To what extent sentiment analysis of Twitter is able to forecast electoral results? Evidence from France, Italy and the United States

Luigi Curini – Università degli Studi di Milano, luigi.curini@unimi.it
Andrea Ceron – Università degli Studi di Milano
Stefano M. Iacus – Università degli Studi di Milano

Work in progress – Do not quote without authors’ permission

The growing diffusion of social media raises the possibility to delve into the web to explore and track the political preferences of citizens. Monitoring social media during an electoral campaign can become therefore a useful complement of traditional off-line polls. Some scholars, however, go even further than that, claiming that by analyzing social media we can produce a reliable forecast of the final result. Relying on a proper methodology for sentiment analysis remains a crucial issue in this respect. In this work, we apply the recent supervised method proposed by Hopkins and King (2010) to analyze the voting intention of Twitter-users in three different contexts: France (for the 2012 Legislative election, first round), Italy (for the two-rounds of the centre-left 2012 primaries) and the United States (for the 2012 Presidential election). Our analysis shows a remarkable ability of Twitter to forecast electoral results, in some cases improving on mass surveys results. On the contrary, more “naïve” analyses of social media provide results less satisfactory than ours.

**Keywords:** sentiment analysis, text mining, text analytics, social media, electoral forecast

Paper to be presented at the Panel:
“Partisanship and the New Media: The Effect of the Internet on the Stability of Political Attitudes”

7th ECPR General Conference
Sciences Po, Bordeaux
4 - 7 September 2013
Introduction

The exponential growth of social media and social network sites, like Facebook and Twitter, have started to play a growing role on real world politics in recent years. Social networks have been used, for example, to organize demonstrations and revolts during the ‘Arab spring’ (Ghannam, 2011; Cottle 2011); to engage individuals in mobilizations (Segerberg and Bennett, 2011; Bennett and Segerberg, 2011); to build social movements and political parties, like the Pirate Party in Sweden and Germany or the Italian ‘Movimento 5 Stelle’, which use the web to set the party line and to select candidates.¹

The diffusion of social media raises however also the possibility to delve into the web to explore and track the political and electoral preferences of citizens (Madge et al., 2009; Woodly, 2007). As a matter of fact, scholars have recently started to explore social media as a device to assess the popularity of politicians (Gloor et al., 2009), to track the political alignment of social media users (Barberá 2012, Conover et al. 2011) and to compare citizens’ political preferences expressed on-line with those caught by polls (O’Connor et al., 2010). Analyzing social media during an electoral campaign can indeed be a useful supplement/complement of traditional off-line polls for a number of reasons (Xin et al., 2010). Besides being cheaper and faster compared to traditional surveys, a social media analysis allows to monitor day-by-day (at the extreme, hour-by-hour) an electoral campaigning. Through that, the possibility to nowcast the campaign, that is to track in real time trends and capture (eventual) sudden change (so called “momentum”: Jensen and Anstead 2013) in public opinion well before of what can be done via traditional polls (as a result, for example, of a TV debate: see below), becomes a reality.³ Some scholars, however, go even further than that, claiming that analyzing social media allows a reliable forecast of the final result (Tjong and Bos, 2012). This is quite fascinating cause forecasting an election is one of the few exercises on social events where an independent measure of the outcome that a model is trying to predict is clearly and indisputable available, i.e., the vote-share of candidates (and/or parties) at the ballots.

¹ For a more skeptical view on the role that social media can play in organizing revolts, with respect in particular to protests related to the 2009 Iranian elections, see Morozov (2009).
² During the EU elections held in 2009, the Pirate Party won 7.1% of votes in Sweden, gaining 1 seats in the EU parliament. In Germany, it received 2% of votes in the 2009 German Federal Election. It subsequently obtained positive results in German regional elections. In Italy, the Movimento 5 Stelle also reported surprising results during local elections held between 2009 and 2012, before obtaining a striking 25.1% in the 2013 general elections.
³ In addition, traditional surveys pose solicited questions and it is well known this might inflate the share of strategic answers (Payne, 1951). Conversely sentiment analysis does not make use of questionnaires and just focus on listening to the stream of unsolicited opinions freely expressed on internet. In other words it adopts a bottom-up approach, at least if compared with the top-down approach of off-line surveys.
Some of these works rely on very simple techniques, focusing on the volume of data related to parties or candidates. For instance, Véronis (2007) proved that the number of candidate mentions in blog posts is a good predictor of electoral success and can perform better than election polls. Along the same line, some scholars claimed that the number of Facebook supporters could be a valid indicator of electoral fortunes (Upton, 2010; Williams and Gulati, 2008), while Tumasjan et al. (2010) compared party mentions on Twitter with the results of the 2009 German election and argued that the relative number of tweets related to each party is once again a good predictor for its vote share.

It has been also noted that the mere count of mentions or tweets can be a rather crude way to provide an accurate foresight (Chung and Mustafaraj, 2011). Accordingly, other studies tried to improve this stream of research by means of sentiment analysis. Lindsay (2008), for example, built a sentiment classifier based on lexical induction and found correlations between several polls conducted during the 2008 presidential election and the content of wall posts available on Facebook. O’Connor et al. (2010) show similar results displaying correlation between Obama’s approval rate and the sentiment expressed by Twitter users. In addition, sentiment analysis of tweets proved to perform as well as polls in predicting the results of both the 2011 (Tjong Kim Sang and Bos, 2012) and the 2012 legislative elections in the Netherlands (Sanders and den Bosch 2013), while the analysis of multiple social media (Facebook, Twitter, Google and YouTube) was able to outperform traditional surveys in estimating the results of the 2010 UK Election (Franch, 2012).

Still not all enquiries succeeded in correctly predicting the outcome of the elections (Gayo-Avello et al., 2011; Goldstein and Rainey, 2010). For instance, it has been shown that the share of campaign weblogs prior to the 2005 federal election in Germany was not a good predictor of the relative strength of the parties insofar as small parties were overrepresented (Albrecht et al., 2007). In a study about Canadian elections, Jansen and Koop (2005) failed in estimating the positions of the two largest parties. Finally, Jungherr et al. (2011) criticized the work of Tumasjan et al. (2010) arguing that it does not satisfy the “freedom from irrelevant alternatives” condition (e.g., including the German Pirate Party into the analysis would have had yielded a negative effect on accuracy of the predictions).

