Appendix to Baylis (2017)

Patrick Baylis

February 24, 2017

Contents

1 Word lists

In this section I present samples of the word lists used to construct the sentiment scores in Baylis (2017).

Table 1: AFINN word-score examples

Positive Affect		Neutral Affect		Negative Affect	
superb	5	combat	-1	betraying	-3
thrilled	5	apologizes	-1	agonises	-3
hurrah	5	exposing	-1	destroying	-3
outstanding	5	oxymoron	-1	swindle	-3
breathtaking	5	provoked	-1	abhors	-3
roflcopter	4	limited	-1	humiliation	-3
wowow	4	escape	-1	chastises	-3
rejoicing	4	unconfirmed	-1	victimizing	-3
lifesaver	4	passively	-1	bribe	-3
winner	4	blocks	-1	lunatic	-3
miracle	4	poverty	-1	scandal	-3
triumph	4	attacked	-1	outrage	-3
fabulous	4	gun	1	betrayed	-3
roflmao	4	feeling	1	terror	-3
euphoric	4	intrigues	1	abuse	-3
heavenly	4	alive	1	greenwash	-3
fantastic	4	protected	1	falsified	-3
ecstatic	4	unified	1	douche	-3
funnier	4	relieves	1	agonized	-3
winning	4	fit	1	criminals	-3
masterpiece	4	restore	1	defects	-3
masterpieces	4	relieve	1	idiotic	-3
stunning	4	greeting	1	woeful	-3
godsend	4	yeah	1	acrimonious	-3
lmfao	4	cool	1	nuts	-3
lmao	4	vested	1	swindles	-3
rotflmfao	4	clearly	1	lost	-3

Notes: Raw scores shown. Standardized scores used in analysis. Full list includes 2,477 total word-score mappings and can be obtained here: `http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=6010`.

Table 2: Hedonometer word-score examples

Positive Affect		Neutral Affect		Negative Affect	
laughter	8.5	fui	5.08	suicide	1.3
happiness	8.44	gilbert	5.08	terrorist	1.3
love	8.42	hart	5.08	rape	1.44
happy	8.3	hij	5.08	murder	1.48
laughed	8.26	hun	5.08	terrorism	1.48
laugh	8.22	indonesia	5.08	cancer	1.54
laughing	8.2	jo	5.08	death	1.54
excellent	8.18	john	5.08	died	1.56
laughs	8.18	juan	5.08	kill	1.56
joy	8.16	knee	5.08	killed	1.56
successful	8.16	laws	5.08	torture	1.58
win	8.12	listed	5.08	arrested	1.64
rainbow	8.1	manhasset	5.08	deaths	1.64
smile	8.1	marion	5.08	raped	1.64
won	8.1	martinez	5.08	killing	1.7
pleasure	8.08	medicaid	5.08	die	1.74
smiled	8.08	medicine	5.08	jail	1.76
rainbows	8.06	meyer	5.08	terror	1.76
winning	8.04	might	5.08	kills	1.78
celebration	8.02	morgen	5.08	fatal	1.8
enjoyed	8.02	morris	5.08	killings	1.8
healthy	8.02	nas	5.08	murdered	1.8
music	8.02	necessarily	5.08	war	1.8

Notes: Raw scores shown. Standardized scores used in analysis. Full list includes 10,223 total word-score mappings and can be obtained here: `http://hedonometer.org/words.html`

Table 3: LIWC word examples

Positive emotion	Negative emotion
love	hurt
nice	ugly
sweet	nasty

Notes: LIWC is a commercial product, selected examples are described in Tausczik and Pennebaker (2010). Full list includes 905 total words.

Table 4: Emoticons list

Positive emotion	Negative emotion
:-)	>:[
:)	:-(
:D	:(
:o)	:-c
:]	:c
:3	:-<
:c)	:?C
:>	:<
=]	:-[
8)	:[
=)	:{
:}	:'-(
:^)	:'(
:?)	
:-D	
8-D	
8D	
x-D	
xD	
X-D	
XD	
=-D	
=D	
=-3	
=3	
B^D	
:-))	
:))	

Notes: List of emoticons obtained from Wikipedia: `https://en.wikipedia.org/wiki/List_of_emoticons`.

