Tweets and Votes, a Special Relationship

The 2009 Federal Election in Germany

Andreas Jungherr
University of Bamberg, Germany
Feldkirchenstrasse 21
96045 Bamberg, Germany
andreas.jungherr@gmail.com

ABSTRACT

As the microblogging service Twitter becomes an increasingly popular tool for politicians and general users to comment on and discuss politics, researchers increasingly turn to the relationship between tweets mentioning parties or candidates and their respective electoral fortunes. This paper offers a detailed analysis of Twitter messages posted during the run-up to the 2009 federal election in Germany and their relationship to the electoral fortunes of Germany's parties and candidates. This analysis will focus on four metrics for measuring the attention on parties and candidates on Twitter and the relationship to their respective vote share. The metrics discussed here are: the total number of hashtags mentioning a given political party; the dynamics between explicitly positive or explicitly negative mentions of a given political party; the total number of hashtags mentioning one of the leading candidates, Angela Merkel (CDU) or Frank-Walter Steinmeier (SPD); and the total number of users who used hashtags mentioning a given party or candidate. The results will show that during the campaign of 2009 Twitter messages commenting on parties and candidates showed little, if any, systematic relationship with subsequent votes on election day. In the discussion of the results, I will raise a number of issues that researchers interested in *predicting* elections with Twitter will have to address to advance the state of the literature.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; J.4 [**Computer Applications**]: Social and Behavioral Sciences—*Sociology*

Keywords

Computational Social Science; Electoral Forecasting; Twitter

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
PLEAD'13, October 28, 2013, San Francisco, CA, USA.
Copyright 2013 ACM 978-1-4503-2418-2/13/10
http://dx.doi.org/10.1145/2508436.2508437 ...$15.00.

1. INTRODUCTION

With the ever increasing use of Twitter by parties, politicians, political supporters and the general public during political campaigns, researchers have become increasingly interested in the question if success on Twitter (usually measured by the total or relative volume of Twitter messages mentioning a given party or candidate) can be taken as indicator for electoral success. The seeming success of some early studies in finding correlations between some Twitter metrics and some measures of electoral success led some researchers to enthusiastically proclaim the potential to predict election results by simply counting Twitter messages [26, 7, 6, 2]. This position has come increasingly under attack, be it because of the lack of a theoretical connection between the metric of choice (Twitter messages) and the outcome of interest (electoral success) or be it because of the methods its proponents used to collect and analyse the underlying data [15, 18, 12, 9, 10, 13, 21, 20, 11, 8, 24].

This paper adds to the growing literature on the relationship between Twitter messages and electoral success by examining Twitter messages that were posted during the run-up to the 2009 federal election in Germany and their relationship to the electoral fortunes of Germany's parties and candidates. This paper will discuss four metrics for measuring the attention on parties and candidates on Twitter and will show how these metrics correspond with the actual vote share of each party. These metrics are:

- The total number of hashtags mentioning a given political party;
- The dynamics between explicitly positive or explicitly negative mentions of a given political party;
- The total number of hashtags mentioning one of the leading candidates, Angela Merkel (CDU) or Frank-Walter Steinmeier (SPD);
- The total number of users who used hashtags mentioning a given party or candidate.

2. WHY GERMANY?

The 2009 federal election in Germany offers an interesting case to examine the relationship between Twitter success and electoral fortunes for two reasons: In 2009, Germany witnessed the growing success of a new political party, *Die Piratenpartei* (The Pirate Party). In the spring and summer of 2009 the German Pirate Party experienced a sudden

increase in membership as plans by the outgoing governing coalition of CDU and SPD to block access to websites suspected of hosting child pornography met with heavy resistance. The pinnacle of which was a very popular and highly publicised online-petition to the German parliament to stop the proposed law [14]. Suddenly Internet regulation was a popular topic in Germany's policy debate. The Pirate Party profited massively by this sudden attention on Internet policy. Especially for researchers interested in the relationship between online buzz and subsequent electoral results the Pirate Party offers an interesting case. Here we have a party which, at least in the 2009 federal election, did not have significant success at the polls. Still, because of its roots in online activism, the Pirate Party dominated the political online sphere and had a heavy presence in the traditional media during the run of the campaign. A minor actor in the traditional political sphere was a major actor in the political sphere online at the same time. How did this constellation affect the linkage between a party's online buzz and its subsequent electoral success?

A second reason, why Germany offers an interesting case, is that one of the early studies examining potential connections between Twitter success and subsequent electoral success focused on the campaign for the 2009 federal election in Germany [26]. The authors state very confidently that by mere counting of messages mentioning political parties one would be able to successfully predict the election results of these parties. This claim has been contested based on the decision of the authors to include in their analysis only political parties that managed to collect enough votes to be represented in parliament. This choice led the authors to exclude mentions of the above mentioned Pirate Party. If the Pirate Party had been included in the analysis the precision of the prediction would have suffered significantly. The exclusion of the Pirate Party raises the question why one would examine the connection between the share of verbalised public attention online with votes cast on election day but a priori exclude the political group with the strongest share of voice online from the analysis. Also, the precision of the predicted election results was shown to be dependent on the time span included in the analysis [18].

Both these reasons–the existence of a party that showed a strong online presence but only limited success at the polls and the existence of a highly influential early study–make the case of the 2009 federal election in Germany an interesting one to examine the relationship between Twitter messages and electoral success.

