A Bad Workman Blames His Tweets

The Consequences of Citizens' Uncivil Twitter Use when Interacting with Party Candidates*

Yannis Theocharis[1] Pablo Barberá[2] Zoltán Fazekas[3] Sebastian Adrian Popa[1]

[1]*Mannheim Centre for European Social Research*

[2]*New York University*

[3]*University of Southern Denmark*

Abstract

The recent emergence of microblogs has had a significant effect on the contemporary political landscape. The platform's potential to enhance information availability and make interactive discussions between politicians and citizens feasible is especially important. Existing studies focusing on politicians' adoption of Twitter have found that far from exploiting the platform's two-way communication potential, they use it as a method of broadcasting, thus wasting a valuable opportunity to interact with citizens. We argue that citizens' impolite and/or uncivil behaviour is one potential explanation for such decisions. Social media conversations are rife with trolling and harassment practices and politicians are often a prime target for such behaviour, a phenomenon altering the incentive structures of engaging in dialogue on social media. We use all Spanish, Greek, German and UK candidates' tweets sent during the run-up to the recent EU election, along with the responses they elicited, and rely on automated text analysis to measure their level of civility. Our contribution is an actor oriented theory of the political dialogue that incorporates the specificity of the social media platform, further clarifying how and why democratic promises of such social media platforms are fulfilled or limited.

*Paper prepared for the *2015 Annual Meeting of the American Political Science Association*, 3-6 September, San Francisco, CA.

"Just confirmed: the next Star Wars film will be made in the UK. Great news for our creative industries. May the force be with us?"
– George Osborne, UK Chancellor (10:50 pm. 10 May 2013)

"**@GeorgeOsborne** Please kill yourself"
– Mlake9, Twitter user (10:51 pm. 10 May 2013)

"**@GeorgeOsborne** As though you've seen Star Wars you mealy mouthed prick"
– notchrisevans, Twitter user (10:53 pm. 10 May 2013)

"**@GeorgeOsborne** Wanker"
– rebecarrr, Twitter user (10:54 pm. 10 May 2013)

"**@GeorgeOsborne** Is Darth Vader your dad?"
– doctorstuffandguff, Twitter user (10:55 pm. 10 May 2013)

"We suck at dealing with abuse and trolls on the platform and we've sucked at it for years"
– Dick Costolo, Twitter CEO

Over the past decade, social media have been integrated and widely used by by politicians worldwide (Williams and Gulati, 2010; Barberá et al., 2014; 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 voters (Larsson, 2015; Lee and Oh, 2012; Williams and Gulati, 2010; Barberá et al., 2014; 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). This feature would not only allow the establishment of a public space in which which citizens can deliberate with their representatives, but also a deliberative space that enhances democracy (Papacharissi, 2002).

Despite this promise, neither the adoption, nor the use of Twitter by politicians managed to live up to these theoretical and 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, and in particular candidates who, entering the battle for office have a great interest in displaying an engaging profile to the public, may be using the platform in ways seemingly inconsistent with the promotion of democratic deliberation. Rather than focusing on the supply side only, as most previous studies have done (holding, as a result, politicians accountable for failing to live up to citizens' expectations), 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 all information and communication technologies have built-in features 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, 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 such as Twitter in the first place, is subject to decisions involving whether some sort of civilised and constructive dialogue can take place. While extreme cases of uncivilised behavior have often led to penalties and even imprisonment of political Twitter trolls (BBC, 2014), most of everyday trolling is seemingly ignored and probably considered unavoidable. However, this by no means indicates that the presence of such responses should not alter how candidates approach their social media communication. Politeness and civility are a fundamental requirement for democratic discourse (Mutz and Reeves, 2005; Papacharissi, 2004) and the anonymity behind which many users choose to hide themselves allows for limitless abuse (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 all Spanish, British, Greek and German candidates who stood for a seat during the 2014 elections to the European Parliament. We account for impoliteness, adopting a definition that relies mostly on abusive vocabulary use and bad manners, but also for incivility, thus examining the more far-reaching and lasting democratic consequences of abusive behaviour that are associated with morality, and are perceived as undermining collective traditions of democracy (Shils, 1992).

Our theoretical contribution is threefold. Incorporating specific features of the social media platform, 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 (i.e. the interaction between demand and supply side). In this sense, we explain why 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 to the use of social media in politics, but emphasize systematic differences contingent on candidate characteristics. Finally, we highlight the importance of the conceptual and empirical distinction between impolite and uncivil reactions and their limiting effects on democratic

dialogue between candidates and the public.

Candidates on social media: Engaging vs. broadcasting communication

Social networking sites such as Facebook and microblogging platforms such as Twitter 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; Barberá et al., 2014; Koc-Michalska et al., 2014; Vergeer and Hermans, 2013; Vergeer et al., 2011b). 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) which lend them special importance for specific functions for the candidates and the public as well. For example, while Facebook is more suitable for political event organisation, member management, and the communication of relatively long messages, 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 can not 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. This is particularly important due to the fact that, as Veenstra and Lyons (2014) note, emotional closeness felt towards a candidate has traditionally been among the highest campaign priorities because it can encourage support for a candidate. Lee and Shin's (Lee and Shin, 2012) study, which examined the effects of politicians' interactivity on Twitter in an experimental setting, points towards the same direction. According to the authors, a political figure's Twitter page consisting mostly of interactive tweets (i.e. frequent responses to other users), as opposed to seeing the same page with non-interactive tweets, elicited feelings of direct conversation to introvert participants and positive evaluations of the politician. In the same vein, Veenstra and Lyons (Veenstra and Lyons, 2014, p.13) found that compared to an unengaging, broadcast-focused politician, one who includes conversational cues – such as @replies, second-person pronouns, and the same hashtags their followers use in the discussion–, is likely to be viewed more

positively overall.

