Election forecasts with Twitter
How 140 characters reflect the political landscape

Andranik Tumasjan[1], Timm O. Sprenger[2], Philipp G. Sandner[3], Isabell M. Welpe[4]

Technische Universität München
TUM School of Management
Chair of Strategy and Organization,
Leopoldstr. 139, 80804 Munich, Germany
Phone: ++49 89 28924800
Fax: ++49 89 28924805

Working paper version
21 January, 2011

This is the working paper version of the following article:

Tumasjan, A., Sprenger, T. Sandner, P., & Welpe, I. (2010). Election forecasts with Twitter: How 140 characters reflect the political landscape. *Social Science Computer Review*. Advance online publication. doi: 10.1177/0894439310386557

[1] E-mail: andranik.tumasjan@wi.tum.de
[2] E-mail: sprenger@tum.de
[3] E-Mail: philipp.sandner@tum.de
[4] E-mail: welpe@tum.de

Election Forecasts with Twitter
How 140 Characters Reflect the Political Landscape

Andranik Tumasjan
Timm O. Sprenger
Philipp G. Sandner
Isabell M. Welpe
Technische Universität München

This study investigates whether microblogging messages on Twitter validly mirror the political landscape offline and can be used to predict election results. In the context of the 2009 German federal election, we conducted a sentiment analysis of over 100,000 messages containing a reference to either a political party or a politician. Our results show that Twitter is used extensively for political deliberation and that the mere number of party mentions accurately reflects the election result. The tweets' sentiment (e.g., positive and negative emotions associated with a politician) corresponds closely to voters' political preferences. In addition, party sentiment profiles reflect the similarity of political positions between parties. We derive suggestions for further research and discuss the use of microblogging services to aggregate dispersed information.

Keywords: Twitter; microblogging; information market; prediction markets, election forecasts, politics; elections; sentiment analysis

Introduction

The successful use of social media in the US presidential campaign of Barack Obama has established Twitter, Facebook, MySpace, and other social media as integral parts of the political campaign toolbox. Some analysts attribute Obama's victory to a large extent to his online strategy. Obama's social-networking website helped him set records in terms of donations and grassroot mobilization (Williams & Gulati, 2008). Shortly after his victory, Obama used Twitter to let the web community know how he felt: "This is history". As this example demonstrates, after the rise of candidate websites in 1996, e-mail in 1998 (the Jesse

Ventura campaign), online fund-raising in 2000 (the John McCain campaign), and blogs in 2004 (the Howard Dean campaign; Gueorguieva, 2007), the microblogging platform Twitter has become a legitimate and frequently used communication channel in the political arena as a result of the 2008 campaign.[1] While some political analysts are already turning to the "Twittersphere" as an indicator of political opinion, others have suggested that the majority of the messages are "pointless babble" (pearanalytics, 2009). Therefore, the purpose of our study is to answer the question whether microblogging messages can actually inform us about public opinion and the political landscape in the offline world.

In particular, our study explores four aspects of this research question in the context of the 2009 federal election of the national parliament in Germany. First, we examine whether Twitter is used as a vehicle for political deliberation[2] by looking at how people use microblogging to exchange information about political issues. Second, we evaluate whether Twitter messages reflect the political preferences and the political landscape offline in a meaningful way. Third, we investigate whether individual accounts show evidence of party bias. Finally, we analyze whether the content of Twitter messages can be used to forecast the election result.

Background

The German election

In our study, we examine more than 100,000 tweets published in the weeks leading up to the federal election in Germany, which took place on September 27th, 2009. After 4 years in a grand coalition with the social democrats (SPD), Chancellor Angela Merkel - member of the conservatives (CDU) - was running for reelection, but favoring a coalition with the liberals (FDP). Many commentators have called the parties' campaigns uninspiring due to the

unwillingness of the main candidates to attack their then-coalition partners. The left side of the political spectrum was fragmented by the rise of the socialist party (Die Linke). The SPD publicly rejected Die Linke as a possible coalition partner, thus limiting its options to build a governing coalition. The potential coalition of CDU and FDP was leading by a slight majority in most polls and was ultimately able to form a center-right government after the election.

Related work and research questions

Recently, the exponential growth of Twitter has started to draw the attention of researchers from various disciplines. There are several streams of research investigating the role of Twitter in social media, product marketing, and project management. One stream of research concentrates on understanding microblogging usage and community structures (e.g. Honeycutt and Herring (2009)). In sum, this research demonstrates that the intensity of Twitter usage varies considerably. Market researchers have reported that in June 2009 (only a couple of weeks before the German federal election) 71% of all 1.8 million German users had visited Twitter only once and 15% of them at least 3 times (Nielsen Media Research, 2009). Honeycutt and Herring (2009) showed that Twitter is used not only for one-way communication but often serves as a means of conversation. In their study exploring conversation via Twitter, they find that 31% of a random sample of tweets contain an "@"-sign and that the vast majority (91%) of those were used to direct a tweet to a specific addressee. While these findings have provided us with a general understanding of why and how people use microblogging services, they have not explored the use of this new communication device in specific contexts such as, for instance, corporate public relations or the political debate online. This strand of literature only considers the online world and refrains from assessing its roots in the offline world. Another stream of research focuses on

corporate applications of microblogging such as the company-internal use for project management or the analysis of Twitter as electronic word of mouth in the area of product marketing (e.g., Jansen, Zhang, Sobel, & Chowdury, 2009). In their study, Jansen et al. (2009) have found that 19% of a random sample of tweets contained mentions of a brand or product and that an automated classification was able to extract statistically significant differences of customer sentiment (i.e., the attitude of a writer towards a brand). While this study provides reason to believe that sentiment may also be embedded in tweets covering other topics besides branding, Twitter sentiment analysis has not yet been applied to research regarding the political debate online. While several scholars have debated the potential of weblogs as a forum for democratic debate, "empirical research on deliberative democracy has lagged significantly behind theory" (Delli Carpini, Cook, & Jacobs, 2007, p. 316). A few researchers have empirically examined internet discussion boards as a vehicle for political deliberation (e.g., Jansen and Koop, 2005). Koop and Jansen (2009) have defined the exchange of substantive issues as an indicator of deliberation and the equality of participation as a measure of the deliberative quality of blog-based discussion. While they have found discussion boards and blogs to be dominated by a relatively small number of users, it is unclear whether these findings also apply to the political debate on Twitter. Recent scholarly work on political blogs has focused on their effect on real world politics, such as complementing the watchdog function of the mainstream media and mobilizing supporters, but largely ignored the reflection of offline politics in the digitally enhanced public sphere. However, there are a few studies exploring the reflection of the political landscape in "traditional" weblogs and social media sites. For instance, Williams and Gulati (2008) have found that the number of Facebook supporters can be considered a valid indicator of electoral success. Sunstein (2008) is more pessimistic and questions the ability of blogs to aggregate dispersed bits of information. Next to social media, in which the reflection of the political

