Prediction markets, social media and information efficiency

Article

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.
Published version at: http://onlinelibrary.wiley.com/doi/10.1111/kykl.12119/full
To link to this article DOI: http://dx.doi.org/10.1111/kykl.12119

Publisher: Wiley

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Prediction Markets, Social Media and Information Efficiency

Leighton Vaughan Williams*   J. James Reade†

May 17, 2016

Abstract

We consider the impact of breaking news on market prices. We measure activity on the micro-blogging platform Twitter surrounding a unique, newsworthy and identifiable event and investigate subsequent movements of betting prices on the prominent betting exchange, Betfair. The event we use is the *Bigotgate* scandal, which occurred during the 2010 UK General Election campaign. We use recent developments in time series econometric methods to identify and quantify movements in both Twitter activity and Betfair prices, and compare the timings of the two. We find that the response of market prices appears somewhat sluggish and is indicative of market inefficiency, as Betfair prices adjust with a delay, and there is evidence for post-news drift. This slow movement may be explained by the need for corroborating evidence via more traditional forms of media. Once important Tweeters begin to Tweet, including importantly breaking news Twitter feeds from traditional media sources, prices begin to move.

JEL Classification: G14, L83, D72.
Keywords: Information, Efficiency, Twitter, Social media, Betting.

1   Introduction

A central concern of economists is the information efficiency of markets; how quickly information is reflected in market prices. In this paper we investigate the breaking of a prominent news story and its subsequent impact on market prices. On 28 April 2010 in the midst of a general election campaign, unaware that he was still linked to a microphone, then UK Prime Minister Gordon Brown expressed his disgust with a voter he had just met who had mentioned immigration negatively, referring to her as a 'bigoted woman'. This event was swiftly publicised by the television company which picked up the microphone feed, and

---

*Professor of Economics and Finance, Nottingham Business School, Nottingham Trent University, Burton Street, Nottingham, NG1 4BU, United Kingdom. Email: leighton.vaughan-williams@ntu.ac.uk.

†Lecturer in Economics, Department of Economics, University of Reading, Reading, RG6 6AA, United Kingdom, and Programme for Economic Modelling (EMoD), Institute for New Economic Thinking at the Oxford Martin School, Oxford, United Kingdom. Email: j.j.reade@reading.ac.uk. Reade would like to thank the Open Society Foundations (OSF) and Oxford Martin School (OMS) for financial support in the writing of this paper.

became known as the *Bigotgate* scandal. The event was almost immediately recognised to be a gaffe: "An unintentional act or remark causing embarrassment to its originator; a blunder".[1] It quickly became the only news story of note related to the general election on that day, and is widely regarded as having had an impact on the eventual election outcome.

An efficient market reflects all relevant information in its price, assuming an absence of information processing costs. In this paper we use prices agreed on the commercial prediction market *Betfair*, and data from *Twitter*, an online social networking website, in order to trace the impact of the event on market prices. Twitter enables its users to share short messages (140 characters) publicly, and has become renowned for breaking news stories, while prediction markets allow participants to trade contracts contingent on the verifiable outcome of events such as political elections. As such we have detailed data on the original event, on the propagation of news relating to it, and on prices relevant to the wider context within which the event occurred, and we use this to investigate market efficiency. We find that the impact of the event on market prices appears somewhat sluggish, and consider a number of explanations for this.

In Section 2 we introduce the events of 28 April 2010, setting out a timeline, before in Section 3 we consider the theoretical context of information efficiency, introducing both Twitter and Betfair. In Section 4 we introduce our datasets, in Section 5 we analyse the Bigotgate event in terms of Twitter and Betfair activity, and in Section 6 we conclude.

2 Timeline of Events

It will be helpful to set out explicitly the timeline of events of April 28 2010, the day of the #bigotgate episode.[2] That morning the then UK Prime Minister Gordon Brown was campaigning in Rochdale in Greater Manchester, a seat thought to be marginal between Labour, Brown's party, and the Liberal Democrats. Whilst talking with local residents, Brown got into a conversation with a local lifelong Labour voter, Gillian Duffy. The conversation touched upon immigration, and the general tone appeared quite hostile from Duffy, who pressed Brown in particular on the perception that Eastern Europeans were taking lots of local jobs away from local people.

After the conversation was over, Brown was ushered into his car in order to make the drive to central Manchester for an interview with Jeremy Vine at the BBC. On getting in the car at about 10:45am (GMT), Brown's microphone remained switched on, and comments he made to party aides in the car could clearly be overheard. He referred to the meeting with Duffy as a disaster before referring to her as a 'bigot'. Sky News, who received the feed from the microphone, appear then to have made use of this information. Twenty minutes later, at 11:05am a Sky News journalist tweeted the first reference to 'bigot' that mentioned Brown, simply typing out the transcript of Brown's comments. At 11:08am the second tweet was sent, providing more context, and at 11:10am a third tweet provided yet more information about the encounter. From the Live

---

[1]See the Oxford English Dictionary's definition at http://www.oxforddictionaries.com/definition/english/gaffe.

[2]Live blog accounts can be found from the Guardian at http://goo.gl/Mb7Ifh, the Telegraph at http://goo.gl/56bWl6 and Channel 4 News at http://goo.gl/4Hcb1i.

Table 1: Timeline of events of #bigotgate episode on April 28 2010.

Time	Event
c.10:45am	Brown calls voter 'bigot'.
11am	Brown arrives in central Manchester for start of Jeremy Vine show on BBC Radio 2.
11:05am	First tweet referencing Brown event (transcript of comments).
11:08am	Second tweet (expressing shock at events).
11:09am	Event referred to in Channel 4 News Live Blog of day's election-related events.
11:10am	Third tweet (emphasises Brown comments candid).
11:11am	First retweet.
c.11:45am	Brown confronted live on air about calling voter bigot, live TV cameras record Brown's reaction, all news channels covering story live on air.
12 noon	Hourly news bulletins air on radio and TV stations, headlining Brown's gaffe.
1:57pm	Brown returns to Rochdale to meet with Duffy and apologise.
2:42pm	Brown emerges from Duffy home to apologise to gathered media.

Blogs mentioned (with hyperlinks) in footnote 2, at 11:09am Channel 4 News referred to the incident without revealing any details, at 11:15am the Telegraph noted that "News channels now reporting Brown has been caught on tape", and at 11:18am the Guardian introduced the event as a "big story". At this time, Brown was at the BBC's studios for the Jeremy Vine show on BBC Radio 2, which began at 12 noon (11am BST); at about 11:45am Vine confronted Brown live on air about his comments to Duffy. At that time, a camera had been placed in the radio studio and focussed in on Brown's reaction, which was to put his head in his hands. This footage and the radio interview was streamed live on the BBC's News 24 channel, and at this stage was being picked up by various news outlets.

Later in the afternoon, once Brown's Vine interview was over, he travelled back to Rochdale to meet with Duffy in her home. After a lengthy meeting, Brown left and was interviewed outside Duffy's home, expressing his regret and apologising for his actions. In Table 1 we summarise this timeline.

3   Theory, Information, Betfair and Twitter

At its simplest, the economic theory of perfect competition implies that all relevant information about a good is reflected in its market price. With a large number of buyers and sellers, none of whom is able to influence the market price, information regarding the relative scarcity or abundance of a good's supply or demand is reflected in the price and any movements that occur. This aspect of the functioning of a market is referred to as information efficiency. Weak-form market efficiency states that any market price reflects all information contained in past prices and price movements. Semi-strong form efficiency states that in addition markets reflect all publicly available information, and strong-form efficiency implies further that markets respond to even privately held information

(Bachelier, 1900; Fama, 1965, 1970, 1998). In order for a market to be efficient, Fama (1970) sets out sufficient conditions: (i) absence of transactions costs, (ii) all information costlessly available, and (iii) all market participants agree on the implication of current information for the market price. If markets are inefficient it suggests there exist some barriers to participants trading on available information, such as transactions costs or information processing costs.

A particular difficulty in appraising the information efficiency of markets is measuring the point at which news 'broke'; while on the stock market a firm may issue a profits warning and see a subsequent share price fall, it could be that via informal networks this news had broken earlier (see, for example, Fama et al., 1969; Ball and Brown, 1968; Asquith, 1983). Gil and Levitt (2007) and Croxson and Reade (2014) circumvent this problem by using sports bets traded on prediction markets, markets that allows tweeters to trade contracts whose value is contingent on some future event occurring. Sports bets have a number of attractive properties that both studies exploited; they are traded over a short time horizon (no time discounting), the events are repeated numerous times (allowing appraisal of the true value of contracts), and news breaks (goals in football matches) at very distinct (yet almost totally unpredictable) points in time, moves markets (on average 20%), and spreads very quickly. Gil and Levitt (2007) find evidence for post-news drift and hence information inefficiency, while Croxson and Reade (2014) find that prediction markets are semi-strong form efficient. However, Brown (2014) finds that mispricing exists as agents have information processing constraints which bind when multiple contracts exists for single events, and when numerous events take place concurrently.

Sports betting markets afford an opportunity to observe the reasonably clean breakage of information, most pertinently in the arrival of goals, which are highly unpredictable. In our case we use Twitter, the micro-blogging social networking website, to identify when news broke, and Betfair, a prediction market, to evaluate how quickly this news was reflected in market prices.

We choose a specific event during the election campaign that had great salience but also simplicity and interpretability. During an election campaign, then UK Prime Minister Gordon Brown was caught on microphone referring to a voter he had just met as a 'bigot'. The event has since become known as *bigotgate*, or in Twitter hashtag vernacular, #bigotgate. This was not the only notable event that occurred during the campaign; the 2010 General Election was also marked by the first live televised debates involving the leaders of the then three major parties in the UK. However, live debates, being 90-minute long discussions between the leaders, had multiple narratives emerging from them, while #bigotgate was much simpler in nature. As we will show with our data analysis, the debates and #bigotgate exerted notable influence on market prices and hence we do not consider #bigotgate to be any less of a major event than the debates.[3]

Very quickly in the course of the day's events, Brown's gaffe generated a term that hitherto had never been used, *bigotgate*, making the event very easy to trace.[4] From the moment news broke until election day the term 'bigotgate'

---

[3]The Wikipedia page on the 2010 General Election (http://en.wikipedia.org/wiki/United_Kingdom_general_election,_2010 as of 9 October 2014) devotes an entire section to the episode, suggesting its gravity.

[4]The first reference of the term #bigotgate on Twitter comes at 11:20am, approximately 35 minutes after the event occurred.
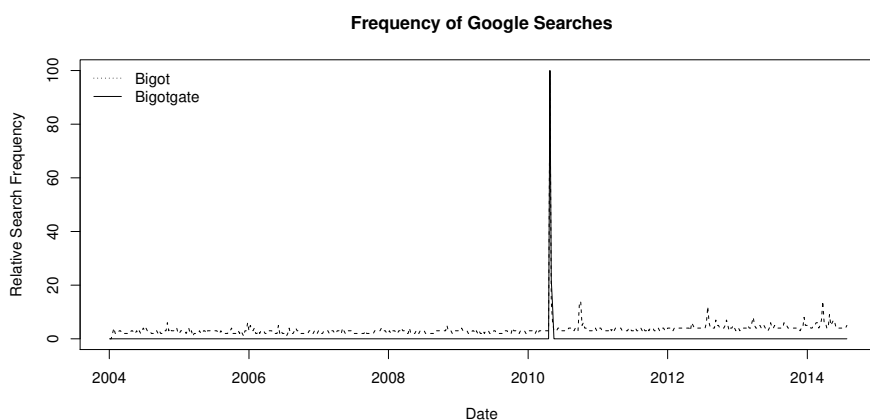
Frequency of Google Searches

Figure 1: Frequency of searches on Google for 'bigot' and 'bigotgate' since 2004; scale is relative to week with highest search volume over sample period. Source: Google Trends.

was mentioned 15,560 times on Twitter but before that, the term was never used. Figure 1 shows the frequency of searches on Google for the terms 'bigot' and 'bigotgate' over a much wider time interval, showing that bigotgate itself was not searched for at any point before the week containing April 28 2010, and after the election finished was not searched for again. Even the more general term 'bigot' was searched for 20 times more than usual in the week containing 28 April 2010.

It also became *the only* election-related story of any note on that day, which (Iannucci, 2010) notes, and a cursory glance at any online Live Blog from the day testifies to.[5][6] While undoubtedly politicians make gaffes on a frequent basis, some are more important than others, and the #bigotgate episode had such gravity that a BBC article for the 2015 UK General Election listed as number one in its list of "what not to do on the election trail" to be "don't leave your mic on".[7]

In this paper we use the uniqueness of the #bigotgate term to identify when news broke using Twitter, and then compare this to Betfair to learn about the nature of the market response to information.