Gayo-Avello (2011, 2012) pinpoints several theoretical problems with predicting elections based on tweets. First, he stresses how several of the quoted words are not predictions at all, given that they generally present post-hoc analysis after an election has already occurred. This, inter-alia, also increases the chances that only good results are published, inflating the perceived ability of using social media to correctly forecast election (see also Lewis-Beck 2005). Second, he underlines the difficulty to catch the real meaning of the texts analyzed, given that political discourse is plagued with humor, double meanings, and sarcasm. Third, he highlights the risk of a spamming effect, i.e., not all the tweets are trustworthy given the presence of rumors and misleading information. Finally, in most of the previous studies demographics are neglected: not every age, gender, social, or racial group is in fact equally represented in social-media.

---

4 Sentiment analysis consists in analyzing texts to extract information.
In the present paper we show how both the second and the third of the above concerns can be addressed by relying on a proper methodology for sentiment analysis. With respect to the first concern, on the other side, all the analyses that we discuss here have been conducted (and published on media) before the day of the election, so they can be considered as real predictions. More in details, by employing the method proposed in Hopkins and King (2010) (HK, from now on), we will analyze three different countries: France, considering the 2012 Legislative election, first round; Italy, considering the first and second round of the primary elections held by the centre-left coalition in November 2012; the United States, focusing on the electoral campaign for the 2012 Presidential election. In each context we will contrast our results with the ones obtained through traditional off-line surveys as well as to actual electoral results. The variety of cases so analyzed has been deliberately pursued to better investigate the strength and the limits of monitoring social media through a method that, for a number of different reasons explained below, seems to be a clear advance compared to previous analyses. In the conclusion we advance some suggestions for future research.

1. How to scrutinize voters’ preferences through social-media

Nowadays, Internet access is available to a wider audience of citizens (and voters). In turn, the usage of social media is growing at very fast rates. Around 35 people out of 100 got access to the web, all over the world, in 2011 (approximately 2.5 billion people). Among them, 72% of the internet population is active on at least one social network, like Facebook (over 1.1100 million of users) or Twitter (over 500 million of users).

Given the wide amount of data about public opinion available on-line (and its growing relevance), monitoring this flow of preferences becomes a relevant task. The problem is to select the kind of method more appropriate in this regard. While earlier studies, as already discussed, focused mainly on the volume of data (related, for instance, to each party or candidate), here we aim to catch the attitude of internet users going beyond the mere number of mentions. To this aim we will employ the method recently proposed in Hopkins and King (2010).

---

7 A second stream of research in the literature on social media adopts a more “political supply-side” approach, analyzing how internet and the diffusion of social-media have affected the content of the electoral campaigning and the political communication by candidates and parties. While some of the initial hope for e-democracy have been unfulfilled (Chadwick, 2008; Hilbert, 2009), internet still provides new opportunities linked with electoral campaign (Larsson and Moe, 2012; Smith, 2009) such that politicians can engage with the wider public (Gibson et al., 2008).
The HK method presents two specific advantages compared to other methods, especially when it comes to relate social media and (the results of) elections: a better interpretation of the texts and more reliable aggregate results. The former advantage is mainly linked to the fact that HK performs a supervised sentiment analysis. The traditional approach to sentiment analysis is in fact based on the use of ontological dictionaries: this means that a text is assigned to a specific opinion category if some pre-determined words or expressions appear (or do not) in the text (see Grimmer and Stewart 2013 on this point). The benefit of this approach is, of course, the possibility to implement a totally automated analysis (once the dictionary has been defined). The strong drawback, on the other side, is the difficulty in classifying opinions expressed through ironic or paradoxical sentences, or in appreciating all the language nuances (specific jargons, neologisms, etc.): the informal expression “what a nice rip-off!” is quite ambiguous from the viewpoint of an ontological dictionary, because it includes both a positive and a negative term.

The HK method, on the contrary, is based on a two-stage process. The first step involves human coders and consists in reading and coding a subsample of the documents downloaded from some Internet source. This subsample – with no particular statistical property, see below - represents a training set that will be used by the HK algorithm to classify all the unread documents, in the second stage. Human coders are of course more effective and careful than ontological dictionaries in recognizing all the previously discussed language specificity and the author’s attitude to the subject (Hopkins and King, 2012). Moreover, human coding is better suited to identify the (ever-present) problem of spamming in social communication. This is of course important, given that spamming can have an impact on the accuracy of the final result. At the second stage, the automated statistical analysis provided by the HK algorithm extends such accuracy to the whole population of posts, allowing for properly catching the opinions expressed on the web. Indeed, the expected error of the estimate is around 3%.8

The methodology is based on the assumption that the opinion of people posting on social networks can be deduced by all the terms they use: not only the terms explicitly related to the topic they talk about, but also the “neutral” part of the language commonly used. Therefore, in order to characterize the different opinions, the single units (blogs, posts) in the data set are decomposed into their own single words: consequently, each unit is represented by the vector of the terms used, which we call “word profile” of the unit.9

The formal background of the method is simple (for further details see: Hopkins and King, 2010). Indicate by $S$ the word profiles used in the text units and by $D$ people’ opinions expressed in the texts. The target of estimation is $P(D)$, i.e. the frequency distribution of the opinions over the posting population.

---

8 [http://www.crimsonhexagon.com/quantitative-analysis/](http://www.crimsonhexagon.com/quantitative-analysis/). From our replications, the root mean square error of the estimates drops until 1.5% when the number of hand-coded documents increases up to 500.

9 In other terms, a “word profile” is a vector made of 0’s and 1’s: 0 when a term does not appear in the unit (but it is used in some other units) and 1 when a term appears in the unit.
The standard statistical approach is to decompose \( P(D) \) in the following way:

\[
P(D) = P(D|S) P(S) \quad (1)
\]

\( P(S) \) corresponds to a tabulation of frequencies of word profiles in the whole population of texts. \( P(D|S) \) is estimated from the training set as \( P_T(D|S) \), i.e. the conditional frequency distribution of word profiles inside the training set, using any standard classifier (multinomial regression, classification trees, random forests, support vector machines, etc). Through this approach, each individual classification of posts in the test set (i.e. post belonging to the corpus of texts but not to the training set) is assigned to some category \( D_i \) with some probability, i.e., for a text \( j \) in the test set, with word profile \( S_j \), its category is estimated through \( P_T(D_i|S_j) \) for \( i=1, 2, ..., k \). Then, the aggregated distribution of opinions \( P(D) \) of all texts in the corpus is obtained by aggregating individual classifications, each with its own misclassification error. As a result, the individual misclassification error do not vanish due to aggregation but may easily propagate up to the extent that, in many applications with thousands or millions of texts, one could see the error raising up to 15-20%. This is clearly quite problematic if one is mainly interested in estimating some kind of aggregate measure through the analysis of social media, as it happens with all the researches that wants to map, somehow, tweets into votes.

