The relationship between unexpected real-world events, Twitter-inferred surprisingness and risk-taking behaviours

Student id: 2185276
Supervisors: Lukasz Walasek, Mikhail Spektor
Target journal: Psychological Science
Word count: 5,542

The relationship between unexpected real-world events, Twitter-inferred surprisingness and risk-taking behaviours

Abstract
Social media posts present a unique opportunity to study real-time emotional responses to current events on a massive scale. At the same time, changes in emotional states have been found to influence cognitive processing ability and hence behaviours, such as the propensity to take risks. Two largely separate bodies of literature have studied the effect of unexpected real-world events, through the lens of their impact on affective states. The first finds that unexpected events create seemingly extraneous effects- including increased gambling behaviour, whilst the second studies the influence of unexpected events on Twitter-inferred mood, finding that they result in larger changes in sentiment than those that are in-line with expectations. In this paper, we explore the interconnect between these two strands of literature. First, we examine the relationship between unexpected sports outcomes of five New York based teams and the degree of surprisingness displayed in a large dataset of tweets. In support of the literature, we find that unexpected sports outcomes do indeed predict daily sentiment prediction errors. Second, we ask whether Twitter-inferred surprisingness is related to real-world gambling, measured as lottery ticket sales, finding no evidence in our dataset of such a relationship.

Our expectations of life-events shape our emotional reaction to their outcomes. This is the reason for the salient nature of underdog stories, such as *The Tortoise and the Hare* – the less we expect the Tortoise to beat the Hare, the stronger our affective response is when it does. Changes in affective state influence behaviours such as risk-taking (Arkes et al., 1988; Ashby et al., 1999; Bassi et al., 2013), possibly due to an optimism bias: being in a good mood may lead us to overestimate the likelihood of positive risky outcomes (Sharot, 2011).

Studying the effects of mood changes in a laboratory setting is often limited by experimental design, whereas large-scale emotional responses to current events can be inferred in a naturalistic setting from analysis of social media posts. This has enabled researchers to quantify population-level affective responses to real-world events and, consequentially, to study behavioural changes from large-scale changes in mood (e.g. Bhatia et al., 2019).

The objective of this paper is to further explore the intersection between unexpected events, risk-taking behaviour and Twitter-inferred affective states. We will study the relationship between the surprisingness inferred from sports team-related tweets, unexpected sporting outcomes and lottery ticket sales.

Drawing comparisons between expectations and realised outcomes is commonplace in day-to-day life. Businesses compare sales figures to forecasts, just as children might compare their school progress against standards set by siblings. Such comparisons between *a priori* beliefs and actual outcomes are also abundant in reference-dependent models in psychology (e.g. Kahneman & Tversky, 1979; Kőszegi & Rabin, 2006; Tversky & Kahneman, 1991), and neurological reinforcement learning models (Ottenheimer et al., 2020; Schultz et al., 1997). Intuitively, when an outcome exceeds or falls short of expectations it is likely to elicit a stronger emotional response: a better-than-expected outcome will lift mood to a greater extent than one that aligns with expectations and vice versa for those that are worse-than-expected (Bhatia et al., 2019; Rutledge et al., 2014; Villano et al., 2020). Imagine the difference in your emotional response to receiving a £100 gift out of the blue compared to getting the same amount when you were given a week's notice.

Our mood ebbs and flows with daily life (Clark & Watson, 1988; Kuppens et al., 2010) and as it does so it alters our cognitive processing abilities – such as working memory (Eich & Macaulay, 2006; Martin & Kerns, 2011), emotional recognition (Chepenik et al., 2007) and risk-taking (Arkes et al., 1988; Ashby et al., 1999). Consequentially (and unsurprisingly) mood swings permeate into day-to-day choices (Kim et al., 2021; Lerner et al., 2004; Manucia et al., 1984). For example, whilst in high spirits, we are more likely to recall positive memories, be generous and help others (Forgas, 2002). It is through the channel of mood that seemingly extraneous events – including a sunny day and sports team results – influence a variety of decision-making environments, from financial markets (Bassi et al., 2013; Cao & Wei, 2005; Hirshleifer & Shumway, 2003; Saunders, 1993) to court house rulings (Chen & Loecher, 2019; Danziger et al., 2011; Heyes & Saberian, 2019).

Several studies have indeed investigated the interplay between surprising events, changes in mood and their seemingly exogenous knock-on effects. In a study of Louisiana juvenile court rulings, an upset loss for the local college football team (Louisiana State University) was linked to harsher judicial decisions – with average sentences given in the week that followed increasing by an estimated additional 1,332 days (Eren & Mocan, 2018). Unexpected sporting victories, which are a common focus of such studies due to the frequent nature of fixtures and sporting outcomes' known influence on mood (Edmans et al., 2007; Wann et al., 1994), have also been linked to increases in the number of votes that incumbent US politicians receive (Healy et al., 2010) as well as, alongside surprising fair weather (Fu et al., 2021), increases in risk-taking behaviour in the form of lottery ticket purchasing (Otto et al., 2016).

Increased risk-taking as a result of unexpected events provides the first half of the motivation for this research proposal. The second half draws from research that finds that unexpected events elicit greater than proportional changes in the sentiment of social media posts. For example, Bhatia et al. (2019) studied changes in the sentiment of tweets from before and after the 2014 US Senate elections and 2014/15 NFL American football games. They found that unexpected outcomes, in both winning and losing directions, elicited greater than proportional changes in Twitter sentiment. Other studies also focusing on the realm of Twitter-inferred mood states have similarly found stronger responses in tweets from unexpected events (Barnaghi et al., 2016; Lucas et al., 2017). It therefore seems that changes in Twitter sentiment are reflective of the surprisingness of real-world events.

Combining these two motivations of thought (that unexpected events increase risk-taking behaviour and unexpected events elicit larger changes in Twitter sentiment) it is reasonable to believe a relationship between Twitter inferred surprisingness and risk-taking behaviour may exist, through each element's respective relationship with unexpected events.

The relationship between the sentiment of tweets and risk-taking behaviour is thus far largely unexplored. However, in a study that harnessed sentiment analysis of over five million tweets, Otto and Eichstaedt (2018) did indeed identify that Twitter-inferred citywide mood was influenced by unexpected sporting and weather events, and simultaneously that changes in Twitter inferred city-wide mood influenced lottery ticket sales.

Following this research, we studied the sentiment of tweets relating to New York sports team over the course of 2012 to understand whether surprisingness conveyed in sports-specific tweets relates to both unexpected sports outcomes and New York lottery ticket sales. We examined two measures of Twitter-inferred surprisingness – changes in daily sentiment (sentiment-based prediction error), and the amount of the emotion 'surprise' conveyed in tweets (surprise scores). First, we examined whether unexpected sporting fixtures – measured by sports prediction errors (the difference between actual and predicted fixture outcomes)— would lead to an increase in the surprisingness of tweets. For example, whether an unexpected loss for the New York Giants resulted in a large drop in the average sentiment of related tweets on that day. Second, following evidence that Twitter inferred mood states related to lottery ticket sales (Otto & Eichstaedt, 2018) – we examined whether the surprisingness of sports-related tweets would positively correlate with next-day lottery ticket sales.

Lottery ticket sales act as a viable proxy for risk-taking behaviour for two reasons. First, the widespread nature of lottery ticket purchasing means that it is possible to make population-level behavioural inferences from changes in ticket sales. Second, each lottery ticket represents a fixed-odds gamble with identical chances of winning, meaning that the expected value of each ticket is identical and thus any changes in the number of tickets purchased can be assumed to be extrinsic to the lottery ticket itself. We study New York lottery ticket sales data, acquired by Otto et al. (Otto et al., 2016), over the course of 2012.

Methods

Twitter Data

Twitter data was scraped through Twitter's API, which allows researchers to freely access historical Twitter feeds. Anonymous tweets for each day of 2012 containing hashtags (e.g. "#giants") relating to one of seven New York sports teams (the Giants, Devils, Knicks, Mets, Yankees, Islanders and Rangers) were collected using the using the python package 'tweepy' (Roesslein, 2020). Tweets relating to two of the teams, the Rangers and Islanders, were removed from the dataset due to sparseness of Twitter data. The final Twitter dataset comprised of 262,427 tweets, with an average of 143 tweets per day for each team.