Despite these advantages over other platforms, empirical evidence shows that Twitter is scarcely ever used in an interactive way by politicians. In a study of EU parliamentarians, Larsson (2015) found very limited evidence that Twitter is used much outside electoral campaigns with differences between MEPs of different party groups being insignificant. Graham and colleagues (Graham et al., 2013), in a study of British MPs' tweets, found that politicians made largely unidirectional use of the platform, a finding corroborated by Lilleker and colleagues (Lilleker et al., 2011) who added that the platform was mainly used for self-promotion. The vast amount of tweets sent by US national legislators and examined by Golbeck and colleagues (Golbeck et al., 2010) were used for broadcasting information, while most of the tweets sent by Australian MPs in the study of Grant and colleagues (Grant et al., 2010) were primarily aimed at broadcasting rather than interacting with the public. With few exceptions (Enli and Skogerbø, 2013), most studies in the field report very similar findings (Glassman et al., 2010; Small, 2011; Larsson and Moe, 2011).

Although much research has focused on factors explaining the presence or frequency of policy elites' activity on Twitter, less attention has been paid on what may be explaining, especially candidates', *style* of use. Yet, whether a candidate is making broadcasting or engaging use of the platform is not only important from the perspective of campaign effectiveness and voter mobilisation but, most crucially, from a democratic point of view as it can theoretically encourage public voice and repair the damaged communication between disillusioned citizens and their representatives (Stoker, 2006; Hay, 2007).

Understanding candidates' 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. 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, 2014) 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, 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 oppor-

tunity to post messages in frames that they (or their consultants) think that present them in a positive light to their followers (Veenstra and Lyons, 2014), and allows them to present the content in a personal and direct way. In a time where politicians are advised, and even expected, to self-represent themselves as public and private persons exposing their personal traits, Twitter represents an incredibly powerful tool for building a public image and for revealing a public side as well. In all, the benefits derived from increased visibility and personalisation go hand in hand: providing more information to the electorate while enabling a feeling of direct interaction on the side of the voters – although, as empirical evidence has shown, especially for affiliated individuals (Lee and Oh, 2012).

Twitter has been considered the quintessential social media platform for *mobilising* citizens for political events with numerous studies assessing its mobilising potential and effects, especially with regards to protesting and social movement action (Lotan et al., 2011; González-Bailón et al., 2011; González-Bailón and Barberá, 2013; Theocharis et al., 2015). However, the platform is also ideally suited for voter mobilization by allowing the fast diffusion of, for example, speech announcements, invitations to campaign events, donation requests and volunteering requests at a very low cost (Williams and Gulati, 2010). Extant research shows that Twitter has been an especially effective mobilising tool for a specific cast of policy makers[1], with some studies suggesting that more intensive online activity even pays off at the polls - at least in the context of EU elections (Vergeer et al., 2011a).

With positive outcomes for both the candidate (visibility, personalisation, votes) and democracy (higher turnout), marketing and mobilisation are already two strongly and sufficiently beneficial incentives for using Twitter in the first place. Most importantly, both yield benefits without necessitating the adoption of an engaging style of communication with the citizens – a much more demanding style of tweeting that not only requires more time for following-up on discussions but also entails the risk of attracting trolls. Our baseline hypothesis reflecting this incentive structure 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). From the perspective of the candidate, interaction can provide direct input from voters and improve political communication while establishing someone as a "Web 2.0" politician, bringing them closer to citizens prone to engaging through new participatory avenues. This is especially the case during elections whereby political campaigns become sites of political renewal in a democracy and celebrate citizens' role in it (Stromer-Galley, 2014, p.5). At the same time, direct communication with the voters can play a significant role

[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, 2015).

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, 2014). 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 on the incentive structure and relates to both 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. As Stromer-Galley putts it "the rhetoric of participation and the presence of interactive gadgets, blogs, Facebook profiles create merely a spectacle of interactivity" (Stromer-Galley, 2014, p.5). Early research on websites relying on interviews with politicians, has showed that the reasons why they are hesitant to use the interactive features of their websites lie not only in strenuous work schedules and limited time – as other studies have shown (Coleman and Blumler, 2009) – but also out of fear of losing control over the content and, strategically, due to concerns of losing intentional ambiguity over policy positions by having to specify claims or policy positions (Stromer-Galley, 2000; Chadwick, 2006). Yet, admittedly, this risk is substantially reduced on Twitter. The platform's word limit, set on 140 characters, allows for greater control of the content (than e.g. blogs, websites or even Facebook), while the platform-imposed laconicism is ideal for strategic ambiguity. These properties clearly counteract two of the major inhibitors for directly engaging with the public; loss of control and ambiguity of campaign communication (Stromer-Galley, 2000). This said, dialogue does come with responsibility. If one decides to engage, one must be prepared to follow-up, and this implies that one must also be ready to engage with multiple members of the public, which 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 undermines fundamental discussion norms. Despite the high level of control that Twitter messages enable, especially prominent politicians (but certainly not only them) become often victims of abuse, with heavy insults directed at

