Journal of Public Economics 184 (2020) 104161

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Temperature and temperament: Evidence from Twitter☆

Patrick Baylis

University of British Columbia, Vancouver School of Economics, Iona 113, 6000 Iona Dr., Vancouver, BC V6T 2E8, Canada

ARTICLE INFO

Article history:
Received 15 May 2019
Received in revised form 30 December 2019
Accepted 18 February 2020
Available online 16 March 2020

ABSTRACT

How do people value their climate? This paper demonstrates a new approach to estimating preferences for nonmarket goods using social media data. I combine more than a billion Twitter updates with natural language processing algorithms to construct a rich panel dataset of expressed sentiment for the United States and six other English-speaking countries around the world. In the U.S., I find consistent and statistically significant declines in expressed sentiment from both hot and cold temperatures. To better understand how preferences may adapt, I document heterogeneity in both regional and seasonal responses. I complete the U.S. analysis with a suite of validation exercises to understand the magnitude of these effects and two methods to estimate willingness-to-pay for climate amenities. Finally, I document similar relationships between temperature and expressed sentiment for four out of the six non-U.S. countries I examine.

As the possibility of substantial changes to Earth's climate has become more certain, economists have grown increasingly interested in calculating the full scope of benefits and costs resulting from these changes. While much of the literature has focused on changes in welfare due to the effects of increases in temperature on aggregate incomes, morbidity and mortality, civil conflict, and agricultural profits, among others, relatively little work has examined individuals' underlying preferences for different climates and, by extension, the amenity value implications of climate change itself. This gap is the result of a classic problem in nonmarket valuation: because ambient temperature is non-rival and non-excludable, there are no direct markets from which researchers might infer preferences for different climates.

Instead, the existing literature has relied on hedonic valuation approaches to indirectly estimate individuals' willingness-to-pay for different climate characteristics by observing how housing prices vary with climatic conditions. These approaches generally find that individuals would pay between 1% and 4% of their annual incomes to avoid projected end-of-century increases in temperature (Sinha et al., 2018; Albouy et al., 2016). However, because the climate to date has varied relatively little across time, these values are necessarily identified using cross-sectional differences in climate. Preferences for climate that are robust to potential cross-sectional contamination are key parameters for the development of local, national, and global public policies relating to climate change.

This paper demonstrates a new method to estimate preferences over nonmarket goods that allows researchers to include controls for unobservables across both time and space: I construct a spatially and temporally rich dataset on daily expressed sentiment, or emotional state, and estimate the relationship between sentiment and outdoor ambient temperature.

Section 1 discusses how this approach relates to existing methods designed to elicit preferences for nonmarket goods. Section 2 describes the construction of the dataset, which begins with a geographically and temporally dense collection of more than a billion geocoded social media updates (hereafter, "tweets") from the online social media platform Twitter. I measure the expressed sentiment of each tweet using a set of natural language processing (NLP) algorithms designed to extract sentiment, or emotional state, from unstructured text data. For computational tractability and to account for noise in the estimation of expressed sentiment, the primary analysis takes daily averages for each Core-Based Statistical Area (hereafter, CBSAs) as the unit of observation, although I also estimate a model with individual tweets as the unit of observation to test for compositional effects. Because of the uncertainty inherent in estimating underlying emotional state from language, I compile four separate measures of sentiment using word lists constructed using previous research in NLP and, in three of four cases, specifically intended to extract sentiment from "microblogs" such as tweets.

☆ I am grateful to Maximilian Auffhammer, Severin Borenstein, and Solomon Hsiang for their invaluable suggestions, as well as to Michael Anderson, Hunt Allcott, Judson Boomhower, Josh Blonz, Marshall Burke, Fiona Burlig, Tamma Carleton, Richard Carson, Aluma Dembo, Meredith Fowlie, Walter Graf, Sylvan Herskowitz, Nick Obradovich, Elizabeth Sadoulet, seminar participants at UC Berkeley, the AERE Annual Conference, the CU Environmental and Resource Economics Workshop, the Heartland Workshop at Illinois, and several anonymous referees. This paper was previously circulated under the title "Temperature and Temperament: Evidence from a billion tweets."

E-mail address: patrick.baylis@ubc.ca.

https://doi.org/10.1016/j.jpubeco.2020.104161
0047-2727/

The analysis in Section 3 then uses the geographic information attached to the tweets in my dataset to match measures of sentiment to daily weather conditions at the location of the user. The identifying assumption in the econometric model I estimate is that temperature realizations are as good as random after accounting for spatial and temporal fixed effects. Allowing temperature to enter the model flexibly, I find consistent evidence of an upside-down "U" shape: a roughly symmetric decline in sentiment away from moderate temperatures, with peak sentiment occurring roughly around 21.0 C (69.8 F). The point estimate of the difference in expressed sentiment between 20–25 C and above 40 C is a statistically significant and between 0.1 and 0.2 standard deviations (SD), depending on measure used. The responses of expressed sentiment to temperature are markedly similar across choice of measure, and both qualitatively and quantitatively consistent across a range of different specification choices. To discern the mechanism by which sentiment responds to temperature, I also estimate the relationship between online profanity and temperature and find a U-shaped relationship there as well, suggesting that aggression is at least part of the explanation for the decline in expressed sentiment in both hot and cold temperatures.

To better understand how preferences for climate are formed and updated, I extend the baseline results by examining regional and seasonal heterogeneity in the response of sentiment to temperature. I find notable differences in both. Regionally, areas that are colder tend to have stronger response to warm temperatures, and vice-versa. Seasonally, the responses suggest preferences for cooler temperatures in summer and fall and warmer temperatures in winter, with relatively little sentiment response to temperatures on the spring.

Section 4 documents a set of validation and valuation exercises designed to aid interpretation of the main results, as well as a projection exercise to estimate future damages. I begin by showing how average sentiment changes over the course of the week, and that the difference in expressed sentiment between a Sunday and Monday is roughly 0.1 SD. Next, I present additional empirical exercises identifying the effect of plausibly random variation in hurricanes, American football outcomes, quarterly wages, and receipt of parking or speeding tickets on observed social media sentiment. I conclude by using the value of expressed sentiment from the last two exercises to value of the amenity losses due to climate change, which I find to be between 0.6% and 2.5% of present-day annual income by end of century.

Finally, Section 5 estimates the relationship between temperature and expressed sentiment in Australia, India, South Africa, the Philippines, Kenya, and Uganda, six English-speaking countries for which I am able to obtain sufficient data from Twitter and with adequate temperature variation. Compared to the United States, I estimate similar preferences for temperature in Australia, India, South Africa, and the Philippines, but not in Kenya or Uganda, suggesting that the effects I estimate are not isolated to the unique case of the United States, but that caution should be taken when extrapolating these results to the entire world.

This paper makes several contributions to the literature. It is the first to identify sentiment-analyzed social media posts as a source of information on latent individual preferences for environmental goods.[1] I identify a response function of sentiment to temperature that concurs qualitatively with existing work and is robust to a wide range of statistical specifications. The methods I document provide a tractable roadmap for future work estimating preferences for nonmarket goods from social media posts. Second, the paper identifies sharply diverging seasonal and regional preferences for temperature, suggesting an important adaptive component to the baseline observed response. Third, the set of validation and valuation exercises I employ serve as novel empirical exercises in their own right and provide initial steps in valuing the impacts observed here. Fourth and finally, the estimates of similar patterns in four out of the six English-speaking countries are the first of their kind and suggest that the responses to climate that I observe may well be global.

1. Background

Economists have studied the economic impacts of climate change for more than two decades (Nordhaus, 1991; Cline, 1992), but the increasing availability of a range of panel datasets have made possible the identification of the causal effects of changes in temperature on a diverse set of economic outcomes, including crop yields, economic production, civil conflict, mortality, migration, and many others (Carleton and Hsiang, 2016). In the absence of historical changes in long-run climate, researchers have used estimates of the changes in these outcomes resulting from plausibly exogenous historical variation in temperature to predict future damages from climate change (Dell et al., 2014). The assumptions required for this extrapolation are formally derived in Hsiang (2016), but intuitively the central consideration is whether or not adaptation behaviors require large fixed costs. This question is likely to be answered differently for different sectors, but for those in which it is econometrically possible to estimate a "long-differences" approach, e.g., Burke and Emerick (2015), few differences have been found between estimates produced using short-run or long-run variation in temperature.

This work has had an impact on public policy. Many of the estimated outcomes contribute, directly or indirectly, to aggregate measures of the total cost of climate change produced by summary reports (Stern, 2006; Houser et al., 2014) and integrated assessment models (IAMs), which in turn are inputs to the United States government's estimate of the social cost of carbon, or SCC (Interagency Working Group on Social Cost of Carbon, 2013).[2] By 2014, the then-central value of $36 per ton of CO_2 equivalent had been incorporated into 79 U.S. regulations as part of required benefit-cost analyses conducted in the course of the federal rulemaking (United States Government Accountability Office, 2014).[3]

Different areas of the world will experience climate change in very different ways. Coastal areas will face rising sea levels and major economic impacts from typhoons or hurricanes (Hsiang, 2010). Farmers are likely to experience substantial changes in the yields of major crops (Schlenker and Roberts, 2009). Areas in the developing world where subsistence farming is a major source of calories could experience catastrophic droughts and resulting food security crises (Lobell et al., 2008). For other areas, the impacts of climate change will be more subtly felt: instead of increases in large-scale natural disasters or acute economic crises, most of the world will simply experience a steady increase in average temperatures (IPCC, 2014). Prior work has projected the impact of these gradual changes on income (Deryugina and Hsiang, 2014), crime (Ranson, 2014), mortality (Deschênes and Greenstone, 2011), and other outcomes. This paper focuses instead on the welfare cost of changes in amenity values resulting from rising outdoor temperatures.[4]

Traditional approaches to calculating the welfare impact of a policy change date back as far as Marshall (1890) and rely on knowledge of either the demand curve, the supply curve, or both. For private goods with well-established markets, the shapes of these curves can be estimated using plausibly exogenous supply or demand shifters and from these curves the in policy can be calculated. Estimating changes in welfare due to changes in the allocation of public goods, or nonmarket goods more generally, has proven to be more challenging due to the absence

[1] In a paper drafted after this one, we extend this method to include Facebook posts as well (Baylis et al., 2018).

[2] Three IAMs are used to derive the social cost of carbon: DICE, FUND, and PAGE. Diaz (2014) provides a more detailed comparison of how these models are built.

[3] The Environmental Protection Agency, under direction from the Trump administration, reduced the SCC to between $1 and $6 per ton, largely because it used only damages to the United States in its accounting for the costs of carbon emissions.

[4] To date, only the DICE model directly incorporate estimates of the costs or benefits from climate as an amenity. Nordhaus and Boyer (2000) finds that 2.5 C warming will result in a *gain* of about 0.3% of GDP.

of available markets. Nevertheless, a handful of approaches to this problem have emerged, many within the environmental economics literature (Pearce, 2002).

Climate can be viewed as a public amenity[5]: it is non-rival (a single person's consumption of climate does not reduce the amount of climate available to anyone else) and non-excludable (no person cannot be prevented from consuming climate), and although individuals can alter their local climates at home and at work, the outdoor ambient temperature is determined by factors outside of their control. Hedonic price and discrete choice methods provide a method to value amenities like climate: recent works by Sinha et al. (2018) and Albouy et al. (2016) identify implicit values for different climates using observed household decisions about where to live. These approaches are useful in part because it is straightforward to back out monetary valuations of different climates from the model estimates. However, since historical changes in climate have thus far been fairly modest, the estimates from these models must be identified using cross-sectional variation. As a result, unobserved spatial variation such as cultural norms, geographic factors like proximity to oceans or mountains, or other unobserved amenities that correlate with climate could bias these estimates in unknown directions.