landscape may be only a byproduct, prediction markets deserve mention as a special electronic platform that is designed to aggregate information on political elections. Prediction markets are similar to financial markets and allow trading in virtual securities tied to the outcome of a particular event (e.g., a candidate winning the election). Market prices can be interpreted as predictions (e.g., the share of the vote). The most cited and best known example of a prediction market is the Iowa Electronic Market (IEM), which was established in 1988 by the University of Iowa. It runs markets in federal and state elections. When compared directly to the corresponding large-scale polls, IEM prices were more accurate 76 percent of the time (Berg & Rietz, 2006). However, individual users of both social media and prediction markets show a significant bias in their party orientation. Adamic and Glance (2005) provide evidence of the fragmentation or polarization of the political blogosphere. They found that linkage patterns among bloggers reflect the blogosphere along party lines with 91% of all links directed to like-minded websites. Studies of the IEM found traders to be biased by their party preference, which was reflected in both their trading activity and their portfolio holdings (Forsythe, Rietz, & Ross, 1999). Despite previous research providing evidence that "traditional" social media content can be used to validly predict political outcomes despite individual party biases, we know little about the predictive power of Twitter for political debates and outcomes. Previous scholarly examinations of social media may not be easily transferrable to Twitter for the following reasons: First, tweets are much shorter and contain much less content than, for instance, news articles and traditional blogs. Hence, their informational value is less clear-cut. One marketing consultancy has even suggested that up to 40% of all Twitter messages are "pointless babble" (pearanalytics, 2009). Second, only part of the information conveyed is found in the words themselves because 19% of all messages contain links to other websites (Zarrella, 2009). Thus, a basic question is whether 140-character-messages can contain differentiated information regarding the

electorate's political preferences. Preliminary results from two recent reports suggest that microblogging content may be a good predictor of election results. A conference paper analyzing the correlation between candidate mentions on Twitter and the results of the Japanese national election reports that in more than 80% of all constituencies the most mentioned candidate won the election (Suenami & Yutaka, 2010). A similar survey of candidate mentions on Twitter during the 2010 UK election, presented by a website that aggregates political tweets, finds the predictions of the national share of vote to be better than most opinion polls with an average error of only 1.75 percentage points (Tweetminister, 2010). While these studies indicate that political microblogs may hold intriguing information to describe the political landscape, both are largely limited to the evaluation of the frequency of candidate mentions. We extend these findings by examining not just the number of mentions, but also the information content of the actual messages through linguistic sentiment analysis. We examine the nature of the political debate on Twitter and explore party biases of individual users. Next to the election results, we also investigate the relationship of sentiment profiles to election programs and likely coalitions.

Although the reference to tweets in many political commentaries shows that even analysts are already using Twitter as an indicator of political opinion, to the best of our knowledge, there are no scientific studies systematically investigating the sentiment in political microblogs. Therefore, the present study aims at addressing this general question in the following four ways: First, we examine whether Twitter provides a platform for political deliberation online. Second, we evaluate how accurately Twitter can inform us about the electorate's political preferences and the political landscape offline. Third, we investigate whether individual accounts show evidence of party preference. Fourth, we explore whether Twitter can serve as a predictor of the election result with respect to both share of vote as well as likely coalitions.

Microblogging forums as information markets

According to the American Association of Public Opinion Researchers (AAPOR), about $2 billion were spent on online research in 2009, 85% of which replaces traditional survey methods. However, the AAPOR concludes that "researchers should avoid nonprobability online panels when one of the research objectives is to accurately estimate population values [and] claims of representativeness should be avoided when using these sample sources" (p. 5). The AAPOR focuses on online surveys and largely ignores user-generated-content. We agree that many online samples, including the data used in this study, are not representative[3] and that representative results can only come from a survey of a representative sample. However, user-generated-content is not the result of a survey process, but the collection of a broad information exchange, in which users not only post their own opinions, but reflect on and discuss the comments of others and external sources (e.g., the press). We argue that microblogging forums allow users to weigh information and can thus produce accurate predictions, even if the users are not representative of the general population. Similar to information produced by other “unrepresentative” sources of “accurate” opinion, such as think tanks or financial markets, results can be accurate if information is aggregated efficiently. Our study investigates whether we can turn to microblogging content as an acceptable alternative to traditional surveys.

So what warrants an investigation of microblogging forums as a viable source of public opinion next to the growth of online opinion mining and advantages relative to traditional survey methods? The focus of our study is on the comparison of the information extracted from social media with accepted benchmarks such as election results. However, next to this empirical focus, there are also theoretical arguments that support the hypothesis that social media content can produce accurate predictions. We can approach this phenomenon with theories on crowd wisdom and collective intelligence. In a summary of related studies,

Surowiecki (2004) has suggested a number of conditions for the crowd to be wise, i.e. a large group of people to come to an accurate judgment: diversity (e.g., Hong & Page, 2001), independence (e.g., Bikhchandani, Hirshleifer, & Welch, 1998), and decentralization combined with a mechanism to aggregate dispersed bits of information (e.g., Hayek, 1945). Like financial markets, many social media networks, including Twitter, largely comply with these properties. The information structure among social media users is not very different from financial markets, which are widely accepted as information aggregation mechanisms. Whereas financial markets aggregate information through the price system, microblogging forums lack a *natural* aggregation mechanism. Sunstein (2008, p. 88) suggests that "the blogosphere cannot operate in Hayekian fashion, because it lacks […] the price system (or any reasonable analogue to it)". He argues that "participants in the blogosphere [are not encouraged to produce reliable and unbiased information because they] lack an economic incentive" (Sunstein, 2008, p. 90). However, a study of prediction markets has shown that both play- and real-money markets predicted outcomes equally well (Servan-Schreiber, Wolfers, Pennock, & Galebach, 2004) which indicates that non-monetary incentives such as the position in a publicized ranking may encourage earnest participation. Most microblogging forums have their own implicit rankings which may have similar effects: So even if Twitter does not have an explicit mechanism for aggregating information, the size of the followership and the rate of retweets may represent the Twittersphere's "currency" and provide it with its own kind of a "price system". Studies have shown that, despite the abundance of available information and considerable noise, Twitter users follow the accounts to which they subscribe closely and are highly attentive to their content. A study of a Twitter account making directional forecasts of the stock market has shown the number of followers to be correlated with the accuracy of the published information (i.e., the forecasts of the stock market; Giller, 2009). The fact that users not only notice these subtleties in the dense

information stream they are exposed to, but also act on them by maintaining or terminating their subscription is only one powerful example of how the quality and accuracy of content determines the number of followers. In addition, new and valuable pieces of information are retweeted more often, providing Twitter with a mechanism to weigh the importance of information. Even if this mechanism does not provide us with a signal as clear as a market price, the display of word clouds or the ranking of trending topics illustrate successful attempts to improve the information aggregation. These mechanisms can certainly inform us about what people find important because, even with hundreds of thousands of tweets being sent every day, time and again, newsworthy messages from private individuals with no more than a few dozen followers (such as the one making the first report of the plane crash in the Hudson River in 2009) bubble to the surface and get spread to a wider audience.