Pre-processing tweets

Pre-processing involves the cleaning of text data to remove unnecessary noise, normalise words with similar meaning and remove semantically uninformative terms (e.g. stop words and user tags). It both helps to improve computational processing speed, by reducing the size of text data as well as aiding sentiment analysis techniques by creating more consistent usage of language. In our case, this involved removal of stop words (words that carry no semantic information, such as "it" and "the"); lemmatisation of words (grouping inflected forms of words by their lemma (for example "going" and "went" are changed to "go"); converting emojis into to text representation (e.g. A smiley face emoji is converted to "smiley face"); Twitter user tags and URL links are converted to "@USER" and "HTTPURL" respectively and finally laughter was normalised (e.g. "hahahaha" is converted to "haha"). The sentiment analysis tools used in this paper are capable of understanding punctuation,

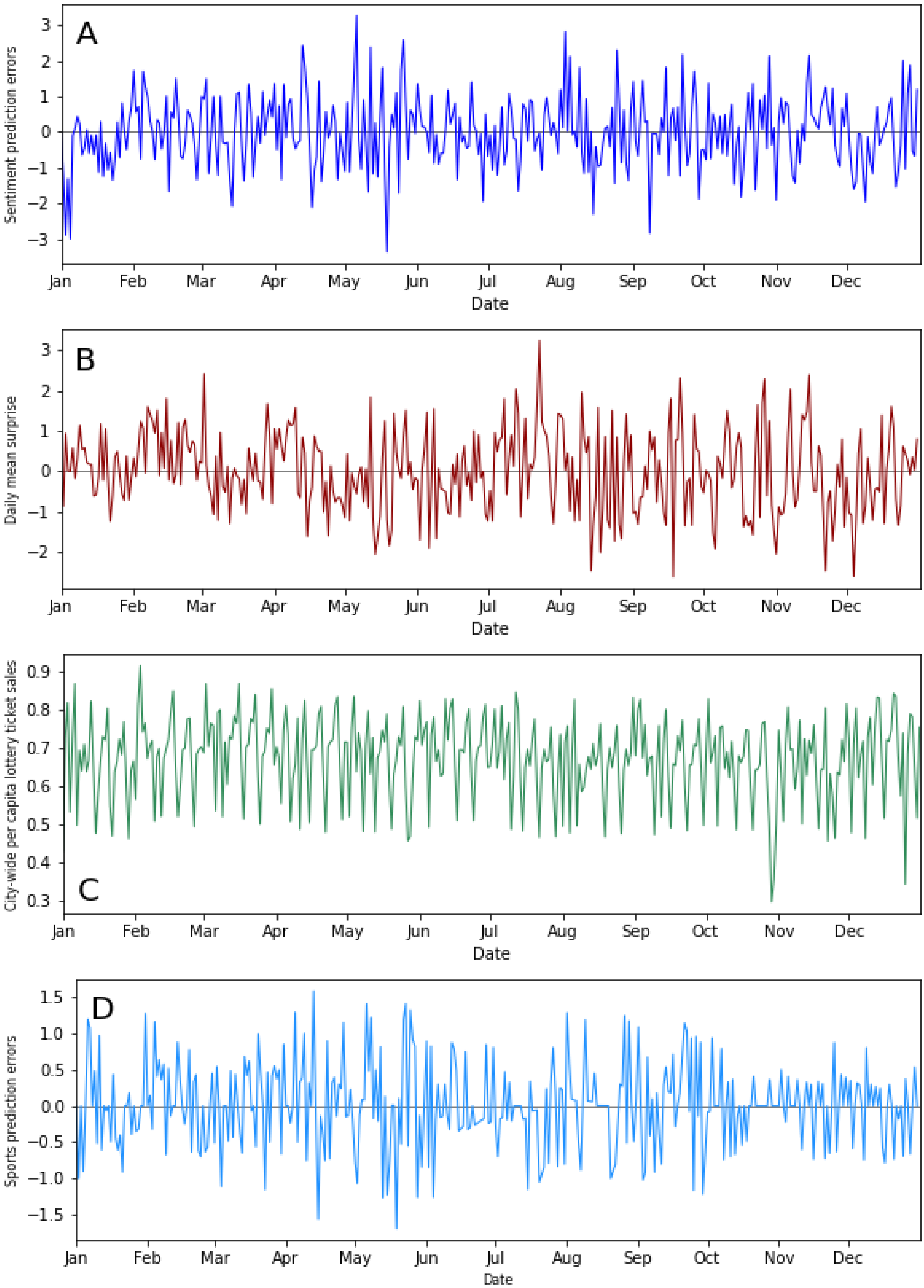

Fig 1: Time-series of Twitter inferred surprisingness scores, city-wide per capita lottery ticket sales and summed sports prediction errors. **A**: The z-scores of the sum of daily ('pysentimiento') sentiment-based prediction errors relating to each of the five NYC sports teams (the Giants, Knicks, Mets, Devils and Yankees). **B**: The z-score of the sum of mean daily 'surprise' scores of tweets relating to each of the five NYC sports teams. **C**: New York city-wide daily per capita lottery ticket sales. **D**: Sum of daily sports fixture prediction errors relating to the five NYC sports teams included in the final Twitter dataset.

therefore punctuation was left in as it may provide semantic richness to short tweets (e.g. "!!!" implies increased valence of emotions).

Twitter inferred surprisingness

Two measures are used to infer the degree of surprisingness expressed towards New York sports teams in the Twitter data. The first is a prediction error of mean daily sentiment scores (the difference between actual mean daily sentiment and expected mean daily sentiment) which is used to quantify *changes* in daily sentiment, given by:

$$PE_t = \left[O(t) - \overline{sent(t)}\,\right]$$

Where $\overline{sent(t)}$ is the expected mean sentiment on day t (see formula below) and $O(t)$ is the actual mean sentiment of tweets on day t. The expected daily sentiment is calculated as an exponentially weighted daily average of sentiment, given by:

$$\overline{sent(t+1)} = \overline{sent(t)} + \alpha\left[O(t) - \overline{sent(t)}\,\right]$$

Where α is a recency parameter, that dictates the weight that more recent days' sentiment contributes to expected sentiment. Following the method used by Otto et al. (2016), the recency parameter, α, was set to 0.1. A large deviation from expected sentiment might reflect the affective response to a real-world unexpected sporting event.

We examined changes in daily sentiment to infer the degree of surprisingness in affective responses to real-world events, rather than simply sentiment itself, due to literature indicating that unexpected events elicit larger changes in sentiment (e.g. Barnaghi et al., 2016; Bhatia et al., 2019). Prediction errors are suitable for quantifying changes in sentiment due to their incorporation of the exponentially weighted mean sentiment (i.e. expected sentiment). Calculating expected sentiment in this way factors in the trend of sentiment from prior days, whilst making no assumptions about influences of the trend (aside from the value of the recency parameter).

The second sentiment-related measure used to infer the degree of surprisingness expressed in the Twitter data is simply a measure of the amount of the emotion 'surprise' that tweets convey. Cognitive literature, as does common sense, tells us that the emotion 'surprise' comes in response to unexpected events (e.g. Meyer et al., 1997). Computational language models are capable of measuring the degree of 'surprise' in bodies of text, thus making it an interesting variable to examine in the context of this paper.

Both methods (sentiment-based prediction errors and surprise scores) were measured from tweets relating to individual teams (e.g. tweets containing "#giant"). For each team, mean daily sentiment scores and surprise scores were calculated and a prediction error was then calculated from the mean daily sentiment scores. It should be noted that there were several days where there were fewer than 5 tweets relating to the Devils and Knicks. This is likely due to too few tweets being made on these days, which we assume means there were no events, such as fixtures or player drafts, happening on these days. On these days, we replaced sentiment-based prediction errors and surprisingness scores with the respective mean for each team from days with more than five tweets. We then aggregated these team-specific prediction errors and surprise scores, producing a sum of daily sentiment-based prediction errors and surprise scores, for use in regression analysis.

Sentiment analysis techniques

Deep learning models – including Google's BERT model (Devlin et al., 2019) – have emerged as state-of-the-art approaches in computational linguistics in recent years. Deep learning methods extract and understand complex dimensions of text data – such as the context within which particular words appear— and from

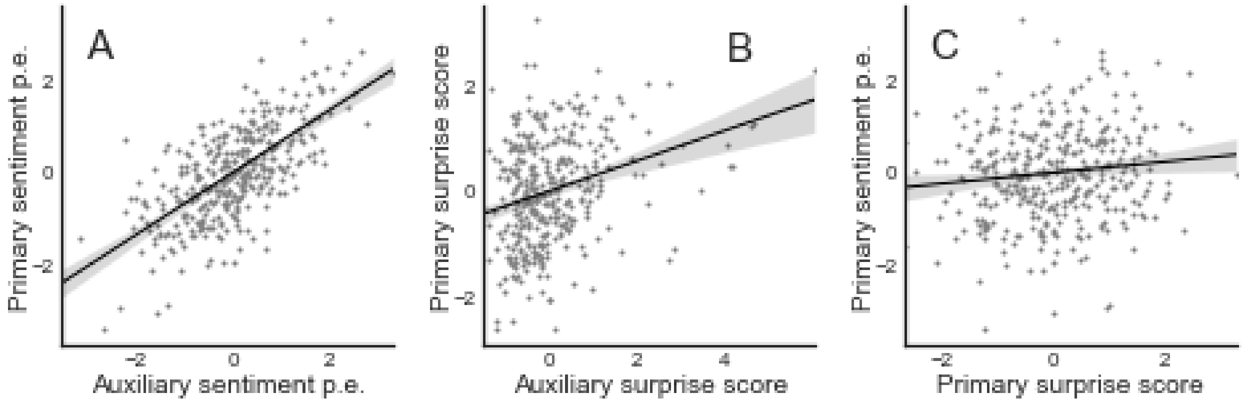

Fig 2: Comparing z-scores of primary and confirmatory Twitter-inferred surprisingness variables. **A:** The z-scores of the daily sentiment prediction errors derived from the primary ('pysentimiento') sentiment analysis technique against prediction errors derived from VADER. **B:** The z-scores of the daily mean surprise scores derived from the primary ('DistilRoBERTa') against the 'distilbert' model. **C:** The z-scores of primary sentiment-based p.e. against z-scores of the primary surprise scores.

that can classify the sentiment ("positive", "negative" or "neutral") and emotion of text data. Deep learning models are pre-trained on vast quantities of data, for example, BERT was trained using a corpus of over 2,500m Wikipedia words. This exposes the models to varying topics and text structures, allowing them to understand text data from a variety of sources. Fine-tuned models are deep learning models that have already been trained on vast quantities of text data and are then exposed to a text of a particular style (e.g. social-media posts) – allowing the model to understand the nuances of different text structures better. Pre-trained models can be accessed through huggingface's 'Transformers' python package (Wolf et al., 2020). We used two such pre-trained deep learning models – "pysentimiento" (Pérez et al., 2021) and "Emotion English DistilRoBERTa-base" (Hartmann, 2022) as primary methods of sentiment analysis to produce sentiment-based prediction errors and surprise scores respectively. Two additional methods were used to cross-reference the outputs from the two primary methods (Hutto & Gilbert, 2014; Sanh et al., 2020).

