

Theocharis Y, Barberá P, Fazekas Z, Popa SA, Parnet O.

[A Bad Workman Blames His Tweets: The Consequences of Citizens' Uncivil
Twitter Use When Interacting With Party Candidates.](#)

Journal of Communication 2016, 66(6), 1007-1031.

Copyright:

This is the peer reviewed version of the following article: Theocharis Y, Barberá P, Fazekas Z, Popa SA, Parnet O. A Bad Workman Blames His Tweets: The Consequences of Citizens' Uncivil Twitter Use When Interacting With Party Candidates. *Journal of Communication* 2016, 66(6), 1007-1031, which has been published in final form at <https://doi.org/10.1111/jcom.12259>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

DOI link to article:

<https://doi.org/10.1111/jcom.12259>

Date deposited:

01/12/2017

Embargo release date:

28 October 2018

Trolling Alone? Incivility in Interactions with Candidates on Twitter in the 2014 European Elections

Abstract

Existing studies focusing on politicians' adoption of Twitter have found that they use it primarily as a broadcasting tool. We argue that citizens' impolite and/or uncivil behaviour is one possible explanation for such decisions. Social media conversations are rife with harassment and politicians are a prime target. This alters the incentive structure of engaging in dialogue on social media. We use Spanish, Greek, German and UK candidates' tweets sent during the run-up to the recent European Parliament (EP) elections, and rely on automated text analysis and machine learning methods to measure their level of civility. Our contribution is an actor-oriented theory of political dialogue that incorporates Twitter's specific affordances, clarifying how and why Twitter's democratic promise may be limited.

Keywords: political communication, machine learning, social media, Twitter, civility, politeness, automated text analysis

Trolling Alone? Incivility in Interactions with Candidates on Twitter in the 2014 European Parliament Elections

Over the past decade, social media have been integrated and widely used by politicians worldwide (Williams and Gulati, 2010; Grant et al., 2010). The ease of adoption, the capacity to bypass the mainstream media and create a personal publicity channel, and the limitless opportunities for personalised communication, have made them important campaign tools that candidates can use as a permanent form of communicating with the electorate (Larsson, 2015; Lee and Oh, 2012; Williams and Gulati, 2010; Grant et al., 2010). Twitter, perhaps the most widely adopted platform by politicians and one with the capacity to enable a more direct and interactive engagement with the public, was supposed to open the door for more citizen voice and participation in the political process via different means, counteracting one of the main inhibitors of political involvement – the fact that “nobody asks” (Rosenstone and Hansen, 1993).

Despite this promise, neither the adoption, nor the use of Twitter by politicians managed to live up to these normative expectations. Even though this is often seen as a supply-side problem, attributed to politicians’ tendency to not take advantage of the platform’s interactive opportunities and their persistence on using the platform in a broadcasting style (Grant et al., 2010; Graham et al., 2013), fewer explanations have taken into consideration the interaction between the supply and demand side, as well as the platform’s own limitations and “dark sides”.

In this paper we address the question of why politicians may be using the platform in ways seemingly inconsistent with the promotion of democratic deliberation. But rather than focusing on the supply side only, we take advantage of the unique asymmetrical relationship structure of Twitter and advance existing literature by proposing an explanation that lies in the *interaction between the supply and demand sides*. Specifically, taking into consideration that information and

communication technologies have built-in features and affordances that can both enable and constrain social relationships (Latour, 2005), we investigate the possibility that the demand side, i.e. the users, bolstered by Twitter's wall of anonymity and the platform's weak capacity to deal with harassment and trolling (Hern, 2015), may also be falling short of their responsibilities as counterparts in political discourse. We argue that not only the style of a candidate's engagement with their followers, but also their decision to interact with someone in a public space is subject to decisions and trade-offs involving whether some sort of civilised and constructive dialogue can take place.

Although most of online interactions are civilised, online spaces are rife with incivility and abuse (Mason, 2016). While extreme cases of uncivil behavior have often led to penalties and even imprisonment of political Twitter trolls (BBC, 2014), most of everyday trolling is probably considered unavoidable. However, this by no means indicates that the presence of such responses to candidates' tweets should not alter how they approach their social media communication. Politeness and civility are fundamental requirements for democratic discourse (Mutz and Reeves, 2005; Papacharissi, 2004) and the anonymity behind which many users choose to hide themselves allows for limitless abuse (Cheng et al., 2015; Davis, 2009), which can ultimately influence the motivations behind the communication style of candidates.

We empirically test this argument with data from the Twitter communication of Spanish, British, Greek and German candidates who ran for a seat during the 2014 European Parliament (EP) election. Our analysis provides evidence of a positive relationship between candidates' engagement on Twitter and exposure to attacks and harassment from citizens. Our theoretical contribution is twofold. We extend prior research by shifting attention to potential disincentives grounded in the behavior of the *public* that ultimately influences how candidates make use of Twitter. In this sense,

we explain why the use of social media might not be able to live up to its own promise for politics. Second, we show that this explanation fits within an actor-oriented approach of the use of social media in politics, but emphasize systematic differences contingent on candidate characteristics. By revealing the trade-offs that candidates for office face when articulating their communication strategies on social media, our study yields important insights about how the use of these platforms may affect the quality of public discourse and voters' knowledge of their options in the voting booth.

Candidates on Social Media: Engaging vs. Broadcasting Communication

Social networking sites and microblogging platforms have been put to use as everyday channels for reaching the public, and have been strategically embedded in local, national and supranational electoral campaigns (Gibson, 2013; Gulati and Williams, 2013; Koc-Michalska et al., 2014; Vergeer and Hermans, 2013; Vergeer et al., 2011b; Nulty et al., 2016). The sharp rise in social media adoption by candidates stems from the quick realisation that there are significant benefits in adopting these tools for enriching traditional political communication practices and enhancing the much-strained relationship with voters (Wattenberg, 2002). It has also given the opportunity to candidates in party-centered systems to engage in personal promotion outside the auspices of their parties (Larsson and Moe, 2011; Karlsen and Skogerbø, 2015).

Much of the recent literature on the political properties of social media has focused on social media platforms' different "affordances" (Earl and Kimport, 2011): for example, contrary to Facebook, Twitter is particularly suitable for an active, engaged style of messaging a candidate's followers due to the embedded asymmetrical structure of relationships that allows for direct interaction between unknown people (Grant et al., 2010). This makes Twitter of particular interest as it cannot only facilitate genuine engagement from the public but it can also have important benefits for

candidates. Lee and Oh (2012), for example, argue that directly addressing followers on Twitter can stimulate feelings similar to those of face-to-face communication, overcoming the depersonalising effects of digital communication and enhancing one's feelings of presence, ultimately increasing emotional closeness felt towards the candidate and eliciting positive evaluations (Lee and Shin, 2016). In the same vein, Veenstra and Lyons (2016, p.13) found that compared to an unengaging, broadcast-focused politician, one who includes conversational cues is likely to be viewed more positively overall. Despite these advantages over other platforms, empirical evidence, with few exceptions (Enli and Skogerbø, 2013), shows that Twitter is scarcely ever used in an interactive way by politicians (Golbeck et al., 2010; Glassman et al., 2010; Graham et al., 2013; Grant et al., 2010; Golbeck et al., 2010; Larsson and Moe, 2011; Small, 2011). Although much research has focused on factors explaining the presence or frequency of policy elites' activity on social media (see Obholzer and Daniel, 2016; Nulty et al., 2016; van Dalen et al., 2015 for recent examples), less attention has been paid on what may be explaining candidates' *style* of use.

Incentive Structure of Adopting an Engaging Communication Style on Twitter

A crucial first step for understanding why candidates may not engage in dialogue with users on Twitter is to examine motivations for using Twitter that, despite not requiring interaction, can nevertheless offer them concrete benefits. Our overall approach can be situated within Benoit's functional theory of political discourse and Stromer-Galley's controlled interactivity thesis, both of which posit that candidates' communication and messaging tactics are predominantly aimed at achieving one goal: winning the election (Benoit, 2007, p. 32; Stromer-Galley, 2014, p. 2). Building on this research, as well as on the conceptualization of Enli and Skogerbø, we posit that in an actor (candidate) oriented framework, there are three clear, universal motives for investing

resources – in this context referring to personal time and money – on social media: *marketing*, *mobilisation* and *dialogue* (Enli and Skogerbø, 2013, p.763).

