Temperature and Temperament: Evidence from a billion tweets

Patrick Baylis*

November, 2015†

Job Market Paper

Abstract

What is the welfare cost of environmental stress? The change in amenity values resulting from temperature increases may be a substantial unaccounted-for cost of climate change. Because there is no explicit market for climate, prior work has relied on cross-sectional variation or survey data to identify this cost. This paper presents an alternative method of estimating preferences over nonmarket goods which accounts for unobserved cross-sectional and temporal variation and allows for precise estimates of nonlinear effects. Specifically, I create a rich dataset on hedonic state: a geographically and temporally dense collection of updates from the social media platform Twitter, scored using a set of both human- and machine-trained sentiment analysis algorithms. Using this dataset, I find limited evidence of temperature effects on hedonic state in low temperatures and strong evidence of a sharp decline in hedonic state above 70°F. This finding is robust across all measures of hedonic state and to a variety of specifications.

*University of California, Berkeley. Energy Institute at Haas; 207 Giannini Hall, Berkeley, CA 94720-3310; Phone: (507) 581-1807; E-mail: pbaylis@berkeley.edu. I am grateful to Maximilian Auffhammer, Severin Borenstein, and Solomon Hsiang for their invaluable suggestions, and to Michael Anderson, Judson Boomhower, Josh Blonz, Marshall Burke, Fiona Burlig, Tamma Carleton, Richard Carson, Aluma Dembo, Meredith Fowlie, Walter Graf, Sylvan Herskowitz, Elizabeth Sadoulet, and to seminar participants at UC Berkeley, the AERE Annual Conference, the CU Environmental and Resource Economics Workshop, and the Heartland Workshop at Illinois. Errors are entirely my own.

†The latest version of this paper is available at: http://are.berkeley.edu/candidate/patrick-baylis.

1 Introduction

Acute environmental stressors like typhoons, hurricanes, and other marked changes in the external environment are known to have large economic costs (Hsiang and Jina 2014). However, slower-moving changes in the environment, such as temperature increases due to climate change, tend to have subtler economic effects. The empirical climate impacts literature has set out to estimate the size of these effects, largely focusing on estimating the indirect impacts of climate change, *e.g.* temperature-induced changes in income, crime, or natural disasters.

Because temperature is a nonmarket good, estimating the "direct" impacts of climate change has proven to be more challenging.[1] Prior work estimates that individuals would be willing to pay between 1% and 3% of their incomes to avoid a one °F increase in summer temperatures (Cragg and Kahn 1997; Sinha and Cropper 2013; Albouy et al. 2013). However, these costs are almost exclusively identified using cross-sectional variation in climate and therefore rely on important assumptions about unobservable variation in climate preferences. A separate literature uses subjective well-being surveys in order to estimate preferences for temperatures. While these papers do not estimate costs directly, they are able to account for some unobserved cross-sectional variation by using fixed effects (Levinson 2012; Feddersen, Metcalfe, and Wooden 2012), but yield conflicting results due to limited statistical power.

This paper estimates preferences over nonmarket goods using an alternative approach that addresses both the identification and statistical power concerns described above. I construct a geographically and temporally dense collection of more than a billion geocoded social media updates from the platform Twitter. To estimate preferences for temperature, I code each tweet using a set of sentiment analysis algorithms designed to extract hedonic state from natural language.[2] Using more than a billion Twitter updates, or "tweets", I

[1] "Direct" here refers to the hypothesized welfare impact of changing average daily while holding the other indirect impacts of temperature constant. This can also be viewed as the amenity value of changes in climate.

[2] Since climate change is projected to manifest primarily as changes in average temperature for most of the world (IPCC 2014), I focus specifically on temperature as the nonmarket good of interest. Still, this approach generalizes to many other nonmarket goods that are experienced heterogeneously across space and time.

resolve identification concerns by accounting for correlated unobservables at the county, neighborhood, and even individual level with an extensive set of fixed effects and while simultaneously accounting for unobserved state-specific seasonal variation.

I define hedonic state as a one-dimensional measure of mood ranging from negative to positive. The four measures I use span a range of sentiment analysis techniques designed to elicit mood from natural language. Two measures are specified using expert- and crowd-sourced dictionaries that map words to numerical scores. A third measure scores tweets by whether or not they contain profanity. The final measure trains a machine-learning algorithm using the Twitter updates that contain emoticons, *e.g.* ":)" or ":(", to predict the emotional content of the full set of tweets. I validate these measures by demonstrating their change across day of week and hours of day, and, following Card and Dahl (2011), as a result of nearby NFL teams' wins or losses.

Using geographical information attached to the Twitter updates, I match the measures of emotional state to daily weather conditions at the precise location of the user. My identifying assumption is that temperature draws are as good as random after accounting for spatial and seasonal fixed effects. Allowing temperature to enter the econometric model flexibly, I find limited evidence of temperature effects on hedonic state in low temperatures and strong evidence of a sharp decline in hedonic state above 70°F. The difference in hedonic state between 60-70°F and 80-90°F is significant and comparable in size to the average difference in hedonic state between Sundays and Mondays.

I conduct a series of robustness checks to further explore the results and to test for potential sources of bias. First, I demonstrate consistent effects in both direction and standardized magnitude across all measures of hedonic state, indicating that the results are not driven by measure design. I additionally confirm that the observed effects are not generated by correlated compositional changes in the sample across temperatures by estimating a model with individual fixed effects. Next, I examine heterogeneity in the response by hour of day and document that the baseline results are driven by temperatures experienced during daylight hours. To understand the discrepancy between the estimates of winter temperature preferences in my results and prior work, I document heterogeneity in the effects by season. I exploit human sensitivity to humidity to examine the effect of temperatures outside the bulk of the support of historical data, finding a remarkable decrease in hedonic state resulting from the combination of high temperatures and humidity. I consider the effects of adaptation by comparing the slope of the heat response function across regions with different historical temperatures and use downscaled climate projection data to estimate the projected effects of changes in temperature on hedonic state across the United States. Following prior work, I implement a back-of-the-envelope calculation to back out the monetary costs implied by my estimates.

Sections 2 and 3 sketch the conceptual framework and review the related literature. Section 4 describes the data and sentiment analysis algorithms I use and section 5 lays out the empirical approach and identifying assumptions. Section 6 reports the baseline results, section 7 documents robustness checks and extensions, and section 8 concludes.

2 Conceptual framework

A simple conceptual framework helps illustrate the problem of estimating the costs of climate change. Consider a representative consumer with a utility function defined over temperature T, a composite of goods whose consumption utility is affected by temperature c_T, and a composite of goods whose consumption utility is unaffected by temperature c_N. Let this consumer choose the quantity of c_T and c_N she consumes, subject to their prices p_T and p_N and income I. T is assumed to be exogenous to the consumption choice[3] and thus does not enter the budget constraint. The consumer's problem is as follows:

$$\max_{c_T, c_N} U = U(T, c_T, c_N) \text{ s.t. } p_T c_T + p_N c_N \leq I \tag{1}$$

[3]A two-period model would allow consumers to choose T by changing location, in doing so alter the prices and utility value of both c_T and c_N. I focus on the simpler model for clarity.

To maximize utility, the consumer chooses c_T^* and c_N^* optimally such that $\frac{\partial U}{\partial c_T} = \lambda p_T$ and $\frac{\partial U}{\partial c_N} = \lambda p_N$, where λ is the shadow value of relaxing the budget constraint by one unit. Note that c_N^* is implicitly a function of T through the budget constraint, since changes in T may alter c_T^*. Consider two types of exogenous shocks: a change in T and a change in I.

$$\frac{dU}{dT} = \frac{\partial U}{\partial T} + \frac{\partial U}{\partial c_T^*}\frac{\partial c_T^*}{\partial T} + \frac{\partial U}{\partial c_N^*}\frac{\partial c_N^*}{\partial T} \tag{2}$$

$$\frac{dU}{dI} = \frac{\partial U}{\partial c_T^*}\frac{\partial c_T^*}{\partial I} + \frac{\partial U}{\partial c_N^*}\frac{\partial c_N^*}{\partial I} \tag{3}$$

Combining these, the monetary cost of a unit change in temperature is the compensating variation x that keeps the consumer on her original indifference curve:

$$\frac{dU}{dT} + x\frac{dI}{dT} = 0 \tag{4}$$

$$\frac{\partial U}{\partial T} + \frac{\partial U}{\partial c_T^*}\frac{\partial c_T^*}{\partial T} + \frac{\partial U}{\partial c_N^*}\frac{\partial c_N^*}{\partial T} + x\left[\frac{\partial U}{\partial c_T^*}\frac{\partial c_T^*}{\partial I} + \frac{\partial U}{\partial c_N^*}\frac{\partial c_N^*}{\partial I}\right] = 0 \tag{5}$$

In principle, a researcher could estimate x using a choice experiment in which consumers are asked to state their willingness to pay to avoid a degree rise in average temperature. In reality, multiple market failures make this design infeasible. First, information is not perfect: the costs of climate change are incompletely understood even by researchers in the field, and likely less so by the average consumer (IPCC 2014). Moreover, even with perfect information, present-day consumers may have a discount function that is inappropriate to capture the full costs of climate change, since those costs will likely be endured mostly by generations who have yet to be born.[4] Third, the choice experiment as presented suffers from a collective action problem, since the benefits of climate change mitigation are spread across the entire world.

Instead, in practice, the literature estimates the effect of temperature on different sectors of the economy and calculates the cost of climate change to be the sum of the value of the

[4] The problem of how to properly discount future climate damages is particularly thorny one. See Stern (2006) and Nordhaus (2007) for two views of this question.

projected changes in those sectors. As an example, let c_T^C be crime risk, which has been documented by Ranson (2014) to increase in temperature. Researchers estimate $\frac{\partial c_T^C}{\partial T}$ and multiply by estimates of willingness to pay to avoid crime. Integrated Assessment Models (Hope 2006; Nordhaus and Sztorc 2013; Antoff and Tol 2014) and the Social Cost of Carbon (United States Government 2013) aggregate $\frac{\partial c_T}{\partial T}$ for all possible impacts, combine and sum over these impacts and multiply by expected temperature changes to get the net benefit of climate change.[5]

The climate impacts literature has historically focused on estimating $\frac{\partial c_T}{\partial T}$, which I refer to as the "indirect" effects of climate change. Because these effects on welfare are driven through other factors, measuring indirect impacts relies on the combination of measurement of preferences for these indirect factors and predicted changes in these factors due to climate change, but not measurement of direct preferences for temperature itself. This paper instead measures $\frac{\partial U}{\partial T}$, the "direct impacts" of climate change. $\frac{\partial U}{\partial T}$ can be thought of as the amenity value of temperature, or the marginal change in hedonic state associated with a marginal change in temperature.[6]

3 Background

Economists have studied the economic impacts of climate change for more than two decades (Nordhaus 1991; Cline 1992), but the recent availability of panel datasets and advanced econometric techniques have made possible the identification of the causal effects of changes in temperature on a wide variety of outcomes (Dell, Jones, and Olken 2014).

[5] This is, of course, a highly simplified and incomplete description of how IAMs and the SCC are constructed. For more complete descriptions see the listed citations or the summary in Diaz (2014). This framework does not imply that the net benefit must be less than zero, but most current estimates find this to be the case empirically.

[6] It is reasonable to argue that this paper too examines an "indirect impact", since psychological changes, for example, could be viewed as a kind of mechanism. I use the term "direct" here to refer to mechanisms in which weather alters individuals' day-to-day experience of the world. I make use of the fact that the main drivers of hedonic state are an individual's underlying hedonic state and transient changes in the state of the world (Kahneman and Krueger 2006). This suggests that the primary effects I observe are likely to correspond closely with the prior literature's definition of amenity value.

Early work in the climate impacts literature focused on identifying the effects of changes in climate on agricultural output (Mendelsohn, Nordhaus, and Shaw 1994; Schlenker, Hanemann, and Fisher 2005; Deschênes and Greenstone 2011). One notable finding from this literature is that the response function of yields to temperature changes contains important non-linearities: yields tend to increase slightly up to a threshold, after which they decrease sharply, implying severe negative effects on yields under many climate change scenarios (Schlenker and Roberts 2009).

Recently, scholars have directed their attention to non-agricultural impacts of climate change. Dell, Jones, and Olken (2012) use country-level data to identify the effect of weather variation on aggregate economic outcomes, and find that higher temperatures reduce economic growth in poor countries. Using county-level data on U.S. incomes, Deryugina and Hsiang (2014) conduct a similar analysis in the United States and document the negative impacts of warm weekday temperatures on county income, and provide suggestive evidence that these effects are driven by changes in the productivity level of basic economic units such as workers and crops. Burke, Hsiang, and Miguel (2015b) expand these findings to the global scale, providing evidence that economic productivity declines in high temperatures for both rich and poor countries.

Other work has examined the effect of temperature on economic productivity. Graff Zivin and Neidell (2014) study the effect of temperature on time allocation using county-level data, finding that the quantity allocated to labor decreases in higher temperatures. In related work, Graff Zivin, Hsiang, and Neidell (2015) study the effect of temperature on cognitive performance, using a panel of test scores to find statistically significant decreases in math (but not reading) performance when the temperature rises above 79°F.

A substantial literature has examined the relationship between climate and conflict. Burke, Hsiang, and Miguel (2015a) conduct a meta-analysis of the available estimates and find that one standard deviation increase in temperature increases interpersonal and intergroup violence by 2.4% and 11.3%, respectively.

Other work has looked at the relationship between temperature and electricity usage, or load. Auffhammer and Mansur (2014) review the existing literature and document the need for additional panel data studies to properly control for unobserved cross-sectional variation. Existing panel data studies, such as Deschênes and Greenstone (2011) find a significant increase in energy consumption due to high temperatures using state-level averages, while Auffhammer and Aroonruengsawat (2011) use detailed billing data from California to document within-state heterogeneity in load responses.

Individuals without access to air conditioning are more susceptible to the effects of temperature changes. Understanding the adoption of temperature-regulating technology informs predictions about future effects of climate change. Auffhammer (2013) uses a two-stage model to estimate both intensive and extensive margin increases in air conditioning due to climate change. Relatedly, Davis and Gertler (2015) study air conditioner adoption in Mexico, predicting close to full adoption within a few decades, primarily due to adoption resulting from income growth rather than changes in climate.

Climate-induced changes in mortality have been studied by Deschênes and Greenstone (2011) and Barreca et al. (2013), among others. The first estimates a 3% increase the age-adjusted mortality rate in the United States, while the second documents the importance of air conditioning in mitigating the temperature-mortality relationship observed in the first half of the 20th century.

Many of the estimates described 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 (United States Government 2013). In particular, three IAMs are used to construct this estimate. They are the Dynamic Integrated Climate-Economy Model (Nordhaus and Sztorc 2013), or DICE, the Climate Framework for Uncertainty, Negotiation, and Distribution (Antoff and Tol 2014), or FUND, and the Policy Analysis of the Greenhouse Effect (Hope 2006), or PAGE. IAMs integrate economic and ecological models to weigh the costs and benefits of global warming.[7] The link between warming and damages (or benefits) is modeled in each using either a single damage function or a set of damage functions.