them seconds after they post[2]. Racist, homophobic, shaming or ridiculing remarks are hardly inspiring conversation starters and, as research has shown, may have strong negative consequences even for those simply observing an online discussion (Gervais, 2015; Veenstra and Lyons, 2014; Anderson et al., 2013)[3]. Indeed, much of the content that is addressed to politicians on Twitter goes far beyond robust discussion (Bartlett, 2015) [4], being, at best, impolite and, at worst, uncivilised. Although impoliteness and incivility tend to be conflated due to their conceptual resemblance (Papacharissi, 2004, p.260), there have been important conceptual clarifications in the literature which have rendered a clear distinction between the two. According to Papacharissi, to capture incivility one needs to move beyond rudeness and poor manners. Uncivil remarks involve impolite behaviour with direct democratic consequences, such as when people offend individuals or social groups by denying their personal freedoms and stereotyping them. Thus, impolite, and especially uncivil, discourse can can have a widespread poisonous and polarising effect in discussions (Prior, 2007; Anderson et al., 2013) even to those simply reading, but not participating in, the discussion (known also as "lurkers"), thus providing a disincentive for engaging in dialogue. This is further corroborated in studies showing that exposure to uncivil political talk induces feelings of anger and aversion which in turn reduces satisfaction with the message board discourse (Gervais, 2015). Similarly, and along the same lines as Anderson and colleagues (Anderson et al., 2013), Veenstra and Lyons (Veenstra and Lyons, 2014, p.14) found that, if a politician's message on Twitter is viewed unsympathetically, and presumably commented upon in an uncivil manner so as to reflect this, the entire discussion surrounding it may collapse.

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. Or:

$$\text{Democracy} \rightsquigarrow \text{dialogue} > \text{mobilisation} > \text{marketing}$$

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,

[2]Impolite remarks are presumably not always perceived as discouraging. Indeed, often competing with the Habermasian ideal of rational and enlightened discussion (Habermas, 1989), the literature on democratic deliberation has long stressed that politeness is far from a necessary element of democratic deliberation (Schudson, 1997). Excessive politeness and strict adherence to social norms is seen by some as negative, and as an attitude that leads to the avoidance of disagreement, and to token agreement through the pursuit of safe topics (Papacharissi, 2004, p.262). Such conversational attitude may come at the expense of passionate, robust and heated discussion which is often perceived as the true emancipatory aspect of democratic deliberation that furthers democratic ideals (Lyotard, 1984; Schudson, 1997; Fraser, 1992)

[3]Individuals respond negatively to incivility directed at them or their views, and it may even influence the formation of negative attitudes about the issue at hand (Hwang et al., 2008). Moreover, incivility in online exchanges makes participants perceive uncivil statements as less fair, informative and credible (Brooks and Geer, 2007; Ng and Detenberg, 2005) (But see the study by Thorson and colleagues (Thorson et al., 2010) who, however, operationalise incivility only as derogatory, bad-mannered comments)

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

with as less responsibilities as possible and, at the same time, with as great a benefit as possible. Or:

Candidate $\rightsquigarrow$ marketing $>$ mobilisation $>$ dialogue

One cannot fail to notice a tension in the above as things may not be as straightforward when it comes to adopting an engaging style on Twitter. Given the clear benefits of directly addressing people on Twitter, some candidates may be willing to take the risk - or to invest the necessary time - to engage the public. Furthermore it is plausible that structural constraints apply too and that, for example, candidates in countries where political elites and institutions enjoy high levels of citizen trust may be less likely to be harassed online and thus more comfortable in frequently engaging the public. Previous research has shown that there is variation among incumbents and challengers when it comes to Twitter adoption and frequency of using the platform, while studies have even revealed a geographical divide with between more active Northern European politicians and less active Southern European ones (Vergeer and Hermans, 2013, p.142).

Against the background provided above, it is reasonable to assume that there will be variation when it comes to broadcasting and engaging use of the platform too. For example, a high-ranked candidate from a resourceful party who has strong presence in mainstream media and other social media platforms, as well as specialised staff dealing with each one of those accounts, will probably has less incentives to engage the public – especially if a large amount tweets addressing her/him is abusive. To the contrary, young and upcoming candidates who are in a greater need to attract voters, will be more eager to show willingness in addressing the public and its needs through direct conversation, and thus have higher incentives to use the platform for engaging the public regardless of the amount of impolite or uncivil remarks they receive. Indeed, research has shown that candidates who are behind in the polls and have little to lose during electoral campaigns, are more likely to experiment in involving the public and supporters online than candidates leading the polls (Stromer-Galley, 2014, p.34). The challenge, as Stromer-Galley notes, is to determine which risks with digital campaigning methods are worth the rewards. Clearly, then, for some candidates, the benefits of adopting an engaging style - will outweigh the cost of being addressed in an impolite or uncivil manner. Following this rationale, we formulate the following hypothesis:

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

The hypothesis is in line with the thinking that engaging use of Twitter is at the bottom of candidates' incentives because (the risk of) harassment - and thus absence of possibility to have a constructive discussion - reduces their willingness to directly engage in discussion. Following this rationale, we also

expect that when citizens decide to make an engaging use of Twitter, that will actually have a negative impact on their electoral chances, since it attracts criticism and vitriol.