Marketing reflects the most obvious benefit as it allows for increased visibility (Lassen and Brown, 2011; Veenstra and Lyons, 2016) and provides ample opportunities for political message personalisation (Enli and Skogerbø, 2013). Maintaining a Twitter profile leads to greater reach and thus expands candidates' visibility during, but also outside, electoral campaigns. It functions not only as a personal publicity channel, allowing candidates to distinguish themselves from other candidates without depending on the news media (Benoit, 2007, p.35), but also as a method of rapidly reacting to critical ongoing political developments, communicating with the press, and responding to the spread of questionable information or personal attacks without being limited by gatekeepers. At the same time, Twitter affords candidates the opportunity to post messages in frames that they (or their consultants) think that present them in a positive light to their followers (Veenstra and Lyons, 2016), and allows them to present the content in a personal and direct way. Consequently, Twitter represents an incredibly powerful tool for building a public image and for revealing a public side as well. Twitter has been considered the quintessential social media platform for *mobilising* citizens for protest events (Lotan et al., 2011; González-Bailón et al., 2011; Barberá et al., 2015b; Theocharis et al., 2015), but is also ideally suited for voter *mobilization*. It allows the fast diffusion of speech announcements, invitations to campaign events, donation requests and volunteering requests at a very low cost (Williams and Gulati, 2010), with some studies suggesting that more intensive online activity even pays off at the polls, at least in the context of EP elections¹ (Vergeer et al., 2011a).

With positive outcomes for both candidates and democracy, marketing and mobilisation are already two strongly and sufficiently beneficial incentives for using Twitter. Most importantly, both yield

benefits without necessitating the adoption of an engaging style of communication with the citizens. A baseline hypothesis thus is:

Hypothesis 1 (H1): Politicians make broadcasting rather than engaging use of Twitter

Despite the clear benefits of broadcasting use, however, engaging in *dialogue* with citizens has consistently been the most desirable and revolutionary, from a normative point of view, aspect of the internet; one thought of as being able to benefit both the politician and, most crucially, democracy (Rheingold, 1993; Barber, 2004; Etzioni, 1993; Stromer-Galley, 2014). Perhaps the most well-documented cause for citizens' disconnection from politics is that they feel that they have no say in what government does and no influence in political affairs, primarily because there is little public dialogue and discussion with politicians, and because politicians do not listen to them (Henn et al., 2002; Coleman & Blumler, 2009; Hay, 2008; Stoker, 2009). The possibility of two-way interaction between citizens and political actors is, thus, seen as a major step towards re-establishing democratic accountability and facilitating public participation.

Two-way interaction on social media can thus not only reinvigorate democracy but can provide direct input from voters and improve political communication. Direct communication with the voters can play a significant role in repairing the damaged relationships between voters and politicians in general, in reinstating some level of trust through greater intimacy, and in facilitating the emergence of a democratic online public sphere by opening up a new avenue for citizen voice and deliberation. Importantly, and beyond the theoretical and normative benefits, empirical evidence shows that there are real gains in adopting an engaging (as opposed to broadcasting) style of Twitter – both for the candidate who makes the extra effort to engage the public, and for democracy in general (Lee and Shin, 2012; Lee and Oh, 2012; Veenstra and Lyons, 2016). Why, then, do candidates continue to use Twitter in a one-directional manner?

(Im)politeness and (In)civility as an Inhibitor of Engaging Use of Twitter

We argue that part of the explanation lies in the incentive structure and relates to trade-offs, risks and responsibility on the part of the candidate. Engaging citizens online has long been considered a risky business for politicians, and it has been supported that political campaigns do not use digital media to genuinely engage citizens and supporters but merely to create a “spectacle of interactivity” (Stromer-Galley, 2014, p.5). Early research has showed that the reasons why politicians were hesitant to use the interactive features of their websites lay not only in strenuous work schedules and limited time (Coleman and Blumler, 2009) but also out of fear of losing control over the content and due to concerns of losing intentional ambiguity over policy positions by having to specify claims or policy positions (Stromer-Galley, 2000). Yet this risk is substantially reduced on Twitter. The platform’s word limit allows for greater control of the content (than e.g. blogs, websites or even Facebook) and this laconicism is ideal for strategic ambiguity. These properties counteract two of the major inhibitors for directly engaging with the public – loss of control and ambiguity of campaign communication (Stromer-Galley, 2000). However, dialogue does come with responsibility. If one decides to engage, one must be prepared to follow-up (i.e. engage with multiple members of the public). This, due to the higher resources required may, *unless there are clear gains*, bring dialogue to the bottom of the incentive list.

We suggest that, in the outlined incentive structure, engaging in dialogue on Twitter comes at the bottom of a candidate’s list because much of the content addressed to them is democratically damaging, or undermines fundamental discussion norms. Despite the high level of control that Twitter messages enable, especially prominent politicians are often victims of abuse, with heavy insults directed at them seconds after they post². Extant psychological research on the side-effects of anonymous computer-mediated communication shows that sometimes communicators tend to

be more susceptible to group influence, social attraction, stereotyping, gender typing, and discrimination (Postmes et al., 1998, Postmes et al., 2001). At the same time, due to the internet's affordances, "harassers can take advantage of being unidentifiable, anonymous, and invisible, in addition to having immediate, easy-to-execute, almost untraceable escape route mechanisms" (Barak, 2005. p. 83). This implies that trolling in an environment such as Twitter is not only a very low-cost, but also a low-risk activity. Thus, the more a candidate attempts to engage, the more material will provide to potential trolls.

Previous research has found that impolite and uncivil, discourse can have a widespread poisonous and polarizing effect on discussions (Anderson et al., 2013, Veenstra and Lyons, 2016). To our knowledge, there is no study examining how candidates or political actors in general, react to impolite and uncivil language. Although most existing research on incivility³ focuses mainly on candidates' attacks on one another (Mutz and Reeves, 2005; Mutz, 2015; Brooks and Geer, 2007), in this study we argue that much of the content that is addressed to politicians on Twitter also goes far beyond robust discussion (Bartlett, 2015), being, at best, impolite and, at worst, uncivil. Impoliteness and incivility tend to be conflated due to their conceptual resemblance (Papacharissi, 2004, p.260). Specifically, even though for some scholars uncivil discourse is defined by "communication that violates the norms of politeness for a given culture" (Mutz, 2015, p. 6), we agree with Papacharissi that to capture incivility one needs to move beyond rudeness and poor manners. We thus provide a more fine-grained measure of incivility that involves impolite behaviour with direct democratic consequences, such as when people offend individuals or social groups by denying their personal freedoms and stereotyping them. This implies that we conduct a stricter test for incivility than previous studies. Based on these theoretical considerations, from a democratic point of view, engaging use of Twitter, which mainly involves dialogue with citizens, should be prioritised over broadcasting use that involves mobilisation and marketing. As from the

candidate's point of view conflict aversion should be prioritised, the above incentive structure changes so as to reflect a style of tweeting which leaves the candidate less exposed to risk, with less responsibilities and, at the same time, with as great a benefit as possible.

Against this background, it is reasonable to assume that there will be variation when it comes to broadcasting and engaging use of the platform by, say, a high-ranked candidate from a resourceful party who has strong presence in the media as well as dedicated staff, and young and upcoming candidates who are in a greater need to attract voters and thus have higher incentives to use the platform for engaging the public. Following this rationale, we formulate the following hypothesis:

Hypothesis 2 (H2): Engaging style of tweeting is positively related to impolite or uncivil responses

Data Collection and Case Selection

The data used in this paper was collected as part of the European Election Study 2014, Social Media Study. The study identified and collected the candidates list of all major parties competing in the 2014 EP election. Afterwards, starting from January 2014, the study created a list with all the Twitter handles and Facebook user names for candidates who were active on social media. This list was updated right before the election in May 2014. All in all, we found that across the entire space of the European Union a total of 2,482 out of 15,527 identified MEP candidates (16%) had a presence on Twitter (for a more detailed discussion see Nulty et al, 2016 and Supplementary Material C).