3. DATA COLLECTION

The data for this paper were collected in close collaboration with Pascal Jürgens from the University of Mainz. We regularly queried the Twitter API during the months directly preceding the 2009 federal election in Germany (mid-June to early October 2009). We decided to focus on collecting all messages of a sample of politically vocal Twitter users, those who, during the run-up to the campaign, had at least once used one of 19 pre-determined politically relevant hashtag (e.g. names of parties and candidates or campaign related hashtags). Once a user had posted a message with one of these hashtags we then collected all her previous and future messages. We did this on one hand to be able to assess the role politics played in relation to all other messages of these users, and on the other hand to capture other potentially politically relevant messages that did not happen to contain one of the political hashtags identified by us prior to the data collection. In doing this, we thus identified all users who used the names of the leading candidates or the political parties in their messages at least once by using a hashtag. Between June 18 and October 1, 2009, 32.731 users used one or more of the pre-identified political hashtags in their messages. During this time span these users posted a total of 10.085.982 messages.

There is the possibility that, by focusing only on political mentions in hashtags and not on topically relevant words, we created a dataset biased in favour of experienced Twitter users, those who knew how to use hashtags to contribute to the issue public commenting on the campaign. This might have created a dataset that underestimates the total number of politically relevant Twitter messages. Still, we believe this tradeoff is justified as the focus on hashtags reduced the risk of polluting the data by a large number of false positives (i.e. messages using politically relevant words or phrases, such as "Angela" or "Wahl", that can also refer to topics outside of politics) and spam.

To get a better, more structured understanding of the communication on Twitter, it helps to differentiate between messages that were posted in support of parties or candidates and messages posted in opposition to them. In the past some studies have used manually coded support vector machines (SVM) to identify the tonality of political Twitter messages [3]. For the purposes of this analysis, I propose a simpler approach enabled by a special feature of the German twitter-sphere. Starting in early 2009 the German communication consultant Sascha Lobo introduced the service *Wahlgetwitter* . The website tracked the use of names of parties in combination with the suffixes + and - in hashtags. The name of a party followed by a plus-sign was meant to signal support of a party (e.g. #piraten+), while the name of a party followed by a minus-sign was meant to signal opposition to it (e.g. #piraten-). A website calculated and documented the total sums of positive and negative mentions of political parties per day. This tool was communicated widely on the web and in traditional media. This led to the growing adoption of the usage convention during the campaign. In the dataset, we find a total of 11.212 users who posted messages with a #partyname followed by a plus or a minus. 224.551 messages contain hashtags using this convention.[1] The ratio between the use of supportive and opposing hashtags can be used as an indicator for the nature of political commentary during the campaign for the 2009 federal election in Germany.

4. TWITTER SUCCESS AND ELECTORAL FORTUNES

Papers examining the relationship between politically relevant Twitter messages and subsequent electoral results usually focus on the number of messages a political actor, be it a party or a candidate, was mentioned in and the subse-

[1]This number is the sum of all messages that contained at least one of the following hashtags: #cdu+, #cducsu+, #csu+, #spd+, #fdp+, #diegrünen+, #gruene+,#gruenen+, #grüne+, #grünen+, #dielinke+, #linke+, #linkspartei+, #piraten+, #piratenpartei+, #cdu-, #cducsu-, #csu-, #spd-; #fdp-, #die_gruenen-, #gruene-, #gruenen-, #grüne-, #grünen-, #die_linke-, #dielinke-, #linkspartei-, #piraten-, #piratenpartei-.

quent number of votes this actor actually received. Thus, this metric is a promising starting point for the examination of the relationship between Twitter messages and election results. In the following analysis, I focus exclusively on those cases when a political actor was named explicitly in a hashtag (e.g. #cdu, #spd, #grüne, #piraten, #merkel, #steinmeier et al.). All potential mentions of political actors in the text of a Twitter message without a hashtag are thus excluded from the analysis. This might lead to an underestimation of the total number of messages commenting on parties or candidates but it should not impact the identification of relative dynamics between mentions of different political actors. Since there was significant variation in the spelling of political actors, I collected the most prominent hashtag variations referring to each actor in encompassing concepts (e.g. #grüne, #gruene, #bündnis, #buendnis et al.). These concepts sum up all appearances of relevant hashtags commenting on the respective political actor.[2] For this analysis, the occurrence of all hashtags collected in these concepts were summed up for each day between June 18 and October 1, 2009.

The basic argument of optimistic studies on the relationship between Twitter messages and election results is: more messages on Twitter mentioning a party or a candidate are indicative of more votes on election day. Is this true for the 2009 federal election in Germany? Figure 1 shows the comparison between the share of hashtag mentions selected parties received and the actual vote share the parties received on election day.[3] The plot on the left shows the share of hashtag mentions each party concept received from June 18 to October 1, 2009 relative to the sum of mentions all six received. The plot on the right shows the votes each of these parties received on election day September 29, 2009 relative to the sum of all votes the six parties could collect.[4]

Figure 1 clearly shows that mentions of the Pirate Party dominated by far the political discourse on Twitter. If the simple argument *more tweets mean more votes* would be true, from 2009 to 2013 the Pirate Party would have been the leading partner in German government. Still, one could argue that maybe the Pirate Party should be treated as an outlier. We also see that a prognosis based on Twitter messages would get the order of parties wrong. CDU/CSU, SPD and FDP were the parties with the most mentions on Twitter (apart from the Pirate Party) and the three parties that gathered the most votes, but for Bündnis 90/Die Grünen and Die LINKE this relationship is not stable. Bündnis 90/Die Grünen received more mentions on Twitter than Die LINKE but less votes on election day. So even if we exclude the Pirate Party from our analysis, the relationship *more tweets mean more votes* is not stable.[5]