> **Hypothesis 3 (H3)**: Engaging style of tweeting is negatively related to candidates' likelihood of being elected

Data collection and case selection

The data used in this paper was collected as part of the European Elections Study 2014, Social Media Study. The study identified and collected the candidates list of all major parties competing in the 2014 EP elections. Afterwards, starting from January 2014, a number of research assistants checked which of the identified candidates had a Twitter and/or a Facebook accounts and created a list will all their corresponding Twitter handles and Facebook user names. Since we expected that some candidates would create social media profiles for the specific purpose of these elections, we also updated our list of social media accounts right before the election in May 2014.[5]. 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.

Based on this list, our partners at TNS Opinion – having direct access to the Twitter firehose – collected all the social media communication centered around the candidates. In the specific case of Twitter, this implied every tweet, re-tweet 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 around the 2014 EP elections. In Table 1 we provide a summary of what this implies for the 28 EU countries.

Due to limited resources, for the specific purpose of this paper we choose to concentrate only on four countries. These were chosen based on the degree of support for the EU[6] 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 Elections campaign. The general expectation is that the level of politeness and civility would vary depending on these two factors. To be more specific we generally expect a higher use of impolite and uncivil language in countries that were severely affected by the crisis (i.e. countries that received substantial international support) and in countries

[5] Still as we only had a finite resources and a limited period for collecting this information, we do not exclude the distinct probability that we missed some of the candidates having a social media account.

[6] Support for the EU is measured based on the EES 2014 Voter study item QP7 (i.e. "Generally speaking, do you think that (OUR COUNTRY)'s membership of the EU is...?": "A good thing"; "A bad thing"; "Neither a good thing nor a bad thing") and represents the proportion of the country samples that considers the EU to be "a good thing".

Table 1: Data coverage per country

Country	Lists	Candidates	on Twitter	Tweets
Austria	6	249	65	40,016
Belgium	8	76	23	23,430
Bulgaria	6	102	11	2,244
Croatia	6	66	20	1,822
Czech Republic	9	244	41	5,035
Cyprus	9 + Ind.	48	11	1,279
Denmark	8	101	54	19,885
Estonia	8 + Ind.	88	23	2,397
Finland	14	249	157	43,439
Germany	9	501	123	86,777
Greece	9	359	99	18,709
Hungary	6	211	16	536
Italy	9	653	299	336,258
Ireland	10 + Ind.	41	30	39,779
Latvia	14	170	55	10,782
Lithuania	15	257	29	642
Luxembourg	9	54	13	0
Malta	7	33	20	2,502
Netherlands	18	344	221	133,377
Poland	8	1035	165	57,477
Portugal	5	105	23	11,810
Romania	10	398	41	470
Slovakia	9	117	25	1,598
Slovenia	11	118	35	13,250
Spain	11	648	221	463,937
Sweden	17	373	233	87,927
UK	28	733	304	273,886

Note: each column indicates the number of lists and candidates competing in the election, the number of candidates with a Twitter account, and the total of tweets related to the election included in our dataset for each country.

where there are strong anti-EU feelings.

Table 2: Case Selection

	Received bailout	Did not receive bailout
High support for EU	Spain (55.4%)	Germany (68.5%)
Low support for EU	Greece (43.8%)	UK (41.4%)

Table 2 offers a summary of our case selection depending on the two factors. The upper left quadrant is represented by Spain, a country that although was severely affected by the EU crisis still has relatively high level of support for the EU, i.e. 60% of the respondents in Spain consider the EU to be a "good thing" compared to 55.4% the mean for the pooled data[7]. On the upper right quadrant the obvious choice is Germany, the country with one of the highest levels of support for the EU – 68.5% – and the main "donor country" during the crisis. The bottom right corner (i.e. low support and not bailout received) is occupied by the UK, a country where the strong level of opposition for the EU

[7]An alternative for this quadrant would have been Ireland where support for the EU is even more widespread (69.3% of the sample considers EU to be "a good thing"), still we consider Spain to be a more emblematic case as it had the highest degree of Twitter usage during the EP campaign (see Table 2).

is denoted by both the low percent of the samples which considers the EU to be a "good thing" (i.e. 41.4%, with lower levels being recorded only in Cyprus and the Czech Republic) and also by the current political landscape (i.e. UKIP where the winners of the 2014 EP elections and the referendum on whether the UK should remain a member of the EU is scheduled to take place before the end of 2017). Finally Greece is the best example of the country that was severely affected by the economic crisis and received the most consistent financial aid package, while also having very low levels of support for the EU, i.e. 43.8%[8] . All in all we consider that our case selection offers sufficient coverage on the two dimension of interest in orders to be able to generalize the results of this study across the EU.