These considerations encourage us to believe that the information stream on Twitter can be aggregated in a meaningful fashion in order to make accurate, albeit not necessarily representative, predictions and that we can leverage Twitter as an information market.

Data set and methodology

We examined 104,003 political tweets, which were published on Twitter's public message board between August 13th and September 19th, 2009, prior to the German national election. We systematically collected all tweets that contained the names of either the 6 parties represented in the German parliament (CDU/CSU, SPD, FDP, B90/Die Grünen, and Die Linke) or those politicians of these parties who are regularly included in a weekly survey on the popularity of politicians conducted by the research institute "Forschungsgruppe Wahlen". CDU and CSU, often referred to as the "Union", are sister parties which form one faction in the German parliament. Our query resulted in roughly 70,000 tweets mentioning one of the 6 major parties and 35.000 tweets referring to their politicians.

Given the samples size, we decided to use sentiment analysis, an automated mechanism to quantify the information contained in these messages. In the domain of natural language processing, the term sentiment analysis is "used in reference to the automatic analysis of evaluative text and tracking of the predictive judgments therein" (Pang & Lee, 2008, p. 10). This analysis includes the extraction of the polarity (either positivity or negativity), but, more generally, refers to the computational extraction of information from a given text sample. To extract the sentiment of the tweets objectively and systematically, we used LIWC2007 (Linguistic Inquiry and Word Count; Pennebaker, Chung, & Ireland, 2007), a text analysis software developed to assess emotional, cognitive, and structural components of text samples using a psychometrically validated internal dictionary. This software calculates the degree to which a text sample contains words belonging to empirically defined psychological and structural categories. Specifically, it determines the rate at which certain cognitions and emotions (e.g., future orientation, positive or negative emotions) are present in the text. For each psychological dimension, the software calculates the relative frequency with which words related to that dimension occur in a given text sample (e.g., the words "maybe", "perhaps", or "guess" are counted as representatives of the construct "tentativeness"). LIWC has been used widely in psychology and linguistics, but also for topics related to political science. Examples include studies of the sentiment levels in US Senatorial speeches (Yu, Kaufmann, & Diermeier, 2008), the linguistic differences between positive and negative political ads and television interviews of presidential candidates. LIWC-based analyses have also been used to examine shorter text samples such as instant message conversations, which are similar in length to postings on, and the Twitter accounts of gubernatorial candidates in various US state elections (for a comprehensive overview of related studies, see Tausczik & Pennebaker, 2009). We use the following 12 LIWC dimensions in order to profile sentiment in political tweets: Future orientation, past orientation, positive emotions, negative

emotions, sadness, anxiety, anger, tentativeness, certainty, work, achievement, and money. These categories have either been successfully used in prior studies of political text samples or seemed best suited to profile messages in the political domain by covering both emotions (e.g., anger, anxiety) as well as content dimensions (e.g., work, money). Following the methodology used by Yu et al. (2008) we concatenated all tweets published over the relevant timeframe into one text sample to be evaluated by LIWC. Our sample was restricted to German language tweets, which were translated into English and then processed by the LIWC English dictionary.

It is important to distinguish online sentiment, as described above, and offline sentiment. The political sentiment offline may become apparent in the form of party preference (i.e. election results) or positions on individual policy issues or people (i.e. left vs. right-wing positions or positive/negative emotions with respect to a candidate). To avoid confusion with online sentiment, we refer to political sentiment offline as political preferences.

Results

Twitter as a platform for political deliberation

In this section, we will evaluate our sample along two widely accepted indicators of blog-based deliberation, the exchange of substantive issues and the equality of participation (Koop & Jansen, 2009).

Table 1 shows the number of mentions and a randomly selected tweet for all parties in our sample. These messages are only supposed to provide a glance at the underlying data. While this is only a small selection of the information stream in our sample, these messages illustrate that tweets can contain a lot of relevant information. So, despite their brevity, substantive issues can be expressed in 140 characters or less.

[Table 1 about here]

Next, we analyze the level of addressivity in the messages as an indication of the exchange of ideas on Twitter. About one third of all tweets in our sample (30.8%) contain an "@"-sign which is in line with previous research that has also suggested that the vast majority of "@"-signs are used to direct a tweet to a specific addressee (Honeycutt & Herring, 2009). A more conservative measure of direct communication are direct messages to another user starting with an "@"-sign. Roughly 10% of the messages in our sample are direct messages indicating that people are not just using Twitter to post their opinions, but also engage in interactive discussions. Many users on Twitter forward messages to their followership. These so-called retweets often contain information that the sender finds noteworthy such as links to other websites. Consequently, the rate at which messages are retweeted indicates whether information is considered being interesting. According to Zarrella (2009), only 1.44% of all tweets are retweets. In our sample, however, that share of both retweets and messages containing a hyperlink are significantly higher: 19.1% of all messages were retweets and 54.2% contain a link to a website. Summarizing, our results indicate that people are finding interesting political information on Twitter, which they share with their network of followers.

We now turn to the analysis of the equality of participation. While we find evidence of a lively political debate on Twitter, it is unclear whether this deliberation is led by a few "political junkies" rather than the wider general public. Jansen and Koop (2005) found less than 3% of all users on the political message board *BC Votes* to be responsible for almost a third of all posted messages. Table 2 shows the share of users and the share of messages across various user groups for our sample according to the frequency with which a user posts messages. We adopted the categorization from Jansen and Koop (2005).

[Table 2 about here]

While the distribution of users across user groups is almost identical with the one found by Jansen and Koop (2005), we find even less equality of participation for the political debate on Twitter. There is a high concentration of messages in the groups of heavy (23.1%) and very heavy users (21.2%). These make up 3.3% and 0.6% of the users, respectively. So roughly 4% of all users accounted for more than 40% of the messages. In sum, Twitter is used as a forum for political deliberation, and this forum is dominated by a small number of heavy users.