Twitter is a micro-blogging website which enables any tweeter to broadcast very short (140-character) messages globally by simply typing the message and clicking *tweet*. All twitter users (henceforth *tweeters*) have their own 'twitter feed', a webpage where all the tweets sent by tweeters they *follow* appear the instant they are sent. Following another tweeter is a unilateral action; the

---

[5]See footnote 2 for links to Live Blogs from the *Guardian*, *Telegraph* and *Channel 4 News*. All three first reference the event at around 12:15 (note the hour difference between the timings in this paper, which are in Greenwich Mean Time (Universal Standard Time), and British Summer Time).

[6]We get a sense of this from our Twitter dataset, which is all tweets globally that mention the term 'bigot', and contains 78,147 tweets, of which 39,979 occur on April 28, the day of the #bigotgate events, and 39,603 of those after 11:05am, when the news was first broken on Twitter.

[7]See Simm (2015).

5

tweeter being followed need not reciprocate. Hence any tweeter who writes and sends a tweet ensures that at least momentarily that message appears on the twitter feeds of all the tweeters that follow them. As such, Twitter enables private information to be costlessly made public by tweeters for their particular purpose, be it information about themselves or others (Java et al., 2007; Naaman et al., 2010). It thus enables us to argue that information in our case study is costlessly available, as the theory of information efficiency requires. How widely that message is viewed depends to an extent on how important the tweeter is, in terms of their number of followers. While all tweets can be searched hence anybody can see the tweet, the seemingly more likely route for wider exposure is via *retweets*; other twitter tweeters can *retweet* a tweet, akin to forwarding a message, which is then seen by their followers, widening the audience of the message.

Already academic research using Twitter is flourishing. In addition to understanding the nature of tweeter usage (Java et al., 2007; Naaman et al., 2010), the credibility of information (Castillo et al., 2011), the spread of it (Lerman and Ghosh, 2010), its newsworthiness (Sankaranarayanan et al., 2009; Popescu and Pennacchiotti, 2010) and its geolocation (Sakaki et al., 2010) have all been studied, in particular developing machine learning techniques to automatically detect Twitter activity of interest. Information need not be neutral, and Thelwall et al. (2011) consider the sentiment of tweets surrounding popular events, finding negative sentiment to be proportional to the popularity of an event. Twitter sentiment can help predicting stock market indices (Bollen et al., 2011; Porshnev et al., 2013; Rao and Srivastava, 2012; Sprenger et al., 2013; Nardo et al., 2015; Sprenger et al., 2014), the Euro/US dollar exchange (Janetzko, 2014), the success of box-office movies (Treme and VanDerPloeg, 2014), the outcome of elections (Tumasjan et al., 2011) and even NFL scores (Sinha et al., 2013).

Feng and Yang (2011) consider the adoption of Twitter by politicians, while the use of Twitter has been investigated during political election campaigns in the Australia (Burgess and Bruns, 2012), Sweden (Larsson and Moe, 2012), Germany (Tumasjan et al., 2011), the UK (Anstead and O'Loughlin, 2014), and the US (Wang et al., 2012). Anstead and O'Loughlin (2014) discuss the extent to which extracting sentiment from tweets compares to political polling in attempting to reflect public opinion. To the extent that political polls are samples selected by pollsters whereas Twitter is populated by self-selected individuals, the two are likely to offer contrasting reflections on public opinion.[8]

Twitter also affords an opportunity to identify and measure the breaking and spread of new information (see, e.g. Lerman and Ghosh, 2010). One application of this is the resulting speed at which market prices respond to such news; how soon after tweets on a subject begin to appear do relevant market prices begin to adjust? The various forecasting studies already referred to above implicitly assume information contained in tweets is relevant for the updating of economic variables including prices (Bollen et al., 2011; Porshnev et al., 2013; Rao and Srivastava, 2012; Sprenger et al., 2013; Janetzko, 2014). For a more explicit investigation of information efficiency, what is required is

[8]Duffy (2014) noted that the weight of opinion expressed on Twitter for the Scottish independence referendum (#indyref) in September 2014 was dramatically different to the votes expressed at the polls. On Twitter, 73% of tweets were pro-independence, but in the referendum only 45% of voters voted for independence.

a suitable market price to observe. Meinusch and Tillmann (2015) consider the impact of beliefs regarding monetary policy expressed on Twitter on various macroeconomic variables associated with monetary policy. Their analysis, however, is aggregated to the daily level and although it provides evidence of the impact of Twitter on economic variables, is unable to provide information on the speed of the market response beyond that at a daily interval.

We focus on Betfair as the relevant market price in the #bigotgate episode. Betfair is a betting exchange, allowing exchange participants to trade contracts whose value is contingent on specified future events occurring. As such, by tracking the prices at which such contracts specific to the 2010 UK General Election trade, we have a market price that ought to react to the events of #bigotgate. It can be argued that Betfair prices should be efficient; agents trading financially on future events should rationally be informed about those events, and evidence suggests Betfair markets are semi-strong form efficient (Croxson and Reade, 2014). Betfair charges users a rate of commission of up to 5% of net winnings; as such, referring to the second sufficient condition for market efficiency, transactions costs do exist, but they are small. Additionally, other studies such as Croxson and Reade (2014) have concluded that Betfair is an efficient market despite these small transactions costs.

Semi-strong form efficiency suggests that the Betfair price ought to react instantaneously to any relevant information that becomes public. Hence we can use Twitter to detect the speed with which public information is reflected in prices on Betfair. If the market is strong-form efficient, and if the gaffe was relevant to election outcomes, the market ought to react to the events before any tweets are sent, since the moment the gaffe occurred it was private knowledge for a period of time. If the market is not efficient then we would expect sluggish updating. As we have established costless information availability, and trivial transactions costs, we thus focus on the third sufficient condition for examining market efficiency, namely that market participants agree on the impact of the news on the market price.

Hence it is important to identify when the news broke, what determines the credibility of new information, and what constitutes an instantaneous, or quick, reaction in market prices. It is well regarded that Twitter is a source of breaking news, both from official and unofficial sources.[9] The first Tweet sent relating to the events of #bigotgate could be defined to be the point that the news broke.

As only a small subset of Tweeters would have seen the original Tweet, and the relevance of the information may not have been initially clear, it is important that we analyse the nature of the breaking of information on Twitter. In particular, the third sufficient condition for market efficiency, that all participants agree about the impact of information on the market price, must be established. While it is true that both Twitter and Betfair participants are non-random samples from the general population, we argue that active Betfair market participants will be consumers of all available information and hence will either access their information from Twitter or from another credible news source more directly, which will happen at essentially the same time as the news breakage on Twitter.

---

[9]Elliott (2013) notes nine events broken on Twitter from the Royal Wedding to the Hudson river plane crash in New York, both of which occurred before #bigotgate.

As new information appears, market participants must judge the credibility of the information, and the relevance and hence impact of such information. The very first tweet on #bigotgate is simply the transcript of the audio recorded whilst Brown was unaware, having seemingly left the public stage and got into his car. Without that additional context, it would have been unclear that this was anything other than a random tweet involving the Prime Minister. The efficient markets hypothesis assumes costless information processing such that this stage ought to be instantaneous. As Twitter enables any user anywhere (with an internet connection) to post anything, it would seem likely that information processing on Twitter requires non-trivial effort due to the sheer volume of information.[10] Three minutes after the first tweet, a second tweet is posted which provides some context, namely that Brown had called someone he met on the campaign trail a "bigoted woman", and a further minute and a half later the third tweet makes it clear that Brown was unaware he was being recorded. Thus it is not until almost five minutes after the first tweet that an agent using solely Twitter information could have, in principle, learned of the full nature of the gaffe.

Two alternative theories are proposed with regard to information processing and Twitter (and technology more generally); the first is that Twitter reduces information asymmetries and ultimately will eradicate them (see, e.g. Tabarrok and Cowen, 2015), while the second is that Twitter increases the salience of information. Both ought to aid information efficiency as they decrease the cost of processing information for all. Nonetheless, considering again the first three tweets on #bigotgate, one would have to be following all three tweeters in order to see the tweets, else have the foresight to be continually searching for tweets including the term 'bigot' within them. While technology like Twitter must reduce information processing costs, it does not eradicate them. Furthermore, based on these initial few tweets, two are from the same news organisation that recorded the remarks, and the third by the head of a business journalism outlet; would this be sufficient for agents to view the news as credible and likely to affect election outcomes? Castillo et al. (2011) consider the credibility of information on Twitter, evaluating the impact of a number of characteristics of any tweet posted. They find that the certainty with which information is presented alongside the existence of corroborating evidence such as hyperlinks to relevant external material, are particularly important.

Reflecting secondly on the salience of information, there are multiple facets of this. Is it that media makes certain topics more salient, or that media responds to topics that are salient? In the political elections arena this matters, since it is often suggested that traditional media outlets such as newspapers and television stations exert influence by which topics they choose to make salient. Gelman and King (1993) suggest that voters become more "enlightened" by "the information they have learned during the campaign", suggesting a role for the media, but that equally they vote based on "ideology and party identification". Considering Twitter, a central question is the extent to which the sentiment expressed on the platform reflects that amongst the wider populace, and this question rests on whether Twitter adds salience to particular topics rather than reflecting their salience more generally. Tumasjan et al. (2010) suggest that

---

[10]It is essentially costless to post information on Twitter. Posting via text message, or SMS, enabled users to tweet before mobile internet was as prevalent as it has since become.

Twitter political activity accurately reflected more general activity outside of Twitter.

Applying specifically to our event, #bigotgate, this is an event that occurred offline, and had an effect on expected election outcomes, as we will show, via Betfair. It also generated Twitter activity in keeping with the magnitude of the event, and as such it could be argued that the salience of #bigotgate lay in the nature of the original event rather than the response on Twitter making it salient.

Considering the speed of updating, we require some metric to judge whether its speed. Croxson and Reade (2011) consider betting exchange prices for European Championship football matches in 2008 and find that they update within a minute whereas bookmakers are slower when goals occur. Croxson and Reade (2014) consider a wide range of matches, and consider the reaction 1–5 minutes after a goal has occurred, focussing in particular on the reaction after 1–2 minutes, and find no evidence of any post-news drift suggesting prices have already updated. The important aspect of their work when considering the depth of our betting markets in this study is that many of the markets they study are much thinner than our markets yet they find no evidence of more sluggish updating. Buraimo et al. (2008) find that prices on Betfair update within a minute.

4 Data

We introduce our two data sources separately; first Twitter (Section 4.1), and then Betfair (Section 4.2).

4.1 Twitter

We have all tweets from April 28th until May 7th that contain the word 'bigot'; this constitutes 78,147 tweets.[11] Our dataset contains the text of each tweet, the timestamp at which it was sent, the tweeter's self-described biography and location, and the method through which the tweet was sent (mobile phone, web browser, SMS).

While this doesn't give a particularly long pre-event baseline from which to observe the impact of #bigotgate, Figure 2 plots the frequency of tweets per minute on April 28 2010 and displays a clear difference in tweet volume before and after the #bigotgate event (just before 11am). In using the search term 'bigot', it is clear we will pick up, after 11:05am, the time of the first #bigotgate related tweet, a tiny amount of false positives; on average before 11:05 there is about a tweet every other minute, and between 7am and 11am there are just 45 tweets. In the hour following the initial tweet at 11:05 there are 3597 tweets posted, a rate of one per second.

At 11:05:49 Darren McCaffrey of Sky News tweeted "More Brown ... Just ridiclous [sic]...There will be ... Aagh, everything, she was a sort of bigoted woman .... said she used to be a ....". By 11:11:06 there were eight tweets, six of which were by Tweeters identifying themselves with Sky News, the last tweet being a retweet. Between 11:05 and 11:20 there were 329 tweets mentioning

---

[11]This dataset was purchased from *Texifter* via *Gnip*, and initially analysed using *Texifter*'s *DiscoverText* resource.

the term 'bigot'. The initial peak volume occurs between 11:37 and 11:38 when there are 240 tweets, and between then and 15:30 there was, on average, 92 tweets per minute mentioning the term 'bigot'. There is then a slow decline in the volume of tweets, albeit with a surge between 9:30pm and 10pm, likely coinciding with topical news programmes and late night news flashes. Our dataset does run through until the day following the election, but as our interest here is the breaking of information, we focus only on the day of April 28.

In our dataset there are 38,710 individual tweeters, hence on average two tweets sent per tweeter. However, only about a third of these tweeters (12,592) actually tweet more than once in our entire dataset, and only 2% tweet more than ten times, with one particular tweeter tweeting over a thousand times.

