Supplemental Materials for *Fine-scale Twitter Big Data Reveals Neighborhood Inequalities in Mental Health Effects of High Temperatures*

S1 Alternative Sentiment Algorithms

S1.1 Hedonometer

In addition to the Valence Aware Dictionary for sentiment Reasoning (VADER) sentiment analysis method [**gilbert˙vader˙2014**], we also used the hedonometer method [**Dodds2011Dec**]. This sentiment metric yielded very similar findings to those presented in the main paper with the VADER method.

The Hedonometer [**dodds˙temporal˙2011**] is a corpus-based technique to get sentiment score from multi-lingual texts. The core steps are (1) building human evaluations of the happiness of a set of individual words, and (2)using a naive algorithm for scaling up from individual words to texts. For the English corpus, Dodds and others [**dodds˙temporal˙2011**] collect 10,222 unique words and used crowd-sourcing platform Amazon Mechanical Turk to get human evaluation of happiness degree for each word in an integer scale from 1 to 9, representing a sad to happy spectrum. Score 5 represents neutral words. Each word will be calculated average score and then the word and happiness score are compiled into a dataset (labMT 1.0). Some illustrative example of words are:

$$h_{avg}(\text{laughter}) = 8.50$$

$$h_{avg}(\text{the}) = 4.98$$

$$h_{avg}(\text{hate}) = 2.34$$

The sentiment score for single text will be the mean happiness score for all words in the text.
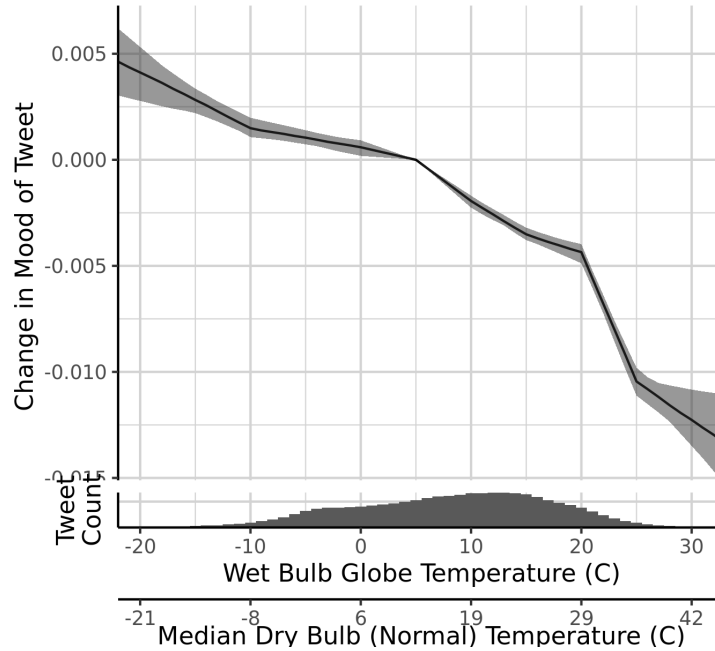

