**Winter storms and unplanned school closure announcements on Twitter: Comparison between the states of Massachusetts and Georgia, 2017-18**

**Online Supplementary Materials**

Haley I. Evans 1†, Maya T. Handberry 1†, Kamalich Muniz-Rodriguez 1, Jessica S. Schwind 1, Hai Liang 2,

Bishwa B. Adhikari 3, Martin I. Meltzer 3, Isaac Chun-Hai Fung 1,4\*

† Co-first authors

\*Correspondence should be addressed to Isaac Chun-Hai Fung, PhD, Department of Biostatistics, Epidemiology & Environmental Health Sciences, Jiann-Ping Hsu College of Public Health, P.O. Box 7989, Georgia Southern University, Statesboro, GA 30460-7989, U.S.A. Telephone: +1 912 478 5079; Fax: +1 912 478 0171. Email: cfung@georgiasouthern.edu ORCID: 0000-0001-5496-2529

1 Department of Biostatistics, Epidemiology and Environmental Health Sciences, Jiann-Ping Hsu College of Public Health, Georgia Southern University

2 School of Journalism and Communication, Chinese University of Hong Kong

3 Health Economics and Modeling Unit, Division of Preparedness and Emerging Infections, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention

4 Guest Researcher, Health Economics and Modeling Unit, Division of Preparedness and Emerging Infections, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention

Disclaimer: The opinions expressed in this article are that of the authors and do not necessarily represent the official positions of the Centers for Disease Control and Prevention or the United States Government.

**Outline**

Part I: Data Collection, Documentation, Codebook and

Part II: Further Discussion

Part III: Supplementary Tables

Part IV: Additional Figures

**Acronyms**

API: Application Programming Interface

GA: Georgia

MA: Massachusetts

NCES: National Center for Education Statistics

USCA: Unplanned School Closure Announcements

**Part I: Data Collection, Documentation, and Codebook**

Manual Codebook Creators: Haley I. Evans and Maya T. Handberry

Lead Manual Data Collectors: Haley I. Evans and Maya T. Handberry

Project PI (and liaison with CDC): Isaac Chun-Hai Fung

Project Co-PI(s): Bishwa B. Adhikari, Jessica S. Schwind, Martin I. Meltzer

**Project Objective**

The purpose of this project is to quantify and compare the rates of unplanned school closure announcements on Twitter between public school districts in Massachusetts and Georgia during the Winter Season of 2017-2018.

**Procedure Data Collection, Processing, and Editing**

**Component description**

Data downloaded from the NCES website (https://nces.ed.gov/) contained demographic school information such as district name, county the district is located in, total number of students, and student-teacher ratio, etc. Each of the following spreadsheets contained the school name, district name, school address, locality (city, suburb, town, or rural), Twitter account, Twitter announcements, etc.

• Georgia Public School Districts

• Massachusetts Public School Districts

**Phase 1**: Public School District Twitter Handle Collection

a) Identified winter storms that affected the two states

b) Downloaded public-school district data from the NCES website.

c) Manually searched for Twitter accounts for each public-school district.

d) Verified Twitter handles by clicking on the links on the school district’s webpage.

e) If the webpage was not found, OR a Twitter handle was not located on the webpage, searched for the school/district name and address on Google with the word “Twitter”.

f) Verified Twitter handle by checking the school address/location.

g) If the above methods were still unsuccessful, used the search function on the Twitter homepage (www.twitter.com) to search the name and address of the district.

h) Only Twitter handles owned by the superintendent or school district were accepted.

**Phase 2:** Tweet Extraction and Data Management

a) Using the Twitter handles, the Tweets were identified for the public school districts using API

b) Merged the data obtained from NCES and API

c) Determined the correct amount of days for the school year, the winter season, and the affected days for Winter Storms for the two states

d) Subsetted the data into three data sets for each state: school year, winter season, and Unplanned School Closure Announcements

d) Created a new variable for the rate of frequency of tweets by the amount of days

e) Determined the correct locale label from NCES locality data

e) Created a new variable for locale variable from NCES locality variable

f) Converted all relevant variables to numeric variables for further use

g) In the Unplanned School Closure Announcements data set for the two states, created new variables for the log10 version of relevant tweets

h) Merged the Unplanned School Closure Announcements data for the two states

**Keywords:** Closed, Storm\* Snow\*, Quinn, Riley, Skylar, Toby, Grayson, Benji, Inga, Hunter, Closure\*, Weather, Air Quality, Power, Outage\*, Temperature\*, Road\*, Condition\*, Ice, Cold, Icy, Blizzard\*, Closing\*, Flooding, Flood\*, Cancellation\*, Cancel, Cancelled, Snowfall\*, Safety, Dismissed, Prepare, Emergency, Preparedness, Snowstorm **\***plural form of the specific words