DICE uses a global damage function that is built from separate, sector-level damage functions. The author uses a time of use survey to value nonmarket amenities, resulting in a quadratic damage function between temperature and amenity value. This formulation estimates *benefits* from changes in amenity value that actually exceed the total market impacts in the United States (Nordhaus and Boyer 2000). PAGE includes damage functions for both economic and noneconomic changes, the parameters of which are generated from the findings of the third IPCC report (Hope 2006), which did not include nonmarket amenity values directly (IPCC 2001). FUND uses a set of damage functions, but these do not include a separate function for nonmarket amenities (Antoff and Tol 2014).

That the direct effect of climate change could entail a significant welfare cost follows from the observation that people have preferences over weather. Still, estimating these preferences and the cost associated with shifting the temperature distribution has been challenging, due primarily to the fact that there is no market for temperature. Two main approaches emerge, the first using hedonic price models and the second using life satisfaction surveys.

The hedonic price approach recovers willingness-to-pay (WTP) for climate amenities by comparing cross-sectional differences in wages and climate amenities after controlling for other covariates (for an early example, see Hoch and Drake (1974)). Cragg and Kahn (1997) model the locational choices of migrants and finds that movers are willing to pay about about 1.5% of annual income for an additional one °F in winter and -1.2% of annual income for an additional °F in summer.[8] Sinha and Cropper (2013) also look at migration decisions using a discrete model of location choice to estimate the rate of substitution between wages and climate amenities. The authors estimate that the marginal WTP for a one °F increase is

[7]For a detailed review of the three IAMs listed, see Diaz (2014) or Rose (2014).

[8]The authors split results up by age and estimate different of WTP. Estimates are the unweighted average of the estimates in Table 7 of Cragg and Kahn (1997), adjusted for a one °F increase and divided by the annual household income of the movers in their sample.

between 1% and 5% of income in winter, and between -3% and -1.5% of income in summer. Finally, Albouy et al. (2013) use a hedonic framework and data from the 2000 census to find a marginal WTP for a one °F increase in winter to be between 0.5% and 1% of income, and in summer between -2.5% and -1% of income.[9]

The hedonic approaches described above are appealing because they identify implicit demand for climate using households' observed choices on where to live. Using estimates of the differential between wages and costs of living, they are also able to back out a WTP for climate. However, because the models estimate the effect of climate characteristics, which are mostly stable across time, the coefficients are identified using cross-sectional variation. This approach requires the assumption that there is no unobserved variation that is correlated with both climate and with the differential between wages and costs of living, an assumption that may be violated by the existence of unobservable cultural factors, for example.

The survey approach uses 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, Peasgood, and White 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 important unobserved variation across time and space. For example, responses to questions about one's well-being may depend on cultural factors that differ across people and geographies and could be driven by the interaction between the interviewer and the interviewee.

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[10]. Only two control for unobservable cross-sectional variation using panel data mod-

[9] I take the estimates of MWTP for a day at 40°(80°) F from Table 3 in Albouy et al. (2013) and divide by the distance between 40 (80) and 65 to get the MWTP for one degree at that temperature.

[10] Howarth and Hoffman (1984) collect data from 24 Canadian male university students over a period of 11 days and find that higher temperatures improve hedonic state. Keller et al. (2005) study the effect of weather on both cognition and hedonic state and find that pleasant weather, *i.e.* moderate temperature or barometric pressure, is associated with higher hedonic state, although they find that higher temperatures in the summer are associated with lower hedonic state. Dennisenn et al. (2008) also find that higher temperatures reduce hedonic state, while Klimstra et al. (2011) follow nearly 500 adolescents and find large individual differences

els. Levinson (2012) uses 6,035 surveyed individuals from the General Social Survey to find a inverse-U shaped relationship between temperature and happiness, though the paper is primarily focused on the effects of pollution. Feddersen, Metcalfe, and Wooden (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, while temperature is not found to have an impact.

The mixed results in this literature suggest that statistical power is constrained by the combination of the high variance in SWB responses driven by non-temperature factors and relatively small sample sizes. Most studies in this area have either relied heavily on small sets of repeated samples, which limits external validity, or large sets of non-repeated samples, which raises concerns about unobserved cross-sectional variation. Additionally, since these are survey-based approaches, it is possible that the size of the effects could be driven in part by the interaction between interviewer and subject, if those interactions change in warmer weather.

Temperature preferences are likely to be correlated with unobservable factors that vary across both space and time, and may be small relative to preferences for other goods and services. To control for both geographic and temporal variation while maintaining sufficient power to identify small, non-linear effects would require a prohibitively expensive survey of subjective well-being. Instead, I use sentiment analysis algorithms to detect hedonic state from a large set of Twitter data.

Sentiment analysis is a natural language processing technique designed to elicit subjective feeling from textual data. There are a small number of a studies in computer science and computational linguistics that have used sentiment analysis techniques on Twitter data. Dodds and Danforth (2010) create an dictionary-based algorithm that scores individual

in their responses to hedonic state. Lucas and Lawless (2013) find little effect of temperature on hedonic state using state-level data.

tweets using a mapping of more than ten thousand English words to scores of hedonic state. The authors demonstrate that although the algorithm sometimes misclassifies individual sentiments, in aggregate it produces plausible results (Mitchell et al. 2013). Other work uses machine learning techniques to predict the sentiment of tweets (Pak and Paroubek 2010). Related work has used sentiment analysis on Twitter data to predict economic outcomes of interest. Notably, Bollen, Mao, and Zeng (2011) find that collective hedonic state can help predict the stock market, Eichstaedt et al. (2015) use measures of county-level hedonic state to predict heart disease mortality, and Gerber (2014) shows that local Twitter hedonic state can improve local predictions of crime. To my knowledge, no studies have used sentiment-analyzed Twitter data in a causal setting.

By collecting a large, geographically and temporally detailed dataset, I am able to account for unobserved variation across both time and space. The size of my sample and the empirical techniques I use allow me to precisely estimate the effect of temperature in the midst of substantial unrelated variation in hedonic state. Additionally, I am able to identify non-linearities in the temperature response function and previously unexplored dimensions of heterogeneity. The sentiment analysis methods I use are applied identically across space and time and not subject to the same potential biases inherent in administering or taking surveys.

4 Data

I generate four measures of hedonic state using data from Twitter and match these to weather data at the tweet level. Table 1 describes sample characteristics. The first panel shows the count, mean, median, minimum, and maximum of the measures of hedonic state I describe later in this section, the second and third panel describe the weather data used, and the fourth panel summarizes the number of tweets by individual, grid cell, and county in the data.

Twitter data

Created in 2006, Twitter is a social networking site built around the public exchange of short (<140 characters) Twitter updates. Since its founding, Twitter has become one of the most popular websites on earth, with 288 million active users sending over 500 million tweets per day.[11] Tweets are considered to be in the public domain.

Twitter's Streaming API[12] 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.[13] I collect the vast majority of geolocated tweets produced within my sample period, which ends in October 2015.

Geo-located tweets are those for which the user has consented to have his or her location information shared. The location information is either produced using the exact latitude and longitude of the user if the tweet is sent from a phone, or from a reverse-geocoding algorithm that derives the latitude and longitude from location information entered by the user. In principle, Twitter limits the total number of tweets delivered 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 rarely comes to more than 1% of the total tweets worldwide (geocoded and otherwise). Over the course of the sample I collect, the percentage of missed tweets is fewer than 0.01% of the total available. A sampling of the tweets is available in appendix. Figure 1 is a map of Twitter update density where the shading for each pixel represents the log of the total number of tweets in the dataset for each grid cell, a 4 km^2 area. The distribution of tweets closely resembles the population distribution in the United States.

To construct a measure of hedonic state, I rely on the sentiment analysis techniques

[11] Population summary statistics from https://about.twitter.com/company.

[12] https://dev.twitter.com/streaming/overview.

[13] There are two gaps, from June 26th to July 12th, 2014, and from September 18th to October 27th, 2014, corresponding to periods of time when the streaming client was unable to connect to the Streaming API.

described in section 2. As previously discussed, no single measure of hedonic state will perfectly capture the hedonic state of the individual at time of update. Accordingly, I construct four separate measures of hedonic state from the text in the Twitter updates: Expert, Crowd-sourced, Profanity, and Emoticon measures.

Table 1 shows the raw measures of hedonic state in the sample. Count is the total counts of Twitter updates in the dataset, irrespective of whether or not covariate data was obtained for those tweets.[14] Note that although the Profanity and Emoticon scores are binary variables and thus would be expected to have median zero or one, the table displays the median of the average measure in a grid-cell day, weighted by count of tweets. The descriptive statistics are constructed using the raw measures, but the difference in means and scales suggests that standardization will be important for empirical comparison. As such, the measures are standardized (mean zero and unit standard deviation) for the empirical estimation described in section 5. The fourth panel shows the number of tweets per individual, grid cell, and county in my dataset over the entire sample. There is considerable variation in the tweet volume across these groups. Los Angeles county, for example, is responsible for more nearly 5% of the sample, while a single user accounts for nearly a quarter million tweets.

Table 2 shows the correlations between the four measures. As expected, all of the measures are positively correlated with each other, reflecting general agreement. Some of the correlations are low, particular those between the Profanity measure and the other measures, likely reflecting the considerable differences in the ways these measures are constructed. The complexity of measuring hedonic state, as demonstrated by the relatively limited agreement of the measures presented here, suggests the importance of considering the effects across all measures rather than just one. I next detail the construction of each measure.

[14] A proportion of tweets in my sample came from locations just outside the continental United States, which is outside the range of the meteorological data I use.

Expert measure

The Expert measure is constructed by using an expert-created dictionary that maps from words to scores of hedonic state. The AFINN-111 dictionary contains 2,477 words scored using integers between -5 and 5, where -5 indicates negative hedonic state and 5 indicates positive hedonic state. The dictionary focuses on words that are indicative of hedonic state, and was created by Nielsen (2011) to analyze language typically used in microblogging. The dictionary is refined from an earlier dictionary built by psychologists to assess the affective state (the psychological equivalent concept to hedonic state) of written texts Bradley and Lang 1999. The measure is constructed using the following procedure:

1. Tweets are cleaned of extraneous punctuation, URLs, hashtags, and other nonsense characters.
2. Tweets are checked for weather-related stopwords to avoid a mechanical correlation generated by individuals discussing aberrant weather patterns. If a stopword is found, the given tweet is scored as missing.
3. For each word in a tweet that matches an entry in the AFINN dictionary, the corresponding measure of hedonic state is retrieved.
4. The overall score for a given tweet is the average score for word matched in step 3. If no words in the tweet matched to the dictionary, then the measure is scored as missing.

Let $j = 1..J$ index words w_j in a cleaned tweet and let $k = 1..K$ index the tuple (w_k, s_k), which are the word-score pairings in the dictionary. The Expert measure E^E for a given tweet is:

$$E^E = \frac{\sum_{j=1}^{J} \sum_{k=1}^{K} \mathbb{1}[w_j = w_k] \times s_k}{\sum_{j=1}^{J} \sum_{k=1}^{K} \mathbb{1}[w_j = w_k]}$$

The AFINN-111 dictionary is specifically designed to include only words that are indicative of emotional state. For example, the tweet “happy anniversary mom and dad” has five words, but only “happy” is included in the AFINN-111 dictionary, and has rating $s_{\text{happy}} = 3$. The overall score for the tweet is just the average across scored words, which in this case is just $E^E = 3$ for this tweet, since only “happy” was scored. Similarly, the tweet “i can’t watch matt cry" is given $E^E = -1$, since the word "cry" has $s_{\text{cry}} = -1$. More examples of words with positive, neutral, and negative sentiment are available in the appendix.

Crowd-sourced measure

The Crowd-sourced measure E^C is constructed similarly to the Expert measure, but the dictionary used is that provided by and described in Dodds and Danforth (2010). The authors crowd-source a dictionary of more than 10,000 words by using the Mechanical Turk service, which outsources tasks to external users. Users were asked to rate each word on a scale from 1 to 9, where 1 indicated negative emotional state and 9 indicated positive emotional state, and scores were averaged across users to get a single score for each word.

Unlike the Expert-measure, the Crowd-sourced measure scores most commonly-used words regardless of whether they are likely to be indicative of underlying hedonic state. Taking the same example tweets from the section above, "happy anniversary mom and dad" has $E^C = 6.976$, since the words in the tweet have scores of 8.3, 6.7, 7.64, 5.22, and 7.02, respectively. "i can't watch matt cry" has $E^C = 4.428$ with word scores of 5.92, 3.42, 5.7, 5.26, and 1.84 for each word in the tweet, respectively. More examples of words with positive, neutral, and negative sentiment are available in the appendix.

Emoticon measure

While lexical affinity approaches such as the Expert and Crowd-sourced methods are frequently used in the sentiment analysis literature, they can be sensitive to the particular word-sentiment score mapping chosen by the researcher. To complement these approaches, I construct a measure of hedonic state that classifies tweets as positive or negative using a small set of assumptions and machine learning techniques.

Emoticons are text-based translations of common facial expressions. In general, emoticons can indicate positive moods, e.g. ":)" or ":-)", or negative moods, e.g. ":(" or ":-(". One possible approach would be to limit the sample to tweets that contain either a positive or a negative emoticon. However, since emoticons appear in only about 2% of the sample, this approach substantially limits power. Since most Twitter updates with emoticons contain words as well, researchers in computational linguistics have employed machine learning techniques to leverage the subset of tweets with both emoticons and words to predict the sentiment of the entire set of tweets (Go, Bhayani, and Huang 2009; Kouloumpis, Wilson, and Moore 2011).

I collect a training dataset consisting of all tweets containing either positive or negative emoticons. For this training dataset, I code the hedonic state as binary and assume its polarity (1 if positive, 0 if negative) is indicated by the attached emoticon. For a full list of the emoticons used to collect this dataset, see the appendix. Next, I train an effective, computationally efficient machine learning classifier, Multinomial Naïve Bayes,[15] to estimate whether particular words are more likely to be associated with positive or negative emoticons. Finally, I use this classifier to compute the Emoticon measure E^M of the population of tweets.

Why Naïve Bayes?

Developing a predictive model as described above could be done using a variety of tools, ranging in complexity from ordinary least squares to ensemble techniques that incorporate multiple machine-learning algorithms. I select Naïve Bayes, a relatively simple machine learning approach, as my predictive modeling technique of choice for a few reasons. First, Naïve Bayes has been shown to be as, if not more, effective than more complex machine learning techniques for text classification tasks (Go, Bhayani, and Huang 2009). Second, Naïve Bayes is computationally efficient, an important considering when using a dataset with a billion observations.