"Pysentimiento'

'pysentimiento' is an 'out-of-the-box' fine-tuned RoBERTa model used for sentiment analysis. The model is trained on a large body of human-labelled Twitter data, thus making it suitable for determining the sentiment of sports-related tweets here. Each tweet is assigned a likelihood score that the tweet is of "positive", "negative" and "neutral" sentiment, outputting the most likely sentiment as well as the assigned probability. For our sake, "positive" tweets were encoded as $+1 * prob.(positivte)$, negative scores as $-1 * prob.(negative)$ and neutral as 0. The daily mean of the sentiments of tweets relating to each team was then used to form sentiment-based prediction errors.

'DistilRoBERTa'

"*Emotion English DistilRoBERTa-base"* (Hartmann, 2022) is a fine-tuned DistilRoBerta-base model that is used for classifying text as one of Ekman's 6 basic emotions: anger, disgust, fear, joy, neutral, sadness and surprise. It is fine-tuned on a mixture of labelled emotions data from Twitter, Reddit and TV dialogue. The model was used here to predict the likelihood of tweets displaying the emotion 'surprise'. Mean daily surprise scores, summed across teams, were used directly in regression analysis.

Sports prediction errors

Sports prediction errors reflect the gap between the expected and actual results of a fixture. Thus, the bigger the prediction error, the more unexpected the outcome of the fixture. Sports prediction error data for

New York sports teams was sourced from prior research by Otto et al. (2016). They gathered fixture results for the same New York sports teams analysed here from ESPN.com for the 2012 regular and post-season and calculated daily prediction errors according to:

$$PE(t) = O(t) - p_{win}(t)$$

Where t represents the day, $O(t)$ is the outcome of a fixture on date t (win = 1, loss = 0 and draw = 0.5) and $p_{win}(t)$ represents the exponentially weighted mean result, calculated by:

$$p_{win}(t+1) = p_{win}(t) + \alpha[O(t) - p(t)]$$

Where, as previously described, α is a recency parameter set at 0.1. Evidence indicates that prediction errors relating to real-world outcomes (e.g. a student's expectations of exam grades compared to actual exam grades) are more predictive of affective responses than the outcomes themselves (Mellers et al., 1997; Rutledge et al., 2014; Villano et al., 2020).

New York City Lottery Ticket Data

Lottery ticket sales data for New York was also taken from a publicly available dataset provided by Otto et al.'s (2016) prior research. They collected daily lottery ticket sales figures from the New York State Gaming Commission, grouped by 174 New York state ZIP codes. Ticket sales from 14 different fixed-odds lottery games were summed together and divided by the population size of ZIP code areas– giving a daily per capita sales figure. Due to a lack of ZIP code granularity in this paper's Twitter data, we were not able to analyse the relationship between Twitter inferred surprisingness and lottery ticket sales by ZIP code. Therefore, per capita lottery ticket sales were aggregated for the New York state area as a whole and log-transformed. We expected that if any effect of Twitter inferred surprisingness would materialise in next day lottery ticket sales, hence log-transformed next-day per capita lottery ticket sales were used in regression analysis.

Nuisance variables

Following the method adopted by Otto et al. (2016), nuisance variables were included in regression analysis to control for possible exogenous effects on New York lottery ticket sales and Twitter data. These included binary encoded variables for days of the week, the month, national holidays and other significant calendar dates that fell on Mondays or Fridays (e.g. New Year's Day, Columbus Day and Valentine's Day), pay-days (assumed to fall on the 1st and 15th of each month) and severe-weather days (e.g. Hurricane Sandy made landfall in New York on 29/10/12). Inclusion of these variables in regression models controls for the separate influence on lottery ticket sales each date may have had – for example, average per capita city-wide lottery ticket sales was only 0.48 during the days around Hurricane Sandy (29/10/12 – 1/11/12), down from an annual mean of 0.67. Twitter sentiment is similarly likely to be influenced by the date, for example, it has been found that Twitter inferred happiness is lowest on Mondays (Dodds et al., 2011) and thus nuisance variables were controlled for in regressions using Twitter inferred surprisingness as a dependent variable.

Regressions Models

Linear regression models were carried out in the R programming language to analyse 1) whether real-world sports prediction errors were predictive of Twitter-inferred surprisingness, and 2) whether Twitter-inferred surprisingness was predictive of city-wide lottery ticket sales. Surprisingness variables were included as z-scores in all regressions. In the first set of regressions, each surprisingness variable was used as the dependent variable and the daily sum of sport prediction errors and nuisance variables were used as independent variables. In the second set of linear regressions, the log-transformed

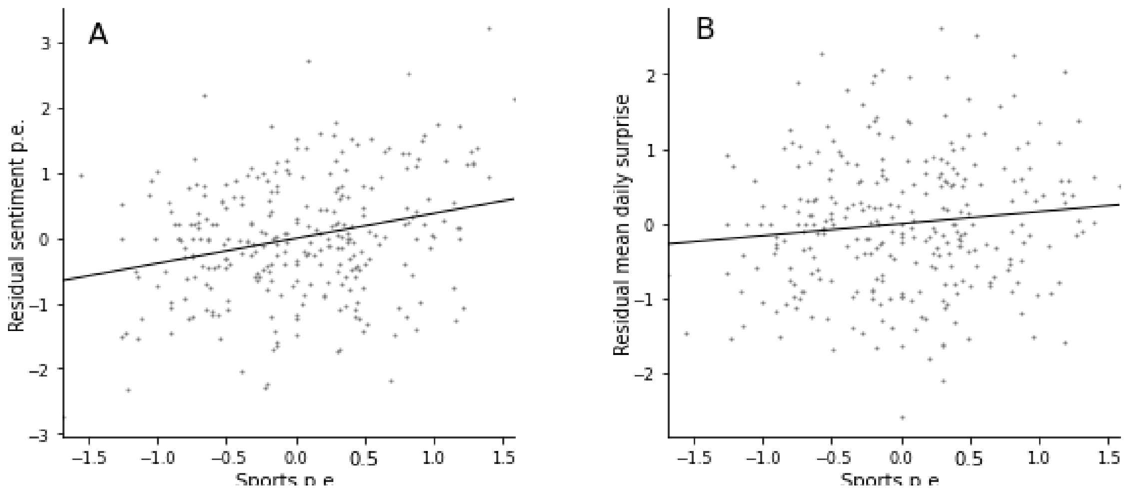

Fig 3: A: Residual sentiment-based prediction errors against real-world sports prediction errors. **B:** Residual average daily surprise scores against real-world sports prediction errors. Residual values are taken from linear regression models regressing Twitter-inferred surprisingness variables against a variety of cyclical nuisance variables (see nuisance variables).

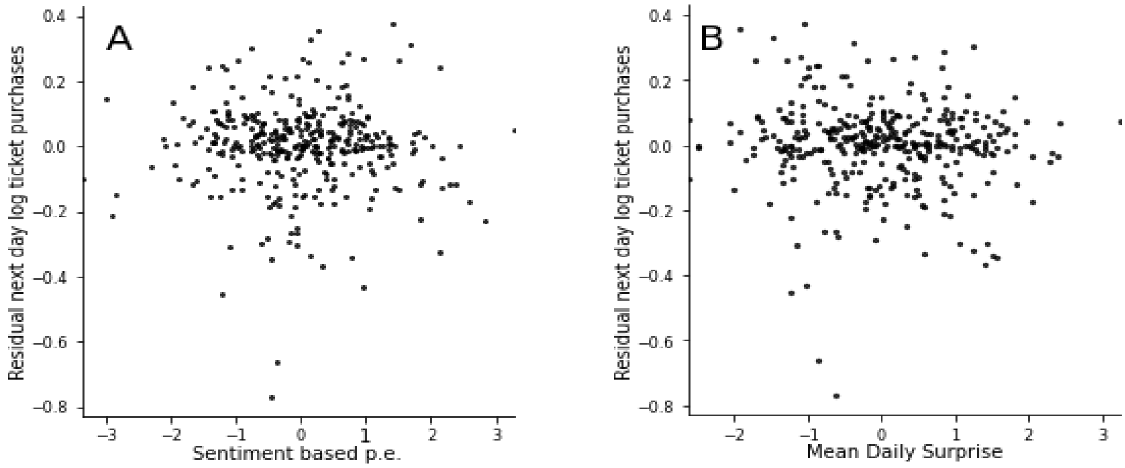

Fig 4: Log of residual next day per-capita lottery ticket sales against **A** the z-scores of sentiment-based prediction errors and **B** the z-scores of the mean daily surprise scores. Residual values are taken from linear regression models regressing Twitter-inferred surprisingness variables against a variety of cyclical nuisance variables (see nuisance variables).

next-day city-wide per capita lottery ticket sales were used as the dependent variable with the two surprisingness variables (in separate regressions) and nuisance variables as independent variables. A summary of regression coefficients is included in regression tables in the appendix. The dependent variables were also regressed solely on the nuisance variables, and residual values from these models were used to depict the relationship between the variables of interest, excluding the effects of nuisance variables (see fig. 3 and 4).