Figure S1: Effects of temperature on sentiment, measured with the hedonometer method.
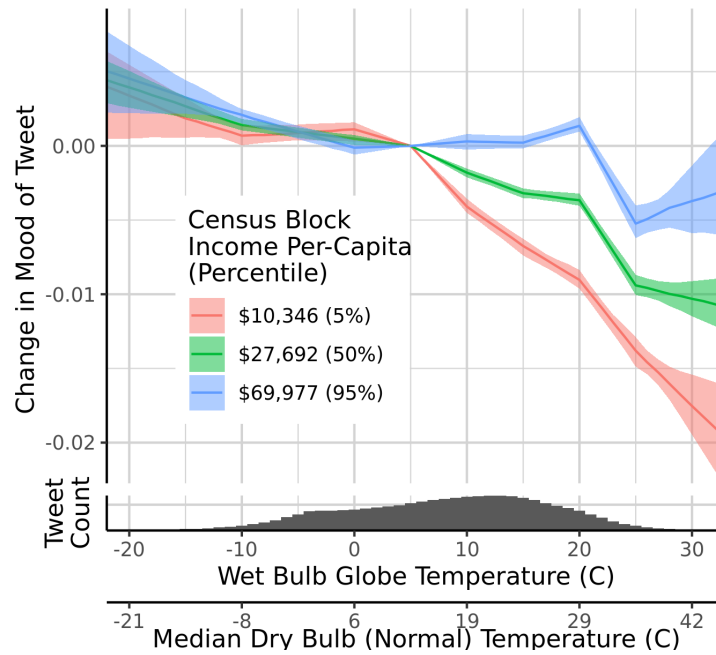
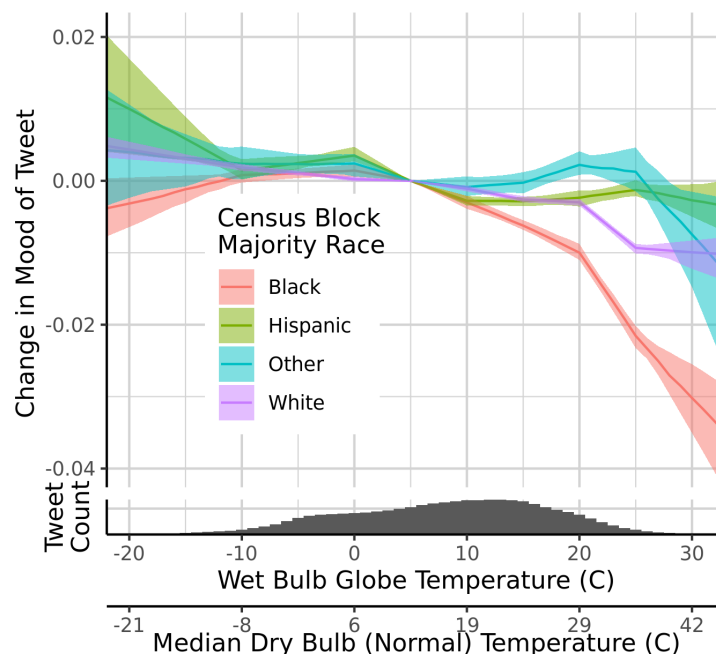Figure S2: Effects of temperature on sentiment and moderated by income, measured with the hedonometer method.

Figure S3: Effects of temperature on sentiment and moderated by race, measured with the hedonometer method.