Table 5: ML features

Word	Score	Word	Score
our latest	6.90	miss my	-3.50
the follow	3.76	o NUMBER	-3.25
for following	3.53	hurts	-2.76
latest	3.31	ugh	-2.70
URL HASHTAG	3.30	i miss	-2.67
check out	2.81	dont wanna	-2.50
HASHTAG in	2.68	poor	-2.45
this HASHTAG	2.63	wanna go	-2.42
welcome	2.56	im sorry	-2.41
thanks for	2.46	wish i	-2.40
opening	2.41	just wanna	-2.31
MENTION thanks	2.38	just want	-2.31
MENTION thank	2.34	i wish	-2.27
click	2.32	upset	-2.25
in HASHTAG	2.27	rip	-2.19
erful	2.20	sucks	-2.16
great day	2.16	miss	-2.04
thank you	2.16	really want	-1.99
MENTION happy	2.15	want a	-1.96
thanks	2.10	i wanna	-1.96

Notes: Raw scores shown, standardized scores used in analysis. Score is the difference between the log predicted probability that the sentence is positive versus negative, conditional on having observed the presence of the given n-gram.

Table 6: Sample of tweets with sentiment scores

Tweet (first 50 characters)	(1)	(2)	(3)	(4)
Yes! Chicago today	1	6.27	0	0.50
What a great way to end 2014 @ SNOW IN TUCSON WHAT	3	5.46	1	0.99
@GottliebShow does he get that money plus whatever	0	5.52	1	-0.70
This was hella worth it	2	5.23	0	-0.25
Happy New Years! Had the best time @WattsBarChurch	3	5.83	3	0.82
@jarlenykillz *Turns straight*	1	5.63	0	0.00
@dandandempsey haven't cut mine in 4	-1	4.76	-1	-0.26
Not a very nice chef in Jcpenny http://t.co/2Gsm7	3	5.72	0	0.44
Go follow my boy. @oLewisss he's a sick GFX artist	-2	5.34	0	-0.96
@love_the_Ks so awesome! If you need any beer rec	3	5.72	1	0.94
happy to be this http://t.co/XMebBnIVGF	3	6	1	0.42
@g_jackson24 I better see you nigggaaaaa	2	6.3	1	0.23
Feeling and love are both over rated they are best	2	5.69	2	-0.01
@wasabisauce lets waste time chasing squirrels	-1	4.68	0	0.47
someone please get me out of my house rn	1	5.78	0	-0.94
As long as my momma told me Happy New Year Idc bou	1	5.48	1	-0.61
Tonight is going to be good	3	5.77	1	-0.71
makes me happy seeing close friends having a good	3	6.36	2	-0.55
All my friends are drunk rn	-2	5.87	0	-0.89
MSU for the win	4	6.11	1	0.77
I want chipotle	1	5.81	0	-0.95
@smarcher_ I have light canceling curtains up but	-3	5.34	-1	-0.90
@nnakedd Lmao dude I got you	4	5.87	1	-0.44
Happy new year	3	6.83	1	0.76
First morning of 2015, I check my phoneI go to sn	0	5.83	2	0.56
I wish you cared.	1	6.48	1	-0.98
i hate pitbull	-3	4.13	-1	-0.96
But not really feeling it	1	4.93	0	-0.88

Table 7: Weather-related stopwords

blizzard	frostbite	precipitation
breeze	frosty	rain
chilly	gail	rainbow
clear	gust	showers
clouds	hail	sleet
cloudy	heat	snowflakes
cold	hot	soggy
damp	humid	sprinkle
dew	hurricane	sunny
downpour	icy	thunder
drizzle	lightning	thunderstorm
drought	misty	typhoon
dry	moist	weather
flurry	monsoon	wet
fog	muddy	wind
freezing	overcast	windstorm
frigid	pouring	windy

Notes: Source: author construction.

2 Empirical checks

In this section I document a series of checks intended to test the robustness of the result to different sample selection criteria and model specifications.

2.1 Weather-related tweets

As described in the paper, I remove tweets related to weather from my sample before estimating the main results. I do this in order to ensure that the effects reflect shifts in general sentiment rather than weather-related changes. While weather-related tweets constitute around 1.3% of my sample, it is possible that sentiment surrounding descriptions of current weather could bias the results. Table 8 documents the sentiment response of non-weather tweets, weather tweets, and the pooled sample. The model that includes weather tweets documents a much more pronounced response to temperature, indicating that the sentiment of weather-related tweets responds much more sharply to changes in temperature.

