THE ONLINE SOCIAL NETWORKS AS A TOOL FOR ESTIMATING VOTE

Fernando Moreno Reyes, PhD in Marketing.
Phone +34 678 50 55 29, Email:femoreno@ucm.es / fermoreno1@msn.com

University Complutense of Madrid, Spain.

Maria Puelles Gallo, PhD in Marketing.
Email: mpuelles@ucm.es

University Complutense of Madrid, Spain.

Alejandro Kress, PhD Candidate in Industrial Engineering.
Email: luis.alejandro.kress@estudiant.upc.edu

University Politecnica of Catalunya. Spain.

© Fernando Moreno Reyes
The online social networks as a tool for estimating vote

This research stems from the idea that raises the article written by Jon Gertner "Social Media as Social Index" (2010) from The 10th Annual Year in Ideas of The New York Times. This paper compares different studies using the messages launched by Twitter users and their consequences as social index. This fact is significant, but it is even more relevant that Cameron Marlow, head of Facebook data scientists’ team, sentences: "So it seems like it's improbable that the social media won’t be the way that we acquire opinion research". Gertner (2010) states that the Internet and the social networking sites could be very useful tools not only for market research, but for all areas of marketing and sales, also for estimating the vote of an electoral process.

The Social Networks and the Internet.

The social networks have existed since the beginning of mankind, the research about them is evident in the Theory of Six Degrees (Karinthy, 1930; Watts, 2004) and The Problem of Small World (Milgram, 1967). Mitchell (1973) grouped the two schools (the graph theory and the structural functionalism) that laid the foundations of social network analysis. With the emergence of Internet and the World Wide Web, the Search Marketing (search engine marketing or Internet marketing) (Boughton, 2005) and RSS (Really Simple Syndication files. Rss. xml to indicate or share content on the web) (Murugesan, 2007), there were numerous applications including the social networking sites, whose business plan is based on the theories previously formulated. Friendster was the first social network to appear in 2002, later on the Friendster’s structure was replicated by MySpace in 2003, Facebook in 2004 and Twitter in 2006.

According to O'Reilly (2004), the Web 2.0 is divided into communication applications (blogs [short for web log or web log], social networking, social networking aggregation), collaboration (wikis [websites that can be edited by users], social bookmarking sites news, podcasts [media file or dissemination of multimedia files on the Internet], forums), multimedia (video sharing, photo sharing, live casting [broadcast audio, video, live internet], audio and share music) and entertainment (virtual worlds, share games, online games). There are several platforms that work with these 4 features. The best known supports are Video Search (Gibbon & Liu, 2008), Podcasting (Edison Research, 2009), Blog (Nardi et al, 2004), Mobile Search Marketing (marketing or search engine marketing Internet on mobile phone
The online social networks as a tool for estimating vote

carrier) (comScore, 2011; eMarketer, 2010) and social networking / social media (O'Reilly, 2005; Boyd & Ellison, 2007).

These platforms are certainly optimal tools for communication and political marketing, because for a politician or a political party they bring several strengths and opportunities that do not offer other supports like flexibility, interactivity, more direct communication and the opportunity of monitoring the information received (Artusi & Maurizzi, 2010). It would be extremely naive to think that such applications or sites are used for communication among users only. They are companies that receive a lot of data, making the mining, analysis, extraction and use of these data their rationale. Edward Snowden revealed that the National Security Agency or NSA and the PRISM program were monitoring Facebook (NSA collecting phone records of millions of Verizon customers daily, 2013; NSA Prism program taps into user data of Apple, Google an others, 2013). These data are reversed in marketing goals or resold to third parties to perform similar actions, questioning the laws on privacy (O'Brien, 2012; Barrett, 2010; Andrews, 2012; Clemente, 2011).

Concluding the analysis of the social networks and their application to the Internet, it is concluded that they are closely related to the behavior, influence, image, productivity, conduct, results and in relation to various sectors such as anthropology, sanitation, geography or economics. This fact makes feasible the study of social networking sites as tools to influence political behavior. From the point of view of this research, these data could be useful to predict the estimated vote analyzing the users of these social networks during an election call, beyond the polls and the dynamics that the social network sites are. Twitter is a microblogging network (whose blog post has a maximum of 140 characters), and it seems very appropriate for this analysis. To perform this analysis it is mandatory to investigate whether social networking sites have the penetration necessary to carry out this work, and to review the literature linking the Internet and politics.

The Social Networks sites in Spain. Twitter.

The studies of ONTSI (2011), The Cocktail Analysis (2012; 2013), IAB (2012; 2013) were analyzed. The conclusions are that Spain is one of the leading countries in the use of social networks and Twitter is the leader microblogging network; also Twitter is the third most used social network in Spain and the third preferred social network to be used in the
The online social networks as a tool for estimating vote

future. Twitter is currently growing at 33% annual average; its purpose is to gain leisure/professional profiles based on personal/professional and microblogging profiles. Twitter is an open social network with horizontal level of integration.

Chart 1.

Using social networking sites 2014.