Twitter sentiment as a reflection of the political landscape offline

The fact that users are discussing political issues online does not mean that we can necessarily extract meaningful information from this debate. To explore this question, we aggregated the information stream about politicians and parties and compared the resulting profiles with evidence from the press and election programs. We investigate the reflection of two aspects of the political landscape in the messages: First, whether the messages reflect the voters' political preferences (e.g., positive or negative emotions with respect to the main candidates based on press reports and anecdotal evidence from the campaign) and, second, whether the tweets mirror the ideological proximity of political parties (e.g., similarity of the parties' political agendas based on their election programs).

In order to analyze the sentiment of the political tweets, we generated multi-dimensional profiles of the politicians in our sample using the relative frequencies of LIWC category word counts (i.e., the percentage of words in all tweets about a particular candidate, which are related to the 12 chosen dimensions according to the LIWC dictionary). It is important to note

that these sentiment profiles are not policy or ideological measures. They merely represent an aggregated linguistic profile of the messages associated with a particular party or politicians. Figure 1 shows these profiles for the leading candidates of the 5 main parties. Overall, positive emotions clearly outweigh negative emotions. This is in line with Yu et al. (2008) who find that positive emotions outweigh negative emotions by more than 2 to 1 in an LIWC-based analysis of 18 years of congressional debates. Only liberal party leader Guido Westerwelle and socialist party leader Oskar Lafontaine show more distinctive deviations from this profile on some dimensions. The dimension of perceived anger, for example, is most prominent in the case of these two politicians who, as free-market advocate and socialist leader, represent two contrasting political programs in the political spectrum. Messages regarding Frank-Walter Steinmeier, who at the time of our recording was sending mixed signals regarding potential coalition partners for his party after the election, reflect more tentativeness than those of other politicians. The higher share of tentative messages also corresponds to findings indicating that tentativeness correlates to lower status and rank (Tausczik & Pennebaker, 2009). Compared to acting chancellor Merkel, this profile is in line with Steinmeier's role as vice chancellor and clear runner-up in the polls. As can be seen in Figure 1, while small in absolute terms, the sentiment embedded in tweets does reflect nuanced differences between the politicians in our sample. To conclude, one can say that Twitter messages can be considered a plausible reflection of voters' sentiment. [4]

[Figure 1 about here]

Next we explore whether the tweets mirror the ideological proximity of political parties. Since it is not easy to visually analyze the profiles using the radar charts, we computed a measure of similarity for the LIWC profiles of all combinations of two parties (Table 3).

Hinich and Munger (1997) suggest the use of Euclidian distances along multiple policy dimensions to measure the political distance between parties.

The distance of party *k* and party *l* can be calculated as

$$(1) \qquad d_{kl} = \frac{1}{S}\sqrt{\sum (x_{kj} - x_{lj})^2} \quad ,$$

where x_{kj} represents the position of party *k* on the policy dimension *j* and *S* is the number of parties that were included in the calculation. In our case, we use LIWC sentiment dimensions instead of policy dimensions. The measure d_{LIWC} represents the root of the sum of all squared differences between the 12 LIWC dimensions across a sample of parties. The higher the value of d_{LIWC}, the higher the heterogeneity of the LIWC profiles of the included parties. In other words, the lower the value of d_{LIWC}, the more similar the LIWC profiles. In order to evaluate whether the LIWC profiles reflect the political landscape offline, we compare them to an objective measure of political similarity. As indicated, the distance measure has been used widely in political science, even in the context of German politics (see, for example, Bräuninger & Debus, 2009, who refer to it as ideological heterogeneity). To provide an objective benchmark for the similarity of linguistic profiles we have constructed the same distance measure based on data from the Comparative Manifesto Projects (CMP). The CMP conducts quantitative content analyses of election programs of parties from more than 50 countries covering all free democratic elections since 1945 (Klingemann, Volkens, Bara, Budge, & McDonald, 2006). The purpose is to measure political preferences of parties. The CMP database provides frequency tables that indicate how many sentences in its election program a party dedicates to each of 56 categories. Examples of these categories include human rights, free enterprise, economic goals, welfare state expansion/limitation, education expansion/limitation and traditional morality. CMP provides data for the election programs of the Union faction (CDU/CSU), SPD, FDP, Grüne and Die Linke for the German federal

election 2009. Thus, we were able to calculate the ideological heterogeneity of the parties in our sample according to the CMP (d_{CMP}) using the measure of heterogeneity outlined above. For simplicity, we have limited this analysis to all combinations of two parties for which the CMP provides data.

As can be seen in Table 3, the heterogeneity of LIWC profiles of the major parties corresponds closely to their political proximity. Most notably, the distance measure confirms the tight fit between the Union faction of sister parties CDU and CSU (d_{LIWC} =0.34). Both measures indicate a fair degree of correspondence between the CDU and the other main parties, except for the socialist party Die Linke (d_{LIWC}=1.48, d_{CMP}=1.38). The LIWC profile of the Green party is most similar to their former and desired coalition partner SPD. Both measures show Die Linke to be farthest from the center-right parties CDU/CSU and FDP and closer to the left-of-center parties SPD and Green party. Overall, the similarity of LIWC profiles is a plausible reflection of the political proximity of the parties' election programs in the weeks before the federal election.

[Table 3 about here]

Party bias of individual users

In this section, we will examine whether individual users show evidence of a party bias in the volume or the sentiment of their postings. Studies of prediction markets have linked biases in trading activity and portfolio holdings to traders' party preference (Forsythe, Rietz, & Ross, 1999). Many politicians and regional party branches maintain Twitter accounts, which contain the party acronym in their account name (e.g., "SPDBerlin"). By focusing on the subset of accounts, which contain the name of the two largest parties (CDU and SPD), we

were able to identify a group of 82 SPD- and 57 CDU-labeled accounts that are likely associated with one of the two parties. We filtered the messages posted through these accounts by the parties or candidates mentioned (Table 4). As can be seen, the party accounts dedicate roughly 80% of their mentions to their own party. Also, they make about three times as many references to their party's candidate compared to the opponent. With respect to the sentiment contained in these messages, we find that negative emotions are clearly correlated with party affiliation. Positive emotions, on the other hand, do not follow the same pattern. These results are consistent with a large body of literature that shows negative information to be processed more thoroughly and to have more impact than positive information ("negativity bias"; e.g., Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). In summary, we can conclude that individual users show a party bias with respect to the volume5 as well as the negative sentiment of the messages they post. The analysis of party-affiliated accounts suggests that this bias is linked to party preference.

