Why and How to Create a Panel of Twitter Users

Julien Boyadjian, Marie Neihouser

Boyadjian Julien, University of Montpellier, julien.boyadjian@hotmail.fr

Neihouser Marie, University of Montpellier, neihouser.marie@gmail.com

Why and How to Create a Panel of Twitter Users

Abstract:

Twitter offers a new space for citizens to air their political views. Every day, thousands of individuals comment on politics via this social network. Although the literature has demonstrated clearly the relevance of the media context on the level of political messages posted, little research has yet been done on users' social and political characteristics and the impact of these characteristics on the number and frequency of posts. To determine the social logics behind political tweets, I devised a method for creating a panel of the social network. The study results show the decisive impact of sociological variables on individuals' propensity to talk about politics on the social network.

Keywords: Twitter; big data; political participation; political discussion

Introduction

Twitter, the microblogging web service founded in 2006 in the United States, has quickly become a valuable field of investigation for analyzing the political behavior of ordinary citizens via the Internet (Jungherr, 2014). Users of this social network can post messages of no more than 140 characters on any subject they choose, including politics. These political messages can thus be considered true research materials in the same right as answers to questionnaires and interview transcriptions. The thousands of tweets posted daily on the network are all evidence of practices, attitudes and opinions expressed by Internet users on a variety of topics, including politics. This new material has certain specific features, however. Unlike answers recorded in the context of a questionnaire, the messages posted on social networks are not data artificially generated through fieldwork interaction. Internet users

do not post messages because a researcher has asked them to do so. Therefore, they do not adapt their discourse according to a survey protocol. A second aspect that characterizes messages posted on the social media is the fact that they are time-stamped. On Twitter in particular, each tweet carries the precise date and time of the post. Using web-tracking software, which makes it possible to collect and store huge amounts of digital data, researchers can thus study messages several months and even several years after their authors have posted them. Internet users' reactions, discourses and representations regarding past facts or events as they occurred at the time of the events can thus be studied without having to rely on the respondents' memory, which is often faulty. Web-tracking software thus makes it possible to objectively measure the level of political message posts and identify the exact time at which these messages were produced. A third characteristic has to do with the potentially exhaustive nature of the computerized analysis of tweets. As web-tracking software makes it possible to collect vast data sets, it is no longer necessary to sample the target population studied to obtain statistically representative results. The digital nature of the field of study makes it technically possible to conduct an exhaustive analysis of the social phenomena studied on line.

While the above characteristics present real heuristic advantages and, as will be seen, partly condition the way in which the data has been viewed in the literature, digital posts have the considerable drawback of not being systematically associated with an author having clearly defined socio-demographic characteristics. Even if people sometimes give information about their sex, age, occupation or place of residence, this information is only fragmentary and is not enough to define a user's social profile. Contrary to questionnaire surveys, it is impossible to situate socially the authors of the messages collected using web-tracking analysis tools. Researchers who wish to perform a quantitative analysis of the political messages posted on the social media, particularly on Twitter, can then basically adopt two different research strategies. They can opt for a computerized analysis of the social network (big data) and thus take advantage of the features offered by the nature of tweets (unsolicited time-stamped messages and exhaustive data). But the social characteristics of the authors of the messages studied cannot be identified in a computerized analysis of this sort. Or investigators can choose a more classic approach using a questionnaire survey and thus identify the sociological characteristics of the message producers. This approach, however, makes it impossible to take advantage of web-tracking analysis.

Recent academic studies devoted to the political uses of Twitter mostly rely on computerized analyses of very large data sets. This research position prompts investigators to conduct an analysis that is more focused on message content and the context in which they were posted than on the social logics underlying their production.

State of the Art

Considerable research has been devoted to the political uses of Twitter within a very short time period. In a review article published in 2014, Andreas Jungherr inventoried no fewer than 115 studies on this topic (Jungherr, 2014). Aside from their sheer number and novelty, another feature of these studies is the disciplinary diversity of their authors. Researchers from a variety of disciplines generally having little exchange between them, namely the information and communication sciences, computer science and to a lesser extent Anglophone political science, have entered this new "digital" research field.

The original discipline of these authors (most of them in computer science), probably explains their inclination to prefer computerized web-tracking analysis tools and to suggest very empiricist and descriptive approaches to the social network. These studies therefore count the number of political messages posted (Ausserhofer and Maireder, 2013), analyze their content and tonality (Linh et al., 2013), sometimes comparing them to opinion survey data or election results (Gayo-Avello, 2013). Only a fraction of the studies set out to interpret empirical results observed from the standpoint of a theoretical framework for analysis (Jungherr, 2014). The theoretical references used are drawn mainly from political communication research focusing on media effects, in particular through issues of agenda-setting (McCombs and Shaw, 1972) and media framing (Gamson, 1989)Twitter emerges as an object of study particularly suited to these research questions. The social network is a space that is highly reactive to news stories and televised events. Moreover, Internet users' reactions can be situated in a specific media and political context, as each message posted is date/time-stamped.