Multinomial Naïve Bayes

The principle behind Naïve Bayes that uses Bayes' Theorem to estimate the probability

[15] I use the scikit-learn implementation of the Multinomial Naive Bayes classification algorithm (Pedregosa et al. 2011).

that a given word (called a unigram) or set of words (called bigrams, trigrams, etc.) are associated with a particular sentiment. In particular, I use a version of the technique called Multinomial Naïve Bayes, which works well with collections of words such as tweets. Pang, Lee, and Vaithyanathan (2002) report that unigrams perform as well or better than bigrams, and described the Naïve Bayes classification as follows: sentiment class $s^* \in \{0, 1\}$ is assigned to tweet d, where

$$s^* = \arg\max_s P(s|d)$$
$$P(s|d) = \frac{P(s)\prod_{m=1}^{M} P(w_m|s)}{P(d)}$$

$P(s|d)$ is the probability that tweet d has sentiment s. w_m represents a particular unigram (word) out of a total of M possible words. $P(s)$ is the overall average sentiment, estimated in the training set, while $P(w_m|s)$ is the likelihood of observing word w given sentiment s, estimated in the training set. Laplacian smoothing is used to ensure that $P(w_m|s) \neq 0$. $P(d)$ is the probability of observing a particular tweet d, but since it is a scalar it does not affect the choice of s^* and is therefore not included in the estimation procedure. The predicted sentiment obtained from the represent a simple scoring system: tweets whose content is predicted to be positive are scored 1, while those with negative content are scored 0.

Other machine learning techniques I also test other machine learning classification algorithms. To do so, I train different classifiers using a random subsample of the training set of tweets with emoticons, then cross-validate the predicted sentiment classification using the remainder of the training set. I test Multinomial Bayes, Stochastic Gradient Descent (SGD), and Support Vector Machines[16] (SVM), and find that Multinomial Bayes performs as well or better as SGD and SVM, which are more complicated techniques. See the appendix

[16] For detailed descriptions of Stochastic Gradient Descent and Support Vector Machines, see Pedregosa et al. (2011).

for details. I find that Multinomial Bayes achieves accuracy of around 80%, which happens to match the observed percentage with which human raters of sentiment tend to agree (Wilson, Wiebe, and Hoffmann 2005).

Profanity measure

Finally, to provide a measure with a more intuitive interpretation, I compile a list of more than 300 profanities and scored each tweet for the presence or absence of these profanities.[17] In the sentiment analysis literature, this approach is called a "keyword spotting" approach. I calculate the Profanity measure as follows: $E^P = \mathbb{1}[Profanity \in Tweet]$. The assumption that drives the Profanity measure is that, in general, profanities indicate negative hedonic states. As a result, it should estimate the opposite relationship relative to the other measures.

Validation exercises

I conduct a series of validation exercises to tie the measures to phenomena that most readers will find intuitive. Figure 2 shows the measures by day of week. Since the measures use different scales, they are standardized to have mean = 0 and standard deviation = 1. The measures move in concert, though the Profanity measure has the opposite sign, as expected. To calibrate the results later in the paper, it is useful to note that the average difference in sentiment score between Sunday and Monday is approximately 0.01σ across measures.

Following Card and Dahl (2011), I conduct a separate validation exercise using 2014 National Football League (NFL) game outcomes. Twitter users within 80 kilometers of an NFL stadium are matched to their home team, and their average hedonic state in the remainder of a day following a win or loss is measured. The results are shown in Figure 3. The difference between a win and a loss is approximately 0.01σ across all measures, though the difference is larger in the Expert measure and smaller in the Profanity measure. This corresponds roughly to the difference in hedonic state observed between Sundays and

[17] List of profanities available from http://www.noswearing.com/dictionary, which maintains a comprehensive database of swear and curse words.

Mondays.

Weather data

This work focuses primarily on the effects of temperature, but some specifications will include other weather variables such as precipitation, cloud cover, humidity, and wind speed.

Temperature and precipitation

I use daily data on minimum temperature, maximum temperature, and precipitation at 4 km^2 grid cell across the contiguous United States. These data are from PRISM Climate Group's AN81d dataset and are produced using the Parameter-elevation Relationships on Independent Slopes Model, which interpolate measurements from more than 10,000 weather stations (Daly et al. 2002). The data capture a high degree of both spatial and temporal heterogeneity in weather. The second panel in Table 1 describes sample statistics for the PRISM data, weighted by tweet volume.

Other weather data

Prior work suggests that other weather variables besides temperature and precipitation may be important determinants of hedonic state (Dennisenn et al. 2008; Levinson 2012). I collapse hourly data on proportion of day that was overcast, visibility in kilometers, relative humidity, station pressure, and wind speed from 2,162 weather stations included in the Quality Controlled Local Climatological Data (QCLCD) data from NOAA to the daily level. I drop any station-months in which more than 10% of the observations were missing. To fill in the remaining observations, I compute the inverse-distance weighted quantile of a given measure from nearby stations and estimate the value of that measure for the station with the missing data using the cumulative distribution function of that station. This gives me a balanced panel of weather station observations. I then use inverse distance weighting to impute these measures of weather on a grid similar to that of the PRISM data. Maps of average daily measurements from within my sampling frame are available in the appendix. All measures of weather show substantial geographic and temporal heterogeneity. The third panel in Table 1 describes sample statistics for the QCLCD data, weighted by tweet volume.

5 Empirical specification

I estimate a panel fixed effects model to identify the effect of temperature on hedonic state. As is standard in the climate impacts literature, the model is identified under the assumption that temperature is as good as random after accounting for unobserved cross-sectional and seasonal variation (Dell, Jones, and Olken 2014). To this end, I include PRISM grid cell and state-by-month of year fixed effects in my empirical specification. Following prior work that estimates marked non-linearities in weather impacts across multiple economic outcomes (Schlenker and Roberts 2009; Ranson 2014; Graff Zivin, Hsiang, and Neidell 2015), I estimate the effects on hedonic state as a non-linear function of temperature by including temperature in the model using a set of ten °F bins. Following standard practice, 60-70° F is the omitted category, such that the coefficient on, say, 80-90° F should be interpreted as the effect on hedonic state caused by replacing a 60-70° F with a day which has an average daily temperature of between 80-90° F (Barreca et al. 2013; Albouy et al. 2013). The empirical model I estimate is given by:

$$\overline{E}_{gd} = \sum_{b \neq 60\text{-}70}^{B} \beta_b T_{gd}^b + \phi_g + \phi_{sm} + \varepsilon_{gd} \tag{6}$$

Let g, s, d, m index grid cell, state, day, and month of year, while b is an index over temperature bins. $\overline{E}_{gd}$ is the grid cell-day average of one of the four measures of hedonic state described in section 4. Because my temperature measure varies at the grid cell-day, taking the grid-cell day average of the hedonic state measures and weighting by the total number of tweets in that grid-cell day estimates the same point estimates and standard errors as would be estimated using a model where each observation represented a single tweet (Wooldridge 2002), while reducing computation time substantially.

T_{gd}^{b} is a dummy variable = 1 if the daily average temperature in a grid cell falls within the associated ten-degree bin b. I estimate a similar model with precipitation in bins as the primary right-hand side variable, where the zero precipitation bin is the omitted category.

The grid cell fixed effects ϕ_g control for time-invariant unobservables across space. For example, individuals with higher income tend to have higher levels of life satisfaction (Easterlin 2001) and may be inclined to locate in areas with generally pleasant climate. By including ϕ_g, I identify the coefficients of interest using within-cell variation over time. I also include state-by-month fixed effects ϕ_{sm} to account for state-specific within-year trends between temperature and hedonic state, *e.g.* the well-known seasonal variation of human emotion and seasonal changes in weather.

The coefficients β_b are identified using within-grid cell variation in weather that is not absorbed by state-month fixed effects and map out a non-linear response function between temperature and hedonic state[18]. Per the discussion in Dell, Jones, and Olken (2014) and in Hsiang, Burke, and Miguel (2013), I focus my analysis on temperature as the exclusive weather variable in my primary specification to avoid the over-controlling problem, though in other specifications I include a large set of weather covariates. To allow for spatial and temporal correlation in the data, I cluster the standard errors two ways, by state (48)[19] and by week of sample (50)[20].

[18]Because of the high dimensionality of both the grid-cell and state-by-month fixed effects, estimates are obtained using the *reghdfe* module in Stata (Correia 2014).

[19]I exclude Alaska and Hawaii due to limitations of the Twitter Streaming API and because the PRISM weather data are confined to the continental United States.

[20]I also run a model that allows for spatial correlation up to 16 km and temporal correlation of up to 7 days using spatial standard errors as described by Conley (2008) and implemented using code from Hsiang (2010). The standard errors are smaller than those obtained using the two way clustering described here, suggesting that the confidence intervals presented here may be conservative.

6 Baseline results

This section presents the main results of the paper, estimates of the non-linear relationship between temperature and hedonic state. Results from model (6) are tabulated for the Expert and Emoticon measures are displayed in column (4) of Tables 4 and 5, respectively, and plotted for all measures in Figure 4. The point estimates β_b are the conditional means of hedonic state relative to the omitted category, 60-70°F. In order to make the scores comparable, all measures are standardized such that they have zero mean and unit standard deviation.

Each column in Tables 4 and 5 displays point estimates and standard errors for increasingly robust sets of fixed effects and controls. Column (1) is the ordinary least squares (OLS) estimate, which finds a strong negative effect of high temperatures. There is also mixed evidence of effects in colder temperatures, though the point estimates are inconsistent across measures: negative for the Expert measure and positive for the Emoticon measure. However, the coefficients in this model could suffer from the classical omitted variables bias problem for many reasons, such as endogenous sorting, different word choice norms, income levels, and seasonal variation in temperature and hedonic state. For example, the northern United States tends to be more affluent and experiences lower average temperatures. If affluence has a positive effect on hedonic state, this would introduce a downward bias in the coefficients on high temperatures.

To account for these unobservables, column (2) adds county and month fixed effects, standard in the climate impacts literature (Dell, Jones, and Olken 2014). These estimates are identified using within-county fluctuations in temperature, after accounting for seasonal variation, to identify the effect of temperature on hedonic state. The point estimates for the higher temperature bins are approximately a quarter of the size to those estimated in column (1), strongly suggesting that unobserved variation was responsible for a portion of the OLS estimates, while the estimates for the lower temperatures are not statistically significant.

Since seasonal variation may be substantially different across geographies, column (3) replaces the month fixed effects with state-by-month fixed effects, allowing for different seasonal trends by states. The point estimates and standard errors are not substantially altered by this additional.

Column (4) replaces the county fixed effects with grid cell fixed effects. Since the observations in this model are grid cell-day averages, it is possible that unobservable within-county variation could bias the point estimates in the prior models. This is more likely in regions of the country with large county sizes, such as California. For example, Los Angeles county is more than 10,000 km^2 and contains a large number of PRISM grid cells. Returning again to the potential confound of affluence, if wealthier individuals tend to live in areas with more pleasant temperature, the warmer temperature coefficients could be biased away from zero. To account for this possibility, the preferred specification includes grid cell fixed effects. There is some slight attenuation of the point estimates relative to the second and third models, but the difference is not substantial and may be driven by the absorption of weather variation in between grid cells.

Column (5) adds a robust set of weather controls, following Feddersen, Metcalfe, and Wooden (2012), who found that the addition of other weather covariates into their regression model removed the effect of temperature, suggesting that other weather variables, correlated with temperature, may be the actual drivers of mood. Using data from both the PRISM and QCLCD datasets described in section 4, I control for the daily temperature spread, precipitation, cloudiness, visibility, station pressure, relative humidity, and average wind speed. In principle, adding weather controls may introduce a "bad controls" problem, since weather covariates may themselves by causally influenced by temperature[21]. In practice, doing so has little effect on the point estimates or standard errors of the temperature variables in the model. Of the added weather controls, only cloudiness and relative humidity have a statistically significant effect on hedonic state, both negative.

[21] The bad controls problem is described in Angrist and Pischke (2008) and discussed in the context of weather and climate regressions in particular by Dell, Jones, and Olken (2014) and in Hsiang, Burke, and Miguel (2013).

Column (6) adds day-of-week fixed effects, motivated by the possibility that there may be compositional differences in the Twitter user base by day of week that could confound the results. For example, more temperature-sensitive users using Twitter on Mondays could induce a bias in the coefficients. Again, the results are not substantially changed. For this reason, the remainder of the analysis focuses on results using variations on equation (6), the preferred specification.

Turning to Figure 4, the four measures of hedonic state all clearly reject the null of no effect of temperature on hedonic state, and provide strong evidence of a negative relationship between hedonic state and temperature above an average daily temperature 70° F. There is some suggestion that temperatures slightly below 70° F are preferred. This may be the result of averaging the maximum and minimum temperature: while an instantaneous temperature of 70° F is generally viewed as pleasant, days with average temperature of 70° F typically contain maximum temperatures of around 80° F. Below 60° F, the effect of temperature on hedonic state is flat with wide confidence intervals, suggesting substantial heterogeneity in the response to cooler temperatures.

For higher temperatures, the negative relationship between temperature and hedonic state resembles that estimated by Albouy et al. (2013) and other work in the locational choice literature, who find that individuals would pay to avoid warm temperatures in summer. However, it differs in lower temperatures, where they also observe a willingness to pay for warmer temperatures in winter. Some portion of this puzzle may be explained by seasonal variation in temperature preferences. I explore the possible sources of this heterogeneity in section 7.3.

All measures estimate that the difference between a 60-70° F day and an 80-90° F day to be approximately 0.01σ. As a point of comparison, this difference is approximately comparable to the average difference in hedonic state between tweets sent on Sunday versus tweets sent on Monday (see Figure 2).

The effect is highly robust across measures. The plots in Figure 4 use identical axes, demonstrating that the point estimates are remarkably similar. The size of the confidence intervals differ in expected ways: the Expert and Crowd-sourced measures have the largest confidence intervals due to the fact that, unlike the Emoticon and Profanity measures, they do not code the full sample of tweets. The Profanity measure effect is slightly smaller at the right tail of the temperature distribution than the other measures; this attenuation may be attributable to the use of profanities to capture both positive and negative sentiment.

7 Robustness checks and extensions

The previous section established the baseline results. This section extends the results with a series of robustness checks and extensions: I account for possible endogenous selection into sample using individual fixed effecst, examine differences in winter and summer responses to temperature, disaggregate the response by hour of day, document increased effects of humidity, discuss adaptation using heterogeneity in the response by average historical temperature, and use a preliminary method to estimate a willingness-to-pay for temperature from these data.[22]

7.1 Accounting for endogenous sample selection

Including grid cell fixed effects in the empirical model accounts for sorting even by micro-climates, since PRISM grid cells are 4 km^2 geographic areas. By contrast, most panel datasets in the climate impacts literature are limited by their geographic detail to the county or state level. In this respect, model (6) is highly robust to unobserved variation. However, since participation in Twitter, and social media in general, is a voluntary choice on the part of a given user, failing to account for potential endogeneity of Twitter participation may induce a sample selection bias (Heckman 1979). In this setting, the selection bias of greatest concern

[22] In the appendix, I include a series of additional robustness checks. I estimate the model using five degree bins, cubic splines, minimum and maximum daily temperature rather than average temperature, and lagged temperature variables.

is compositional sorting: samples at different temperatures may include sets of users with different unobservable characteristics. For example, if individuals with higher or lower native affect become more likely to create Twitter updates in different temperatures, the coefficients could be capturing this compositional change in the sample rather than a change in average hedonic state.