Using the Twitter firehose, we collected all the social media communication centered around the candidates resulting in a dataset containing every tweet, retweet, and response of a candidate as well as all the responses to these tweets. Furthermore, we also collected all the tweets that mentioned the candidates in any form. The data collection procedure lasted for 4 weeks from the

5th of May 2014 until June 1st 2014, covering the last 3 weeks of the electoral campaign and the week following the elections. The final outcome is a database of approximately four million tweets that we believe accurately depicts the Twitter communication in the 2014 EP election.

[Table 1 about here]

For the specific purpose of this paper we choose to concentrate only on four countries: Spain, Germany, UK, and Greece. These were chosen based on the degree of support for the EU and whether or not the countries received financial aid during the public debt crisis in the Eurozone, while also taking into account the use of Twitter during the campaign 2014 EP Election campaign (see more details in Supplementary Material X). As incivility is related more to contextual than to habitual factors (Herbst, 2010), our general expectation is that the level of politeness and civility would vary depending on these two contextual factors. To be more specific, in the online environment “weightier” frames are prone to generate a higher percentage of uncivil responses (Coe et al., 2014). Thus, the interplay of strong anti-EU feelings and severe economic conditions (i.e. receiving substantial financial international support) is expected to add more “weight” and also increase the level of conflict (i.e more polarization) around the EP elections (Popa et al., 2016; Hobolt and de Vries, 2015, 2016), resulting in a more frequent use of impolite and uncivil language in the social media communication around the EP elections. In Table 1 we provide a summary of our dataset used for our analyses with Twitter communication during the 2014 EP elections in all 4 countries included in this study.⁹

Automatic Classification of Social Media Posts

Generating a Labeled Dataset

In order to test our hypotheses, it was necessary to classify tweets along various different dimensions of interest, such as their level of politeness, or how engaging they are. We achieved this goal by selecting for labeling a random sample of 7,000 tweets in each country, which we then used to train a machine learning classifier that predicts the category to which all tweets in our dataset correspond. The coding scheme used in the labeling process was developed by the authors and contains the following three categories related to the tweet content:

1. **Communication style** is the dependent variable of this study and differentiates between *broadcasting* tweets (i.e. tweets that simply depict statement or an expression of opinion) and *engaging* tweets (i.e. tweets that are directed to someone else/another user or are a direct response to a previous tweet).
2. **Polite vs impolite** distinguishes between tweets that are written in a well-mannered and non-offensive way vs. tweets that are ill-mannered, disrespectful or contain offensive language.
3. **Morality/Democracy** refers to whether the tweet contains a reference to moral and/or democracy issues, which are roughly covered by the Freedom and Democracy Domain and the Social Fabric Domain present in the European Parliament Election Study 1979-2009, Manifesto Study (Braun et al., 2015).

In addition, we also constructed a measure of **incivility** for each tweet combining the information in these two last categories. We consider incivility as a subcategory of impolite tweets that also refer to moral issues or democracy (e.g. tweets that make reference to one of the following topics: freedom and human rights, traditional morality, law and order, social harmony, freedom and human rights, democracy, constitutionalism). The basic assumption that guides our operationalization is that impolite remarks with direct democratic consequences constitute an uncivil tweet. To be more

specific, by making impolite remarks such tweets stereotype and offend individuals/social groups and/or challenge their freedoms/rights, disrespecting thus collective democratic traditions. Further details of the coding scheme and examples for each category can be found in the supplementary material A.

The overall results of the coding process, including estimates of intercoder reliability and summary statistics for these variables are in the Supplementary Material B. We obtain above 80% coder agreement across the board despite the unbalanced distribution in terms of tone and content, resulting in lower reliability scores. Finally, we also emphasize that these results should not be considered as measures of reliability in the traditional sense (for concept measurement), as the classification stage incorporates any disagreement at the human coding level into the estimation. The data resulted from the coding procedure is supplemented by a number of other variables that will mainly serve as control in our analyses. These refer to both candidate (i.e. gender, incumbency status in the EP, electoral viability¹¹, estimated ideological position) and party features (size of party, incumbency status, placement on left-right and pro-anti EU dimension).

Training Machine Learning Classifiers

Using the dataset of labeled tweets from each country, we then constructed machine learning classifiers that allow us to estimate the probability that each individual tweet in each country in our dataset corresponds to one of the three categories of interest. Our analysis is divided in three steps: text preprocessing, training and validation of the classifiers, and application to our entire corpus (see Hastie et al., 2009 for a more technical description of this method, and Barberá *et al* (2016) for an application to media texts). As we describe in Supplementary Material B, in most cases we find levels of accuracy (percentage of tweets correctly predicted by our classifier) that outperforms the benchmark of just choosing the modal category for each variable.¹⁴ The performance of this

method is similar in magnitude to the intercoder reliability among the coders of the labeled set, which suggests that our classifier is able to approximate the quality of human coding.¹⁵

[Figure 1 about here]

As additional validation for the outcome of our automated classification method, we also examined that the terms that the model identifies as being most associated with each category indeed correspond to our expectations. As an example, we found that the classifier predicts as *engaging* those tweets that indicate direct communication (e.g. an @ sign followed by what could be the beginning of a message, such as “thank you” or “hi”), as *impolite* those tweets that contain insults and expletives, and as mentioning *moral and democracy* issues those tweets with words such as “freedom”, “democracy”, “peace” or “rights”. We also validated that our estimate of a given tweet being engaging is accurate by relying on our behavioral expectations. In particular, we tested whether tweets sent by candidates with a higher probability of being classified as engaging are receiving more responses by ordinary citizens, under the assumption that a good measure of whether candidates are reaching to voters is observing that voters are indeed reacting to that message. Figure 1 displays the results of this validation exercise. Here, we use a Poisson model where we regress the number of responses to each tweet on the predicted probability of that same tweet being considered engaging, and display the predicted number of responses and a 95% confidence interval. The results confirm our expectation and strengthen our claim that the automated classification method we employ is accurately measuring our dimensions of interest.

The final step in our analysis is to predict the labels for all the tweets in our dataset. To do so, we apply the same text-preprocessing procedure to the text of the tweets, construct the feature matrix, and compute the predicted probability that each tweet corresponds to one category or the other. Finally, we aggregate these probabilities at the candidate level, both for the tweets that he or she

sent, and for the tweets that he or she received (that contained a mention to twitter handle).

Analysis

In order to test our two hypotheses, we conduct two separate sets of analyses. First, we use the dataset described in the previous section to test if it is indeed the case that most candidates use Twitter as a tool to broadcast content, as our first hypothesis states. Second, we examine if part of the variation in candidates' willingness to engage with citizens on social media is related to their exposure to *impolite* tweets, our main independent variable. To test whether there is a positive relationship between these two variables, we estimate three complementary regression models, each of them using data at a different level of aggregation: across candidates, within candidates and over time, and across individual tweets. First, we aggregate all tweets at the candidate level and use multivariate linear regressions to demonstrate that candidates who send more engaging tweets are also more likely to receive more impolite responses, holding all else constant. Second, we adopt a dynamic perspective to provide evidence that candidates who send more engaging tweets in a given week are more likely to be exposed to more impolite responses the following week. Finally, we focus on individual tweets and rely on multilevel regression models that reveal that tweets that are classified as engaging also tend to receive more impolite responses.

Results

Do Politicians Make Broadcasting or Engaging use of Twitter?

Table 1 lists the means of our main dependent and independent variables (with distribution figure in Supplementary Material D). We find substantive variation in our variables of interest both across and within countries. Candidates in the United Kingdom and Spain tend to send more tweets that are directed to the users, although still a large proportion of tweets in these countries are classified

as broadcasting (47% and 55%, respectively), which provides support for our first hypothesis. Greece and Germany lie at the other extreme of this distribution – here, for most candidates less than 25% of their tweets engage with citizens in any way, and the total of broadcasting tweets is 81% and 74%, respectively. The variation within countries also appears to correspond to our expectations: candidates that belong to the Pirate Party in the UK, Spain, and Germany are clear positive outliers, with the highest average proportion of engaging tweets (68%, 61%, and 58%, respectively).