The first column in Table 8 is the baseline specification. As noted above, the sample for this estimate excludes tweets with weather-related words (see appendix for the list of weather-related words I use). The second column estimates the same model, but the sample is limited instead to only the weather-related tweets. The estimates from this column are similar in sign but an order of magnitude larger than those from the weather-excluded sample. The third column estimates the same model using all tweets, whether or not they contain weather related terms. As expected from the first two columns, the estimates from this sample are larger than those in the baseline model, reflecting the addition of the weather-related tweets to the sample. That weather has a more notable effect on the sentiment of weather-related posts is little surprise, and justifies the exclusion of weather-related tweets from the baseline sample. By focusing on non-weather related tweets I hope to isolate the subconscious effect of weather on sentiment.

Table 8: Weather-related tweets

	Baseline	Weather-related	Pooled
Max temperature T			
$T \leq 0$	-0.010^{*}	-0.203^{***}	-0.014^{**}
	(0.006)	(0.012)	(0.006)
$T \in (0{,}5]$	-0.006	-0.150^{***}	-0.009^{*}
	(0.005)	(0.007)	(0.005)
$T \in (5, 10]$	-0.003	-0.116^{***}	-0.006
	(0.003)	(0.006)	(0.003)
$T \in (10, 15]$	0.0001	-0.075^{***}	-0.001
	(0.003)	(0.004)	(0.003)
$T \in (15, 20]$	-0.0005	-0.027^{***}	-0.001
	(0.001)	(0.003)	(0.001)
$T \in (20, 25]$	0	0	0
$T \in (25, 30]$	-0.002^{*}	-0.021^{***}	-0.002^{**}
	(0.001)	(0.003)	(0.001)
$T \in (30, 35]$	-0.006^{***}	-0.086^{***}	-0.008^{***}
	(0.001)	(0.004)	(0.001)
$T \in (35, 40]$	-0.010^{***}	-0.150^{***}	-0.013^{***}
	(0.002)	(0.008)	(0.002)
$T > 40$	-0.011^{***}	-0.210^{***}	-0.014^{***}
	(0.003)	(0.010)	(0.003)
Observations (m)	2,087,861	1,149,283	2,093,032
Twitter updates (m)	1,443,406,925	19,846,044	1,463,252,969

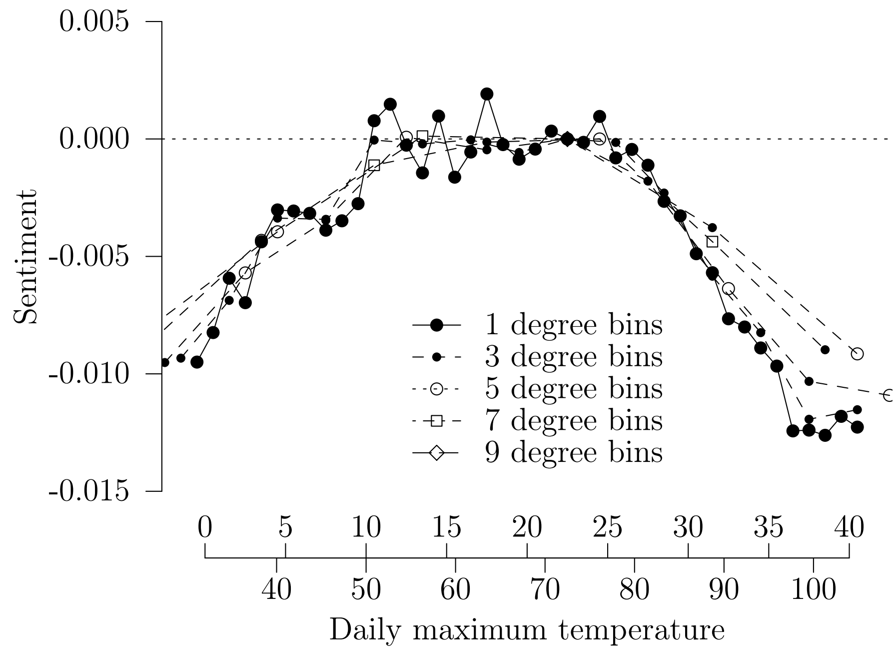
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Dependent variable is the Sentiment index in a county-day. Coefficients represent the change in standard deviations of sentiment between a day with maximum temperature in the associated temperature bin and a day with temperature $T \in [20{,}25)$, the omitted category. All models include precipitation P and the listed fixed effects, standard errors clustered by county-month of sample and date.