Several authors have thus shown that the number of political tweets posted on the social network fluctuated considerably from one day to the next and that spikes in production coincided with political events receiving a high level of media attention (Larsson and Moe, 2013). In this regard, televised political events (debates, campaign rallies, etc.) are believed to

generate the most online political reactions (Jungherr and Jürgens, 2014). The level of political activity on Twitter is thus thought to be determined by the level of politicization of news items: the more airtime the media devotes to political subjects, the more Internet users "talk" about politics online. These results would appear to confirm Mc Combes and Shaw's (1972) hypotheses regarding the media's ability to determine citizens' concerns and topics of discussion. Some authors have sought to test this hypothesis at a more precise level of analysis by comparing the airtime devoted to various political candidates and the number of Twitter messages mentioning these candidates. These authors have observed a positive correlation between the two phenomena: the candidates mentioned most often on Twitter are also those most often mentioned in the media (Jungherr, 2014). The more media coverage a politician is given, the more Internet users tend to talk about the person online. Although Internet users' discussion topics thus seem partly determined by the media, some authors have wondered if online discussions mobilize the same interpretive frameworks. The results pertaining to this question are more nuanced. While political tweets indeed refer mostly to events the major media outlets have put on the agenda, many of these messages also criticize how the media frames them. Some users for instance endeavor to correct factual errors made by journalists or politicians in real time (Ampofo, Anstead and O'Loughlin, 2011; Anstead and O'Loughlin, 2011). Although it is a highly reactive space to mainstream media, Twitter nevertheless seems to be used at least by some subscribers as a space for critical discussion of this very media.

Agenda setting and media framing theories used in this literature are thus particularly suited to the nature of the research topic and help to make sense of the irregular patterns in posting political messages on the network. The number of political tweets posted is thus believed to be determined by the level of politicization of topics on the media agenda. However, are all Internet users equally inclined to comment on political topics in the headlines, whatever their social background and level of politicization? Do individuals who are the most interested in politics also post more messages during periods of high politicization, or does this phenomenon apply only to the least politicized individuals? The research reviewed above does not help to answer these questions. The social dispositions of individuals who express political opinions on Twitter are very rarely analyzed, and this for a relatively simple reason: the socio-demographic characteristics of the message authors cannot be determined, at least not solely, through web-tracking analysis, the principal tool used in this research.

Conversely, the rare academic studies that have chosen to use a questionnaire survey over computerized analysis of the social network have identified Internet users' sociological characteristics, though without examining the political and media context in which the messages were produced (Bekafigo and McBride, 2013; Dimitrova et al., 2014). These authors have endeavored to measure the number of Twitter users in the overall population, generally concluding that they make up only a slim proportion: in 2010, Twitter users made up only 8% of the American Internet using population (Rainie et al., 2012). In the United States, the majority of political tweet producers tend to be well-educated, white, male, Democrats and Independents (Gainous and Wagner, 2014). These results, while they seem plausible (and fairly consistent with the data for France), should nevertheless be interpreted with caution. Given the low level of Internet users claiming to have a Twitter account, these questionnaire surveys (conducted on representative samples of 1,000 to 2,000 American adults) are often based on very small samples (100 to 300 individuals at most). More fundamentally, another limit to these surveys is that they do not compare the users' socio-demographic data with their posting activity. It is thus impossible to measure to what extent the volume and frequency of the political messages an individual posts is determined by his or her social and political characteristics and to determine what influence the media context has on this production. Do a user's age, education and political orientation determine the number of political messages posted? If so, do these variables exercise the same influence whatever the media and political context?

The literature offers little by way of answers to these questions. Only a study combining web-tracking analysis and a classic questionnaire survey can highlight the dispositional and contextual factors that preside over the production of political messages on the network. It is to this end that I devised my own methodological system for creating a panel of Twitter users.

Method

By administering a short questionnaire to a representative sample of Twitter users and then, using special software, collecting all of the messages posted by these accounts over several months, it is possible to profit from the advantages of web-tracking analysis (unsolicited talk not elicited by the researcher, time-stamped messages, etc.), without being hampered by its main limitation (authors having poorly identified social

characteristics). However, putting together a truly representative panel of users likely to post political messages raises some methodological difficulties. A brief description of the various stages in forming the panel of users of the social media will serve to illustrate these.