A related approach to understanding preferences is to use surveys of subjective well-being (SWB) to estimate preferences over temperature. These surveys ask respondents to assess their well-being on a single dimensional scale (Diener, 2000; Dolan et al., 2008). Kahneman and Krueger (2006) and Mackerron (2012) discuss the merits and weaknesses of these studies: a common challenge is that measurements of SWB are, by definition, subjective and likely to include unobserved variation across time and space. For example, responses to questions about one's well-being may depend on regional dialects or norms, or could be driven by the interaction between the interviewer and the interviewee, which may itself be affected by temperature.

The estimates of the effect of temperature on SWB vary widely within the literature. Most studies use cross-sectional variation or follow a very small group of individuals over time. To my knowledge, only two studies control for unobservable cross-sectional variation using panel data methods. Feddersen et al. (2012) use nearly 100,000 observations from Australian SWB surveys to compare the effects of short-term weather and long-term climate on life satisfaction. Since individuals are observed more than once in their data, they are able to control for individual fixed effects for some specifications. They find that weather affects reported life satisfaction through solar exposure, barometric pressure, and wind speed, but do not find impacts from changes in temperature itself. Dennisenn et al. (2008) use an online survey to find that weather impacts are variable across individuals, but that those variations do not correspond to observable characteristics.

A small literature attempts to assign a monetary value to environmental goods using self-reported happiness data (Welsch and Kühling, 2009). For example, Rehdanz and Maddison (2005) estimate the relationship between climate and self-reported happiness, and include a valuation method based on country-level GDP. Levinson (2012) conducts a similar exercise to estimate willingness-to-pay to avoid pollution using happiness data, but includes weather as a covariate. Because these studies implicitly rely on income as an exogenous driver of happiness, this approach could induce bias if that assumption does not hold (Mackerron, 2012).

The method of assessing preferences for nonmarket goods I describe in this paper relies on the assumption that contemporaneous changes in expressed sentiment deliver insights into individuals' underlying preferences for these goods. I have described previous work that assesses these preferences and the challenges researchers face in controlling for unobservable sources of cross-sectional variation. The approach in this paper mitigates the problem of unobserved correlates over time and space, allows for flexible estimation of non-linear effects, and provides sufficient data richness to examine geographic and seasonal variation in the response of expressed sentiment to temperature change in order to better understand adaptation.

Conceptually, one way to view expressed sentiment is as an estimate of “experienced utility”. The concept of experienced utility predates the modern neoclassical definition of utility, which for clarity and following Kahneman and Sugden (2005), I refer to hereafter as “decision utility”. Whereas decision utility is an ordinal description of the value obtained from bundles of goods, experienced utility follows is an instantaneous measure of pleasure and pain (Bentham, 1789). Discussions of which measure of utility is the appropriate metric for welfare analyses is beyond the scope of this paper, but for the purposes here it is sufficient to view expressed sentiment, like subjective well-being or experienced utility, as a useful proxy for individual preferences. In the following section, I describe how this paper estimates expressed sentiment and document descriptive statistics that suggest its relationship to underlying preferences.

2. Data

While it would be prohibitively expensive to estimate daily sentiment across the United States using a survey, publicly available updates on social media provide a low-cost alternative. By combining a large set of geolocated tweets with sentiment analysis algorithms (NLP algorithms designed to reveal emotional state), I am able to measure daily variation in expressed sentiment across the United States. In this paper, I combine this data with meteorological observations to estimate the sentiment response to temperature. Previous work in computer science has estimated models that related expressed sentiment to meteorological variables (Hannak et al., 2012), but, to the best of my knowledge, this is the first paper to do so in a causal framework and in order to elicit underlying preferences for temperature. The following section describes the construction of the measures of sentiment and the weather covariates. Table 1 summarizes the variables included in the empirical model. The first panel shows the count, mean, median, minimum, and maximum of the measures of sentiment, the second panel describes the weather data used, and the third panel summarizes the number of tweets by CBSAs and by individuals in the data.

2.1. Twitter data

Created in 2006, Twitter is a social media platform where users exchange brief updates, otherwise known as tweets. Since its founding, Twitter has become one of the most popular such platforms worldwide, with 288 million active users sending over 500 million tweets per day as of 2015.[6]

Twitter's Streaming API is designed to give developers access to the massive amount of data generated on the Twitter platform in real time. Starting in June 2014, I began collecting geolocated Twitter updates from within the continental United States using a client that is continuously connected to the Streaming API.[7] I collect the vast majority of geolocated tweets matching my specifications and produced within my sample period, which ends in October 2016.

Geolocated tweets are those for which the user has consented to have his or her location information shared for that post. The location information is either produced using the exact latitude and longitude or from a reverse-geocoding algorithm that derives the latitude and longitude from a general location (e.g., a neighborhood) entered by the

[5] I use the term “public amenities” here to indicate that changes in the climate can be either goods or bads, depending on the region and sector of the world in which they occur.

[6] Per Twitter's website, accessed September 2015.

[7] More details on the collection process are given in Appendix A.1.1. There are two substantial gaps in the time series, from June 26th to July 12th, 2014, and from September 18th to October 27th, 2014, and a small number of gaps of a few days. These gaps correspond to periods of time when the streaming client was unable to connect to the Streaming API.

Table 1
Sample characteristics.

	Count	Mean	Median	Min	Max
A: Sentiment measures					
AFINN-111	598,750,185	0.5	0.4	−5.0	4.0
Hedonometer	1,061,098,510	5.5	5.5	2.6	8.3
LIWC	1,092,491,096	0.3	0.2	−5.0	5.0
Vader	1,160,617,577	0.1	0.1	−1.0	1.0
B: Weather covariates					
Minimum temperature (C)	1,160,617,577	9.4	11.3	−34.7	31.9
Maximum temperature (C)	1,160,617,577	21.2	24.2	−23.5	47.4
Precipitation (mm)	1,160,617,577	2.9	0.0	0.0	434.2
C: Twitter updates per...					
CBSA	908	1,395,932	205,419	15,224	78,506,336
User	11,659,619	106	11	1	17,957

Notes: First panel summarizes unstandardized measures of expressed sentiment: AFINN-111, Hedonometer, LIWC, and VADER. Second panel summarizes weather covariates obtained from PRISM. Third panel summarizes the number of tweets per CBSA and user.

user. Geolocations are assigned on a by-post basis. In principle, Twitter limits the total number of tweets delivered to any client through the Streaming API to 1% (Morstatter et al., 2013) of the total tweets created. Since I request only geolocated tweets from within the United States, this total infrequently comes to more than 1% of the total tweets worldwide (geocoded and otherwise). As a result, over the course of the days in which the streaming client was operational, the percentage of missed tweets is fewer than 0.01% of the total geolocated tweets within the United States. I include tweets only from users who tweet fewer than 25 times a day to reduce the incidence of non-human accounts in the sample. The left panel of Fig. 1 maps the total tweet volume in my sample across the United States, where pixel shading represents the logged volume of tweets. There is considerable spatial variation in Twitter activity, and most activity occurs in cities. The map also captures the extent to which this activity follows human movement patterns: along with cities, major highways and roads are readily visible in the map.

The translation of unstructured text data into quantitative data is known as "natural language processing", or NLP. Within NLP, the set of techniques designed specifically to quantify expressed sentiment is called "sentiment analysis." At the time of this writing, there are more than fifty publicly available algorithms and/or wordlists used to conduct sentiment analysis (Medhat et al., 2014). Because the method by which these measures are constructed can differ substantially, analyses using expressed sentiment should ideally demonstrate reasonable consistency across multiple measures. In this paper, I translate tweet content into four measures of expressed sentiment derived from prior work: AFINN (Nielsen, 2011), Hedonometer (Dodds and Danforth, 2010), LIWC (Pennebaker et al., 2015), and VADER (Gilbert and Hutto, 2014). The construction for each measure is similar: each uses a word list, or dictionary, which contains sentiment scores that correspond to English-language words. The overall measure of sentiment in a piece of text is the average of all scored words within that piece of text.

Panel A in Table 1 describes the unstandardized sentiment measures in the sample, although I standardize the measures prior to analysis for comparability. Following prior work, I pre-process each tweet before scoring in order to increase the precision of the NLP algorithms (Pak and Paroubek, 2010). I remove punctuation, URLs, hashtags (e.g., "#job"), and mentions (e.g., "@person") to isolate the word selection in the tweet. Because the independent variable of interest is weather, I remove tweets that contain any weather-related terms (see Table A.4 for the list of weather teams I exclude) to ensure that the responses do not capture the sentiment of observations about the weather, only changes in general sentiment due to weather. Once the tweets have been pre-processed, I score them for sentiment using the pre-existing dictionary (AFINN, Hedonometer, and LIWC), with the exception of VADER, which contains its own pre-processing routines. Appendix A.1.2 gives background and additional detail for each measure. Finally, in addition to the sentiment measures, I include a profanity measure intended to capture the use of online vulgarity, changes in which are reported as a percentage of average profanity used in the sample.[8]

Table 2 shows the correlations between the five measures at the CBSA-date level.[9] All of the measures are strongly positively correlated with each other, except the measure of profanity which is negatively correlated with all measures. The measures capture substantial geographic heterogeneity: the right panel of Fig. 1 documents average sentiment (as measured by VADER) by CBSA across the country. As in the state averages, urban CBSAs and CBSAs in the northern part of the United States show higher average sentiment relative to rural CBSAs and CBSAs in the southern part of the country.

2.2. Weather data

To obtain local estimates of daily weather across the contiguous United States, I use the PRISM Climate Group's AN81d gridded weather dataset. These data provide daily measures of minimum temperature, maximum temperature, and precipitation at roughly 4×4 km grid cells for the entire United States. The data are produced using a model that interpolates measurements from more than 10,000 weather stations and corrects for altitude and other influences on local climate (Daly et al., 2002). The second panel in Table 1 describes sample statistics for the PRISM data. I aggregate the gridded data to the CBSA level using population weights to ensure that the weather covariates reflect the average weather experienced by individuals within each CBSA.

Prior work suggests that other weather variables besides temperature and precipitation may be drivers determinants of emotional state (Dennisenn et al., 2008). Accordingly, I also gather daily data on the proportion of day that was overcast, relative humidity, station pressure, and wind speed from 2162 weather stations included in the NOAA Quality Controlled Local Climatological Data, or QCLCD.

3. Estimating preferences for temperature

This section empirically estimates the expressed sentiment response to temperature in order to identify preferences. Section 3.1 describes the empirical approach and identification strategy, while Section 3.2 documents the findings across a range of measures of expressed sentiment, empirical specifications and sampling frames. Section 3.3 uses the richness of the data to explore whether climate preferences are likely to adapt over time.

3.1. Empirical approach

I identify the causal effect of temperature on expressed sentiment using a panel fixed effects model, with temperature entering the regression using a flexible functional form. This flexibility is justified for the following reasons: first, prior work estimating temperature has documented non-linearities across a wide array of responses to temperature (Carleton and Hsiang, 2016); second, an appropriate flexible functional form should reveal the shape of the underlying response function, linear or otherwise (Hsiang, 2016); and third, intuition suggests that there is some bliss point for temperature, if only because temperatures which threaten human survival are clearly not preferable. The value of the panel nature of the dataset is that it allows me to control for unobservable cross-sectional

8 Figure A.2 shows state average expressed sentiment, ordered from lowest to highest. While cross-sectional differences in expressed sentiment provide limited causal insights, the plot suggests that colder regions express higher sentiment on average: the top six states are Montana, New Hampshire, Vermont, Wyoming, Minnesota, and South Dakota.

9 Table A.12 shows cross-sectional correlations between the measures at the state level, as well as comparisons to other measures of subjective well-being.