Results

Twitter-inferred surprisingness

In the absence of labelled Twitter to evaluate the performance of the two sentiment analysis methods ('pysentimiento' and 'DistilRoBERTa') used in this study, we cross-referenced the sentiment prediction errors and surprise scores with those produced using additional ('auxiliary') sentiment analysis methods.

We compared mean daily prediction errors produced using the 'pysentimiento' model with those produced by a lexicon-based sentiment analysis technique that is tuned towards analysing microblog posts (Hutto & Gilbert, 2014). We observed a relatively strong correlation between the two mean daily sentiment prediction error variables ($r(364) = .68,\ p < 0.001$), indicating that the different methods of sentiment analysis broadly agreed with one another (see fig. 2A).

Similarly, we cross-referenced surprise scores produced using the 'Emotion English DistilRoBERTa-base' model and another deep learning model (Sanh et al., 2020), also fine-tuned on Twitter data and found a significant, yet lower in magnitude Pearson correlation coefficient ($r(364) = .29,\ p < 0.001$), indicating agreement to a lesser extent in 'surprise' scores yet nonetheless a positive relationship.

We also examined the relationship between the z-scores of the 'pysentimiento' sentiment-based prediction errors and the 'surprise' scores to understand whether both measures mirrored the same underlying information from the Twitter data. Interestingly, we observed a weak correlation ($r(364) = .12,\ p = 0.023$) indicating that whilst the two types of 'surprisingness' measures are related, they are likely on the whole to be made up of different features from the Twitter data.

Twitter-inferred surprisingness and sport prediction errors

We obtained tweets relating to five New York sports teams (each tweet included a team's official hashtag, e.g. "#giants") for each day in 2012. Two separate measures of Twitter inferred surprisingness were drawn from this Twitter dataset: sentiment-based prediction errors and surprise scores (see Methods). To measure whether real-world unexpected sporting events influenced Twitter-inferred surprisingness, sports prediction errors relating to the same five teams were sourced from data made available by Otto et al. (2016). Prediction errors were used to measure the unexpectedness of an outcome on account of support from the literature that they are more reflective of emotional responses than actual outcomes (Rutledge et al., 2014; Villano et al., 2020). Following research that has found that unexpected outcomes are reflected in larger changes in Twitter sentiment (Barnaghi et al., 2016; Bhatia et al., 2019; Lucas et al., 2017), we hypothesised that unexpected sporting wins would result in an increase in Twitter inferred surprisingness—that is to say that sports prediction errors would positively relate to Twitter inferred surprisingness.

Linear regression, using Twitter inferred surprisingness as the dependent variable and sports prediction errors as independent variables (whilst controlling for nuisance variables) uncovered a significant and positive effect of sports prediction errors upon the sentiment-based prediction errors ($\beta = 0.45, p < 0.001$) and a positive, yet insignificant, effect of sports prediction errors upon surprise scores ($\beta = 0.18,\ p = 0.057$) (see tables 1 and 2 in the appendix).

To examine whether the effect differed for better-than-expected sports outcomes and worse-than-expected outcomes, we carried out the same regressions on days with positive and negative prediction errors separately (See tables 7-10 in the appendices). Interestingly, the only significant effect was between sentiment-based prediction errors and positive sports prediction errors ($\beta = 0.74\ \ p = 0.002$).

Additionally, we created sentiment-based prediction errors based on tweets that were either of "positive" or "neutral" sentiment. This was done based on a broad assumption that "negative" tweets were more likely to be made by fans of opposition teams. Fans of opposition may tweet with sentiment in the opposite direction than we'd expect – for example if the Giants had an unexpected win (resulting in a positive prediction error) their tweets might include "#giants" and be gloomy, whilst we'd expect celebratory tweets. Therefore, excluding "negative" tweets might result in tweets more reflective of the outcomes of teams included in this paper. We then ran the same regressions—studying the different effects on days with positive and negative sports prediction errors (see tables 11 and 12). It's important to note that sentiment-based prediction errors still reflect *changes* in the "neutral" and "positive" tweets here. Again, we found that sports prediction errors had a significant effect on sentiment-based prediction errors on days with positive sports prediction errors ($\beta = 1.07\ p < 0.001$), but not on days with negative prediction errors ($\beta = .39\ p = 0.078$).

Twitter inferred surprisingness and lottery ticket sales

The significant and positive effect of sports prediction errors upon the sentiment-based prediction errors provides evidence that unexpected sporting outcomes are reflected in Twitter-inferred surprisingness. To investigate our second research question, whether Twitter inferred surprisingness would positively affect New York lottery ticket sales, daily citywide per capita lottery ticket sales for 2012 were sourced from data made publicly available by Otto et al. (2016). It was expected that if a high degree of surprisingness was expressed, possibly reflecting that an unexpected event had occurred, the effect on gambling behaviour would surface on the following day. Therefore, we regressed next-day log-transformed citywide lottery ticket sales against each Twitter inferred surprisingness measure separately, whilst controlling for nuisance variables. The effect of both Twitter inferred surprisingness measures on lottery ticket sales was insignificant (sentiment-based prediction errors effect: $\beta = 0.00, p = 0.639$,; surprise scores: $\beta = -0.011, p = 0.181$).

Discussion

We harnessed two methods of sentiment analysis on a large dataset of tweets to study the relationships between Twitter-inferred surprisingness, real-world unexpected sporting outcomes and lottery ticket sales. We found evidence on the one hand that sport prediction errors are predictive of sentiment prediction errors, whilst on the other that they are not predictive of the amount of 'surprise' expressed in tweets. Furthermore, we found no evidence of a relationship between lottery ticket sales and either of our two Twitter inferred surprisingness variables.

Due to a lack of labelled data upon which to evaluate how well the two sentiment analysis techniques performed, we generated the same two Twitter-inferred mood variables using alternative sentiment analysis techniques and compared the outcomes. The logic to this was that if two different sentiment analysis techniques produce consistent results, it is would act as reassurance that sentiment analysis methods worked as intended.

Cross-referencing each Twitter-inferred mood variable in this way revealed positive significant correlations between both sets of variables. Interestingly though, a relatively weaker correlation between 'surprise' scores highlighted ambiguity around the reliability of the 'surprise' score sentiment analysis techniques (see fig. 2B). The black-box nature of deep learning models muddies the water as to the cause of variation in 'surprise' scores; however, literature has found surprise to be an ambiguous emotion in itself. For example, it may be a mildly pleasant emotion (e.g. a surprise birthday party), yet many people "don't like surprises" and studies have found it to be both slightly negative (Noordewier & Breugelmans, 2013) and neutral (Marmolejo-Ramos et al., 2017) in valence. Ambiguity around how 'surprise' is experienced may contribute to inconsistencies found here.

To understand whether the sentiment-based prediction errors and surprise scores reflected similar features of the underlying Twitter data, we calculated the Pearson correlation coefficient between the two sets of variables. A low, yet significant correlation implied that these variables represented different underlying features of the Twitter data. A possible explanation of this is that a change in sentiment corresponds to a change in the valence of emotions (the pleasantness/unpleasantness of emotions) displayed in tweets. Multiple emotions can exist at similar degrees of valence – e.g. “happy”, “satisfied” and “calm” are all relatively pleasant—whereas the emotion ‘surprise’ is a single emotion in itself, making the variables inherently different.

We found an encouraging relationship between sports prediction errors and sentiment-based predictions. This finding implicates changes in sentiment as being reflective of the affective responses to unexpected sporting events and aligns with findings from a study carried out by Bhatia et al. (Bhatia et al., 2019). Measuring Twitter sentiment on days before and after an event, they found that unexpected American Football and US senate election outcomes resulted in larger changes in Twitter sentiment than fixture/election outcomes with more likely odds.

