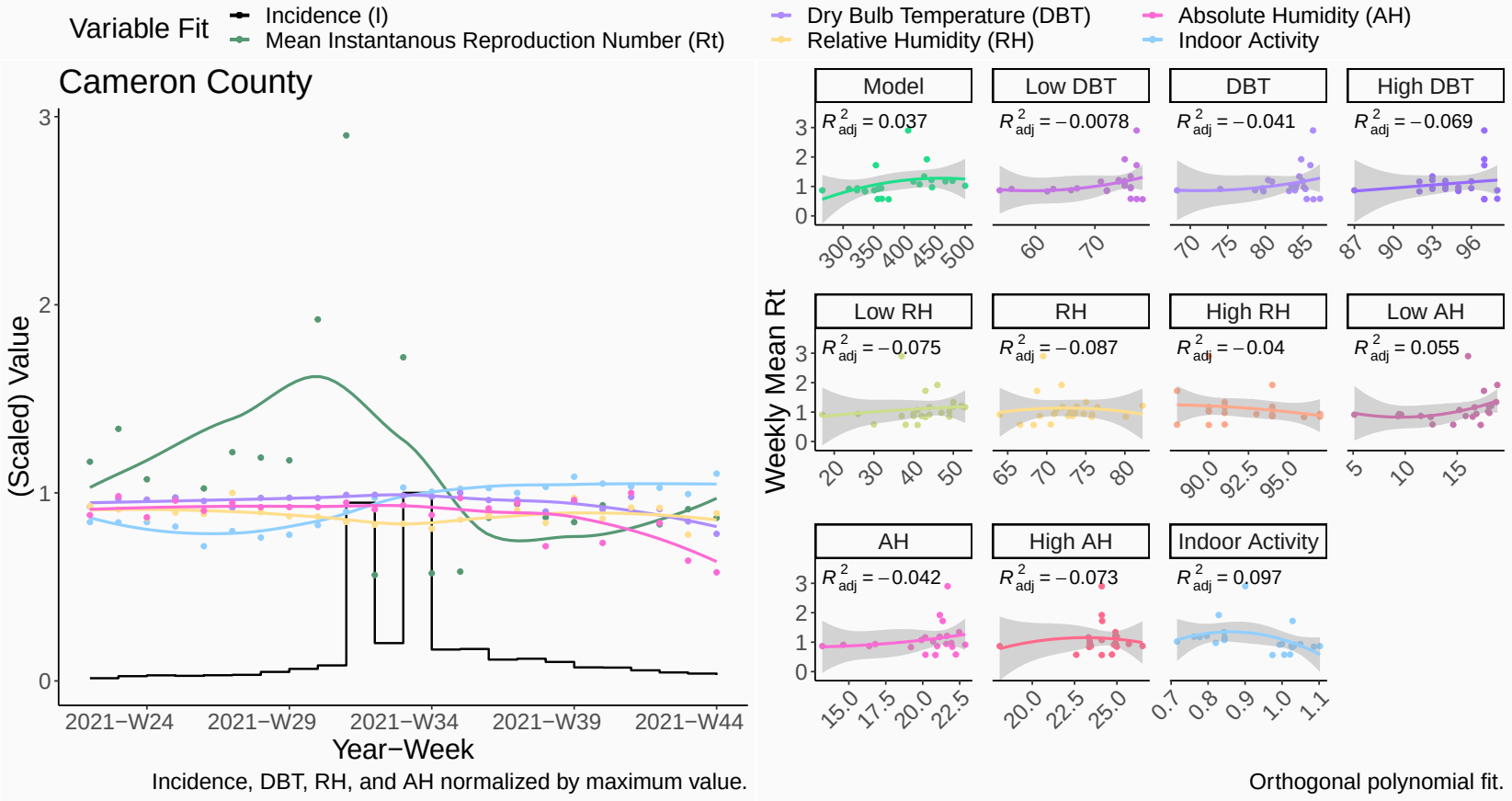


Incidence, DBT, RH, and AH normalized by maximum value.



Orthogonal polynomial fit.

Figure 14: Cameron County



Incidence, DBT, RH, and AH normalized by maximum value.