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>No Internet</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td>52%</td>
<td>33%</td>
<td>15%</td>
</tr>
<tr>
<td>USA</td>
<td>50%</td>
<td>29%</td>
<td>21%</td>
</tr>
<tr>
<td>Russia</td>
<td>50%</td>
<td>9%</td>
<td>42%</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>49%</td>
<td>30%</td>
<td>22%</td>
</tr>
<tr>
<td>Spain</td>
<td>49%</td>
<td>30%</td>
<td>21%</td>
</tr>
<tr>
<td>Poland</td>
<td>40%</td>
<td>17%</td>
<td>42%</td>
</tr>
<tr>
<td>Brazil</td>
<td>40%</td>
<td>9%</td>
<td>51%</td>
</tr>
<tr>
<td>France</td>
<td>39%</td>
<td>36%</td>
<td>25%</td>
</tr>
<tr>
<td>Italy</td>
<td>38%</td>
<td>24%</td>
<td>38%</td>
</tr>
<tr>
<td>Turkey</td>
<td>35%</td>
<td>8%</td>
<td>56%</td>
</tr>
<tr>
<td>Lebanon</td>
<td>34%</td>
<td>15%</td>
<td>51%</td>
</tr>
<tr>
<td>Tunisia</td>
<td>34%</td>
<td>7%</td>
<td>57%</td>
</tr>
<tr>
<td>Germany</td>
<td>34%</td>
<td>46%</td>
<td>20%</td>
</tr>
<tr>
<td>Mexico</td>
<td>33%</td>
<td>3%</td>
<td>63%</td>
</tr>
<tr>
<td>China</td>
<td>31%</td>
<td>18%</td>
<td>50%</td>
</tr>
<tr>
<td>Japan</td>
<td>30%</td>
<td>36%</td>
<td>34%</td>
</tr>
<tr>
<td>Egypt</td>
<td>30%</td>
<td>7%</td>
<td>63%</td>
</tr>
<tr>
<td>Greece</td>
<td>29%</td>
<td>19%</td>
<td>51%</td>
</tr>
<tr>
<td>Jordan</td>
<td>29%</td>
<td>6%</td>
<td>65%</td>
</tr>
<tr>
<td>India</td>
<td>6%</td>
<td>1%</td>
<td>89%</td>
</tr>
<tr>
<td>Pakistan</td>
<td>35%</td>
<td>3%</td>
<td>94%</td>
</tr>
</tbody>
</table>


Among the main activities of Twitter arises take part in advocacy / civic mobilization with a frequency of very common and quite often 18%. This activity did not exist in 2009 and 2010. The users use these networks as tools of influence and a way to express themselves with no censorship (12% of users), also as a vehicle for awareness / mobilization with a frequency of very common to quite often. These data confirms the importance of these tools as potential sources of rich and diverse information in reference to current political issues, political parties or electoral process. These results confirm the enormous potential as a tool for prediction and estimation of the vote for any electoral process, Twitter is the social network with more use as social mobilization tool to support a political call.
Therefore, the users of the social networks have diversified the main purpose of them, from a playful communication tool to share, to a communication tool where more and more users want not only interact or communicate, but to use them for two-ways communication that allows them to be present in a globalized world, influencing events directly, recommending and reviewing all kinds of global and local issues. This democratization of influence in the network is a potential market that is already exploiting various brands but not always with good results, a phenomenon that is seen in the different frequencies of use. It is also a potential tool for political parties for measuring the estimation of the vote, so very relevant findings are confirmed in these studies: the growth of the social networks as tools of mobilization, expression, social awareness and policy; and Twitter as the chosen support for them. These findings propel a literature review of the Internet and politics.

The Internet and Politics.

The advent of the era of information and technology in post-industrial societies has changed how to execute the actions of political communication. The formulation of concepts as Third Age of Political Communication and Third Age of Political Communication (Blumber & Kavanagh, 1999), Postmodern Campaign (Norris, 2000), Permanent Campaign (Nimmo, 1999), Politics as usual (Margolis & Resnick, 2000) and Postburocratic Policy (Bimber, 2003); summary that we are in a new era of political communication where campaigns become postmodern thanks to the Internet, we are always under a political campaign because the Internet is considered a policy tool for everyday communication. There is no doubt that these formulations show the change of the political communication experience. The new technologies influence the society in a more precise manner, delivering messages to a target audience increasingly segmented, activating and mobilizing. Therefore the marketing is more agile, one to one, where interactivity plays a fundamental role. So the political marketing and the technologies generate content, opinion; they could estimate the voting and policies trends from various technological tools.

The new tools of communication technology (SMS or Short Message Service messages), Internet (Web 2.0), social networking sites, ... make the political communication more flexible (Mylona, 2008; Ward, 2003; Lusoli, Ward & Gibson, 2006), one way or more direct way (Anduiza, 2009; Ward, 2003; Gibson, Ward & Lusoli, 2002), an evolution of the communication from one to many to many to many (Gibson & Rommele, 2008), with
simplicity of the message (Mylona, 2008), no editorial control (Ward, 2003), less top down and more bottom up (Mylona, 2008; Gibson, Ward & Lusoli, 2003; Gibson & Rommele, 2008), fostering the ability to mobilize/local and global protests (Lance Bennett, Breunig & Givens, 2008; Gibson, Ward & Lusoli, 2002; Artusi & Maurizzzi, 2010; Vaccari, 2008; Bimber & Davis, 2003), low cost (Mylona, 2008; Ward, 2003), accessible, closer to the young voter (Mylona, 2008; Blais, Gidengil & Nevite, 2004; Leppäniemi et al, 2010; Norris, 2003), more interactive (Norris, 2000; Farrel & Webb, 2000; Mylona, 2008), to promote e-discussion or discussion in the network (Bimber, 2001), propitiate less political influence in the media (Blumber & Kavannagh, 1999), conducive or cyber activism (Blumber & Kavannagh, 1999; Sherman & Schiffman, 2002; Gibson, Ward & Lusoli, 2003), allows transmit messages in a limited way or narrowcasting (Gibson, Ward & Lusoli, 2002). All these features are conditioned by the digital divide or penetration of this technology in the population (Mylona, 2008).