Our interest will be studying characteristics of tweets after the #bigotgate events, in order to understand more about the nature of the market reaction to information. To this end, we extract information from each tweet and aggregate to the minute level. Castillo et al. (2011) list a range of characteristics of tweets that they hypothesise may matter for determining the credibility of information contained within a tweet, such as the presence of a hyperlink, the number of other users mentioned, whether it is a retweet, the proportion of text in upper case, and the sentiment of the text (positive or negative). They also refer to the credibility of the tweeter, which is established, amongst other things, by the number of followers they have. Our dataset does not contain information on the number of followers a tweeter has, nor how many times that tweet was retweeted or favourited (generally regarded as a measure of the credibility of a tweet). Nonetheless, as we will detail, we are able to extract many of the measures Castillo et al. (2011) use, and hence form a clear picture of the nature of the breaking of information on Twitter.

Just less than half of our tweets (37,497) contain the '@' symbol, which indicates that a tweeter has tweeted a message specifically mentioning another tweeter. There are two categories of tweets mentioning another tweeter; one is messages directed specifically at another tweeter, which we refer to as *directed tweets*, and the other is *retweets*. Considering the latter, just under a third of tweets (23,639) contain the abbreviation 'RT', which is short for *retweet*, and indicates that a tweeter saw fit to re-post a tweet sent by another tweeter. While the distinct interpretation of a retweet remains a matter for discussion (see, for example Boyd et al., 2010; Macskassy and Michelson, 2011; Nagarajan et al., 2010), fundamentally the retweet represents propagation of information, as that tweet then appears on the twitter feeds of all those that follow the retweeter.

Of the former category of tweets using the '@' symbol, directed tweets, there are 14,303 that are not retweets and hence tweets directed at one or more other tweeters; such messages may form part of conversations, rather than simple broadcast messages. It may be that many of these 14,303 messages are only attempted conversations, rather than actual conversations; it may be that the tweeter mentioned in a tweet never responds to the initial tweet(s). Greetham and Ward (2012) study the nature of conversations on Twitter, finding that two-person conversations are relatively balanced, whereas those involving more than two tweeters tend to be dominated by one or two participants.

Conversations are generally considered distinct from information sharing as uses of Twitter, being more informal in nature (Kwak et al., 2010). Conversation tweets must be identified within any dataset, but as only 2% of tweeters tweet more than ten times in our entire dataset, this implies a very small number
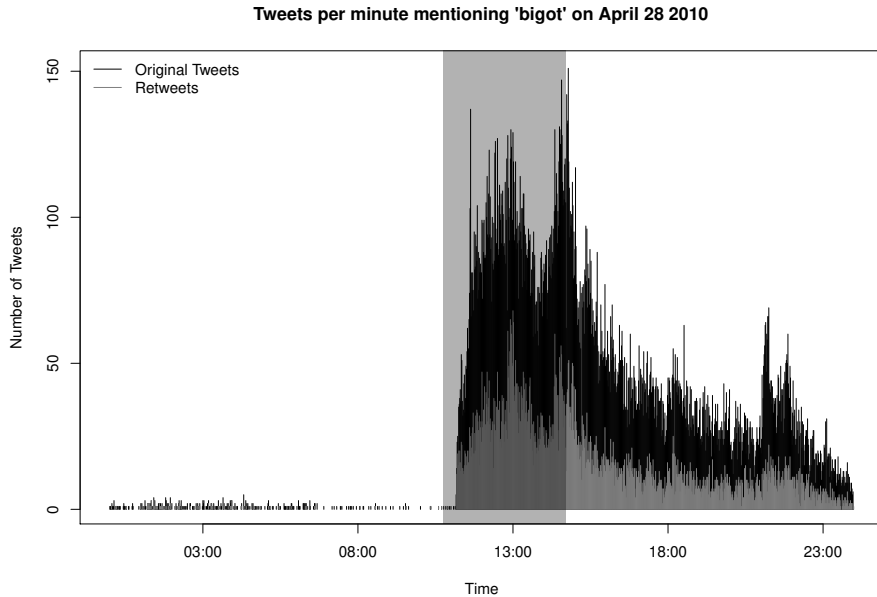
Tweets per minute mentioning 'bigot' on April 28 2010

Figure 2: Number of tweets per minute mentioning the word 'bigot' on April 28 2010. Source: Twitter.

of conversations and suggests that our dataset better reflects an information-sharing episode. If we define a conversation to involve the same group tweeters (as in, no additional participants enter the group) and involve at least two tweets, then we find there are just 2,004 tweets that form part of a conversation, only 1240 groups of tweeters producing those tweets, and 2113 individual tweeters participating. This, however, could simply be one tweeter tweeting another repeatedly; requiring that more than one tweeter sends tweets in a group reduces the number of conversations to 569, with only 816, or 2% of all tweeters taking part in these conversations. Of these, 372 and 124 respectively occur in the five hours after 11:05am on April 28.

4.2 Betfair

We have prices for four Betfair markets related to the 2010 general election. The first was the market for a Labour majority, with contracts stipulating the buyer pay the seller if Labour won a majority in the election; a similar market existed for the Conservatives (Tories), paying out if the Tories won the most seats in the election. A further market existed for any other party to win a majority (although realistically, the only party this market covers is the Liberal Democrats, at the time the third largest party in the UK), and a final market for no overall majority — if no party won enough seats to command a majority in the House of Commons. Over all four markets, just over £6m was traded.[12]

---

[12]Of that £6m, £4.75m was traded during the statutory campaigning period between April 6 and May 6, £3.66m between the day of #bigotgate and election day, and £101,571.10 on the day of #bigotgate itself.
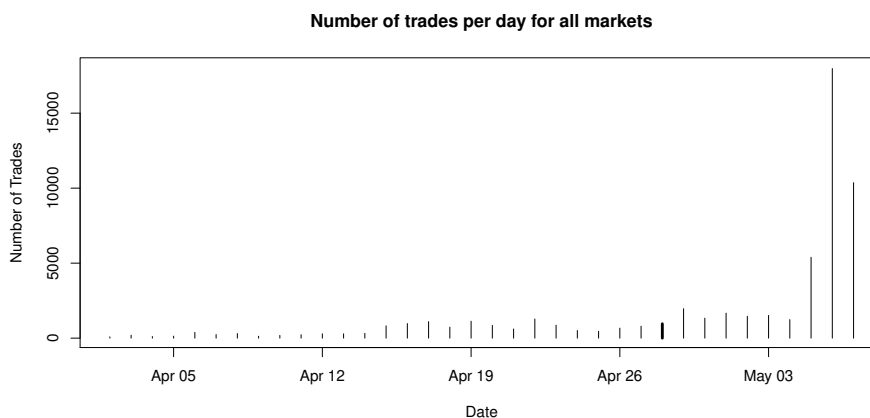
Figure 3: Number of trades per day, broken down by Betfair market.

Before considering the data further, it may be helpful to provide some context on the UK electoral system, which is a First Past the Post system. Candidates participate in elections in 650 individual constituencies around the country to be the representative in Parliament for that area. The party that gains an overall majority, hence wins 326 of the seats, has the opportunity to form a government. An important characteristic of this system is that a party can win 326 seats on a relatively small share of the overall vote; for example in the May 2015 General Election, the Conservative Party won 330 seats and a majority, yet only won 36.9% of the vote, while the Scottish National Party won 56 seats despite winning 4.7% of the vote, and the United Kingdom Independence Party won 1 seat despite winning 12.7% of the vote. If no party wins 326 seats, there is *no overall majority*. In this situation, the largest party is usually afforded the opportunity to form a government, but may require either a formal coalition, or some other kind of arrangement with some other parties in order to form a majority.

Figure 3 shows that election day, May 6, was the day with the heaviest trading, followed by the days either side of it.[13] The day of the bigotgate scandal, April 28, is not particularly outstanding; more trading occurs than on the previous day, reflecting the salience of the event, but less than the following day, which was distinct for being the day of the final televised debate between the party leaders. The two most heavily traded markets (not shown) are for a Tory majority and no overall majority, reflected by their status as the most likely two events, shown in Figure 4 which plots the implied probability since mid-2009.[14] From mid-2008 onwards, these two events take up more than 80% of the probability space for the election.

Focussing on the final few weeks of the campaign, the likelihood of a Labour majority was viewed as highly unlikely according to the market. There is no necessary implication that #bigotgate had minimal impact simply because

---

[13] As polling stations close at 10pm and counting begins immediately, the result is not known until early morning of the day after election day.

[14] Betfair prices are quoted in decimal odds. The probability of outcome implied by those odds is given by the reciprocal of them. We do not correct for the overround in our plots.

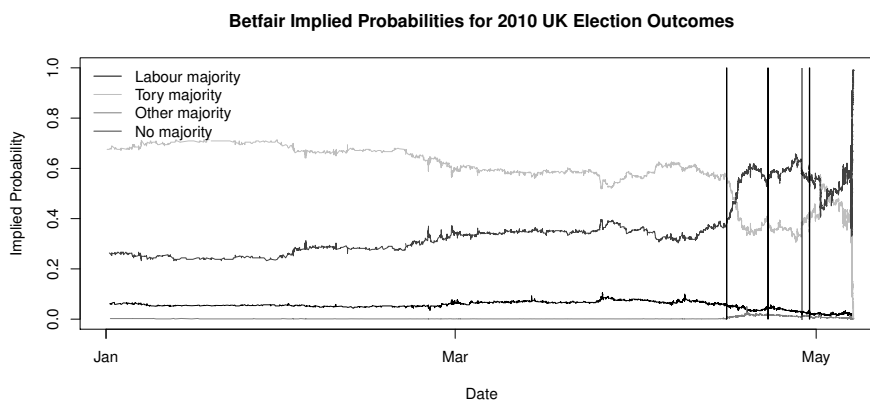**Betfair Implied Probabilities for 2010 UK Election Outcomes**

Figure 4: Implied probability of particular election outcomes throughout existence of markets. Marked on as vertical bars are the times of the three live election debates and the shaded area is #bigotgate day.

Brown and Labour were unlikely to win; while a Labour majority was unlikely, it remained plausible that Labour would win enough seats to ensure there would be No Overall Majority.[15] If #bigotgate reduced the number of seats Labour would win, this increased the likelihood of a Conservative majority and hence can be argued to have been a significant event.

One way to measure whether the event was significant is to consider its impact on the Betfair markets. Assuming the market is at least weak-form efficient, then it follows a random walk and the impact of any price movement is permanent on the market price. Hence if a clear shift occurred around the time of #bigotgate, this would indicate that the event was significant. Considering Figure 4, the first election debate is shown as the first vertical line, and is notable for apparently setting in motion a downward trend in the Conservative majority probability, and increasing the likelihood of no overall majority (green). The new level that prices settled at, a few days after the first debate, remained in place through the second debate (second vertical line) until April 28, the day of Bigotgate, which is the shaded area. There is an immediate fall in the implied probability of no majority, and an equally large increase in the implied probability of a Conservative majority, appearing to occur instantaneously. This then appears to be followed by a period of drift which continues through the final debate the following day (third vertical line), albeit with a blip in both prices around that time. Hence based on an ocular analysis, we might conclude that Bigotgate was a significant event, as it appears to have prompted price movements.

Zooming in on the day further, in Figure 4.2 we plot individual Betfair transactions on the day of the Bigotgate scandal, starting on the left at 8am that morning, and ending on the right hand side at 10pm. Each point is an agreed bet, with the horizontal axis being the time the bet was agreed at, and

---

[15]In that situation Labour might have made a coalition with the third party, the Liberal Democrats, thus keeping the Conservatives from office.