[Table 4 about here]

Twitter as a predictor of the election result

In order to understand whether the activity on Twitter can serve as a predictor of the election outcome we examine two aspects. First, we compare the share of attention the political parties receive on Twitter with the result of the 2009 German federal election. Second, we analyze whether tweets can inform us about the ideological ties between parties and potential political coalitions after the election. It is important to distinguish this analysis based on party mentions from the sentiment profiles in the previous section. The number of mentions is not a

sentiment measure, but merely measures the overall attention a party garners, i.e., the "buzz" it generates.

Table 5 shows the number of tweets mentioning a particular party. As can be seen, the ranking by tweet volume (i.e., the number of tweets) and the ranking by share of vote in the election results are identical. In fact, the relative volume of tweets mirrors the results of the federal election closely. If we consider the number of tweets to be a predictor of the election result, the mean absolute error (MAE) of this prediction is 1.65%. The MAE is a measure of forecast accuracy and has been widely applied to compare the accuracy of political information markets relative to election polls (see Berg et al., 2008).

[Table 5 about here]

To understand how the above-mentioned prediction based on message volume compares with traditional methods to collect this data, we compare Twitter with a number of election polls and the IEM prediction market. The MAE of 6 research institutes, which published election polls in our sample period, ranges from 1.1% to 1.7%. Thus, Twitter comes close to these accepted benchmarks. The predictive accuracy is even more impressive when compared to the track record of the IEM, a prediction market set up with the explicit purpose to predict election results. The IEM produced a MAE of 1.37% in US presidential elections and 2.12% in non-US elections based on election eve market prices (Berg, Forsythe, Nelson, & Rietz, 2008). In conclusion, the mere number of tweets mentioning a political party can be considered a plausible reflection of the vote share and its predictive power even comes close to traditional election polls.

While the predictive qualities of Twitter mentions confirm reports of similar findings (Suenami & Yutaka, 2010; Tweetminister, 2010), other studies have not explored the

information content of microblogs with respect to likely coalitions after the election. Therefore, after extracting the absolute strength of each party from the message volume, we now turn to the relationships between the parties. This is all the more relevant, as all parties were far from an absolute majority in the weeks preceding the federal election and a coalition government was on the horizon. In order to assess the predictive accuracy of the information content of political microblogs with respect to likely coalitions, we need a benchmark. We can look at two sources to provide this benchmark. Based on pre-election polls, government formation on state and federal level since 1990 and political heterogeneity of political programs, Bräuninger and Debus (2009) have developed econometric models to predict the likelihood of various coalitions after the federal election 2009. They come to the conclusion that, if CDU/CSU and FDP can gain a majority in the election, a coalition of the two is almost certain. A remake of the grand coalition is the most likely scenario, if the center-right fails to gain a majority. According to Bräuninger and Debus (2009) a left-of-center coalition is very unlikely whereas a so-called Jamaica-coalition consisting of CDU/CSU, FDP and the Green party is only slightly less likely than a center-right coalition. In addition to this academic study, we also provide the results of a survey conducted by research institute TNS Emnid, which provided the most accurate pre-election poll among the 6 research institutes cited above. In a survey of 1,000 people, 43% of the respondents selected the center-right coalition of CDU/CSU and FDP, 20% the grand coalition (CDU and SPD), 6% a coalition of CDU and Green party, and 5% a "Jamaica"-coalition (CDU/CSU, FDP, Green party) as the most likely coalition to be formed after the election. Although both studies confirm conventional wisdom among political analysts in Germany, they provide an objective benchmark for our results, especially for readers not familiar with the German political system.

While some tweets mention only one particular party (sole mentions), many messages refer to several parties (joint mentions). We investigate whether joint mentions reflect prevailing or even upcoming political ties. To make the comparison easier and the interpretation more straightforward, we focus on tweets mentioning only two parties. Based on the overall probability that any one party is mentioned in these tweets, a conditional probability that two parties are mentioned together can be computed. If all combinations were equally likely, this conditional probability should equal the observed share of tweets mentioning these two parties. Due to different base rates, we divide the observed share of joint mentions by the conditional probability to derive a comparative measure. If *share*(*CDU*, *CSU*) represents the share of observed joint mentions of these two parties, the relative frequency *f*, is calculated as follows:

$$(2) \qquad f = \frac{share(CDU, CSU)}{\left[P(CDU|CSU) + P(CSU|CDU)\right]/2}$$

The relative frequency illustrates how often two parties are mentioned together relative to the random probability based on the overall "share of voice" of the individual parties. If *f* equals 1.5 the share of observed joint mentions is 50% higher than pure chance would suggest.

Table 6 shows the relative frequency for all combinations of two parties based on all tweets mentioning more than on party (n = 61.700). Not surprisingly, the combined mentioning of sister parties CDU and CSU was the most frequent (f = 1.25), whereas CSU and the left-of-center parties (SPD, Grüne, and Die Linke) were mentioned together the least. While the governing coalition of CDU and SPD are naturally mentioned jointly quite frequently, the Union parties (CDU and CSU) are associated most closely with its desired coalition partner at that time, the FDP. The parties of the left side of the political spectrum are associated with each other more often than with the right-of-center parties (CDU, CSU, and FDP). In sum,

the joint mentions of political parties accurately reflect the political ties between the parties. We conclude that, despite the fact that the Twittersphere is no representative sample of the German electorate, the activity prior to the 2009 German election seems to validly reflect the election outcome.

[Table 6 about here]

Conclusion

Discussion of results

We analyzed over 100,000 Twitter messages mentioning parties or politicians prior to the German federal election 2009. Overall, we found that Twitter is indeed used as a platform for political deliberation. The mere number of tweets reflects voters' preferences and comes close to traditional election polls, while the sentiment of political Twitter messages closely corresponds to the electorate's sentiment and evidence from the media coverage of the campaign trail. With respect to our first research question, we found more than one third of all messages to be part of a conversation indicating that Twitter is not just used to spread political opinions, but also to discuss these opinions with other users. While we find evidence of a lively political debate on Twitter, this discussion is dominated by a small number of users: only 4% of all users accounted for more than 40% of the messages. With respect to our second research question, we found the multi-dimensional sentiment profiles of politicians and parties to plausibly reflect many nuances of the election campaign. Overall, the similarity of profiles between parties matches the similarity of their political agendas. With respect to our third research question, we have found a party bias in individual user accounts with