2.2 Bin widths

Because statistical models employing bin specifications can sometimes be affected by the selection of bin width, I estimate models with 1, 3, 5 (baseline), 7, and 9 degrees C in Figure I

Figure I: Bin width



Notes: Comparison of the sentiment response to temperature across bin width.

2.3 Date fixed effects

In Baylis et al. (2017) our primary model includes date fixed effects. In this paper, I estimate the models using a more limited set: county and state by month-of-sample fixed effects. I do this because of this paper's objective of projecting the effects of future climate change, which in some areas may cause more than 5 degrees C in warming by end-of-century. Because finer sets of fixed effects (like date fixed effects) remove more variation from the data, in this paper I choose to use the minimum set of fixed effects that defensibly satisfies the conditional unconfoundedness assumption. In other words, I argue that after accounting for for county-level and state by month-of-sample unobservables, weather realizations are as good as random. This assumption is consistent with a broad body of work in the climate impacts literature (Dell, Jones, and Olken 2014; Carleton and Hsiang 2016).

Nevertheless, it is useful to observe the effect of including fixed effects on my results. Table 9 documents this exercise.

Columns (1) and (3) reproduce the estimates from columns (3) and (4) in Table 4 in the main paper, respectively. Columns (2) and (4) add date fixed effects to columns (1) and (3). In both cases, the qualitative result that expressed sentiment declines in extreme temperatures holds. However, the magnitudes shift, and instead of finding roughly equal magnitudes between the coldest temperature bins and the warmest temperature bins, the models with date fixed effects have larger magnitudes and smaller standard errors, while the magnitudes of warmer bin coefficients shrink.

Column (3) remains the baseline model because it more than satisfies reasonable requirements for conditional unconfoundedness and retains the largest possible amount of variance.

2.4 Lags and leads

I also estimate a model that includes leads and lags. For ease of presentation, I use a quadratic model instead of bins. Table 10 documents the lagged results.

I find that, with the exception of a single squared lagged coefficient, the model is unchanged by the addition of one, two, or three days of leads and lags. Because standard errors increase slightly, the level of temperature is no longer statistically significant with the addition of lags, but the squared terms remains signifant in all models.

Table 9: Date fixed effects

	(1)	(2)	(3)	(4)
Max temperature T				
$T \leq 0$	-0.022^{***}	-0.027^{***}	-0.010^{*}	-0.012^{***}
	(0.004)	(0.002)	(0.006)	(0.001)
$T \in (0,5]$	-0.017^{***}	-0.023^{***}	-0.006	-0.010^{***}
	(0.003)	(0.002)	(0.005)	(0.001)
$T \in (5,10]$	-0.012^{***}	-0.017^{***}	-0.003	-0.008^{***}
	(0.003)	(0.002)	(0.003)	(0.001)
$T \in (10,15]$	-0.005^{**}	-0.010^{***}	0.0001	-0.004^{***}
	(0.002)	(0.001)	(0.003)	(0.001)
$T \in (15,20]$	-0.002^{*}	-0.004^{***}	-0.0005	-0.002^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
$T \in (20,25]$	0	0	0	0
$T \in (25,30]$	0.0003	0.002^{***}	-0.002^{*}	0.0001
	(0.001)	(0.001)	(0.001)	(0.0004)
$T \in (30,35]$	-0.002	0.003^{***}	-0.006^{***}	-0.002^{***}
	(0.002)	(0.001)	(0.001)	(0.001)
$T \in (35,40]$	-0.008^{***}	-0.002	-0.010^{***}	-0.005^{***}
	(0.002)	(0.002)	(0.002)	(0.001)
$T > 40$	-0.014^{***}	-0.010^{***}	-0.011^{***}	-0.006^{***}
	(0.004)	(0.003)	(0.003)	(0.002)
Observations (m)	2,087,866	2,087,866	2,087,866	2,087,866
County FE	Yes	Yes	Yes	Yes
M-o-s FE	Yes	Yes		
State × M-o-s FE			Yes	Yes
Date FE		Yes		Yes