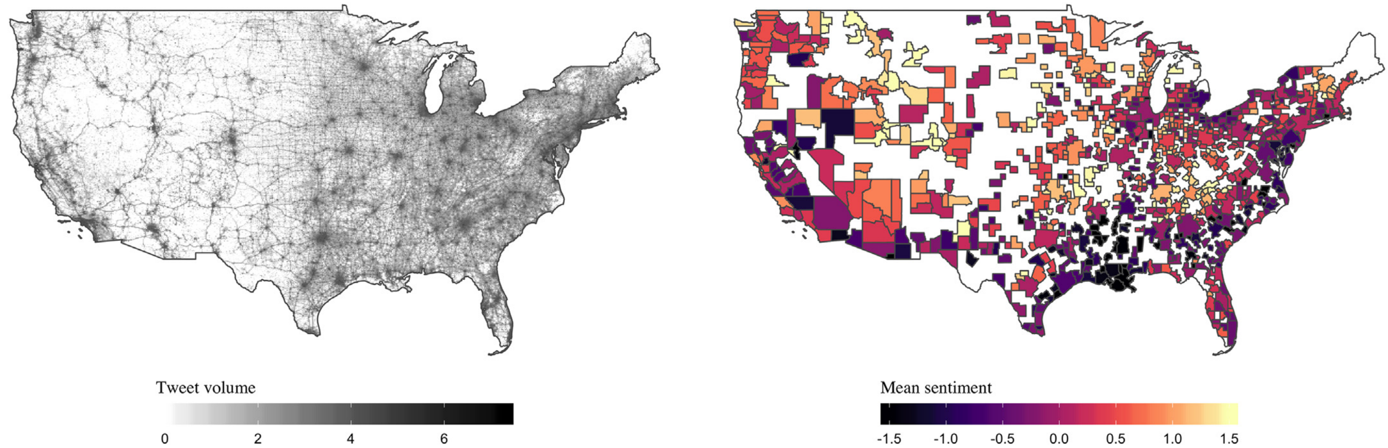

Fig. 1. Tweet density and average sentiment by CBSA *Notes:* Left panel: pixel shading represents log (base 10) of count of tweets in sample. Right panel: Mean standardized VADER score by CBSA for CBSAs with more than 100 tweets in sample.

and seasonal variation. Specifically, I estimate the following statistical model:

$$\overline{S}_{cd} = f(T_{cd}) + P_{cd} + \phi_c + \phi_{\text{time}} + \varepsilon_{cd} \tag{1}$$

Let c and d index CBSA and date. $\overline{S}_{cd}$ is the CBSA-date average of one of the four measures of sentiment described in Section 2. T_{cd} is the maximum daily temperature in a CBSA, and $f(T_{cd})$ is a flexible function of temperature, which I implement in practice using a binned model specification to allow for nonparametric responses of expressed sentiment to temperature. In particular, I let $f(T_{cd}) = \sum_b^B \beta_b T_{cd}^b$, where T_{cd}^b is an indicator variable equal to one if T_{cd} falls in the given bin b. β are the coefficients of interest. P_{cd} is daily precipitation. ϕ_c represents CBSA fixed effects, and ϕ_{time} represents a set of additional temporal controls, including month, year, day of week, and holiday fixed effects, as well as state-specific time trends. ε_{cd} is the idiosyncratic error term, clustered by both CBSA and date. I estimate the model using weighted least squares, where the weights are the average number of tweets in the CBSA.

T_{gd}^b specifies one, three, or five degree bins running between 0 and 40 degrees C, with edge bins for all observations with maximum temperature less than 0 or greater than 40.[10] I include both three and five degree versions of this model as part of the main results I present in the paper, and a comparison of all three bin widths in the appendix. For all bin widths, I choose the bin that contains 22.5 C as the omitted category. This choice does not alter the shape of the estimated response function, since relative differences between conditional means are preserved, but it does reflects the prior finding that Americans prefer 65 F (18.3 C) average daily temperature (Albouy et al., 2016). In my sample, because I use daily maximum temperatures rather than average temperatures, this corresponds to the omitted bin that I choose.

As shown earlier, the right panel of Fig. 1 documents cross-sectional variation in sentiment. Although all regions have a mix of high and low-sentiment CBSAs, visual inspection suggests that there is substantial regional variation in expressed sentiment. Additionally, prior evidence suggests that individuals with higher incomes tend to experience higher levels of life satisfaction and can afford to locate in areas with generally pleasant climate (Easterlin, 2001). If this regional variation, which may result from cultural or economic factors, correlates with regional weather differences, a naïve estimate of the relationship between weather and expressed sentiment is likely to be biased. To account for this regional variation in sentiment, I include CBSA fixed effects ϕ_c. These fixed effects ensure that the model is estimated on deviations from CBSA averages rather than on cross-sectional differences in climate, which could correlate with average sentiment or lexical patterns that register as different sentiments. Intuitively, the implication of this modeling choice is that the estimates represent a weighted average of within-CBSA comparisons, e.g., the difference in sentiment in Madison, WI on a hot day versus a cold day.

A second concern addressed by this identification strategy is the seasonality of both sentiment and temperature. To account for this possibility, I include month of year fixed effects as part of ϕ_{time}. Intuitively, this choice of fixed effects implies that the model coefficients represent a weighted average of the differences in sentiment on hot days versus cold days within a particular area and month, e.g., Chicago, IL in June. State time trends and year fixed effects account for potentially correlated trends in both temperature and sentiment that in the sample, while day of week and holiday fixed effects remove statistical noise related to within-week and by-holiday variation in expressed sentiment.

The combination of these fixed effects defines the identification strategy: at most, I assume that deviations in weather are as good as random after accounting for unobserved variation by CBSA, month of year, and year. This assumption is typical of the climate impacts literature (Hsiang, 2016). I also estimate alternative specifications with differing sets of fixed effects. Conditional on the assumptions given above, the coefficients of interest β_b can be interpreted as the average change in sentiment resulting from replacing a day in the omitted bin with a day in temperature bin b.

3.2. Findings

I find statistically significant declines in expressed sentiment resulting from both low and high temperatures. Section 3.2.1 documents the baseline findings from Eq. (1) across a range of measures of expressed sentiment and specification choices. Section 3.2.2 undertakes a disaggregated analysis using tweets as the unit of observation in order to test for compositional sorting, and Section 3.2.3 identifies the degree

Table 2
Correlations of expressed sentiment measures.

	AFINN-111	Hedonometer	LIWC	Vader	Profanity
AFINN-111	1.00	0.65	0.73	0.77	−0.56
Hedonometer		1.00	0.59	0.72	−0.34
LIWC			1.00	0.76	−0.38
Vader				1.00	−0.39
Profanity					1.00

Notes: Pairwise correlations of CBSA-date means of measures of standardized expressed sentiment and profanity measure.

[10] Because three does not multiply evenly into 40, the upper limit for the three degree bin specification is 39C.

to which the use of profanity responds to temperature. I discuss each in turn.

3.2.1. Baseline estimates

For expositional clarity, I first present the main result for each sentiment measure in Fig. 2. I show that the shape of the response functions is remarkably similar across the different measures of sentiment. Second, Table 3 tabulates the response of the VADER measure under a range of different choices of fixed effects.

Fig. 2 documents the temperature response of all four measures of sentiment estimated using Eq. (1). Because each outcome measure is standardized to have mean zero and unit standard deviation, the point estimates β_b represent the change in the conditional mean of expressed sentiment, measured in standard deviations, expected as a result replacing a day with a high of 21–24 C with a day with a high in bin *b*. I include a histogram underneath each plot to demonstrate the support of the temperature distribution. Each panel includes all four sets of point estimates, with the darker line indicating the central estimate, as indicated in the subtitle, and the dotted lines indicating the 95% confidence interval for the central estimate. For comparison, the other estimates are included as light gray lines without confidence intervals.

The upper-left panel documents a decline in the AFINN sentiment measure below 12 C and above 30 C. The difference in sentiment between days with the coldest temperatures (<3 C) and days in the omitted bin is around 0.15 SD, similar to the difference in sentiment between very hot days (>39 C) and days in the omitted bin. Confidence intervals are slightly wider for cooler temperature estimates but the point estimates are statistically different from zero at both ends of the temperature range. The AFINN measure estimates the second largest cold-weather effect and the largest warm weather effect.

The upper-right panel estimates a similar response shape for the Hedonometer measure, both in shape and in magnitude. Point estimates are statistically different from zero on both ends of the temperature scale. The Hedonometer measure estimates the second largest warm-weather effect and the largest cool weather effect.

The bottom-left panel documents the response for the LIWC measure. While still within the confidence intervals of the other estimates, LIWC documents more limited impacts of cold temperatures on sentiment. The point estimates are statistically significant and similar to the other measures for warm weather temperatures, but smaller and not statistically significant for colder temperatures. This difference from the other measures may result from the measure's lack of suitability for the microblogging format.

The bottom-right panel documents the response as measured using VADER. Like the AFINN and Hedonometer measures, VADER estimates a decline in sentiment below 12 C and above 30 C that reaches about 0.2 SD at maximum. VADER estimates a similar response to AFINN or Hedonometer but a larger response than LIWC.

Each outcome measure in Fig. 2 documents a statistically significant negative relationship between sentiment and hot temperatures, relative to a day with moderate temperatures. The magnitudes of the effect sizes differ, ranging from about 0.1 SD to more than 0.2 SD for the hottest temperature bin. The relationship between sentiment and cold temperatures is slightly less precisely estimated, and one of the four measures

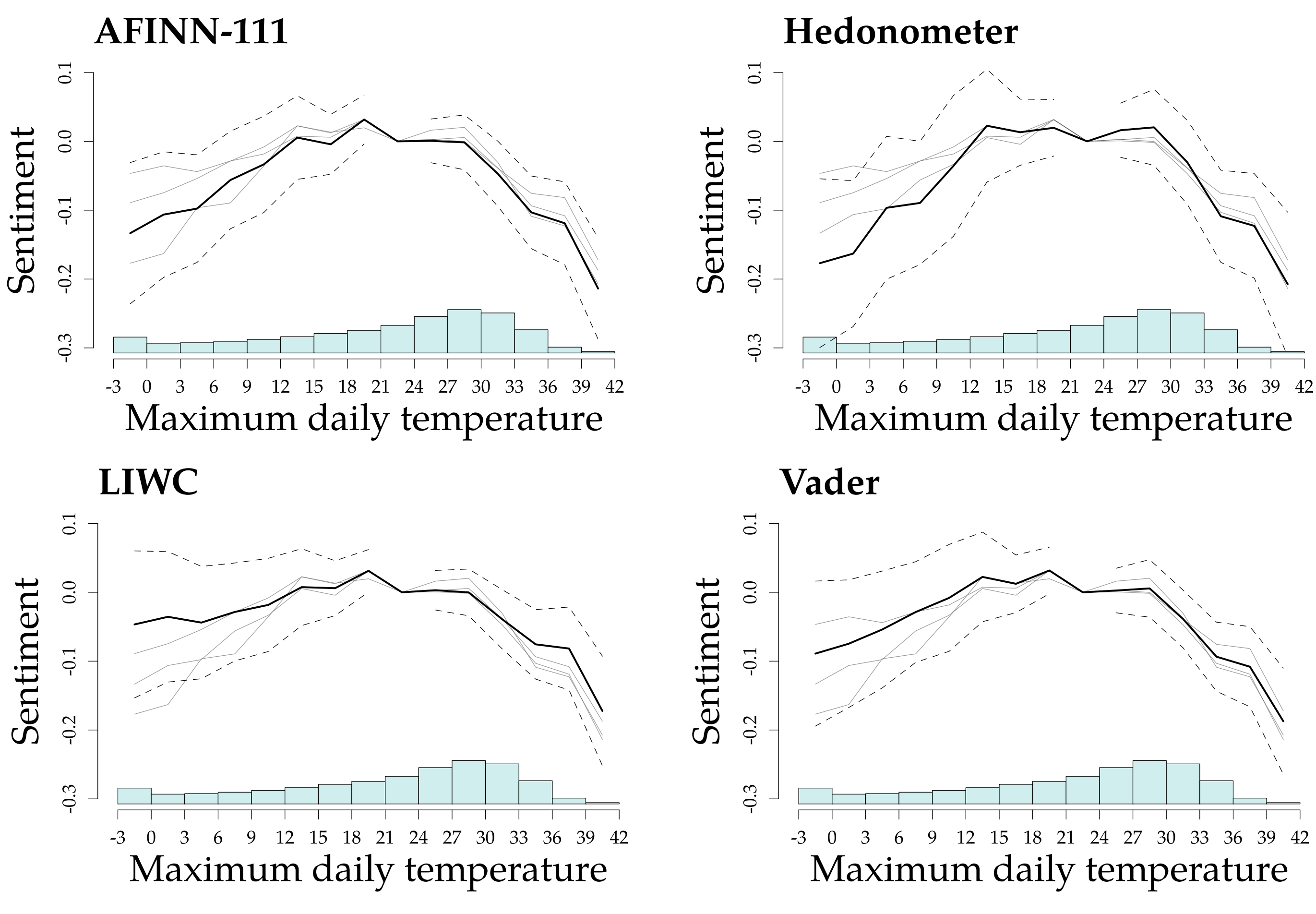

Fig. 2. Effect of temperature on expressed sentiment by measure *Notes:* Panels document the temperature response for each of the four standardized measures of sentiment. Solid lines show the regression coefficients on temperature and represent the difference (measured in standard deviations) in CBSA-date sentiment for the temperature bin T_b relative to 21–24 C, controlling for state time trends and fixed effects for CBSA, day of week, holiday, month, and year fixed effects. Dotted lines are 95% confidence intervals, with standard errors clustered by CBSA and date.