Since the data I collect include an identifier for the unique user responsible for a given tweet, I control for potential composition sorting in my sample using user fixed effects. To do so, I estimate the following model:

$$E_{it} = \sum_{b \neq 60\text{-}70}^{B} \beta_b T_{gd}^{b} + \phi_i + \phi_{sm} + \varepsilon_{it} \tag{7}$$

This model substitutes user fixed effects, ϕ_i, for the grid cell fixed effects, ϕ_g, in model 6.[23] The model requires the use of the entire unaggregated sample of observations in my dataset; because the right-hand side of model (7) includes variation at the individual level, it not possible to compute the same coefficients using grid cell-day averages. Let i and t be the user and the time a status update was sent, respectively. E_{it} is one of the four measures of hedonic state.

I estimate similar results using model 7, which includes user-level fixed effects and the full, non-aggregated sample of tweets. For comparison, I overlay the estimates with those obtained using model 6 in Figure 5. I include only the Expert and Emoticon measures for brevity, but results for other measures are similar and included in the appendix. The shape of the estimated temperature response is similar, and the point estimates do not notably change between models, suggesting that compositional sorting is not a confounding feature in this setting. For this reason, and because running the model on the non-aggregated sample is computationally costly, I conduct the remainder of the analysis using grid cell-day averages

[23] A possible concern with model (6) is that the same individual tweeting from different locations may be endogenously determined with weather, *e.g.* a family choosing to vacation in California to avoid a cold snap in Minnesota. To address this bias, I estimate a specification that also includes PRISM grid cell fixed effects alongside the individual fixed effects. The results, available in the appendix, are qualitatively the same.

as the observations.

7.2 Effect by hour of day

To better understand the sources of variation in the data, I disaggregate the response by hour of day. This serves two purposes: first, it is a check that temperature is the main driver of the result: because people are most exposed to temperature during daytime hours, the effect of temperature on hedonic state should be smaller in magnitude at night. Second, the hourly distribution of effect size may help to inform our understanding of likely adaptation margins and potential policy responses. For example, if individuals only express strong preferences for temperature during non-work hours, *e.g.* during lunchtime or after work hours, then this could suggest that workplace adaptation has already fully mitigated the effect of air conditioning.

For ease of comparison, I restrict the sample to days in which the mean temperature is greater than 60°F and replace the temperature bin variables used in model (6) with a single temperature variable.[24] I capture the heat response by hour by estimating 24 different models, where each sample contains only observations from the given hour of the day in local time. Each econometric model takes the following form:

$$E_{gd} = \gamma_q T_{gd} + \phi_g + \phi_{sm} + \mu_{gd} \text{ where } T_{gd} \geq 60^\circ \tag{8}$$

The models are identified by comparing tweets within a given hour in the same grid cell on warm days to tweets within the same hour on cooler days, after accounting for seasonal variation. Figure 6 plots the results, where each point represents a single point estimate and the thick horizontal bar shows the across-hour response mean. Notably, the effect sizes are

[24] I also estimate a separate set of models that use the entire sample, and find qualitatively similar results. However, due to the non-linear of the response function and the limitations on statistical power imposed by running each hourly regression separately, the effects are statistically insignificant.

larger during daylight hours, supporting the interpretation of the baseline results as driven by changes in temperature. For all measures, the effects are largest between 7 a.m. and 8 p.m., with no notable difference at noon or after 5 p.m. If the primary driver of hedonic response is exposure to temperature, then the lack of a higher response during lunch or after close of business (when workers are more likely to be outside) is puzzling. One explanation is that the noon to 1 p.m. hour is not sufficiently consistent for lunch to estimate an effect, and that the increased exposure induced by the end of the workday is counterbalanced by individuals returning to their homes. However, given the size of the standard errors, even a relatively large effect could be masked by noise in the estimates.

7.3 Response by season

Model (6) estimates an average response function over the entire year. One notable difference between the results in figure 4 and those in the hedonic literature is that I do not find a negative effect of cold temperatures on hedonic state. Pooling the response over both winter and summer months could mask seasonal heterogeneity in the response, since individuals may respond differently to a relatively warm day in winter than they would in summer.

The richness of my data allow me to capture this heterogeneity by specifying a model that allows the effects to differ by winter and summer months. I define the winter period as running from November through April, and the summer as May through October.[25]

$$\overline{E}_{gd} = \sum_{b \neq 60\text{-}70}^{B} \beta_b T_{gd}^b + \sum_{b \neq 60\text{-}70}^{B} \delta_b T_{gd}^b \times \mathbb{1}[\text{Winter}]_m + \phi_g + \phi_{sm} + \varepsilon_{gd} \tag{9}$$

Figure 7 documents the response function by seasons for the Expert and Emoticon measures[26] by plotting β_b (summer) and $\beta_d + \delta_b$ (winter) on top of each other. While colder temperatures have little effect on hedonic state in the winter, they drive positive emotions in the summer.

[25] To estimate sufficient overlap in the temperature distributions across seasons, I choose bin breaks that run from 40 to 80 by ten °F.

[26] As before, results using other two measures reflect the same qualitative result and are included in the appendix.

The relationship between warm temperatures and hedonic state is negative in both seasons. The lower panels show the estimate of the difference between the two seasons. The difference between summer temperatures and winter effects of 40°to 60°F is statistically significant.

This result accounts for the unusual finding in the baseline results that temperatures cooler than 60-70° F are preferred. The finding appears to be driven entirely by cooler temperatures in summer, while cooler temperatures in winter do not appear to affect hedonic state. Still, that there is no estimable negative relationship between cooler temperatures and mood even in the winter suggests that the *ex post* hedonic experience of temperature may differ from *ex ante* decision to pay for warmer winter temperatures estimated by the locational choice literature.

7.4 Humidity

A major challenge in projecting impacts of climate change is that historical temperatures provide relatively little support for projections on the right tail of the temperature distribution. Since increases in the mean and variance of climate distributions are likely and since some of the impacts demonstrate important threshold effects, a better understanding of this right tail is important. Because human preferences are the objects of interest in my setting, I can exploit a feature of human physiology, susceptibility to humidity, to estimate preferences for extremely high temperatures.

Higher humidity makes people feel warmer, all else equal. Humidity reduces the rate at which perspiration evaporates, which in turn reduces the rate at which humans can cool themselves by sweating; accordingly, humidity has been found to be an important factor in workplace productivity (Kjellstrom, Holmer, and Lemke 2009). I investigate whether the influence of humidity is observable in the data by replacing the piecewise function of temperature with a piecewise function of heat index as the right-hand side variable in model (6). Heat index is a measure computed from air temperature and relative humidity that represents the apparent temperature, or the equivalent temperature without humidity. It is equal to air temperature when air temperature is less than 80°F[27].

The estimates obtained from this regression are plotted in Figure 8. Below a heat index of 100, the results are qualitatively similar to the main result. Above 100, however, hedonic state declines sharply across all measures. Since this effect occurs at the tail of the heat index distribution, I interpret these results as clear evidence that high humidity at high temperatures causes substantial negative emotions on hedonic state and suggestive evidence that people would strongly disprefer very high temperatures, holding humidity constant.

7.5 Adaptation

The extent to which individuals adapt to changing climate regimes is an important input to understanding the cost of climate change (Barreca et al. 2013; Burke, Hsiang, and Miguel 2015a). Since hedonic state is known to adapt to changes in circumstances, it is possible that the hedonic response to temperature could fully adjust to changes in the mean of the climate distribution. Put another way, if the change in hedonic state due to temperature is solely a function of the distance from the mean temperature, then the change in the mean of the climate distribution will have no effect on welfare. With sufficient data, one way to test for this possibility would be to use a long differences approach similar that implemented by Burke and Emerick (2015). Because my data are a much shorter time series, I provide suggestive evidence of future adaptation by comparing temperature response functions across areas with different climates. Because observable cross-sectional variation could bias this result, it is not a definitive test prediction of adaptation, but finding no statistically significant inter-regional difference in response slope could be considered evidence of no adaptation. I estimate linearized responses by quintiles of grid cell historical average temperature.

$$\overline{E}_{gd} = \sum_{q=1}^{5} \gamma_q T_{gd} \times \mathbb{1}[\text{Quint} = q] + \phi_g + \phi_{sm} + \mu_{gd} \text{ where } T_{gd} \geq 60^\circ \tag{10}$$

[27] The heat index calculation I use is from the National Weather Service, see http://www.wpc.ncep.noaa.gov/html/heatindex_equation.shtml for details.

Specification 10 differs from specification 7 only in the functional form with which temperature enters the estimating equation. The results in Figure 9 suggest a weaker response to changes in temperature in areas with more extreme climate, both warm and cold. The change in average temperature between bins is roughly 5-6°F, which corresponds to predicted end-of-century global average increases in temperature under the highest emissions scenarios. I interpret these findings as suggestive evidence in favor of incomplete adaptation to warmer temperatures, a careful consideration in modeling the potential adaptive capability of individuals to higher temperatures.

I caution that this comparison is cross-sectional and should be taken as suggestive evidence at best. This could be reflective of multiple factors: first, areas with higher temperatures are also more likely to have higher air conditioner penetration. Second, these areas are likely to attract (or at least deter less) individuals with a tolerance for warmer temperatures. To the extent that climate change is a slow process, this comparison may be accurate, since changes in temperature could induce a similar sorting over the course of the coming century. However, if these differences capture cultural or geographic factors that will not be altered by climate change (e.g. warmer areas' access to water, for example), then these estimates may overestimate the extent of adaptation.

7.6 Climate projections

The projected effects of climate change are, on average, an increase in the mean of the climate distribution. To better understand the future impacts of climate change on hedonic state, I combine the estimates documented above with projected changes in United States climate. The thought experiment I perform is as follows: if the predicted end-of-century effects of climate change were to take place tomorrow, how should we expect hedonic state to change? By using downscaled climate data, I am able to account for likely geographic heterogeneity in climate impacts and observe how different regions of the United States may be affected. I emphasize that the projections exercises are not meant to be direct predictions of future changes in hedonic state; they are instead meant to illustrate ways in the which the amenity costs of temperature could be differentially altered in the United States. I conduct three separate projection exercises.[28]

First, I use the average response function across the United States as the basis of projection, holding that response function constant over time. The projected damages are products of the coefficients estimated in Figure 4 and the expected change in the number of days in a given bin, summed over all bins. The result of this exercise is mapped in the top left panel of Figure 10. In general, southern areas of the United States experience the greatest losses of hedonic state. This finding is driven by the findings of the climate models, which predict a large increase in the number of very hot days in this region. Because the most severe impacts of hedonic state are found in higher temperatures, these regions are most profoundly affected.

However, important differences in the response function by geographic area could have different implications for climate damages. For example, if areas in the south are already more adapted to higher temperatures, it may not be the case that they will be most adversely affected by climate change. To allow for this possibility, I estimated disaggregated response functions by quintile of average historical temperature, again using bin breaks of 40, 50, 60, 70, and 80 °F to allow for sufficient overlap across quintiles.[29] I then conduct separate projection exercises for each quintile, and aggregate the results. The top right panel of Figure 10 contains the combined map. This exercise is partly limited by statistical power in some of the quintiles, but suggestive that disaggregation is important: the most heavily affected areas are more in the middle of the country than in the south.

For the third projection exercise, I use the same disaggregated function but allow grid cells to adapt to a new temperature regime by adopting a response function their new quintile, using the historical quintile breaks. To fix ideas, suppose that there is a grid cell in Minnesota

[28]What follows is a brief description of the projection exercises I conduct. Additional detail on projection methodology is available in the appendix.

[29]See appendix for plots of the response functions by quintile.

that is in the lowest historical daily average temperature quintile. After allowing for climate change, this grid cell would now fall into the second lowest quintile using the historical temperature cutoffs. I project the effect of climate change using the response function of the second lowest quintile, which would, for example, include Kansas. This exercise allows Minnesota's response function to adjust to look more like Kansas' response function. The lower panel of Figure 10 contains this final projection exercise. This map suggests that the most affected regions are likely to be in the northern part of the country.

I emphasize that these projections are reliant on strong assumptions, in particular regarding future technological change, migration, and adaptation. I attempt to provide a margin for adaptation, both past and future, in the second and third exercises. With that in mind, these estimates suggest large changes in hedonic state due to climate change. Returning to the calibration exercise, for some areas this change would be the equivalent of replacing every Saturday and Sunday in a year with a Monday. Given the strong assumptions required to obtain this estimate, I instead focus on the important regional differences in the projected outcomes produced by varying aggregation levels and allowances for adaptation. This setting is likely not the only area in which these regional differences are important, and suggests the importance of both accounting for these differences and using them to infer adaptation behavior.

7.7 Estimating a willingness-to-pay for temperature

The evidence provided thus far demonstrates a clear relationship between hedonic state and temperature. However, to compare the magnitude of these cost of changes in hedonic state to the magnitude of costs in other sectors, it is necessary to convert the changes in hedonic state into monetary damages.[30] Following prior work, I present a *highly preliminary* method for this conversion. I emphasize that this method relies on strong assumptions and should

[30] Conversion into a monetary cost is also important for inclusion in Integrated Assessment Models (Hope 2006; Nordhaus and Sztorc 2013; Antoff and Tol 2014) or the social cost of carbon (United States Government 2013).

be interpreted as a back-of-the-envelope calculation at best.

The technique I use follows Train (2002) and Levinson (2012), the latter of which implements it to estimate the monetary cost of changes in air quality on reported life satisfaction. I estimate the following model:

$$\overline{E}_{gd} = \beta T^b_{gd} + \gamma I_b + \phi_{sm} + \varepsilon_{gd} \tag{11}$$

The major addition to the model is I_b, Census Block Group median income in thousands. β can be interpreted as the change in hedonic state induced by a one °F change in temperature, while γ is the change in hedonic state associated with a \$1,000 dollar increase in the income of an individuals Census Block Group.

I estimate and totally differentiate the above, holding $dE = 0 \rightarrow \frac{\partial I}{\partial T} = -\frac{\hat{\beta}}{\hat{\gamma}}$. This estimate can be interpreted as the willingness to substitute between a degree of temperature change and \$1,000 increase in median income. The results of this regression are displayed in Table 3. Computing the willingness to substitute across all four measures yields estimates of \$548, \$875, \$2096, and \$816 for the Expert, Crowd-Sourced, Emoticon, and Profanity measures, respectively. These estimates are largely driven by the size of the denominator γ, and constitute a 1-2% change in income relative to the median in my sample, which is in line with other results estimated in the locational choice literature.

I emphasize that this procedure requires two strong assumptions. First, it requires that $dE = 0 \Rightarrow dU = 0$, or that holding hedonic state constant is equivalent to holding utility constant. Second, it requires that within state, between-Census Block Group differences in income are as good as random. The results of this exercise should be interpreted with appropriate caution.