To determine the sociological characteristics of individuals who post political messages on Twitter, I decided to administer a short questionnaire to a representative sample of users of this service. Unlike web-tracking approaches by which it is possible to analyze the entire population having Twitter accounts, administration of a questionnaire requires a population sampling procedure, and this for practical reasons. To perform this operation the investigator must have in his or her possession a complete list of the users of the social media. To my knowledge, such a list does not exist, or at least is not made public by Twitter. On the other hand, it is possible to inventory all of the users likely to post political messages by collecting all the tweets that include one or more political keywords over a given period of time.[1] I thus collected all the tweets posted on the network between March 1 and 31, 2012 that had the name (spelled correctly or incorrectly),[2] nickname, or campaign hashtag of the candidates in the first ballot of the French presidential election held on April 22, 2012. During this heavily politicized period, over 2,800,000 tweets were collected from 248,628 individual accounts.[3]

From this set I then extracted a random sample of 20,000 accounts on which a "manual" eligibility analysis was performed. To be eligible for the panel, the accounts had to have an individual author (and not an organization, a media outlet or a corporate entity) and be written primarily in French. I then "followed" 10,229 eligible accounts so that they would consult my profile inviting them to answer an online questionnaire containing some twenty socio-demographic items (sex, age, occupation and socio-professional category, education, social background, etc.) and political questions (political orientation, voting history, advocacy activity, etc.). In the wake of a follow-up campaign, 658 individuals (6.4 % of eligible accounts) answered the questionnaire and 608 (6.3 %) agreed to be part of the panel (the questionnaire specified that their tweets would be collected and analyzed anonymously by a team of researchers). Since that time, there has been no further interaction with the respondents.

---

[1] According to the definition of political given above, the keywords are the names of the main political parties, actors and institutions.

[2] In an effort to prevent any social discrimination.

[3] It can be assumed, although it is not verifiable, that users who posted no political message during this period of highly politicized ordinary discussions were highly unlikely to post any political messages at a later date.

To measure the extent to which the individuals who agreed to answer the questionnaire were representative of all the eligible accounts and thereby to minimize the selection bias inherent in any questionnaire survey, in addition to the "respondent panel," I also formed a control panel composed of a random sample of 620 individuals who had not answered the questionnaire. I then sought to characterize the members of this "non-respondent panel" socially and politically, using fragments of information that the users sometimes include in the social media or their blogs. This "non-respondent panel" was thus made up of both individuals for whom I had one or more socio-demographic indicators as well as individuals for whom I had none.

Using the software, I archived and analyzed all the tweets posted by the "respondent" and "non-respondent" panels over an 11-month period, from March 1, 2013 to January 31, 2014. During this period, nearly two million tweets (including nearly 125,000 messages containing occurrences of the political keywords) were posted by the members of these two panels. Before discussing the results of the statistical analysis of the panels, I would like to mention the advantages of this methodology.

The method of creating a panel of Twitter users for this study combines the principal merits of both web-tracking analysis and questionnaire surveys. As in the case of research based solely on web-tracking analysis, the messages posted by panel members were not elicited by a researcher. While there was indeed interaction with the respondents at the outset of the study to administer the questionnaire, no further contact was undertaken. It is thus highly likely that between the date when the respondents answered the questionnaire (in March 2012) and the date when computerized analysis of their messages began (in March 2013, i.e. one year later), the respondents forgot about the investigation. Furthermore, like studies based on web-tracking analysis, the messages recorded by the software used are systematically time-stamped, making it possible to analyze the distribution by day of the number of tweets posted and situate this production in a specific political context.

The method for creating a social media user panel has two major differences compared to web-tracking research methods. For one, it does not enable the researcher to make an exhaustive study of all the messages posted on the social network. Nevertheless, the panel forms a statistically representative sample of the people likely to post political tweets on the social network. The size of the sample ($n = 658$) and the randomness of its selection ensure representativeness, justifying statistical processing of the data. Moreover, as previously noted,

to control possible selection bias induced by the questionnaire, a control panel was formed of people who did not respond. By comparing the data drawn from the respondent panel and the control panel, the representativeness of the results observed can be measured regularly. Secondly, unlike with web-tracking analysis, my method for creating the panel makes it possible to systematically associate messages collected by the software with socio-demographic and political data on their authors. By administering the questionnaire I was able to identify the sociological characteristics of each member of the panel (sex, age, socio-professional category, interest in politics, etc.). This information was programmed into the software so that each message collected was associated with the social and political characteristics of its author. This data then made it possible to establish a precise sociological profile of the people expressing their political views on Twitter and thus investigate the hypothesis of a "democratic" discussion forum. It also enabled me to compare the volume and frequency of message posts among the various categories of users (employees and managers, little politicized and highly politicized individuals, voters on the left and right, etc.). In other words, I could thus measure the impact of sociological variables on the number of political tweets posted. Thirdly, it enabled me to objectivize the influence of the political and media context on these various categories of users.

These hypotheses were verified by longitudinal analysis of the "respondent" and "non-respondent" panels, and the dispositional and contextual factors governing over the posting of political messages were thus identified.

Key Findings

Three main series of results can be drawn from the information gleaned from the questionnaire as well as the software-generated data. First, it will be shown that people who post political messages on Twitter are a highly select social group. Second, these individuals' social and political characteristics have been found to determine their level of political message posting. Third, the political context influences the level of political message posting for all categories of users.