HK theory is effective in that it reverses the previous approach and, instead of estimating the individual opinion and the aggregating, it aggregates all word profiles and estimates the aggregated distribution of opinion directly, leading to an error of the order of 2-3%. Accordingly, being an approach for analyzing texts that does not aim to classify the individual documents into categories, but to measure directly the proportion of documents in each category, represents the second (statistical) crucial advantage of the HK method (see also the discussion in Grimmer and Stewart 2013). More in details, the frequency distribution of the terms \( P(S) \) can be expressed as:

\[
P(S) = P(S|D) P(D) \quad (2)
\]

The frequency distribution \( P(S) \) can be evaluated tabulating all the texts posted, and it requires only some computer time and no debatable assumption. The conditional distribution \( P(S|D) \) cannot be observed, and must be estimated by the hand-coding of the training set of texts. The hand-coding of the training text, in fact, allows for calculating \( P_T(S|D) \), i.e. the conditional frequency distribution of word profiles inside the training set. The assumption – and the reasonable requirement – of the method is that the texts of the training set are homogeneous to the whole data set, i.e. they come from the same “world” the rest of the dataset comes, such that one can assume that:

\[
P_T(S|D) = P(S|D) \quad (3)
\]
If this is the case, the frequency distribution of the opinions can be consistently estimated, because both \( P(S) \) and \( P_T(S|D) \) are observable. Therefore, by equation (2) and noticing that \( P_T(S|D) \) and \( P(S|D) \) are both matrices, we have

\[
P(D) = P(S|D)^{-1} P(S) = P_T(S|D)^{-1} P(S) \quad (4)
\]

where \( P_T(S|D)^{-1} \) is the inverse matrix of \( P_T(S|D) \), similarly for \( P(S|D)^{-1} \).

It is worth remarking that – while the homogeneity of the training set to the dataset is required – no statistical property must be satisfied by the set: in particular, the training set is not a representative sample of the population of texts.\(^{10}\)

In all the cases discussed below, we analyze social media (in particular Twitter: see below) to try to predict the final electoral result. In this respect, \( P(D) \) refers to the propensity to vote for each candidate/party.\(^{11}\) We considered in particular a tweet as casting a “vote” for a candidate/party only if at least one of the following three conditions is satisfied: a) the tweet includes an explicit statement related to the intention to vote for a candidate/party; b) the tweet includes a statement in favor of a candidate/party together with an hashtag\(^{12}\) connected to the electoral campaign of that candidate/party; c) the tweet includes a negative statement opposing a candidate/party together with an hashtag connected to the electoral campaign of another candidate/party. Considering not simply a generic positive statement, but a positive statement plus an hashtag permits to focus on those signals that by being more “costly” in terms of self-exposition by the Twitter user (including an hashtag in a message denotes after all a clear stance) are also more credible (on this point, see the large literature on signaling games: Banks 1991). On the other hand, condition c) allows to reduce the arbitrariness in the “supervised” stage of the analysis. This applies also to a (largely) two-candidate case such as the U.S. Presidential race. For example, if a tweet says “do not vote for Romney” this does not necessarily imply that the person who wrote that post will then vote for Obama. He could decide to vote for a third candidate or also to abstain. In a multi-party race, of course, this problem just strengthens. However, going back to the previous example, an hypothetical tweet such as “do not vote for Romney. #fourmoreyears” would be counted, according to our classification, as a vote in favor

\(^{10}\) Besides, while classic web analyses allow only distinguishing between positive and negative references to a particular topic, the HK method enables to catch also the reasons behind such opinions expressed: as an example, the opinion on a new party can be “I like that party’s leader, but his party is too radical for me”, or “Their manifesto is just perfect to me. The problem is that party will never pass the electoral threshold”. The human-coding phase can identify all these categories; the subsequent automatic classification estimates their relative weight in the overall opinion distribution.

\(^{11}\) In all the analysis we also considered the categories “Others” and “Uncertain”.

\(^{12}\) An hashtag is a word or a phrase prefixed with the symbol # that provides a means in Twitter of grouping all the messages including that word or phrase. Through that, one can search on-line for the hashtag and get the set of messages that contain it.
of Obama, given that #fourmoreyears has been one of most largely used hastag supporting the Obama’s electoral campaign.

Similarly, we counted all the retweets (i.e., a message rediffusion by a Twitter user of a message posted by another user) that satisfy the previous conditions as a “vote” for a candidate/party. Despite retweeting, strictly speaking, does not imply the production of new information, it implies that someone else thought a communication was valuable for herself (Jensen and Anstead 2013). On the other side, if it is true that the act of re-tweet does not necessarily imply an “endorsement” by the user that re-tweets, it is also true that when the re-tweet includes a text in which an intention to vote a given candidate/party is clearly expressed or where an identifiable hashtag connected to a candidate/party is presented, it becomes a costly act, exactly for the same reasons already noted above. As a consequence, it should happen only when a Twitter user shares to a large extent the content of the tweet and the underlying connected “vote”.

Broadly speaking, there are several social-media that could be analyzed. Here we will focus, as already noted, on Twitter, a social network for microblogging (Jansen et al., 2009) that experienced a sharply growth in the last years. When we come to the countries analyzed in this work, we observe that in 2012 Twitter was the second most-used social network in the United States (around 23 millions of users) and Italy (5 millions) and the third most-used in France (6 millions). A further crucial advantage of Twitter, that makes it so popular in the literature on social-media analysis, is that all the posts by users can be freely accessible, contrary to other social networks. Moreover, it also allows to geo-localize the origin of the tweet, therefore permitting a more fine-grained analysis (as we will see below). To download the data employed in the present paper we have relied on the social media monitoring engine Voices from the Blogs (http://www.voicesfromtheblogs.com/)\textsuperscript{13}, while the analysis have been run in R.

2. Electoral campaign and social media (1): the 2012 French Legislative election

The first political context in which we explore the usefulness of a social media analysis concerns the first round of the 2012 French Legislative election, held on 10 June 2012. Given the large number of parties competing in that election, this forecasting exercise represents a rather ambitious exercise. We gathered in this respect 79,300 tweets released during the last week before the elections to predict the national share of votes of the main parties. We then applied to those tweets the HK procedure discussed above.