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

Table 10: Lags and leads

	Baseline	One day	Two days	Three days
T	0.001^{***}	0.001	0.001	0.001
	(0.0003)	(0.0004)	(0.0004)	(0.0004)
T^2	-0.00003^{***}	-0.00002^{*}	-0.00002^{**}	-0.00002^{**}
	(0.00001)	(0.00001)	(0.00001)	(0.00001)
T_{t-1}		0.0001	−0.0002	−0.0001
		(0.0003)	(0.0003)	(0.0003)
T^2_{t-1}		-0.00001^{**}	−0.00001	−0.00001
		(0.00001)	(0.00001)	(0.00001)
T_{t-2}			0.0004	0.0001
			(0.0003)	(0.0003)
T^2_{t-2}			−0.00001	−0.00000
			(0.00001)	(0.00001)
T_{t-3}				0.0003
				(0.0004)
T^2_{t-3}				−0.00000
				(0.00001)
T_{t+1}		0.0005	0.0001	0.0001
		(0.0005)	(0.0005)	(0.0004)
T^2_{t+1}		−0.00000	−0.00000	−0.00000
		(0.00001)	(0.00001)	(0.00001)
T_{t+2}			0.0004	0.0004
			(0.0004)	(0.0003)
T^2_{t+2}			−0.00000	−0.00000
			(0.00001)	(0.00001)
T_{t+3}				0.00001
				(0.0004)
T^2_{t+3}				−0.00000
				(0.00001)

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

2.5 Individual tweet-level analysis

The ecological fallacy is the observation that the properties of aggregated groups may not reflect properties of the individuals in the underlying populations (Robinson 1950). In this setting, it may be that county-level responses reflect a form of selection bias: since participation in Twitter is a 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 is compositional sorting: samples of tweets at different temperatures may reflects different sets of users with different unobservable characteristics. For example, if individuals with higher or lower native affect become more likely to compose Twitter updates in different temperatures, the coefficients could be capturing this compositional change in the sample rather than a change in average emotional state.

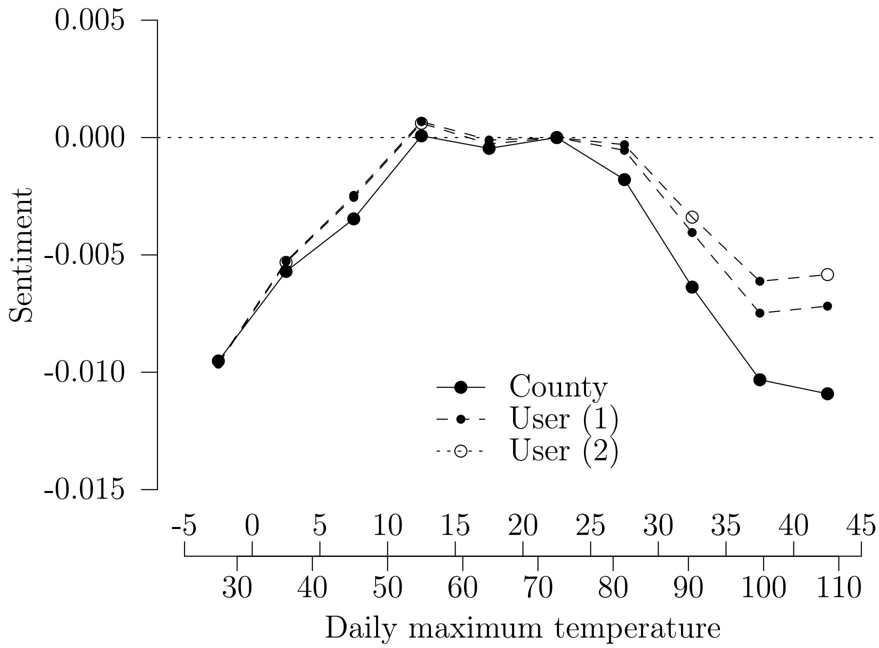
Since the data I collect include an identifier for the tweet creator, I account for compositional sorting in my sample using user fixed effects. To do so, I estimate the following model:

$$E_{id} = \sum_{b \neq 20\text{-}25}^{B} \beta_b T_{cd}^{b} + \phi_i + \phi_c + \phi_{sym} + \varepsilon_{id} \tag{1}$$

This model adds user fixed effects, ϕ_i to the baseline model. The model requires the use of the disaggregated sample of tweets in my dataset; for computational reasons, I use a 20% subsample of users to estimate the following results.