Table 3
Effect of temperature on expressed sentiment by specification.

	(1)	(2)	(3)	(4)	(5)
Maximum daily temperature T					
$T \le 5$	−0.15	−0.16	−0.05	−0.19	−0.09
	(0.06)	(0.05)	(0.07)	(0.05)	(0.05)
$T \in (5,10]$	−0.09	−0.07	−0.005	−0.11	−0.03
	(0.05)	(0.04)	(0.05)	(0.04)	(0.04)
$T \in (10,15]$	−0.03	−0.02	0.02	−0.03	0.002
	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)
$T \in (15,20]$	−0.01	0.004	0.02	−0.003	0.01
	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)
$T \in (25,30]$	−0.001	0.01	−0.01	0.01	−0.003
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
$T \in (30,35]$	−0.04	−0.03	−0.04	−0.03	−0.06
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)
$T \in (35,40]$	−0.11	−0.10	−0.09	−0.09	−0.12
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)
$T>40$	−0.23	−0.20	−0.11	−0.18	−0.21
	(0.04)	(0.05)	(0.06)	(0.04)	(0.04)
Other controls					
CBSA	Yes	Yes	Yes	Yes	Yes
Month	Yes	Yes			Yes
Year	Yes	Yes	Yes		Yes
DOW, Hol		Yes	Yes	Yes	Yes
S×M			Yes		
MOS				Yes	
State trends					Yes
N (millions)	0.7	0.7	0.7	0.7	0.7
Tweets (millions)	1160.6	1160.6	1160.6	1160.6	1160.6

Notes: Table shows estimates of weather on expressed sentiment using different statistical specifications. Expressed sentiment measured using the VADER sentiment analysis algorithm. Statistical specifications vary according to choice of fixed effects and other controls, as given in the "Other controls" panel. All specifications include daily precipitation. Coefficients represent the difference (in standard deviations) in expressed sentiment between a day with maximum temperature in the associated temperature bin and a day with temperature $T \in (20,25]$, the omitted category. N (in millions) is the number of observations (CBSA-dates). Tweets (in millions) is the number of posts represented by each regression. Standard errors (in parentheses) clustered by CBSA and date.

fails to reject the null of no difference between cold and moderate temperatures, although the consistent decline of the point estimates provides suggestive evidence of a negative effect in low temperatures. Despite these differences, the results of this exercise are markedly similar across measures: each exhibits the same upside-down "U" shape, each reaches similar magnitudes on both the cold and warm temperature ends of the temperature spectrum, and each is statistically significant at both of those ends (with the exception of LIWC in cooler temperatures).

Because the response functions are consistent across measures, the remainder of the paper focuses on results obtained using VADER. Table 3 estimates the effect of temperature on expressed sentiment using five degree C bins and across a range of choices of fixed effects. All columns include CBSA fixed effects.

Column (1) reflects the baseline specification, which also includes month and year fixed effects. As in Fig. 2, I observe negative and statistically significant effects below 10 and above 30. Column (2) adds day of week and holiday fixed effects to absorb weekly variation in sentiment (see Fig. 5) and variation related to holidays, but the estimated effects are virtually unchanged. Column (3) introduces state-by-month fixed effects to account for regionally distinct seasonal trends. This set of fixed effects is particularly restrictive, as the model is now identified solely off of variation within a state-month and because most states have only a few CBSAs. Qualitatively, I find that the point estimates again identify an upside-down U-shape impact of temperature on expressed sentiment. However, the impact of cooler temperatures is not statistically different from zero under this specification, and the point estimates for both cold and hot temperatures are partly attenuated towards zero. These estimates likely result from the influence of the more restrictive state-month fixed effects, which reduce the residual variation available in the model more than any other specification (see Fig. A.6). As described in Angrist and Pischke (2008), classical measurement error in fixed effect models can lead to a biasing of the coefficients towards zero as more and more restrictive fixed effects groupings are considered. To further investigate this claim, column (4) replaces the state-by-month fixed effects with month of sample fixed effects. Under this specification, the cold temperature estimates are again large and significant. Finally, column (5) replaces the state-by-month fixed effects from column (3) with a more flexible set of state trends and finds largely similar results, although the point estimates are somewhat larger for the cold temperature estimates.

The negative relationship between temperature and sentiment below 12 C and above 30 C resembles that estimated by Albouy et al. (2016), who find that individuals pay to avoid warm temperatures in summer and cold temperatures in winter. The preferred model estimates the magnitude of the difference between a moderate day and an extremely cold or hot day to be around 0.1 SD and 0.2 SD, respectively.

Broadly, I find qualitatively similar results across a range of specifications. Both hot and cold temperatures have negative effects on expressed sentiment. In addition to the test for composition sorting described in Section 3.2.2, the appendix includes more sensitivity checks, including the inclusion of additional weather variables (Table A.7), variations on bin width (Fig. A.5), different sampling frames (Table A.8), and weighting choices (Table A.9), none of which qualitatively alter the baseline results.

3.2.2. Compositional sorting

Because Twitter users choose when — and when not — to tweet, the selection mechanism into the sample could induce a compositional bias in the estimates observed in Fig. 2, a sample selection effect akin to that described by Heckman, (1979). This can also be viewed as a form of the ecological fallacy: the observation that the properties of aggregated groups may not reflect properties of the individuals in the underlying populations (Robinson, 1950). To fix ideas, imagine two types of Twitter users: positive and negative. Positive users create only positively-scored tweets, while negative users create only negative tweets. Because neither type will change the content of its tweets in response to temperature, the true underlying effect of temperature on their sentiment in zero. However, suppose as well that positive users choose to put their phones away when it's very cold or very hot, whereas negative users are unaffected. An econometric approach using CBSA averages that does not control for the type of user could misidentify compositional sorting as part of the true effect.

Since the data I collect include an identifier for the tweet creator, I can account for compositional sorting in my sample using post-level data and user fixed effects. To do so, I estimate the following model:

$$E_{id} = \sum_{b \neq 20-25}^{B} \beta_b T_{cd}^b + \phi_i + \phi_m + \phi_y + \varepsilon_{id} \tag{2}$$

This model replaces CBSA fixed effects with user fixed effects ϕ_i in equation Eq. (1). The model requires the use of the disaggregated sample of tweets in my dataset; for computational reasons, I focus on users who tweet during 25% of the days in my sample but who produce fewer than 25 tweets per day, resulting in 432 million tweets. Table 4 compares the results between models using CBSA averages and individual posts as the unit of observation.

The fourth column documents the estimates using the individual sample. For computational reasons, this estimate is obtained using user, CBSA, month, and year fixed effects, but does not include state trends or day of week and holiday fixed effects. In order to compare this estimate to the main sample, the first three columns document results using the CBSA-by-date aggregated dataset, where the first column in Table 4 is equivalent to the fifth column in Table 3. The second column drops the state trends and day of week and holiday fixed effects, but maintains the baseline sample. The third column restricts the CBSA sample to the sample used in the individual-level estimate.

I find qualitatively similar results for these measures using the individual sample, although the estimates for higher temperatures are slightly attenuated in the individual fixed effects model relative to the models using CBSA-date averages. It is possible that this is evidence of some compositional sorting at higher temperatures, but could also be the result of measurement error driven by using a sparser source of variation as a result of the user fixed effects. In either case, the shape of the results is largely similar across the two models and the results do not appear to be primarily driven by compositional changes resulting from temperature variation.

3.2.3. Profanity response

A large literature has documented the impact of climate on conflict (Burke et al., 2015). One possible mechanism is the finding that warm temperatures encourage aggressive behavior (Kenrick and MacFarlane, 1986). To understand whether the expressed sentiment response to temperature is due in part to this aggression mechanism, I estimate the relationship between temperature and expressions of profanity. Using a list of more than 300 profanities, I estimate Eq. (1) with the occurrence of tweets in a CBSA-date that contain a profanity as the outcome of interest. One concern with this approach may be that if users are simply using more profanities to reflect their mood, these effects could represent same decline in emotional state captured by Fig. 2. To investigate this possibility, I also construct an “aggressive profanity” metric that counts only the number of tweets that include a popular, aggressively profane phrase. Fig. 3 plots the results, scaling each measure by its own mean to obtain the percent change in tweets using the selected profane terms.

I find that use of profanity and aggressive profanity rise in both hot and cold temperatures. Previous work on both conflict (Burke et al., 2015) and on violent crime (Ranson, 2014) find that both increase during periods of high temperatures. That I document a similar effect for hot temperatures is consistent with the hypothesis that increases in temperature induce violence by making individuals more aggressive. However, I also find that cold temperatures induce more profane text than moderate temperatures. This finding is in contrast to previous work on temperature and aggressive behavior, which has not typically found an increase in crime or conflict during periods of cooler temperatures (Ranson, 2014; Burke et al., 2015). It may be that aggressive impulses increase in response to temperature discomfort of both kinds, but that cooler temperatures limit opportunities to act on that aggression.

Table 4
Effect of temperature on expressed sentiment by unit of observation.

Unit	CBSA-date averages			Posts
	Baseline	+Spec.	+Sample	Individual
Maximum daily temperature T				
$T \leq 5$	−0.09	−0.15	−0.08	−0.08
	(0.05)	(0.06)	(0.03)	(0.02)
$T \in (5,10]$	−0.03	−0.09	−0.06	−0.04
	(0.04)	(0.05)	(0.02)	(0.01)
$T \in (10,15]$	0.002	−0.03	−0.02	0.01
	(0.03)	(0.04)	(0.02)	(0.01)
$T \in (15,20]$	0.01	−0.01	−0.01	0.01
	(0.01)	(0.02)	(0.01)	(0.01)
$T \in (25,30]$	−0.003	−0.001	−0.01	−0.001
	(0.02)	(0.02)	(0.01)	(0.01)
$T \in (30,35]$	−0.06	−0.04	−0.01	−0.02
	(0.02)	(0.03)	(0.02)	(0.01)
$T \in (35,40]$	−0.12	−0.11	−0.07	−0.04
	(0.03)	(0.03)	(0.02)	(0.01)
T>40	−0.21	−0.23	−0.15	−0.10
	(0.04)	(0.04)	(0.03)	(0.02)
Other controls				
CBSA FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
DOW, Hol FE	Yes			
State trends	Yes			
User FE				Yes
N (millions)	0.7	0.7	0.7	432.6
Tweets (millions)	1160.6	1160.6	432.6	432.6

Notes: Table shows estimates of weather on expressed sentiment using different units of observation. Coefficients represent the difference (in standard deviations) in expressed sentiment between a day with maximum temperature in the associated temperature bin and a day with temperature $T \in [20,25]$, the omitted category. First three columns use CBSA-date averages as the unit of observation, the final column uses individual posts. “Individual” is estimates on tweets from users who posted on at least 25% of the days in the sample and averaged no more than 25 posts per day. “Baseline” column reproduces column (5) from Table 3. “+Spec.” column matches the “Individual” column specification. “+Sample” additionally matches the “Individual” column sample. All specifications include daily precipitation. N (in millions) is the number of observations (CBSA-dates or posts). Tweets (in millions) is the number of posts represented by each regression. Standard errors (in parentheses) clustered by CBSA and date, except in the case of the “Individual” column, where they are clustered by CBSA.