As shown in Figure 1 below, at the national level our prediction is close to the actual results. This is true for almost every party. In particular we made an accurate forecast concerning UMP, Greens, minor moderate parties and, to a lesser extent, the Socialist Party. On the contrary,

\textsuperscript{13} The population of tweets collected consists of all the tweets posted during the temporal-period considered (see below) that include in their text at least one of a set of keywords (the name of the political leaders/parties covered by each of our analysis in France, Italy and the United States, as well as the most popular hashtags characterizing each candidate/party’s electoral campaign).
we overestimated far left parties (FdG, NPA and others) while the vote share of National Front has been underestimated. A possible explanation for such discrepancy may be linked to the presence of FN voters (more) reluctant to publicly express their voting behavior on-line due to some kind of “socially desiderability” bias (Knigge, 1998), a factor that seems to explain why normally also electoral-surveys tend to underestimate the support for the extreme-right in France (Jerome et al., 1999; Durand et al., 2004). Similarly, it could be that the far-left parties have tended to be the ones more heavily affected by strategic voting, so that a (radical) left-wing internet user has expressed her sincere preference on-line, but not at the polls (and indeed the moderate-left parties in Figure 1 tends to be slightly under-estimated overall). Albeit in a run-off electoral system, such as the one applied in the French Legislative election, the incentives to vote strategically are stronger in the second round, they are not absent also in the first round (Cox, 1997).

Note that such incentive to express a sincere vote on-line and then to vote differently does not exist by definition when we have just two main parties/candidates running at the polls. Therefore, in such case we expect ex-ante a better result in terms of our forecast (see below).

That noted, on average the Mean Absolute Error (MAE) of our prediction remains quite low being equal to 2.38%, which is not far from those displayed by the surveys held during the last week before the elections. On average survey polls registered a MAE equal to 1.23%, ranging from 0.69% to 1.93%.

The data on the French Legislative election allow also to explore a further point, linked to the possibility to assess the main sources of bias that may alter the accuracy of our prediction. In this respect, we use data about local constituencies. We exploited the geo-tagging service made available through Twitter to gather preferences within 13 local areas: Bordeaux, Dijon, Le Havre, Lille, Lyon, Marseille, Montpellier, Nice, Rennes, Saint Etienne, Strasbourg, Toulouse, Toulon. Then we ran 13 analysis to get social-media prediction within each area and we compared such estimates with the actual results in the 46 local districts connected to those cities. We measured the MAE, which represents our dependent variable and varies between 2.7 and 8.2. Then we tried to assess which elements increase or decrease the MAE of our prediction. We have estimated two different models. In the first one, we include our main independent variables: Number of Tweets, the number of comments released in each area, which expresses the

---

14 The paradoxical result of the first round of the French Presidential election in 2002 (when the Socialist candidate was not able to get access to the second round because too many extreme-left voters voted sincerely for small, but losing, radical candidates: see Laver et al. 2006) could have taught such lesson to the French extreme-left electorate.

15 The percentage of geotagged tweets is usually a portion of the overall number of tweets posted every day. As such, the sample we drew upon was necessarily global in nature, while not being necessarily representative of those using Twitter.

16 We excluded Paris due to its broad size that makes it harder to establish a link between the origin of each post and the electoral districts existing in the city.
information available, and *Abstention*, the percentage of district voters who decided to abstain, together with three control variables: *Le Pen Votes Share*, the share of votes gained in the district by the far-right candidate during the 2012 Presidential elections (used to identify those areas where the extreme right is strongest); *Mélenchon Votes Share*, the share of votes gained in the district by the candidate of the Front de Gauche, during the 2012 Presidential elections (as a proxy for the ‘red’ districts), and *Incumbent*, a dummy variable equal to one when the incumbent MP is running to seek the re-election. We added *Le Pen Votes Share* and *Mélenchon Votes Share* to control for the tendency of our analysis (as depicted in Figure 1) to underestimate the far-right and overestimate the far-left. With respect to *Incumbent*, on the other side, it has been argued that elections are referenda on the incumbent (Freeman and Bleifuss, 2006) and these candidates may outperform in the pre-electoral surveys compared to the actual results due to name recognition, affecting therefore positively the value of our dependent variable. Table 1 reports the results.

Table 1

From Model 1 we can observe that any growth of the information available on-line improves our predictive skills. For instance, an increase of 1,000 tweets analyzed lowers our error roughly by a quarter point. Conversely, the MAE is greater when the *Abstention* growths (the same concerns affects traditional off-line pre-electoral polls: see Crespi, 1988). This could happen because some citizens can easily express their opinion on-line though refusing to cast a vote (perhaps because they feel that their choice will not alter the results or because after all the act of voting is costly: Downs, 1957). Social media analysis then seems less able to provide accurate predictions when voters tend to abstain at a higher rate (a 10% increase in *Abstention* raises MAE by 1.2 additional points), while the accuracy should be greater when forecasting elections with a higher turnout. With respect to our control variables, neither of them appears to be significant.

In Model 2, we test the possible interacting effect of *Number of Tweets* and *Abstention*, to assess whether the effect of having additional information about citizens’ preferences is conditional on the likelihood that citizens do actually cast a vote. As can be seen, the coefficient of the interaction term is highly significant and allows us to depict a more fine grained relationship lying behind our model. In this respect, Figure 2 reports the marginal effect of *Number of Tweets* as the level of *Abstention* increases\(^\text{17}\): as it clearly appears, having more information on citizens’ voting preferences decreases the error when the turnout rate is sufficiently high (higher than 55%, a common finding in most elections). Up to such threshold, our predictive skills are enhanced by any increase in the number of comments about voting intentions, and such effect enlarges as turnout growths (i.e., *Abstention* decreases). Conversely, when voters tend to abstain at a really large rate (i.e., an abstention rate higher than 45%), having

\(^{17}\) We also superimpose a histogram portraying the frequency distribution for *Abstention* (the scale for the distribution is given by the vertical axis on the left-hand side of the graph)
more information about their (declared) voting choice ceases to affect significantly the accuracy of our predictions. Finally, at admittedly rather extreme values of Abstention, at least for Western-European democracies (i.e., an abstention rate higher than 55%), increasing Number of Tweets positively affects our error: in this case we face a kind of “information overload”, i.e., given that voters would express themselves on Twitter instead of casting a real vote, the MAE tends to increase. This finding is quite interesting given that it clearly shows the strict relationship incurring between what happens on-line and off-line in terms of our ability to extract reliable measures from social media analysis.

Figure 2

3. Electoral campaign and social media (2): the selection of the centre-left coalition leader in Italy

We double-checked the predictive skills of a social media analysis by applying this technique to forecast the first and second rounds of the Primary elections held by the centre-left coalition “Italia Bene Comune”, to select the leader of the alliance. This is an intriguing exercise, given the particularly complex environment of a primary election, that makes the possibility of a forecast particularly burdensome (AAPOR 2009): primary elections are indeed a contest that involve typically a partisan electorate, a larger number of viable candidates, with less ideological differentiation than one finds in the general election (a fact that makes more costly the electoral choices by voters), and lower turnout (Jensen e Anstead 2013). The fact that in Italy the primary elections are not legally recognized as such makes things just harder.