A very socially selective political discussion forum

People who more or less regularly post political messages on Twitter are a highly select social group. University educated, managerial level males are overrepresented in this population. Keenly interested in politics, regular voters and activists, authors of political tweets also show a very high level of politicization. However, unlike the usually most politicized and participative populations, authors of political tweets exhibit a certain sociological specificity: the younger population categories are very clearly overrepresented.

The most outstanding characteristic of the panel members that first comes to light is their young age. Whereas in 2012 the 18-25 age group made up 9.7 % of the French population, 39.8 % of the questionnaire respondents and 45.8 % of the non-respondent group (whose age could be identified) belong to this age group.

Table 2. "Respondent" and "non-respondent" panel populations by age (%)

Age group	Respondent panel (*n = 658*)	Non-respondent panel (*n = 240**)	French population INSEE 2012
Under 18	8.1	19.6	22.2
18 to 25	39.8	45.8	9.7
26 to 34	20.1	14.2	11.1
35 to 45	18.8	14.6	14.7
46 to 60	9.3	4.6	20.0
61 and over	4.0	1.3	22.2
Total	100.0	100.0	100.0

*Individuals whose age could be identified (38.8% of non-respondent panel randomly drawn)

This result may seem counter-intuitive, as the younger population categories are precisely those that vote the least. The 18-24 age group contains the lowest proportion of people who are not registered to vote and the highest number of non-voters. Yet it is indeed the 18-24-year-olds in the panel who express themselves politically the most on Twitter. In the social media, contrary to what is observed in elections, younger individuals are those who "participate" the most. On examining these results, one might then assume that young people are perhaps not less politicized but simply less drawn to classic forms of political participation (voting, working for a party, etc.) and find in the social media a new form of commitment that better suits them. The data taken from the questionnaire would appear to qualify this

hypothesis. Younger individuals who have posted political tweets are precisely those who voted in the first ballot of the presidential election. While French people between the ages of 18 and 24 are proportionally those who voted the least on April 22, 2012 (only 72.4 % of this age group voted), almost 92 % of those of the same age who posted political tweets *declared* having voted. This is the highest turnout rate by age group of the entire "respondents" panel. Aside from their young age – which therefore cannot be interpreted as a sign of the democratization of political participation – the people who posted political messages on Twitter actually show very similar social characteristics to the people belonging to the most politicized segments of the population.

Authors of political tweets thus occupy fairly high positions in the social space. Senior managers and professions are significantly overrepresented in the sample. Whereas managers make up only 14.8 % of the French working-age population no longer enrolled in a course of study, 49.2 % of the people who answered our questionnaire and 71.1 % of the non-respondent group (and whose social position could be identified)[4] belong to this socio-professional category. On the other hand, manual workers and employees are underrepresented. Only 2.4 % of the respondent panel and 1.0 % of the non-respondent panel (whose occupation could be identified) are workers, whereas this socio-professional category represents 19.1 % of the French working-age population no longer enrolled in a course of study.

---

[4] Regarding the non-respondent panel, the socio-professional category could be identified in 32.4 % of the cases. It should be noted, however, that certain categories are easier to identify than others. The socio-professional category of non-respondents whose occupation could be identified are not distributed the same was as the socio-professional category of non-respondents whose occupation could not be identified. Indicating one's own profession on the Internet is a socially situated activity that is distributed differently according to the various socio-professional categories. Managers thus tend more than average to indicate their profession on their Twitter profile or on other social media. Their presence on "professional" networks such as LinkedIn and Viadeo (more geared toward skilled jobs), where user profiles often amount to online *curricula vitae*, facilitates identification of their profession. On the other hand, technicians and associate professionals, manual workers and employees rarely specify their occupation on the Internet. For example, 74 % of the managers in the "respondents" panel had filled in their Twitter profile as opposed to only 33 % of the manual workers.

Table 3. Panel members by socio-professional category (%)

Socio-professional category (excluding students)	Respondents (*n* = 378)	Non-Resp. (*n* = 201)**	INSEE 2012
Farmers	0.3*	0.0*	1.0
Sole traders, self-employed and business owners	2.4*	4.0*	3.4
Managers and higher intellectual professions	49.2	71.1	9.6
Technicians and associate professionals	25.7	16.9	13.3
Employees	10.6	3.5*	16.0
Manual workers	2.4*	1.0*	12.4
Retired	5.8	1.0*	26.5
Other no occupation	3.7*	2.5*	17.7
Total	100	100.0	100.0

Source: INSEE, 2012; IFOP, 2010

*Low numbers **Individuals whose occupation could not be identified ***Only the unemployed and housewives/husbands.

Aside from their occupation, the panel members display other social characteristics that make them similar to the generally most politicized populations. One example is the overrepresentation of men. Whereas the French Internet-using population is made up of men and women in equal proportions,[5] women are clearly underrepresented in the sample. 59.4% of the authors of political tweets who answered the questionnaire and 66.4% of the non-respondents are men.