Table 1: Hashtag share between June 18 and October 1, 2009 compared to vote share (excluding mentions of and votes for the Pirate Party from totals of hashtag and vote share)

party	# share	vote share	difference
CDU/CSU	31	36	-5
SPD	28	25	+3
FDP	19	15	+4
Bündni90/Die Grünen	14	11	+3
Die LINKE	8	13	-5

Table 1 shows the relationship between mention and vote share while excluding the Pirate Party from the analysis. We see that estimating the vote share of SPD, FDP and Bündnis 90/Die Grünen based on their hashtag share would have led to an overestimation of their strength on election day while an equivalent estimation of CDU/CSU and Die LINKE would have led to an underestimation of their vote share. Taking these findings into account–the absolute dominance of the online-sphere by the Pirate Party while this party remained an insignificant actor on election day, the unstable relationship between the ranking of parties based on mentions and votes and finally the clear under- and overestimation of vote share based on prognoses based on Twitter mentions–this shows that the total amount of mentions on Twitter had, if at all, a very weak connection to election results of political parties.

A crucial question researchers trying to determine the relationship between Twitter messages and votes have to address is which messages to include in their analysis and which to leave out. While the question of topical relevance can be easily addressed (for example by including all messages using politically relevant words, phrases or hashtags) the question of relevant time spans is much harder to answer. Is the relationship of mentions to vote share stable over a long period of time, or do we have to focus only on messages close to the election? Researchers have to explicitly address why they focus on messages posted during a given time span and not another. Unfortunately this question is raised seldom in the existing literature. Only if the share of mentions of a given political actor would remain stable over time one could ignore this point. The data used in this analysis offer an interesting perspective on the dynamics of the daily

[2]The following hashtags were collected in concepts: CDU/CSU: #cdu, #cducsu, #csu; SPD: #spd; FDP: #fdp; Bündnis 90/Die Grünen: #buendnis90, #bündnis, #bündnis90, #bündnis90diegrünen, #bündnis90grüne, #bündnisgrüne, #bündnisgrünen, #die_gruenen, #die_grünen, #diegrünen, #gruene, #grüne, #grünen; Die LINKE: #die_linke, #dielinke, #linke, #linkspartei; Piratenpartei: #piraten, #piratenpartei; Angela Merkel: #angie, #merkel, #angie_merkel, #angelamerkel, #angela_merkel; Frank-Walter Steinmeier: #steinmeier, #fws, #frank_walter_steinmeier, #steini, #frankwaltersteinmeier, #frank_steinmeier. This collection might still exclude some more exotic or more ambiguous spelling variations of the political actors in question. Still, this collection should account for the vast majority of the hashtags referring to the political actors in question and thus should offer a comprehensive view on the dynamics between them.

[3]The parties selected for this comparison are five parties that managed to gather enough votes to enter the German parliament (i.e. CDU/CSU, SPD, FDP, Bündnis 90/Die Grünen, Die LINKE) and the Piratenpartei. The Pirate Party was included in the analysis despite its subsequent failure to enter parliament since it dominated the political online-sphere and received extensive media coverage in the run-up to the election.

[4]The amount of votes used in this analysis is based on the official vote-count reported by the *Bundeswahlleiter* [5].

[5]In the past, some studies have used regression analysis to determine the strength of the relationship of tweets to votes. The German case does not offer itself for such an analysis since the number of parties, and therefor also the number of observations, is too small to built a meaningful regression model.

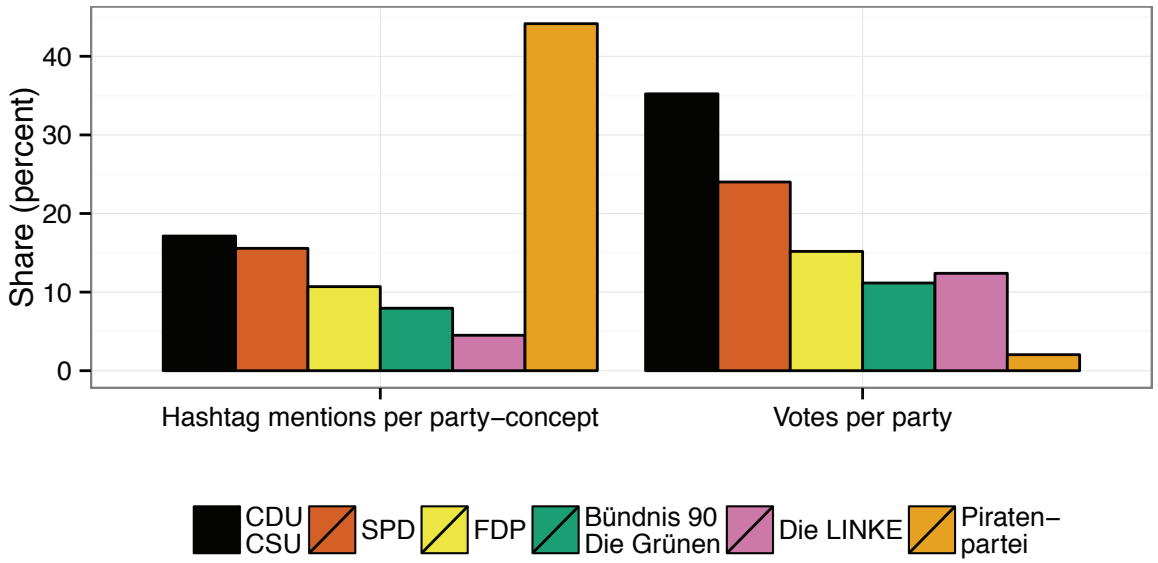

Figure 1: Hashtag share compared to vote share

mentions of the respective parties during the three months preceding the election.