We also find variation across countries and within countries in our second variable of interest.¹⁷ Greece is by far the country with most impolite tweets: on average, 18% of all tweets mentioning a candidate were classified as impolite (vs 6% in Germany, 4% in Spain, and 5% in the UK). An examination of some of the outliers within each country corresponds to our expectations: e.g. 10% of tweets mentioning UKIP's Nigel Farage were impolite, and 20% of tweets mentioning German extreme right-wing activist Ricarda Riefeling were impolite. As we show in Supplementary Material D, these differences are stable over the campaign.

Do Engaging Tweets Receive More Impolite and Uncivil Responses?

Table 2 shows the results of our first approach to testing our second hypothesis: a set of multivariate linear regressions of the proportion of engaging tweets sent by each candidate on the proportion of impolite tweets they receive,¹⁸ weighing our observations by the number of tweets sent by each of them.¹⁹ We find clear support for our hypothesis. In the first two models, where we add country fixed effects and our main set of control variables, we find a positive partial correlation between engaging tweets sent and impolite tweets received: the model predicts that an increase of 25 percentage points in engaging tweets sent (which is similar to a change from the 25th to the 75th percentile in this variable) is associated with an increase in impolite tweets received of 0.76

percentage points, which corresponds to around 12% of the std. deviation in this variable. In other words, the results suggest that candidates who try to engage in conversations with voters receive more vitriol.

[Table 5 about here]

This result is robust to the inclusion of other potential covariates that might explain the relationship between these two variables, such as the number of followers, the vote share for the party they belong to, and their expected success according to their position on the party list. Although we did not have specific hypotheses regarding the effect of these covariates, the results are consistent with conventional wisdom: candidates from small parties and with few followers, as well as female candidates, appear to receive more impolite responses, even though these two last effects are not statistically significant. However, Model 3 shows that the magnitude of the estimated effect decreases when we control for the position on the left-right and European integration dimensions, which we measured by scaling the follower networks of the MEP candidates and the national MPs in each country.²⁰ We find that right-wing and pro-Europe candidates are more likely to receive impolite tweets.

In Model 4 we explore country-level heterogeneity by interacting our main independent variable with the country dummies. After computing the marginal effects of the number of engaging tweets sent, we find that the estimate has the expected sign in the UK (0.06, $p = 0.28$), Spain (0.02, $p = 0.32$), and Greece (0.06, $p < 0.01$), but not in Germany (-0.005, $p = 0.36$), and it is only significant in the Greek case. However, part of this pattern could be due to not having enough sample size to properly estimate country-level differences. Finally, we also try to disentangle the effects of impoliteness vs civility by replicating our analysis using as dependent variable a measure of the proportion of uncivil tweets received by candidates. In particular, this variable is the product of the

proportion of impolite tweets received by the proportion of tweets received that mention morality or democracy issues. As we discuss earlier in the paper, we consider incivility as impolite behavior with direct democratic consequences, because it features attacks on social groups and their rights. Here, we still find a statistically significant effect of engaging tweets sent on uncivil tweets received, and of similar relative magnitude: an increase from the 25th to the 75th percentiles in the independent variable is associated with an increase in uncivil tweets received of 0.11 percentage points (around 15% of the standard deviation in this variable).²¹

One of the limitations of our analysis is the possibility that candidate-specific characteristics such as their ideological positions explain *both* how often they engage with citizens on Twitter and the type of response they receive. To overcome this limitation, we now turn to a time-series analysis of how candidates' tweeting behavior changed during the campaign. We split the tweets sent by each candidate and the tweets mentioning each candidate by week, into three groups: tweets sent in the third week before the election, the second week before the election, and the week before.²² For each of these weeks, we then compute again the average probability that tweets by the candidate are classified as engaging, and also that tweets mentioning the candidate are impolite, which results in a panel dataset where the unit of analysis is candidate \times week.²³

Using this new dataset, we examine the relationship between candidates' communication style on Twitter and their exposure to impolite messages by estimating a bivariate linear regression with candidate fixed effects. Since our comparison is now within candidates, it is not necessary to control for other variables in the previous analyses, which remain constant. More specifically, we regress the *change* in the proportion of impolite tweets received on the *lagged* proportion of engaging tweets sent by that candidate. This allows us to observe whether candidates who interact with their followers more often are more likely to increase the levels of harassment they are

exposed to as a result, during the following week. Table 6 displays the results of this analysis, first pooling all data together and then for each of the four countries we consider. We find strong support for our hypothesis in the pooled model: candidates who are more engaging in their communication style tend to receive more impolite tweets as the campaign progresses. In particular, we estimate that a one-standard-deviation positive change in the proportion of engaging tweets (around 19 percentage points) increases impolite tweets received by 5.2 percentage points (around 73% of the standard deviation in this variable). As it was the case before, when we disaggregate by country we find coefficients in the expected direction, but generally not reaching conventional levels of statistical significance.

[Table 6 about here]

We turn to our third type of analysis, where we offer a more fine-grained examination of how individuals react to candidates by taking tweets as our unit of analysis. We consider only those tweets sent by candidates (134,330 during our period of analysis), and look up in our full dataset any tweet by citizens that was a direct response to each of these tweets.²⁴ We then aggregate the predicted probability of each response being classified as impolite to compute a measure of the level of negativity that candidates are exposed to, after they post each individual tweet.

Figure 4 provides a first look at the relationship between these two variables at the tweet level. Here, we display the predicted impoliteness in responses to candidate tweets, as a function of how engaging they are estimated to be, in a linear regression fitted separately for each country. In all cases we find strong, significant evidence that candidates' efforts to reach directly to voters tend to generate higher levels of impoliteness in citizens' responses.

[Figure 4 about here]

Part of this relationship could be due to candidate- or country-specific characteristics. In order to show that this result is robust to controlling for some of these other covariates, we now offer estimates from multilevel regression models with varying slopes where our key covariate is the probability that each tweet by the candidate is classified as engaging. This approach allows us to model the structure of the data (tweets nested within candidates) and is flexible enough to estimate whether the effect of engagement on impoliteness varies across candidates.

Table 7 displays the results of this analysis, which confirms our result that tweets that are classified as being engaging receive many more impolite responses. In particular, according to the results in Model 2, we find that an increase in the probability of a candidate tweet being engaging from 0.17 to 0.78 (25th and 75th percentiles in this variable, respectively) increases the average impoliteness in the responses to that tweet in 14 percentage points, which is equivalent to slightly over 3 standard deviations in this variable. As in the previous analyses, we find some heterogeneity across countries, but in this case we find positive and statistically significant effects in all cases. In particular, the estimated marginal effects are: 0.017 ($p < 0.01$) in Germany, 0.019 ($p < 0.01$) in Greece, 0.005 ($p < 0.01$) in Spain, and 0.011 ($p < 0.01$) in the UK.

[Table 7 about here]

Discussion

Social media have become an important platform for electoral campaigning. Twitter in particular, due to its asymmetrical structure, is an important campaigning tool for candidates. Not only is it an efficient tool for a politician's image promotion and policy position distribution, it also offers a platform for voter mobilization and provides a space on which candidates can present a more personal side, reducing the emotional distance with citizens. Most importantly, from a democratic

point of view, Twitter provides an incredible opportunity for interactive communication between candidates and citizens. On Twitter, candidates can listen to citizens' feedback directly, while they also have the opportunity to respond using a platform whose laconic conversational structure allows for short and concise messages that enable strategic ambiguity and reduce the danger of loss of content control. Interactive use has been shown to have benefits for both sides, with politicians standing to especially benefit by being generally seen more positively when they interact with the public than when they don't.

Despite these important advantages for the politician at a personal level and for democracy in general, previous research has shown clearly that politicians, even when generally active on the platform, choose to make broadcasting, rather than interactive use. We have confirmed this pattern in this paper as well. Extant literature has offered little insight as to why this may be the case on Twitter - a platform whose affordances shield candidates from many of the potential dangers highlighted in the literature (Stromer-Galley, 2000).