3.3. Understanding adaptation

As I discuss in Section 1, the climate impacts literature has identified a range of settings in which variation in temperature has had both statistically and economically significant impacts on economic outcomes of interest. The question of whether and to what extent these impacts can be extrapolated to climate change is critically important for projecting cumulative economic impacts. The spread of humans across the planet suggests that, in the long run at least, humans are highly capable of surviving in a wide range of environments. The relatively slow pace of climate change invites the possibility that many of the measured impacts could be partly mitigated by either adaptive responses or by sorting.[11] Empirical estimation of adaptation has presented substantial challenges for researchers working in this area: direct, causally identified models usually rely on long-differences methods as in Burke and Emerick (2015), which in turn rely on sufficient long-run variation in temperature and the outcome of interest across a large geographical area. For most studies, including this one, the requirement of a multidecadal panel dataset for proper estimation of long-run effects is unattainable. Even for studies with such a dataset available, the research design effectively reduces the number of available observations to the number of observed geographical units, which restricts statistical power and reduces the ability of researchers to strongly reject large portions of the parameter space.

As an alternative to providing direct evidence on adaptation or sorting, in this section I take advantage of the richness of the dataset to run two empirical tests designed to suggest whether preference adaptation or sorting is likely to occur in this setting. Below, I estimate the degree of heterogeneity in temperature-sentiment responses both across the four quartiles of average annual temperatures and across the four seasons of the year, finding important differences in the sentiment response across these dimensions. While these are not sharp tests of adaptation, they do help to inform the extent to which temperature preferences do or do not adapt over time.

Fig. 4 estimates separate splined models by region, where each panel identifies the response for the given region. Regions are split by quartiles of average annual temperature, with labels given in order as “Coldest”, “Cold”, “Warm”, and “Warmest”. In order to mitigate the loss of statistical power that results from estimating regional models, I use a splined model of temperature (separately estimated for the full sample in Fig. A.3). To document uncertainty, I bootstrap these estimates with 1000 iterations of the same specification, sampling from the full dataset with replacement. For each region, the red line indicates

[11] There is a burgeoning literature on understanding adaptation. For more complete discussions of the subject across a range of areas, see Auffhammer et al. (2013), Houser et al. (2014), Graff Zivin et al. (2018), Auffhammer (2013), and Dell et al. (2014).

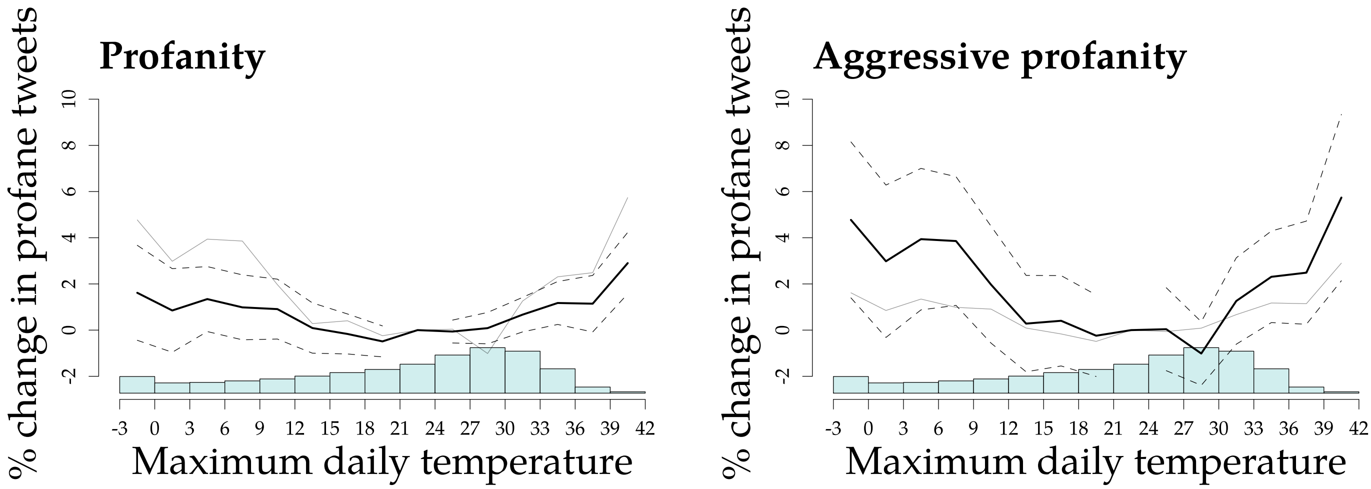

Fig. 3. Effect of temperature on profanity *Notes:* Figure documents occurrence of profanity in response to temperature, using a list of 311 profanities. Point estimates represent the difference (measured in standard deviations) in CBSA-date profanity occurrence for the temperature bin T_b relative to 21–24 C, conditional on CBSA and state by month of sample fixed effects. 95% confidence intervals represent standard errors clustered by CBSA and date.

the response identified for the entire sample for that season and the lighter gray lines indicates the bootstrapped responses. All lines are normalized such that their maximum value is equal to 0. Clear differences in the sharpness of the sentiment response to temperature can be observed across regions: colder regions have attenuated responses to cold temperatures, while warmer regions have attenuated responses to warm temperatures. Conversely, colder regions respond more to high temperatures and warmer regions less to cold temperatures.

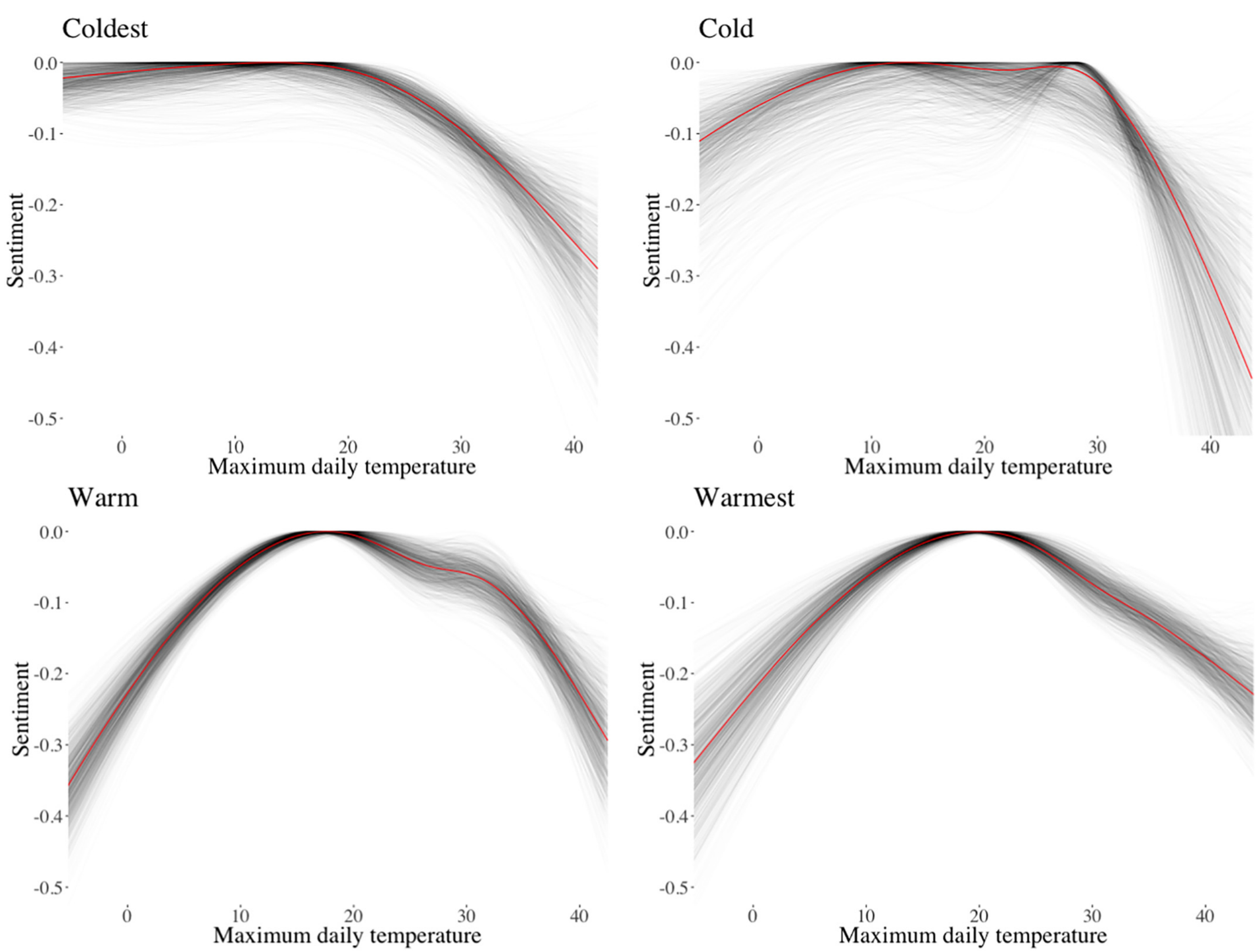

Fig. 4. Sentiment responses to temperature differ by region.

The regional heterogeneity in responses is evidence of either preference adaptation, technological adaptation, sorting, or some combination of all three, but I cannot distinguish between these. Individuals may have adapted their preferences to accommodate their climatic zones, they may have a greater degree of technologies available to mitigate extreme temperatures to which they have become accustomed (e.g., air conditioning and indoor heating), or individuals with stronger preferences around lower or higher temperatures may have chosen to vote with their feet, so to speak.

To help distinguish between these explanations, I also estimate the differences in seasonal responses, where the possibilities for technological adaptation and sorting are more limited. Appendix A.2.2 documents differences that are consistent with common intuition: cooler temperatures are preferred in fall and summer, warmer temperatures are preferred in winter, and moderate temperatures are preferred in spring. The combined regional and seasonal dependencies of preferences for temperature are suggestive of adaptive possibilities for temperature, and emphasize the importance of projecting climate impacts in a way that allows preferences to change in response to shifting average temperatures, as I do in Section 4.3.

4. Interpreting changes in expressed sentiment

By using expressed sentiment as a proxy for underlying preferences for temperature, I am able to mitigate the identification concerns that arise when using hedonic or discrete choice models to estimate the value of climate, as described above. The dataset I construct also allows me to estimate the underlying relationship between temperature and expressed sentiment non-parametrically, and to estimate region- and season-specific responses. These benefits must be weighed against the major drawback of this approach: the challenge of interpretation. This is a problem also faced by others in the body of literature that uses measures of life-satisfaction as the outcome of interest (Mackerron, 2012): how much is one unit of expressed sentiment or reported life satisfaction worth?

The advantage of the hedonic and discrete choice approaches is that that valuation is relatively straightforward (Albouy et al., 2016; Sinha et al., 2018). By contrast, backing out estimates of the value of climate damages using changes in expressed sentiment requires additional assumptions. However, doing so is important for several reasons: first, assigning a monetary value grounds the size of these effects in a metric that is more likely to be consistently interpreted by different readers; second, monetary calibration of the effect of changes in temperature on emotional state allows researchers and policy analysts to compare the size of these estimates to other documented effects of climate change; and third, monetary estimates are critical for inclusion in the three Integrated Assessment Models currently used by the United States Government to estimate the social cost of carbon (Rose, 2014).

As a first step towards identifying willingness-to-pay for climate amenities using these data, I provide a range of approaches designed to give meaning and context to the magnitude of the results I observe and to guide future work in this area. Section 4.1 describes three validation exercises: these demonstrate how other types of external variation impact the measures of expressed sentiment that I record. Section 4.2 demonstrates two valuation exercises to identify the monetary value of the shifts in sentiment I observe. Section 4.3 combines one of these estimates with predictions of future climate to project the damages from climate change.