Last but not least it is important to mention that in May 2015 we re-checked the list in the four countries analyzed. The check revealed 16 candidates in Greece, 39 in the UK, 35 in Germany and 38 in Spain that although had a twitter account at time of the campaign were not covered by our analysis.[9] Assuming that this further check helped us to identify the population of candidates with a twitter account, 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.

Coding

In order to classify tweets according to our categories of interest, we labeled a large random sample of tweets, 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 different categories related to the content of the tweet, such as its sentiment, politeness, communication style, political issues mentioned, and geographic level. Here we focus on three categories relevant to the scope of our paper:

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 or disrespectful and contain offensive language.

[8] Portugal and Cyprus would also fit in this quadrant, still the socio-political impact of the crisis is not as devastating as in Greece.

[9] Additionally 6 candidates in Greece, 9 in the UK, 15 in Germany and 9 in Spain created a twitter account after 15th of April 2014.

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 sub-category 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 illustrative example for each category can be found in Appendix A.

Having compiled the codebook, we recruited six coders that would each code around 7000 tweets. Our goal was to have around 7000 tweets coded in the language of each of the four countries, of which approximately 3500 tweets were coded by two coders so that we can assess inter-coder reliability.

The coding process started with a training session that took place in early March 2015 in which the coders were introduced to the coding scheme, the software used for coding (i.e. CrowdFlower[10]) and went through a number of short exercises (i.e. coding around 40 English language tweets). After the training session all coders were assigned the same 160 English language tweets as a follow-up exercise. This was used to evaluate the overall reliability across all six coders, offer feedback to the coders, and for minor adjustment of the codebook. Given that for the coding of the respective tweets the average reliability was satisfactory across all categories we went further with assigning the country-specific tweets. As a first step in mid-March the coders were asked to analyze 1000 tweets. After this stage was finalized, the reliability across all countries was re-assessed and in the cases[11] where the reliability indicators were not satisfactory the coders received detailed feedback. At this point we also introduced the language sub category to the filter question as we noted that in the case of Spain there were a number of tweets in Catalan and Basque, and also in the case of Germany the presence of two least leading candidates among the EP candidates (i.e. Martin Schulz for the S&D groups and Ska Keller for the Green) meant that a large proportion of the tweets addressed to them were not in German. Following this clarification the coders received the last batch of 6000 Tweets in early April. This was subsequently supplemented with 2000 tweets for Germany and 1000 tweets for Spain in order to compensate for

[10]It is worth mentioning that the coding scheme (see Appendix A) was embedded in the coding platform allowed so that the coders had always had an easy access to it.

[11]These were the polite vs. impolite category for Greece and the political vs. personal, EU Vs. national vs. subnational tweet; political content (which has 13 possible sub-categories) for Germany.

the language issue mentioned above. The coding process was finalized in mid-May, the overall results are summarized in Table 3.

Table 3: Coding process: summary statistics

		UK	Spain	Greece	Germany
Valid tweets coded (total)		6721	7465	6211	6348
Valid tweets coded by 2 coders		3413	1965	3193	3063
Communication	Broadcasting	1437	2537	3297	3283
Style	Engaging	5284	4928	2914	3155
	Reliability	85%	78%	85%	79%
	(agreement/alpha)	0.62	0.66	0.70	0.58
Polite vs.	Polite	6343	7293	4790	5916
impolite	Impolite	379	172	1421	522
	Reliability	80%	96%	80%	92%
	(agreement/alpha)	0.25	0.17	0.25	0.28
Morality	Moral	355	609	355	421
	Other	5000	5441	5000	2962
	Reliability	95%	94%	95%	92%
	(agreement/alpha)	0.52	0.40	0.52	0.50

Notes: the total number of valid tweets is less than 7,000 because here we exclude tweets we classified as "spam" or in other languages. As measures of intercoder reliability, we report the percent agreement between the coders for those tweets coded by two coders, and Krippendorff's alpha

Last but not least the data resulted from the coding procedure is supplemented by a number of other variables that would mainly serve as control in our regressions. These refer to both the individual traits of the candidate (i.e. gender, incumbency status in the EP, electoral viability,[12], estimated ideological position) and characteristics of the party (size of party, incumbency status, placement on left-right and pro-anti EU dimension). A complete description of these variables can be found in Appendix B.

Classifier training

Using the dataset of labeled tweets from each country, we then constructed a machine learning classifier that allows us to estimate the probability that all the tweets in each country correspond to one of the three categories of interest.

Our analysis is divided in three steps. First, we processed the text of the labeled tweets by removing stopwords in each of the four languages, converting to lowercase, transliterating all characters to ASCII (e.g. replaced *á* by *a*) to avoid problems with accentuation differences, stemmings all the words to convert them into tokens, and splitting the text into unigrams (tokens) and bigrams (sets of two

[12]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.

tokens). We kept all hashtags as they were published, but we substituted all Twitter handles by just an @ sign to avoid overfitting.[13] To further remove extremely rare and extremely frequent, which are likely to add noise to our classifier, we only consider n-grams that appear in two or more tweets, and in less than 90% of all tweets.