Orthogonal polynomial fit.

Results x

Table 1: Summary of R^2_{adj} and optimal lag

State	Region	R^2_{adj}	Lag (weeks)
NY	Queens County	0.52	0
	Erie County	0.60	0
VA	Thomas Jefferson Health District	0.28	-1
	Virginia Beach Health District	0.26	-2
TX	Harris County	0.48	0
	Cameron County	0.037	2

Results xi

- U-shaped association with weekly mean R_t during the “Delta phase”
  - Weak & non-significant associations during the “Alpha phase”
- Latitudinal trend $\longrightarrow$ stronger further from equator
- Population differences $\longrightarrow$ stronger in larger populations

Conclusions

Conclusions

DBT, RH, and indoor activity are significant drivers of SARS-CoV-2 transmission in the **near-absence** of preventative measures.

- Inform predictions of future COVID-19 outbreaks
- Guide public policy planning in affected communities

Future Directions

- Fit model to more counties across the US
- Analyze the Omicron wave
- Control for lags in symptom onset and reporting
- Incorporate demographic parameters
    - Age, risk factors, etc.

Acknowledgements

Special thanks to:

- Annabel Coyle for her collaboration
- Linsey Marr for her mentorship
- David Schmale and Shane Ross for supporting this program

This research was supported in part by the National Science Foundation (NSF) under grant number 1922516.

Questions?

References i

S. T. Ali, B. J. Cowling, J. Y. Wong, D. Chen, S. Shan, E. H. Lau, D. He, L. Tian, Z. Li, and P. Wu.
Influenza seasonality and its environmental driving factors in mainland china and hong kong.
Science of The Total Environment, 818:151724, 2022.

A. Cori, Z. Kamvar, J. Stockwin, T. Jombart, E. Dahlqwist, R. FitzJohn, and R. Thompson.
EpiEstim v2.2-3: A tool to estimate time varying instantaneous reproduction number during epidemics.
https://github.com/mrc-ide/EpiEstim, 2021.

H. Nishiura, N. M. Linton, and A. R. Akhmetzhanov.
Serial interval of novel coronavirus (covid-19) infections.
International journal of infectious diseases, 93:284–286, 2020.

References ii


A. L. Pineda Rojas, S. M. Cordo, R. I. Saurral, J. L. Jimenez, L. C. Marr, and E. Kropff.
Relative humidity predicts day-to-day variations in covid-19 cases in the city of buenos aires.
Environmental Science & Technology, 55(16):11176–11182, 2021.


R Core Team.
R: A language and environment for statistical computing, 2021.


D. Rubin, J. Huang, B. T. Fisher, A. Gasparrini, V. Tam, L. Song, X. Wang, J. Kaufman, K. Fitzpatrick, A. Jain, et al.
Association of social distancing, population density, and temperature with the instantaneous reproduction number of sars-cov-2 in counties across the united states.
JAMA network open, 3(7):e2016099–e2016099, 2020.

References iii

Z. Susswein, E. C. Rest, and S. Bansal.
Disentangling the rhythms of human activity in the built environment for airborne transmission risk.
medRxiv, 2022.

J. Tamerius, M. I. Nelson, S. Z. Zhou, C. Viboud, M. A. Miller, and W. J. Alonso.
Global influenza seasonality: reconciling patterns across temperate and tropical regions.
Environmental health perspectives, 119(4):439–445, 2011.

Assumption Validation

Assumptions i

Table 2: Shapiro-Wilk test of normality

State	Region	p-value
NY	Queens County	0.0061
	Erie County	0.0831
VA	Thomas Jefferson Health District	0.0578
	Virginia Beach Health District	0.0845
TX	Harris County	0.140
	Cameron County	0.0001

Weekly mean R_t is *mostly* normally distributed ($p > 0.05$).

Assumptions ii

For linear regression ...

- Linearity
- Homoscedasticity
- Errors are independent of x
- Errors are normal
- Mean of errors is 0

Assumptions iii

Autocorrelation present in residuals

Heteroskedasticity present in data

Attempts to control for them have proved unfruitful