4.1. Validation exercises

This section documents the relationships between three non-temperature sources of variation with expressed sentiment. Section 4.1.1 documents changes in expressed sentiment by day of week, while Sections 4.1.2 and 4.1.3 identify the causal impacts of hurricanes and football game outcomes, respectively, on expressed sentiment.

4.1.1. Expressed sentiment by day of week

First, I conduct a validation exercise that examines how sentiment changes over the course of the days of the week. First, Fig. 5 shows the standardized measures by day of week. The weekly variation in matches prior work (Dodds et al., 2011) and common intuition: weekends and Fridays are preferred to non-Friday weekdays, with the lowest measures of affect occurring on Mondays and the highest on Saturdays. To help interpret to the results shown earlier, note that the average difference in sentiment measure between Sunday and Monday is between 0.1 SD and 0.2 SD across the measures, or roughly the difference between experiencing a day with maximum temperature between 20 and 25 C and a day with maximum temperature between 35 and 40 C in Table 3.

4.1.2. Impact of hurricanes on expressed sentiment

The high winds, heavy rains, storm surges, and the distress and uncertainty hurricanes create results in both substantial economic losses and difficult-to-quantify human hardship (Hsiang and Jina, 2014). Although hurricanes mostly affect areas on the Eastern seaboard between June and November, their appearance and path of destruction tend to be unpredictable in the short term. For these reasons, estimating the impact of hurricanes on expressed sentiment serves as a useful benchmark for the baseline results in this paper. I collect data on hurricane occurrence from the Atlantic Hurricane Database (Landsea and Franklin, 2013) and combine it with CBSA-date averages of expressed sentiment, measured using VADER. To identify the causal impact of a nearby hurricane on expressed sentiment, I estimate a model similar to Eq. (1), but replace the weather variables with an indicator for proximity to a hurricane. Table 5 documents the results of this estimation for all hurricanes (first two columns) and for hurricanes with an average wind speed greater or equal to 40 m/s (second two columns), using two different sets of fixed effects.

On average, CBSAs experience a daily reduction in expressed sentiment of nearly 0.4 SD from any nearby hurricane, while high speed hurricanes cause a reduction of expressed sentiment of around 0.7 SD. This effect is of the same sign and between 2 and 7 times are large as the effect estimated in Fig. 2, indicating that the impact of nearby hurricanes is more pronounced than the impact of extremely hot or cold temperatures.

4.1.3. Impact of football game outcomes on expressed sentiment

I estimate a similar model to the above using the outcomes of NFL football games during my sample, following in the spirit of Card and Dahl (2011), who show that unexpected football losses correlate with

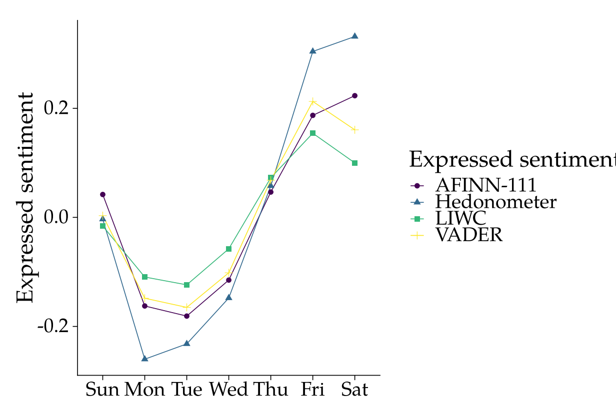

Fig. 5. Expressed sentiment by day of week *Notes:* Lines represent average standardized measures of expressed sentiment by day of week. Standardization is conducted using the weighted mean and variance of the CBSA-date averages.

Table 5
Impact of hurricanes on expressed sentiment.

	All	All	40+	40+
Hurricane in area	−0.41	−0.42	−0.65	−0.63
	(0.15)	(0.19)	(0.29)	(0.20)
Fixed effects				
CBSA	Yes	Yes	Yes	Yes
DOW	Yes	Yes	Yes	Yes
Month	Yes		Yes	
Year	Yes		Yes	
MOS		Yes		Yes

Notes: Table documents impact of nearby hurricanes expressed sentiment for nearby CBSAs. Outcome variable is standardized CBSA daily average of expressed sentiment. All regressions control for day of hurricane along with listed fixed effects. First and second columns include all hurricanes, third and fourth columns only includes those with a wind speed of 40 m/s. Standard errors clustered by CBSA and date.

Table 6
Impact of football outcomes on expressed sentiment.

	100 km	100 km	25 km	25 km
Nearby team won	0.10	0.10	0.08	0.08
	(0.03)	(0.03)	(0.05)	(0.05)
Fixed effects				
CBSA	Yes	Yes	Yes	Yes
DOW	Yes	Yes	Yes	Yes
Month	Yes		Yes	
Year	Yes		Yes	
MOS		Yes		Yes

Notes: Table documents impact of football outcomes on expressed sentiment for nearby CBSAs. Outcome variable is standardized CBSA daily average of expressed sentiment. All regressions control for day of game along with listed fixed effects. First and second columns include CBSAs within 100 km of a football team's stadium, third and fourth columns include CBSAs within 25 km. Standard errors clustered by CBSA and date.

family violence. I map each CBSA to a nearby football team, where being within either 100 or 25 km of a team's home stadium qualifies as "nearby." The model I estimate also includes an indicator for whether the team was playing on a given day. Table 6 documents the findings, where the first two columns take a nearby team as one within 100 km, and the second two take a nearby team as one within 25 km, with two different sets of fixed effects for each.

As expected, a nearby team suffering a loss causes a negative impact on expressed sentiment. The size of the impact is roughly equal to the negative impact of a very hot day in Fig. 2, and is slightly smaller when estimated using teams that are within 25 km relative to 100 km.

4.2. Valuing expressed sentiment

This section describes two approaches to convert the estimates of changes in expressed sentiment presented above into monetary value in order to estimate the value of different climates. This approach follows in the spirit of Levinson (2012), who converts reported life satisfaction into a dollar value by dividing the response of life satisfaction to pollution levels by the response of life satisfaction to changes in income.[12] Here I conduct a similar exercise using ambient temperature and expressed sentiment.

I repeat, as stated at the outset, that obtaining credible estimates of the value of expressed sentiment is important but requires strong assumptions. First, it must be the case that these changes in income are exogenously determined with respect to expressed sentiment, and second, that expressed sentiment represents an appropriate proxy for utility. With respect to the first assumption, the following two approaches represent attempts to isolate plausibly exogenous variation in incomes. The second assumption is more is difficult to test and the results that follow should be interpreted in light of that consideration.

To identify the relationship between income and expressed sentiment, I first estimate the degree to which quarterly expressed sentiment responds to changes in quarterly wages from the Quarterly Census of Employment and Wages (QCEW) and use this response to estimate a per-SD value of $196.77. Appendix A.4.1 documents the construction and estimation procedure in detail. The second approach identifies plausibly exogenous variation in income by focusing on the population of users in my dataset who received parking or speeding tickets and who noted that receipt on Twitter. Using this approach, I find a per-SD value of $78.60. See Appendix A.4.2 for more details on the construction and estimation of this approach.

To estimate the value of these temperature changes, I multiply the per-SD values from the two approaches above with the base estimates of the change in expressed sentiment caused by temperature from Table 3, column (5). Table 7 documents the results. Interpreted literally, these estimates imply, for example, that the average individual in my sample would be willing to pay $11.94 or $4.77 (depending on whether the wage regression or the parking ticket response value is used) to exchange a day between 30 and 35 degrees with a day between 20 and 25 degrees.

Because I estimate a higher per-SD value using the relationship between sentiment and quarterly wages, all of the estimates in the first column are larger (in absolute value) than those in the second column. Given the uncertainties outlined above in these valuation strategies, I proceed conservatively: the following section, which projects the costs of climate change using expected changes in expressed sentiment, relies on the smaller per-SD value ($78.60) obtained using the parking ticket approach.

4.3. Projecting damages from climate change

This section combines the estimates from Section 3 with projections of future climate and the valuations obtained in Section 4.2 to value the change in expressed sentiment we might expect from climate change. First, I project the annual amenity cost of rising temperatures across the United States on amenity value, measured in the change in SD of sentiment. I multiply these estimates with the

Table 7
Value of temperature.

	Value from wage regression	Value from parking ticket response
Maximum daily temperature T		
$T \leq 5$	−17.72	−7.08
	(8.90)	(3.55)
$T \in (5,10]$	−6.73	−2.69
	(7.04)	(2.81)
$T \in (10,15]$	0.46	0.18
	(6.42)	(2.56)
$T \in (15,20]$	2.27	0.91
	(2.92)	(1.17)
$T \in (25,30]$	−0.62	−0.25
	(3.15)	(1.26)
$T \in (30,35]$	−11.94	−4.77
	(4.03)	(1.61)
$T \in (35,40]$	−23.61	−9.43
	(5.69)	(2.27)
T>40	−42.02	−16.78
	(8.53)	(3.41)

Notes: Table identifies the value of daily temperature. Reproduces column (5) in Table 3 with coefficients multiplied by valuation of a unit SD change in standardized VADER sentiment. Value of an SD change in the first column uses regression of expressed sentiment on quarterly CBSA wages, as described in Appendix A.4.1. Value of an SD change in the second column is calculated by dividing the median observed ticket cost ($100) by the total cumulative sentiment loss from a parking or speeding ticket (1.27 SD) in Fig. A.11, as described in Appendix A.4.2.

[12] Allcott et al. (2019) conduct a similar exercise in their investigation of the optimal soda tax.

per-SD value of changes in expressed sentiment to obtain the dollar value of these changes.

The projection exercise can be described mathematically as follows:

$$\int^{T} f(t)\Delta g(t) v dt \tag{3}$$

where $f(\cdot)$ represents the damage function (valued in SD of expressed sentiment), such as the one estimated in Fig. 2, $\Delta g(\cdot)$ is the change in the distribution of climate, and v is the value of a single SD change in dollars. By integrating the product of f and g over the range of temperature T and multiplying by v I obtain the total damages. Empirically, I estimate the shape of $f(\cdot)$, combine climate and weather data to obtain $\Delta g(\cdot)$, and draw v from Section 4.2 in order to numerically approximate Eq. (3).

I conduct two exercises using this framework, referred to hereafter as the "baseline" and "adaptive" projection exercises. The baseline exercise projects damages using a single function for f, the estimate obtained by the splined model in Fig. A.3. The adaptive exercise follows Auffhammer and Aroonruengsawat (2011) and uses the regionally-specific damage functions estimated in Fig. 4 to project changes due to adaptation. For each projection, I assume that CBSAs respond with the regional damage function that corresponds to the annual average temperature it contemporaneously experiences. In other words, if climate changes are projected in an area to the extent that the area moves from the "Warm" quartile into the "Warmest" quartile, then its damages are estimated using the "Warmest" quartile's damage function.

$g(\cdot)$ is estimated using the ensemble average from the output of 20 downscaled climate models,[13] I compile average projections for each CBSA for the years 2006–2099. In order to de-bias the projections, I follow the prescriptions of Auffhammer et al. (2013) and add the difference between projected monthly decadal averages starting in 2026 and projected monthly averages from 2006 to 2025, then add those differences to the historical weather data from 2006 to 2015 to simulate future weather regimes for each decade while retaining historically observed variance in temperature. I estimate the difference in the distributions between baseline climate and the given future climate to obtain $\Delta g(\cdot)$.

Fig. 6 documents the evolution of per-person annual damages over time, averaged over CBSAs and presented separately for RCP 4.5 and RCP 8.5, two different climate forcing scenarios of intermediate and high warming, respectively (IPCC, 2014), and using both the baseline and adaptive methods described above.