Within Democracy 2.0. and Government 2.0 (Dominguez, 2009), the new tools of communication technology, the Internet, the social networking sites; they not only have the characteristics or functions from above, but also they promote the digital democracy or e-democracy (Demertzis et al, 2005), reinforce the review (Bimber & Davis, 2003), allow microsegment or segment to influence the vote (Artusi & Maurizzzi, 2010; Leppäniemi et al 2010), raise funds or fundraising (Towner & Dulio, 2012; Bimber & Davis, 2003), capture and engage voters/engagement (Vaccari, 2008; Lusoli, Ward & Gibson, 2006; Leppäniemi et al, 2010), to persuade the undecided voter (Vaccari, 2008).

From a scientific point of view, the political communication in this context presents more opportunities than threats, but the reality is that politicians and political parties do not exploit the participatory power of the Internet, operating mainly through top-down communication strategies, low interconnectivity, wrong using of digital political communication and only to promote themselves; especially in the major parties (Anduiza, 2009; Lillekerd, 2010; Demertzis et al, 2005; Margolis & Resnick, 2000; Golbeck, Grimes & Rogers, 2010). This fact is a weakness against the youth vote because they are a very active segment of voters in the new technologies (Blais, Gidengill & Nevite, 2004; Leppaniemi, Karjalioto & Goman, 2010). But nevertheless, new technology, the Internet, the Web 2.0, the social networking sites are an opportunity for minority parties (Norris, 2003; Lillekerd, 2010; Gibson, Ward & Lusoli, 2002; Lusoli, Ward & Gibson, 2006), because they can reach more
people at minimal cost, especially young people, matching the major parties without the editorial control of the traditional media, mainly used by the major parties. The difference between the U.S. and Europe regarding this topic is clear, due to technological development and the European political pluralism.

Regarding the investigation of politics and the social networking sites, especially Twitter, there are several areas of research in which different models and research methods envision different outcomes. These fields or categories of research are influence, mobilization, sentiment, data, credibility, prediction and research on events that occurred in different countries and institutions. The research focused on the Influence (Subramani & Rajagopalan, 2003) concludes that social networking sites are more compelling and persuasive than personal interactions; they affect a number of individuals with minimal effort and more flexibility. Regarding mobilization, the online political mobilization works because the policy itself induces to the expression and also to the collection of information and validation of the vote, but may also discourage the voting (Bond, Fariss & Jones, 2012; Schmitt-Becka & Mackenrodt, 2010). Feeling or semantic analysis of data in the social networks (Bermingham & Smeaton, 2012; Bollen, Pepe & Mao, 2010; O'Connor et al, 2010) concludes that the politicians and the political, cultural and economic events are correlated with significant levels of mood, although the techniques and analysis must be substantially improved. André et al (2012) argue that the happy feelings expressed in the tweets are rated positively, while negative messages or tweets do not like. The data extracted from the social networking sites, specifically Twitter, are valid for exercising marketing actions (Mislove et al, 2011; Jansen et al, 2009; Weng et al, 2012), regarding credibility, they question that all data published on social networking sites are credible (Mustafaraj & Metaxas, 2010; Honey & Herring, 2009; Boyd & Crawford, 2012).

Referring to political prediction, there is more divergence as some authors and researchers argue that Twitter is a tool for estimating the vote, while others refute this thesis. Congosto, Moro & Fernández (2011) highlight the growing influence of the users of the social networks, and Twitter as a main tool for communication in real time. They also point out how competitive the role of Twitter is, in relation to the election polls, because Twitter acts as a probe to measure the opinion continuously in a much longer period; also they emphasise the more transparent nature of Twitter when sending messages or opinions because Twitter’s users want their free speech messages to be spread. Sang & Bos (2011) predict the
results of the Senate election in the Netherlands using Twitter. They investigate two variants of each political party, the abbreviation of the name and the name. They remove the tweets mentioning more than one party and they only keep the first tweets of each user. A combination of the two is the counted data: storing the first tweet that mentions a political party of each user. They then analyze manually the sentiment, drawing if the messages are positive or negative in relation to each political party. The result is that the analysis resembled the actual election results with a margin of 1.7%, but the prediction of the three major parties varied 1 point, with an error of 29% higher than the results of the two polls.