Table 4. Population by sex (%)

Sex	Respondent Panel (*n = 658)*	Non-respondent panel (*n = 590)*	Internet pop. IFOP 2012 (*n = 2005)*	French pop. INSEE 2012
Male	59.4	66.4	49.7	48.5
Female	40.6	33.6	50.3	51.6
Total	100.0	100.0	100.0	100.0

Sources: IFOP, INSEE

Beyond the "historical" differences in political competence between men and women (Norris, 1993), the gender gap in political participation on the social network can also be explained by the predominance of management-level positions in the sample. Each socio-professional category is defined by a specific sex ratio. In 2012, among the French population age 15 and over, the "employee" category was thus made up of 76.6 % women, whereas in

---

[5] On Facebook, the social media with the highest penetration rate among the Internet user population, women on the other hand are a majority: 53.1 % of the registered users are women (source: IFOP, "Observatoire des réseaux sociaux," 2012 wave).

managerial positions, it was only 40.2 %. Overrepresentation of managers in the sample thus provides an additional explanatory factor for the statistical underrepresentation of women.

The last sociological characteristic of people posting political messages on Twitter is their very high level of education. Individuals who agreed to answer our questionnaire are considerably better educated than the average French person.[6] This overrepresentation is true not only of individuals enrolled in a course of study, as noted previously. If students are excluded from the sample, three-quarters of the tweet authors have a degree in higher education (74.9 %) and nearly half even have a Master's degree or higher 44.5%.

Table 6. Respondent population by education (%)

Last degree obtained (excluding students)	Respondent panel (*n* = 378)
Less than the Baccalaureate	11.4
Baccalaureate or equivalent	13.8
Associate's degree, vocational degree or equivalent	14.8
Bachelor's or equivalent	14.6
Master's or equivalent	44.5
Doctorate	3.7
Overall Total	100.0

People posting political messages thus exhibit atypical social characteristics: they are primarily highly educated male students and managers. These characteristics help to explain the very high level politicization of this population. 56 % of the questionnaire respondents state they are "very" interested in politics, a finding that is considerably higher than the results of surveys taken among representative samples of the French electorate. Thus according to an exit poll taken in June 2012, only 21 % of French citizens of voting age declare they are "very" interested in politics, a rate that is 35 percentage points lower than the present study findings (CEVIPOF, 2012). Beyond this very high level of politicization, authors of political tweets are also characterized by a fairly homogenous political profile: the leftwing electorate is predominant on Twitter. 56 % of the questionnaire respondents position themselves "very"

---

[6] It should be noted, however, that the underrepresentation of the older population categories in the sample automatically explains the high educational level, as the chances of having a Master's degree in 2015 is not the same for the over 65 and the 25-49 age group.

(19 %) or "rather" (37 %) on the left, compared to only 23 % who position themselves "very" (4 %) or "rather" (19 %) to the right.

The influence of social and political variables on tweeting level

From March 1, 2013 to January 31, 2014, the members of the "respondents" panel posted 840,251 tweets. 81,606 of these messages contained political occurrences,[7] which is 9.7 % of tweets posted. The statistical distribution of these messages shows very irregular tweeting patterns. Differences were noted not only with regard to the points in time of the postings – the number of political tweets posted by the panel varies considerably from one day to the next – but also among individuals.

The number of political messages posted by Twitter users is in fact very uneven from one person to another. Over the entire period, 3% of the panel members alone produced over half the political messages posted. The number of political messages posted therefore turns out to be very unevenly distributed over the population. Statistical analysis isolated several variables that help to explain this unevenness. The first variable is age. A significant correlation was noted between the users' age and the total number of political tweets they posted.

Table 10. Average number of political tweets posted by age group

Age Group	n =	Average number of tweets
Under 18	43	25
18 to 25	230	78
26 to 34	108	75
35 to 45	103	216
46 to 60	52	229
61 and over	22	921
Total	558	146

---

[7] A list of 121 political terms was drawn up, including the name of the main parties, institutions and politicians.

The number of political tweets posted thus increases with age. Although it has already been pointed out that the older age groups are considerably underrepresented in the sample population (the over 61 make up 4 % of the "respondents"), they are nevertheless the age group that produces the most political tweets on average. In other words, there are fewer people over the age of 61 who produce political tweets, but this minority of tweet producers is the group that produces the most tweets on average. The older age groups are thus the biggest message producers. These individuals seem to combine four salient characteristics at once, which explains their high level of message production: a greater age than the average age of Internet users (41 in 2012), a lack of professional and family constraints and thus more free time, a high level of politicization and mastery of the Internet and computer technology – often through self-training.

A second variable significantly correlated with number of political messages posted is the respondents' level of politicization. A correlation was thus noted between self-declared level of interest in politics and the number of political tweets posted. Respondents who said they are "very" interested in politics post on average eight times more political tweets than those who declare "little" or "no" interest and four times as many as those who claim to have "some" interest.