From October 6th to November 25th (the day the election was held) we analyzed no less than 500,000 tweets on several points in time to assess the voting intention of internet-users toward the 5 different candidates: Pierluigi Bersani (head of the Democratic Party), Matteo Renzi (Mayor of Florence), Nichi Vendola (governor of Apulia and leader of the left-wing party Left Ecology and Freedom), Laura Puppato (Democratic Party whip in Veneto), and Bruno Tabacci (Milan budget councillor). Figure 2 reports the fluctuation of voting intentions according to our analysis. We reported our estimates measured in 10 different days by analyzing around 40,000-50,000 tweets released in a time spam that ranges from 10 days (at the beginning of the electoral campaign) to one week (since Nov. 12th). 18

Figure 3

Figure 3 allows to monitor how the distribution of preferences has changed over the campaign, as well as the different “momentums” that characterized such campaign. To start with,

18 The results of this analysis have been published on the home-page of the Italian newspaper Corriere della Sera and are still accessible from the following link: http://www.corriere.it/politica/speciali/2012/primarie-centrosinistra/document_testatac.shtml
Bersani expected votes share has been always around the 40% while we observed an increase during the last week, when the voting intentions grew up to 47.6% (note that a similar trend was reported also by several polls). However in the last two days before the elections this value shrinks back to 43%, closer to the actual result (we mistaken by 1.9% only). Renzi was second-ranked always second since the beginning. His support has been on average around 31%, with a peak on Nov. 12\textsuperscript{th}, when the candidates were involved in a debate on SKY television, followed by a loss after the debate. Then his expected votes share started to grow again after the convention (called “Leopolda 2012”) held in Florence by Matteo Renzi and his supporters. Nichi Vendola, who actually ranked third, started his campaign only in November after he was discharged from a prosecution.\textsuperscript{19} This combination of events attracts him an initial large share of support, which declined in a few days when such effect vanished, bringing Vendola expected votes share around 18%. Finally, the two minor candidates, Puppato e Tabacci, retained only a low share of votes (according to our forecast) during the whole campaign, except after the debate when they took advantage from an outstanding public visibility.

Figure 3 also reports the actual votes share and highlights in red, per each candidate, the absolute difference between our last forecast and the final results. The gap between our estimates and the results is very narrow being on average below 2%. This error is in line with the average error provided by the surveys polls issued in the last week, which is 1.9%. This also appears from Figure 4 where we display, per each candidate, the absolute error of our prediction along with those of polls. What is more, our technique succeeded in predicting the gap between the two foremost candidates, Bersani and Renzi, better than the traditional survey polls. Indeed, according to our results, the gap between the two candidates was 10.5% while Bersani was leading by 9.4 points after votes have been counted. This means a difference of 1.1% while on average the polls mistaken the magnitude of the gap by 3%.

We were able to produce a similar accurate forecast of the Italian centre-left Primaries also in the second round. Table 2 below reports the actual results, our forecast by using Twitter and the results according to several surveys published in the last week before the second round. In this case, our analysis on almost 25 thousand tweets posted between Thursday 29\textsuperscript{th} of November and Saturday 1\textsuperscript{st} of December (the night ahead of the second round) predicted a clear victory for Bersani (58.4% vs. 41.6% for Renzi). At the ballot, Bersani won with 61.1% of votes against 38.8% for Renzi.

\textsuperscript{19} Note that on October 10th the campaign for the Primary elections was not officially started yet. Therefore our estimates included also some potential candidates who later decided to do not run for nomination.
According to these results social network sites confirm themselves as sources of valuable information that can be exploited to carry out electoral forecasts. However the goodness of such forecasts seem also to depend on the technique adopted. Indeed, several social media analyses conducted on the same Primary elections, but considering just the volume of mentions by candidates, were clearly (and consistently) highlighting a strong advantage for the actual second-ranked candidate, that is Renzi, in both the first and the second round.\(^{20}\) In this sense, the way we choose to analyze social media seems indeed to make a difference. This is a point on which we will go back, even more clearly, in the next section.

4. Electoral campaign and social media (3): the 2012 US Presidential election. Too close to call?

The last attempt to adopt social media as devices to predict the outcome of an election has been carried out with respect to the 2012 US Presidential election. From the September 28\(^{th}\) to November 6\(^{th}\) we monitored the voting intention expressed on Twitter towards the four main candidates: Barack Obama (Democratic Party), Mitt Romney (Republican Party), Gary Johnson (Libertarian Party), and Jill Stein (Greens). In this lapse of time we estimated the political preferences of American voters on a daily basis by analyzing more than 50 millions of tweets, a bit more than 1 million of tweets per day.\(^{21}\) Each data is calculated as a moving average along those 7 days.\(^{22}\) These results are summarized in Figure 5.

The fluctuation of the preferences expressed on-line closely follows the main events that happened during the electoral campaign (note that this link between the political agenda and the shifts in public opinions is similar to those observed during the campaign for the centre-left nomination). For instance, while in September Obama retained a wide margin over his main opponent, the wind started to change at the beginning of October, when Romney was able to reduce the distance from the Democratic candidate. As Obama performed poorly during the first TV debate, Romney overcame the former President in the voting intention of Twitter-users (note that traditional survey polls revealed a similar trend only in the following days, given the relatively slow pace of polling). The second debate represents another turning point that arrested the loss of support for Obama. In the last days before the election some political scandals (like the “Benghazi gate” or the statement against the abortion in case of rape, pronounced by the


\(^{21}\) Even the results of this analysis have been published on a daily basis on the home-page of the Italian newspaper Corriere della Sera throughout the American electoral campaigning.

\(^{22}\) Applying a shorter moving average (i.e., 3 days) does not change qualitatively any of our results.
member of the “Tea Party”, Richard Mourdock and exogenous events (the Sandy hurricane) jumped into the campaign wielding advantages to one candidate or the other.

In the very last days the on-line sentiment highlighted a positive trend toward Obama, and our final prediction made on November 6th forecasted a victory for Obama in the popular votes with a clear and safe margin of 3.5%. As the real gap in the share of votes was 3.9%, our forecast proved to be more accurate than those made by traditional survey polls that on average assigned only a narrow margin in favor of Obama (+0.7) claiming that the race was “too close to call”.