For this step of the analysis, I calculated the sum of all hashtag mentions for Germany's established parties (i.e. CDU/CSU, SPD, FDP, Bündnis 90/Die Grünen and Die LINKE). Figure 2 shows the daily share each of the five parties had of the sum of their mentions. We see different trends in the day-to-day development of this mention share for different parties. Some parties remain relatively stable in their share of hashtag mentions over time (i.e. FDP and Die LINKE). In contrast the shares of CDU/CSU and SPD are wildly fluctuating from day to day. A third pattern is seen with Bündnis 90/Die Grünen. At the beginning of the time series, the party holds a somewhat higher share of the daily hashtag mentions. Towards the end of the time series, the party loses somewhat in the daily share. Whatever the reasons for these different dynamics might be, the heavily fluctuating shares in hashtag mentions show that any attempt to draw conclusions of coming election successes based on the share of Twitter messages is highly dependent on the time span one uses to calculate the share of Twitter messages. Any relationship between the number of mentions on Twitter and votes received on election day is thus far from stable.

Up until now, this analysis focused only on neutral hashtags mentioning political parties. Without any further knowledge about the content of the messages containing one of the neutral hashtags mentioning a party, one could assume that the positive and negative mentions of parties would be roughly equal. Given this assumption, one could assume that high hashtag counts of parties would also speak for high support. A feature special to the political Twitter sphere in Germany allows us to test this assumption. As mentioned above, during the campaign of 2009 users started to explicitly mark their messages in support or in opposition of one of the parties in the race. These hashtags were a combina-

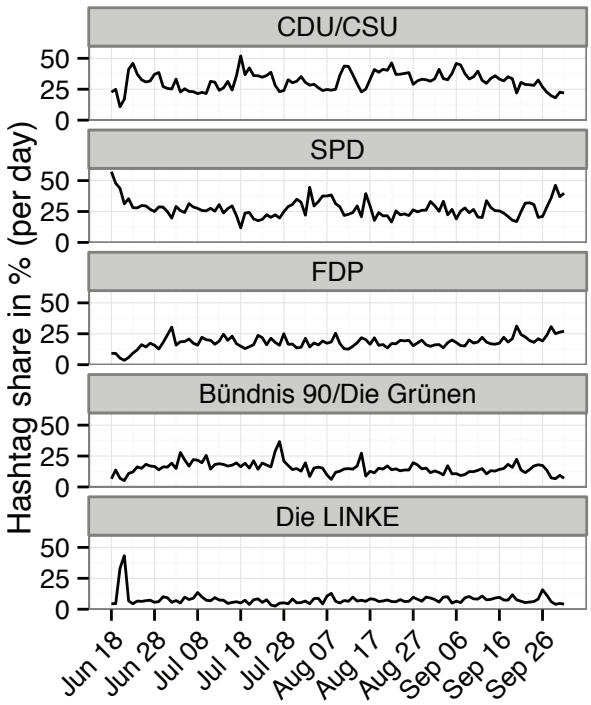

Figure 2: Hashtag share per party over time

tion of the name of a party followed by a + or a - sign (e.g. #cdu+ or #fdp-). This convention helps us to understand the relationship between positive and negative mentions on Twitter. Analogue to the summation of hashtags referring to the same party, I constructed party concepts for explicitly positive or negative mentions.[6] Figure 3 shows the results.

Examining Figure 3 we see at once, most mentions of parties in combination with a plus or a minus sign are identifying tweets in opposition to a party. The negative mentions of each party vastly outnumber positive mentions. This is especially true on election day. Most mentions of parties on Twitter thus refer to a party in a negative context. This is true for all parties, even the Pirate Party. While the dynamic between #partyname+ and #partyname- is probably no valid base to extrapolate from this to the relationship between positive or negative comments in tweets only identified by a neutral hashtag, it is still an interesting indicator. Twitter is a channel where users voice political opposition much more frequently than political support. It seems reasonable to assume that any attempt at using the total count of party mentions in neutral hashtags on Twitter to estimate votes on election day thus counts a significant amount of messages that were posted in opposition to a party. Heavy opposition to a party online would thus become an indicator for its success offline. This is a paradox that advocates of this approach should at least address in their work.