13

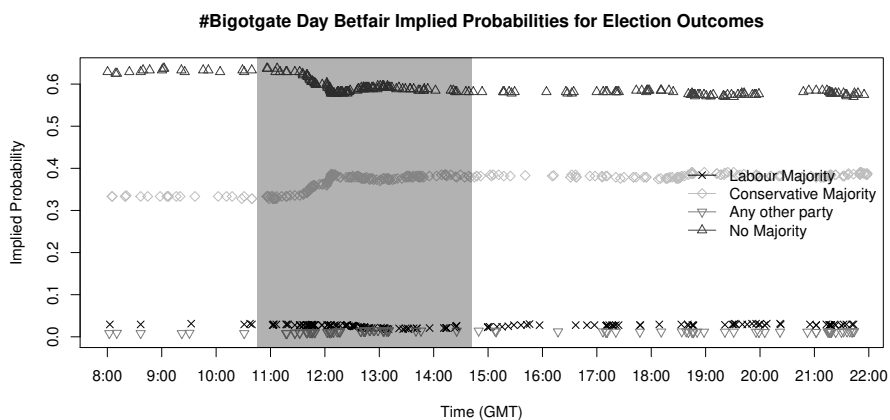**#Bigotgate Day Betfair Implied Probabilities for Election Outcomes**

Figure 5: Betfair implied probabilities for Labour majority, Conservative majority, other majority and no majority between 8am and 10pm April 28 2010. Each point represents a trade at a particular point in time (horizontal axis) at a particular implied probability (vertical axis). Marked on with vertical lines and shaded areas are important timeline events from the day (see Table 1 for more details).

the vertical axis being the implied probability the bet was agreed at. The green points are those bets agreed on no majority, the blue points for a Conservative majority, red for Labour and yellow for other majority. We additionally plot a number of vertical bars for particular events outlined in the timeline in Table 1; the first line is at 10:45, around the time Brown was recorded. The next three lines are the first three tweets referencing the recording, between 11:05 and 11:10. The subsequent two lines are 11:45, the time when Jeremy Vine confronted Brown live on BBC radio and television about his comments, and finally at 12 noon, when hourly news bulletins occurred on radio and television. The implied probabilities for a Conservative majority and no overall majority show clear movement after the events have become public knowledge, and appear to remain at their new levels after the event without displaying any drift. In the next section we examine the statistical significance and timeliness of that movement.

5 Analysis

Our analysis proceeds in a number of steps. First, we must establish that indeed #bigotgate was a newsworthy event and second that no other political event at the time could have caused the price movements observed. Following this, we consider the speed of the market reaction, and the nature of it.

5.1 Methods of Analysis

Although we consider only one event within one election campaign, we can nonetheless ensure our analysis is rigorous by employing time series statistical methods. We aggregate relevant data series from our two datasets (e.g. prices, number of tweets) to the minute-by-minute level, collating information into minute-long intervals. Although this represents a loss of information, it nonetheless creates regular time series that can be analysed using standard time-series methods, allowing us to make statements regarding the statistical significance of changes we observe in Twitter activity and Betfair prices.

We are considering the presence of shift changes in time series data: permanent shifts in prices, and temporary shifts in various types of activity on Twitter. Significant shifts and trend changes in Betfair and Twitter activity show the impact of the #Bigotgate event, and also its timing.

As such, time series methods related to detecting outliers and structural breaks can be employed (see, for example Bai, 1997; Bai and Perron, 2003). We employ a particularly flexible and powerful method, namely *indicator saturation* Hendry et al. (2008, 2013); Castle et al. (2015).[16] Indicator saturation adds indicator variables for outliers, step shifts and trend shifts, and selects between them using the general-to-specific model selection algorithms associated most commonly with David F. Hendry (Hoover and Perez, 1999; Hendry and Krolzig, 1999; Doornik, 2009).[17] The method includes indicator variables for all observations, and hence must do so in batches, carrying out general-to-specific (see, for example Campos et al., 2005) model selection over each included batch before selecting over the union of all retained indicators from the batches.[18] The method imposes no restriction on the number of outliers, shifts or trend breaks that will be found, unlike other common methods such as that proposed by Bai and Perron (2003). We enter impulse indicator variables to detect outliers, step indicator variables to detect step shifts, and trend functions to detect more gradual shifts or drift in variables.[19][20]

In Figure 6 we plot some examples of how indicator saturation works. Each example is designed to correspond to a particular case when considering information efficiency, but of course the method applies much more generally to observed shifts in general data series. The generated data are plotted in grey, with the set of indicator variables found using indicator saturation plotted in black. Henceforth we refer to this set of indicator variables plotted in this manner as the *coefficient path*. In the simplest case, top left, we have a data

---

[16]Indicator saturation is commonly abbreviated to IS. However, the current geopolitical climate renders such acronyms highly unpalatable.

[17]Indicator variables (also known as dummy variables), are variables in an analysis that are constructed by the statistician to satisfy particular design requirement. Usually they take the value 1 for a subset of observations, and zero otherwise, although time trends are also commonly used. Trend shifts are what Pretis et al. (2014b) refers to as *design functions*, showing the flexibility of the procedure.

[18]We use the *gets* R package of Pretis et al. (2014a) to carry out indicator saturation.

[19]Step indicator variables are variables which are zero up to some point, and take the value one for all observations after that point, and trend indicator variables are variables which are zero up to some point, and increase by one for each observation after that point.

[20]Temporary shifts, as we might expect in some Twitter behaviours around the #bigotgate events, can be represented by considering the difference of two step indicator variables, since the resulting variable takes the value one for only a number of observations before returning to zero.
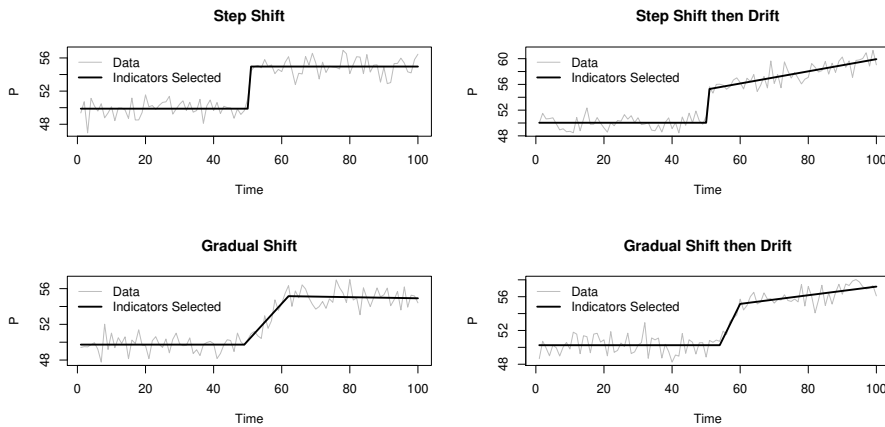
Figure 6: Generated examples of indicator saturation; each case corresponds to a fictitious price series with particular characteristics, and the original data (solid line) plus the discovered steps or trends (dashed line) are plotted.

series with a step shift after 50 observations, representing semi-strong form information efficiency, then in the top right panel the case of a step shift at 50 observations followed by an upward drift, representing post-news drift and semi-strong form information inefficiency, is plotted. The bottom two cases also refer to inefficient price adjustments, with the bottom left representing a gradual shift, and the bottom right panel a gradual shift followed by upward drift.

In all three cases, the coefficient paths found are almost exactly what were specified to generate the data plotted, and are hence indicative of the hypothetical kinds of shifts in data series we are interested in detecting both in our Betfair and Twitter data series. The statistical power of indicator saturation has been shown both theoretically and via simulation (see Johansen and Nielsen, 2009; Doornik, 2009) to be high, giving confidence that the procedure can pick up significant changes in the properties of time series. Considering size, or the likelihood of false positives, on average αT indicator dummies are incorrectly retained for a significance level of α and a sample size T. As such, the selection of α is important.

We run two variants of indicator saturation:

1. Week-long saturation with step shifts and ultra-low significance level: We run this on our Betfair series for the week around the #bigotgate event in order to establish the *existence* of a break, hence the *newsworthiness* of the event. A newsworthy event should generate a shift in market prices. We run this saturation with an ultra-low α level of $\alpha = 1 \times 10^{-30}$, in order that the likelihood of spuriously retaining a shift is absolutely negligible at $\alpha T = 1.008 \times 10^{-26}$.

2. Five-hour (10:30–15:30) saturation with step and trend shifts and $1/T$ significance level: We run this on both our Betfair and Twitter series in order to determine the *timing* of breaks that occur. As highlighted

16

in Figure 6, if a break occurs gradually then simply using step indicator saturation may provide a inaccurate estimate of the timing of the break.

We note a computational problem that arose in the running of indicator saturation on large datasets, specifically with trend indicators. With large values of T, the matrix of explanatory variables populated by trend indicators becomes essentially collinear, causing the procedure to fail. In order to circumvent this problem we carry out trend indicator saturation but include trend dummies at a lower frequency than every minute. For the five-hour saturation we included step dummies for every five minutes, which need not impact the timing of shifts, since we still include step dummies for every observation, hence any steeper shift can be proxied by a step shift.

5.2   The Newsworthiness of #bigotgate

The information contained in a tweet need not provoke a market response. If the information is not credible it should not be reflected in prices, and if the information is insufficiently important, or salient, it should not spur any market movement. From the argument for weak-form market efficiency, a precondition for this analysis is that market prices follow a random walk: that relevant historical information remains impounded in market prices. This implies that Betfair prices must follow a random walk, and hence we first test for this in Table 2. The table has rows for each market, and columns for two different tests. On the left panel, the test statistic and p-value for the Augmented Dickey-Fuller test is presented (Said and Dickey, 1984), and on the right panel the KPSS test (Kwiatkowski et al., 1992). The two tests differ in presentational terms in their null hypothesis; the the ADF test the null hypothesis is a random walk, or non-stationary series, but for the KPSS test the null hypothesis is that the data is a stationary series.[21] For Conservative, Labour and No Majority, the null hypothesis of non-stationarity is not rejected, and very comfortably. It is only in the case of Other Majority that the decision is borderline. However, the KPSS test provides evidence that all four prices are non-stationary, as the null hypothesis of stationarity is heavily rejected in all four cases. We thus conclude that indeed the four prices follow a random walk and the market is weak-form efficient. Consequently we can investigate whether permanent step shifts occur around the time of the #bigotgate event, as such breaks imply that the event was salient and newsworthy.

In order to determine whether #bigotgate was newsworthy, we apply indicator saturation to Betfair market prices for the week surrounding April 28; if market prices exhibit a clear and sizeable shift change that remains, we can conclude #bigotgate is newsworthy. We have already noted that by all accounts, #bigotgate was the only news story pertaining to election outcomes on April 28, so any substantial movements in prices we can reasonably conclude were caused by the #bigotgate events; we use indicator saturation to determine whether any significant shifts did occur.

We use week-long saturation to determine salience of the event. In Figure 7 we plot for each market the coefficient path retained on the same set of axes, scaled to zero at 10:45am on April 28, the moment Brown referred to Duffy

---

[21]It is generally noted that unit-root tests have low power, particularly for marginal cases, and as such it is common to present multiple unit root tests.

Table 2: Unit root testing of all four Betfair prices (aggregated to minute-by-minute level).

	ADF Test		KPSS Test	
	Test Statistic	p-value	Test Statistic	p-value
Conservative Majority	-1.243	0.900	12.228	0.010
Labour Majority	-2.157	0.512	42.246	0.010
Other Party Majority	-3.640	0.028	9.193	0.010
No Majority	-1.792	0.667	43.886	0.010

as a "bigotted women".[22] The shaded area corresponds to the time between 10:45am, approximately the time of the #bigotgate events, and 2:42pm, the time when Brown emerges from Duffy's house to face the assembled media; see Table 1 for the timeline of events on that day.

Considering the existence of a break and hence the newsworthiness of the event, clear shifts occur for the Conservative and No majority markets not long after the #bigotgate events occur. Table 3 contains the timings of all the breaks; in the Conservative market, the first shift occurs at 11:31am, with a second shift at 12:01pm. The total size of the shift is 4.6 percentage points. In the No majority market, the first shift occurs at 11:32am with a subsequent shift 12:31pm, and the total shift is 5.4 percentage points. Although these shifts are not the largest in the seven-day interval (a larger shift occurs on May 1), they are nonetheless proportionately large movements, and very significant (t-statistics of between 27.4 and 39.0). The shifts in the Labour and Other majority markets are smaller in absolute terms, yet relatively large and statistically significant. In the Labour market the shifts occur at 11:30am and 12:32pm and total 0.98 percentage points, and in the Other majority market at 11:32am and 12:31pm, totalling 0.5 percentage points. Given that at the point of the #bigotgate events, these two probabilities were at 3% and 0.9% respectively, these are proportionately large price movements, the largest observed over this seven day period. This is evidence regarding the newsworthiness of the event.

We report the results for each market and our two indicator saturation variations in Table 3; with the exception of the Labour price, the picture appears to be that Betfair prices react some time after the #bigotgate events occur and have been made public, and the reaction is gradual rather than instantaneous, with the steeper price movements occurring between 11:25am and 12:25pm. This would seem to be *prima facie* evidence suggestive of post-news drift, suggesting markets are not semi-strong form efficient. Nonetheless, we should consider the dynamics of information as contained on Twitter before making such conclusions.