Here we have argued that, in the design of communication strategies on social media platforms, candidates face an important trade-off between what is normatively desirable and what can be advantageous during an election campaign. On the one hand, using social media websites like Twitter or Facebook to connect with the electorate and establish a constructive dialogue with them is normatively desirable, and at least *a priori* also what voters prefer. On the other, this type of behavior implies giving up some communication control in order to reap the interactive benefits of social media, and is thus risky: it can attract the vitriol of citizens who, protected by the apparent anonymity of the platform, harass or attack the candidate, downgrading the quality of the debate and discouraging others from participating while potentially also destroying the candidate's reputation. From this perspective, although as our study shows the majority of tweets addressed to

candidates do not include harassment, perhaps a strategy of just using social media as a one-way communication device, useful to bypass traditional media outlets and reach directly the electorate could actually improve candidates' electoral performance.

In this paper, we have provided evidence of the existence of this trade-off. Relying on a large dataset of social media posts related to the elections to the European Parliament in four different countries, and exploiting recent advances in automated classification of text, we have been able to measure the extent to which candidates engage in conversations with citizens, and also their levels of exposure to impolite and uncivil messages. Our results support the hypothesis that these two types of behavior are positively related: candidates with more engaging messages are also more exposed to criticism and harassment. This finding has important ramifications for political communication research, especially normative accounts about the benefits of two-way interaction, at least as far as Twitter is concerned. Placing less emphasis on what can be called *misaffordances* of social media, extant literature has generally assumed that (a) the two-way interaction enabled by platforms such as Twitter will be crucial for democratic reinvigoration, that (b) if citizens are given the opportunity to engage they will do so conforming to generally accepted standards of politeness, and that (c) candidates' efforts to reach the public are, more often than not, insufficient. Our study shows that, often, the opposite is true as the deliberative democratic potential offered by the platform's own affordances may be inhibited not because of the potential lack of willingness on the part of candidates, but because of the ways citizens often tend to behave in largely anonymous online contexts, and because of the constraints imposed (or at least not prevented) by the platforms themselves. These insights call for more research on the misaffordances of new digital tools, as well as on how to integrate the kind of citizen behaviours they give rise to – and their consequences, within our broader understanding of political and campaign communication dynamics. Importantly, moving the focus away from the well-studied phenomenon of incivility *among*

politicians, our findings urge that more attention in political communication research should be paid to the demand-side and the consequences of direct interactions between citizens and politicians.

Our analysis also illustrates the large potential of automated text analysis techniques applied to the study of social media platforms. Although the iterative process to develop a codebook and train coders required a significant effort, the data generated proved to be useful in training supervised learning algorithms that allowed us to code the content of hundreds of thousands of tweets with accuracy that matches human coding. In combination with recent development in crowdsourcing techniques (Benoit et al., 2015), we believe our approach will enable researchers to answer standing questions in the study of political communication that up to now required an extensive and expensive coding process. However, our analysis is not without shortcomings.

First, we are not able to establish whether these relationships are causal. We cannot distinguish whether candidates who send more engaging tweets attract more “trolls”, or whether they send this type of messages more often precisely because they are responding to such attacks. Our analysis of how candidates’ behavior evolves during the campaign partially addresses this concern, although an experimental setup would be more useful. Second, although we have tried to make a distinction between impoliteness and incivility, in our analysis we did not find any meaningful differences in their effects on candidates’ behavior. However, this could be due to the difficulty of distinguishing these two dimensions empirically, and not necessarily because they have identical effects. Finally, given the party-centered electoral system of the EP elections, we were not able to examine the effect of different campaign strategies at the candidate level of their subsequent electoral success, which is clearly a missing piece in the puzzle of why candidates may decide to make only broadcasting use of social media platforms.

References

- Anderson, A. A., Brossard, D., Scheufele, D., Xenos, M., and Ladwig, P. (2013). The "Nasty Effect": Online Incivility and Risk Perceptions of Emerging Technologies. *Journal of Computer-Mediated Communication*, 19(3): 373–387.
- Barak, A. (2005). Sexual Harassment on the Internet. *Social Science Computer Review*, 23 (1): 77-92.
- Barber, B. (2004). *Strong democracy: Participatory politics for a new age*. University of California Press, Berkeley.
- Barberá, P., Popa, S.A., Schmitt, H. (2015) Prospects of Ideological Realignment(s) in the 2014 EP Elections? Analyzing the Common Multidimensional Political Space for Voters, Parties, and Legislators in Europe. *Paper presented at the 2015 meeting of the Midwest Political Science Association*.
- Barberá, P., Wang, N., Bonneau, R., Jost, J.T., Nagler, J., Tucker, J., González-Bailón, S. (2015) The Critical Periphery in the Growth of Social Protests. *PLOS ONE*, 10 (11).
- Barberá, P., Boydston, A., Linn, S., McMahon, R., and Nagler, J. (2016) Methodological Challenges in Estimating Tone: Application to News Coverage of the U.S. Economy. *Paper presented at the 2016 meeting of the Midwest Political Science Association*.
- Bartlett, J. (2015). Which party leader gets the most abuse on Twitter? Retrieved May 9, 2016, from <http://goo.gl/LPk0y4>
- BBC (2014). Twitter troll: What I said was utterly appalling and disgusting. Retrieved May 9, 2016, from <http://goo.gl/NIfgXE>
- Benoit, W.L. (2007). *Communication in Political Campaigns*. New York, Peter Lang.
- Benoit, K., Conway, D., Lauderdale, B. E., Laver, M., and Mikhaylov, S. (2015). Crowd-sourced

text analysis: reproducible and agile production of political data. *American Political Science Review*.

Braun, D., Schmitt, H., Wüst, A. M., Popa, S. A., Mikhaylov, S., and Dwinger, F. (2015).

European parliament election study 1979-2009, manifesto study. GESIS data archive, Cologne.

Brooks, D. and Geer, J. (2007). Beyond Negativity: The Effects of Incivility on the Electorate. *American Journal of Political Science*, 51(1):1–16.

Cheng, J., Danescu-Niculescu-Mizil, C., and Leskovec, J. (2015). Antisocial behavior in online discussion communities.

Coe, K., Kenski, K., & Rains, S. A. (2014). Online and Uncivil? Patterns and Determinants of Incivility in Newspaper Website Comments. *Journal of Communication*, 64(4), 658–679.

Coleman, S. (2004). From Service to Commons: Re-Inventing a Space for Public Communication. in Tambini, D. and Cowling, J. (eds) *From Public Service Broadcasting to Public Service Communications*. London: IPPR: 88-98.

Coleman, S. and Blumler, J. (2009). *The Internet and Democratic Citizenship: Theory, Practice and Policy*. Cambridge University Press, Cambridge.

Danescu-Niculescu-Mizil, C., Sudhof, M., Jurafsky, D., Leskovec, J., and Potts, C. (2013). A computational approach to politeness with application to social factors

Davis, R. (2009). *Typing Politics: The Role of Blogs in American Politics*. Oxford University Press, New York.

Earl, J. and Kimport, K. (2011). *Digitally Enabled Social Change: Activism in the Internet Age*. MIT Press, Cambridge.

Enli, G. and Skogerbø, E. (2013). Personalized Campaigns in Party-Centred Politics. *Information, Communication & Society*, 16(5):757–774.

Etzioni, A. (1993). *The Spirit of Community: Rights, Responsibilities, and the Communitarian*

Agenda. Crown Publishers, New York.

Gervais, B. (2015). Incivility Online: Affective and Behavioral Reactions to Uncivil Political Posts in a Web-based Experiment. *Journal of Information Technology & Politics*, 12(2):167–185.

Gibson, R. (2013). Party Change, Social Media and the Rise of Citizen-initiated? Campaigning. *Party Politics*.

Glassman, M., Straus, J., and Shogan, C. (2010). Social Networking and Constituent Communications: Member Use of Twitter During a Two-month Period in the 111th Congress. *Journal of Communication Research*, 2(2):219–233.