The preceding discussion showed that, at least during the campaign for the 2009 federal election in Germany, a party's relative success on Twitter (measured in hashtag mentions) did not necessarily translate to electoral success. But is that also true for hashtag mentions of the leading candidates of CDU/CSU and SPD, Angela Merkel and Frank-Walter Steinmeier? Did the dynamics of their hashtag mentions offer early insight in the electoral results? For this analysis, I collected various hashtags referencing one of the candidates in two candidate concepts. In this analysis, I only included neutral hashtags.[7] If we focus only on the total amount of hashtag mentions we find that Angela Merkel was mentioned 7.583 times between June 18 and September 26, 2009 (the day before the federal election). Frank-Walter Steinmeier was mentioned 5.781 times during the same time span. If we are content with stating that the candidate who ultimately won the election was mentioned more often during the run-up to the election, this could count as a success. Still, a closer examination of when the candidates were mentioned

[6]The following hashtags were collected in party concepts: CDU/CSU+: #cdu+, #cducsu+, #csu+; SPD+: #spd+; FDP+: #fdp+; Bündnis 90/Die Grünen+: #diegrünen+, #gruene+, #gruenen+, #grüne+, #grünen+; Die LINKE+: #dielinke+, #linke+, #linkspartei+; Piratenpartei+: #piraten+, #piratenpartei+. CDU/CSU-: #cdu-, #cducsu-, #csu-; SPD-: #spd-; FDP-: #fdp-; Bündnis 90/Die Grünen-: #die_gruenen-, #gruene-,#gruenen-, #grüne-, #grünen-; Die LINKE-: #die_linke-, #dielinke-, #linkspartei-; Piratenpartei-: #piraten-, #piratenpartei-. This list somewhat diverges from the hashtags collected in the original party concepts. The reason for this is that not all hashtags that were used to identify parties were used in combination with + or - signs.

[7]The following hashtags were included: Angela Merkel: #angie, #merkel, #angie_merkel, #angelamerkel, #angela_merkel; Frank-Walter Steinmeier: #steinmeier, #fws, #frank_walter_steinmeier, #steini, #frankwaltersteinmeier, #frank_steinmeier.

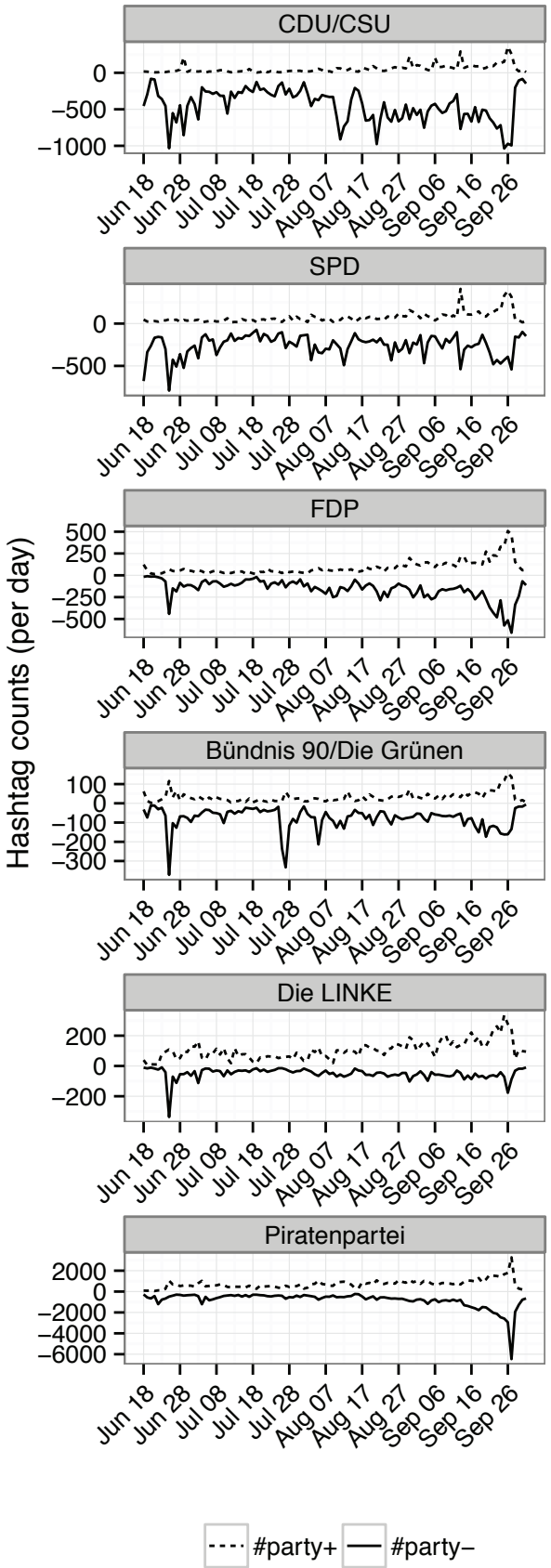

Figure 3: Hashtags count of #party concept+, #party concept- over time

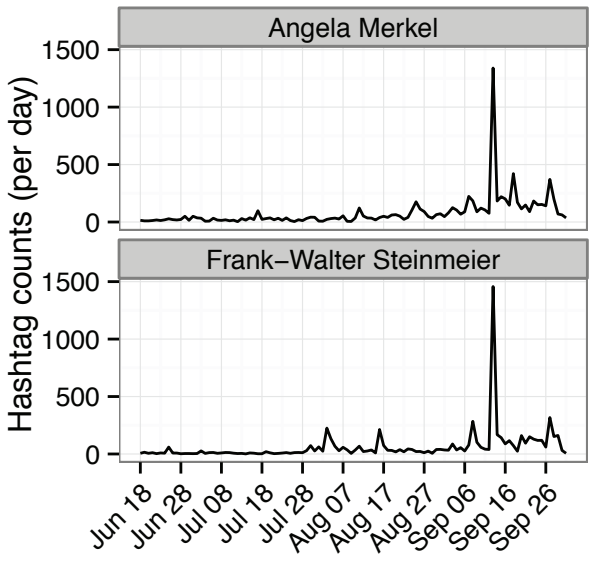

Figure 4: Hashtag count of candidate concepts over time

offers a more detailed view into the dynamics of candidate mentions in hashtags.