---

[22]We include further tabular and graphical output from indicator saturation applied to the four prices in the appendix (see Table 5 and Figure 16), and invite interested readers to make use of the R code we make available with this paper, but we present in Figure 7 the coefficient path discovered for each market, and the size of their impact on the implied probability.
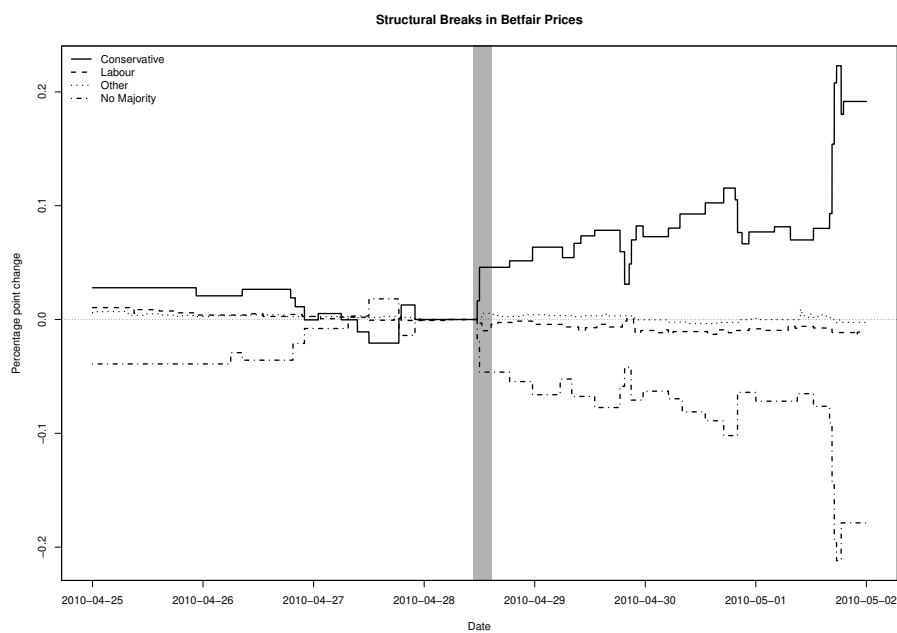
Figure 7: Coefficient paths for all four Betfair markets (Conservative, Labour, other and no majority). The plot is of the mean price as dictated by step shifts located using indicator saturation. The shaded area corresponds to the events on the day, starting at 10:45am, approximately the time of the #bigotgate events, and ending at 2:42pm, the time of Brown's re-emergence from Duffy's house to face the assembled media; see Table 1 for timeline of events on that day.

Variant	Conservative	No	Labour	Other
7-day	11:31	11:32	11:30	11:32
5-hour	11:35	11:35	10:25/12:31	11:25

Table 3: Different times for time of first break according to different sample sizes and significance levels.

5.3 The Speed of Market Reaction

We have established in the previous section that there was a statistically significant market reaction to the #bigotgate events. Next we need to establish with more accuracy the timing of the movements in Betfair prices. We again employ indicator saturation; this time we use step and trend indicators over a five hour period that includes all the events of the day, and a conventional significance level as we seek to find the right number of break terms.

We plot the coefficient paths in Figure 8, following the same format as in Figure 7. A more refined picture is presented here, however, as we can capture drifting prices rather than simple step shifts. In general there are small drifts in either direction shown by prices throughout, but the large shifts after 10:45 are quite distinct, and occur at around 11:45am. In particular, for the Conservative Majority, the trend shifts up at 11:35 from about 0.01 percentage points per minute to 0.1 percentage points a minute before a step shift occurs at 11:44, and another at 12:04. For No Overall Majority, the first three trend shift changes appear to be cancelling each other out and hence arguably spurious, before at 11:35 a steep downward trend of 0.2 percentage points per minute begins, which increases to 0.35 percentage points per minute at 12:05 before beginning a correction at 12:10.

In the Other Majority market, a relatively steep upward trend begins at 11:25, and continues until 12:55 when the trend turns negative and an apparent correction begins.

The Labour Majority market is the hardest one to understand, as a general negative trend starts before the #bigotgate events occur, at about 10:25am, and continues with step shifts at 10:30, 11 and 11:30 amounting to a downward shift of 0.4 percentage points by 12:31, at which point there is a sharp downward fall in the price down to 1.2 percentage points below its pre-#bigotgate level at 13:15. After this point, the price quickly recovers to roughly its pre-crash negative trend level. As this price was at about 3% before the events began, these are considerable price movements, but also reflect a relatively illiquid market.

Hence the market reaction appears to be at about 11:35am, about 50 minutes after the events themselves, about 30 minutes after the initial tweet, but around about the time of the first peak in Twitter activity we noted earlier, between 11:37 and 11:38am. About 1000 tweets had been sent by 11:35am referring to the event.

Between 10:45 and 11:35 there were 90 trades conducted on Betfair, spread reasonably evenly between the four markets.[23]

We now apply indicator saturation to various measures of Twitter activity

---

[23]There are 28 trades on the other majority market, 25 for no overall majority, 22 for Labour majority and 15 for Conservative majority.
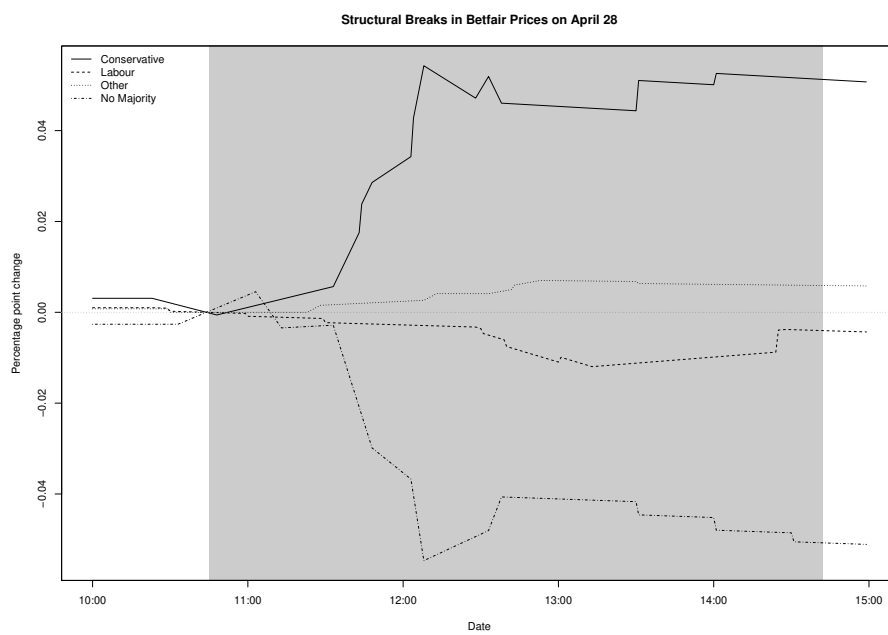
Figure 8: Coefficient paths for all four Betfair markets (Conservative, Labour, other and no majority). The plot is of the mean price as dictated by step shifts located using indicator saturation. The shaded area corresponds to the events on the day, starting at 10:45am, approximately the time of the #bigotgate events, and ending at 2:42pm, the time of Brown's re-emergence from Duffy's house to face the assembled media; see Table 1 for timeline of events on that day.

in order to time the Twitter response to the events in relation to the Betfair, market, response. We again use a five-hour interval as with the second set of Betfair saturation cases. In Figure 9 we plot the coefficient path for the number of tweets per minute in total on the bottom panel, with the coefficient paths for the Betfair markets on the top panel. We plot the negative of the No Majority and Labour Majority prices in order that they lie above zero. The first shift up in Twitter activity occurs as a trend shift at 11:09am and for the following ten minutes there is an extra four tweets per minute mentioning 'bigot'. This trend then slows to only about an extra 0.5 tweets per minute at 11:19am before a shift up at 11:37am of 68 tweets per minute, before a shift back down at 11:39am. In these two minutes there is around 120 tweets each minute, the high point of Twitter activity.

As such, the significant changes in the time profile of all Twitter activity and the significant price movements do not appear particularly well aligned; Betfair price movements in the two most important markets (Conservative and No Majority) occur at around 11:35, yet this is 26 minutes after the original upward trend materialises in Twitter activity, and 17 minutes after the last Live Blog to report the event (see footnote 2). Peak Twitter activity at 11:37–11:38 thus comes after the initial price movements, albeit at the same time that much of the price movement is occurring. The Conservative and No Majority prices stabilise just after 12 noon, while Twitter activity is still increasing more moderately at around 0.5 tweets per minute until 12:39pm.

Thus market prices appear to start moving considerably after Twitter activity has significantly increased, and seem to bear little relation to peaks in Twitter activity either. Information regarding the incident appears on Twitter more than 20 minutes before any price reaction takes place. As already established, the information was costlessly available, with minimal transactions costs involves, meaning that agreement on the implications of the information for market prices must be why adjustment didn't occur earlier. We now consider a number of characteristics of tweets in order to shed more light on this. We firstly consider characteristics of tweeters, then characteristics of tweets.

5.4 Characteristics of Tweeters

In this section we distinguish between types of tweeters; as with any network, some agents have more influence than others. Tweets from some users will carry more authority than tweets by others, and we attempt to measure this. Unfortunately our dataset does not carry detailed information about tweeters, meaning we cannot use the number of followers they have to determine the level of influence of a tweeter. An alternative and simple way of measuring the authority of a tweeter is the number of times they are mentioned by others; as a subset of this, we get how often that tweeter is retweeted, since a retweet requires a mention of the tweeter whose tweet is being retweeted. This measure can be seen as a very simple implementation of Google's PageRank algorithm, which ranks webpages based on how many other pages link to that page.

Our dataset only pertains to tweets mentioning the term 'bigot' in the week prior to the 2010 general election and as such we cannot use wider information about the influence of tweeters. This need not be a disadvantage; it means that our measure better reflects those actively engaged in the discussion surrounding the events. We consider two mention-related measures of importance; the first
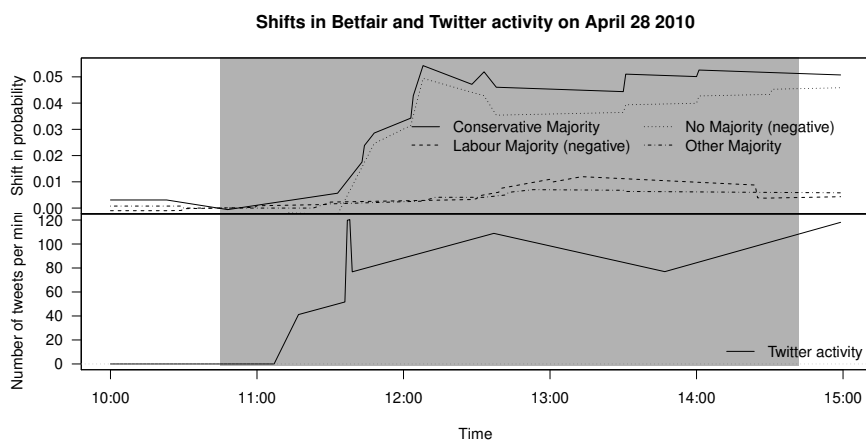
Shifts in Betfair and Twitter activity on April 28 2010

Figure 9: Coefficient paths for all four Betfair markets (top panel) plus Twitter activity (number of tweets per minute containing the word 'bigot', bottom panel). The two negative price reactions (Labour and No majority) are plotted in mirror image around the horizontal axis.

is simply the number of mentions, and the second is a discrete dummy taking the value 1 if a tweeter has more than two mentions elsewhere.[24] We consider the average number of mentions tweeters have each minute during our sample, and the proportion of tweeters with at least two mentions.

We plot the resulting coefficient paths in Figure 10, in the same format as Figure 9. The significant increases in the number of tweets by more influential tweeters are as follows: at 11:25am the average number of retweets of the author of a tweet sent increases from zero to 0.51, and at 11:35am the proportion of tweeters with at least two retweets amongst authors of tweets increases from zero to 2.1. These are both much more aligned with the observed timings of price movements, and hence suggest that it is tweets from authoritative tweeters that hold more influence.

Our data also includes a biography that users can submit to attach to their profile. This biography is easily found when viewing tweets, and hence plausibly forms part of the information set a viewer might use to establish the credibility of the information in a tweet. Castillo et al. (2011) consider the simple existence of a user biography as potentially important in determining the credibility of a tweeter and hence of information. In addition to the existence of a bio, we consider from those bios three types of (self-identified) tweeters: news organisations providing "breaking news" Twitter feeds, journalists, and bloggers. Considering each in turn, as Twitter is viewed as a source of breaking news, many Twitter accounts identify themselves as a source of such breaking news; journalists as a result are likely to be users of Twitter as they seek to determine the credibility of breaking news stories for their employers, and finally particularly in the political scene, blogs are viewed as a credible source of information (Johnson et al., 2007).

We consider per minute the proportion of tweets from these three groups

---

[24]From our dataset, 11% of tweets are sent by users with more than two mentions elsewhere.

