Climate Change and Physical Activity: Estimated Impacts of Ambient Temperatures on Bikeshare Usage in New York City

Alexandra K. Heaney, Daniel Carrión, Katrin Burkart, Corey Lesk, and Darby Jack

Table of Contents

Table S1. Sociodemographic characteristics of Citi Bike service area compared to NYC.

Table S2. Segmented regression estimates.

Table S3. The total number of Citi Bike rides stratified into gender and age groups.

Table S4. Total hours segmented regression results for men and women.

Table S5. Average distance segmented regression results for men and women.

Table S6. Total hours segmented regression results for weekday vs. weekend rides.

Table S7. Average distance segmented regression results for weekday vs. weekend rides.

Figure S1. Daily time series plots from July 2013 to September 2017. (a) maximum temperature, (b) minimum temperature, (c) average temperature, and (d) heat index.

Figure S2. Proportions of total hours ridden on Citi Bikes. (A) Proportion ridden on each day of the week. (B) Proportion ridden each hour of the day, shown for both weekdays and weekend days.

Figure S3. GAM and segmented regression results using average temperature. (a) Non-linear dose-response curve produced by GAM predicting total hours ridden using average temperature. The curve is shown as the effect of daily average temperature on differences between hours ridden and average hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily average temperature is represented by the red histogram. (b) The estimated segmented linear relationship between average temperature and total daily hours ridden. (c) and (d) as for (a) and (b) but predicting the daily average distance ridden.
**Figure S4. GAM and segmented regression results using minimum temperature.** (a) Non-linear dose-response curve produced by GAM predicting total hours ridden using minimum temperature. The curve is shown as the effect of daily minimum temperature on differences between hours ridden and average hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily minimum temperature is represented by the red histogram. (b) The estimated segmented linear relationship between minimum temperature and total daily hours ridden. (c) and (d) as for (a) and (b) but predicting the daily average distance ridden.

**Figure S5. GAM and segmented regression results using heat index.** (a) Non-linear dose-response curve produced by GAM predicting total hours ridden using heat index. The curve is shown as the effect of daily heat index on differences between hours ridden and average hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily average heat index is represented by the red histogram. (b) The estimated segmented linear relationship between average heat index and total daily hours ridden. (c) and (d) as for (a) and (b) but predicting the daily average distance ridden.

**Figure S6. GAM and segmented regression results using 2 degrees of freedom.** (a) Non-linear dose-response curve produced by GAM predicting total hours ridden using maximum temperature and a spline for time with 2 degrees of freedom. The curve is shown as the effect of daily maximum temperature on differences between hours ridden and average hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily average maximum temperature is represented by the red histogram. (b) The estimated segmented linear relationship between maximum temperature and total daily hours ridden. (c) and (d) as for (a) and (b) but predicting the daily average distance ridden.

**Figure S7. GAM and segmented regression results using 4 degrees of freedom.** (a) Non-linear dose-response curve produced by GAM predicting total hours ridden using maximum temperature and a spline for time with 4 degrees of freedom. The curve is shown as the effect of daily maximum temperature on differences between hours ridden and average hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily average maximum temperature is represented by the red histogram. (b) The estimated segmented linear relationship between maximum temperature and total daily hours ridden. (c) and (d) as for (a) and (b) but predicting the daily average distance ridden.

**Figure S8. GAM and segmented regression results using 8 degrees of freedom.** (a) Non-linear dose-response curve produced by GAM predicting total hours ridden using maximum temperature and a spline for time with 8 degrees of freedom. The curve is shown as the effect of daily maximum temperature on differences between hours ridden and average hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily average maximum temperature is represented by the red histogram. (b) The estimated segmented linear relationship between maximum temperature and total daily hours ridden. (c) and (d) as for (a) and (b) but predicting the daily average distance ridden.
Figure S9. GAM and segmented regression results using 10 degrees of freedom. (a) Non-linear dose-response curve produced by GAM predicting total hours ridden using maximum temperature and a spline for time with 10 degrees of freedom. The curve is shown as the effect of daily maximum temperature on differences between hours ridden and average hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily average maximum temperature is represented by the red histogram. (b) The estimated segmented linear relationship between maximum temperature and total daily hours ridden. (c) and (d) as for (a) and (b) but predicting the daily average distance ridden.