Figure 4 shows the hashtag mentions of Angela Merkel and Frank-Walter Steinmeier. Both time series show similar patterns. We see that both candidates have a largely stable baseline of mentions. Before the televised debate on September 13 both candidates are usually mentioned in hashtags less than 100 times per day. After the televised debate this baseline fluctuates roughly between 100 and 200 mentions. We see that Angela Merkel is mentioned somewhat more often than Frank-Walter Steinmeier but only in aggregate these differences amount to much. Another prominent feature of both time series are sudden spikes in mention volume. On some days both candidates are mentioned much more often than the baseline would lead us to expect. The most prominent examples are September 13, the day of the televised debate between the two leading candidates and September 27, the day of the election. Other strong outliers are documented in Table 2 with corresponding events, relevant to the campaign.

Table 2 shows that the events that drive the use of the candidates' names in hashtags are predominantly TV appearances, controversies and pseudo-events, big events created by the campaign for the purposes to create media coverage[1]. Towering above all these mediated events are the spikes in messages during the televised debate and election day. Both these events correspond with the concept of media events, scheduled events intensely covered by traditional media with ritualistic social meaning[4]. So, while optimistic researchers claim that there is a relationship between more tweets and more votes, what they really seem to be saying is: there is a relationship between more TV coverage and more tweets. This would make Twitter data a research instrument for determining the resonance of political TV coverage and the detection of political controversies relevant to Twitter

Table 2: High volume of hashtags mentioning Angela Merkel or Frank-Walter Steinmeier and corresponding events during the campaign

date	candidate	event
July 17	Merkel	Merkel's birthday
August 3	Steinmeier	announcement of program
August 11	Merkel	controversial poster
August 16	Steinmeier	TV interview
August 25	Merkel	Ackermann controversy
September 3	Merkel	party convention
September 7	Merkel	TV interview
September 8	Steinmeier	TV interview
September 13	both	TV debate
September 18	Merkel	declines TV interview
September 27	both	election day

users, not so much an instrument to determine subsequent electoral fortunes of a candidate.

Instead of focusing on how often parties were referred to in hashtags, one could also count the number of users referring to a given party and then check if parties referred to by more users ended up having more votes. Figure 5 shows how many people used hashtags referring to a given party. In total 19.258 users referred to a party by posting at least one of the neutral hashtags listed above. The figure shows that this metric is an even less reliable indicator of a parties strength at the polls on election day than counting messages. Again, we find the Pirate Party dominating the communication space with 11.747 users commenting on it. The second strongest party is the SPD with 7.155 users directly followed by the FDP (6.694 users) and the CDU (6.679 users).[8] This order is not even close to the order of parties based on vote share.

For sake of completeness, Figure 6 shows the number of users who referred to Angela Merkel and Frank-Walter Steinmeier by hashtags. Only 4.252 users mentioned either candidate by using hashtags in their messages. As was the case in the total number of messages, Angela Merkel (3.140) was mentioned by more users than Frank-Walter Steinmeier (2.380).

5. CONCLUSION AND IMPLICATIONS FOR FURTHER RESEARCH

The results presented above raise a number of issues that should be addressed in further research into the relationship between Twitter messages and election results. First, we saw that for Germany's 2009 federal election the number of hashtag mentions a party received during the campaign was no valid indicator for the votes it would receive on election day. The strongest deviation between hashtag mentions and votes was seen in the case of the Pirate Party, a party that dominated the political discourse on Twitter but was all but insignificant on election day. The strength of a party offline was not mirrored by its strength online. This contradicts the expectations of the normalisation thesis, which expects that over time strong political actors offline would also dominate online discourse [19]. This is an interesting indicator

[8]Since some users referred to more than one party in their messages, the user counts of the parties do not sum up to the total number of users.