Beside the popular vote, we also tried to predict the results in the ‘swing states’, i.e., those where the race is usually very close and few thousands of votes can alter the balance between the candidates and the outcome of the whole Presidential election. In 2012 the surveys focused on 11 ‘swing states’ and among these they considered Florida, Ohio, and Virginia as the main battlegrounds. We therefore paid attention to those races. In Table 3 we report the gap between Obama and Romney in each state, according to three different measures. The first one (S) consists in the method of sentiment analysis discussed so far. For the “swing states” estimates, we replicated our method considering the pools of tweets geotagged in each State only. The second (R) is the gap displayed on November 6th on the web site www.realclearpolitics.com, which recorded the average of the survey polls issued on the last week before the election. The third one (V) represents the actual gap between the two candidates after votes have been counted. Then we display the difference between the forecasts (made either through sentiment analysis or surveys) and the actual votes. Finally, we highlight what prediction has been the best one according to the ability to correctly predict the winner. When both sentiment analysis and survey polls predicted the same winner, we discriminate by measuring the difference between the expected and the actual gap.

Table 3

---

23 The Benghazi gate concerns to the murder of the US ambassador in Libya. According to the Republicans, Obama may have omitted the truth when talking about this event in public speeches. On the other side, Richard Mourdock, who was running for the Senate in Indiana for the Republican party, said that abortion should be neglected in case of rape, because even that birth is a God’s will.

24 See for instance the interesting article written by Andrew Gelman, director of the Applied Statistics Center at Columbia University: http://campaignstops.blogs.nytimes.com/2012/10/30/what-too-close-to-call-really-means/?smid=tw-share
Overall, analyzing social media correctly predicts the winner in 9 out of 11 swing states, Colorado and Pennsylvania being the only two exceptions. Furthermore, in a plurality of states (7 against 2), our data \((S)\) proved to be more accurate than the average of polls (these states are: Florida, Iowa, Virginia, Nevada, New Hampshire, Michigan, Wisconsin), while in the remaining two swing states (Ohio and North Carolina) the different forecasts (social media vs. surveys) performed in a similar manner.

The most interesting results concern the three main battlegrounds: Ohio, Virginia and Florida. In Ohio both tools predicted similar results. On the contrary, in Virginia our data proved able to catch the trend pro-Obama the emerged in the last days (and in the last hours) when the Democratic staff mobilized the partisan voters (even during the electoral night they pushed voters to stay in line by means of Twitter messages). The same happened in Florida, where our prediction was the only one to claim a victory for Obama, with a safe margin. Eventually these results could be explained by our ability to measure the voting intention of the Hispanic voters who could be less likely to answer to survey polls.

Two further issues deserved to be highlighted with respect to the US Presidential race. First of all, it is interesting to note that at the beginning of the electoral campaign, Obama had almost 16.8 million followers on Twitter, while Romney hadn’t even hit 600,000. Despite such (huge) disparity, our results underlined a different story, that was not only remarkably in line with the actual votes, as we have discussed, but that also illustrated a social media support for the two main competitors that was much more volatile compared to what we could have expected by looking at the number of followers only (see Figure 5 above). In this sense, our results confirm that the number of Facebook friends or Twitter followers on their own are largely misleading as predictors of election outcomes (see Cameron et al. 2013 on this point). This happens also because Twitter users are often about divided between those who follow leaders they agree with and those who also follow political figures they disagree with (see Pamelee and Bichard 2011).

Second, while the HK supervised method for sentiment analysis adopted throughout this paper was able to catch the voting intention of US citizens, other computerized tools like the Twitter Political Index (Twindex: [https://election.twitter.com/](https://election.twitter.com/)) developed by Topsy and Twitter itself, estimated the day ahead of the election a wide margin for Obama (+15 points) in terms of positive sentiment compared to Romney. Besides, throughout the entire electoral campaign, according to Twindex Obama was almost always leading, with a rather comfortable overall average lead (i.e., more than 6 points). It is worth to mention that Twindex was constructed for the two candidates as the proportion of the total number of positive tweets versus the total number of positive and negative tweets (contrary to our index). Besides, to determine positive and negative feelings, Topsy used ontological dictionaries rather than a supervised method. In this sense, both aspects could probably explain the different results between our analysis and the ones (less accurate, at least when compared to the final electoral results) provided by Twindex.\(^{25}\)

\(^{25}\) Another difference is that Twindex analyzed less data than what done here (around 2 million tweets a week). URL: [http://usatoday30.usatoday.com/news/politics/twitter-election-meter](http://usatoday30.usatoday.com/news/politics/twitter-election-meter). The
Conclusion

The growing usage of social media by internet users, who express their opinions on a wide variety of topics, has raised the interest about the opportunity to exploit this information to better understand the link between preferences and political behavior. Not surprisingly, in the last years we assist a dramatically increase in the number of works that analyzes social media in order to assess the opinions of internet users and to check whether the attitudes expressed on-line can be eventually used to forecast the voting behavior of the whole population of voters. For all these reasons, being able to rely on techniques apt to measure on-line public opinion becomes a pressing topic.

In this paper we have applied in three (very) different political scenarios a statistical method recently introduced in the literature that performs a supervised sentiment analysis on social networks, and that improves on traditional methods yielding more accurate results. From the results of our empirical analyses, all conducted before elections occurred, we can raise some general claims. First of all, our analysis shows a consistent correlation between social media results and the ones we could obtain from more traditional mass surveys as well as a remarkable ability of social media to forecast electoral results (a so careful prediction that could not be due simple to chance). This seems to be true for both “single issue” elections (such as Presidential race), in which the preference eventually expressed by an internet user involves only a choice among two single options, as well as for more difficult situations to forecast such as the ones in which internet users can choose to express a preference among a (large) number of different targets (such as political leaders or political parties). This – together with the fact that the political scenarios here analyzed come from three different countries (Italy, France, and the United States) in which the traits of internet-users are not supposed to be necessarily identical – is clearly important for the robustness of our results. True: sentiment analysis of social media seems to provide more accurate predictions when focusing on the most popular leaders or on mainstream parties, but on average our results seem very promising for future research.