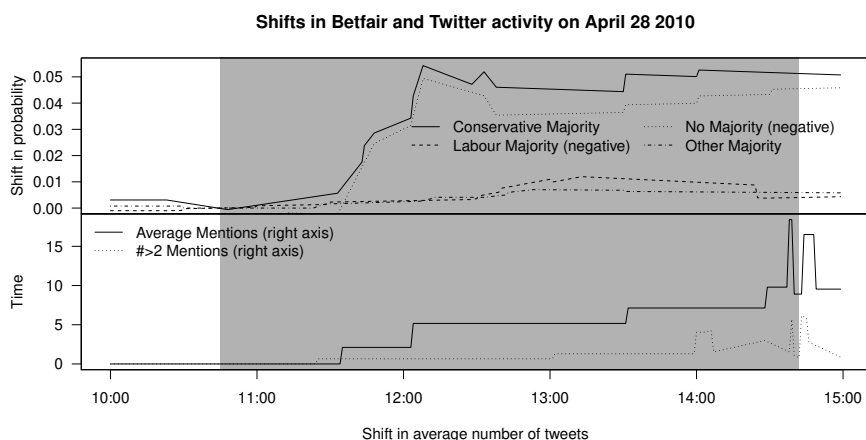**Shifts in Betfair and Twitter activity on April 28 2010**

Figure 10: Coefficient paths for all four Betfair markets (top panel) plus Twitter activity (average number of mentions each tweeter has and proportion of tweeters with at least two mentions, bottom panel). The two negative price reactions (Labour and No majority) are plotted in mirror image around the horizontal axis.

of tweeters in Figure 11. Tweets by all three types of tweeters doesn't provide much insight into the price movements. Ignoring the spikes before 11am (which are driven by there being less than one tweet per minute), the frequency of tweets by journalists and bloggers shift up too early at 11:11am and 11am respectively. Similarly (not plotted) the proportion of tweets by tweeters without any biographic entry yields little insight, increasing at 11:15 hence still 20 minutes before any price movements. Tweets by breaking news sources, on the other hand, potentially offer some insight; there are three minutes with spikes at 11:10am, 11:16am and 11:25am, but there is a sustained upward shift at 11:31am which persists until 11:48am, overlapping with the main price movements.

Taken together, our findings regarding the characteristics of tweeters suggests that it is authorative tweets, in particular from renowned breaking news sources that drive markets.

5.5 Characteristics of Tweets

As our analysis of the characteristics of tweeters has not produced any explanation regarding the gap between the Twitter reaction and Betfair reaction, we turn to the characteristics of tweets. We split this analysis into three: first we consider categories of tweets, then we conduct a sentiment analysis of tweets, and finally consider a number of key words contained within tweets.

5.5.1 Categories of Tweets

We can categorise tweets by their contents; we already considered mentions of other tweeters, but here we more specifically consider retweets. The retweet acts to propagate the information contained within a tweet; regardless of the intent behind the retweet, the action of retweeting causes the tweet to be seen

Shifts in Betfair and Twitter activity on April 28 2010
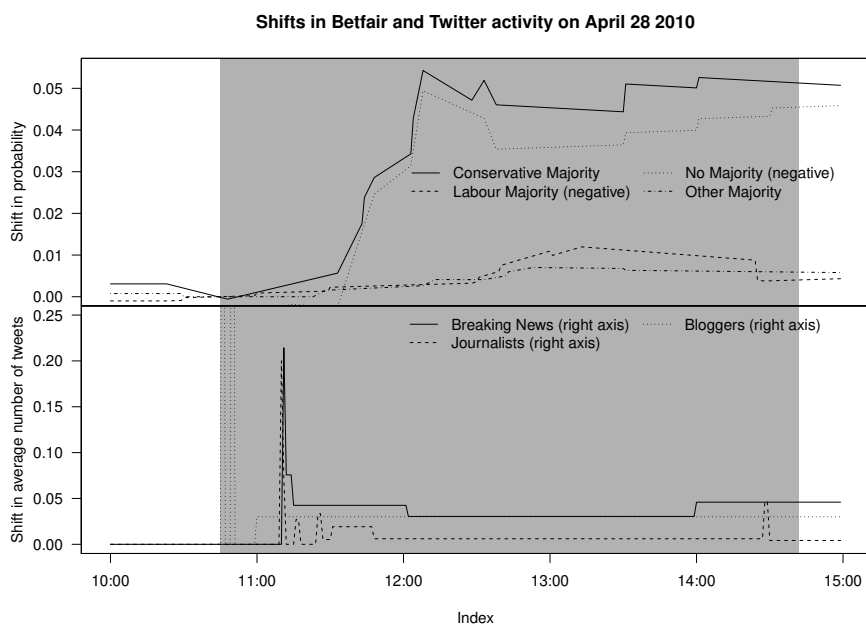
Figure 11: Coefficient paths for all four Betfair markets (top panel) plus Twitter activity (proportion of all tweets per minute by breaking news feeds, journalists and bloggers, bottom panel). The two negative price reactions (Labour and No majority) are plotted in mirror image around the horizontal axis.

Bigotgate Day Betfair Prices for Conservative/No Majority and Types of Tweets
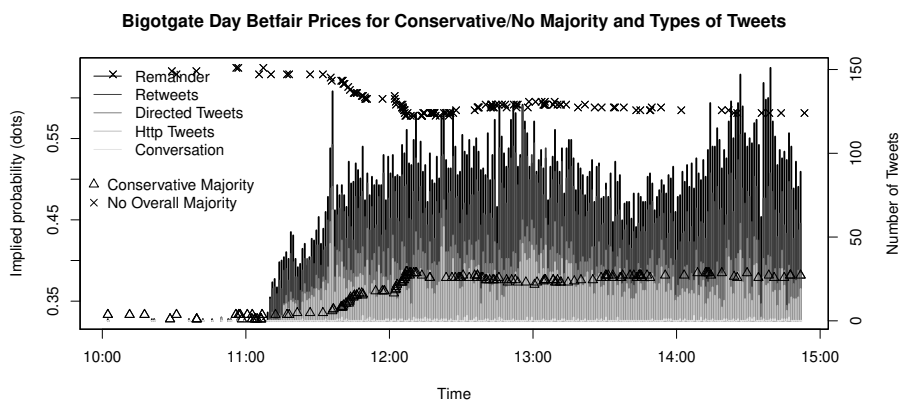

Figure 12: Different types of tweets alongside movements in Betfair prices for Conservative/No Majority between 11am and 1pm on April 28 2010.

by potentially all of the followers of the retweeter. We also consider tweets where a tweeter is mentioned, but the tweet is not a retweet, and we refer to these as directed tweets. Such directed tweets may form part of a conversation if the tweeter the tweet was directed at responds; more generally if we identify a number of tweets with a common group of tweets, we consider these to be conversations. Finally, tweeters can include in tweets links to other content outside of Twitter via hyperlinks.

Figure 5.5.1 plots retweets in red, tweets with hyperlinks in blue, directed tweets in green, and conversation tweets in light blue, alongside the Betfair prices for Conservative and No Majority. The numbers of each tweet type are stacked on top of each other.

We run indicator saturation on retweets, hyperlinked tweets, directed tweets and hyperlinked retweets; the resulting step and trend shifts are plotted in Figure 13 in the same style as Figures 9 and 11. The bottom panel shows the coefficient path, which reflects the significant changes in Twitter activity, and all four begin to trend upwards, or shift upwards, at between 11:10am and 11:11am, hence again long before Betfair market movements. Thus the simple act of retweeting, or posting hyperlinks, by twitterers does not in itself appear to influence market prices.

5.5.2 Sentiment Analysis

The simple volume of tweets, or particular types of tweets, may not be the appropriate indicator with which to determine the weight of new information entering the public sphere. News, by and large, is non-neutral, being positive or negative for parties involved. As such, the early tweets, such as the first one at 11:05, may simply have been factual rather than attaching any judgement regarding the nature of the news to any of the parties involved in the election.

We construct a sentiment index for our tweets in order to determine whether the balance of words used in the tweet is positive or negative. Our sentiment
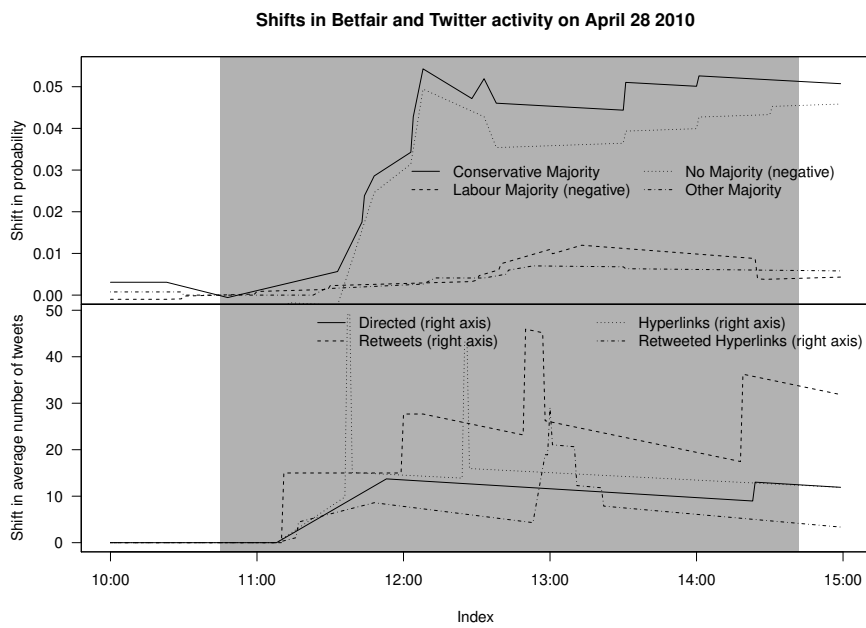
26

Figure 13: Coefficient paths for all four Betfair markets (top panel) plus Twitter activity (proportion of all tweets per minute that are retweets, have hyperlinks, both, or are directed tweets, bottom panel). The two negative price reactions (Labour and No majority) are plotted in mirror image around the horizontal axis.

27

index is very simple; we use Hu and Lui's opinion lexicon index to assign words a positive or negative score, depending on the nature of that word. We remove stopwords and focus on adjectives in constructing our index. We then sum up the scores of the words and get a score for a given tweet. There are many criticisms of sentiment indices, as Anstead and O'Loughlin (2014) point out; by considering words in isolation contextual information is lost, so for example sarcasm will be missed by such measures. We nonetheless use a simple index in order to determine whether even in such a basic form it can help understand the sluggish response to news breaking on Twitter.

Figure 14 plots the standard coefficient paths as in Figures 9–13, with the addition of the raw sentiment index, which is plotted in grey and can be seen moving around the shift and trend indicators plotted. Once the #bigotgate story breaks after 11am, the sentiment index undergoes a clear structural shift, and is persistently negative, albeit becoming less so throughout the period. This is picked up as a negative step shift at 11:08am, and at 11:15am a very slight positive trend begins, with a step shift up at 12:30pm.

As with many other series we have considered, however, this occurs more than twenty minutes before any price movement begins, and as such negative sentiment alone cannot shed light on observed prices. We split sentiment up between those who identified themselves as left- or right-leaning, and we also calculated the disagreement measure proposed in Meinusch and Tillmann (2015) for positive or negative tweets, and we also apply a sentiment index constructed by Di Fatta et al. (2015) for the 2015 UK General Election. We found essentially the same results in each case; basic sentiment analysis yields little insight into observed price movements.

5.5.3   Information Content of Tweets

Sentiment indices give some idea about the nature of news; it may be helpful to consider the text context of tweets in a slightly different manner: What new information is being provided in each new tweet? New information is a collection of words, the collection of which constitutes important information, but the constituent parts of which may contain words not previously used in this context, or may add emphasis via repetition. The premise of this investigation is the salience of the term 'bigot', but equally the appearance of particular nouns (Brown, Duffy, gaffe, apology etc) and adjectives (e.g. big, huge, massive, awful), marks the breaking of new information. We have already considered generic retweets, but it may be that if particular words appear particularly frequently this helps foster some sense of agreement regarding the impact of the news on market prices.

We break this analysis into two sections; firstly we consider the appearance of previously unused words, before we then consider the use of particular salient words in tweets.