S1.2 AFINN

S2 Alternative Interaction Variables

For our main analysis, we treated income as a continuous (log-transformed) variable and race as a categorical variable. We ran similar analyses with income as a categorical variable and race as a continuous variable (the percent of a census block's population that was non-white/minority).
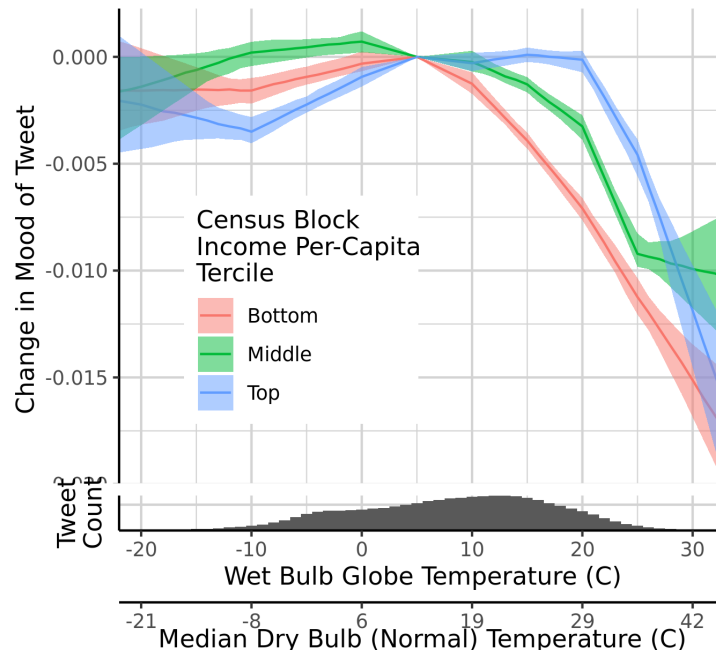
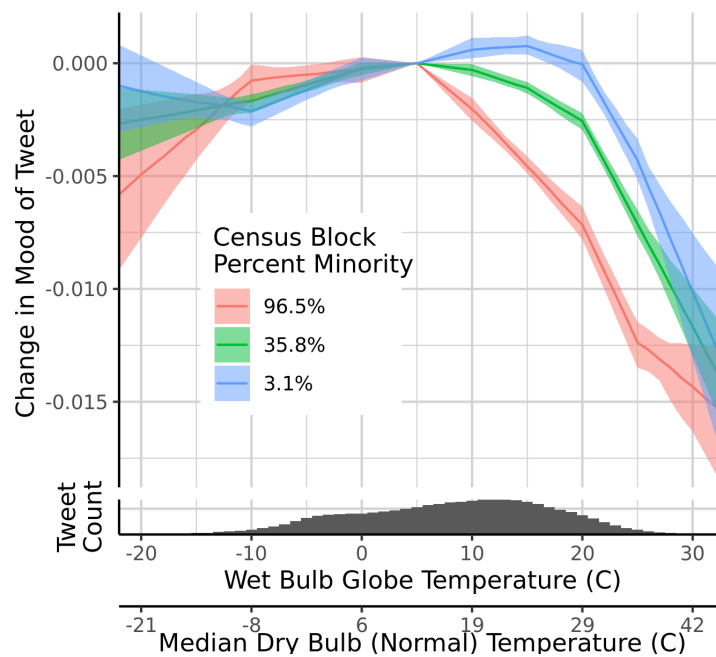Figure S4: Effect of temperature on sentiment and moderated by income, with income as a categorical variable.

Figure S5: Effect of temperature on sentiment and moderated by race, with race as a continuous variable.