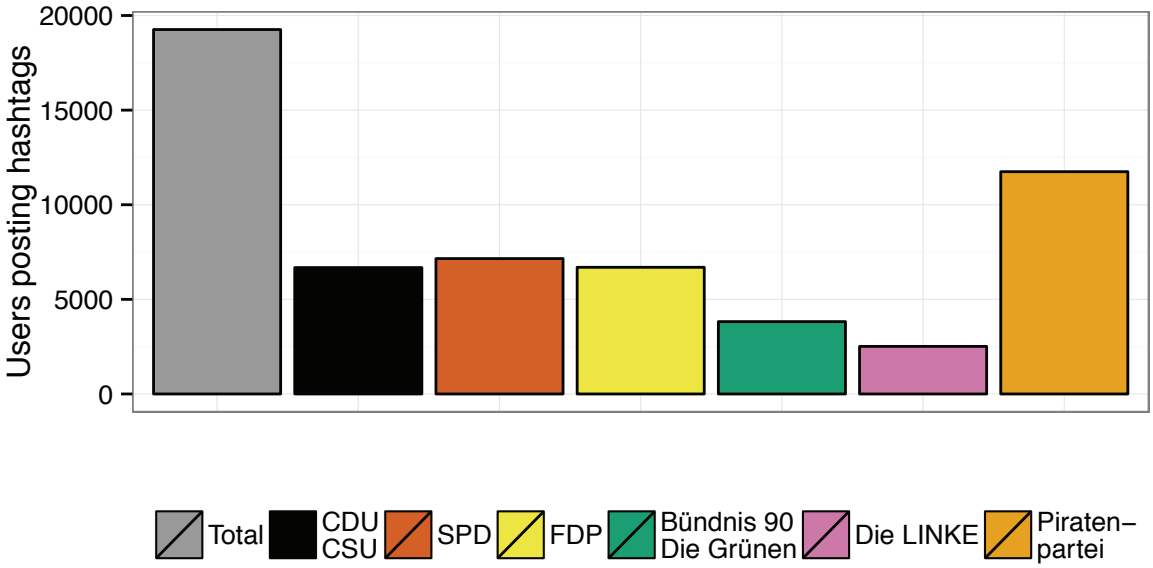

Figure 5: Users referring to parties in hashtags

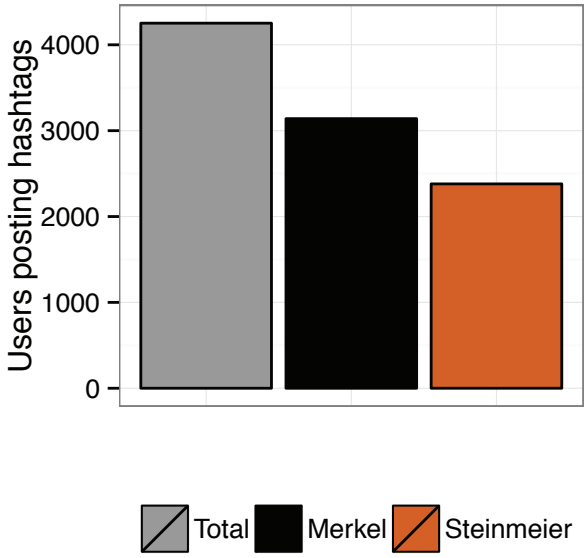

Figure 6: Users referring to candidates in hashtags

that the potential for *predicting* election results in a given political environment by counting Twitter messages heavily depends on a successfully completed normalisation process. This potential precondition should be further examined.

We also saw, that the share of hashtag mentions parties received was fluctuating from day to day. Any attempt to connect Twitter messages posted during a given time span to election results thus has to provide a discussion on why the respective time span was chosen. Potential patterns in these fluctuations might also offer interesting research objects, especially since a growing number of studies suggest that political offline events leave traces in the fluctuation of Twitter messages [15, 17, 16, 25].

The analysis also showed, once one focused on explicitly positive and explicitly negative comments on political parties, the negative mentions heavily outnumbered the positive ones. Thus, Twitter seems to be a medium used predominantly to voice political opposition rather than support. This raises an interesting paradox for researchers trying to construct systematic connections between Twitter messages and votes. Why should high counts of messages predominantly published in opposition to a party be indicative of its subsequent success on election day? This question shows that research in this field has to explicitly address possible theoretical links between the volume of Twitter messages (be they supportive or critical) mentioning a political party and its subsequent vote share. As of now, this discussion is largely absent from the literature.

When examining the use of the candidates' names in hashtags, we saw that the leading candidate of the party that won the election, Angela Merkel, was indeed mentioned more often than her challenger Frank-Walter Steinmeier (SPD). What at first might seem like good news for researchers interested in picking political winners based on tweet-counts was shown to be not quite so clear cut. Hashtags referring to the names of leading candidates were used predominantly in

reaction to TV appearances of the candidates, controversies or staged campaign events. Tweets thus become an indicator of TV appearances and other related campaign activity. The candidate who is appearing more often on TV, is campaigning more intensely and is creating more controversies thus seems to be the candidate receiving more mentions on Twitter. While all these characteristics might be sufficient conditions for a candidate to win on election day, clearly these are no necessary conditions.

Also, the number of users who referred to political parties by hashtags was not indicative of their electoral fortunes. Only when focusing on the number of users that mentioned either Angela Merkel or Frank-Walter Steinmeier in hashtags did I find that more users commented on the leading candidate of the party with the most votes, Angela Merkel, than her challenger, Frank-Walter Steinmeier.

All in all, this analysis shows that, at least for Germany's 2009 federal election, Twitter data did not prove to be a strong indicator for the electoral fortunes of parties or candidates mentioned on Twitter. The discussion showed that the literature on the relationship between tweets and votes has to address a series of fundamental issues. As of now, most of the literature focuses on documenting correlations between some measures of the attention a political actor (be they parties or candidates) received on Twitter and some measures of electoral success (be it the number of votes or vote share). The novelty factor of these results stems from the fact that these correlations seem somewhat counterintuitive. Why should there be a systematic relationship between the number of times a political actor was mentioned on a social media service, Twitter, and her later electoral fortunes? Especially, since, as of now, there is no indicator that in any country Twitter's user base is a representative sample of its whole population. The relevant literature has largely ignored this question. Maybe it would be possible to postpone work on this question if correlations between tweets and votes were shown to be stable. But if, as this paper has shown, these correlations are highly dependent on arbitrary selections by researchers (i.e. on which time intervals to focus or which political actors to include in the analysis) it becomes of crucial importance to address the implicit mechanism that should create a systematic relationship between tweets and votes. In this, research into the relationship between tweets and votes illustrates the limits of an empiricistic, exclusively data driven approach in the social sciences. An approach that recently has gathered some steam under the term *big data*. Especially in data rich contexts, and thus contexts with a high probability of spurious correlations, research has to be grounded in the theoretical development and data based examination of social mechanisms that lead to the emergence of specific data patterns. The discussion of these mechanisms is largely missing from the literature on the *prediction* of elections based on Twitter messages. This discussion is necessary before we can claim that showing some correlations between some Twitter messages and some election results is more than a surprising data artefact. For the developing discussion on the relationship between tweets and votes, the case of Germany's 2009 federal election offers a valuable counterpoint to some recent, more optimistic studies.