New Words   We observe tweets containing one particular word 'bigot'; on average the tweets in our dataset contain 17–18 words, and as such we must consider the nature of the remaining words. If, as one imagines with a word such as 'bigot', there is no single conversation on Twitter and different users are engaged in potentially multiple conversations, or simply posting thoughts about bigots in general, it seems likely that the words used in any given tweet,

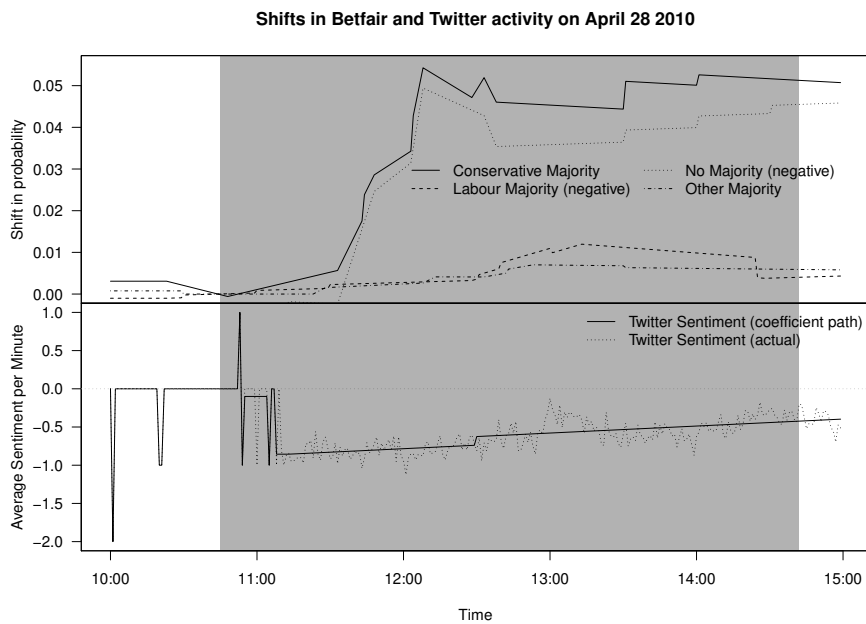**Shifts in Betfair and Twitter activity on April 28 2010**

Figure 14: Coefficient paths for all four Betfair markets (top panel) plus Twitter sentiment (average sentiment per tweet, bottom panel). The two negative price reactions (Labour and No majority) are plotted in mirror image around the horizontal axis.

29

chronologically ordered, will be quite different. Conversely, if discussion is solely focussed on one single topic, then it seems likely that the number of words used in any given tweet will be similar to those previously used. In a situation of very few new words per tweet, no new information is contained within tweets, although the salience of the existing information must be increasing. Additionally, it may be that new words represent the interpretation of existing news, and hence we seek to consider the extent to which the number of new words in tweets aligns with Betfair price movements.

We create a list of stop words, namely words that are very commonly used in everyday situations ('the', 'and', 'is', 'at', etc.), and add each word in any tweet (chronologically ordered) that does not exist in the list of stop words. Whilst doing this, we count how many words in each tweet are not in the growing list of stop words. This provides us with the number of words previously not used on April 28 in any given tweet, a measure of the uniqueness of the information contained within that tweet. Considering the whole of April 28 as the time frame in which to judge new information is arbitary, and hence equally we consider only words used in the previous hour when augmenting the list of stop words. Hence if discussion returns to a topic previously discussed more than an hour previously, this counts as new information, based on the assumption that the topic would not have been returned to were there not new information being released.

Figure 15 plots the average number of new words per minute on the standard indicator saturation plot. A higher number of new words per minute is indicative of greater variation in topics being mentioned, and hence this characterises the time period before the #bigotgate story broke, with spikes for each minute where tweets occur being quite high (6–10). As the #bigotgate story breaks, this falls to about one word per minute, and continues to fall to about 0.5 words per minute by about 2pm (about the time Brown was returning to Duffy's house to apologise). The shift to the post-#bigotgate trend occurs at 11:11am, and hence again occurs before the major price movements.

Salient Words Rather than consider the mention of new words, which may not necessarily represent the point at which they become widely known, or salient, it may be more informative to consider a subset of words or grammatical terms used in relation to the story, and whether their usage at any particular time point may have prompted the price movements observed. As already mentioned, one way in which Twitter may impact the information efficiency of markets is by increasing the salience of events.

We consider words based on their frequency between 11am and 3pm, ruling out references to Gordon Brown (PM, prime minister) and Gillian Duffy (voter, pensioner, woman). This leaves words with the stem 'call' (calling, called, calls, call), references to the microphone (mic, microphone), caught, election, news and gaffe as the most commonly used. Of these, most might be described as factual, describing events, whereas gaffe represents a judgement on the events and their importance within the election campaign, and hence market prices related to election outcomes. By and large, gaffes are considered to be harmful to a politician's electoral chances.

Additionally, the use of exclamation marks may be indicative of strength of feeling about the content of a tweet, as may be the use of multiple exclamation
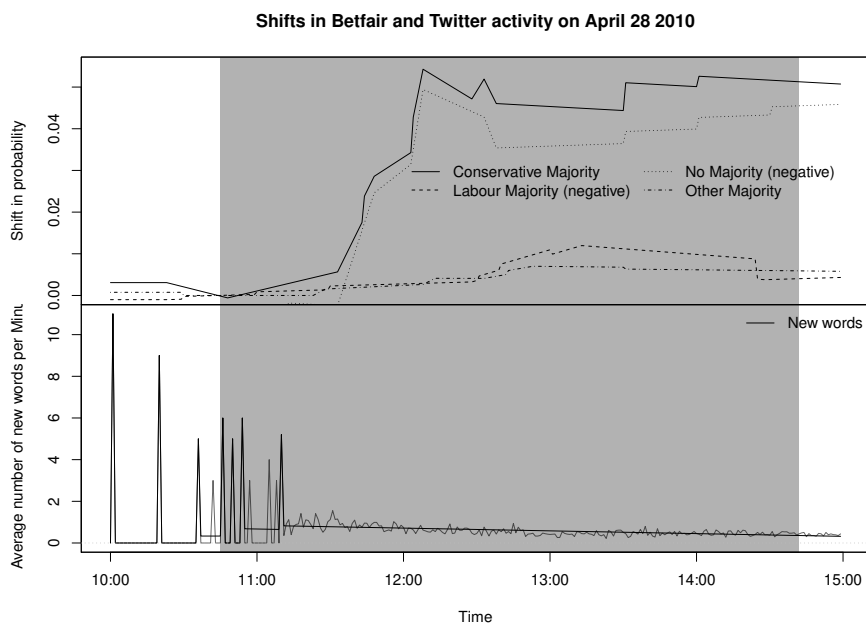
Figure 15: Coefficient paths for all four Betfair markets (top panel) plus new words Twitter (average number of new words per tweet, bottom panel). The two negative price reactions (Labour and No majority) are plotted in mirror image around the horizontal axis.

31

Table 4: Time of first significant trend or step shift dummy found for indicator saturation looking at the frequency of listed words or grammatical terms.

Word	Time	Word	Time
"watch"	11:16am	exclamation mark	11:13am
"bbc"	11:16am	exclamation marks	11:15am
"radio"	11:15am	question mark	11:17am
"tv"	11:29am	question marks	11:10am
"gaffe"	11:10am	proportion upper case	11:11am

marks. Question marks in Tweets may reflect uncertainty about the meaning of information being conveyed. We also consider the use of exclamation marks and question marks using indicator saturation, alongside a small number of commonly used words. We found a very similar pattern with almost all of these words, namely that the change in Twitter activity occurred too soon to be reasonably linked to price movements. In Table 4 we report the main structural shift for each word or grammatical term(s). With the exception of the word "tv", all significant shifts detected are around 11:15am, hence around 20 minutes before price movements occurred. With "tv", the change at 11:29am is very small, and from a visual investigation of the time profile of such tweets it appears that the upward shift probably began slightly earlier than 11:29am, and in addition many of the references to "tv" are retweets of a tweet about Duffy answering her phone whilst being interviewed on TV.

As such, it seems unlikely that a detailed analysis of particular words may yield insights, since despite different time profiles of different types of Twitter word activity, all of the significant shifts in levels and trends occur well in advance of any price movements.

6 Conclusions

In this paper we consider the speed with which markets respond to new information, a fundamental aspect of market efficiency. We use as our lens of observation the response on Betfair's prediction market to news of a significant political event in the UK during the 2010 General Election campaign, using Twitter to capture the news regarding the event, as it broke. We use modern time series econometric methods to determine the existence, and timing, of structural breaks in both social media behaviour and market prices. We find that markets appear to have responded surprisingly sluggishly to new information; around a thousand tweets are sent and twenty minutes pass before any price reaction is noted. This is considerably slower than previous investigations of market efficiency using Betfair have found (Croxson and Reade, 2014; Buraimo et al., 2008). Further analysis suggests that influential users of Twitter, as well as breaking news feeds from major 'old media' networks, drive the price movement that occurs, which supports the work of Castillo et al. (2011) on the credibility of information on Twitter. Once credible sources of information begin to tweet in significantly greater volumes, then prices begin to move. Our results in this way contribute to the literature on the efficiency of markets with respect to the release of new information, and in particular the way in which that information is disseminated and certified.
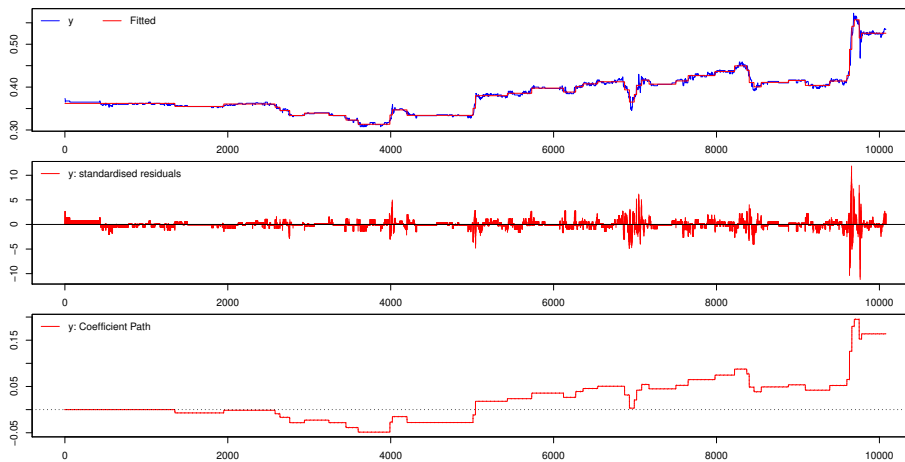
Figure 16: Graphical output from indicator saturation on seven days of the Conservative majority Betfair implied probability. 0 is midnight April 25, 10,000 is 22:39pm on May 1. The #bigotgate timeline occurs between observations 4965–5202. The top panel shows the actual values and fitted values (the step shifts and trend changes detected), the middle panel shows the residuals from the same regression model, and the bottom panel shows the step and trend terms on their own. For additional output, see the R code made available by the authors.

A   Supplementary Material

B   ISAT Results

References

N. Anstead and B. O'Loughlin. Social Media Analysis and Public Opinion: The 2010 UK General Election. *Journal of Computer-Mediated Communication*, 2014.

P. Asquith. Merger Bids, Uncertainty and Stock Holder Returns. *Journal of Financial Economics*, 11:51–83, 1983.

L. Bachelier. Theorie de la speculation. In P. Cootner, editor, *The Random Character of Stock Market Prices*. MIT Press, 1964, Cambridge, MA, 1900.

J. Bai. Estimation of a Change Point in Multiple Regression Models. *Review of Economics and Statistics*, 79(4):551–563, 1997.

J. Bai and P. Perron. Computation and Analysis of Multiple Structural Change Models. *Journal of Applied Econometrics*, 18(1):1–22, 2003.

R. Ball and P. Brown. An Empirical Evaluation of Accounting Income Numbers. *Journal of Accounting Research*, 6:159–178, 1968.

J. Bollen, H. Mao, and X. Zeng. Twitter Mood Predicts the Stock Market. *Journal of Computational Science*, 2(1):1–8, 2011.

33

Table 5: Table of output from indicator saturation on each Betfair price (run with step shifts added only every hour to condense output — see R code made available for (much) larger version of this table). Each column is a Betfair price, and the rows reflect retained step shift dummy variables for a particular time.