Table 11. Declared interest in politics and number of political tweets posted
(Spearman test)

Variable	Group	$n =$	Average number of tweets	Test t	p-value	$p < 0.05 = $ *
Number of political tweets posted	Little or no interest	72	29.14	0.272	0.000	*
	Some interest	171	53.59			
	Very interested	315	223.31			
	Total	558	146.25			

People the most interested in politics are those who post the most political messages. While this finding is hardly surprising, it is nevertheless interesting to note that interest in politics, which is a self-reported variable, objectively determines the number of political messages posted. In the case at hand, the data gleaned from observation corroborates the data generated by the survey questionnaire. Likewise, membership in a political party also determines the number of political tweets produced. Party activists thus posted twice as many

political tweets on average during the period (225 tweets, or 20 tweets per month) than non-activists (109 tweets, or 10 tweets per month).

Table 12. Comparison of the number of political tweets posted between activists and non-activists (Spearman test)

Variable	Group	n =	Average number of tweets	Test *t*	*p*-value	*p* < 0.05 = *
Number of political tweets posted	Activists	178	225.30	2.256	0.024	*
	Non-activists	380	109.22			
	Total	558	146.25			

The third variable that significantly influences the number of political messages posted is the users' level of activity on the social media platform. A number of indicators can serve to measure this level of activity, in particular the total number of messages posted. The total number of tweets posted counts all messages posted by the respondents, whether political or not. A high number of messages then suggests that the user is very active on the network and is used to posting messages regularly and communicating with other social media users. Postulating that the total number of tweets posted determines the number of political tweets may at first seem tautological, as political tweets are necessarily included in the total number of tweets posted. The opposite does not prove to be true, however: posting a large number of tweets does not necessarily imply posting a large number of political tweets. The total number of tweets posted turns out to be the variable the most strongly correlated with number of political tweets produced. It would appear that the necessary (but not sufficient) condition for producing political tweets is the regularity of tweeting. Being politically "talkative" on Twitter presupposes being "talkative" on the social media platform in general.

Table 13. Total number of tweets and number of political tweets (Pearson's test)

Population	Variable	n =	r	*p*-value	*p* < 0.05 = *
Respondents	Number of political tweets	558	0.437	0.000	*
	Total number of tweets				
Non-respondent panel	Number of political tweets	517	0.309	0.000	*
	Total number of tweets				

Users' age, level of politicization and level of activity on the social network are thus all variables that explain strong individual discrepancies in posting political messages.

Impact of the political context on tweeting level

The number of political tweets posted is not only very uneven from one person to another, but it is also very irregular from one day to the next. The number of political tweets, as well as the number of individuals who post them, varies considerably according to the date the messages are posted on the social network.

Figure 1. Distribution of political tweets posted by the panels by day from March 1, 2013 to January 31, 2014
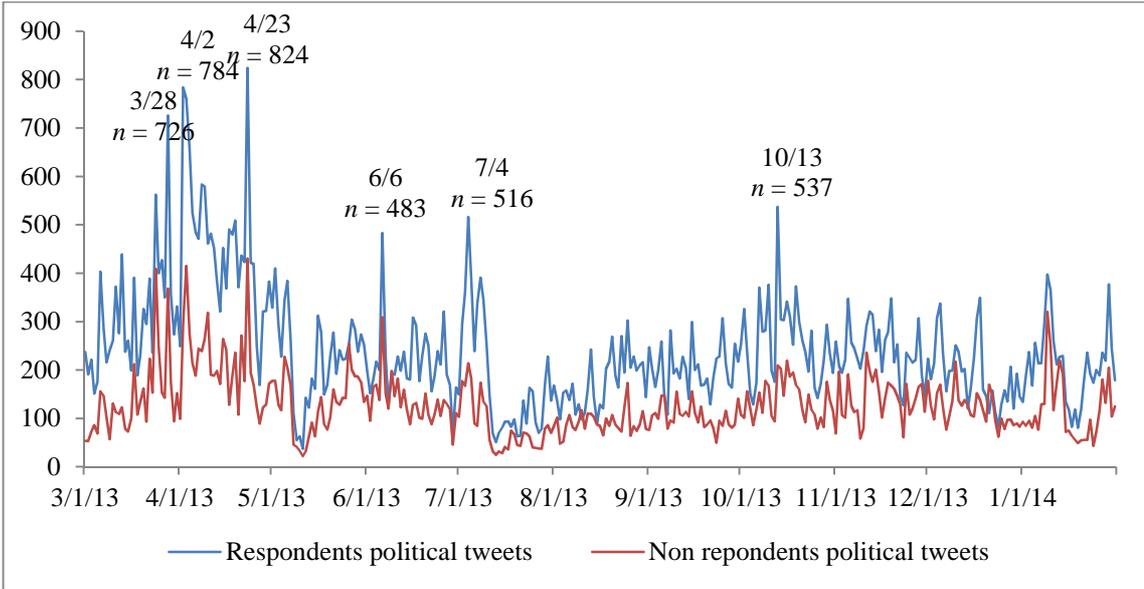
As is the case in the previously cited research, the above figure indicates clear spikes in political message tweets. Except for the month of April, these peaks are fairly spaced out in time and are also very short in duration: the volume of tweets generally plummets right after the spike occurs. Eight different peaks were isolated and tweet samples posted on those days were examined. Each of the peaks in fact coincides with a political "event" that occupied headlines at the time (scandals, revelations regarding the private life of the head of state, law on same-sex marriage, demonstrations against this law). In France as elsewhere, political topics discussed by social media users seem to be partly defined by the mainstream media.