Figure S10. GAM regression results using maximum temperature lagged 0-2 day. (a) Non-linear dose-response curve produced by GAM predicting total hours ridden using maximum temperature unlagged. The curve is shown as the effect of daily maximum temperature on differences between hours ridden and average hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (b) and (c) as for (a) but using maximum temperature lagged 1 and 2 days. (d)-(f) as for (a)-(c) but predicting the daily average distance ridden.

Figure S11. Total hours GAM and segmented regression results for male subscribers. (A) Non-linear dose-response curve produced by GAM predicting total hours ridden using maximum temperature within the male subpopulation. The curve is shown as the effect of maximum temperature on differences between total hours and average total hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and total hours ridden in the male subpopulation. The threshold maximum temperature is represented by the dashed vertical line.

Figure S12. Total hours GAM and segmented regression results for female subscribers. (A) Non-linear dose-response curve produced by GAM predicting total hours ridden using maximum temperature within the female subpopulation. The curve is shown as the effect of maximum temperature on differences between total hours and average total hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and total hours ridden in the female subpopulation. The threshold maximum temperature is represented by the dashed vertical line.

Figure S13. Average distance GAM and segmented regression results for male subscribers. (A) Non-linear dose-response curve produced by GAM predicting average distance ridden using maximum temperature within the male subpopulation. The curve is shown as the effect of maximum temperature on differences between distance and average distance ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and average distance ridden in the male subpopulation. The threshold maximum temperature is represented by the dashed vertical line.
Figure S14. Average distance GAM and segmented regression results for female subscribers. (A) Non-linear dose-response curve produced by GAM predicting average distance ridden using maximum temperature within the female subpopulation. The curve is shown as the effect of maximum temperature on differences between distance and average distance ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and average distance ridden in the female subpopulation. The threshold maximum temperature is represented by the dashed vertical line.

Figure S15. Total hours GAM and segmented regression results for subscribers ages 18-25. (A) Non-linear dose-response curve produced by GAM predicting total hours ridden using maximum temperature within the subpopulation aged 18-25. The curve is shown as the effect of maximum temperature on differences between total hours and average total hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and total hours ridden in the subpopulation aged 18-25. The threshold maximum temperature is represented by the dashed vertical line.

Figure S16. Total hours GAM and segmented regression results for subscribers ages 25-35. (A) Non-linear dose-response curve produced by GAM predicting total hours ridden using maximum temperature within the subpopulation aged 25-35. The curve is shown as the effect of maximum temperature on differences between total hours and average total hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and total hours ridden in the subpopulation aged 25-35. The threshold maximum temperature is represented by the dashed vertical line.

Figure S17. Total hours GAM and segmented regression results for subscribers ages 35-45. (A) Non-linear dose-response curve produced by GAM predicting total hours ridden using maximum temperature within the subpopulation aged 35-45. The curve is shown as the effect of maximum temperature on differences between total hours and average total hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and total hours ridden in the subpopulation aged 35-45. The threshold maximum temperature is represented by the dashed vertical line.

Figure S18. Total hours GAM and segmented regression results for subscribers ages 45-55. (A) Non-linear dose-response curve produced by GAM predicting total hours ridden using maximum temperature within the subpopulation aged 45-55. The curve is shown as the effect of maximum temperature on differences between total hours and average total hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and total hours ridden in the subpopulation aged 45-55. The threshold maximum temperature is represented by the dashed vertical line.
Figure S19. Total hours GAM and segmented regression results for subscribers ages 55-65. (A) Non-linear dose-response curve produced by GAM predicting total hours ridden using maximum temperature within the subpopulation aged 55-65. The curve is shown as the effect of maximum temperature on differences between total hours and average total hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and total hours ridden in the subpopulation aged 55-65. The threshold maximum temperature is represented by the dashed vertical line.