Golbeck, J., Grimes, J., and Rogers, A. (2010). Twitter Use by the U.S. Congress. *Journal for the American Society for Information Science and Technology*, 61(8):1612–1621.

González-Bailón, S., Borge-Holthoefer, J., Rivero, A., and Moreno, Y. (2011). The Dynamics of Protest Recruitment through an Online Network. *Nature*, 1(197).

Graham, T., Broersma, M., Hazelhoff, K., and van't Haar, G. (2013). Between Broadcasting Political Messages and Interacting With Voters. *Information, Communication & Society*, 16(5):692–716.

Grant, W., Moon, B., and Grant, J. (2010). Digital Dialogue? Australian Politicians' use of the Social Network Tool Twitter. *Australian Journal of Political Science*, 45(4):579–604.

Gulati, G. and Williams, C. (2013). Social Media and Campaign 2012: Developments and Trends for Facebook Adoption. *Social Science Computer Review*, 31:577–588.

Habermas, J. (1989). *The Structural Transformation of the Public Sphere: An Inquiry into a Category of a Bourgeois Society*. MIT Press, Cambridge.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning*. Springer.

Hay, C. (2007). *Why We Hate Politics*. Polity Press, Cambridge.

- Henn, M., Weinstein, M. and Wring, D. (2002). A Generation Apart? Youth and Political Participation in Britain. *The British Journal of Politics and International Relations*, 4 (2): 167-192.
- Herbst, S. (2010). *Rude Democracy: Civility and Incivility in American politics*. Philadelphia, PA: Temple University Press.
- Hern, A. (2015). Twitter CEO: We suck at dealing with trolls and abuse. Retrieved May 9, 2016, from <https://goo.gl/6jhvJP>.
- Hobolt, S. B., & de Vries, C. (2015). Issue Entrepreneurship and Multiparty Competition. *Comparative Political Studies*, 48(9), 1159–1185.
- Hobolt, S. B., & de Vries, C. (2016). Turning against the Union? The impact of the crisis on the Eurosceptic vote in the 2014 European Parliament elections. *Electoral Studies*.
- Hwang, H., Borah, P., Namkoong, K., and Veenstra, A. (2008). Does Civility Matter in the Blogosphere? Examining the Interaction Effects of Incivility and Disagreement on Citizen Attitudes. In *58th Annual Conference of the International Communication Association*, Montreal.
- Karlsen, R. and Skogerbø, E. (2015). Candidate Campaigning in Parliamentary Systems. *Party Politics*, 21(3):428–439.
- Koc-Michalska, K., Lilleker, D., Surowiec, P., and Baranowski, P. (2014). Poland's 2011 Online Election Campaign: New Tools, New Professionalism, New Ways to Win Votes. *Journal of Information Technology & Politics*, 11(2):186–205.
- Larsson, A. (2015). The EU Parliament on Twitter - Assessing the Permanent Online Practices of Parliamentarians. *Journal of Information Technology & Politics*, 12(2):149–166.
- Larsson, A. and Moe, H. (2011). Studying Political Microblogging: Twitter Users in the 2010 Swedish Election Campaign. *New Media & Society*, 14(5):729–747.
- Lassen, D. and Brown, A. (2011). Twitter: The Electoral Connection? *Social Science Computer*

Review, 29(4):419–436.

Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press, Oxford.

Lee, E.-J. and Oh, S. (2012). To Personalize or Depersonalize? When and How Politicians' Personalized Tweets Affect the Public's Reactions. *Journal of Communication*, 62(6):932–949.

Lee, E.-J. and Shin, S. (2012). Are They Talking to Me? Cognitive and Affective Effects of Interactivity in Politicians' Twitter Communication. *Cyberpsychology, Behavior, and Social Networking*, 15(10):515– 520.

Lilleker, D., Koc-Michalska, K., Schweitzer, E., Jacunski, M., Jackson, N., and Vedel, T. (2011). Informing, Engaging, Mobilizing or Interacting: Searching for a European Model of Web Campaigning. *European Journal of Communication*, 26(3):195–213.

Lorenzo-Rodríguez, J. and Garmendia Madariaga, A. (2016). Going Public Against Institutional Constraints? Analyzing the Online Presence Intensity of 2014 European Parliament Election Candidates. *European Union Politics*, 17(2): 303-323.

Lotan, G., Graeff, E., Ananny, M., Gaffney, D., Pearce, I., and Boyd, D. (2011). The Revolutions Were Tweeted: Information Flows During the 2011 Tunisian and Egyptian Revolutions. *International Journal of Communications*, 5.

Mason, R. (2016). Senior Tory calls on government to overhaul internet abuse laws. Retrieved May 9, 2016, from <https://goo.gl/xqLPfL>

Mutz, D.C. (2015). *In-Your-Face-Politics: The Consequences of Uncivil Media*. Princeton, Princeton University Press.

Mutz, D. and Reeves, B. (2005). The New Videomalaise: Effects of Televised Incivility on Political Trust. *American Political Science Review*, 99(1):1–15.

Ng, E. and Detenberg, B. (2005). The Impact of Synchronicity and Civility in Online Political

Discussions on Perceptions and Intentions to Participate. *Journal of Computer-Mediated Communication*, 10(3).

Nulty, P., Theocharis, Y. Popa, S.A., Parnet, O. & Benoit, K. (2016) Social Media and Political Communication in the 2014 Elections to the European Parliament. *Electoral Studies*.

Obholzer, L. and Daniel, W. T. (2016). An online electoral connection? how electoral systems condition representatives' social media use. *European Union Politics*.

Papacharissi, Z. (2002). The Virtual Sphere: The Internet as a Public Sphere. *New Media & Society*, 4(1):9–27.

Papacharissi, Z. (2004). Democracy Online: Civility, Politeness, and the Democratic Potential of Online Political Discussion Groups. *New Media & Society*, 6(2):259–283.

Popa, S.A., Rohrschneider, R., Schmitt, H. (2016). Polarizing without legitimizing: The effect of lead candidates' campaigns on perceptions of the EU democracy. *Electoral Studies*.

Postmes, T., Spears, R. and Lea, M. (1998). Breaching or Building Social Boundaries? SIDE-Effects of Computer-Mediated Communication. *Communication Research*, 25 (6): 689-715.

Postmes, T., Spears, R., Sakhel, K. and de Groot, D. (2001). Social Influence in Computer-Mediated Communication: The Effects of Anonymity on Group Behavior. *Personality and Social Psychology Bulletin*, 27 (10): 1243-1254.

Rheingold, H. (1993). *The Virtual Community: Homesteading on the Electronic Frontier*. Addison- Wesley, Reading, Mass.

Rosenstone, S. and Hansen, J. (1993). *Mobilization, Participation, and Democracy in America*. Macmillan, New York.

Shils, E. (1992). Civility and Civil Society. In Banfield, E., editor, *Civility and Citizenship*, pages 1–16. PWPA, New York.

Small, T. (2011). What the Hashtag? A Content Analysis of Canadian Politics on Twitter.

Information, Communication & Society, 14(6):872–895.

Sobieraj, S. and Berry, J. (2011). From Incivility to Outrage: Political Discourse in Blogs, Talk Radio, and Cable News. *Political Communication*, 28(1):19–41.

Stoker, G. (2006). *Why Politics Matters: Making Democracy Work*. Palgrave Macmillan, London.

Stromer-Galley, J. (2000). On-line interaction and why candidates avoid it. *Journal of Communication*, 50(4):111–132.

Stromer-Galley, J. (2014). *Presidential Campaigning in the Internet Age*. Oxford University Press, New York.

Stroud, N.J., Scacco, J.M., Muddiman, A. and Curry, A. (2014). Changing Deliberative Norms on News Organizations' Facebook Sites. *Journal of Computer-Mediated Communication*. 20 (2): 188-203.

Theocharis, Y., Lowe, W., van Deth, J.W. & Garcia-Albacete, G. (2015). Using Twitter to Mobilize Protest Action: Online Mobilization Patterns and Action Repertoires in the Occupy Wall Street, Indignados and Aganaktismenoi Movements. *Information, Communication & Society*, 18(2): 202-220.