**Part II: Further Discussion**

As presented in the main text, increases in Twitter usage was seen during all winter storms in both Massachusetts and Georgia except for Winter Storm Toby in Massachusetts. A possible reason was that Winter Storm Skylar occurred the week prior to Toby. Schools could have already been closed and fewer tweets from the school district Twitter accounts were needed to inform about the school closure. It could be that only one tweet was sent out saying schools will remain closed instead of multiple tweets including reminders and other weather information.

**Part III: Supplementary Tables**

Table S1. Descriptive statistics of school districts in Georgia and Massachusetts, and among those with Twitter handles, the number of tweets per district.

|  |  |  |
| --- | --- | --- |
|  | Georgia | Massachusetts |
| Number of school districts | 232 | 431 |
|  |  |  |
| **National Center for Education Statistics data** |  |  |
| Number of students, median (range) | 7,122 (110 – 178,214) | 1,868 (95 – 66,194) |
| Student-teacher ratio, median (range) | 15 (1 – 18) | 36 (1 – 102) |
|  |  |  |
| **Twitter data of school districts with Twitter handles** |  |  |
| Number (%) of school districts with Twitter handles\* | 66 (28.45) | 176 (40.84) |
| Number of tweets per district in academic year 2017-18 | 689 | 347 |
| Number of tweets per district in the 2017-18 winter season tweets per district (% of the whole academic year) | 245 (35.56) | 134 (38.62) |
| Number of 2017-18 winter storm related unplanned school closure announcement tweets per district (% of the whole academic year)  | 18 (2.61) | 23 (6.63) |

\* A chi-square test was run to determine significant difference between the percentage of districts with Twitter handles among GA and MA. The p-value was 0.04.

Table S2. District and tweet count (%) by Twitter activity and by state in Georgia and Massachusetts in the academic year of 2017-18.

|  |  |  |
| --- | --- | --- |
|  | Georgia | Massachusetts |
| Category of Tweet Count | District Count (%) | Total Tweet Count (%) | District Count (%) | Total Tweet Count (%) |
| Less than 250 | 19 (28.8) | 1,946 (4.3) | 116 (65.9) | 11,494 (18.8) |
| 251-500 | 14 (21.2) | 5,195 (11.4) | 22 (12.5) | 8,260 (13.6) |
| 501-750 | 15 (22.7) | 9,138 (20.1) | 13 (7.4) | 8,017 (13.2) |
| 751-1,000 | 3 (4.5) | 2,570 (5.7) | 11 (6.3) | 9,664 (15.9) |
| 1,001-1,250 | 2 (3.0) | 2,271 (5.0) | 3 (1.7) | 3,258 (5.4) |
| 1,251-1,500 | 2 (3.0) | 2,752 (6.1) | 0 (0) | 0 (0) |
| 1,501-1,750 | 4 (6.1) | 6,333 (13.9) | 6 (3.4) | 9,370 (15.4) |
| 1,751-2,000 | 2 (3.0) | 3,658 (8.0) | 1 (0.6) | 1,966 (3.2) |
| 2,001-2,250 | 2 (3.0) | 4,082 (9.0) | 3 (1.7) | 6,280 (10.3) |
| 2,250-2,500 | 2 (3.0) | 4,723 (10.4) | 0 (0) | 0 (0) |
| More than 2,500 | 1 (1.5) | 2,793 (6.1) | 1 (0.6) | 2,526 (4.2) |
| Total | 66 (100) | 45,461 (100) | 176 (100) | 60,835 (100) |

Table S3. Median (interquartile range, IQR) of numbers of schools and students, student-teacher ratio, follower count and following count for Georgia and Massachusetts school districts during the 2017-2018 school year among the districts that have Twitter accounts.