Our research differs by measuring Twitter sentiment on a daily basis, without any prior knowledge of whether a sporting event occurs on a given day, thus using a different approach to understanding the relationship between Twitter inferred mood and unexpected events. There are 77 days within the data upon which there are no sports prediction errors, due to no fixtures occurring on these days – so to confirm the relationship was changed by days we also studied this relationship excluding these days and the significant effect remains (see tables 5 and 6 in the appendix). This finding also supports evidence that prediction errors are determinant of affective responses to events, as has also been displayed in the literature (Eldar et al., 2016; Rutledge et al., 2014; Villano et al., 2020).

Separately studying the effect of positive and negative sports prediction errors, we found that only positive sports prediction errors had a significant effect on sentiment-based prediction errors. This might be reflective of the tweeting behaviours of the fans—some fans may only tweet when their team wins unexpectedly but refrains from tweeting when they lose.

We also studied sentiment-based prediction errors calculated from only “positive” and “neutral” tweets—on the broad assumption that fans of other teams were more likely to make “negative” tweets about the teams. This analysis uncovered a significant and more positive relationship between these “fans only” sentiment prediction errors on days with positive sports prediction errors, yet not on days with negative prediction errors. Whilst this lends support to the idea that fans may asymmetrically tweet about their teams, the assumptions that non-fans are more likely to tweet with “negative” sentiment is tenuous.

The relationship between sports prediction errors and surprise scores was found to be positive yet insignificant. Whilst surprise is a likely emotional reaction to unexpected events, it seems that this reaction does not materialise in ‘surprise’ expressed in social media posts. It is worth noting that the effect between sports prediction errors and surprise scores is approaching significance, which may mean that analysing a larger Twitter dataset with greater statistical power, or tweets related to unexpected events in a different domain (e.g. weather patterns) may reveal evidence to the contrary.

The lack of relationship between surprisingness of tweets and lottery ticket sales meant that we found no evidence that Twitter-inferred surprisingness would lead to risk-taking behaviour, and thus limiting this papers’ contribution to research into affective states and behavioural change. One unlikely explanation for the lack of a relationship found here might be that no relationship exists between Twitter-inferred affective states and risk-taking behaviour, however, existing research by Otto et Eichstaedt (2018) has found evidence of the contrary. Instead, the limited granularity and size of the Twitter dataset used is likely to be a constraining factor to truly examine this question. The strength of Otto and Eichstaedt's research (2018) lies in the amount of Twitter data analysed as well as collecting locational information of tweets. They collected 12.2 million tweets across 2012 and 2013 and were able to carry out location-specific analysis by assigning locations-of-origin to tweets according to the location field in the users' profiles. This enabled comparison between localised changes in lottery ticket sales and Twitter-inferred mood. Whilst we attempted to scape geo-tagged tweets (where longitudinal and latitudinal coordinates of the location from where the tweet was posted), we were not able to collect enough data, firstly, because a only very small proportion of Twitter user's used the geo-tagging feature (estimated at 0.85% (Sloan et al., 2013)) and secondly, Twitter disabled specific geo-tagging of posts in 2019 – such that specific geo-tagging information can only be acquired for images and posts from third-parties (e.g. Instagram). A possible extension of this report would therefore be to follow Otto and Eichstaedt's data collection method—using locational information.

Whilst this highlights the necessity that to draw meaningful conclusions from *big data* analytics, it is a requirement to have *big* amounts of data, this paper has highlighted that from studying time-series Twitter data, it is possible to make inferences about the expectations of real-world events. Analysing time-series of massive datasets of social-media data has been used predict subjective well-being (Schwartz et al., 2013), identify responses to natural disasters (Yin et al., 2015) and to detect real-time traffic incidences (Dabiri & Heaslip, 2019). Expectations of events like sports matches and elections can be quantified in betting-odds, however, expectations of other events may be more subjective. For example, expectations of the quality of a film or play or expectations about geo-political decisions, such as a leader deciding to go to war. Studying changes in sentiment of large datasets of text data may serve as a tool to quantify expectations in such situations and thus, examine the role of unexpected outcomes in a variety of contexts and their behavioural implications.

Bibliography

Arkes, H. R., Herren, L. T., & Isen, A. M. (1988). The role of potential loss in the influence of affect on risk-taking behavior. *Organizational Behavior and Human Decision Processes*, *42*(2), 181–193. https://doi.org/10.1016/0749-5978(88)90011-8

Ashby, F. G., Isen, A. M., & Turken, A. U. (1999). A neuropsychological theory of positive affect and its influence on cognition. *Psychological Review*, *106*(3), 529. https://doi.org/10.1037/0033-295X.106.3.529

Barnaghi, P., Ghaffari, P., & Breslin, J. G. (2016). Opinion Mining and Sentiment Polarity on Twitter and Correlation between Events and Sentiment. *2016 IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService)*, 52–57. https://doi.org/10.1109/BigDataService.2016.36

Bassi, A., Colacito, R., & Fulghieri, P. (2013). 'O Sole Mio: An Experimental Analysis of Weather and Risk Attitudes in Financial Decisions. *The Review of Financial Studies*, *26*(7), 1824–1852. https://doi.org/10.1093/rfs/hht004

Bhatia, S., Mellers, B., & Walasek, L. (2019). Affective responses to uncertain real-world outcomes: Sentiment change on Twitter. *PLOS ONE*, *14*(2), e0212489. https://doi.org/10.1371/journal.pone.0212489

Cao, M., & Wei, J. (2005). Stock market returns: A note on temperature anomaly. *Journal of Banking & Finance*, *29*(6), 1559–1573. https://doi.org/10.1016/j.jbankfin.2004.06.028

Chen, D. L., & Loecher, M. (2019). *Mood and the Malleability of Moral Reasoning* (SSRN Scholarly Paper No. 2740485). https://doi.org/10.2139/ssrn.2740485

Chepenik, L. G., Cornew, L. A., & Farah, M. J. (2007). The influence of sad mood on cognition. *Emotion*, *7*(4), 802–811. https://doi.org/10.1037/1528-3542.7.4.802

Clark, L. A., & Watson, D. (1988). Mood and the mundane: Relations between daily life events and self-reported mood. *Journal of Personality and Social Psychology*, *54*(2), 296–308. https://doi.org/10.1037/0022-3514.54.2.296

Dabiri, S., & Heaslip, K. (2019). Developing a Twitter-based traffic event detection model using deep learning architectures. *Expert Systems with Applications*, *118*, 425–439. https://doi.org/10.1016/j.eswa.2018.10.017

Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences*, *108*(17), 6889–6892. https://doi.org/10.1073/pnas.1018033108

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* (arXiv:1810.04805; Version 2). arXiv. https://doi.org/10.48550/arXiv.1810.04805

Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., & Danforth, C. M. (2011). Temporal Patterns of Happiness and Information in a Global Social Network: Hedonometrics and Twitter. *PLOS ONE*, *6*(12), e26752. https://doi.org/10.1371/journal.pone.0026752

Edmans, A., García, D., & Norli, Ø. (2007). Sports Sentiment and Stock Returns. *The Journal of Finance*, *62*(4), 1967–1998. https://doi.org/10.1111/j.1540-6261.2007.01262.x

Eich, E., & Macaulay, D. (2006). Cognitive and clinical perspectives on mood-dependent memory. In *Affect in social thinking and behavior* (pp. 105–121). Psychology Press.

Eldar, E., Rutledge, R. B., Dolan, R. J., & Niv, Y. (2016). Mood as Representation of Momentum. *Trends in Cognitive Sciences*, *20*(1), 15–24. https://doi.org/10.1016/j.tics.2015.07.010

Eren, O., & Mocan, N. (2018). Emotional Judges and Unlucky Juveniles. *American Economic Journal: Applied Economics*, *10*(3), 171–205. https://doi.org/10.1257/app.20160390

Forgas, J. P. (2002). Feeling and Doing: Affective Influences on Interpersonal Behavior. *Psychological Inquiry*, *13*(1), 1–28. https://doi.org/10.1207/S15327965PLI1301_01

Fu, H.-N., Monson, E., & Otto, R. (2021). The Relationship Between Unexpected Outcomes and Lottery Gambling Rates in a Large Canadian Metropolitan Area. *Critical Gambling Studies*, *2*(1), 55–67. https://doi.org/10.29173/cgs28

Hartmann, J. (2022). *Emotion English DistilRoBERTa-base*. https://huggingface.co/j-hartmann/emotion-english-distilroberta-base/

Healy, A. J., Malhotra, N., & Mo, C. H. (2010). Irrelevant events affect voters' evaluations of government performance. *Proceedings of the National Academy of Sciences*, *107*(29), 12804–12809. https://doi.org/10.1073/pnas.1007420107

Heyes, A., & Saberian, S. (2019). Temperature and Decisions: Evidence from 207,000 Court Cases. *American Economic Journal: Applied Economics*, *11*(2), 238–265. https://doi.org/10.1257/app.20170223

Hirshleifer, D., & Shumway, T. (2003). Good Day Sunshine: Stock Returns and the Weather. *The Journal of Finance*, *58*(3), 1009–1032. https://doi.org/10.1111/1540-6261.00556

Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291. https://doi.org/10.2307/1914185

Kim, D., Lam, J., Kutz, A., & Yoon, K. L. (2021). Punishment sensitivity and risk taking in depressed mood. *Motivation and Emotion*, *45*(1), 122–130. https://doi.org/10.1007/s11031-020-09860-4