Variable	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic	Coefficient	t-statistic
1 mconst	0.360	2047.722	0.040	748.971	0.016	391.311	0.593	3563.967
10 2015-04-25 08:00					-0.003	-65.237		
2 2015-04-25 18:00			-0.005	-53.044				
3 2015-04-26 05:00			-0.002	-20.177				
4 2015-04-26 17:00					-0.001	-31.429		
5 2015-04-26 20:00	-0.022	-59.881					0.028	81.649
6 2015-04-27 09:00	-0.020	-40.965	-0.003	-54.195	-0.001	-13.961	0.023	48.581
7 2015-04-27 18:00							-0.017	-37.887
8 2015-04-27 19:00	0.019	41.079						
9 2015-04-27 22:00					-0.002	-47.881		
11 2015-04-28 12:00	0.045	100.136	-0.005	-23.613	0.004	33.282	-0.046	-109.977
12 2015-04-28 13:00			-0.003	-11.786	0.002	12.345		
13 2015-04-28 15:00			0.006	34.206	-0.002	-27.529		
14 2015-04-28 23:00	0.019	41.030					-0.017	-44.487
15 2015-04-29 06:00			-0.003	-39.638				
16 2015-04-29 16:00	0.006	14.110						
17 2015-04-29 19:00			0.005	37.270				
18 2015-04-29 22:00			-0.010	-69.301	-0.004	-75.444		
19 2015-04-30 05:00					-0.003	-51.962		
20 2015-04-30 08:00	0.027	63.090						
21 2015-04-30 09:00							-0.022	-58.787
22 2015-04-30 20:00					0.002	13.046		
23 2015-04-30 21:00	-0.022	-52.338	0.002	27.310	0.001	12.245	0.018	44.576
24 2015-05-01 09:00					0.003	54.461		
25 2015-05-01 16:00							-0.044	-38.820
26 2015-05-01 17:00	0.119	231.854			-0.005	-86.179	-0.071	-60.020

D. Boyd, S. Golder, and G. Lotan. Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In *43rd Hawaii International Conference on System Sciences (HICSS)*, pages 1–10. Institute of Electrical and Electronics Engineers, 2010.

A. Brown. Information processing constraints and asset mispricing. *The Economic Journal*, 124(575):245–268, 2014.

B. Buraimo, D. Peel, and R. Simmons. Gone in 60 Seconds: The Absorption of News in a High-Frequency Betting Market. 2008.

J. Burgess and A. Bruns. (Not) the Twitter election: the dynamics of the# ausvotes conversation in relation to the Australian media ecology. *Journalism Practice*, 6(3):384–402, 2012.

J. Campos, N.R. Ericsson, and D.F. Hendry. General-To-Specific Modeling: an Overview and Selected Bibliography. International Finance Discussion Papers 838, Board of Governors of the Federal Reserve System (U.S.), 2005.

C. Castillo, M. Mendoza, and B. Poblete. Information credibility on twitter. In *Proceedings of the 20th international conference on World wide web*, pages 675–684. ACM, 2011.

J.L. Castle, J.A. Doornik, D.F. Hendry, and F. Pretis. Detecting Location Shifts during Model Selection by Step-Indicator Saturation. *Econometrics*, 3(2):240–264, April 2015. URL http://ideas.repec.org/a/gam/jecnmx/v3y2015i2p240-264d48166.html.

K. Croxson and J.J. Reade. Exchange vs. Dealers: A High-Frequency Analysis of In-Play Betting Prices. Discussion Papers 11-19, Department of Economics, University of Birmingham, January 2011.

K. Croxson and J.J. Reade. Information and Efficiency: Goal Arrival in Soccer Betting. *Economic Journal*, 124:62–91, March 2014.

G. Di Fatta, J.J. Reade, Jaworska S., and A. Nanda. Big Social Data and Political Sentiment: the Tweet Stream during the UK General Election 2015 Campaign. Proceedings, 8th IEEE International Conference on Social Computing and Networking (SocialCom 2015), 2015.

J.A. Doornik. Autometrics. In J.L. Castle and N. Shephard, editors, *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*. Oxford University Press, Oxford and New York, 2009.

R. Duffy. If #Indyref was decided by tweets, it would've been a landslide Yes vote. Technical report, thejournal.ie, September 2014. URL `http://www.thejournal.ie/indyref-tweet-results-1679617-Sep2014/`. Last accessed 9 October 2014.

A.-M. Elliott. 9 Breaking News Tweets That Changed Twitter Forever. Technical report, mashable.com, October 2013. URL `http://mashable.com/2013/10/31/twitter-news/`.

E. F. Fama. The Behaviour of Stock Market Prices. *The Journal of Business*, 38:34–105, 1965.

E. F. Fama. Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2):383–417, 1970.

E. F. Fama. Market efficiency, long-term returns and behavioral finance. *Journal of Financial Economics*, 49(3):283–306, 1998.

E. F. Fama, L. Fisher, M. C. Jensen, and R. Roll. The adjustment of stock prices to new information. *International Economic Review*, 10:1–21, 1969.

C. Feng and N. Yang. Twitter Adoption in Congress. *Review of Network Economics*, 10(1):1–46, March 2011. URL `http://ideas.repec.org/a/bpj/rneart/v10y2011i1n3.html`.

A. Gelman and G. King. Why are American Presidential Election Campaign Polls so Variable when Votes are so Predictable? *British Journal of Political Science*, 23(4):409–451, October 1993.

R. Gil and S. Levitt. Testing the Efficiency of Markets in the 2002 World Cup. *The Journal of Prediction Markets*, 1:255–270, 2007.

D.V. Greetham and J.A. Ward. Structure, pace and balance in twitter conversations. (MPS-2012-25), November 2012.

D. Hendry, S. Johansen, and C. Santos. Automatic Selection of Indicators in a Fully Saturated Regression. *Computational Statistics*, 23(2):337–339, April 2008.

D.F. Hendry and H.-M. Krolzig. Improving on 'Data Mining Reconsidered' by K.D. Hoover and S.J. Perez. *Econometrics Journal*, 2(2):202–219, 1999.

D.F. Hendry, J.A. Doornik, and F. Pretis. Step-indicator Saturation. Economics Series Working Papers 658, University of Oxford, Department of Economics, June 2013. URL http://ideas.repec.org/p/oxf/wpaper/658.html.

K.D. Hoover and S.J. Perez. Data Mining Reconsidered: Encompassing and the General-to-Specific Approach to Specification Search. *Econometrics Journal*, 2:167–191, 1999.

A. Iannucci. The Duffy affair turned the media into a pack of shrieking gibbons. *The Independent*, 4 May 2010. URL http://www.independent.co.uk/voices/commentators/armando-iannucci-the-duffy-affair-turned-the-media-into-a-pack--of-shrieking-gibbons-19 html.

D. Janetzko. Predictive modeling in turbulent times — What Twitter reveals about the EUR/USD exchange rate. *NETNOMICS: Economic Research and Electronic Networking*, pages 1–38, August 2014.

A. Java, X. Song, T. Finin, and B. Tseng. Why we Twitter: Understanding microblogging usage and communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*, pages 56–65. Association for Computing Machinery, 2007.

S.J. Johansen and B. Nielsen. An Analysis of the Indicator Saturation Estimator As a Robust Regression Estimator. In J.L. Castle and N. Shephard, editors, *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, Oxford and New York, 2009. Oxford University Press.

T.J. Johnson, B.K. Kaye, S.L. Bichard, and W.J. Wong. Every Blog Has Its Day: Politically-interested Internet Users' Perceptions of Blog Credibility. *Journal of Computer-Mediated Communication*, 13(1):100–122, 2007.

H. Kwak, C. Lee, H. Park, and S. Moon. What is Twitter, a social network or a news media? In *Proceedings of the 19th International Conference on World Wide Web*, pages 591–600. Association for Computing Machinery, 2010.

D. Kwiatkowski, P.C.B. Phillips, P. Schmidt, and Y. Shin. Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root. *Journal of Econometrics*, 54(1–3):159–178, 1992.

A.O. Larsson and H. Moe. Studying political microblogging: Twitter users in the 2010 Swedish election campaign. *New Media & Society*, 14(5):729–747, 2012.

K. Lerman and R. Ghosh. Information Contagion: An Empirical Study of the Spread of News on Digg and Twitter Social Networks. In *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, pages 90–97. Association for the Advancement of Artificial Intelligence, 2010.

S.A. Macskassy and M. Michelson. Why do people retweet? Anti-homophily wins the day! In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*. Association for the Advancement of Artificial Intelligence, 2011.

A. Meinusch and P. Tillmann. Quantitative easing and tapering uncertainty: Evidence from twitter. Joint discussion paper series in economics, Universities of Aachen, Giessen, Göttingen, Kassel, Marburg and Siegen, 2015.

M. Naaman, J. Boase, and C.H. Lai. Is it really about me? Message content in social awareness streams. In *2010 ACM Conference on Computer Supported Cooperative Work, New York*, pages 189–192. Association for Computing Machinery, 2010.

M. Nagarajan, H. Purohit, and A.P. Sheth. A Qualitative Examination of Topical Tweet and Retweet Practices. In *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*. Association for the Advancement of Artificial Intelligence, 2010.

M. Nardo, M. Petracco-Giudici, and M. Naltsidis. Walking down Wall Street with a Tablet: A Survey of Stock Market Predictions Using the Web. *Journal of Economic Surveys*, 2015.

A.-M. Popescu and M. Pennacchiotti. Detecting controversial events from Twitter. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, pages 1873–1876. Association for Computing Machinery, 2010.

A. Porshnev, I. Redkin, and A. Shevchenko. Improving prediction of stock market indices by analyzing the psychological states of twitter users. HSE Working papers WP BRP 22/FE/2013, National Research University Higher School of Economics, 2013. URL `http://ideas.repec.org/p/hig/wpaper/22-fe-2013.html`.

F. Pretis, J.J. Reade, and G. Sucarrat. *gets: General-to-Specific (GETS) Modelling and Indicator Saturation Methods*, 2014a. URL `http://CRAN.R-project.org/package=gets`. R package version 0.2.

F. Pretis, L. Schneider, and J.E. Smerdon. Detecting breaks by designed functions applied to volcanic impacts on hemispheric surface temperatures. Discussion paper, University of Oxford, Department of Economics, 2014b.

T. Rao and S. Srivastava. Modeling Movements in Oil, Gold, Forex and Market Indices using Search Volume Index and Twitter Sentiments. Papers 1212.1037, arXiv.org, December 2012. URL `http://ideas.repec.org/p/arx/papers/1212.1037.html`.

S.E. Said and D.A. Dickey. Testing for Unit Roots in Autoregressive-moving Average Models of Unknown Order. *Biometrika*, 71(3):599–607, 1984.

T. Sakaki, M. Okazaki, and Y. Matsuo. Earthquake shakes Twitter users: real-time event detection by social sensors. In *Proceedings of the 19th International Conference on World Wide Web*, pages 851–860. Association for Computing Machinery, 2010.

J. Sankaranarayanan, H. Samet, B.E. Teitler, M.D. Lieberman, and J. Sperling. Twitterstand: News in tweets. In *Proceedings of the 17th ACN Sigspatial International Conference on Advances in Geographic Information Systems*, pages 42–51. Association for Computing Machinery, 2009.

P. Simm. Election 2015: What not to do on the campaign trail. *BBC News*, March 28 2015. URL `http://www.bbc.co.uk/news/election-2015-31522819`.

S. Sinha, C. Dyer, K. Gimpel, and N.A. Smith. Predicting the NFL using Twitter. Technical report, Prague, Czech Republic, September 2013. In Proceedings of the ECML/PKDD 2013 Workshop on Machine Learning and Data Mining for Sports Analytics.

T.O. Sprenger, A. Tumasjan, P.G. Sandner, and I.M. Welpe. Tweets and trades: The information content of stock microblogs. *European Financial Management*, 2013.

T.O. Sprenger, P.G. Sandner, A. Tumasjan, and I.M. Welpe. News or Noise? Using Twitter to Identify and Understand Company-specific News Flow. *Journal of Business Finance & Accounting*, 41(7-8):791–830, 2014.

A. Tabarrok and T. Cowen. The End of Asymmetric Information. *Cato Unbound*, April 2015. URL `http://www.cato-unbound.org/2015/04/06/alex-tabarrok-tyler-cowen/end-asymmetricinformation`.

M. Thelwall, K. Buckley, and G. Paltoglou. Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*, 62(2):406–418, 2011.

J. Treme and Z. VanDerPloeg. The Twitter Effect: Social Media Usage as a Contributor to Movie Success. *Economics Bulletin*, 34(2):793–809, 2014. URL `http://ideas.repec.org/a/ebl/ecbull/eb-13-00821.html`.

A. Tumasjan, T.O. Sprenger, P.G. Sandner, and I.M. Welpe. Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. In *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, pages 178–185. Association for the Advancement of Artificial Intelligence, 2010.

A. Tumasjan, T.O. Sprenger, P.G. Sandner, and I.M. Welpe. Election forecasts with Twitter — How 140 characters reflect the political landscape. *Social Science Computer Review*, 29:402–418, 2011.

H. Wang, D. Can, A. Kazemzadeh, F. Bar, and S. Narayanan. A System for Real-time Twitter Sentiment Analysis of 2012 US Presidential Election Cycle. In *Proceedings of the ACL 2012 System Demonstrations*, pages 115–120. Association for Computational Linguistics, 2012.