S3 Effects of Precipitation and Sunshine

While our main variable of interest was temperature as measured by Wet Bulb Globe Temperature, we also interacted our variables of interest with a binary variable for whether it was raining or not, as well as a continous variable for the amount of shortwave radiation in (W/m^2), or sunshine.
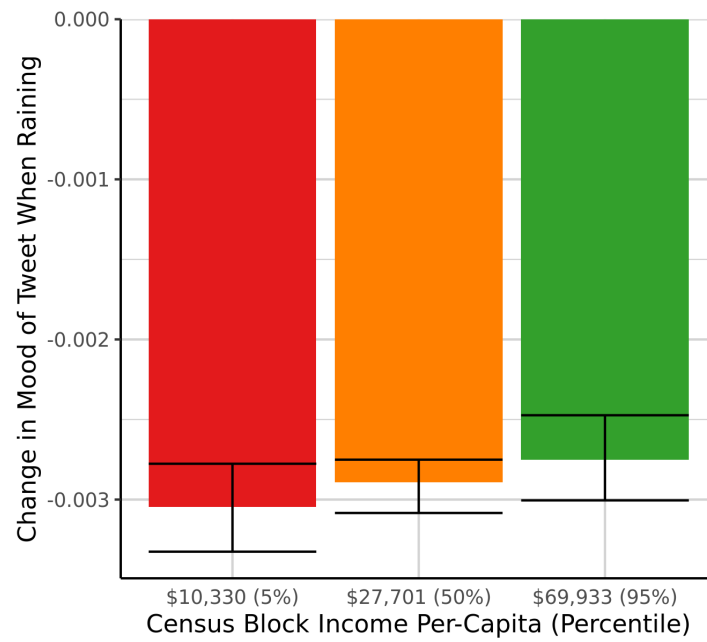
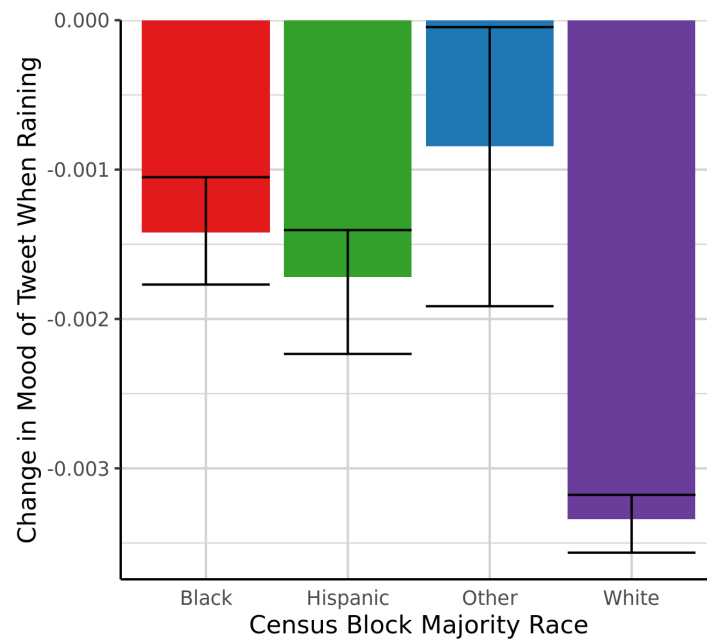Figure S6: Effect of rain on sentiment and moderated by income.

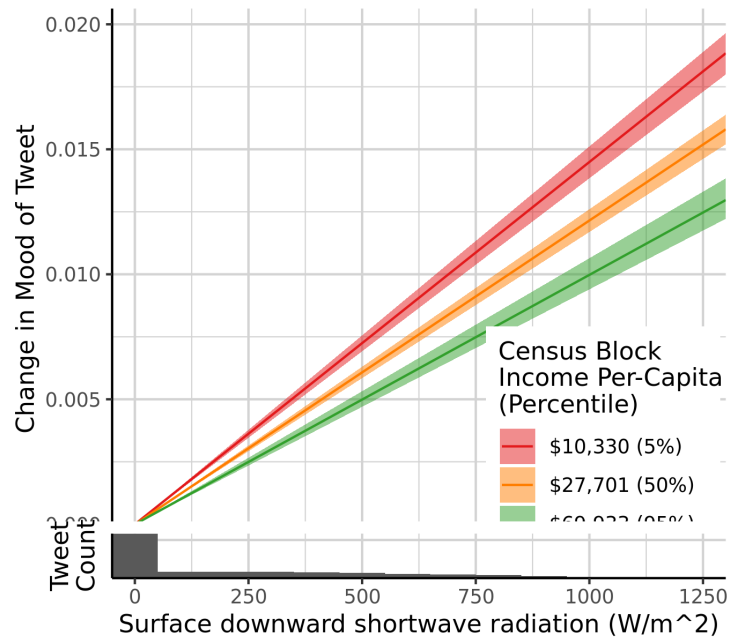Figure S7: Effect of rain on sentiment and moderated by race.

