

# Where There is a Sea There are Pirates: Response to Jungherr, Jürgens, and Schoen

Andranik Tumasjan<sup>1</sup>, Timm O. Sprenger<sup>1</sup>, Philipp G. Sandner<sup>1</sup>, and  
Isabell M. Welpel<sup>1</sup>

## Abstract

In their comment, Jungherr, Jürgens, and Schoen (2010) challenged part of the results presented in our original article (Tumasjan, Sprenger, Sandner, & Welpel, 2010). The present response addresses their points of concern and demonstrates that the conclusions drawn in Tumasjan et al. (2010) are well supported by both data and analyses.

## Keywords

Twitter, microblogging, computational social science, information markets, prediction markets, election forecasts, politics, elections, sentiment analysis

We welcome the comment by Jungherr, Jürgens, and Schoen (2010) to our article “Election forecasts with Twitter: How 140 characters reflect the political landscape” (Tumasjan, Sprenger, Sandner, & Welpel, 2010). We are pleased that our contribution has sparked the interest of and stimulated debate among scholars in this important and emerging field of research. The present response reveals that the conclusions drawn in Tumasjan et al. (2010) are well supported by both data and analyses. With regard to the critical issues raised by JJS, we demonstrate that, contrary to what JJS purport, we specify details of data collection and do not make “arbitrary choices without giving substantive account” (Jungherr et al., 2010) for our reasoning.

In their comment, JJS raise two main points of concern. First, they challenge our study’s selection of political parties. Second, the authors question our choice of the sampling time frame with regard to the replicability of our results. In the following, we respond to both issues raised by JJS.

Regarding the first issue, JJS have asked for additional clarification of our choice to include only the six parties CDU, CSU, SPD, FDP, Die Linke, and the Grüne in our analyses beyond the explanations we already provided in our original manuscript.<sup>1</sup> We have chosen the six parties based on two criteria. First, as specified in the original manuscript, these six parties were selected according to their presence in the German parliament at the time of data collection. Second—and more

---

<sup>1</sup>Technische Universität München, Munich, Germany

## Corresponding Author:

Andranik Tumasjan, Technische Universität München, Munich, Germany  
Email: [andranik.tumasjan@wi.tum.de](mailto:andranik.tumasjan@wi.tum.de)

importantly—our selection of parties and politicians exactly followed the selection of all major German election forecast institutes (i.e., Allensbach, Emnid, Forsa, Forschungsgruppe Wahlen, GMS, and Infratest dimap) who regularly report forecasting data (see e.g., Zicht & Cantow, 2010). Thus, in order to ensure comparability of our data with those of the major German forecasting institutes, we chose exactly the same selection of parties that are reported by name in the forecasts of those institutes, while excluding all other parties for which those institutes report only an aggregated percentage under the label “others.”

Third, following the sample selection criteria outlined above we argue – in contrast to JJS – that the *exclusion* rather than the *inclusion* of the Pirate Party is a more obvious a priori choice, given that they were neither represented in the German parliament nor specifically reported by name in the major election forecasts at the time of our study. By arbitrarily including the Pirate Party in their analyses, JJS build their argument exclusively on one exceptional outlier that has received short-term media attention in 2009.<sup>2</sup> However, before and after the 2009 electoral campaign, the Pirate Party has not received any more attention than any other of the German splinter parties. Hence, it remains theoretically and methodologically questionable to arbitrarily include only this particular party, while at the same time excluding all the other 20 (splinter) parties (e.g., “Allianz der Mitte,” “Freie Wähler Deutschland,” “Deutsche Kommunistische Partei,” “Die Violetten,” “Partei Bibel-treuer Christen,” “Rentner”) that have likewise run for parliament in the 2009 federal elections. To be consistent with their argument, JJS should have collected and analyzed data on all 27 parties that ran for parliament in 2009.<sup>3</sup> However, in their response, they unfortunately do not specify their rationale underlying the selection of only one additional seventh party, while excluding all remaining 20 parties. It would be rather interesting and illuminative to see the prediction results when data on all 27 parties are analyzed. Moreover, we want to call attention to JJS’s selection of keywords which they used to compile their data set as well as the classification of keywords to mentions of the respective party. For instance, while the classification of many keywords is straightforward (e.g., “cdu,” “spd,” “gruene/grüne” would all be classified as mentions of the respective party CDU, SPD, and Bündnis 90/Die Grünen), it is unclear which party the keyword “zensursula”<sup>4</sup> was assigned to. Since JJS do not specify the assignment of their keywords to party counts, we can only speculate about the classification of, for instance, the keyword “zensursula.” Since the “zensursula” campaign was mainly led by the Pirate Party, it could be that JJS counted this keyword as a Pirate Party mention (which could also partly account for the high number of Pirate Party mentions they found). However, this Twitter keyword can by no means be exclusively counted as a mention of the Pirate Party, since it was also heavily used by supporters of other parties (e.g., the Green party). Hence, while JJS are concerned that we “do not specify several details and make arbitrary choices without giving any substantive account” (Jungherr et al., 2010), their own analyses are highly intransparent due to lacking specification and substantiation of data collection and analysis details.