8 Discussion

This paper explores the relationship between temperature on hedonic state as a way to understand preferences for day-to-day temperature. The existing literature estimates large costs due to the change in amenity value driven by climate change, but does so by relying on cross-sectional variation. In this paper, I document a method that allows researchers to estimate preferences over nonmarket goods while accounting for a wide range of unobservable variation across both space and time. I accomplish this by constructing a dataset of text updates from the social media platform Twitter, which I code using human and machine-trained sentiment analysis algorithms from computational linguistics. I combine this geographically and temporally detailed measure of hedonic state with finely gridded weather data to flexibly estimate the effect of weather on mood. I find that hedonic state is unaffected by cooler temperatures, but declines sharply above 70° F. In terms of magnitudes, I estimate a difference of about 0.01σ between a day with mean temperature of 60-70°F and a day with 80-90°F, which is roughly the average difference between observed hedonic state on Sundays relative to Mondays. These results are net of short-term adaptation, *e.g.* air conditioning. Since my data are from the United States, where air conditioner penetration rates are among the highest in the world, it is likely that the relationship between temperature and hedonic state may be even more pronounced in other countries.

The negative effects of warm temperatures strongly resemble qualitative results documented using other approaches. However, the lack of a similar distaste for extremely cold temperatures, even in winter, remains a puzzle. I speculate that this apparent contradiction may illuminate a key difference between *ex ante* preferences for temperature and *ex post* hedonic responses to different temperatures. One important factor may be the relative margins for adjustment to low and high temperatures: cold days can be easily adapted to through additional clothing, but no such margin exists for hot days. Similarly, the greater penetration of heating equipment, relative to air conditioning, could play a role.

The results obtained in section 6 should be interpreted with some caution. First, users of Twitter are a selected sample, though a large one. Moreover, users who choose to enable geolocation services may be yet different from the Twitter user-base at large. The adaptive nature of hedonic state could also imply that the costs of climate change could be overstated by this analysis, though section 7.6 accounts for this possibility and negative impacts remain. Finally, the nature of the results presents challenges to monetary conversion: how much social welfare does the loss of one standard deviation of hedonic state represent? The preliminary method I demonstrate in section 7.7 provides one view, but relies on strict assumptions.

Nevertheless, this paper makes several contributions to the literature. It introduces a new methodology and data source to estimate preferences over nonmarket goods while accounting for possible unobservable cross-sectional and seasonal variation. It demonstrates how an appropriate use of sentiment analysis and machine-learning algorithms can enhance the econometric analysis of large datasets, estimates the relationship between temperature and hedonic state across multiple dimensions of heterogeneity, and suggests a psychological channel through which other impacts of climate change may operate. Additionally, this paper is one of the first to employ social media data in a rigorous causal framework. The projection exercise I conduct is unique in the literature in that I use both aggregated and disaggregated response functions to project future damages, showing that the use of disaggregated response functions and allowing areas to adapt over time substantially modifies the qualitative implications of the projection exercise. Broadly, this work provides supporting evidence that changes in the amenity value of climate are an important component of the overall costs of climate change.

Table 1: Sample characteristics

	Count	Mean	Median	Min	Max
Measures of hedonic state					
Expert	1,077,127,397	0.37	0.38	-5.00	5.00
Crowd-sourced	1,083,068,307	5.51	5.51	1.30	8.44
Profanity	1,083,498,783	0.94	0.94	0.00	1.00
Emoticon	1,083,498,783	0.79	0.80	0.00	1.00
PRISM weather					
Min temperature (F)	943,724,684	53.6	58.0	-33.9	99.3
Mean temperature (F)	943,724,684	63.3	68.4	-22.9	108.7
Max temperature (F)	943,724,684	73.1	78.3	-17.3	123.9
Precipitation (mm)	943,724,684	3.0	0.0	0.0	318.3
QCLCD weather					
Proportion overcast	918,921,992	0.2	0.1	0.0	1.0
Visibility (km)	918,921,992	15.3	15.7	0.2	132.1
Relative humidity	918,921,992	59.6	60.4	2.1	100.0
Station pressure	918,921,992	29.2	29.4	19.9	30.8
Wind speed	918,921,992	7.7	7.3	0.0	74.7
Twitter updates per ...					
Individual	10,227,302	87	9	1	240,045
PRISM grid cell	519,942	2,084	14	1	20,849,368
County	3,102	307,508	33,276	44	45,557,251

Notes: First panel shows statistics for the measures of hedonic state, second and third panels for the weather datasets. For first through third panel, one observation is a single Twitter update. First column in the fourth panel is the total number of individuals, grid cells, and counties in the sample. Second through fifth columns are the means, medians, minimums, and maximums of the count of Twitter updates by individuals, grid cells, and counties, respectively.

Table 2: Measure correlations

	Expert	Crowd-sourced	Emoticon	Profanity
Expert	1.00			
Crowd-sourced	0.59	1.00		
Emoticon	0.35	0.31	1.00	
Profanity	0.39	0.19	0.12	1.00

Notes: Table displays correlations between the four measures of hedonic state described in section 4.

Table 3: Estimating a WTP for temperature

	Expert	Crowd-sourced	Emoticon	Profanity
Mean temperature	-0.000492*	-0.000746*	-0.000784*	0.000607**
	(0.000227)	(0.000297)	(0.000296)	(0.000186)
Income ($1,000)	0.000897***	0.000853*	0.000374	-0.000744**
	(0.000136)	(0.000331)	(0.000288)	(0.000236)
Grid cell-days	17,986,266	15,059,391	18,460,020	18,460,020

Notes: Each column contains coefficients from a regression of a measure of hedonic state on temperature and median Census block group income. Measures of hedonic state described in section 4 and are standardized to have mean zero and unit standard deviation. All regressions include state by month fixed effects and are weighted by the number of tweets in a grid cell-day. Grid cells are 4 km $\times$ 4 km square cells used by the PRISM weather dataset. 95% confidence intervals estimated using two-way cluster robust standard errors on county and day-of-sample.

Table 4: Effect of temperature on hedonic state (Expert measure)

	(1)	(2)	(3)	(4)	(5)	(6)
Daily temperature T						
$T < 20$	–0.011	0.001	0.004	0.001	0.000	–0.003
	(0.007)	(0.005)	(0.008)	(0.008)	(0.008)	(0.007)
$T \in [20, 30)$	–0.013*	0.003	0.006	0.004	0.004	–0.000
	(0.006)	(0.004)	(0.007)	(0.007)	(0.006)	(0.006)
$T \in [30, 40)$	–0.015**	0.002	0.005	0.004	0.004	0.001
	(0.005)	(0.004)	(0.006)	(0.006)	(0.005)	(0.005)
$T \in [40, 50)$	–0.007	0.005*	0.007*	0.006	0.006	0.004
	(0.005)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
$T \in [50, 60)$	0.005	0.002	0.002	0.002	0.002	0.002
	(0.004)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
$T \in [70, 80)$	–0.023***	–0.009***	–0.010***	–0.007***	–0.007***	–0.005***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$T \in [80, 90)$	–0.045***	–0.014***	–0.016***	–0.012***	–0.013***	–0.010***
	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$T \geq 90$	0.006	–0.016***	–0.014***	–0.014***	–0.018***	–0.016***
	(0.007)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Grid cell-days (m.)	19.22	19.22	19.22	19.14	18.51	18.51
Twitter updates (m.)	473	473	473	473	459.9	459.9
County FE	No	Yes	Yes	No	No	No
Grid cell FE	No	No	No	Yes	Yes	Yes
Month FE	No	Yes	No	No	No	No
State×Month FE	No	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	No	No	No	Yes
Weather controls	No	No	No	No	Yes	Yes

Notes: Dependent variable is the average standardized (mean zero, unit standard deviation) Expert measure of hedonic state for a grid cell-day. Independent variables are dummies for temperature (in °F) bins. Each column is a separate regression, coefficients represent the change in standard deviations of hedonic state between a day within the associated temperature bin and a day with temperature $T \in [60, 70)$, the omitted category. Coefficients are estimated conditional on the fixed effects and controls listed. Weather controls include day-level measures of temperature range, precipitation, cloudiness, visibility, station pressure, relative humidity, and average wind speed. Grid cell-days is the count of observations in the regressions in millions. Twitter updates is the count the number of Twitter updates aggregated into the grid cell-days in millions.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

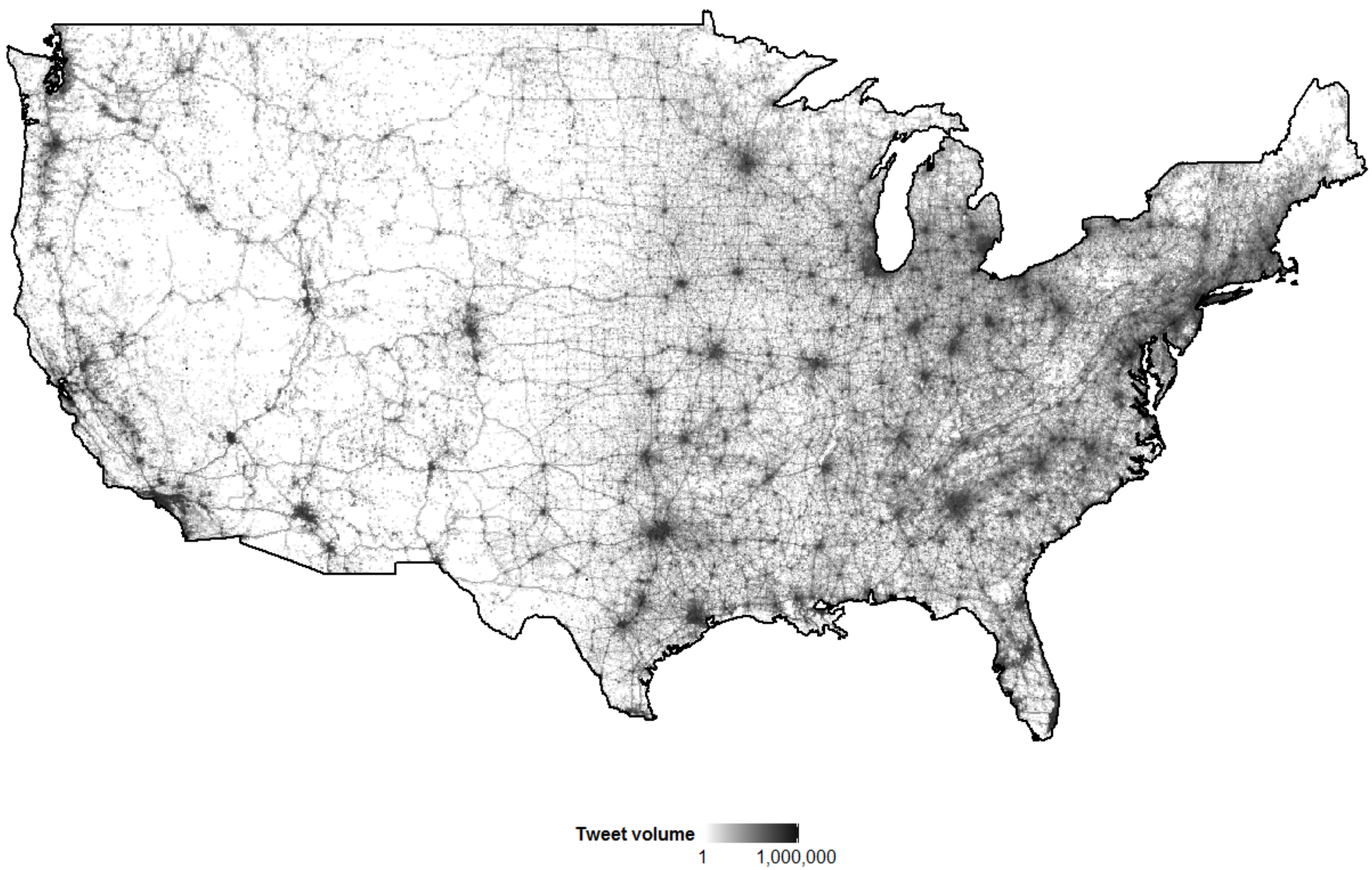
Table 5: Effect of temperature on hedonic state (Emoticon measure)

	(1)	(2)	(3)	(4)	(5)	(6)
Daily temperature T						
$T < 20$	0.010	0.001	0.000	0.000	–0.001	–0.005
	(0.006)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)
$T \in [20, 30)$	0.007	0.005	0.005	0.004	0.003	0.000
	(0.005)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
$T \in [30, 40)$	0.002	0.003	0.003	0.003	0.002	0.000
	(0.005)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)
$T \in [40, 50)$	0.002	0.002	0.002	0.002	0.002	0.001
	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$T \in [50, 60)$	0.008*	0.000	0.001	0.001	0.000	0.001
	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$T \in [70, 80)$	–0.025***	–0.009***	–0.010***	–0.006***	–0.006***	–0.005***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$T \in [80, 90)$	–0.062***	–0.013***	–0.015***	–0.010***	–0.010***	–0.009***
	(0.005)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
$T \geq 90$	–0.025	–0.015***	–0.013***	–0.012***	–0.015***	–0.013***
	(0.014)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Grid cell-days (m.)	23.46	23.46	23.46	23.36	22.59	22.59
Twitter updates (m.)	945.2	945.2	945.2	945.2	918.9	918.9
County FE	No	Yes	Yes	No	No	No
Grid cell FE	No	No	No	Yes	Yes	Yes
Month FE	No	Yes	No	No	No	No
State×Month FE	No	No	Yes	Yes	Yes	Yes
Day of week FE	No	No	No	No	No	Yes
Weather controls	No	No	No	No	Yes	Yes

Notes: Dependent variable is the average standardized (mean zero, unit standard deviation) Emoticon measure of hedonic state for a grid cell-day. Independent variables are dummies for temperature (in °F) bins. Each column is a separate regression, coefficients represent the change in standard deviations of hedonic state between a day within the associated temperature bin and a day with temperature $T \in [60, 70)$, the omitted category. Coefficients are estimated conditional on the fixed effects and controls listed. Weather controls include day-level measures of temperature range, cloudiness, visibility, station pressure, relative humidity, and average wind speed. Grid cell-days is the count of observations in the regressions in millions. Twitter updates is the count the number of Twitter updates aggregated into the grid cell-days in millions.