respect to the volume as well as the sentiment of their political communication on Twitter. Our results suggest that this bias is linked to party affiliation. With respect to our fourth research question, we found that the mere number of messages reflects the election result and that this rather simple metric, with a mean absolute error of 1.65%, even comes close to traditional election polls. This finding is in contrast to previous studies of political deliberation online. In a study on internet message boards by Jansen and Koop (2005), even the positions of the two largest parties were reversed and the party winning an absolute majority was only associated with 27.2% of the party mentions. The authors attributed this phenomenon to the dominance of a few users who "determined the overall ideological 'feel' of the discussion board" (Jansen and Koop, 2005, p. 624). Given that there was even less equality of participation in our sample and a bias in party orientation among individual users, it is all the more surprising that heavy users were unable to impose their views on the discussion and affect the accuracy of aggregate results. However, we strongly believe that our results did not come about by chance of this particular dataset and elaborate on the reasons in our discussion of microblogging forums as information market. Our results provide evidence supporting our theory that microblogging forums provide a mechanism for weighing information and that, despite individual biases, errors can cancel each other out. The predictive accuracy is even more impressive when compared to the track record of the IEM, a prediction market set up with the explicit purpose to predict election results. Our results clearly suggest that Twitter may complement traditional methods of political forecasting (e.g., polls or surveys). There are multiple advantages of extracting public opinion from microblogging content (see O'Connor, Balasubramanyan, Routledge, & Smith, 2010, for more detail). These include cost (because most of the content is freely available and easily accessible), speed (traditional polls often take days to plan and conduct), more recent information (with permanent online access including mobile connections users often share

new opinions instantly when they occur to them), frequency (the density of the information stream on Twitter allows us to draw samples at almost arbitrary intervals at almost no additional costs), unedited expression (natural responses not constrained to predefined topics and by standardized response formats), and a greater variety of topics (microblogging forums covers almost every imaginable topic). Overall, our results demonstrate that Twitter can be considered a valid indicator of the political landscape offline.

Limitations and further research

This study, like others, does not come without caveats. First, the demographics of Twitter users may raise concerns that our sample may not have been representative of the German electorate.[3] While we have explored theoretical reasons for this sample to produce predictions that are accurate, but not representative of the general population, the majority of our paper has been dedicated to the empirical comparison of the microblogging content and the political landscape offline. One of the most critical aspects of further research will be to better understand and investigate empirically the mechanism by which information is weighted and distributed in microblogging forums. Our finding that counting party mentions is different from a survey process provides a starting point for this research. Second, our investigation was based on one particular text analysis software and used an existing dictionary not specifically tailored to classify short political tweets. There are many specifics of communication through microblogging services, including the use of a special syntax and conventions (e.g., the use of emoticons) which are not reflected in our default LIWC dictionary. In addition, categories such as patriotism, xenophobia, parochialism, empathy or humanitarian and philanthropic instincts seem particularly relevant in the realm of political analyses. Since we translated the German language messages into English, some nuances in meaning may have been lost in the translation. However, we believe this effect to be

negligible since LIWC is based on word count only and therefore should not be affected by grammatical errors. Third, due to the requirement of a minimum document length for LIWC to produce meaningful, comparable results, we treated all messages published in a given time frame as one document and were unable, for instance, to distinguish between positive and negative comments. Further research should refine the text analysis to the political discussion and investigate the sentiment of the messages tweet by tweet because Asur and Huberman have shown that "sentiments extracted from Twitter can be further utilized to improve the forecasting power of social media [mentions]" (2010, p. 1). In addition, similar to the stratified sampling approach and weighted designs of traditional polling methods, one could leverage user demographics embedded in the profiles of social media users, to calibrate results derived from online content (O'Connor et al., 2010). Fourth, while we have examined overall sentiment of political tweets, voters' attitudes and opinions may vary depending on specific political issues. Future sentiment analysis could address this issue by conducting a more detailed classification of content. This may allow us to produce results similar to detailed opinion surveys on various political issues. Finally, our study was limited to the federal election in Germany. Although other studies that correlated party or candidate mentions on Twitter with electoral success indicate that similar results can be found in other countries with different electoral systems (even on the level of individual constituencies), we cannot generalize that these results will hold everywhere (e.g., Suenami & Yutaka, 2010; Tweetminister, 2010). To summarize, our results demonstrate that Twitter can be seen as a valid real-time indicator of voters' political preferences. Little research has yet been conducted in this area leaving many questions unresolved. Further research should test whether more advanced text analysis procedures can produce even more meaningful results. Researchers should also try to capture the context of a particular statement in a more comprehensive manner including threads of conversation and links to information beyond the

tweets. Analyzing conversations may be a rich source of data for understanding how users interact, political ideas evolve and arguments are exchanged in online discussions. Including information from external sources, such as news articles linked to in the tweets, may be a starting point to distinguish between the content generated by microbloggers themselves and the content that is merely reflected in the microblogging forum. It would help us understand whether and to what extent the content in microblogging forums provides us with truly new or unique insights or whether Twitter is simply mirroring other sources.

Next to the immediate conclusions from our empirical results, our study also contributes to the understanding of the ability of the blogosphere to aggregate information. Even though we do not yet fully understand how these mechanisms works, our results indicate that information on Twitter actually can be aggregated in a meaningful way. The fact, that even the fairly simple methodology used in our study was able to generate plausible results is encouraging and points to additional possibilities to leverage Twitter as an information market.

Andranik Tumasjan (andranik.tumasjan@wi.tum.de) is a research assistant and PhD student at TUM School of Management at the Technische Universität München (TUM). He holds a master's degree in applied psychology and specializes on topics at the intersection of psychology and business.

Timm O. Sprenger (sprenger@tum.de) is a research assistant and PhD student at TUM School of Management at the Technische Universität München (TUM). His current research focuses on the use of microblogging platforms as information markets.

Philipp G. Sandner (philipp.sandner@tum.de) received his doctoral degree from the Ludwig-Maximilians-Universität Munich (LMU). His published work includes articles on innovation, intellectual property, and entrepreneurship.

Isabell M. Welpe (welpe@tum.de) holds the chair of Strategy and Organization at TUM School of Management at the Technische Universität München (TUM). Her research focus is on behavioral decision making.

References

AAPOR (2010). AAPOR Report on Online Panels. Retrieved June 25, 2010 from http://www.aapor.org/AM/Template.cfm?Section=AAPOR_Committee_and_Task_Force_Reports&Template=/CM/ContentDisplay.cfm&ContentID=2223.

Adamic, L. A., & Glance, N. (2005). The political blogosphere and the 2004 US election: Divided they blog. In *Proceedings of the 3rd International Workshop on Link Discovery*, 36-43. Chicago, IL.

Asur, S., & Huberman, B. (2010). *Predicting the Future With Social Media*. Retrieved June 25, 2010 from http://arxiv.org/pdf/1003.5699

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323-370.