Tumasjan et al (2011) investigate whether one can predict the outcome of an election using Twitter, gauge the political sentiment by the 140 characters that a tweet has. They analyze 104,003 tweets posted in the weeks before the elections in Germany on September 27, 2009 (from August 13 to September 19). The number of tweets reflects the voter preferences and approximates the results of traditional surveys. The feeling of the messages on Twitter corresponds closely to the political agenda, the profiles of the candidates and evidence of media coverage campaign. Twitter messages reflect the results of the elections, even closer than the results of the polls. Twitter can be considered a valid indicator of political opinion and real-time indicator of the political sentiment, as according to these authors the tweets mentioning a political party may be considered a plausible reflection of the percentage of votes, hence concluding that Twitter’s predictive power is more accurate than election polls. Jungherr, Jürgens & Schoen (2011) analyze whether elections can be predicted using Twitter, taking into account the Pirate Party’s victory in the elections in Germany in 2009. This article is a response to the article of Tumasjan et al (2011). The authors reject the conclusion that the tweets mentioning a political party may be considered a plausible reflection of the percentage of votes, hence concluding that Twitter’s predictive power is more accurate than election polls. This hypothesis is rejected for three reasons: 1) it does not indicate or specify the rules for data collection in general, nor the reason of the choice of the parties to discuss, and why they choose that period to collect information, 2) the methodology chosen is not appropriate, 3) the mentions on Twitter are not a valid indicator of the offline political sentiment or future results of the vote.

In relation to the research of social networking activity at certain events occurring in several countries, some researchers point to the importance of the social networks as a mobilizing medium, while others dismiss this idea. Skoric et al (2012) analyze the elections of
2011 in Singapore from April 27 (call) to May 7 (elections). They conclude that there is a moderate force of correspondence between the share of tweets and share of votes nationwide. Regarding the political feeling, Twitter and its data are more relevant at a macro level than a micro level, or specific. The level of quality in the data prediction on Twitter, depends on the level of democracy and censorship in the country. Larsson & Moe (2012) focus on the analysis of Twitter during a general election in Sweden. The analysis was based on 99,382 tweets collected from a month before the election (September 19, 2010) and up to four days after that. They conclude that the active users of Twitter produce more information unlike most of the anonymous users, so Twitter is a tool for disseminating information rather than dialogue, the data extracted from Twitter should be extrapolated to the general use of Twitter by all voters. In general, the authors encourage further research of Twitter as a tool for data extraction, because even their contribution is quite limited to the interaction of Twitter users, rather than the extraction of semantic data.

Gayo-Avello (2011) analyzes Twitter as a tool to predict an election results during the US election in 2008. The author states that Twitter cannot predict an election results. The main conclusions of the research are: not everyone uses Twitter (there is a problem among a sample of Twitter to be considered as a sample of the population of study), not all Twitter users who write posts do it about politics, not all the content on Twitter is true and the feeling expressed on Twitter should be analyzed in a simpler way. He proposes some recommendations for future research: define the sentiment analysis of the political tweets, detect propaganda and disinformation, detect not real users of Twitter, verify credibility, investigate demographic data and assign demographic attributes, basic research participation of the user. Mustafaraj & Metaxas (2010) study Twitter during two months collecting 2,500 topics. The conclusion when it comes to the credibility of the content on Twitter is that the lack of information about the users of Twitter does not allow concluding that the content is real and credible. Asur & Huberman (2010) confirm that it is possible to use the content of social networking sites to predict results in real time. They conducted a study based on Twitter to predict the box office. They apply for the analysis a linear regression model; also they perform an analysis by the subjectivity of feeling (positive and negative tweets divided by neutral tweets) and the polarity (positive feeling tweets divided by negative sentiment). In conclusion, it is confirmed that there is a strong correlation between the amount of attention given to a film and the later position in the box office. After analyzing the sentiment of the
tweets, they confirm their effectiveness in improving predictions once the films have been released. The authors argue that this method can be extrapolated to predict election results.

O Connor et al (2010) investigates the sense of the text in Twitter related to the public opinion. According to the authors, undermining the public opinion of the content published on Twitter could be a faster and a less expensive alternative to traditional surveys, but Twitter only allows writing an average of 11 words per message. For the analysis they used 1 trillion of tweets published in 2008 and 2009. Through a textual analysis they recovered messages and estimated the opinion, to see if these messages express positive or negative opinions or news about a topic by extracting the feeling. Cautiously one could argue that surveys could be replaced by a simple collection of text data generated by a social networking site. But the techniques and analysis should be substantially improved. In this sense, Mislove et al (2011) state that Twitter presents a unique opportunity to examine the public communication of a fraction of the population. The question they ask is whether Twitter is a representative example of society. They analyze 1,8 million tweets sent by 55 million users between March 2006 and August 2009 to compare socioeconomic, educational level and type of job, since there are only data concerning the users name, their location and the text of tweets. The analysis is based on three pillars: geographical distribution, gender and race. The conclusions they reach are: Twitter users significantly represent the most densely regions in the U.S., the users are predominantly male, and they represent a nonrandom sample of classification of race distribution.

Golbeck, Grimes & Rogers (2010) discuss the use of Twitter by the US Congress members. The conclusion they reach is that they use this tool for self-promotion, not for a direct communication between congressmen and citizens. Schmitt-Becka & Mackenrodt (2010) argue that political influence is produced by the interaction of voters with people in their daily environment and the mass media. Aday et al (2011) present a report on the relationship of the new media and the Arab spring that toppled regimes in Tunisia, Egypt, Libya and protests in Bahrain. Thay analyze the codes bit.ly as URL shortener and as a link of a website. They conclude that they served more to disseminate information out of the region than in the same region. So the social networks used by the bit.ly links do not appear to have played an important role in collective action. Burns & Eltham (2009), study the role that Twitter played in public diplomacy and information operators in the crisis of the election in Iran in 2009. It was an important tool to mobilize people and influence the international media.
about what was happening in Iran. The effect of Twitter should have been a catalyst for a peaceful regime change in Iran, but the fact it was not.