I find annual damages increasing over time across all scenarios, with damages from RCP 8.5 exceeding damages from RCP 4.5. For the baseline scenario, I estimate per-person annual damages of \$297 under RCP 4.5 and \$810 under RCP 8.5. Allowing for the possibility of adaptation shifts the estimates downward, to \$195 and \$313 respectively. In all cases, however, these are notable estimates: the annual median income in the United States was \$31,786 in 2017, meaning that the estimates I obtain are between 0.6% and 2.5% of present-day annual income. Less conservative modeling decisions, such as using the per-SD value from the relationship between CBSA and income in Section 4.2, would obtain larger damages. Broadly speaking, the set I give above are in line with other estimates of the amenity cost of climate change. For example, Albouy et al. (2016) estimate end-of-century damages of between 1% and 4% of annual income. Fig. A.12 documents the distribution of damages across space, estimating damages by end of century for each CBSA under RCP 8.5. Under the baseline scenario, damages are increasing in

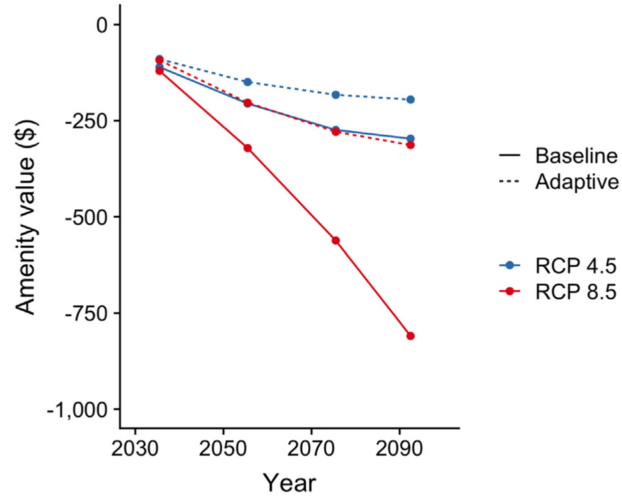

Fig. 6. Projections of changes in amenity value over time *Notes:* Projections of average change in amenity value over time, measured in annual per-person change in SD of expressed sentiment. Line color indicates the warming scenario, or Representative Concentration Pathway, used in the projection data. Line type indicates whether the projection method was the baseline or adaptive method. Further projection details given in Section 4.3.

average temperate and are highest for CBSAs in the South, while under the adaptive scenario, damages are greatest for both the South and for parts of the upper Midwest.

The difference between the baseline and adaptive approaches is also notable. This estimate is made possible by the richness of the data I use, since I am able to estimate separate response functions for each region in Fig. 4. As expected, the ability to adapt one's preferences for temperature as the climate warms mitigates the impact of climate change substantially. Whether this degree of adaptation is likely remains an open question and dependent on whether cross-sectional differences in responses to temperature result from factors that could change in the long run.

5. Impact of temperature on sentiment around the world

For reasons of data availability, the analysis thus far has focused on the United States. However, while all countries will be affected by climate change, extrapolating preferences for climate estimated earlier in the paper to the rest of the world could be inappropriate. In order to provide the first estimates (of which I am aware) of the impact of temperature on expressed sentiment outside of the United States, I collect a similar dataset for six English-speaking countries for which there is adequate temperature variation and for which I am able to obtain a sufficient amount of Twitter data.[14]

To do so, I again use the Twitter Streaming API, but this time request tweets located inside a bounding box that includes all of Africa, parts of Southeast Asia, and Australia. These data, whose collection began later, span from October 2015 until March 2019. For each first-level administrative unit in these countries (equivalent to a U.S. state, hereafter referred to as "states" for brevity), I estimate the average sentiment score given by the VADER sentiment measure described above for all tweets in that administrative unit and day. I keep only countries for which I am able to observe at least one million tweets over the course of the sample.

Because the PRISM dataset is not available outside of the United States, I instead use gridded datasets of maximum daily temperature and total daily precipitation from NOAA.[15] I take population weighted-

[13] Climate forcings drawn from a statistical downscaling of global climate model (GCM) data from the Coupled Model Intercomparison Project 5 (Taylor et al., 2012) using the Multivariate Adaptive Constructed Analogs (MACA; Abatzoglou and Brown, 2012) method with the Livneh (Livneh et al., 2013) observational dataset as training data.

[14] I define "adequate" as requiring that at least four of the five degree bins between 0 and 40 degrees C constitute more than 1% of the total observations.

[15] CPC Global Temperature and Precipitation data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at https://www.esrl.noaa.gov/psd/.

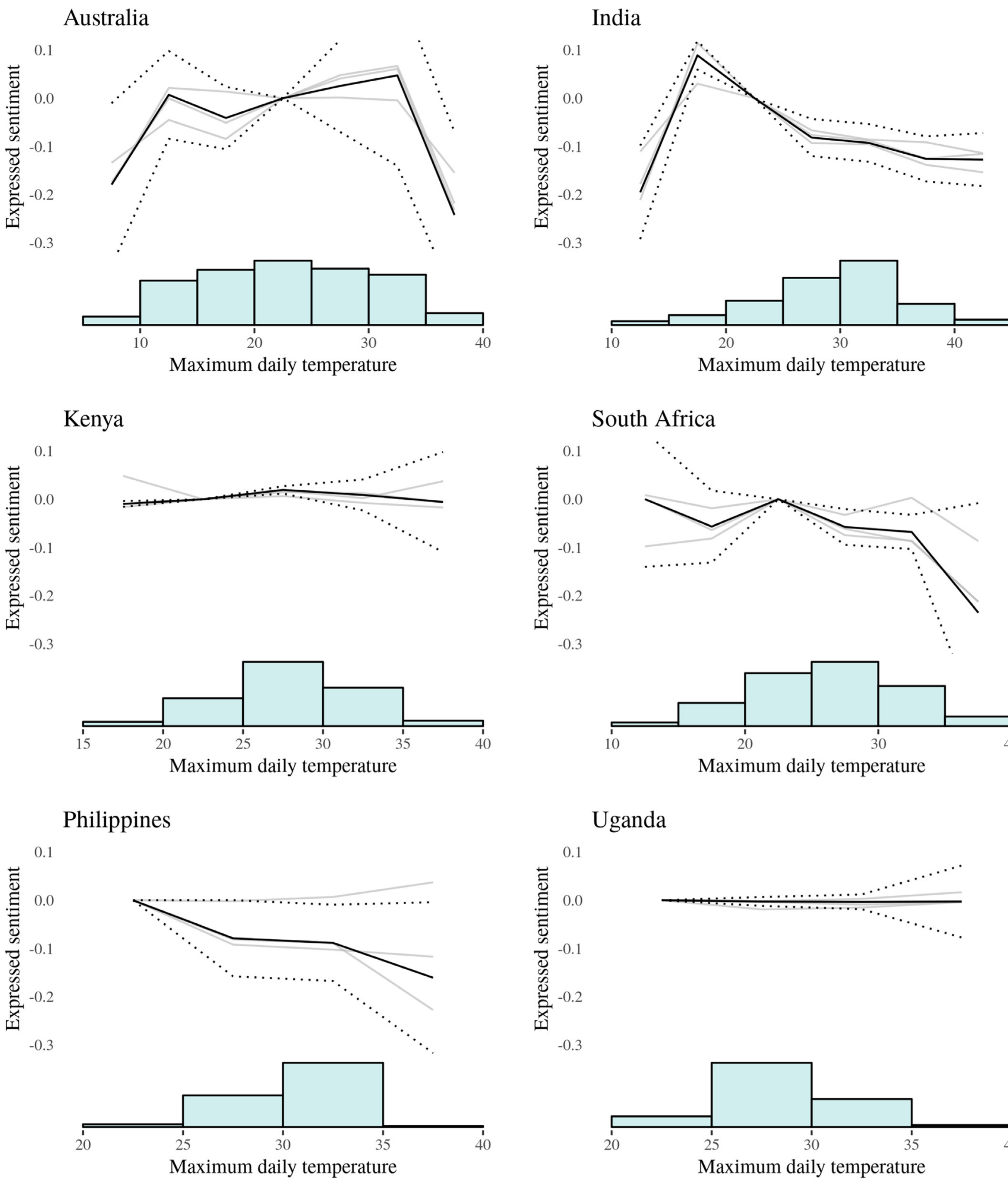

Fig. 7. Effect of temperature on expressed sentiment around the world *Notes:* Subfigures document temperature responses for six countries. Solid lines show the regression coefficients on temperature and represent the difference (measured in standard deviations) in state-date sentiment for the temperature bin T_b relative to 20–25 C, controlling for state, month, and year fixed effects. Light and solid lines represent estimates using alternative specifications of fixed effects including: state plus month of sample, state by month of year plus year, and state plus date. Dotted lines show 95% confidence intervals, with standard errors clustered by state.

averages of these data for each administrative unit in my sample to obtain the temperature and precipitation experienced by an average individual in that administrative unit on that day. After combining these data, I estimate the following specification:

$$\overline{S}_{sd} = f(T_{sd}) + P_{sd} + \phi_s + \phi_m + \phi_y + \varepsilon_{sd} \tag{4}$$

Let s and d index administrative unit, or state, and date. $\bar{S}_{sd}$ is the state-day average of the VADER measure of sentiment described in Section 2. Ts_d is the maximum daily temperature in a state, and $f(T_{sd})$ is a flexible function of temperature, which I implement in practice using five degree bins of maximum daily temperature. ϕ_s, ϕ_m, and ϕ_y represent state, month, and year fixed effects, respectively. ε_{sd} is the idiosyncratic error term, clustered by both state and date of sample, and the regression is weighted using the average number of scored tweets in each state.[16] Fig. 7 documents the findings.

For these countries, I find evidence that moderate temperatures tend to be preferred. Australia, India, and to a limited degree Kenya all exhibit evidence of a preference for moderate temperatures over very hot or very cold temperatures. In South Africa and the Philippines, I find evidence of preferences for moderate temperatures over hot temperatures alone, but no evidence that cold temperature are less preferred. Australia's response seems to most closely resemble that of the United States, while India's response shows a gradual but not sharp distaste for increasingly warm temperatures. Uganda shows no discernible difference in expressed sentiment in response to various temperature, and Kenya's response is fairly limited as well.[17]

I provide these results as evidence of differences in responses for ambient temperature across the world. However, since all of these countries differ from the United States in significant ways, whether these differences also reflect differences in preferences is, to some degree, subject to speculation. Twitter users in other counties may vary in important ways relative to both their own general populations and to the Twitter population in the United States. For example, if only relatively wealthy users participate on Twitter in Uganda, then their exposure to extreme ambient temperatures may be limited due to the use of expensive air conditioning and heating technologies. Because Twitter does not make demographic information available for their users, I am not able to directly evaluate the comparability of samples across countries. Nevertheless, the broad strokes of these findings is that there is some congruence, if not total agreement, in the sentiment response to temperature around the world. I view further pursuit of this question as a valuable avenue for future work.

6. Discussion

By using the contemporaneous responses of expressed sentiment on social media to temperature variation as a proxy for underlying preferences for temperature, I provide an alternative to traditional nonmarket valuation techniques. This approach allows me to estimate nonlinear responses of sentiment to temperature and to account for unobserved variation across both space and time and to identify spatial and seasonal differences of preferences for temperature. The set of validation and valuation exercise I demonstrate provide an interpretive baseline and way forward for future valuation exercises following this method.

This new approach is not without its drawbacks. The formation of emotional state is undeniably complex: the physical, biological, and psychological bases for human emotions remain only partly understood (Russell, 1980), and the distillation of that complexity to a single dimensional affective scale abstracts away from important nuances regarding the formation of emotion, not to mention its relationship between economic definitions of experienced utility. Although I am able to show that the users who choose to geolocate their tweets are not observationally different than the larger set of Twitter users (Appendix A.3.1), I am only able to provide suggestive evidence on the degree to which these findings would generalize to the full population (Appendix A.3.2). Finally, the estimates of the value of expressed sentiment Section 4.2, while important for improving our understanding of climate impacts, unavoidably rely on strong assumptions.