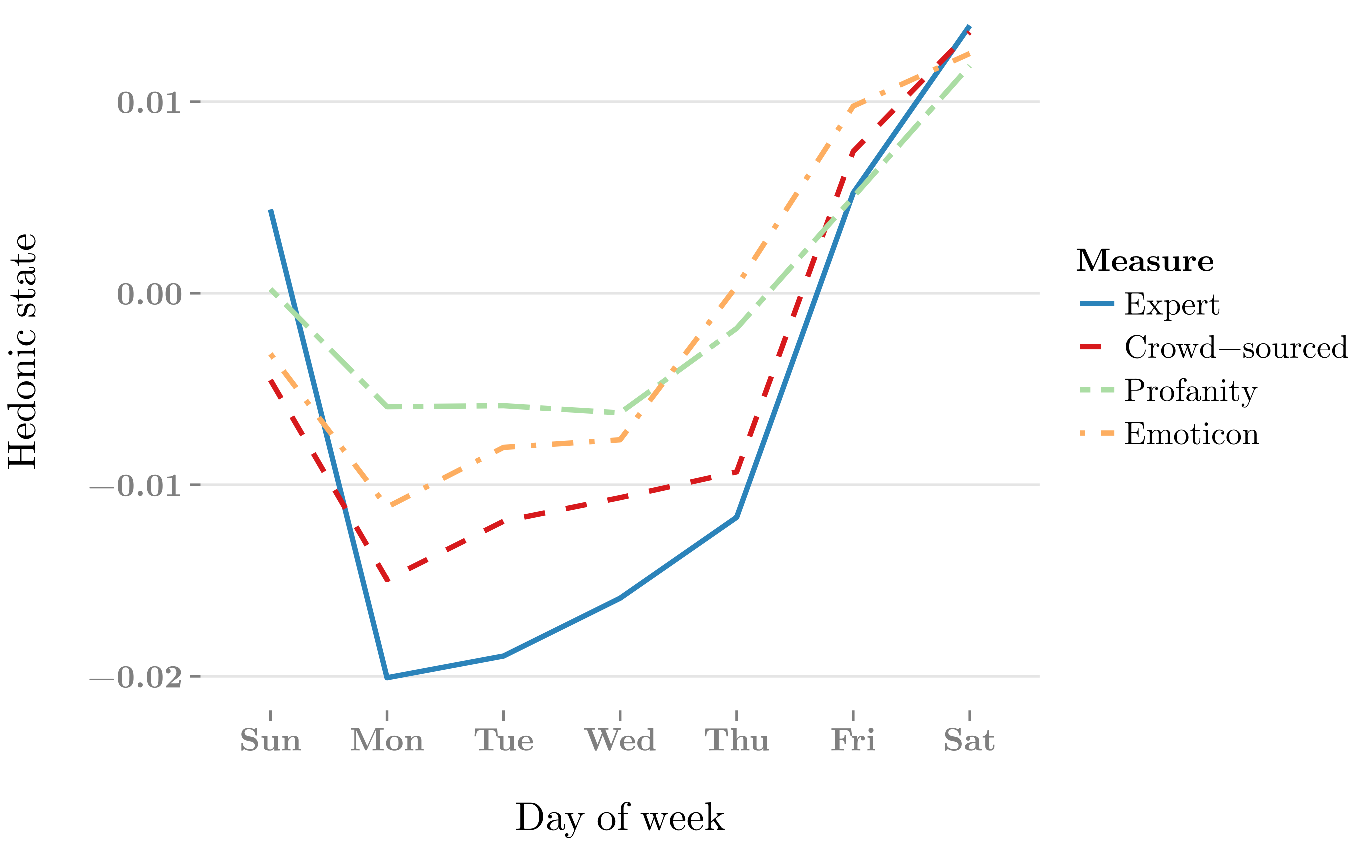
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 1: Tweet density



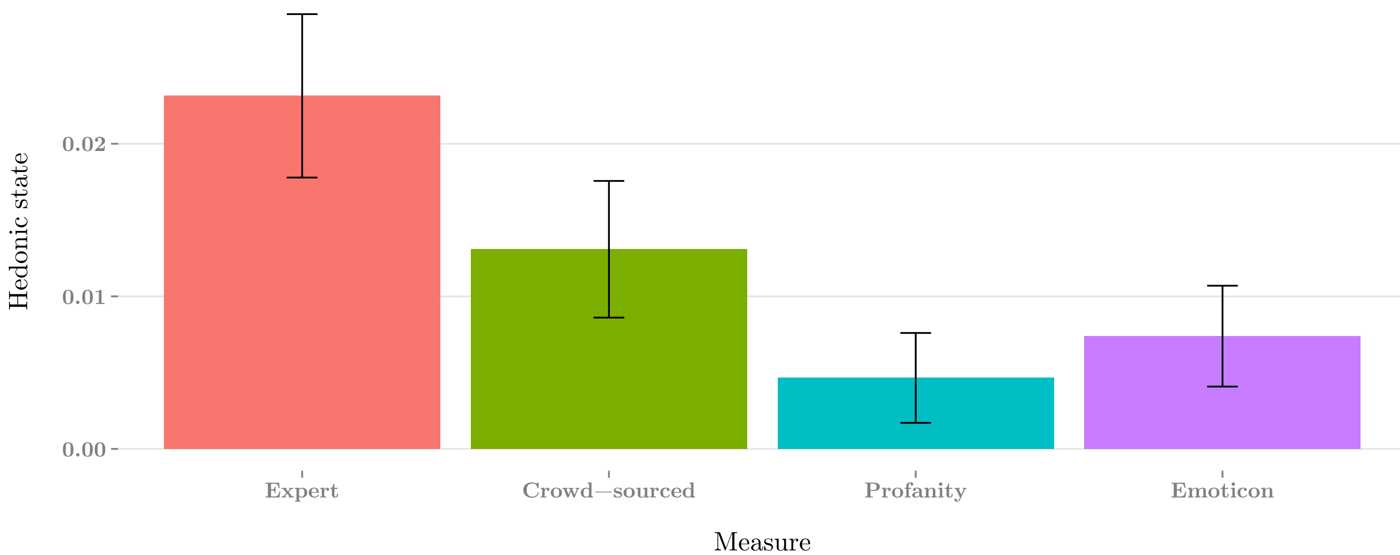
Notes: Darker areas represent higher levels of activity. Each pixel is a 4 km × 4 km grid cell, colored to represent the total recorded number of tweets in that grid cell over the sample period. Color is on a $\log_{10}$ scale.

Figure 2: Hedonic state by day of week



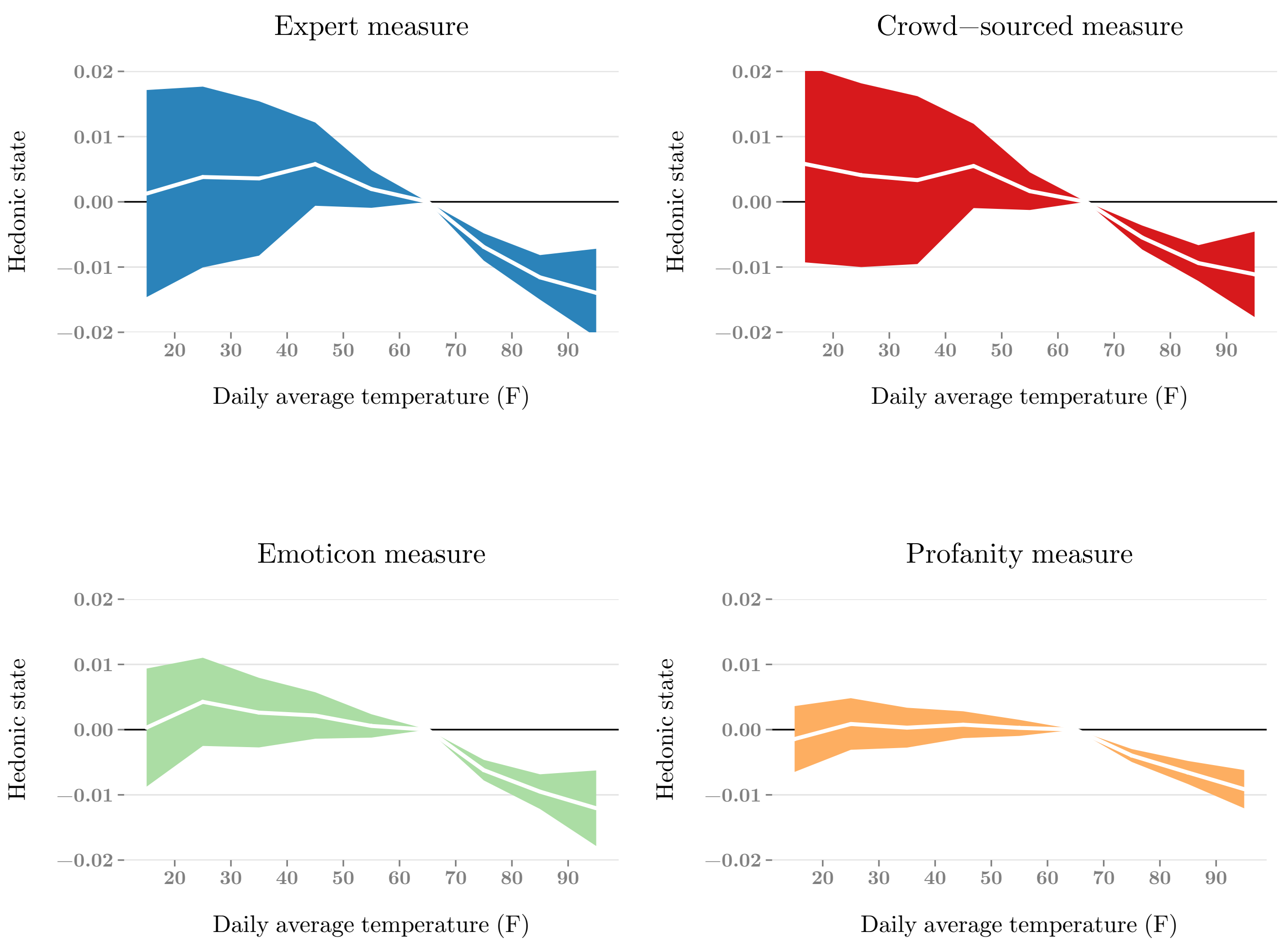
Notes: Each line shows the average hedonic state for each measure described in section 4 by day of week. Measures are standardized to have zero mean and unit standard deviation. Sample excludes major U.S. holidays.

Figure 3: Effect of nearby NFL team win on hedonic state



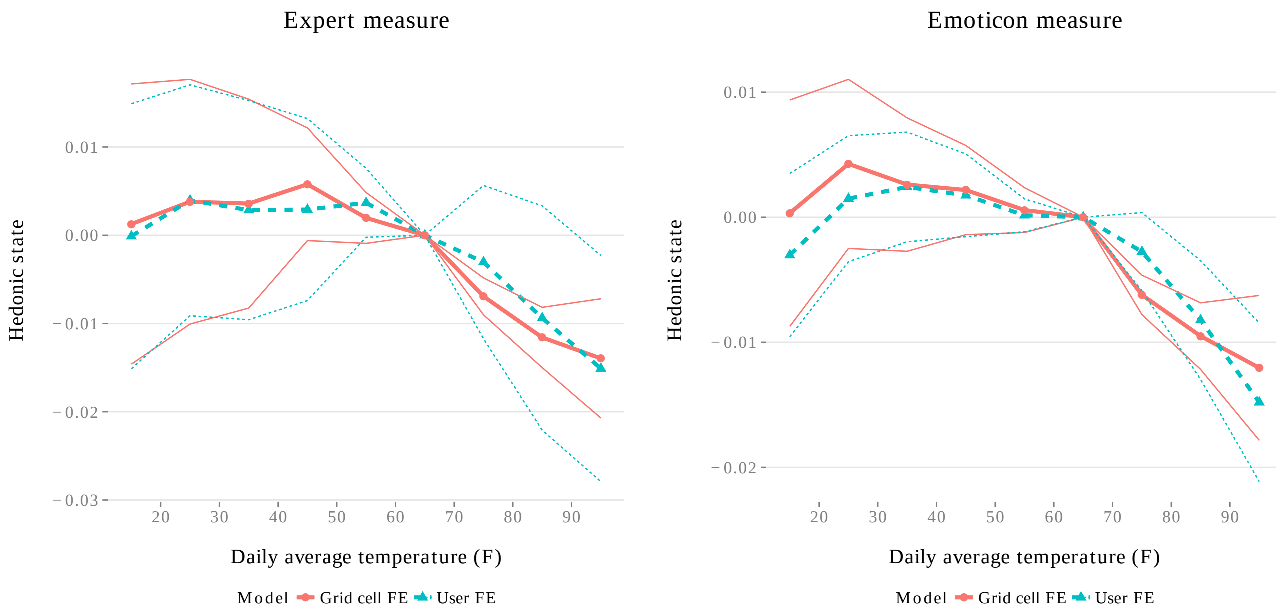
Notes: Height of bars is the change in hedonic state after a win by an National Football League (NFL) team within 80 kilometers. Hedonic response is estimated using the four measures of hedonic state described in section 4. Measures are standardized to have zero mean and unit standard deviation. Sample includes areas within 80 kilometers of an NFL team on Sundays and Mondays during the 2014 season, which ran from September to December. Error bars are the 95% confidence intervals, estimated using two-way cluster robust standard errors on county and day-of-sample.

Figure 4: Effect of temperature on hedonic state



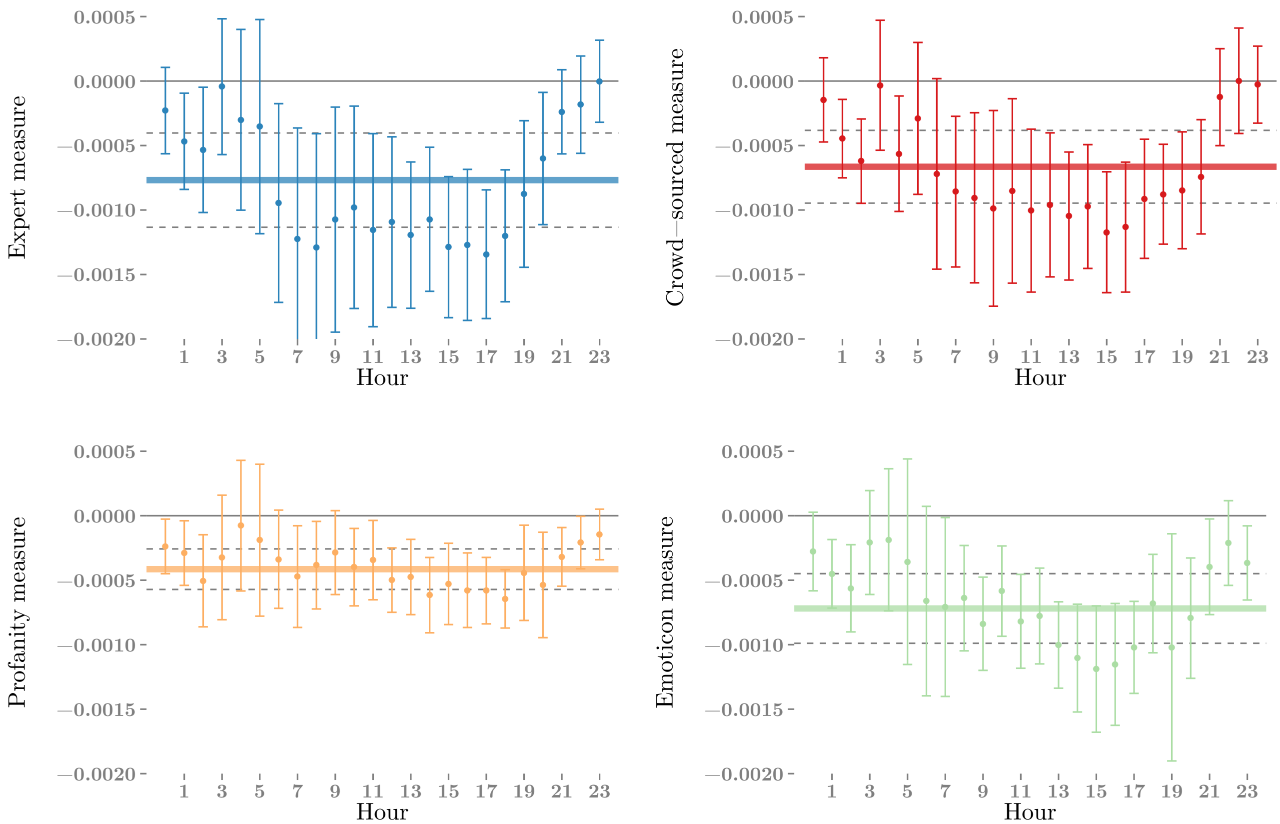
Notes: Plots represent the hedonic response to temperature, where each plot uses a different measure of hedonic state described in section 4. Measures are standardized to have zero mean and unit standard deviation. Each point estimate is the difference in the average grid cell-day hedonic state for the associated ten °F temperature bin relative to the 60-70°F bin (the omitted category), conditional on grid cell and state by month fixed effects and weighted by the number of tweets in a grid cell-day. Grid cells are 4 km × 4 km square cells used by the PRISM weather dataset. 95% confidence intervals estimated using two-way cluster robust standard errors on county and day-of-sample.