As can be seen, the political communication has many advantages in the current context where the information and technology prevalence; but there are political parties, politicians and institutions that do not take this chance, not even one of the tools with more notoriety, the social networking sites. Not all political parties use them in the same way, but what is truly remarkable is the importance that new technologies have had on various political events in different countries like Ukraine (Kyj, 2006) and Egypt (Tufekci & Wilson, 2012). This fact, along with the importance of data for marketing researchers, makes several researchers extract that data from Twitter to predict events, trends and results; because Twitter offers a unique opportunity to discuss the public communication of a fraction of the population (Mislove et al, 2011).

So there are publications that analyze software as IARPA (Goth, 2012) or iScience Map (Reips & Garaizar, 2011). While some researchers reject the hypothesis of Twitter as a tool of political prediction (Gayo-Avello, 2011; Jungherr, Jürgens & Schoen, 2011; Mustafaraj & Chung, 2011), other corroborate it (Tumasjan et al, 2011; Asur & Huberman, 2010; Sang & Bos, 2012; O'Connor et al, 2010).

**Twitter Metrics.**

Twitter is a social network whose business model is advertising and to sale its data to third parties for marketing and sales. But the analysis of this social networking site looms the mechanisms and decrypt algorithms, metrics, techniques and ultimately the Twitter model, concluding that Twitter is a source with a huge potential for measuring trends behaviors, liking, predictions; and probably for the estimation of the vote.

Until now the findings of this article are: Twitter, the microblogging site, is the best tool to investigate the estimation of the vote. Twitter is the fastest growing social network for expressing opinions and to mobilize. The interest of some researchers in Twitter makes clear the interest of the social media as a tool for estimating the vote. Some researchers find Twitter a tool for estimating the vote, confirming this hypothesis, other researchers refute this hypothesis. Twitter has a complex metric to monitor the user behavior. The companies doing market research and election polls in Spain do not use social networking sites as tools for
estimate the vote. With these findings a scientific method and several hypotheses are formulated to confirm or refute that the social networking sites are tools for estimating the vote. The framework of the research is the elections to the Catalan Parliament on November 25, 2012.

The Hypothesis Approach.

Before formulating the hypotheses, it is essential to identify the problem under study and the theoretical statement. The social networks sites as a tool for estimating the vote, is the problem under study. The theoretical statement says that from the commercial and the user point of view, the importance of the social networking sites in Spain and other types of networks is confirmed, also the great future for them, and their performance for data mining (Twitter). In academia several studies has proven the importance of the social networks. With the advent of the Internet, the social networking sites have become increasingly relevant and several researchers are addressing this issue and their potential. In relation to the social networks on the Internet and politics, from the point of view of a company or business, it has been concluded that these tools have been used increasingly for social mobilization, for free speech, what is called the democratization of influence. Twitter is the tool with most estimate growth and it is the leader as a social networking website of microblogging. It is important to stand out the limited use of this medium as a tool for estimating the vote by the market research companies. Several academic researchers are focused on the use of the social networking sites as tools of measurement, estimation and prediction of policies beyond new media, with special emphasis on Twitter trends.

Therefore, the hypotheses deduced are:

H1: Twitter predicts actual trends in society.

H2: Twitter mobilizes politically active users.

H3: Twitter is a tool for estimating the vote.

The Methodology Approach.

The goal is to get relevant information about the changes in trends/voting intentions of one or more specific segments of the electorate, those who frequently use the Internet and the
The online social networks as a tool for estimating vote

social networks, to identify specific sub-segments within them. For this, a selection of a random sample of Twitter users is done during the elections to the Catalan Parliament on November 25, 2012.

The campaign period for the elections to the Catalan Parliament on November 25, 2012; was from November 9 to 22, the 935 Twitter users who write more posts or tweets with the hashtag #25N in Catalunya are selected from November 1 to November 29, 2012. They are instructed to complete a pre-election survey questionnaire about political-ideological and electoral issues (November 1 to 23). Then, the tweets of the 935 users are analyzed during that period. Finally, the users are indicated to participate in a final quiz or post-election survey (November 26 to December 9), so each user of Twitter can explain its voting behavior and whether there has been any change in the voting behavior or attitude towards parties and candidates. Also, to glimpse the role that Twitter may have had while participating.

So, the universe of this study is Twitter users related to electoral political issues in the elections to the Catalan Parliament on November 25, 2012; during the period from 1 to 29 November 2012. The sample is a representation of the previous universe, a random selection of users who write more messages or tweets with the hashtag #25N during the period from 1 to November 29, 2012 at Catalunya environment.

The information is collected in two ways:

- The electoral campaign for the elections to the Catalan Parliament on November 25, 2012 was from November 9, until 24 hours before the election day, November 23. The primary way of collecting information was performed by the servers of the company Aumentha (an adhoc study for this research), where 935 Twitter users were monitored during the period from 1 to 29 of November, including the election campaign attended from day 9 to November 23.