The number of political tweets posted thus depends considerably on the political and media context.

Table 14. Highest daily volume and rate of production by the "respondents" panel

Date	Number of tweets	% of the panel	Event
April 23, 2013	824	27	National Assembly passes the law on same-sex marriage
April 2, 2013	679	25	Jérôme Cahuzac admits having cheated on his taxes
March 28, 2013	726	19	Televised interview with François Hollande on France 2 public television evening news
April 8, 2013	583	19	Asset declarations published by certain ministers
June 6, 2013	483	19	Death of Clément Méric, young extreme leftwing activist
March 24, 2013	562	17	"Manif pour tous" demonstration against same-sex marriage: 300,000 to 1.4 M participants + Legislative by-election UMP/FN Oise
July 4, 2013	516	17	Nicolas Sarkozy's campaign accounts disqualified for partial reimbursement
January 9, 2014	397	15	*Closer* magazine reveals Hollande/Gayet affair

What categories of users are the most sensitive to these political "events"? Are the most politicized the ones who have a keen interest in extraordinary political "events"? Or on the contrary is it the least politicized users who only express themselves politically when political moments coincide with social moments? A comparison of the evolution of the day-to-day tweeting rate of "respondents" who say they are "very" interested in politics with the daily tweeting patterns of the other "respondents" who claim to have "some, a little or no" interest in politics produces two fairly symmetrical curves. In other words, on different scales, the uneven day-to-day political tweeting patterns are the same among the most politicized and the least politicized categories of the panel (who remain, it should be remembered, considerably more politicized than the average French citizen). "Political-media" events thus seem to generate unusual levels of postings among both the most and the least politicized members of the panel. In short, the level of political messages posted on Twitter depends at once on individual variables (age, level of politicization, level of activity on the network) and much more appreciably on the political and media context.

Figure 4.2 Daily evolution of the respondent tweeting rate by declared interest in politics
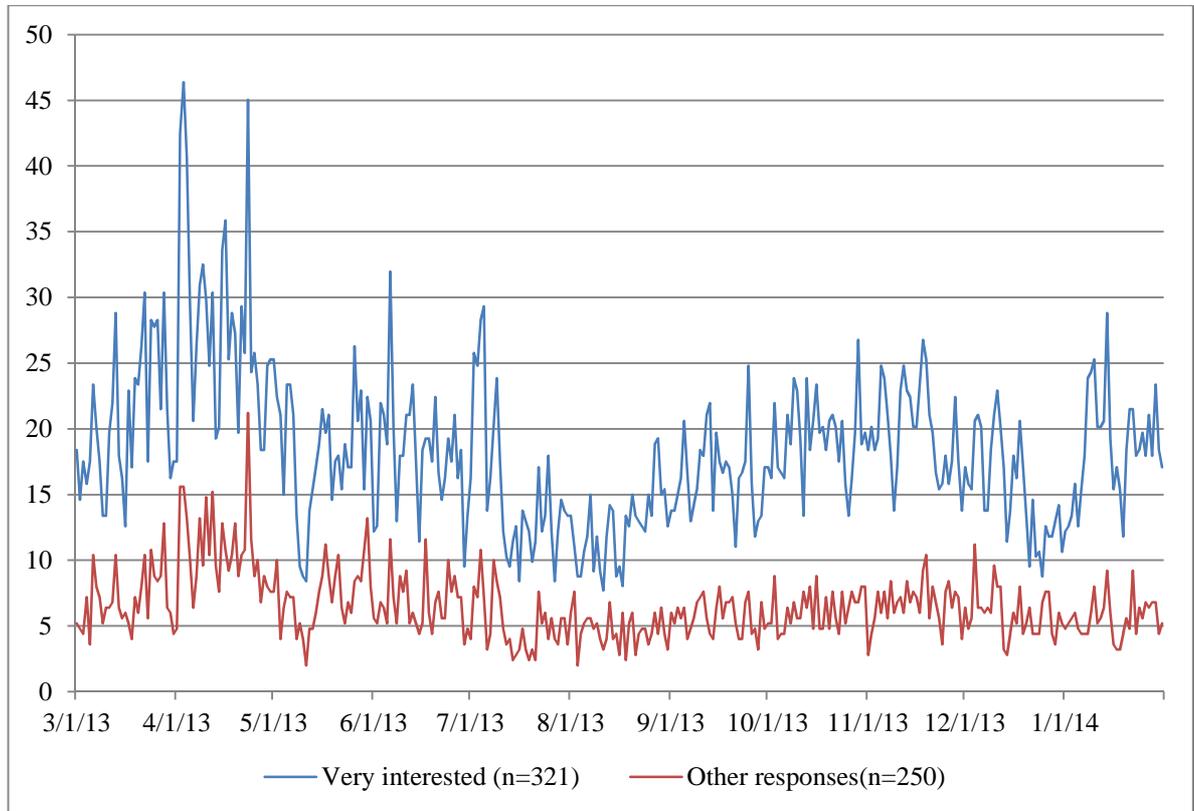
Conclusion