The second step in our analysis is to estimate the parameters of our classifiers. In particular, we use a regularized logistic regression with L2 penalty (ridge regression) that regresses a binary variable indicating whether the tweet corresponds to one or another category on a vector of n-gram counts that indicates the number of times each of the n-grams we consider is mentioned in that tweet.[14] We use regularization in order to deal with the sparseness in our feature matrix (each tweet only contains a few words, and the rest of word counts is zero) and because we have more variables than observations in our dataset (see Hastie et al., 2009 for a more technical description of this method, and Barberá et al., 2014 for an application to media text). Since the tweets in our dataset are written in different languages, we run a different model for each country and variable.

In Table 4 we report different measures of performance for our classifiers in each country. To compute these measures, we use 5-fold cross-validation: we split each dataset randomly into 5 sets ("folds") with 20% of the observations each; we train the classifier with the remaining 80% of the data, predict the labels for the remaining 20%, and compare with their true values; this procedure is repeated 5 times, each time using a different 20% "fold." 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, and similar in magnitude to the intercoder reliability among the coders of the labeled set, which suggests that our classifier is able to reproduce human coding.[15]

To ensure that the predicted values we are estimating correspond to our construct of interest, we also extracted the top predictive n-grams for each category, that is, the n-grams that correspond to the variables with the highest and lowest coefficients in the regularized logistic regression. In Table 5 we report the top 25 n-grams for the three categories of interest in the UK, to illustrate our results. As we expected, the classifier will predict as *engaging* those tweets that indicate direct communication (e.g. an at 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".

The third and final step in our analysis is to predict the labels for all the tweets in our dataset. To

[13] Since we are aggregating tweets at the candidate level, if tweets mentioning the name of a particular candidate are more likely to contain impolite content, then his or her name would be a good predictor of impoliteness, which would induce bias in our analysis.

[14] Note that in the classifier we exclude tweets marked as spam by our coders.

[15] 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.

Table 4: Classifier performance

		UK	Spain	Greece	Germany
Communication Style	Accuracy	0.821	0.775	0.863	0.806
	Precision	0.837	0.795	0.838	0.818
	Recall	0.946	0.890	0.894	0.832
	Baseline	0.752	0.662	0.509	0.549
Polite vs. impolite	Accuracy	0.954	0.976	0.821	0.935
	Precision	0.955	0.977	0.849	0.938
	Recall	0.998	1.000	0.953	0.997
	Baseline	0.949	0.976	0.825	0.937
Morality	Accuracy	0.895	0.913	0.957	0.922
	Precision	0.734	0.665	0.851	0.770
	Recall	0.206	0.166	0.080	0.061
	Baseline	0.879	0.906	0.954	0.919

Notes: *accuracy* is the % of tweets correctly classifier; *precision* is the % of tweets with predicted value of 1 (engaging; polite; related to morality) correctly classified; *recall* is the % of tweets with predicted value of 0 (broadcasting; impolite; not related to morality) correctly classified; *baseline* is the proportion of tweets in the modal category for each variable (engaging; polite; not related to morality

Table 5: Top predictive stemmed n-grams for classifiers

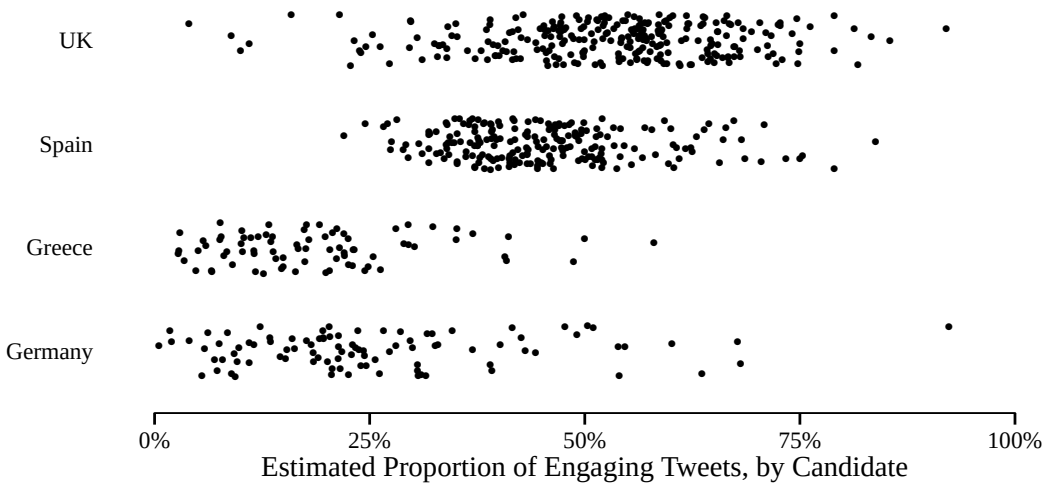
	Communication style
Broadcasting	just, hack, #votegreen2014, :, and, @ ', tonight, candid, up, tonbridg, vote @, im @, follow ukip, ukip @, #telleurop, angri, #ep2014, password, stori, #vote2014, team, #labourdoorstep, crimin, bbc news
Engaging	@ thank, @ ye, you'r, @ it', @ mani, @ pleas, u, @ hi, @ congratul, :), index, vote # skip, @ good, fear, cheer, haven't, lol, @ i'v, you'v, @ that', choice, @ wa, @ who, @ hope
	Politeness
Impolite	cunt, fuck, twat, stupid, shit, dick, tit, wanker, scumbag, moron, cock, foot, racist, fascist, sicken, fart, @ fuck, ars, suck, nigga, nigga ?, smug, idiot, @arsehol, arsehol
Polite	@ thank, eu, #ep2014, thank, know, candid, veri, politician, today, way, differ, europ, democraci, interview, time, tonight, @ think, news, european, sorri, congratul, good, :, democrat, seat
	Morality and democracy
Others	@ ha, 2, snp, nice, tell, eu, congratul, campaign, leav, alreadi, wonder, vote @, ;), hust, nh, brit, tori, deliv, bad, immigr, #ukip, live, count, got, roma
Moral/Dem.	democraci, polic, freedom, media, racist, gay, peac, fraud, discrimin, homosexu, muslim, equal, right, crime, law, violenc, constitut, faith, bbc, christian, marriag, god, cp, racism, sexist