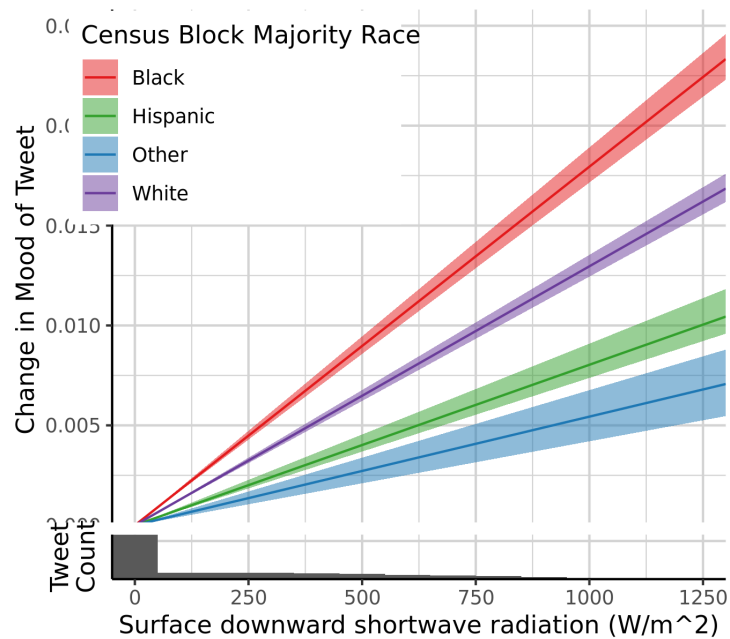Figure S8: Effect of sunshine on sentiment and moderated by income.

Figure S9: Effect of sunshine on sentiment and moderated by race.

S4 Combined Effects of Race and Income

In addition to our models exploring how income and race moderate the effect of heat on sentiment, we fit an additional model with an interaction terms for both race and income. Due to the complexity of a model with spline terms interacting with two different variables, using a categorical variable for race would have yielded a model matrix too large for the largest cloud virtual machine availale to us. Thus, we used a continuous metric for both income as well as race (percent non-white/minority).

To examine how income and racial groups together moderate the effect of heat on sentiment, we

extend our main model to the following form:

$$y = \begin{aligned}&\beta_0 + f_t(t) + f_{m_it}(m_it) + f_{m_rt}(m_rt) + \beta_p p + \beta_{m_rp}m_rp + \beta_{m_ip}m_ip + \beta_{m_im_rp}m_im_rp + \\ &\beta_s s + \beta_{m_is}m_is + \beta_{m_rs}m_rs + \beta_{m_im_rs}m_im_rs + \Phi + \epsilon\end{aligned} \tag{1}$$

Identically to the model in the main body of the paper, y is the sentiment of a tweet, t is the wet bulb globe temperature at the hour of the tweet, p is a binary variable indicating whether it rained at the hour of the tweet, s is the income shortwave radiation, or sunshine, in W/m^2, at the hour of the tweet, Φ is the spatio-temporal fixed effects, ϵ is the normally-distributed errors, and f_t represented the segmented regression for t, with knots at -10°, 0°, 5°, 10°, 15°, 20°, and 25°C WBGT. We estimate the 95% confidence interval of our models using 80-fold bootstrapping.

In this new model, m_i is the log-transformed average income in the census block where the tweet originated, while m_r is the racial composition of the census block, in terms of the percent of the population that was a minority.
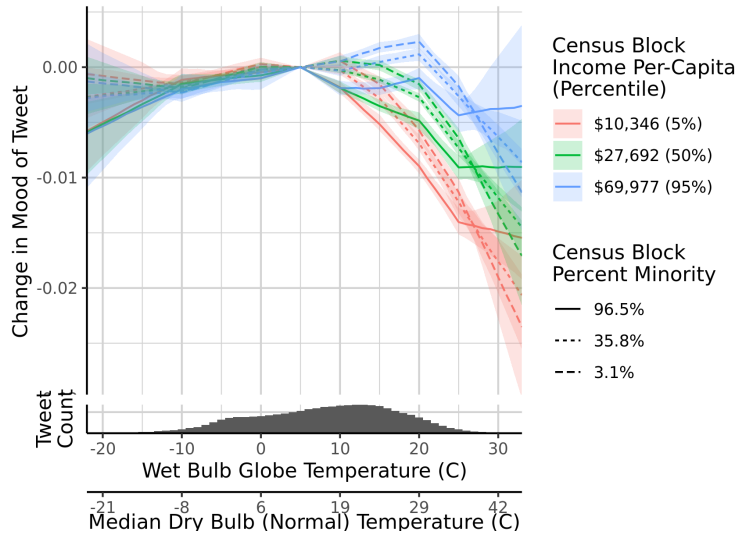

Figure S10: Effect of temperature on sentiment and moderated by both race and income.