|  |  |  |
| --- | --- | --- |
|  | 2017-2018 School Year |  |
|   | Georgia | Massachusetts | P-value\* |
| Number of Schools, Median (IQR) | 15(6-37) | 6(4-9) | <0.0001 |
| Number of Students, Median (IQR) | 10,305 (4,402-41,916) | 3,040 (1,644-5,184) | <0.0001 |
| Student-teacher ratio, Median (IQR)  | 15.3(14.6-15.9) | 35 (24-47) | <0.0001 |
| Followers count, Median (IQR) | 2,808(1,044-16,669) | 1,052(628-2,104) | <0.0001 |
| Following count, Median (IQR) | 100(28-232) | 102(38-261) | 0.27 |

\* The Wilcoxon Signed-Rank Test was used to compare medians between the states.

Table S4. Median (interquartile range) of rate of tweets per day.

|  |  |  |  |
| --- | --- | --- | --- |
|  | 2017-2018 School Year | 2017-2018Winter Season | 2017-2018Winter Storm Unplanned School Closure Announcements |
| Georgia  | 3.84(1.63-5.57) | 4.57(1.77-5.62) | 3.38(1.88-5.25) |
| Massachusetts | 2.36(0.98-4.24) | 1.24(0.55-2.15) | 0.79(0.49-1.64) |
| P-Value\* | 0.003 | 0.69 | 0.31 |

\*The Wilcoxon Signed-Rank Test was used to compare the rate medians.

Table S5. Individual crude rate ratios of unplanned school closure announcements comparing GA and MA.

|  |  |  |
| --- | --- | --- |
| Variable | Rate Ratios (95% CI) | P-value |
| State  |  | <0.0001 |
|  Massachusetts | 0.25 (0.20, 0.33) |  |
|  Georgia | Reference |  |

Negative binomial regression was used and the response variable was the “rate” i.e. the frequency of total tweets/affected days.

Table S6. Adjusted rate ratios of unplanned school closure announcements comparing GA and MA with log10 transformed student number variable.

|  |  |  |
| --- | --- | --- |
| Variable | Rate Ratios (95% CI) | P-value |
| State |  | <0.0001 |
|  Massachusetts | 0.33 (0.25, 0.44) |  |
|  Georgia | Reference |  |
| Log10 (Student number) | 1.47 (1.20, 1.80) | 0.0009 |

Negative binomial regression was used and the response variable was the “rate” i.e. the frequency of total tweets/affected days.

Table S7. Adjusted rate ratios of unplanned school closure announcements comparing GA and MA with log10 transformed student number variable and interaction term.

|  |  |  |
| --- | --- | --- |
| Variable | Rate Ratios (95% CI) | P-value |
| State  |  | 0.06 |
|  Massachusetts | 0.19 (0.04, 0.86) |  |
|  Georgia | Reference |  |
| Log10 (Student number) | 1.36 (1.00, 1.82) | 0.10 |
| Log10 (Student number)\*State | 1.16 (0.77, 1.75) | 0.53 |

Negative binomial regression was used and the response variable was the “rate” i.e. the frequency of total tweets/affected days.

Table S8. Adjusted rate ratios of unplanned school closure announcements comparing GA and MA with log10 transformed student-teacher ratio variable.

|  |  |  |
| --- | --- | --- |
| Variable | Rate Ratios (95% CI) | P-value |
| State |  | <0.0001 |
|  Massachusetts | 0.26 (0.20, 0.35) |  |
|  Georgia | Reference |  |
| Log10 (Student-teacher ratio) | 0.90 (0.62, 1.27) | 0.58 |

Negative binomial regression was used and the response variable was the “rate” i.e. the frequency of total tweets/affected days.

Table S9. Adjusted rate ratios of unplanned school closure announcements comparing GA and MA with log10 transformed student-teacher ratio variable and interaction term.

|  |  |  |
| --- | --- | --- |
| Variable | Rate Ratios (95% CI) | P-value |
| State |  | 0.01 |
|  Massachusetts | 0.11 (0.01, 0.48) |  |
|  Georgia | Reference |  |
| Log10 (Student-teacher ratio) | 0.47 (0.08, 1.45) | 0.30 |
| Log10 (Student-teacher ratio)\*State  | 2.03 (0.62, 12.42) | 0.34 |

Negative binomial regression was used and the response variable was the “rate” i.e. the frequency of total tweets/affected days.