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 mentioned to their twitter handle).

Results

Figure 1 displays the distribution of our dependent variable, the average proportion of tweets that are classified as *engaging* for each candidate to the European Parliament in each country. (To facilitate visualizing the results, we have added some vertical jitter to the plot.) We find substantive variation 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 (42% and 55%, respectively). 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 74% and 63%, 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).

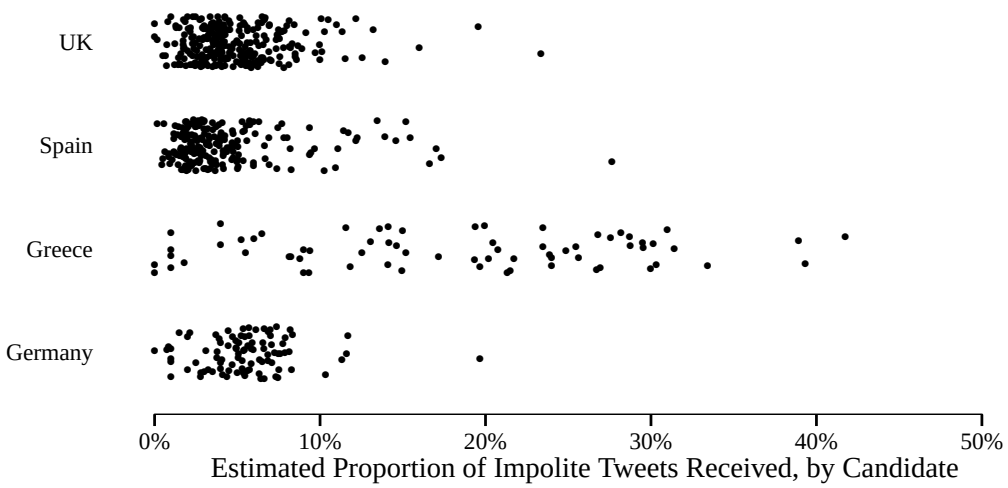
Figure 1: Proportion of engaging tweets, by candidate and country



Our first hypothesis stated that part of this variation is related to candidates' exposure to *impolite* tweets, our main independent variable. In Figure 2 we display the distribution of this variable, which we measure as the average probability that tweets that mention each candidate are classified as ill-mannered or disrespectful. As explained in the previous section, here 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. As it was the case with the previous variable, here we also find variation across countries and within countries. 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 extreme right-wing activist Ricarda Riefling were impolite.

Figure 2: Proportion of impolite tweets received, by candidate and country



To examine whether there is a positive relationship between these two variables, as we expected according to Hypothesis 2, we now show the results of multivariate linear regressions of the proportion of engaging tweets sent by each candidate on the proportion of impolite tweets they receive, weighing our observations by the number of tweets sent by each of them.[16] Table 6 shows our regression estimates.

We find mixed 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 impolite tweets received and engaging tweets sent: the model predicts that an increase of an standard deviation in impolite tweets (5.4 percentage points) is associated with an increase in engaging tweets of 1.6 percentage points. The coefficients for our controls are also consistent with our theoretical framework: incumbent MEPs that belong to large parties (measured as national vote share) are less likely to send engaging tweets.

However, the following three models show that the magnitude of the estimated effect decreases after including of other potential confounders. First, we try to disentangle the effects of impoliteness vs civility by adding an interaction effect with the proportion of tweets received by each candidate 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

[16]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.