Kőszegi, B., & Rabin, M. (2006). A Model of Reference-Dependent Preferences*. *The Quarterly Journal of Economics*, *121*(4), 1133–1165. https://doi.org/10.1093/qje/121.4.1133

Kuppens, P., Oravecz, Z., & Tuerlinckx, F. (2010). Feelings change: Accounting for individual differences in the temporal dynamics of affect. *Journal of Personality and Social Psychology*, *99*(6), 1042–1060. https://doi.org/10.1037/a0020962

Lerner, J. S., Small, D. A., & Loewenstein, G. (2004). Heart Strings and Purse Strings: Carryover Effects of Emotions on Economic Decisions. *Psychological Science*, *15*(5), 337–341. https://doi.org/10.1111/j.0956-7976.2004.00679.x

Lucas, G. M., Gratch, J., Malandrakis, N., Szablowski, E., Fessler, E., & Nichols, J. (2017). GOAALLL!: Using sentiment in the world cup to explore theories of emotion. *Image and Vision*

Computing, *65*, 58–65. https://doi.org/10.1016/j.imavis.2017.01.006

Manucia, G. K., Baumann, D. J., & Cialdini, R. B. (1984). Mood influences on helping: Direct effects or side effects? *Journal of Personality and Social Psychology*, *46*(2), 357–364. https://doi.org/10.1037/0022-3514.46.2.357

Marmolejo-Ramos, F., Correa, J. C., Sakarkar, G., Ngo, G., Ruiz-Fernández, S., Butcher, N., & Yamada, Y. (2017). Placing joy, surprise and sadness in space: A cross-linguistic study. *Psychological Research*, *81*(4), 750–763. https://doi.org/10.1007/s00426-016-0787-9

Martin, E. A., & Kerns, J. G. (2011). The influence of positive mood on different aspects of cognitive control. *Cognition and Emotion*, *25*(2), 265–279. https://doi.org/10.1080/02699931.2010.491652

Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision Affect Theory: Emotional Reactions to the Outcomes of Risky Options. *Psychological Science*, *8*(6), 423–429. https://doi.org/10.1111/j.1467-9280.1997.tb00455.x

Meyer, W.-U., Reisenzein, R., & Schützwohl, A. (1997). Toward a Process Analysis of Emotions: The Case of Surprise. *Motivation and Emotion*, *21*(3), 251–274. https://doi.org/10.1023/A:1024422330338

Ottenheimer, D. J., Bari, B. A., Sutlief, E., Fraser, K. M., Kim, T. H., Richard, J. M., Cohen, J. Y., & Janak, P. H. (2020). A quantitative reward prediction error signal in the ventral pallidum. *Nature Neuroscience*, *23*(10), 1267–1276. https://doi.org/10.1038/s41593-020-0688-5

Otto, A. R., & Eichstaedt, J. C. (2018). Real-world unexpected outcomes predict city-level mood states and risk-taking behavior. *PLOS ONE*, *13*(11), e0206923. https://doi.org/10.1371/journal.pone.0206923

Otto, A. R., Fleming, S. M., & Glimcher, P. W. (2016). Unexpected but Incidental Positive Outcomes Predict Real-World Gambling. *Psychological Science*, *27*(3), 299–311. https://doi.org/10.1177/0956797615618366

Pérez, J. M., Giudici, J. C., & Luque, F. (2021). *pysentimiento: A Python Toolkit for Sentiment Analysis and SocialNLP tasks* (arXiv:2106.09462). arXiv. https://doi.org/10.48550/arXiv.2106.09462

Roesslein, J. (2020). *Tweepy: Twitter for Python!* Https://Github.Com/Tweepy/Tweepy

Rutledge, R. B., Skandali, N., Dayan, P., & Dolan, R. J. (2014). A computational and neural model of momentary subjective well-being. *Proceedings of the National Academy of Sciences*, *111*(33), 12252–12257. https://doi.org/10.1073/pnas.1407535111

Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2020). *DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter* (arXiv:1910.01108). arXiv. https://doi.org/10.48550/arXiv.1910.01108

Saunders, E. M. (1993). Stock Prices and Wall Street Weather. *The American Economic Review*, *83*(5), 1337–1345.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science (New York, N.Y.)*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Schwartz, H., Eichstaedt, J., Kern, M., Dziurzynski, L., Lucas, R., Agrawal, M., Park, G., Lakshmikanth, S., Jha, S., Seligman, M., & Ungar, L. (2013). Characterizing Geographic Variation in Well-Being Using Tweets. *Proceedings of the International AAAI Conference on Web and Social Media*, *7*(1), 583–591.

Sharot, T. (2011). The optimism bias. *Current Biology*, *21*(23), R941–R945. https://doi.org/10.1016/j.cub.2011.10.030

Sloan, L., Morgan, J., Housley, W., Williams, M., Edwards, A., Burnap, P., & Rana, O. (2013). Knowing the Tweeters: Deriving Sociologically Relevant Demographics from Twitter. *Sociological Research Online*, *18*(3), 74–84. https://doi.org/10.5153/sro.3001

Tversky, A., & Kahneman, D. (1991). Loss Aversion in Riskless Choice: A Reference-Dependent Model. *The Quarterly Journal of Economics*, *106*(4), 1039–1061. https://doi.org/10.2307/2937956

Villano, W. J., Otto, A. R., Ezie, C. E. C., Gillis, R., & Heller, A. S. (2020). Temporal dynamics of real-world emotion are more strongly linked to prediction error than outcome. *Journal of Experimental Psychology: General*, *149*(9), 1755–1766. https://doi.org/10.1037/xge0000740

Wann, D. L., Dolan, T. J., McGeorge, K. K., & Allison, J. A. (1994). Relationships between spectator identification and spectators' perceptions of influence, spectators' emotions, and competition outcome. *Journal of Sport & Exercise Psychology*, *16*(4), 347–364.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., … Rush, A. (2020). Transformers: State-of-the-Art Natural Language Processing. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, 38–45. https://doi.org/10.18653/v1/2020.emnlp-demos.6

Yin, J., Karimi, S., Lampert, A., Cameron, M., Robinson, B., & Power, R. (2015, June 27). Using Social Media to Enhance Emergency Situation Awareness: Extended Abstract. *Twenty-Fourth International Joint Conference on Artificial Intelligence*. Twenty-Fourth International Joint Conference on Artificial Intelligence. https://www.aaai.org/ocs/index.php/IJCAI/IJCAI15/paper/view/11210

Appendices

Table 1: Linear regressions results of sentiment-based prediction errors against sports predictions errors, controlling for nuisance variables

Effect	*Estimate*	*SE*	*t*	*p*
Intercept	-.54	.22	-2.42	**.016**
Sport p.e.	.45	.09	4.84	**<0.001**
TUE	-.21	.20	-1.05	.294
WED	-.06	.20	-.32	.748
THU	.03	.20	.15	.884
FRI	.22	.20	1.08	.280
SAT	.21	.20	1.09	.278
SUN	.06	.20	.32	.748
FEB	.66	.26	2.54	**.012**
MAR	.53	.25	2.11	**.036**
APR	.51	.26	1.98	**.049**
MAY	.67	.25	2.63	**.009**
JUN	.48	.25	1.87	.063
JUL	.50	.25	1.96	.051
AUG	.51	.25	2.01	**.045**
SEP	.56	.26	2.17	**.031**
OCT	.48	.26	1.86	.063
NOV	.63	.26	2.44	**.015**
DEC	.27	.26	1.07	.287
FIRST_OF_MONTH	.03	.31	.09	.926
FIFTEENTH_OF_MONTH	.09	.29	.30	.765
HURRICANE	.08	.52	.15	.881
INDEPENDENCEDAY	.37	1.00	.37	.710
THANKSGIVING	1.15	1.00	1.15	.250
CHRISTMASDAY	2.81	1.00	2.80	**.005**
DAYAFTERCHRISTMAS	-.87	1.00	-.87	.386
LABORDAY	-.23	1.00	-.22	.822
EASTER	.73	1.00	.74	.463
NEWYEARSDAY	.52	1.00	.52	.603
COLUMBUSDAY	.84	1.00	.84	.401
MEMORIALDAY	-.32	1.01	-.32	.748
BIRTHDAYOFMARTINLU-THERKINGJR	.68	1.00	.68	.497
VETERANSDAY	.14	1.00	.14	.892
WASHINGTONSBIRTHDAY	.67	1.00	.67	.503
VALENTINESDAY	-.29	1.00	-.29	.770

Table 2: Linear regressions results of mean daily surprise scores against sports predictions errors, controlling for nuisance variables