Despite these limitations, this paper makes several contributions to the literature. It introduces a new method and data source to estimate preferences for and valuations of public goods while simultaneously accounting for possible unobservable cross-sectional and seasonal variation. It reveals previously unobservable geographic and seasonal preferences for temperature and provides suggestive evidence of adaptive capacity in this area. It also demonstrates how NLP can facilitate previously intractable economic analyses of preferences for nonmarket goods and suggests a psychological channel through which other impacts of climate change may operate. Finally, it provides the first estimates of similar impacts in countries other than the United States. Broadly, this work provides additional evidence that changes in the amenity value of climate are an important component of the cost of climate change.

Notes: Panels document the response of the expressed sentiment to temperature for each of the four regions, where regions are defined by quartiles of average daily maximum temperature during the sampling frame. Dark red line is the splined response function estimated from the entire dataset, with knots at the 25th, 50th, and 75th percentile of experienced daily maximum temperatures. Lighter gray lines are splines estimated similarly using bootstrapped samples. All lines are normalized such that the maximum value of each spline is 0. Regressions include CBSA or administrative unit, month, and year fixed effects. Standard errors are clustered by CBSA or administrative unit and date.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpubeco.2020.104161.

[16] Because there are substantially fewer tweets in these countries compared to the United States, the specifications I estimate in this section are necessarily simpler than those given in Equation (1).

[17] Figure A.13 replicates Figure 7 but includes the U.S. response as well for the purpose of direct comparison.

References

Abatzoglou, John T., Brown, Timothy J., 2012. A comparison of statistical downscaling methods suited for wildfire applications. Int. J. Climatol. 32 (5), 772–780.

Albouy, David, Graf, Walter, Kellogg, Ryan, Wolff, Hendrik, 2016. Climate amenities, climate change, and American quality of life. J. Assoc. Environ. Resour. Econ. 3 (1), 205–246.

Allcott, Hunt, Lockwood, Benjamin B., Taubinsky, Dmitry, 2019. Regressive sin taxes, with an application to the optimal soda tax. The Quarterly Journal of Economics 134, no. 3 (May), 1557–1626.

Angrist, Joshua D., Pischke, Jörn-Steffen, 2008. Mostly Harmless Econometrics: An Empiricist's Companion.

Auffhammer, Maximilian, 2013. Quantifying Intensive and Extensive Margin Adaptation Responses to Climate Change: A Study of California's Residential Electricity Consumption (Working Paper).

Auffhammer, Maximilian, Aroonruengsawat, Anin, 2011. Simulating the impacts of climate change, prices and population on California's residential electricity consumption. Clim. Chang. 109, 191–210.

Auffhammer, Maximilian, Hsiang, Solomon, Schlenker, Wolfram, Sobel, Adam, 2013. Using weather data and climate model output in economic analyses of climate change. Rev. Environ. Econ. Policy 7 (2), 181–198.

Baylis, Patrick, Nick Obradovich, Yury Kryvasheyeu, Haohui Chen, Lorenzo Coviello, Esteban Moro, Manuel Cebrian, and James H Fowler. 2018. "Weather impacts expressed sentiment." *PLoS ONE* 13, no. 4 (April): e0195750.

Bentham, Jeremy, 1789. An Introduction to the Principles of Morals and Legislation. *An Introduction to the Principles of Morals and Legislation.* Oxford University Press, January, In *The Collected Works of Jeremy Bentham*.

Burke, Marshall, Emerick, Kyle, 2015. Adaptation to Climate Change: Evidence from US Agriculture. Economic Policy, American Economic Journal.

Burke, Marshall, Hsiang, Solomon, Miguel, Edward, 2015. Climate and conflict. Annual Review of Economics 7, 577–617.

Card, David, Dahl, Gordon B., 2011. Family violence and football: the effect of unexpected emotional cues on violent behavior. Q. J. Econ. 126 (1), 103–143.

Carleton, Tamma, Hsiang, Solomon, 2016. Social and economic impacts of climate. Science 353 (6304).

Cline, William R., 1992. The Economics of Global Warming (Peterson Institute for International Economics).

Daly, Christopher, Gibson, Wayne P., Taylor, George H., Johnson, Gregory L., Pasteris, Phillip, 2002. A knowledge-based approach to the statistical mapping of climate. Clim. Res. 22 (2), 99–113.

Dell, Melissa, Jones, Benjamin F., Olken, Benjamin A., 2014. What do we learn from the weather? The new climate-economy literature. J. Econ. Lit. 52 (3), 740–798.

Dennisenn, J., Butalid, Ligaya, Penke, Lars, Van Aken, Marcel A.G., 2008. The effects of weather on daily mood: a multilevel approach. Emotion 8 (5), 662–667.

Deryugina, Tatyana, Hsiang, Solomon, 2014. Does the Environment Still Matter? Daily Temperature and Income in the United States (NBER Working Paper).
Deschênes, Olivier, Greenstone, Michael, 2011. Climate change, mortality, and adaptation: evidence from annual fluctuations in weather in the US. Am. Econ. J. Appl. Econ. 3 (4), 152–185.
Diaz, Delavane B., 2014. Evaluating the Key Drivers of the US Government's Social Cost of Carbon: A Model Diagnostic and Inter-Comparison Study of Climate Impacts in DICE, FUND, and PAGE (Working Paper).
Diener, Ed, 2000. Subjective well-being. Am. Psychol. 55 (1), 34–43.
Dodds, Peter Sheridan, Danforth, Christopher M., 2010. Measuring the happiness of large-scale written expression: songs, blogs, and presidents. J. Happiness Stud. 11 (4), 441–456.
Dodds, Peter Sheridan, Harris, Kameron Decker, Kloumann, Isabel M., Bliss, Catherine A., Danforth, Christopher M., 2011. Temporal patterns of happiness and information in a global social network: Hedonometrics and twitter. PLoS One 6 (12), e26752.
Dolan, Paul, Peasgood, Tessa, White, Mathew, 2008. Do we really know what makes us happy? A review of the economic literaure on the factors associated with subjective well-being. J. Econ. Psychol. 29 (1), 94–122.
Easterlin, Richard A., 2001. Income and happiness: towards a unified theory. Econ. J. 111 (473), 465–484.
Feddersen, John, Metcalfe, Robert, Wooden, Mark, 2012. Subjective Well-Being: Weather Matters; Climate Doesn't SSRN Electronic Journal. (no. 627).
Gilbert, C.J., Hutto, Eric, 2014. Vader: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text (In Eighth International Conference on Weblogs and Social Media).
Graff Zivin, Joshua, Hsiang, Solomon, Neidell, Matthew, 2018. Temperature and human capital in the short and long run. J. Assoc. Environ. Resour. Econ. 5 (1), 77–105.
Hannak, Aniko, Eric Anderson, Lisa Feldman Barrett, Sune Lehmann, Alan Mislove, and Mirek Riedewald. 2012. "Tweetin'in the rain: Exploring societal-scale effects of weather on mood." In *Sixth International AAAI Conference on Weblogs and Social Media.*
Heckman, James J., 1979. Sample selection Bias as a specification error. Econometrica 47 (1), 153–161.
Houser, Trevor, Robert Kopp, Solomon M. Hsiang, Michael Delgado, Amir Jina, Kate Larsen, Michael Mastrandrea, Shashank Mohan, Robert Muir-Wood, and D. J. Rasmussen. 2014. American Climate Prospectus: Economic Risks in the United States.
Hsiang, Solomon, 2010. Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America. Proc. Natl. Acad. Sci. 107 (35), 15367–15372.
Hsiang, Solomon, 2016. Climate Econometrics (Annual Review of Resource Economics).
Hsiang, Solomon, Jina, Amir, 2014. The Causal Effect of Environmental Catastrophe on Long-Run Economic Growth (NBER Working Paper).
Interagency Working Group on Social Cost of Carbon, 2013. Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866. Technical report, United States Government.
IPCC, 2014. *IPCC Fifth Assessment Report.* Technical Report. IPCC, Cambridge, United Kingdom and New York, NY, USA.
Kahneman, Daniel, Krueger, Alan B., 2006. Developments in the measurement of subjective well-being. J. Econ. Perspect. 20 (1), 3–24.
Kahneman, Daniel, Sugden, Robert, 2005. Experienced utility as a standard of policy evaluation. Environ. Resour. Econ. 32 (1), 161–181.
Kenrick, D.T., MacFarlane, S.W., 1986. Ambient temperature and horn honking: a field study of the heat/aggression relationship. Environ. Behav. 18 (2), 179–191.
Landsea, Christopher W., Franklin, James L., 2013. Atlantic hurricane database uncertainty and presentation of a new database format. Monthly Weather Review 141, no. 10 (October), 3576–3592.
Levinson, Arik, 2012. Valuing public goods using happiness data: the case of air quality. J. Public Econ. 96 (910 (October)), 869–880.
Livneh, Ben, Rosenberg, Eric A., Lin, Chiyu, Nijssen, Bart, Mishra, Vimal, Andreadis, Kostas M., Maurer, Edwin P., Lettenmaier, Dennis P., 2013. A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: update and extensions. J. Clim. 26 (23), 9384–9392.
Lobell, David B., Burke, Marshall, Tebaldi, Claudia, Mastrandrea, Michael D., Falcon, Walter P., Naylor, Rosamond L., 2008. Prioritizing climate change adaptation needs for food security in 2030. Science 319 (5863), 607–610.
Mackerron, George, 2012. Happiness economics from 35000 feet. J. Econ. Surv. 26 (4), 705–735.
Marshall, Alfred, 1890. Principles of Economics. pp. 1–323.
Medhat, Walaa, Hassan, Ahmed, Korashy, Hoda, 2014. Sentiment analysis algorithms and applications: a survey. Ain Shams Engineering Journal 5 (4 (December)), 1093–1113.
Morstatter, Fred, Jürgen Pfeffer, Huan Liu, and Kathleen M Carley. 2013. "Is the sample good enough? comparing data from twitter's streaming api with twitter's firehose." In Seventh international AAAI conference on weblogs and social media.
Nielsen, Finn Årup. 2011. "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs." *arXiv.*
Nordhaus, William D., 1991. To slow or not to slow: the economics of the greenhouse effect. Econ. J. 920–937.
Nordhaus, William D, and J.G. Boyer. 2000. Warming the World: Economic Models of Global Warming. 232. MIT Press.
Pak, Alexander, Paroubek, Patrick, 2010. Twitter as a corpus for sentiment analysis and opinion mining. In LREC 10, 1320–1326.
Pearce, David, 2002. An intellectual history of environmental economics. Annu. Rev. Energy Environ. 27 (1), 57–81.
Pennebaker, James W, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. The development and psychometric properties of LIWC2015. Technical report.
Ranson, Matthew, 2014. Crime, weather, and climate change. J. Environ. Econ. Manag. 67 (3), 274–302.
Rehdanz, Katrin, Maddison, David, 2005. Climate and happiness. Ecol. Econ. 52 (1), 111–125.
Robinson, W.S., 1950. Ecological correlations and the behavior of individuals. Am. Sociol. Rev. 15 (3).
Rose, S., 2014. The Social Cost of Carbon: A Technical Assessment. Technical report, Electric Power Research Institute.
Russell, James A., 1980. A circumplex model of affect. J. Pers. Soc. Psychol. 39, 1161–1178.
Schlenker, Wolfram, Roberts, Michael J., 2009. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. Proc. Natl. Acad. Sci. 106 (37), 15594–15598.
Sinha, Paramita, Caulkins, Martha L., Cropper, Maureen L., 2018. Household location decisions and the value of climate amenities. J. Environ. Econ. Manag. 92, 608–637.
Stern, Nicholas, 2006. The Economics of Climate Change 662.
Taylor, Karl E., Stouffer, Ronald J., Meehl, Gerald A., 2012. An Overview of CMIP5 and the Experiment Design.
United States Government Accountability Office, 2014. Development of Social Cost of Carbon Estimates. Technical report July, United States Government Accountability Office.
Welsch, Heinz, Kühling, Jan, 2009. Using happiness data for environmental valuation: issues and applications. J. Econ. Surv. 23 (2 (April)), 385–406.