- Pre-Election and Post-Election Survey. To encourage the participation of 935 Twitter users in these surveys, a profile on Twitter @estudiUCM was created. A daily account management followed these users and invited them to participate in the surveys; also this Twitter account published daily information and interactions (Community Management). These online surveys were created in the domains http://elections20.com/estudioucm/index-
The online social networks as a tool for estimating vote

Inicio.htm (from November 1, 2012 until the day before the election, November 24) and http://elections20.com/estudioucm/ (from November 26, 2012 to December 9).

In addition to the survey and the analysis of Twitter users, three conditions were also analyzed on the campaign trail: the influence of the media, the speech of the candidates and their Twitter profiles. The candidates are:

- Artur Mas, CIU.
- Pere Navarro, PSC.
- Alicia Sánchez-Camacho, PP.
- Oriol Junqueras, ERC-Cat Sí.
- Joan Herrera, ICV-EUiA.
- Albert Rivera, C’s.
- Alfons López Tena, SI.

The results show a tendency of the vote where ERC would increase considerably in votes above CIU, taking into account the factors of influence of the media during the election campaign, the messages and the election program of the candidates and political parties; and the activity on Twitter. Artur Mas is the candidate most affected by the pressure on the media, this fact distorts his message about independence, he bets for his base voters. For example, he doesn´t have a Twitter account. The results obtained in the elections are 50 seats, a far distant from those results estimated by the published surveys. ERC is the ultimate beneficiary of the news about alleged corruption of CIU, a message clearly in favor of the independence of Catalunya and a winning management strategy in digital marketing, especially in the social media discourse. Oriol Junqueras is the most influential candidate, he won 21 seats, beating all the surveys published.

The Quantitative Analysis.

According to the pre-election survey (352 responses) (compared with the 135 seats of the Catalan Parliament), ERC wins 44 or 45 seats (21 in the 2012 elections), CIU gets 23 or 24 seats (50 in 2012), PSC 4 or 5 Members (20 in 2012), PP 2 or 3 members (19 in 2012), ICV-EUiA 9 or 10 (13 in 2012), C’s 8 (9 in 2012) and SI 23 (0 in 2012). Note that the majority of respondents vote in favor of the independence of Catalonia, if a consultation or referendum is called.
According to the post-election survey (313 responses), ERC wins 52 or 53 seats (21 in the 2012 elections), CIU wins 25 or 26 seats (50 in 2012), PSC gets 9 or 10 seats (20 in 2012), PP 2 or 3 Members (19 in 2012), ICV-EUiA 8 (13 in 2012), C's 8 (9 in 2012), SI 8 (0 in 2012) and CUP 13 or 14 (3 in 2012). As we see there is an absolute disparity with the results of the elections on November 25, but these results are more similar to the barometer of the vote of the Centre d'Estudis d'Opinio (CEO 2013) where ERC would win the election gaining between 37 and 39 members, CiU 34 to 36 deputies, C's between 15 and 17 seats, the PSC between 14 and 16 seats, PPC 13 or 14, ICV-EUiA between 12 and 14, and CUP 6 or 7. The majority of respondents would be in favor of Catalunya as a state of the European Union.

The winner would be ERC or Esquerra Republican. Note the emergence of CUP. The pre-election and post-election surveys confirm the voting trend based on the opportunity that minority parties have. This fact is because of a bottom up political communication, with the Internet as tool for flexible strategies without editorial control with a one way communication that encourages online activism, the mobilization and the commitment of its voters.

Regarding the influence Twitter had on the vote of the users, according to most respondents, the majority of information published during the election campaign on Twitter about politicians, candidates and others reaffirmed the vote, only a minority doubts about their intention to vote upon the information posted on Twitter. For them Twitter is a support for broadcast information like the television. Twitter influenced less than half of the respondents. Therefore, Twitter is a tool that reinforces the intention and the direction of the vote, behaving more like a communication support for candidates and political parties, and a tool for interaction between users because they use frequently this social networking site.

The main use of Twitter by the respondents is to publish personal tweets with the hashtag #25N, posting tweets of personal information and to communicate with friends and strangers. During the election campaign Twitter was used as a tool for self-expression, as a tool for political communication and to mobilize. The respondents estimate that Twitter was influential in the election campaign, also in changing political attitudes and the vote. They think Twitter is a tool for estimating the vote. In conclusion, Twitter is considered a new medium of communication and policy dissemination whose advantage is the interaction between the candidates or political parties with users and between users, because its strength
The online social networks as a tool for estimating vote

is to be a medium that allows free speech in political communication, reinforcing the vote and conditioning it.

**The Qualitative and Quantitative Analysis, and Evolution.**

The analysis is based on the evolution of the tweets posted by a sample of 935 users. This analysis is an ad hoc study made by Aumentha and the members of this research.

- Some data obtained from the analysis are the record number, 445,361 tweets were posted:

  **Chart.**

  **Record Nombre.**

  ![Chart](chart.png)

  *Source: Aumentha.*

- Most used tools and number of tweets posted:

  Web: 139,446  
  Twitter for Android: 100,279  
  Twitter for iPhone: 77,712  
  Twitter for BlackBerry®: 31,638  
  Twitter for iPad: 20,531

- Most popular users and number of mentions.
The online social networks as a tool for estimating vote

@junqueras: 43,875  
@ciu: 31,264  
@solidaritatcat: 28,670

- Most used hashtags and number of times used.