Berg, J., Forsythe, R., Nelson, F., & Rietz, T. (2008). Results from a dozen years of election futures markets research. In C. Plott & V. Smith, *Handbook of Experimental Economics Results*, 742-751, Amsterdam: Elsevier.

Berg, J. and Rietz, T. (2006). The Iowa Election Markets: Stylized facts and open issues. In R. Hahn, Robert and P. Tetlock, *Information Markets - A New Way of Making Decisions*, 142-169, Washington, DC: The AEI Press.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *The Journal of Economic Perspectives*, 12(3), 151-170.

Bräuninger, T., & Debus, M. (2009). Die Regierungsbildung nach der Bundestagswahl 2009: Wie wahrscheinlich ist eine Neuauflage der großen Koalition? (in German). In *DVPW Workshop zur Bundestagswahl 2009*, 1-35, Kiel.

Delli Carpini, M. X., Cook, F. L., & Jacobs, L. R. (2004). Public deliberation, discursive participation, and citizen engagement: A review of the empirical literature. *Annual Review of Political Science*, *7*(1), 315-344.

Forsythe, R., Rietz, T., & Ross, T. (1999). Wishes, expectations and actions: a survey on price formation in election stock markets. *Journal of Economic Behavior & Organization*, 39(1), 83-110.

Giller, G. L. (2009). *Maximum likelihood estimation of a poissonian count rate function for the followers of a Twitter account making directional forecasts of the stock market*. Retrieved June 25, 2010 from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1423628

Gueorguieva, V. (2007). Voters, MySpace, and YouTube: The impact of alternative communication channels on the 2006 Election Cycle and Beyond. *Social Science Computer Review*, 26(3), 288-300.

Hayek, F. V. (1945). The use of knowledge in society. *American Economic Review*, 35(4), 519-530.

Hinich, M., & Munger, M. (1997). *Analytical politics*. Cambridge: Cambridge University Press.

Hong, L., & Page, S. (2001). Problem Solving by Heterogeneous Agents. *Journal of Economic Theory*, 97(2), 123-163.

Honeycutt, C., & Herring, S. C. (2009). Beyond microblogging: Conversation and collaboration via Twitter. In *42nd Hawaii International Conference on System Sciences*, 1-10, Hawaii.

Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology*, 60, 1-20.

Jansen, HJ, & Koop. R. (2005). Pundits, ideologues, and ranters: The British Columbia election online. *Canadian Journal of Communication*, 30(4), 613-632.

Klingemann, H.-D., Volkens, A., Bara, J. & McDonald M. (2006). *Mapping Policy Preferences II: Estimates for Parties, Electors, and Governments in Eastern Europe, European Union and OECD 1990-2003*. Oxford: Oxford University Press.

Koop, R., & Jansen, H. J. (2009). Political blogs and blogrolls in Canada: Forums for democratic deliberation? *Social Science Computer Review*, 27(2), 155-173.

Nielsen Media Research (2009). Das Phänomen Twitter in Deutschland (in German). Retrieved December 15, 2009 from http://de.nielsen.com/news/ NielsenPressemeldung04.08.2009-Twitter.shtml

O'Connor, B., Balasubramanyan, R., Routledge, B., & Smith, N. (2010). From Tweets to polls: Linking text sentiment to public opinion time series. In *Proceedings of International AAAI Conference on Weblogs and Social Media*, 122-129. Washington, DC.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, *2*(1-2), 1–135.

pearanalytics (2009). Twitter study. Retrieved December 15, 2009 from http://www.pearanalytics.com/ wp-content/uploads/2009/08/Twitter-Study-August-2009.pdf

Pennebaker, J., Chung, C., & Ireland, M. (2007). The development and psychometric properties of LIWC2007. Austin, TX.

Servan-Schreiber, E., Wolfers, J., Pennock, D. M., & Galebach, B. (2004). Prediction markets: Does money matter? *Electronic Markets*, *14*(3), 243-251

Suenami, A. & Yutaka, M. (2010). Voter behavior analysis and election campaign strategies using information on the web (in Japanese). In *Proceedings of the JSAI SIG-KBS, No. 87*. 1-6. Tokyo.

Sunstein, Cass. (2007). Neither Hayek nor Habermas. *Public Choice*, 134(1-2), 87-95.

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*. New York, NY: Random House, Inc.

Tausczik, Y. R., & Pennebaker, J. W. (2009). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*.

TNS Emnid (2009). Wahrscheinliche Regierungskoalition nach der Bundestagswahl (in German). Retrieved June 25, 2010 from http://de.statista.com/statistik/daten/studie/29701/umfrage/wahrscheinliche-regierungskoalition-nach-der-bundestagswahl/#info

Tweetminister (2010). Is word-of-mouth correlated to General Election results? The results are in. Retrieved June 25, 2010 from http://www.scribd.com/doc/31208748/Tweetminster-Predicts-Findings

Web Evangelisten (2009). Twitterumfrage 2 (in German). Retrieved June 25, 2010 from http://webevangelisten.de/twitterumfrage/

Williams, C., & Gulati, G. (2008). What is a social network worth? Facebook and vote share in the 2008 presidential primaries. In *Annual Meeting of the American Political Science Association*, 1-17. Boston, MA.

Yu, B., Kaufmann, S., & Diermeier, D. (2008). Exploring the characteristics of opinion expressions for political opinion classification. In *Proceedings of the 2008 international conference on Digital* government research, 82-91. Montreal.

Zarrella, D. (2009). State of the Twittersphere. Retrieved December 15, 2009 from http://blog.hubspot.com/Portals/249/sotwitter09.pdf

Notes

1. The political discourse on Twitter has lead to the establishment of numerous dedicated websites, both in the US (e.g. http://tweetcongress.org, http://www.congressional140.com, and http://govtwit.com) and Germany (e.g. http://parteigefluester.de, http://wahlgetwitter.de).

2. In line with Delli Carpini, Cook, and Jacobs (2007) we use the words deliberation, debate and discussion interchangeably.

3. As far as our data source is concerned, an online survey of 1,707 German Twitter users conducted in November 2009, only 2 months after the German election, has shown that these users are predominantly male (64%), young (31 years) and have a university degree (67%) (WebEvangelisten, 2009). On the other hand, over the last decades the German population eligible to vote has aged continuously and voter turnout has been higher for older age group. Even though we do not have demographic data for the specific users in our sample, these facts indicate that the demographics of political microbloggers and actual voters are probably not skewed in the same, but – to the contrary – in opposite directions relative to the overall population.

4. An analysis of tweets surrounding the TV debate between the two candidates for chancellor showed that political tweets can even track changes in sentiment on a daily basis.