Combining a web-tracking analysis approach with a classic questionnaire survey approach thus enables researchers to profit from the methodological potential the Internet offers while allowing them to situate the data collected in the "real" social space. This methodology has thus made it possible to show the extent to which Twitter users are the product of considerable social selection, as well as the influence of sociological variables on user posting practices and habits. Future research should investigate the impact of sociological variables on the form and content of the messages collected.

Bibliography

AMPOFO, Lawrence, ANSTEAD, Nick and Ben O'LOUGHLIN (2011),"Trust, Confidence, And Credibility: Citizen Responses on Twitter to Opinion Polls during the 2010 UK General Election," *Information, Communication & Society*, 14(6), pp. 850-871.

ANSTEAD, Nick and O'LOUGHLIN, Ben (2011),"The Emerging Viewertariat and BBC Question Time: Television Debate and Real-Time Commenting Online," *The International Journal of Press/Politics*, 16(4), pp. 440–462.

AUSSERHOFER, Julian and MAIREDER, Axel (2013), "National Politics on Twitter: Structures and Topics of a Networked Public Sphere," *Information, Communication & Society*, 16(3), pp. 291-314.

BEKAFIGO, Marija and McBRIDE, Allan (2013),"Who Tweets about Politics? Political Participation of Twitter Users during the 2011 Gubernatorial Elections," *Social Science Computer Review*, 31(5), pp. 625-643.

DIMITROVA, Daniela, SHEHATA, Adam, STRÖMBÄCK, Jesper and NORD, Lars (2014), "The Effects of Digital Media on Political Knowledge and Participation in Election Campaigns: Evidence from Panel Data," *Communication Research*, 41(1), pp. 95-118.

GAINOUS, Jason and WAGNER, Kevin (eds) (2014), *Tweeting to Power: The Social Media Revolution in American Politics*, Oxford, Oxford University Press.

GAYO-AVELLO, Daniel (2013), "A Meta-Analysis of State-Of-The-Art Electoral Prediction from Twitter Data," *Social Science Computer Review*, vol. 31, no. 6, pp. 649-679.

IFOP (2013), "Observatoire des réseaux sociaux".

JUNGHERR, Andreas and JÜRGENS, Pascal (2014), "Through a Glass, Darkly: Tactical Support and Symbolic Association in Twitter Messages Commenting on Stuttgart 21," *Social Science Computer Review*, 32(1), pp. 74-89.

JUNGHERR, Andreas (2014), "The Logic of Political Coverage on Twitter: Temporal Dynamics and Content," *Journal of Communication*, Vol. 64, no. 2, pp. 239-259.

JUNGHERR, Andreas, "Tweets and Votes, A Special Relationship: The 2009 Federal Election in Germany," in WEBER, Ingmar, POPESCU, Ana-Maria and PENNACCHIOTTI,

Marco, (eds), *Proceedings of the 2nd Workshop Politics, Elections and Data*, New York, NY: ACM, pp. 5-14.

JUNGHERR, Andreas, "Twitter in Politics: A Comprehensive Literature Review," Available on line: SSRN: http://ssrn.com/abstract=2402443.

LARSSON, Anders and MOE, Hallvard (2013),"Twitter in Politics and Elections: Insights from Scandinavia," in WELLER, Katrin, BRUNS, Axel, BURGESS, Jean, MAHRT, Marja and PUSCHMANN, Cornelius (eds), *Twitter and Society*, New York, NY: Peter Lang Publishing, pp. 139-330.

LINH, Dang-Xuan, STIEGLITZ, Stefan, WLADARSCH, Jennifer and NEUBERGER, Christoph (2013), "An Investigation of Influentials and the Role of Sentiment in Political Communication on Twitter during Election Periods," *Information, Communication & Society*, 16(5), pp. 795-825.

McCOMBS, Max and SHAW, Donald (1972), "The Agenda-Setting Function of Mass-Media," *Public Opinion Quarterly*, no. 36, pp. 176-187.

NORRIS, Pippa (1993), "The Gender Generation Gap in British Elections," *British Elections and Parties Yearbook*, Vol. 3, Issue 1, pp. 129-142.

RAINIE, Lee, SMITH, Aaron, SCHLOZMAN, Kay, BRADY, Henry and VERBA, Sidney (2012), "Social Media and Political Engagement," *Pew Internet & American Life Project*.