#25N: 197,972  
#debattv3: 57,313  
#totsambelpresident: 27,730

Analysis of the tweets.

The activity of the 935 monitored Twitter users is analyzed. The analysis of the tweets is done taking into account the influence of the media, especially the days 16, 17 and 20 of November of 2012, when the newspaper El Mundo published several articles about the corruption in CIU, the ruling party in Catalunya, who is pressing the Spanish institutions for the independence of Catalunya. The articles are "Police linked accounts of Mas and Pujol in Switzerland with the corruption of CIU" (2012), "The Pujol family have 137 million euros in Geneva, according to the Police" (2012) and "Hacienda investigates the Mas family account in Switzerland" (2012). Also it is taken into account the debate of the 7 major candidates in the Catalan television TV3. After a first analysis, due to the cancellation of 7 user accounts on Twitter, 928 users are analyzed. All those profiles that are public figures, politicians or fake profiles were removed, so 842 valid profiles are semantically analyzed. After every tweet of every user is semantically analyzed, the conclusions of the Twitter users are:

- Political Tendency: Center 11, Right-Center 1, Right 141, Transversalism 93, Left 596.
- Direction of the vote according to the semantic analysis:
The online social networks as a tool for estimating vote

Chart 3.

Direction of the vote according to the semantic analysis.

<table>
<thead>
<tr>
<th>Political Party</th>
<th>Initial Vote</th>
<th>N16</th>
<th>N17</th>
<th>N18</th>
<th>N20</th>
<th>N25</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERC</td>
<td>136</td>
<td>141</td>
<td>147</td>
<td>171</td>
<td>181</td>
<td>247</td>
</tr>
<tr>
<td>CUP</td>
<td>94</td>
<td>97</td>
<td>97</td>
<td>101</td>
<td>107</td>
<td>121</td>
</tr>
<tr>
<td>CIU</td>
<td>90</td>
<td>85</td>
<td>91</td>
<td>95</td>
<td>101</td>
<td>123</td>
</tr>
<tr>
<td>SI</td>
<td>73</td>
<td>79</td>
<td>80</td>
<td>85</td>
<td>83</td>
<td>97</td>
</tr>
<tr>
<td>PSC</td>
<td>50</td>
<td>44</td>
<td>45</td>
<td>44</td>
<td>45</td>
<td>47</td>
</tr>
<tr>
<td>ICV-EUiA</td>
<td>31</td>
<td>30</td>
<td>31</td>
<td>33</td>
<td>35</td>
<td>45</td>
</tr>
<tr>
<td>C’s</td>
<td>13</td>
<td>14</td>
<td>13</td>
<td>14</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>Pirata</td>
<td>11</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>PPC</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Pacma</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>RCat</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>NEC</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>UPyD</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NK/NR</td>
<td>339</td>
<td>337</td>
<td>324</td>
<td>284</td>
<td>261</td>
<td>129</td>
</tr>
</tbody>
</table>

Source: compiled by the author.

A comparison of seats among the post-election survey, the average of polls published in 2013, the average of pre-election polls published, the results of the elections to the Catalan Parliament in 2010 and 2012, the final results of the elections to the Catalan Parliament in 2010 and 2012, and the analysis of the initial and final vote on Twitter.

<table>
<thead>
<tr>
<th>Number of Seats</th>
<th>Post-electoral Survey</th>
<th>Polls 2013 (arithmetic mean)</th>
<th>Polls 2012 (arithmetic mean)</th>
<th>Election Results 2012</th>
<th>Election Results 2010</th>
<th>Twitter Initial Vote</th>
<th>Twitter Final Vote</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIU</td>
<td>25-26</td>
<td>25-36</td>
<td>64-65</td>
<td>59</td>
<td>62</td>
<td>14</td>
<td>20</td>
</tr>
<tr>
<td>ERC-Cat SI</td>
<td>52-53</td>
<td>37</td>
<td>16-17</td>
<td>21</td>
<td>10</td>
<td>22</td>
<td>39-40</td>
</tr>
<tr>
<td>PSC</td>
<td>9-10</td>
<td>15-16</td>
<td>18-19</td>
<td>20</td>
<td>28</td>
<td>8</td>
<td>7-8</td>
</tr>
<tr>
<td>PPC</td>
<td>2-3</td>
<td>14-15</td>
<td>19</td>
<td>19</td>
<td>18</td>
<td>0-1</td>
<td>0-1</td>
</tr>
<tr>
<td>ICV-EUiA</td>
<td>8</td>
<td>13-14</td>
<td>15-16</td>
<td>13</td>
<td>10</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>C’s</td>
<td>8</td>
<td>16-17</td>
<td>17-18</td>
<td>9</td>
<td>3</td>
<td>2</td>
<td>2-3</td>
</tr>
<tr>
<td>CUP</td>
<td>13-14</td>
<td>9-10</td>
<td>2</td>
<td>3</td>
<td>-</td>
<td>15</td>
<td>19</td>
</tr>
<tr>
<td>SI</td>
<td>8</td>
<td>-</td>
<td>0-1</td>
<td>0</td>
<td>4</td>
<td>12</td>
<td>15-16</td>
</tr>
<tr>
<td>Pirata</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PACMA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>RCAT</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NEC</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>UPyD</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NK/NR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>54</td>
<td>20-21</td>
</tr>
</tbody>
</table>

Source: compiled by the author.