Effect	***Estimate***	***SE***	***t***	***p***
Intercept	.07	.22	.32	.750
Sport p.e.	.18	.09	1.91	.056
TUE	-.15	.20	-.74	.459
WED	.17	.20	.84	.404
THU	.15	.20	.75	.454
FRI	.30	.20	1.53	.126
SAT	-.09	.19	-.48	.630
SUN	-.36	.20	-1.86	.063
FEB	.54	.26	2.10	.037
MAR	-.14	.25	-.54	.589
APR	-.02	.25	-.06	.951
MAY	-.34	.25	-1.36	.174
JUN	-.28	.25	-1.12	.265
JUL	.42	.25	1.66	.098
AUG	-.20	.25	-.81	.420
SEP	-.09	.25	-.37	.712
OCT	-.09	.26	-.35	.726
NOV	-.21	.26	-.81	.419
DEC	-.31	.25	-1.23	.219
FIRST_OF_MONTH	-.21	.31	-.68	.497
FIFTEENTH_OF_MONTH	-.01	.29	-.05	.961
HURRICANE	-.72	.51	-1.42	.158
INDEPENDENCEDAY	.33	.99	.33	.741
THANKSGIVING	-2.49	.99	-2.52	.012
CHRISTMASDAY	1.13	.99	1.14	.255
DAYAFTERCHRISTMAS	.45	.99	.46	.648
LABORDAY	-1.12	.99	-1.13	.260
EASTER	1.38	.99	1.39	.164
NEWYEARSDAY	.05	.99	.05	.956
COLUMBUSDAY	.34	.99	.35	.730
MEMORIALDAY	.41	.99	.42	.677
BIRTHDAYOFMARTINLUTHERKINGJR	-.57	.99	-.58	.563
VETERANSDAY	1.54	.99	1.55	.121
WASHINGTONSBIRTHDAY	.71	.99	.71	.476
VALENTINESDAY	-.56	.99	-.56	.574

Table 3: Linear regressions results of the log of next day citywide per capita lottery ticket sales against sentiment-based predictions errors, controlling for nuisance variables

Effect	***Estimate***	***SE***	***t***	***p***
Intercept	-.45	.03	-13.87	**<0.001**
Sentiment p.e.	.00	.01	-.47	.639
TUE	.03	.03	1.13	.260
WED	.09	.03	3.07	**.002**
THU	.16	.03	5.60	**<0.001**
FRI	-.06	.03	-2.11	**.036**
SAT	-.18	.03	-6.19	**<0.001**
SUN	.02	.03	.71	.476
FEB	.05	.04	1.45	.149
MAR	.09	.04	2.52	**.012**
APR	.05	.04	1.23	.220
MAY	.01	.04	.31	.757
JUN	.08	.04	2.07	**.039**
JUL	.01	.04	.37	.714
AUG	.00	.04	-.01	.990
SEP	.03	.04	.74	.463
OCT	-.01	.04	-.36	.721
NOV	.01	.04	.26	.795
DEC	.05	.04	1.22	.223
FIRST_OF_MONTH	.06	.04	1.37	.173
FIFTEENTH_OF_MONTH	.04	.04	1.00	.317
HURRICANE	-.25	.07	-3.33	**<0.001**
INDEPENDENCEDAY	-.02	.14	-.14	.890
THANKSGIVING	-.17	.14	-1.21	.226
CHRISTMASDAY	.00	.14	.01	.990
DAYAFTERCHRISTMAS	.08	.14	.54	.593
LABORDAY	.02	.14	.17	.864
EASTER	-.03	.14	-.22	.822
NEWYEARSDAY	.11	.14	.78	.436
COLUMBUSDAY	.03	.14	.19	.850
MEMORIALDAY	-.05	.14	-.32	.748
BIRTHDAYOFMARTINLU-THERKINGJR	.04	.14	.27	.791
VETERANSDAY	-.02	.14	-.16	.869
WASHINGTONSBIRTHDAY	.03	.14	.23	.815
VALENTINESDAY	.04	.14	.28	.782

Table 4: Linear regressions results of the log of next day citywide per capita lottery ticket sales against mean daily surprise scores, controlling for nuisance variables

Effect	***Estimate***	***SE***	***t***	***p***
Intercept	-.45	.03	-13.98	**<0.001**
Surprise score	-.01	.01	-1.34	.181
TUE	.03	.03	1.11	.268
WED	.09	.03	3.15	**.002**
THU	.16	.03	5.67	**<0.001**
FRI	-.06	.03	-2.02	**.044**
SAT	-.18	.03	-6.28	**<0.001**
SUN	.02	.03	.58	.565
FEB	.06	.04	1.55	.123
MAR	.09	.04	2.45	**.015**
APR	.04	.04	1.19	.237
MAY	.01	.04	.15	.883
JUN	.07	.04	1.96	.051
JUL	.02	.04	.44	.660
AUG	.00	.04	-.12	.903
SEP	.02	.04	.66	.508
OCT	-.02	.04	-.43	.666
NOV	.01	.04	.14	.887
DEC	.04	.04	1.10	.272
FIRST_OF_MONTH	.06	.04	1.32	.189
FIFTEENTH_OF_MONTH	.04	.04	.99	.321
HURRICANE	-.25	.07	-3.44	**<0.001**
INDEPENDENCEDAY	-.02	.14	-.12	.902
THANKSGIVING	-.20	.14	-1.42	.157
CHRISTMASDAY	.00	.14	.03	.979
DAYAFTERCHRISTMAS	.08	.14	.59	.553
LABORDAY	.01	.14	.09	.926
EASTER	-.02	.14	-.14	.888
NEWYEARSDAY	.11	.14	.77	.440
COLUMBUSDAY	.03	.14	.19	.847
MEMORIALDAY	-.04	.14	-.29	.776
BIRTHDAYOFMARTINLU-THERKINGJR	.03	.14	.21	.836
VETERANSDAY	-.01	.14	-.05	.958
WASHINGTONSBIRTHDAY	.04	.14	.27	.787
VALENTINESDAY	.03	.14	.24	.807

Table 5: Excluding days upon which there are zero sports prediction errors (i.e. dates when there are no fixtures): Linear regressions results of sentiment-based prediction errors against sports predictions errors, controlling for nuisance variables

Effect	***Estimate***	***SE***	***t***	***p***
Intercept	-.50	.27	-1.87	.063
Sport p.e.	.42	.09	4.57	**<0.001**
TUE	-.05	.22	-.22	.826
WED	.15	.22	.65	.514
THU	.33	.23	1.44	.150
FRI	.43	.22	1.98	**.049**
SAT	.36	.22	1.62	.106
SUN	.21	.22	.97	.335
FEB	.54	.31	1.77	.077
MAR	.26	.29	.90	.368
APR	.33	.29	1.14	.255
MAY	.53	.28	1.87	.062
JUN	.19	.28	.68	.498
JUL	.17	.29	.58	.560
AUG	.33	.28	1.17	.245
SEP	.42	.29	1.49	.138
OCT	.16	.35	.45	.652
NOV	.45	.33	1.37	.170
DEC	-.09	.33	-.28	.781
FIRST_OF_MONTH	.08	.34	.25	.804
FIFTEENTH_OF_MONTH	.14	.33	.43	.670
INDEPENDENCEDAY	.46	.98	.47	.642
CHRISTMASDAY	1.59	.99	1.60	.110
DAYAFTERCHRISTMAS	-.81	.99	-.81	.416
LABORDAY	.03	.98	.03	.979
EASTER	.79	.98	.81	.420
COLUMBUSDAY	.56	1.00	.56	.577
MEMORIALDAY	.10	.99	.10	.919
BIRTHDAYOFMARTINLU-THERKINGJR	.62	.99	.63	.532
WASHINGTONSBIRTHDAY	.76	.99	.77	.443
VALENTINESDAY	-.35	.99	-.35	.727

Table 6: Excluding days upon which there are zero sports prediction errors (i.e. dates when there are no fixtures): Linear regressions results of mean daily surprise scores against sports predictions errors, controlling for nuisance variables

Effect	***Estimate***	***SE***	***t***	***p***
Intercept	.12	.27	.45	.653
Sport p.e.	.17	.09	1.92	.056
TUE	-.10	.22	-.46	.648
WED	.20	.22	.91	.364
THU	.25	.23	1.09	.278
FRI	.24	.22	1.11	.268
SAT	-.02	.22	-.07	.944
SUN	-.30	.22	-1.39	.167
FEB	.68	.31	2.22	**.027**
MAR	-.22	.29	-.76	.450
APR	-.16	.29	-.54	.587
MAY	-.46	.28	-1.64	.102
JUN	-.37	.28	-1.30	.196
JUL	.32	.29	1.10	.274
AUG	-.17	.28	-.60	.551
SEP	-.35	.28	-1.23	.219
OCT	-.25	.35	-.72	.474
NOV	-.57	.33	-1.72	.086
DEC	-.55	.33	-1.70	.091
FIRST_OF_MONTH	-.02	.33	-.06	.950
FIFTEENTH_OF_MONTH	.19	.33	.58	.559
INDEPENDENCEDAY	.44	.98	.45	.651
CHRISTMASDAY	.21	.99	.21	.831
DAYAFTERCHRISTMAS	.66	.99	.67	.506
LABORDAY	.49	.98	.50	.615
EASTER	1.53	.98	1.57	.118
COLUMBUSDAY	-.53	1.00	-.53	.600
MEMORIALDAY	.03	.98	.03	.978
BIRTHDAYOFMARTINLU-THERKINGJR	-.68	.99	-.69	.490
WASHINGTONSBIRTHDAY	.65	.99	.66	.511
VALENTINESDAY	-.78	.99	-.79	.431