Table S10. Adjusted rate ratios of unplanned school closure announcements comparing GA and MA with log10 transformed following count variable.

|  |  |  |
| --- | --- | --- |
| Variable | Rate Ratios (95% CI) | P-value |
| State |  | <0.0001 |
|  Massachusetts | 0.26 (0.20, 0.33) |  |
|  Georgia | Reference |  |
| Log10 (Following count) | 1.30 (1.12, 1.51) | 0.0009 |

Negative binomial regression was used and the response variable was the “rate” i.e. the frequency of total tweets/affected days.

Table S11. Adjusted rate ratios of unplanned school closure announcements comparing GA and MA with log10 transformed following count variable and interaction term.

|  |  |  |
| --- | --- | --- |
| Variable | Rate Ratios (95% CI) | P-value |
| State |  | 0.004\* |
|  Massachusetts | 0.34 (0.17, 0.69) |  |
|  Georgia | Reference |  |
| Log10 (Following count) | 1.45 (1.08, 1.91) | 0.01\*  |
| Log10 (Following count)\*State | 0.86 (0.62, 1.21) | 0.40 |

Negative binomial regression was used and the response variable was the “rate” i.e. the frequency of total tweets/affected days.

Table S12. Adjusted rate ratios of unplanned school closure announcements comparing GA and MA with log10 transformed followers count variable.

|  |  |  |
| --- | --- | --- |
| Variable | Rate Ratios (95% CI) | P-value |
| State |  | <0.0001 |
|  Massachusetts | 0.40 (0.31, 0.51) |  |
|  Georgia | Reference |  |
| Log10 (Followers count) | 2.19 (1.80, 2.68) | <0.0001 |

Negative binomial regression was used and the response variable was the “rate” i.e. the frequency of total tweets/affected days.

Table S13. Rate ratios of unplanned school closure announcements comparing GA and MA with log10 transformed followers count variable and interaction term.

|  |  |  |
| --- | --- | --- |
| Variable | Rate Ratios (95% CI) | P-value |
| State |  | 0.04 |
|  Massachusetts | 0.30 (0.09, 1.04) |  |
|  Georgia | Reference |  |
| Log10 (Followers count) | 2.10 (1.61, 2.77) | 0.0007 |
| Log10 (Followers count)\*State  | 1.10 (0.74, 1.63) | 0.62 |

Negative binomial regression was used and the response variable was the “rate” i.e. the frequency of total tweets/affected days.

Table S14. Assessing for confounding for the compared GA and MA unplanned school closure announcements.

|  |  |  |
| --- | --- | --- |
|  | Adjusted Rate Ratios (95% CI) for the variable State (primary outcome of interest) | Percentage Change in the Point Estimate |
| Crude Rate Ratio | 0.25 (0.20, 0.33) |  |
| Potential confounding variable |  |  |
|  Log10 (Student number) | 0.33 (0.25, 0.44) | 20.03% |
|  Log10 (Student-teacher ratio) | 0.26 (0.20, 0.35) | 2.83% |
|  Log10 (Following count) | 0.26 (0.20, 0.33) | 2.83% |
|  Log10 (Followers count) | 0.40 (0.31, 0.51) | 33.90% |

The 10% rule was used in order to determine if a variable was a confounder in the relationship between state and tweet rate. The adjusted rate ratios for the variable State (Massachusetts, as compared to Georgia) were compared to the crude rate ratio of 0.25 (95% CI, 0.20, 0.33) using [ln(crude)-ln(adjusted)]/ln(crude). Log10 (Student number) and Log10 (Followers count) were found to be confounders of the relationship between state and unplanned school closure announcements during the winter storm affected days.

**Part IV: Supplementary Figures**

Figure S1. Distribution of Georgia (GA) school district Twitter accounts by frequency of tweets per account.



Figure S2. Distribution of Massachusetts (MA) school district Twitter accounts by frequency of tweets per account.

****

Figure S3. Correlation plot for the GA data



This plot includes the correlation coefficient, scatter plot, and histograms for each variable.

Figure S4. Log transformed correlation plot for the GA data



This plot includes the correlation coefficient, scatter plot, and histograms for each variable.

Figure S5. Correlation plot for the MA data



This plot includes the correlation coefficient, scatter plot, and histograms for each variable.

Figure S6. Log transformed correlation plot for the MA data



This plot includes the correlation coefficient, scatter plot, and histograms for each variable.