Figure 5: User and grid cell fixed effects comparison



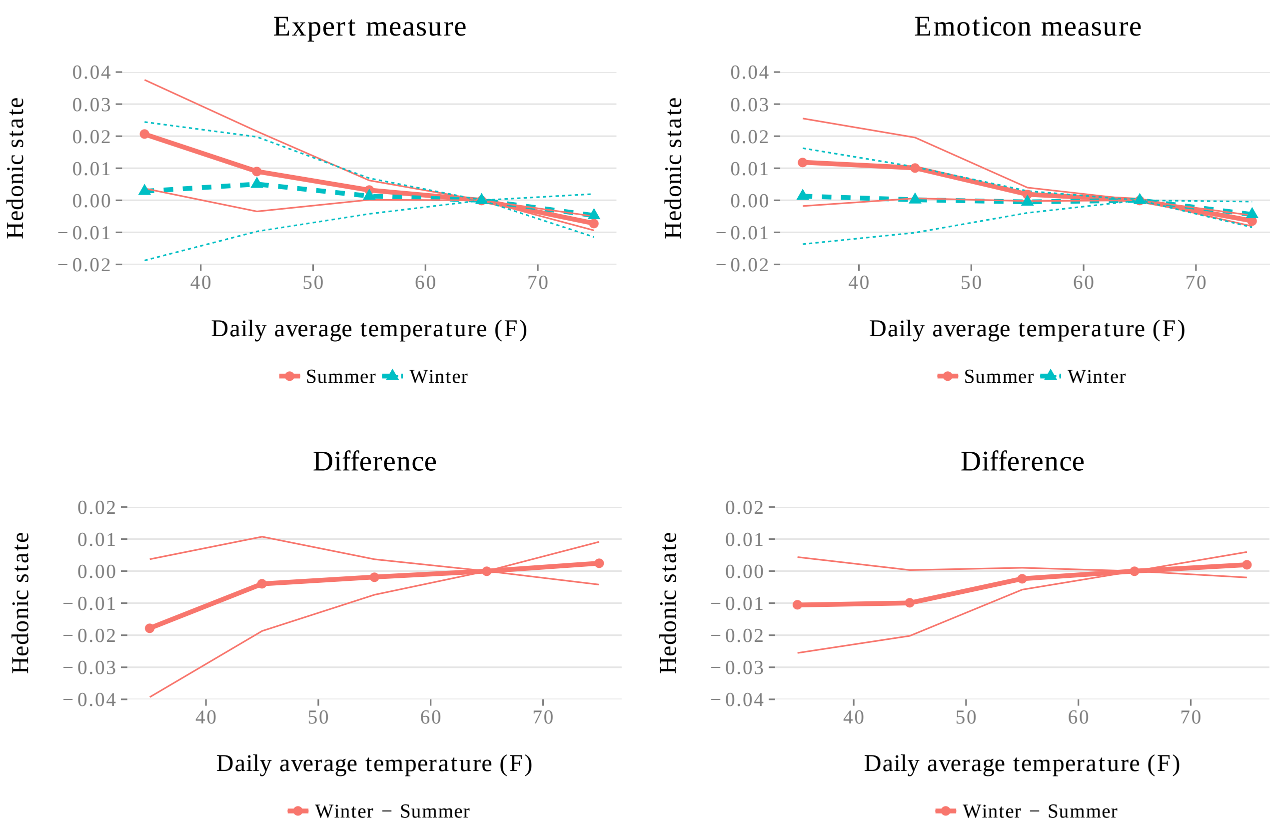
Notes: Plots compares the hedonic response to temperature across two statistical models: the Grid cell FE model include 4 km × 4 km grid cell fixed effects, while the User FE model include fixed effects for each individual in the sample. Hedonic state is measured by the Expert and Emoticon measures described in section 4, which are standardized to have zero mean and unit standard deviation. Both models include state by month fixed effects. Each point estimate is the difference in the average grid cell-day hedonic state for the associated ten °F temperature bin relative to the 60-70°F bin (the omitted category). Grid cells are 4 km × 4 km square cells used by the PRISM weather dataset. 95% confidence intervals estimated using two-way cluster robust standard errors on county and day-of-sample. Results for the other two measures show similar patterns and are available in the appendix.

Figure 6: Response by hour of day



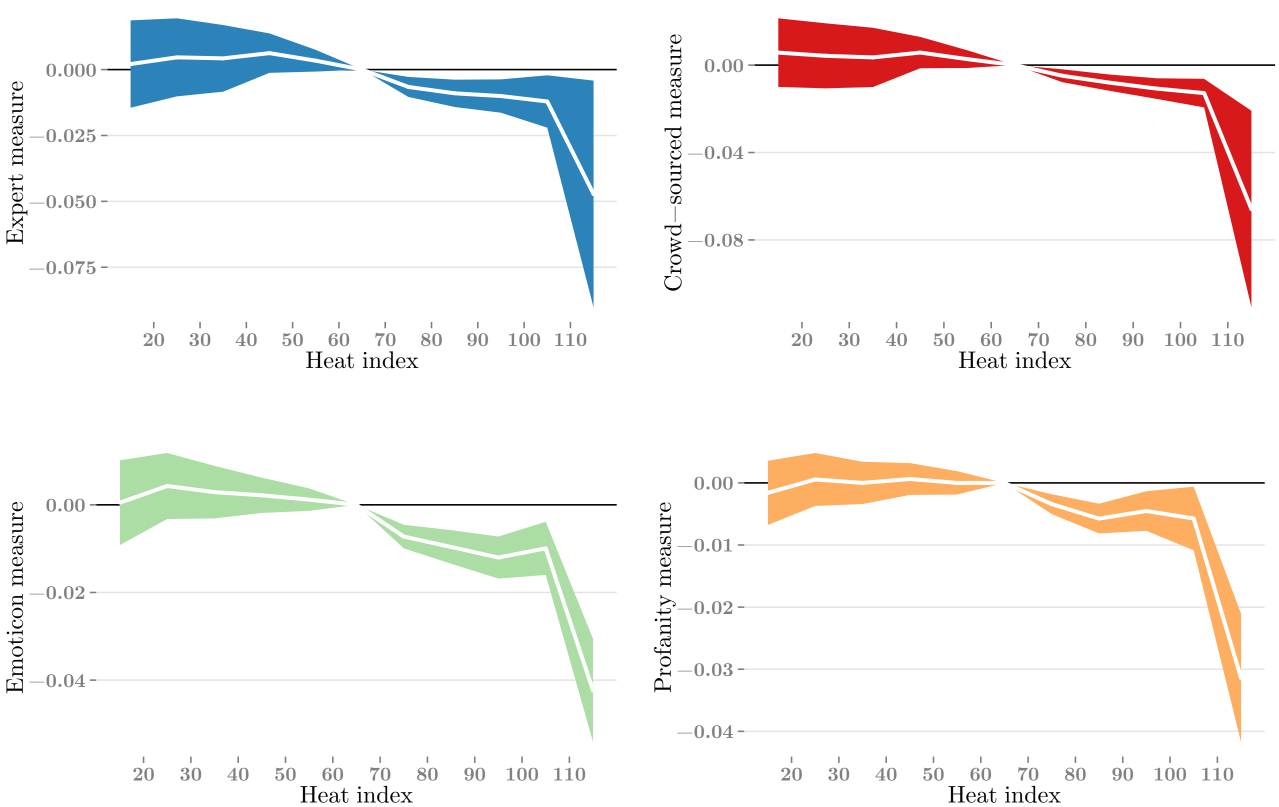
Notes: Plots illustrate hedonic response to high temperatures by hour of day. Measures of hedonic state are as described in section 4 and standardized to have zero mean and unit standard deviation. Sample is limited to days with average daily temperature greater than 60°F. Each point is the coefficient from a separate regression of hedonic state on the daily temperature where the sample is limited to observations in corresponding hour, conditional on grid cell and state by month fixed effects and weighted by the number of tweets in a grid cell-day. Grid cells are 4 km × 4 km square cells used by the PRISM weather dataset. 95% confidence intervals estimated using two-way cluster robust standard errors on county and day-of-sample.

Figure 7: Seasonal response heterogeneity



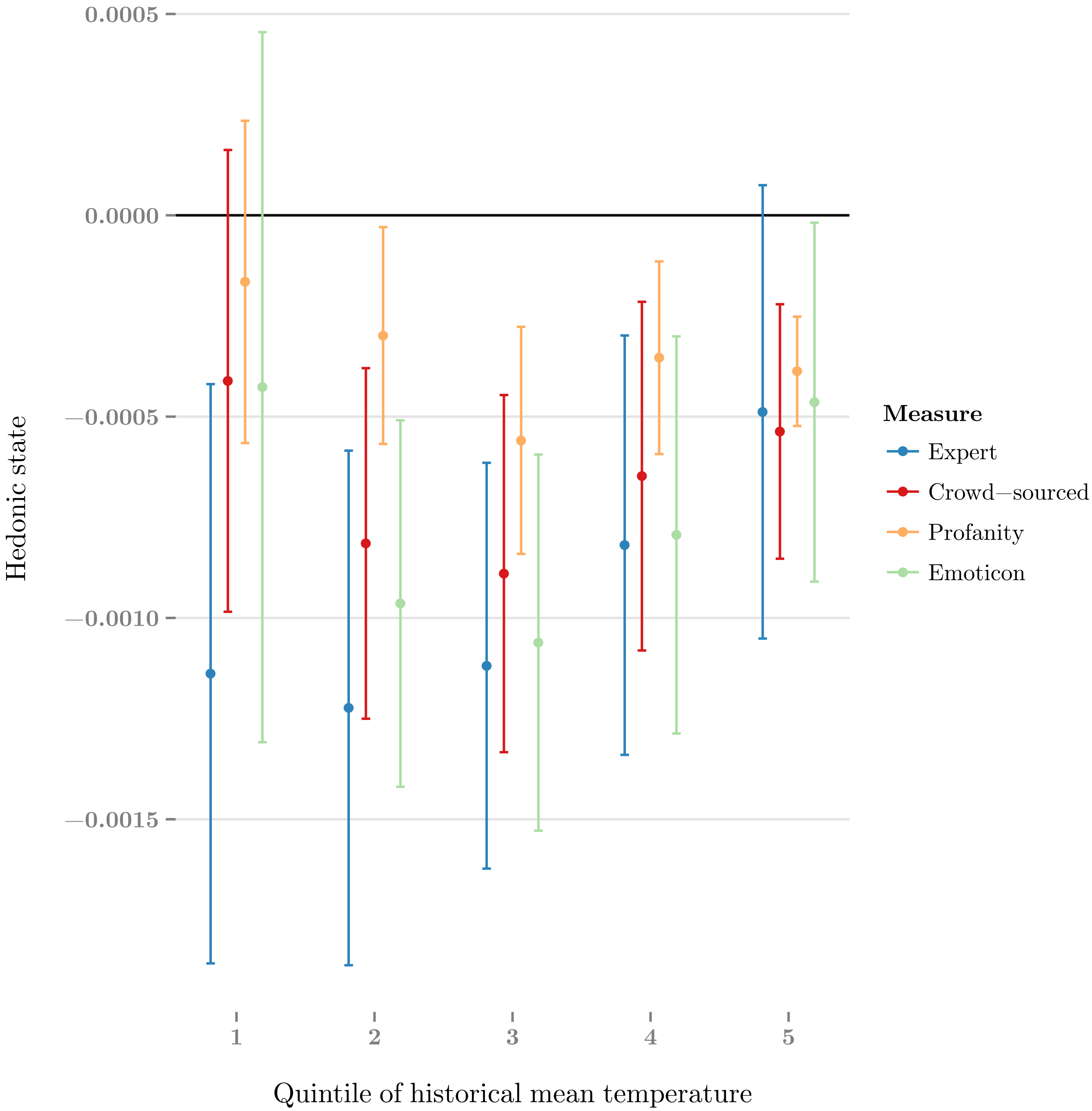
Notes: Plots illustrate heterogeneity in the hedonic response to temperature by season. Hedonic state is measured by the Expert and Emoticon measures described in section 4, which are standardized to have zero mean and unit standard deviation. Top row: each point estimate is the difference in the average grid cell-day hedonic state for the associated ten °F temperature bin relative to the 60-70°F bin (the omitted category), conditional on grid cell and state by month fixed effects and weighted by the number of tweets in a grid cell-day. Bottom row: point estimates are the difference between the corresponding estimates in plot above. Grid cells are 4 km × 4 km square cells used by the PRISM weather dataset. 95% confidence intervals estimated using two-way cluster robust standard errors on county and day-of-sample. Results for the other two measures show similar patterns and are available in the appendix.

Figure 8: Effect of heat index on hedonic state



Notes: Plots represent the hedonic response to heat index, where each plot uses a different measure of hedonic state described in section 4 and heat index is a function of temperature and relative humidity designed to capture how temperature feels to humans. Measures are standardized to have zero mean and unit standard deviation. Each point estimate is the difference in the average grid cell-day hedonic state for the associated ten °heat index bin relative to the 60-70°F bin (the omitted category), conditional on grid cell and state by month fixed effects and weighted by the number of tweets in a grid cell-day. Grid cells are 4 km × 4 km square cells used by the PRISM weather dataset. 95% confidence intervals estimated using two-way cluster robust standard errors on county and day-of-sample.