To compare the results between the user fixed model and the baseline model, I overlay the estimates from each model in Figure II. I find qualitatively similar results for the measures, although the estimates for higher temperatures are attenuated in the individual fixed effects model relative to the baseline model. It is possible that this is evidence of some compositional sorting at higher temperatures, but more likely the result of measurement error driven by using a sparser source of variation. The negative response to cold temperature is nearly identical between models, suggesting that the source of the differential is heterogenous in temperature.

Figure II: County-level and user-level comparison



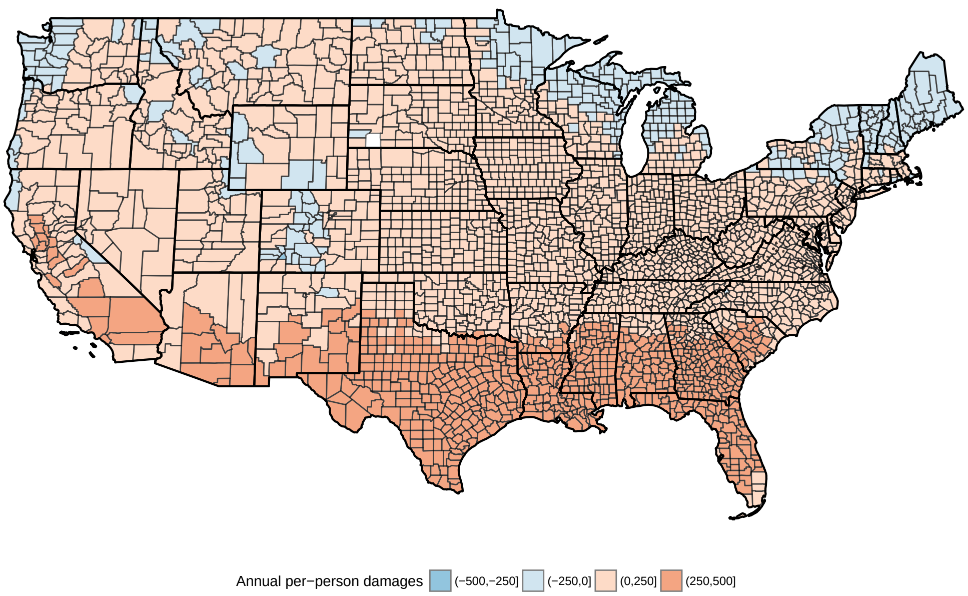
Notes: Plots compares the hedonic response to temperature across two statistical models, one with county and one with user fixed effects. Both models include date fixed effects. Each point estimate is the difference in the average grid cell-day emotional state for the associated five C temperature bin relative to the 20-25 C (68-77 F) bin (the omitted category). 95% confidence intervals estimated using two-way cluster robust standard errors on county and day-of-sample.

3 Projections

3.1 End of century projection (RCP4.5)

The following figure reproduces Figure 8 from the main paper using RCP4.5, a more limited emissions scenario.

Figure III: End of century projection (RCP4.5)



References

Baylis, Patrick. 2017. "Temperature and temperament". *Manuscript submitted for publication.*

Baylis, Patrick, Nick Obradovich, Yury Kryvasheyeu, Haohui Chen, Lorenzo Coviello, Esteban Moro Egido, Manuel Cebrian, and James H. Fowler. 2017. "Weather impacts expressed sentiment". *Manuscript submitted for publication.*

Carleton, T. A., and S. M. Hsiang. 2016. "Social and economic impacts of climate". *Science* 353 (6304).

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2014. "What Do We Learn from the Weather? The New Climate–Economy Literature". *Journal of Economic Literature* 52 (3): 740–798.

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

Robinson, W S. 1950. "Ecological Correlations and the Behavior of Individuals". *American Sociological Review* 15 (3).

Tausczik, Y. R., and J. W. Pennebaker. 2010. "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods". *Journal of Language and Social Psychology* 29 (1): 24–54.