Figure S20. Total hours GAM and segmented regression results for subscribers ages 65+. (A) Non-linear dose-response curve produced by GAM predicting total hours ridden using maximum temperature within the subpopulation aged 65+. The curve is shown as the effect of maximum temperature on differences between total hours and average total hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and total hours ridden in the subpopulation aged 65+. The threshold maximum temperature is represented by the dashed vertical line.

Figure S21. Average distance GAM and segmented regression results for subscribers ages 18-25. (A) Non-linear dose-response curve produced by GAM predicting average distance ridden using maximum temperature within the subpopulation aged 18-25. The curve is shown as the effect of maximum temperature on differences between distance and average distance ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and average distance ridden in the subpopulation aged 18-25. The threshold maximum temperature is represented by the dashed vertical line.

Figure S22. Average distance GAM and segmented regression results for subscribers ages 25-35. (A) Non-linear dose-response curve produced by GAM predicting average distance ridden using maximum temperature within the subpopulation aged 25-35. The curve is shown as the effect of maximum temperature on differences between distance and average distance ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and average distance ridden in the subpopulation aged 25-35. The threshold maximum temperature is represented by the dashed vertical line.

Figure S23. Average distance GAM and segmented regression results for subscribers ages 35-45. (A) Non-linear dose-response curve produced by GAM predicting average distance ridden using maximum temperature within the subpopulation aged 35-45. The curve is shown as the effect of maximum temperature on differences between distance and average distance ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and average distance ridden in the subpopulation aged 35-45. The threshold maximum temperature is represented by the dashed vertical line.
Figure S24. Average distance GAM and segmented regression results for subscribers ages 45-55. (A) Non-linear dose-response curve produced by GAM predicting average distance ridden using maximum temperature within the subpopulation aged 45-55. The curve is shown as the effect of maximum temperature on differences between distance and average distance ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and average distance ridden in the subpopulation aged 45-55. The threshold maximum temperature is represented by the dashed vertical line.

Figure S25. Average distance GAM and segmented regression results for subscribers ages 55-65. (A) Non-linear dose-response curve produced by GAM predicting average distance ridden using maximum temperature within the subpopulation aged 55-65. The curve is shown as the effect of maximum temperature on differences between distance and average distance ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and average distance ridden in the subpopulation aged 55-65. The threshold maximum temperature is represented by the dashed vertical line.

Figure S26. Average distance GAM and segmented regression results for subscribers ages 65+. (A) Non-linear dose-response curve produced by GAM predicting average distance ridden using maximum temperature within the subpopulation aged 65+. The curve is shown as the effect of maximum temperature on differences between distance and average distance ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and average distance ridden in the subpopulation aged 65+. The threshold maximum temperature is represented by the dashed vertical line.

Figure S27. Total hours GAM and segmented regression results for weekday rides. (A) Non-linear dose-response curve produced by GAM predicting total hours ridden using maximum temperature for weekday rides. The curve is shown as the effect of maximum temperature on differences between total hours and average total hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and total hours ridden for weekday rides. The threshold maximum temperature is represented by the dashed vertical line.

Figure S28. Total hours GAM and segmented regression results for weekend rides. (A) Non-linear dose-response curve produced by GAM predicting total hours ridden using maximum temperature for weekend rides. The curve is shown as the effect of maximum temperature on differences between total hours and average total hours ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and total hours ridden for weekend rides. The threshold maximum temperature is represented by the dashed vertical line.
Figure S29. Average distance GAM and segmented regression results for weekday rides.
(A) Non-linear dose-response curve produced by GAM predicting average distance ridden using maximum temperature for weekday rides. The curve is shown as the effect of maximum temperature on differences between distance and average distance ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and average distance ridden for weekday rides. The threshold maximum temperature is represented by the dashed vertical line.

Figure S30. Average distance GAM and segmented regression results for weekend rides.
(A) Non-linear dose-response curve produced by GAM predicting average distance ridden using maximum temperature for weekend rides. The curve is shown as the effect of maximum temperature on differences between distance and average distance ridden (left y-axis, black) with 95% confidence intervals (gray). Density of daily maximum temperature is represented by the red histogram. (B) shows the estimated segmented linear relationship between maximum temperature and average distance ridden for weekend rides. The threshold maximum temperature is represented by the dashed vertical line.