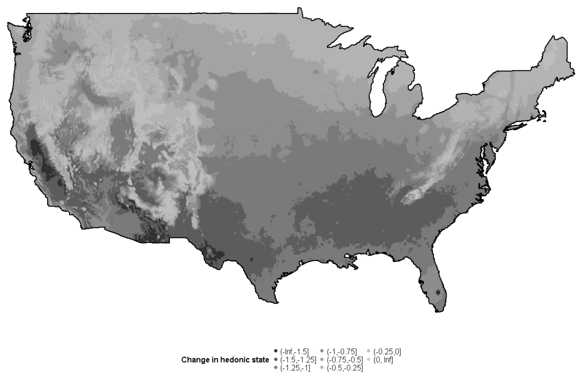
Figure 9: Linearized response by quintiles of historical mean temperature



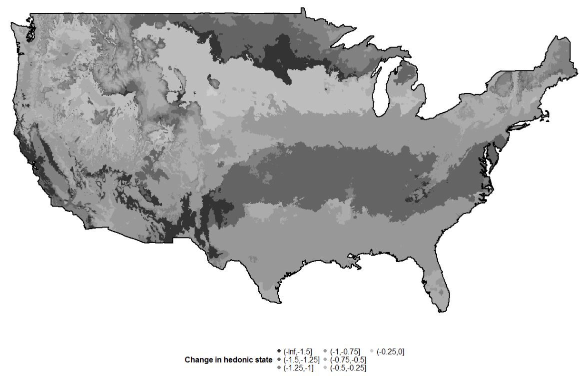
Notes: Points are estimates hedonic response to high temperatures by quintiles of historical mean temperature, where lower quintiles correspond to lower historical temperatures. Measures of hedonic state are as described in section 4 and standardized to have zero mean and unit standard deviation. Sample is limited to days with average daily temperature greater than 60°F. Each point is the coefficient from a separate regression of hedonic state on the daily temperature where the sample is limited to the corresponding quintile, conditional on grid cell and state by month fixed effects and weighted by the number of tweets in a grid cell-day. Grid cells are 4 km × 4 km square cells used by the PRISM weather dataset. 95% confidence intervals estimated using two-way cluster robust standard errors on county and day-of-sample.

Figure 10: Projected changes in hedonic state

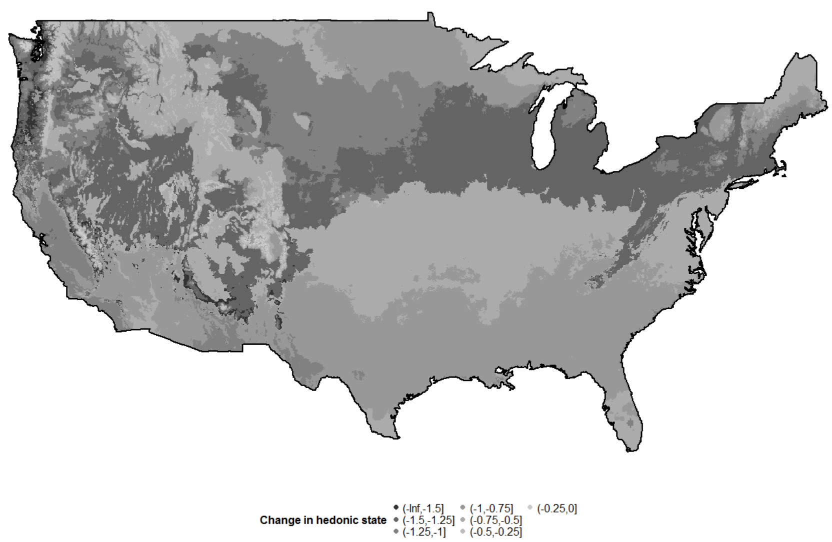
(a) Aggregate response function, no adaptation



(b) Disaggregate response functions, no adaptation



(c) Disaggregate response functions, adaptation



Notes: Darker areas represent larger (in absolute values) annual changes in hedonic state, as measured using the Expert measure described in section 4. Projected changes are computed by taking the difference in the average annual days in a given temperature bin between climate model output of 2086-2099 and 2000-2019, multiplying by the corresponding coefficients in Table 4, and then summing the products. Each pixel is a 4 km $\times$ 4 km grid cell, colored to represent the predicted annual change in standard deviations of hedonic state.

References

Albouy, David, Walter F WF Graf, Ryan Kellogg, and Hendrik Wolff. 2013. “Climate Amenities, Climate Change, and American Quality of Life”. *NBER Work. Pap.*

Angrist, Joshua D., and Jörn-Steffen Pischke. 2008. “Mostly harmless econometrics: An empiricist’s companion”. *An empiricist’s companion*, number March:392.

Antoff, David, and Richard Tol. 2014. *FUND - Climate Framework for Uncertainty, Negotiation and Distribution.*

Auffhammer, Maximilian. 2013. “Quantifying Intensive and extensive margin adaptation responses to climate change: A study of California’s Residential Electricity Consumption”. *Work. Pap.*

Auffhammer, Maximilian, and Anin Aroonruengsawat. 2011. “Simulating the impacts of climate change, prices and population on California’s residential electricity consumption”. *Clim. Change* 109 (SUPPL. 1): 191–210.

Auffhammer, Maximilian, and Erin T. Mansur. 2014. “Measuring climatic impacts on energy consumption: A review of the empirical literature”. *Energy Econ.*0–28.

Barreca, Alan, Karen Clay, Olivier Deschênes, Michael Greenstone, Joseph S. Shapiro, and Olivier Deschenes. 2013. “Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the 20th Century”. *NBER Work. Pap.* Number January:46.

Bollen, Johan, Huina Mao, and Xiaojun Zeng. 2011. “Twitter mood predicts the stock market”. *J. Comput. Sci.* 2 (1): 1–8.

Bradley, Margaret Mm, and Pj Peter J Lang. 1999. “Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings”. *Psychology* Technical (C-1).

Burke, Marshall, Solomon M. Hsiang, and Edward Miguel. 2015a. "Climate and Conflict". *Annu. Rev. Econom.* 7:577–617.

Burke, Marshall, Solomon M Hsiang, and Edward Miguel. 2015b. "Global non-linear effect of temperature on economic production". *Nature* Online Adv.

Burke, Marshall B., and Kyle Emerick. 2015. "Adaptation to climate change: Evidence from US agriculture". *Am. Econ. J. Econ. Policy.*

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

Cline, William R. 1992. "The Economics of Global Warming". *Peterson Inst. Press.*

Conley, Timothy G. 2008. "Spatial Econometrics":741–747.

Correia, Sergio. 2014. "REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects". *Stat. Softw. Components.*

Cragg, Michael, and Matthew Kahn. 1997. "New estimates of climate demand: evidence from location choice". *J. Urban Econ.* 42 (2): 261–284.

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

Davis, Lucas W., and Paul J. Gertler. 2015. "Climate change could drive air conditioning to boost carbon emissions". *Proc. Natl. Acad. Sci.* 112 (19): 5962–5967.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2012. "Temperature Shocks and Economic Growth: Evidence from the Last Half Century". *Am. Econ. J. Macroecon.* 4 (3): 66–95.

— . 2014. "What Do We Learn from the Weather? The New Climate-Economy Literature". *J. Econ. Lit.* 25 (3): 740–798.

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

Deryugina, Tatyana, and Solomon M. Hsiang. 2014. “Does the Environment Still Matter? Daily Temperature and Income in the United States”. *NBER Work. Pap.*

Deschênes, Olivier, and Michael Greenstone. 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”. *Work. Pap.*

Diener, Ed. 2000. “Subjective Well-Being”. *Am. Psychol.* 55 (1): 34–43.

Dodds, Peter Sheridan, and Christopher M. Danforth. 2010. “Measuring the happiness of large-scale written expression: Songs, blogs, and presidents”. *J. Happiness Stud.* 11 (4): 441–456.

Dolan, Paul, Tessa Peasgood, and Mathew White. 2008. “Do we realy 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.

Eichstaedt, Johannes C., Hansen Andrew Schwartz, Margaret L. Kern, Gregory Park, Darwin R. Labarthe, Raina M. Merchant, Sneha Jha, Megha Agrawal, Lukasz A. Dziurzynski, Maarten Sap, Christopher Weeg, Emily E. Larson, Lyle H. Ungar, and Martin E. P. Seligman. 2015. “Psychological Language on Twitter Predicts County-Level Heart Disease Mortality”. *Psychol. Sci.*0956797614557867.

Feddersen, John, Robert Metcalfe, and Mark Wooden. 2012. “Subjective Well-Being: Weather Matters; Climate Doesn’t”. *SSRN Electron. J.* Number 627.

Gerber, Matthew S. 2014. "Predicting crime using Twitter and kernel density estimation". *Decis. Support Syst.* 61 (1): 115–125.

Go, Alec, Richa Bhayani, and Lei Huang. 2009. "Twitter Sentiment Classification using Distant Supervision". *Processing* 150 (12): 1–6.

Graff Zivin, Joshua, and Matthew Neidell. 2014. "Temperature and the Allocation of Time: Implications for Climate Change". *J. Labor Econ.* 32 (1): 1–26.

Graff Zivin, Joshua S, Solomon M Hsiang, and Matthew J Neidell. 2015. "Temperature and Human Capital in the Short-and Long-Run". *NBER Work. Pap.*

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error". *Econometrica* 47 (1): 153–161.

Hoch, Irving, and Judith Drake. 1974. "Wages, climate, and the quality of life". *J. Environ. Econ. Manage.* 1 (4): 268–295.

Hope, Chris. 2006. "The Marginal Impact of CO2 from PAGE2002: An Integrated Assessment Model Incorporating the IPCC's Five Reasons for Concern". *Integr. Assess. J.* 6 (1): 16–56.

Houser, Trevor, Robert Kopp, Solomon M. Hsiang, Michael Delgado, Amir Jina, Kate Larsen, Michael Mastrandrea, Shashank Mohan, Robert Muir-Wood, D. J. Rasmussen, and Others. 2014. "American Climate Prospectus: Economic Risks in the United States". *Rhodium Gr.*

Howarth, Edgar, and Michael S. Hoffman. 1984. "A multidimensional approach to the relationship between mood and weather." *Br. J. Psychol.* 75 (Pt 1) (1): 15–23.

Hsiang, S. M. 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 M., Marshall Burke, and Edward Miguel. 2013. "Quantifying the influence of climate on human conflict." *Science* 341 (6151): 1235367.

Hsiang, Solomon M, and Amir S Jina. 2014. "The Causal Effect of Environmental Catastrophe on Long-Run Economic Growth". *NBER Work. Pap.*

IPCC. 2014. *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Field, C.B., V.R. Barros, D.J. Dokken, K.J.* 1132. Cambridge University Press.

IPCC, International Panel On Climate Change. 2001. "IPCC Third Assessment Report". *Assessment* 5:1–29.

Kahneman, Daniel, and Alan B. Krueger. 2006. "Developments in the Measurement of Subjective Well-Being". *J. Econ. Perspect.* 20 (1): 3–24.

Keller, Matthew C., Barbara L. Fredrickson, Oscar Ybarra, Stéphane Côté, Kareem Johnson, Joe Mikels, Anne Conway, and Tor Wager. 2005. "A warm heart and a clear head: The contingent effects of weather on mood and cognition". *Psychol. Sci.* 16 (9): 724–731.

Kjellstrom, Tord, Ingvar Holmer, and Bruno Lemke. 2009. "during climate change". 1:1–6.

Klimstra, Theo A., Tom Frijns, Loes Keijsers, Jaap JA Denissen, Quinten AW Raaijmakers, Marcel AG van Aken, Hans M. Koot, Pol AC van Lier, and Wim HJ Meeus. 2011. "Come rain or come shine: Individual differences in how weather affects mood." *Emotion* 11 (6): 1495.

Kouloumpis, Efthymios, Theresa Wilson, and Johanna Moore. 2011. "Twitter Sentiment Analysis : The Good the Bad and the OMG !" *Artif. Intell.* 11:538–541.

Levinson, Arik. 2012. "Valuing public goods using happiness data: The case of air quality". *J. Public Econ.* 96 (9-10): 869–880.

Lucas, Richard E, and Nicole M Lawless. 2013. "Does life seem better on a sunny day? Examining the association between daily weather conditions and life satisfaction judgments." *J. Pers. Soc. Psychol.* 104 (5): 872–84.

Mackerron, George. 2012. "Happiness economics from 35000 feet". *J. Econ. Surv.* 26 (4): 705–735.

Mendelsohn, Robert, William D Nordhaus, and Daigee Shaw. 1994. "The Impact of Global Warning on Agriculture: A Ricardian Analysis". *Am. Econ. Rev.* 84 (4): 753–771.

Mitchell, Lewis, Morgan R. Frank, Kameron Decker Harris, Peter Sheridan Dodds, and Christopher M. Danforth. 2013. "The Geography of Happiness: Connecting Twitter Sentiment and Expression, Demographics, and Objective Characteristics of Place". *PLoS One* 8 (5).

Morstatter, Fred, Jürgen Pfeffer, Huan Liu, and K Carley. 2013. "Is the Sample Good Enough? Comparing Data from Twitter's Streaming API with Twitter's Firehose". *Assoc. fo Adv. Artif. Intell.*

Nielsen, Finn Årup. 2011. "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs". *arXiv1103.2903 [cs].*

Nordhaus, William, and Paul Sztorc. 2013. *DICE 2013R : Introduction and User's Manual.*

Nordhaus, William D. 2007. "A Review of the Stern Review on Economics of Climate Change". *J. Econ. Lit.* 45 (3): 686–702.

— . 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, and Patrick Paroubek. 2010. "Twitter as a Corpus for Sentiment Analysis and Opinion Mining." In *Lrec*, 10:1320–1326.

Pang, Bo, Lillian Lee, and Shivakumar Vaithyanathan. 2002. “Thumbs up?: sentiment classification using machine learning techniques”. In *Proc. ACL-02 Conf. Empir. methods Nat. Lang. Process. - EMNLP ’02*, 10:79–86.

Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. 2011. “Scikit-learn: Machine Learning in Python”. *J. Mach. Learn. Res.* 12:2825–2830.

Ranson, Matthew. 2014. “Crime, weather, and climate change”. *J. Environ. Econ. Manage.* 67 (3): 274–302.

Rose, S. 2014. “The Social Cost of Carbon: A Technical Assessment”. *Electr. Power Res. Inst. Tech. Updat. Rep.*

Schlenker, Wolfram, W. Michael Hanemann, and Anthony C. Fisher. 2005. “Will U.S. agriculture really benefit from global warming? Accounting for irrigation in the hedonic approach”. *Am. Econ. Rev.* 95 (1): 395–406.

Schlenker, Wolfram, and Michael J. Roberts. 2009. “Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change.” *Proc. Natl. Acad. Sci. U. S. A.* 106 (37): 15594–15598.

Sinha, Paramita, and M Cropper. 2013. “The value of climate amenities: Evidence from migration decisions”. *NBER Work. Pap.*

Stern, N. 2006. “The Economics of Climate Change”. *Stern Rev.*662.

Train, Kenneth. 2002. *Discrete Choice Methods with Simulation*, 1–388.

United States Government. 2013. “Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis Under Executive Order 12866”, number May:1–21.

Wilson, Theresa, Janyce Wiebe, and Paul Hoffmann. 2005. “Recognizing contextual polarity in phrase-level sentiment analysis”. In *Proc. Conf. Hum. Lang. Technol. Empir. Methods Nat. Lang. Process. - HLT ’05*, 347–354.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. 58:752. 2.