Table 7: Days upon which there are only negative sports prediction errors: Linear regressions results of sentiment-based prediction errors against sports predictions errors, controlling for nuisance variables

Effect	***Estimate***	***SE***	***t***	***p***
Intercept	-.46	.36	-1.29	.200
Sport p.e.	.35	.24	1.50	.137
TUE	.06	.28	.20	.839
WED	.51	.33	1.52	.131
THU	.28	.31	.90	.369
FRI	.47	.28	1.66	.100
SAT	.02	.30	.07	.944
SUN	-.08	.29	-.29	.771
FEB	.64	.42	1.52	.132
MAR	.47	.37	1.25	.215
APR	.44	.37	1.20	.234
MAY	.01	.37	.03	.973
JUN	.43	.34	1.24	.218
JUL	.09	.35	.26	.796
AUG	.36	.42	.84	.401
SEP	.38	.38	.99	.322
OCT	.16	.44	.37	.711
NOV	.64	.46	1.38	.171
DEC	.26	.44	.58	.561
FIRST_OF_MONTH	-.29	.44	-.68	.500
FIFTEENTH_OF_MONTH	-.29	.55	-.52	.607
INDEPENDENCEDAY	.12	.96	.12	.904
CHRISTMASDAY	1.05	.98	1.06	.290
LABORDAY	-.04	.96	-.04	.968
COLUMBUSDAY	.46	.98	.46	.643
MEMORIALDAY	.49	.97	.51	.613
BIRTHDAYOFMARTINLUTHERKINGJR	.54	.96	.57	.573
WASHINGTONSBIRTHDAY	.58	.97	.60	.551

Table 8: Days upon which there are only negative sports prediction errors: Linear regressions results of mean daily surprise scores against sports predictions errors, controlling for nuisance variables

Effect	***Estimate***	***SE***	***t***	***p***
Intercept	-.07	.37	-.20	.841
Sport p.e.	.32	.24	1.33	.187
TUE	.19	.29	.67	.503
WED	.55	.34	1.60	.113
THU	.52	.32	1.65	.101
FRI	.39	.29	1.33	.187
SAT	.20	.31	.66	.510
SUN	-.17	.29	-.57	.567
FEB	.98	.43	2.27	**.025**
MAR	-.08	.39	-.21	.830
APR	.03	.38	.08	.939
MAY	-.77	.38	-2.02	**.046**
JUN	-.02	.35	-.06	.954
JUL	.43	.36	1.20	.234
AUG	-.03	.44	-.06	.949
SEP	.05	.39	.14	.892
OCT	-.25	.46	-.56	.578
NOV	-.39	.48	-.81	.418
DEC	-.36	.45	-.78	.436
FIRST_OF_MONTH	.03	.45	.06	.954
FIFTEENTH_OF_MONTH	-.18	.57	-.31	.756
INDEPENDENCEDAY	.20	.99	.20	.838
CHRISTMASDAY	.02	1.01	.02	.987
LABORDAY	.44	.99	.44	.659
COLUMBUSDAY	-.21	1.01	-.21	.832
MEMORIALDAY	.72	.99	.72	.473
BIRTHDAYOFMARTINLUTHERKINGJR	-.41	.99	-.42	.675
WASHINGTONSBIRTHDAY	.64	1.00	.64	.525

Table 9: Days upon which there are only positive sports prediction errors: Linear regressions results of sentiment-based prediction errors against sports predictions errors, controlling for nuisance variables

Effect	***Estimate***	***SE***	***t***	***p***
Intercept	-1.08	.49	-2.20	**.030**
Sport p.e.	.74	.24	3.14	**.002**
TUE	.09	.38	.25	.806
WED	.29	.34	.86	.394
THU	.62	.36	1.73	.085
FRI	.45	.36	1.26	.209
SAT	.88	.35	2.48	**.014**
SUN	.64	.35	1.79	.076
FEB	.65	.48	1.35	.179
MAR	.23	.46	.49	.623
APR	.21	.48	.43	.669
MAY	.98	.45	2.16	**.033**
JUN	-.12	.49	-.25	.801
JUL	.65	.52	1.24	.216
AUG	.42	.45	.93	.356
SEP	.61	.46	1.32	.189
OCT	.36	.57	.62	.534
NOV	.47	.50	.95	.345
DEC	-.19	.52	-.37	.713
FIRST_OF_MONTH	.73	.54	1.36	.176
FIFTEENTH_OF_MONTH	.57	.42	1.35	.181
DAYAFTERCHRISTMAS	-.40	1.04	-.38	.701
EASTER	.95	1.04	.91	.363
VALENTINESDAY	-.30	1.04	-.29	.774

Table 10: Days upon which there are only positive sports prediction errors: Linear regressions results of mean daily surprise scores against sports predictions errors, controlling for nuisance variables

Effect	***Estimate***	***SE***	***t***	***p***
Intercept	.22	.50	.45	.651
Sport p.e.	.44	.24	1.86	.066
TUE	-.50	.38	-1.31	.194
WED	.07	.34	.20	.845
THU	-.05	.36	-.14	.887
FRI	.03	.36	.08	.935
SAT	-.23	.36	-.64	.521
SUN	-.54	.36	-1.50	.136
FEB	.49	.48	1.01	.313
MAR	-.31	.47	-.66	.510
APR	-.35	.48	-.73	.469
MAY	-.22	.46	-.47	.639
JUN	-.97	.50	-1.94	.055
JUL	.23	.53	.44	.660
AUG	-.16	.46	-.36	.721
SEP	-.68	.47	-1.45	.149
OCT	-.25	.58	-.44	.661
NOV	-.62	.51	-1.22	.227
DEC	-.54	.52	-1.02	.310
FIRST_OF_MONTH	.03	.55	.06	.956
FIFTEENTH_OF_MONTH	.52	.43	1.22	.226
DAYAFTERCHRISTMAS	.57	1.05	.54	.589
EASTER	1.76	1.05	1.67	.096
VALENTINESDAY	-.53	1.05	-.50	.616

Table 11: Sentiment prediction errors calculated excluding all negative sentiment tweets for days upon which there are only negative sports prediction errors: Linear regressions results of sentiment prediction errors against sports predictions errors, controlling for nuisance variables

Effect	***Estimate***	***SE***	***t***	***p***
Intercept	-.52	.33	-1.54	.126
Sport p.e.	.39	.22	1.78	.078
TUE	.11	.26	.40	.687
WED	.24	.31	.77	.444
THU	.10	.29	.36	.719
FRI	.37	.27	1.38	.170
SAT	.26	.28	.92	.361
SUN	.69	.27	2.56	**.012**
FEB	.51	.39	1.30	.197
MAR	.46	.35	1.32	.190
APR	.60	.35	1.73	.087
MAY	.07	.35	.21	.838
JUN	.25	.32	.77	.443
JUL	.13	.33	.40	.688
AUG	.15	.40	.37	.715
SEP	.52	.36	1.45	.150
OCT	.50	.42	1.21	.229
NOV	.68	.43	1.57	.119
DEC	.44	.41	1.07	.285
FIRST_OF_MONTH	-.28	.41	-.68	.497
FIFTEENTH_OF_MONTH	-.28	.52	-.55	.586
INDEPENDENCEDAY	.80	.90	.88	.381
CHRISTMASDAY	2.36	.92	2.56	**.012**
LABORDAY	-.28	.90	-.31	.754
COLUMBUSDAY	-.46	.92	-.50	.621
MEMORIALDAY	-1.14	.91	-1.26	.211
BIRTHDAYOFMARTINLU-THERKINGJR	-.23	.90	-.25	.800
WASHINGTONSBIRTHDAY	.85	.91	.93	.353

Table 12: Sentiment prediction errors calculated excluding all negative sentiment tweets for days upon which there are only positive sports prediction errors: Linear regressions results of sentiment prediction errors against sports predictions errors, controlling for nuisance variables

Effect	***Estimate***	***SE***	***t***	***p***
Intercept	-1.52	.51	-2.96	**.004**
Sport p.e.	1.07	.25	4.30	**<0.001**
TUE	.05	.40	.13	.898
WED	.20	.36	.56	.575
THU	.38	.38	1.02	.308
FRI	.37	.37	.99	.326
SAT	.92	.37	2.49	**.014**
SUN	.62	.37	1.67	.098
FEB	.91	.50	1.81	.073
MAR	.66	.49	1.36	.177
APR	.44	.50	.88	.378
MAY	.75	.48	1.57	.119
JUN	.58	.52	1.11	.268
JUL	.79	.55	1.44	.153
AUG	.78	.47	1.64	.104
SEP	1.13	.48	2.34	**.021**
OCT	1.01	.60	1.68	.095
NOV	.88	.53	1.67	.098
DEC	.03	.55	.05	.963
FIRST_OF_MONTH	.08	.57	.15	.881
FIF-TEENTH_OF_MONTH	.38	.45	.86	.393
DAYAFTERCHRIST-MAS	1.12	1.09	1.03	.306
EASTER	.91	1.09	.84	.404
VALENTINESDAY	-.44	1.10	-.40	.687