Thorson, K., Vraga, E., and Ekdale, B. (2010). Credibility in Context: How Uncivil Online Commentary Affects News Credibility. *Mass Communication and Society*, 13(3):289–313.

Veenstra, A. and Lyons, B. (2016). How (Not) to Talk on Twitter: Effects of Politicians' Tweets on the Whole Twitter Environment. In *Cybersociology, Behavior and Social Networking*, 19(1):8-15.

Van Dalen, A., Fazekas, Z., Klemmensen, R., & Hansen, K. M. (2015). Policy Considerations on Facebook: Agendas, Coherence, and Communication Patterns in the 2011 Danish Parliamentary Elections. *Journal of Information Technology & Politics*, 12(3), 303-324.

Vergeer, M. and Hermans, L. (2013). Campaigning on Twitter: Microblogging and Online Social Net- working as Campaign Tools in the 2010 General Elections in the Netherlands. *Journal of Computer- Mediated Communication*, pages n/a–n/a.

Vergeer, M., Hermans, L., and Sams, S. (2011a). Is the Voter Only a Tweet Away? Micro- blogging During the 2009 European Parliament Election Campaign in the Netherlands. *First Monday*, 16(8).

Vergeer, M., Hermans, L., and Sams, S. (2011b). Online social networks and micro-blogging in political campaigning: The exploration of a new campaign tool and a new campaign style. *Party Politics*, 19(3):477–501.

Wattenberg, M. (2002). *Where have all the voters gone?* Harvard University Press, Cambridge.

Williams, C. and Gulati, G. (2010). Communicating with Constituents in 140 Characters or Less: Twitter and the Diffusion of Technology Innovation in the United States Congress. In *Annual Meeting of the American Political Science Association*.

Endnotes

¹ In terms of party and candidate related differences, smaller or opposition parties have been found to be both early adopters and heavier users of the platform (Vergeer et al., 2011a) while, on average, in Europe younger and incumbent candidates report more activity on the platform (Lorenzo-Rodríguez and Garmendia Madariaga, 2016).

² Impolite remarks are presumably not always perceived as discouraging (see an overview of this argument in Papacharissi, 2004, p.262).

³ For a more complete overview of the effects of incivility on political discussions see Appendix.

⁴ Previous studies have shown that this is the case in other online platforms too. Davis (2009) argued that mockery and derogatory comments are so common on political blogs, that incivility is almost the default condition in such discussion forums – also Sobieraj and Berry, 2011).

⁹ Following a re-checking and validation of the completeness of our sample candidates in May 2015 we can estimate that our data covers: 86% of the Greek candidates, 89% of the UK candidates, 78% of the German candidates and 85% of the Spanish candidates. This re-enforces our belief that the data used in this study offers a very accurate coverage of the social media communication via Twitter at the time of the 2014 EP Elections.

¹¹ Following Hix et al. (2010), we classify candidates as "safe", "doubtful", and "unpromising" based on the candidate's list position relative to the potential number of seats predicted to be won by his or her party. We compute uncertainty about the outcome of the election as the standard deviation between the seats won by each party, and the electoral predictions published by Hix et al. (2014), based on TNS pre-election surveys. Candidates with a list position below the predicted seats minus one standard deviation are classified as "safe". Candidates with a list position above

the predicted seats plus one standard deviation are classified as “unpromising”. All other candidates were classified as “doubtful”. In the case of party lists that are not national (all parties in the UK, and CDU/CSU in Germany), we divided the predicted seats across districts based on their size relative to the total number seats per country.

¹⁴ The only exception is our “morality” classifier, which has low recall (many tweets that are not related to morality or democracy are still classified as such). This is perhaps not surprising given that this concept has a more complex meaning than the other two variables we consider.

¹⁵ We also find levels of accuracy similar to those reported in other studies that applied machine learning methods to the measurement of impoliteness in online settings. For example, Danescu-Niculescu-Mizil et al. (2013) report a maximum of 84% accuracy in coding of impoliteness in conversations on Wikipedia, only slightly below 86% agreement using human coding.

¹⁷ To be clear, we include not only tweets addressed directly to each candidate, but also those that mention them in any way, under the assumption that the candidate will receive a notification every time their name is mentioned, and can thus read what others are saying about them.

¹⁸ We multiply both variables by 100 to facilitate the interpretation of the regression coefficients.

¹⁹ We find substantively similar results if we estimate fractional logit models, which account for the nature of our dependent variables (proportions). However, to facilitate the interpretation of our results, here we report coefficients from OLS regressions.

²⁰ We used estimates provided by Barberá *et al* (2015), which were computed by applying a method similar to latent network modeling to the Twitter networks of individuals who follow each of these politicians. In the supplementary material we offer summary statistics for these variables.

²¹ Given the similar results for both variables, in the remaining analyses in this paper we focus on impoliteness, which is estimated here with a lower degree of measurement error.

²² We exclude from the analysis the tweets after the end of the campaign, since as we saw in Figure 3, they tend to increase in all countries, potentially as a result of factors unrelated to social media activity. In splitting the tweets, we take into account the fact that the EP elections in the UK took place on May 22nd 2014, but on May 25th 2014 in the other three countries.

²³ Note that we only consider weeks in which the candidate sent at least two tweets, in order to reduce measurement error.

²⁴ Direct responses are recorded as such in the metadata that accompanies each tweet as it is captured directly from the Twitter firehose, which allows us to unequivocally determine if a tweet is responding to another tweet or not. The average tweet by a candidate received 0.20 responses, and 90% of tweets receive 0 responses.

Figure 1: Validation: Citizens are more likely to respond to engaging tweets by candidates

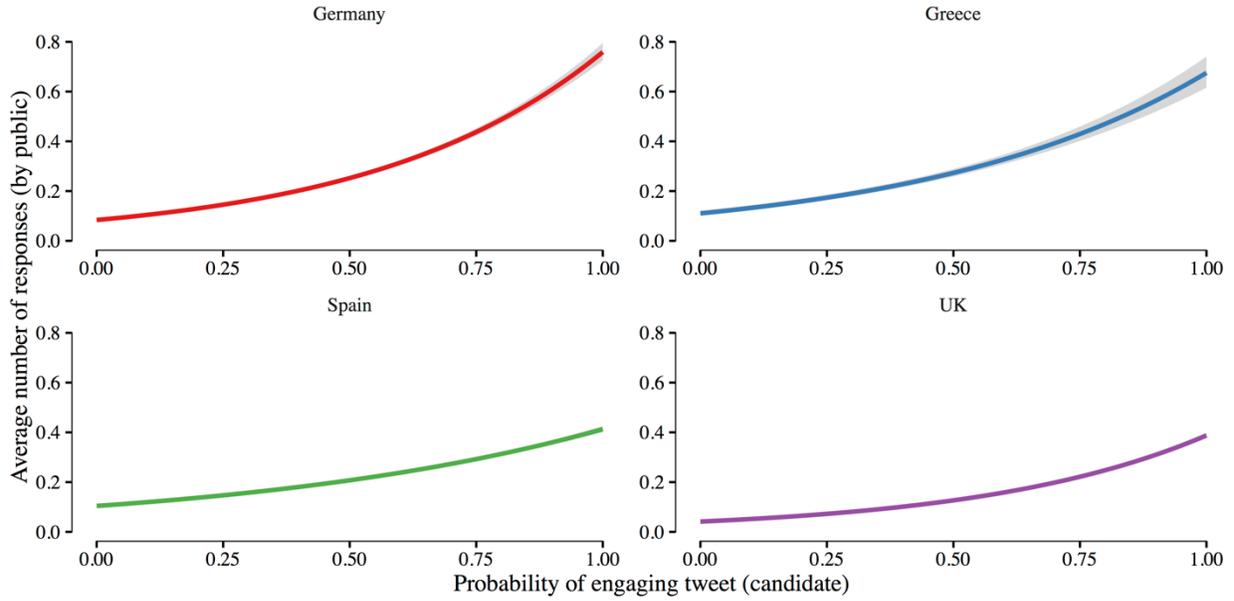


Figure 2: Impoliteness in responses to individual tweets at estimated probabilities of being engaging, by country