Fourth, JJS claim that our “method is dependent on public opinion surveys” and can therefore “hardly complement traditional methods of political forecasting” (Jungherr et al., 2010). However, in our article we explicitly maintain that analyzing Twitter sentiment may “complement”—rather than “replace”—traditional methods of political forecasting. At this point, we also want to clarify our method of data collection, since JJS were concerned about the exact data collection method that we applied. We collected our data by time-interval load balanced queries using the Twitter REST API. As outlined in our article, we “systematically collected all tweets that contained the names of either the six 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’” (Tumasjan et al., 2010, p. 5). Overall, then, we conclude that JJS’s arbitrary inclusion of the Pirate Party and their assignment of keywords to party mentions in their data set remain unclear and unsubstantiated.

The second issue that JJS are concerned about pertains to the sampling period of our data and the resulting replicability of our results. In their response, JJS argue that the mean absolute error (MAE) of predictions varies with the sample period. Their replications using different sampling time frames indicate that the maximum MAE across seven sampling periods is 3.34 percentage points ( $M = 2.07$ ,  $SD = 0.61$ , range: 1.51–3.34), whereas the MAE of our analysis is 1.65 percentage points.<sup>5</sup> In the following, we respond to their critique in detail. First, it comes as no surprise that the MAE varies with the sampling period because increasing the number of measurements (or, for instance, the length of a test) generally decreases the sampling random error. Indeed, the Pearson correlation between the number of days in the sampling period and the MAE in JJS's replication (see their Table 3) reveals a clear association ( $r = -.64$ ) between the two, indicating that the MAE decreases with an increasing sampling period. This result shows that sampling over a longer period does, indeed, result in an overall more accurate prediction. Thus, the varying MAEs that JJS found are a normal result of any empirical research and no particularity of our study.

Second, in the context of political campaigns, it is both absolutely obvious and expected that the results “depend considerably on the period under scrutiny” (Jungherr et al., 2010). This is also the case with any traditional polls, since, obviously, the political sentiment in the population changes over time due to the influence of behaviors and statements of politicians (e.g., speeches, affairs) and political parties (e.g., resolutions, campaigns), especially during an election campaign. On the contrary, it would be rather unrealistic to expect exactly the same results completely independent of the time frame. Nevertheless, as JJS's replication of our analyses demonstrates, the results pattern that we found does not change substantially across different time frames, which, as will be demonstrated in the following, rather speaks in favor of our results.

Third, we argue that the maximum MAE of 3.34 percentage points found by JJS in a sampling period of just two (!) days as well as the fairly constant results pattern underlines rather than questions the robustness of our results. We argue that the maximum MAE is relatively low given the fairly simple count method (as used in previous studies, e.g., Asur & Huberman, 2010) and the extremely short sampling period, especially when taking into account the accuracy of much more sophisticated, expensive, and time-consuming methods of traditional polls. Thus, as outlined above, we see the analysis of Twitter sentiment more as a “fast and frugal” (Gigerenzer & Todd, 1996, p. 3), and, thus, somewhat less accurate method which can indeed only – as we clearly state in our original article – “complement”, but in no case “replace” traditional forecasting methods.

Fourth, although the forecasting quality of Twitter messages is certainly the most arguable part of our article, it is by no means the main focus and, least of all, the only argument that we put forth. The main argument of our article is the much more general notion that the content of Twitter messages can be regarded as a valid indicator of public opinion that plausibly reflects offline political sentiment (rather than the fact that it can always be used to accurately forecast election results and replace traditional polls). As we point out in our original manuscript, our results show that “Twitter can be considered a valid indicator of the political landscape off-line” (Tumasjan et al., 2010, p. 13).