Why does this happen? After all, to forecast election we need to rely on a representative sample, and there is no guarantee that this is something that can be obtained by analyzing social media. On the contrary, socio-economic traits of social media users do not exactly match the results of Table 3 are also more accurate than the ones reported in Choy et al. (2012), where the authors in an ex-ante analysis of the U.S. Presidential Election weight the results of their fully-automated sentiment analysis by some census information to correct the possible bias in the online data. In particular, they consider the pre-existing party affiliations of the American voters. Note that this is a problem also for standard off-line surveys as the poll rates keep falling dramatically in recent years, thanks to mobile phones, caller identification and a rise in phone
actual demographics of the whole population (Tjong Kim Sang and Bos 2012, Wei and Hindman, 2011, Bakker and de Vreese, 2011): people on social media are generally younger and more highly educated, concentrated in urban areas, as well as more politically active overall (Conover et al. 2011; Jensen, et al. 2012). But do we need a representative sample when, for example, 22% of voters spontaneously declared their voting behavior on social network sites, as it happened during the U.S. Presidential campaign (Pew Research 2012b)? Perhaps the sheer magnitude of data available on social media, i.e., the “wisdom of crowds” (Franch 2012), may compensate for this partly unrepresentative information. After all, the crowd to be wise needs to be diverse, independent, while its decisional procedure has to be decentralized (Surowiecki 2004). And this is something that is usually attained in the Big Data world.

Moreover, to cast an accurate forecast we should be more worried about the distribution of political preferences on the web. Previous (albeit quite dated) analyses showed that left-leaning are over-represented, though only marginally (Best and Krueger 2005). We clearly need more (updated) analysis in this regard. Still, some of the potential bias that arise from social media analysis may be softened in the medium (short?) run, as the usage of social network increases: as we have shown, when a growing number of citizens’ express on-line their opinion and/or voting choice, the accuracy of social media analysis increases, provided internet-users act consistently on that (by, for example, confirming their (declared) on-line preference by casting a (real) vote) (on this point see also Ceron et al. 2013, Lenhart et al. 2010). For example, Table 4 below compares the distribution of the ideological self-placement of Italian voters in February 2012 and the sub-sample that declares to be active on social-media. As can be seen, the difference among the two samples is quite trivial.

| Table 4 |

Finally, although the social media population, so far, is not still always representative of one country’s citizenry, there are still some doubts about whether such bias could affect the predictive skills of social media analysis. Indeed, the latter aspect (the predictive skills of social media analysis) does not necessarily need the previous factor (i.e., the issue of representation) to hold true to effectively apply. This can happen, for example, if we assume that internet users act like opinion-makers that are able to influence (or to “anticipate”) the preferences of a wider solicitation, and also the difficulties reaching many population segments (Pew Research Center 2012a, Tourangeau and Plewes 2013, Hillygus 2011, Goidel 2011).

27 Accordingly, one way to improve the social media forecast would be to develop an appropriate set of weights based on the representativeness of certain groups of users (Choy et al. 2011; but see note 26), or, even better, according to the political preferences of social media users, provided this type of information is available (and reliable).
audience (O’Connor et al., 2010), including the ones of the broader media ecosystem\textsuperscript{28}, or to reproduce it as an all. For example, this can be true if Twitter communications are considered to function like a critically engaged interaction system (Ampofo et al., 2011) whose communications about specific issues (such as electoral contests) are thematically representative of larger currents of conversations and preference distributions (Jensen and Anstead 2013). This is clearly another fascinating topic that deserves a further investigation.

Summing up, despite the well-known limits and the troubles faced by social media analysis (Gayo-Avello et al., 2011; Goldstein and Rainey, 2010), our results provide reasons to be optimistic about the capability of sentiment analysis to become (if not to be already) a useful supplement of traditional off-line polls. But the method do matter: Shi et al. (2012) noted that “merely using the volume of tweets […] is not enough to capture public opinions. We need to come up with some sophisticated algorithm and model to make the prediction successfully”. As we have argued along this paper, aggregate supervised techniques, such as the one advanced by HK, seem to grant more accuracy than old-fashioned sentiment analysis precisely in this last respect.

\textsuperscript{28} The fact that quite often journalists are among the most active consumers of social media (Spierings e Jacobs 2013, Lasorsa et al. 2012) could provide an empirical ground to this hypothetical claim.
References


Ceron, Andrea, Curini Luigi, Stefano M. Iacus, Giuseppe Porro (2013). Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens’ policy preferences. An application to Italy and France, New Media & Society, DOI: 10.1177/1461444813480466


Conover, M , B Goncalves, J Ratkiewicz, A Flammini, and F Menczer (2011). Predicting the Political Alignment of Twitter Users. Proceedings of 3rd IEEE Conference on Social


Madge C, Meek J, Wellens J and Hooley T (2009) Facebook, social integration and informal learning at university: It is more for socialising and talking to friends about work than for actually doing work. Learning, Media and Technology 34(2): 141–155.


Tables

Table 1. *OLS regression of Mean Absolute Error*

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of Tweets</strong></td>
<td>-0.000234**</td>
<td>-0.004339***</td>
</tr>
<tr>
<td></td>
<td>(0.000108)</td>
<td>(0.001354)</td>
</tr>
<tr>
<td><strong>Abstention</strong></td>
<td>0.116903*</td>
<td>-0.227582*</td>
</tr>
<tr>
<td></td>
<td>(0.058571)</td>
<td>(0.125282)</td>
</tr>
<tr>
<td><strong>Number of Tweets X Abstention</strong></td>
<td>-</td>
<td>0.000091***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.000030)</td>
</tr>
<tr>
<td><strong>Le Pen Votes Share</strong></td>
<td>0.012887</td>
<td>0.001895</td>
</tr>
<tr>
<td></td>
<td>(0.038121)</td>
<td>(0.034903)</td>
</tr>
<tr>
<td><strong>Mélenchon Votes Share</strong></td>
<td>-0.027611</td>
<td>-0.039590</td>
</tr>
<tr>
<td></td>
<td>(0.104928)</td>
<td>(0.095634)</td>
</tr>
<tr>
<td><strong>Incumbent</strong></td>
<td>-0.383584</td>
<td>-0.635201</td>
</tr>
<tr>
<td></td>
<td>(0.460141)</td>
<td>(0.427129)</td>
</tr>
<tr>
<td><strong>Competiveness</strong></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>1.618636</td>
<td>17.68222</td>
</tr>
<tr>
<td></td>
<td>(2.828653)</td>
<td>(5.880315)</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td><strong>R-squared</strong></td>
<td>0.210</td>
<td>0.361</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 2. *Comparison of the accuracy of Twitter forecast and surveys polls in the Italian Primary election of the centre-left coalition (second round)*