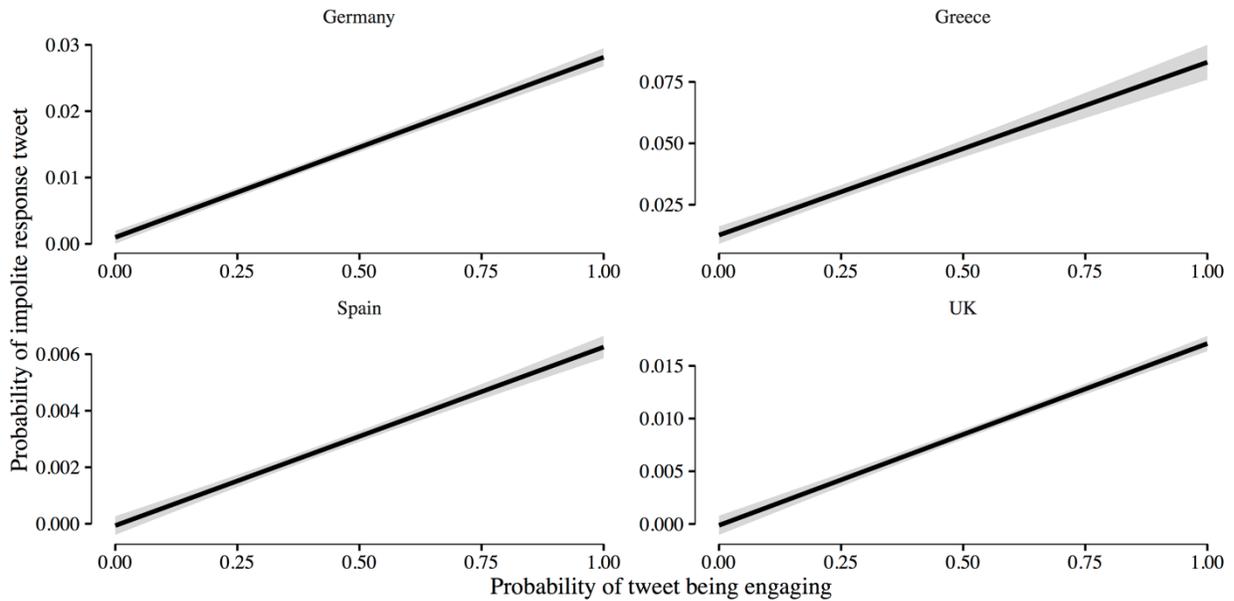


Table 1: Data coverage per country (included in multivariate analyses)

Country	Lists	Candidates	Total Tweets (including public)	Mean engaging tweets (candidates)	Mean impolite tweets (public)
Germany	10	92	80901	26%	6%
Greece	9	79	15057	19%	18%
Spain	12	211	447357	45%	4%
UK	28	271	251421	53%	5%

Table 2: OLS regressions of impolite tweets (Models 1–4) or uncivil tweets (Model 5) received on engaging tweets sent

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	4.56*** (1.02)	4.54*** (1.34)	15.38*** (4.11)	5.17*** (1.62)	0.41 (0.25)
% Engaging tweets sent	0.05*** (0.01)	0.05*** (0.01)	0.03** (0.01)	0.02 (0.03)	0.01*** (0.00)
Greece (dummy)	12.70*** (1.34)	12.74*** (1.34)	15.42*** (1.52)	12.58*** (1.96)	0.29 (0.28)
Spain (dummy)	-2.84** (1.13)	-3.20*** (1.14)	-3.21** (1.31)	-4.74** (1.91)	-0.22 (0.25)
UK (dummy)	-1.55 (1.15)	-1.72 (1.16)	-2.79* (1.46)	-2.65 (2.01)	-0.21 (0.24)
Candidate is incumbent		0.13 (0.49)	-0.20 (0.51)	0.15 (0.49)	-0.07 (0.07)
Viability: Safe		-0.13 (0.55)	-0.05 (0.53)	-0.13 (0.56)	0.05 (0.08)
Viability: Unpromising		0.07 (0.38)	0.01 (0.36)	0.10 (0.38)	-0.03 (0.06)
Candidate is male		-0.30 (0.26)	-0.28 (0.25)	-0.31 (0.26)	-0.07* (0.04)
log(count of followers)		0.14 (0.09)	0.22* (0.13)	0.15* (0.09)	-0.01 (0.01)
Vote share (national)		-5.37 (3.96)	-2.09 (4.48)	-5.40 (4.03)	-0.95 (0.86)
Prime minister (national)		0.08 (1.38)	0.00 (1.44)	0.07 (1.40)	0.07 (0.26)
LR position			-0.60** (0.25)		
EU position			-1.35** (0.59)		
Engaging x Greece				0.00	

				(0.05)	
Engaging x Spain				0.04	
				(0.04)	
Engaging x UK				0.03	
				(0.04)	
Num. obs.: candidates	600	600	455	600	600
Num. groups: parties	58	58	48	58	58
Var: party (Intercept)	3.84	3.76	4.96	3.92	0.22
Var: Residual	1783.59	1784.89	1447.82	1786.86	39.17

*** p < 0.01, ** p < 0.05, * p < 0.1

Table 3: OLS regressions of impolite tweets received on engaging tweets sent, with candidate fixed effects.

	All	UK	Spain	Germany	Greece
% Engaging tweets sent (lagged)	0.28**	0.07	0.35	0.43*	0.41
	(0.14)	(0.05)	(0.36)	(0.22)	(0.30)
Intercept	-0.12**	-0.05*	-0.15	-0.09*	-0.06
	(0.06)	(0.03)	(0.15)	(0.05)	(0.05)
N (candidates)	505	212	187	64	42
N (observations)	907	339	370	123	75
R ²	0.10	0.03	0.13	0.23	0.06

Dependent variable: Change in proportion of engaging tweets sent, by week.

Robust standard errors in parentheses. Sig.: *10% **5% ***1%.

Table 4: Multilevel linear regressions of impolite responses on engaging tweets, at individual tweet level

	Model 1	Model 2	Model 3	Model 4
Intercept	0.55** (0.22)	-0.65 (0.43)	-0.64 (0.61)	-0.87* (0.44)
Prob. tweet is engaging	0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)	0.02*** (0.00)
Greece (dummy)	1.82*** (0.32)	1.86*** (0.31)	3.07*** (0.31)	1.84*** (0.35)
Spain (dummy)	-0.64** (0.26)	-0.77*** (0.26)	-0.99*** (0.23)	-0.30 (0.30)
UK (dummy)	-0.16 (0.26)	-0.14 (0.25)	-0.86*** (0.27)	0.04 (0.29)
Candidate is incumbent		0.08 (0.31)	0.14 (0.24)	0.08 (0.31)
log(count of followers)		0.17*** (0.04)	0.21*** (0.06)	0.17*** (0.04)
Candidate is male		0.15 (0.12)	0.08 (0.09)	0.15 (0.12)
Prime minister (national)		0.15 (0.30)	0.07 (0.25)	0.15 (0.30)
Viability: Safe		-0.09 (0.23)	-0.09 (0.23)	-0.09 (0.24)
Viability: Unpromising		-0.07 (0.22)	-0.06 (0.22)	-0.07 (0.23)
Vote share (national)		-1.08 (0.89)	-1.02 (0.68)	-1.12 (0.89)
LR position			0.04 (0.06)	
EU position			-0.45*** (0.14)	
Engaging x Greece				0.00 (0.01)
Engaging x Spain				-0.01*** (0.00)
Engaging x UK				-0.01 (0.00)
Num. obs.: tweets	134330	134330	120798	134330
Num. groups: candidates	612	612	451	612
Num. groups: parties	59	59	48	59
Var: candidates (Intercept)	3.05	2.96	3.61	2.91
Var: candidates (Engaging tweet)	5.02	4.98	5.92	4.38
Cov: candidates (Intercept, Engaging tweet)	-2.41	-2.34	-3.45	-2.10
Var: party (Intercept)	0.11	0.09	0.00	0.09
Var: Residual	17.01	17.01	16.33	17.02

*** p < 0.01, ** p < 0.05, * p < 0.1