Fifth, we fully agree with JJS that it would indeed be desirable to extend the sampling period in such research. However, as in all empirical research, we had to make a sampling period selection decision based on certain criteria as will be outlined in the following. We chose August 13, 2009, as the beginning of our sampling period because prior to that date there were simply too few Twitter messages with party and politician mentions to assure robust analyses. We set the end of our sampling period on September 19, 2009, 1 week before the actual elections, because most forecasting institutes also do not publish any poll results already well in advance of federal elections. Thus, this date was chosen in order to ensure comparability of our results with the major German forecasting institutes and to test the *predictive* quality of Twitter messages' content. Most important, there was and is no reason to expect entirely different results or result patterns as is demonstrated by JJS's replication.

In sum, JJS's response set out to criticize one part of the results (i.e., the predictive power of Twitter messages) reported in Tumasjan et al. (2010). Their response raised mainly two issues, namely the lack of specification of our data collection as well as our selection of parties and the sampling period, which we have addressed in detail. Overall, we conclude that although we commend the authors for engaging in a scholarly debate on this important and emerging research topic, JJS's critique is largely unjustified and flawed in many respects as demonstrated in the present response.

### Declaration of Conflicting Interests

The authors declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

### Funding

The authors received no financial support for the research and/or authorship of this article.

### Notes

1. "We systematically collected all tweets that contained the names of either the six parties represented in the German parliament (CDU/CSU, SPD, FDP, B90/Die Grünen, and Die Linke)" (Tumasjan et al., 2010, p. 5).
2. Moreover, the Pirate Party's manifesto (election program) predominantly consists of Internet-based topics (e.g., Internet censorship, information privacy, data retention policies) which plausibly explains the relatively high share of Twitter buzz this party has generated in the forefront of the 2009 federal elections.
3. More precisely, 27 parties with state electoral lists (which, however, were not necessarily listed in all German federal states) ran for parliament in 2009.
4. "Zensursula" was a campaign against Ursula von der Leyen, the then German Federal Minister of Family Affairs, Senior Citizens, Women and Youth, as a reaction to her initiative to block child pornography on the Internet with the help of a block list maintained by the German Federal Criminal Police Office.
5. JJS point out that we do not specify the dates of the polls when stating that our MAE of 1.65 percentage points comes close to the "MAE of six research institutes, which published election polls in our sample period, [ranging] from 1.1% to 1.7%." (Tumasjan et al., 2010, p. 11). To specify, we compared our MAE to the average MAE of the six major German forecasting institutes (i.e., Allensbach, Emnid, Forsa, Forschungsgruppe Wahlen, GMS, and Infratest dimap), which published between 2 and 7 polls each, resulting in a total of 32 polls, during our sample period.

### References

- Asur, S., & Huberman, B. A. (2010). *Predicting the future with social media*. Retrieved from <http://arxiv.org/abs/1003.5699v1>
- Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In G. Gigerenzer, P. M. Todd, and the ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 3-34). Oxford: Oxford University Press.
- Jungherr, A., Jürgens, P., & Schoen, H. (2010). 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., & Welppe, I. M. "Election forecasts with Twitter: How 140 characters reflect the political landscape". *Social Science Computer Review*.
- Tumasjan, A., Sprenger, T. Sandner, P., & Welppe, I. (2010). Election forecasts with Twitter: How 140 characters reflect the political landscape. *Social Science Computer Review*. Advance online publication. doi: 10.1177/0894439310386557
- Zicht, W., & Cantow, M. (2010). *Sonntagsfrage Bundestagswahl*. Retrieved from <http://www.wahlrecht.de/umfragen/index.htm>

## Bios

**Andranik Tumasjan** (andranik.tumasjan@wi.tum.de) is a research assistant and PhD candidate at TUM School of Management at the Technische Universität München (TUM). His current research focuses on organizational behavior, entrepreneurial behavior, and social media research.

**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. Welp** (welp@tum.de) holds the chair for Strategy and Organization at TUM School of Management at the Technische Universität München (TUM). Her research interests are in strategy and organizational design and behavior, including the role of digital technologies and social media in and for organizations.