<table>
<thead>
<tr>
<th></th>
<th>Day of publication of the survey</th>
<th>Bersani</th>
<th>Renzi</th>
<th>Gap</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Popular Vote</strong></td>
<td>-</td>
<td>61.1</td>
<td>38.8</td>
<td>Bersani +22.3</td>
</tr>
<tr>
<td><strong>Twitter</strong></td>
<td>12/1/2012</td>
<td>58.4</td>
<td>41.6</td>
<td>Bersani +16.8</td>
</tr>
<tr>
<td><strong>Ipsos</strong></td>
<td>11/29/2012</td>
<td>57.5</td>
<td>42.5</td>
<td>Bersani +15</td>
</tr>
<tr>
<td><strong>Quorum</strong></td>
<td>11/28/2012</td>
<td>56.4</td>
<td>43.6</td>
<td>Bersani +12.8</td>
</tr>
<tr>
<td><strong>SWG</strong></td>
<td>11/28/2012</td>
<td>55</td>
<td>45</td>
<td>Bersani +10</td>
</tr>
<tr>
<td><strong>COESIS</strong></td>
<td>11/28/2012</td>
<td>54</td>
<td>46</td>
<td>Bersani +8</td>
</tr>
<tr>
<td><strong>ISPO</strong></td>
<td>11/27/2012</td>
<td>56.5</td>
<td>43.5</td>
<td>Bersani +13</td>
</tr>
<tr>
<td><strong>IPR</strong></td>
<td>11/26/2012</td>
<td>56</td>
<td>44</td>
<td>Bersani +12</td>
</tr>
<tr>
<td><strong>PIEPOLI</strong></td>
<td>11/25/2012</td>
<td>59</td>
<td>41</td>
<td>Bersani +18</td>
</tr>
</tbody>
</table>
Table 3. US Presidential 2012: Accuracy of the predictions. Comparison between VfB (S) and survey polls (R) estimates with the actual results (V)

| State                | Gap (S)     | Gap (R)     | Gap (V)     | |S-V| |R-V| | Best prediction |
|----------------------|-------------|-------------|-------------|----------|----------|----------|-------------|
| Popular Vote         | Obama +3.5  | Obama +0.7  | Obama +3.9  | 0.4      | 3.2      | S        |
| Florida              | Obama +6.1  | Romney +1.5 | Obama +0.9  | 5.2      | 2.4      | S        |
| Ohio                 | Obama +2.9  | Obama +2.9  | Obama +3.0  | 0.1      | 0.1      | =        |
| Virginia             | Obama +3.5  | Obama +0.3  | Obama +3.9  | 0.4      | 3.7      | S        |
| Colorado             | Romney +1.3 | Obama +1.5  | Obama +5.4  | 4.1      | 3.0      | R        |
| Iowa                 | Obama +4.8  | Obama +2.4  | Obama +5.8  | 1.0      | 3.4      | S        |
| Nevada               | Obama +3.3  | Obama +2.8  | Obama +6.7  | 3.4      | 3.9      | S        |
| New Hampshire        | Obama +3.8  | Obama +2.0  | Obama +5.6  | 1.8      | 3.6      | S        |
| North Carolina       | Romney +3.0 | Romney +3.0 | Romney +2.0 | 1.0      | 1.0      | =        |
| Michigan             | Obama +5.5  | Obama +4.0  | Obama +9.5  | 4        | 5.5      | S        |
| Pennsylvania         | Romney +2.5 | Obama +3.8  | Obama +5.4  | 2.9      | 1.6      | R        |
| Wisconsin            | Obama +7.4  | Obama +4.2  | Obama +6.9  | 0.5      | 2.7      | S        |

Table 4. Distribution of ideological self-placement of Italian voters vs. sub-sample active on social-media

<table>
<thead>
<tr>
<th>Self-ideological placement</th>
<th>All sample</th>
<th>Sub-sample of social media users</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left</td>
<td>10.50</td>
<td>10.22</td>
</tr>
<tr>
<td>Centre-Left</td>
<td>15.75</td>
<td>15.25</td>
</tr>
<tr>
<td>Centre</td>
<td>14.63</td>
<td>13.81</td>
</tr>
<tr>
<td>Centre-Right</td>
<td>11.25</td>
<td>11.51</td>
</tr>
<tr>
<td>Right</td>
<td>4.13</td>
<td>4.32</td>
</tr>
<tr>
<td>None</td>
<td>37.50</td>
<td>38.42</td>
</tr>
<tr>
<td>Do not know/Do not answer</td>
<td>6.25</td>
<td>6.47</td>
</tr>
</tbody>
</table>

Source: IPSOS, February 2012
Figures

Figure 1. Predicted and actual vote shares related to the first round of the 2012 French Legislative elections

<table>
<thead>
<tr>
<th>Party</th>
<th>Twitter Preferences</th>
<th>Actual Votes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Far Left</td>
<td>0.98%</td>
<td>4.4%</td>
</tr>
<tr>
<td>Left Front</td>
<td>6.91%</td>
<td>14.9%</td>
</tr>
<tr>
<td>Socialist Party</td>
<td>4.2%</td>
<td>27.6%</td>
</tr>
<tr>
<td>Greens</td>
<td>5.46%</td>
<td>31%</td>
</tr>
<tr>
<td>Others Left</td>
<td>2%</td>
<td>3.4%</td>
</tr>
<tr>
<td>Modem</td>
<td>2%</td>
<td>1.76%</td>
</tr>
<tr>
<td>New Centre</td>
<td>1.4%</td>
<td>2.2%</td>
</tr>
<tr>
<td>Ump</td>
<td>29.6%</td>
<td>28.36%</td>
</tr>
<tr>
<td>Others Right</td>
<td>3.51%</td>
<td>3.51%</td>
</tr>
<tr>
<td>National Front</td>
<td>8.4%</td>
<td>15.6%</td>
</tr>
<tr>
<td>Others</td>
<td>1.7%</td>
<td>2.62%</td>
</tr>
</tbody>
</table>
Figure 2. Marginal effect of the Number of Tweets on the Mean Absolute Error (with 90% confidence interval)

Figure 3. Candidates share of votes according to Twitter forecast. Comparison with actual results
Figure 4. *Comparison of the accuracy of Twitter forecast and surveys polls in the Italian Primary election of the centre-left coalition (first round)*

<table>
<thead>
<tr>
<th>Polls</th>
<th>Predicted gap Bersani - Renzi</th>
<th>Real gap Bersani - Renzi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ipsos (19/11)</td>
<td>9.4%</td>
<td>9.2%</td>
</tr>
<tr>
<td>Ipsos (21/11)</td>
<td>8.4%</td>
<td>8.4%</td>
</tr>
<tr>
<td>SWG (22/11)</td>
<td>10.6%</td>
<td>10.6%</td>
</tr>
<tr>
<td>Toscot (25/11)</td>
<td>14.0%</td>
<td>14.0%</td>
</tr>
<tr>
<td>YB (25/11)</td>
<td>16.9%</td>
<td>16.9%</td>
</tr>
</tbody>
</table>

Figure 5. *US Presidential 2012: The trend of Twitter-votes for Obama and Romney*