6. ACKNOWLEDGMENTS

The figures were produced using R [22] and the packages ggplot2 [27], scales [28], zoo [29] and xts [23]. I want to thank Pascal Jürgens for his crucial support in transforming the raw data for this study into something fit for analysis. I also want to thank Harald Schoen, Panagiotis Takis Metaxas and Daniel Gayo-Avello for extensive discussions on the topic of this paper.

7. REFERENCES

[1] D. J. Boorstin. *The image: a guide to pseudo events in America*. Harper & Row, 1961.

[2] A. Ceron, L. Curini, and S. M. Iacus. Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New Media & Society*, 2013.

[3] M. D. Conover, B. Goncalves, J. Ratkiewicz, A. Flammini, and F. Menczer. Predicting the political alignment of Twitter users. In *Proceedings of 3rd IEEE Conference on Social Computing SocialCom*, 2011.

[4] D. Dayan and E. Katz. *Media events: the live broadcasting of history*. Harvard University Press, Cambridge, MA, 1992.

[5] Der Bundeswahlleiter. Bundestagswahl 2009: Ergebisse der Wahl zum 17. Deutschen Bundestag, November 2009.

[6] J. DiGrazia, K. McKelvey, J. Bollen, and F. Rojas. More tweets, more votes: social media as a quantitative indicator of political behavior. *Social Science Research Network*, February 21 2013.

[7] F. Franch. (Wisdom of the crowds)2: 2010 UK election prediction with social media. *Journal of Information Technology & Politics*, 10(1):57–71, 2013.

[8] D. Gayo-Avello. Don't turn social media into another 'literary digest' poll. *Communications of the ACM*, 54(10):121–128, October 2011.

[9] D. Gayo-Avello. I wanted to predict elections with Twitter and all I got was this lousy paper: a balanced survey on election prediction using Twitter data. *arXiv*, 2012.

[10] D. Gayo-Avello. A meta-analysis of state-of-the-art electoral prediction from Twitter data. *arXiv*, June, 25 2012.

[11] D. Gayo-Avello. No, you cannot predict elections with Twitter. *IEEE Internet Computing*, 16(6):91–94, Nov.-Dec. 2012.

[12] D. Gayo-Avello, P. T. Metaxas, and E. Mustafaraj. Limits of electoral predictions using Twitter. In *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media*, Menlo Park, CA, 2011. The AAAI Press.

[13] A. Gelman. The tweet-votes curve. *The monkey cage*, April 24 2013.

[14] A. Jungherr and P. Jürgens. The political click: political participation through e-petitions in Germany. *Policy & Internet*, 2(4):131–165, 2010.

[15] A. Jungherr and P. Jürgens. Forecasting the pulse: how deviations from regular patterns in online data can identify offline phenomena. *Internet Research*, 23(5), 2013.

[16] A. Jungherr and P. Jürgens. Stuttgart's black thursday on Twitter: mapping political protests with social media data. In R. Gibson, M. Cantijoch, and S. Ward, editors, *Analyzing social media data and Web networks: new methods for political science*. Palgrave Macmillan, New York, NY, 2013.

[17] A. Jungherr and P. Jürgens. Through a glass, darkly: tactical support and symbolic association in Twitter messages commenting on Stuttgart 21. *Social Science Computer Review*, Online First, 2013.

[18] A. Jungherr, P. Jürgens, and H. Schoen. Why the Pirate Party won the german election of 2009 or the trouble with predictions: a response to Tumasjan, A., Sprenger, T.O., Sander, P.G. & Welpe, I.M. "Predicting elections with Twitter: what 140 characters reveal about political sentiment". *Social Science Computer Review*, 30(2):229–234, 2012.

[19] M. Margolis and D. Resnick. *Politics as usual: the Cyberspace "revolution"*. SAGE Publications, Thousand Oaks, CA, 2000.

[20] P. T. Metaxas and E. Mustafaraj. Social media and elections. *Science*, 338(6106):472–473, October 26 2012.

[21] P. T. Metaxas, E. Mustafaraj, and D. Gayo-Avello. Hot (not) to predict elections. In *IEEE SocialCom Conference, Boston, MA*, pages 165–171, Washington, DC, October 2011. IEEE.

[22] R Core Team. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013.

[23] J. A. Ryan and J. M. Ulrich. *xts: eXtensible time series*. 2013.

[24] H. Schoen, D. Gayo-Avello, P. T. Metaxas, E. Mustafaraj, M. Strohmaier, and P. Gloor. The power of prediction with social media. *Internet Research*, 23(5), 2013.

[25] D. A. Shamma, L. Kennedy, and E. F. Churchill. Peaks and persistence: modeling the shape of microblog conversations. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*, New York, NY, 2011. ACM, ACM.

[26] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe. Election forecasts with Twitter: how 140 characters reflect the political landscape. *Social Science Computer Review*, 29(4):402–418, 2011.

[27] H. Wickham. *ggplot2: elegant graphics for data analysis*. Springer, New York, 2009.

[28] H. Wickham. *scales: Scale functions for graphics*. 2012.

[29] A. Zeileis and G. Grothendieck. zoo: S3 infrastructure for regular and irregular time series. *Journal of Statistical Software*, 14(6):1–27, 2005.