S5 Weather Terms

We excluded tweets that contained the following weather-related terms: aerovane, air, airstream, altocumulus, altostratus, anemometer, anemometers, anticyclone, anticyclones, arctic, arid, aridity, atmosphere, atmospheric, autumn, autumnal, balmy, baroclinic, barometer, barometers, barometric, blizzard, blizzards, blustering, blustery, blustery, breeze, breezes, breezy, brisk, calm, celsius, chill, chilled, chillier, chilliest, chilly, chinook, cirrocumulus, cirrostratus, cirrus, climate, climates, cloud, cloudburst, cloudbursts, cloudier, cloudiest, clouds, cloudy, cold, colder, coldest, condensation, contrail, contrails, cool, cooled, cooling, cools, cumulonimbus, cumulus, cyclone, cyclones, damp, damp, damper, damper, dampest, dampest, degree, degrees, deluge, dew, dews, dewy, doppler, downburst, downbursts, downdraft, downdrafts, downpour, downpours, dried, drier, dries, driest, drizzle, drizzled, drizzles, drizzly, drought, droughts, dry, dryline, fall, farenheit, flood, flooded, flooding, floods, flurries, flurry, fog, fogbow, fogbows, fogged, fogging, foggy, fogs, forecast, forecasted, forecasting, forecasts, freeze, freezes, freezing, frigid, frost, frostier, frostiest, frosts, frosty, froze, frozen, gale, gales, galoshes, gust, gusting, gusts, gusty, haboob, haboobs, hail, hailed, hailing, hails, haze, hazes, hazy, heat, heated, heating, heats, hoarfrost, hot, hotter, hottest, humid, humidity, hurricane, hurricanes, ice, iced, ices, icing, icy, inclement, landspout, landspouts, lightning, lightnings, macroburst, macrobursts, maelstrom, mercury, meteorologic, meteorologist, meteorologists, meteorology, microburst, microbursts, microclimate, microclimates, millibar, millibars, mist, misted, mists, misty, moist, moisture, monsoon, monsoons, mugginess, muggy, nexrad, nippy, NOAA, nor,'easter, nor,'easters, noreaster, noreasters, overcast, ozone, parched, parching, pollen, precipitate, precipitated, precipitates, precipitating, precipitation, psychrometer, radar, rain, rainboots, rainbow, rainbows, raincoat, raincoats, rained, rainfall, rainier, rainiest, raining, rains, rainy, sandstorm, sandstorms, scorcher, scorching, searing, shower, showering, showers, skiff, sleet, slicker,

slickers, slush, slushy, smog, smoggier, smoggiest, smoggy, snow, snowed, snowier, snowiest, snowing, snowmageddon, snowpocalypse, snows, snowy, spring, sprinkle, sprinkles, sprinkling, squall, squalls, squally, storm, stormed, stormier, stormiest, storming, storms, stormy, stratocumulus, stratus, subtropical, summer, summery, sun, sunnier, sunniest, sunny, temperate, temperature, tempest, thaw, thawed, thawing, thaws, thermometer, thunder, thundered, thundering, thunders, thunderstorm, thunderstorms, tornadic, tornado, tornadoes, tropical, troposphere, tsunami, turbulent, twister, twisters, typhoon, typhoons, umbrella, umbrellas, vane, warm, warmed, warming, warms, warmth, waterspout, waterspouts, weather, wet, wetter, wettest, wind, windchill, windchills, windier, windiest, windspeed, windy, winter, wintery, and wintry.