Twitter increased by 18% in account users and use from 2010 to 2011, in 2012 it has over 500 million users worldwide and 5 million in Spain (Twitter, 2012). 2013 closes with Spain as a third world country with the highest penetration (PeerReach, 2014) and a 42%
increase in the use of this social networking website (The Cocktail Analysis, 2013). Twitter is a microblogging social network, which runs through communication and sharing messages of 140 characters per tweet (Clemente, 2011). Each tweet will be embedded in a software package that takes up 10 or 12 times more than the message itself. This is the metadata: data containing valuable information about whom, how, when and where the tweet was originated. A great opportunity for data mining and especially for companies, advertising agencies and marketing. Twitter has great potential as a source of data mining, that is why many researchers have created their own measuring tools or research. There are subsidiaries of Twitter as Crimson Hexagon or Mediasift, they have a contract with Twitter to access the metadata of this social network through Twitter API, and extract extremely rich of information for marketing, or other business/research. Twitter, like other social networking sites, was created to facilitate global communication among the citizens of the world, but they have a part of business based on the information obtained from its users to revert it in marketing for the company or for others (Clemente, 2011; O'Brien, 2012; Barrett, 2010; Andrews, 2012), these bidding practices or circumventing privacy through contract clauses that you agree to register a profile on them. Twitter tenders these privacy practices and circumvents the privacy of the user by the contract clauses that the user agrees when registering. The analysis of the data concluded from Twitter, compared to the average of the surveys, shows certain findings. Twitter predicts the rising of ERC as the most voted party against the lowering overall of CIU. If the results of the analysis of Twitter are compared with the actual results of the elections of 2012, the hypothesis 1 (H1) Twitter predicts actual trends in society is confirmed, comparing the variation with the average of the surveys and barometers of 2013.

According to Bond, Fariss & Jones (2012) the online political mobilization works because it induces to the free expression and it also induces to the collection of information and the validation of the vote. Moreover Schmitt-Becka & Mackenrodt (2010) conclude, regarding the social networking sites, that a person who receives a positive influence in his circle or close group of people about an electoral participation or voter turnout, the likelihood of voting increases. According to different studies (ONTSI, 2012; The Cocktail Analysis, 2012; The Cocktail Analysis, 2013; IAB, 2012; IAB 2013), the users use the social networking sites as tools of influence, a medium to express themselves (35% of users), a vehicle for awareness/mobilization of the civil society with a frequency of very common to
quite often 18%; this activity did not exist in 2009 and 2010. Twitter is the social network with more use for social mobilization to support a political call, 12% use it to express themselves freely or promote/support social or charitable causes. It is the preferred social network for personal expression (20% of users) (The Cocktail Analysis, 2013). Therefore, the users of the social networking sites use them as a communication tool, not only to interact or communicate, but they use them for two-way communication allowing them to be present in a globalized world, influencing events directly, recommending and reviewing all kinds of global and local issues. This democratization of influence is because the growth of the social networks as tools of mobilization, expression, social and political awareness; and Twitter is the social network chosen for it. Regarding the analysis of the data extracted from Twitter for estimating the vote, it should be mentioned two important facts: the technological split in society and the minority parties. Referring to the digital divide, division, or penetration of these technological media in society, it should be emphasized that the analysis of the Twitter users shows that CUP and SI obtain more seats than the real results. This factor is because most people do not have an active Twitter profile, and those who have it tend to support and mobilize for minority parties who use the social networks as an important campaign tool because of its bidirectional communication, low cost, no editorial control, one way, flexible, many to many, bottom up, accessible, near the young voter, acquisition and retention of voter protest of fundraising, cyber activism, Democracy 2.0. Ultimately, the political mobilization on Twitter works because it induces both the users and policy makers to the free expression. So, the hypothesis 2 (H2) **Twitter mobilizes politically active users, is confirmed.**

The most relevant research in Democracy 2.0 and Government 2.0 (in relation to political and social networking sites, especially Twitter) is divided into 7 categories: influence, mobilization, sentiment, data, credibility, credibility, prediction and policy investigation into events that occurred in different countries and institutions. Within the political prediction research, there are investigations focused on Twitter or the social media as tools of political prediction. Some researchers refute the hypothesis of Twitter as a tool of political prediction (Gayo-Avello, 2011; Jungherr, Jürgens & Schoen, 2011; Mustafaraj & Chung, 2011) and other researchers corroborate it (Tumasjan et al, 2011; Asur & Huberman, 2010; Sang & Bos, 2012; O'Connor et al, 2010). The results of the methodology applied to the Catalan Parliament elections of November 25, 2012, regarding the analysis of the tweets of 842 users refute the
hypothesis 3 (H3), **Twitter is a tool for estimating the vote**, mainly because of the digital divide and influence of minor parties.


**Limitations and future research.**

The limitations found in this research are the digital divide, credibility (semantic analysis), to analyze a posteriori is not to estimate, the technology (research versus business) and the period between the election calls. Concerning future research, primarily the recommendation is to choose a group of Twitter users sufficiently large. It should be composed of active and non-politically active users, also it must be ensured they want to participate in a pre-election and post-election survey.
Bibliography.


The online social networks as a tool for estimating vote


The online social networks as a tool for estimating vote


24