5. Additional analyses confirm this volume effect for the whole dataset. The distribution of user attention (i.e., the share of mentions that a user dedicates to the various parties) shows a significant bias. Users put a clear emphasis on the discussion of one particular party.

TABLE 1
Sample tweets by party

Party	**Number of tweets**	**Examples**
CDU	30,886	CDU wants strict rules for internet
CSU	5,748	CSU continues attacks on partner of choice FPD
SPD	27,356	Only a matter of time until the SPD dissolves
FDP	17,737	Whoever wants civil rights must choose FDP!
Die Linke	12,689	Society for Humans Rights recommends: No government participation for Die Linke
Grüne	8,250	After the crisis only Green can help HTTP:[…] Grüne+

Note: Examples were randomly selected from the tweets mentioning each party. Messages were shortened for citation (e.g., omission of hyperlinks).

TABLE 2
Equality of participation and format of communication

User group	**Users**		**Messages**		**Format of communication**			
	Total	Share	Total	Share	Mention	DM	RT	URL
One-time (1)	7,064	50.3%	7,064	10.2%	29.2%	11.7%	16.7%	45.0%
Light (2-5)	4,625	32.9%	13,353	19.3%	29.4%	9.7%	18.6%	48.9%
Medium (6-20)	1,820	12.9%	18,191	26.2%	30.7%	10.2%	19.0%	49.4%
Heavy (21-79)	463	3.3%	15,990	23.1%	32.7%	10.8%	20.1%	55.7%
Very heavy (80+)	84	0.6%	14,710	21.2%	31.1%	10.2%	19.9%	67.6%
Total	**14,056**	**100.0%**	**69,318**	**100.0%**	**30.8%**	**10.4%**	**19.1%**	**54.2%**

Note: Mentions are message containing "@" (includes direct messages), direct messages (DM) start with "@", retweets contain "RT", "via", or "by", URL are messages containing hyperlink (e.g., "http:").

TABLE 3
Heterogeneity of linguistic and policy profiles

Linguistic heterogeneity (LIWC) d_{LIWC}				
	CDU/CSU	**SPD**	**FDP**	**GREEN**
CDU/CSU	0.34*			
SPD	0.86			
FDP	0.90	0.62		
Grüne	0.87	0.65	0.83	
Die Linke	1.48	1.35	1.46	1.00

Ideological heterogeneity (CMP) d_{CMP}				
	CDU/CSU	**SPD**	**FDP**	**GREEN**
CDU/CSU				
SPD	0.88			
FDP	0.81	1.12		
Grüne	1.03	0.68	1.07	
Die Linke	1.38	0.73	1.54	0.76

*Note: * Heterogeneity of Union parties CDU and CSU, the Comparative Manifestos Project (CMP) only provides data for the Union faction as a whole. Due to different base rates, we made the two distance measures (d_{LIWC} and d_{CMP}) more easily comparable by dividing the distances in each table by the average of all values in that table.*

TABLE 4
Share of tweets and sentiment of party-affiliated accounts

Account		**Content (message contains reference to)**			
		CDU	SPD	Merkel (CDU)	Steinmeier (SPD)
CDU	Share of messages	83.9%	12.9%	5.9%	1.6%
	Positive emotions	1.61	1.42	1.35	1.82
	Negative emotions	0.68	0.87	0.79	1.09
SPD	Share of messages	25.5%	81.9%	6.2%	15.7%
	Positive emotions	2.01	1.79	2.79	2.61
	Negative emotions	1.4	0.53	1.06	0.61

Note: The content split addresses all messages containing a particular search term. Since some messages contain multiple party or candidate mentions the sum can be greater than 100%.

TABLE 5
Share of tweets and election results

Party	**All mentions**		**Election**	
	Number of tweets	Share of Twitter traffic	Election result	Prediction error
CDU	30,886	30.1%	29.0%	1.0%
CSU	5,748	5.6%	6.9%	1.3%
SPD	27,356	26.6%	24.5%	2.2%
FDP	17,737	17.3%	15.5%	1.7%
Die Linke	12,689	12.4%	12.7%	0.3%
Grüne	8,250	8.0%	11.4%	3.3%
			MAE:	1.65%

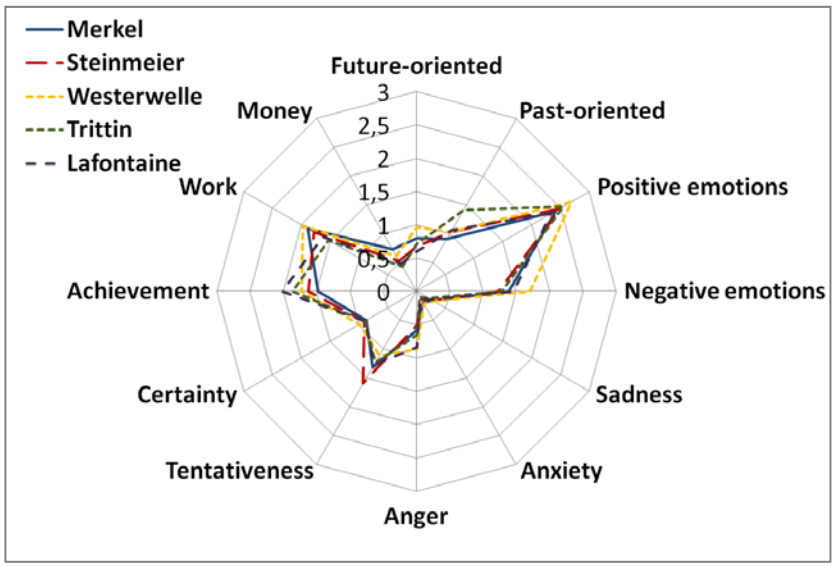
Note: Election result was adjusted to reflect only the 6 main parties in our sample. The number of tweets represents the total number of messages mentioning the party. If a message mentioned several parties, it was counted for each one of those parties. MAE = mean absolute error.

TABLE 6
Relative frequency of joint mentions

	CDU	**CSU**	**SPD**	**FDP**	**Die Linke**
CDU					
CSU	1.25*				
SPD	1.23*	0.71*			
FDP	1.04*	1.01	0.90*		
Die Linke	0.81*	0.79*	1.04*	0.97	
Grüne	0.84*	0.79*	0.98	1.06*	1.18*

*Note: * $p < .05$*

FIGURE 1
Profiles of politicians



Note: The profiles show the percentage of words in all tweets about a particular candidate which are related to the indicated dimension according to the LIWC dictionary. The radar charts connect these percentages for every candidate across all 12 dimensions.