Table 6: OLS regressions of engaging tweets sent on impolite tweets received

	Model 1	Model 2	Model 3	Model 4	Model 5
% Impolite tweets received	0.45*	0.30*	0.25	0.00	2.54*
	(0.13)	(0.13)	(0.20)	(0.14)	(1.29)
Greece (dummy)	−0.17*	−0.13*	−0.12*	−0.01	0.04
	(0.03)	(0.03)	(0.05)	(0.06)	(0.09)
Spain (dummy)	0.09*	0.07*	0.07*	0.09*	0.21*
	(0.01)	(0.01)	(0.03)	(0.02)	(0.08)
UK (dummy)	0.21*	0.21*	0.21*	0.16*	0.33*
	(0.01)	(0.01)	(0.03)	(0.03)	(0.08)
Candidate is incumbent		−0.06*	−0.05*	−0.06*	−0.06*
		(0.02)	(0.03)	(0.02)	(0.02)
Viability: Safe		−0.02	−0.02	−0.02	−0.02
		(0.02)	(0.03)	(0.03)	(0.03)
Viability: Unpromising		−0.03*	−0.03	−0.02	−0.03
		(0.01)	(0.02)	(0.02)	(0.02)
Candidate is male		−0.01	−0.01	−0.00	−0.01
		(0.01)	(0.01)	(0.01)	(0.01)
log(count of followers)		0.01*	0.01	0.02*	0.01
		(0.00)	(0.01)	(0.01)	(0.01)
Vote share (national)		−0.18*	−0.17*	−0.08	−0.15*
		(0.05)	(0.06)	(0.06)	(0.07)
Prime minister (national)		0.04*	0.04	0.05	0.04
		(0.02)	(0.03)	(0.03)	(0.03)
Moral/Democ. mentions			0.21		
			(0.30)		
Moral × impolite			−0.71		
			(2.25)		
LR position				−0.01*	
				(0.01)	
EU position				−0.08*	
				(0.01)	
Impolite × Greece					−2.43
					(1.30)
Impolite × Spain					−2.31
					(1.30)
Impolite × UK					−1.81
					(1.36)
Intercept	0.34*	0.35*	0.34*	0.75*	0.22*
	(0.02)	(0.03)	(0.06)	(0.11)	(0.09)
N	636	600	600	455	600
R^2	0.41	0.46	0.46	0.55	0.47

Robust standard errors in parentheses. * indicates significance at $p < 0.05$

and their rights. Our expectation was that incivility would have a much larger effect on politicians' behavior. However, the coefficient of this interaction is not significant, and in addition reduces the magnitude of the main effect, which suggests perhaps our measures are too noisy to capture this distinction.

Our main effect also disappears 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.[17] We find that conservative and anti-Europe candidates are less likely to send engaging tweets. In a way, this result is not surprising – our descriptive analysis already suggested right-wing candidates are more likely to be harassed on Twitter – and suggests that in order to examine our hypothesis we should turn to a within-candidate analysis, where we examine how their behavior evolves over time in response to changes in the content of the tweets they receive.

Finally, in Model 5 we explore country-level heterogeneity by interacting our main independent variable with the country dummies. After computing the marginal effects of the number of impolite tweets received, we find that the estimate has the expected sign in all countries, and statistically significant in Germany (2.54, $p < 0.01$) and the UK (0.73, $p = 0.01$), but not in Spain (0.22, $p = 0.16$) and Greece (0.16, $p = 0.27$).

As we mentioned above, 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 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 accordingly by week, into three groups: tweets sent in the third week before the election, the second week before the election, and the week before.[18] For each of these weeks, we then compute again the average probability that a tweet is engaging, and also that a tweet mentioning the candidate is impolite, which results in a panel dataset where the unit of analysis is candidate $\times$ week.[19]

Using this new dataset, we examine the relationship between candidates' communication style on Twitter and their exposure to impolite messages by estimating two separate linear regressions with candidate fixed effects. First, 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. Second, we examine whether there is also a positive effect in the reverse direction, by regressing the *change* in the proportion of engaging tweets sent on the *lagged* proportion of impolite tweets received. In contrast with the previous approach, we now focus on whether candidates who are more exposed to harassment tend to respond by adopting a more broadcasting style of tweeting. This would be indicative of a change of strategy: if engaging tweets attract trolls, candidates might respond by avoiding to engage users.

Tables 7 and 8 display 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 models: candi-

[17] In particular, we used the estimates in 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 (Barberá, 2015). In Appendix B we provide summary statistics for these two variables.

[18] 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.

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

Table 7: OLS regressions of impolite tweets received on engaging tweets sent, with candidate fixed effects.

	All	UK	Spain	Germany	Greece
% Engaging tweets	0.28**	0.07	0.35	0.43*	0.41
sent (lagged)	(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. Signif.: *10% **5% ***1%.

Table 8: OLS regressions of engaging tweets sent on impolite tweets received, with candidate fixed effects.

	All	UK	Spain	Germany	Greece
% Impolite tweets	0.69***	0.64*	0.44	1.32	0.45***
received (lagged)	(0.27)	(0.33)	(0.36)	(0.82)	(0.17)
Intercept	-0.05**	-0.04**	-0.02	-0.11*	-0.05
	(0.02)	(0.02)	(0.02)	(0.05)	(0.03)
N (candidates)	522	218	187	69	48
N (observations)	926	345	370	129	82
R^2	0.07	0.04	0.02	0.22	0.18

Dependent variable: Change in proportion of engaging tweets sent, by week. Robust standard errors in parentheses. Signif.: *10% **5% ***1%.

dates who are more engaging in their communication style tend to receive more impolite tweets as the campaign progresses but, contrary to our expectation they, in turn, appear to respond by increasing the proportion of engaging tweets they send, rather than switching to broacasting (which may be an indication that they are trying to deal with trolling behavior directly). 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; and a change of the same magnitude in impolite tweets (around 7 percentage points) in turn increases the proportion of engaging tweets by 5.3 percentage points. When we disaggregate by country, we find that the coefficients have the expected sign and similar magnitude in all cases, although most fail to reach conventional levels of statistical significance, particularly in the first table. One potential explanation for this pattern is that measurement error in these variables leads to attenuation bias in our coefficient estimates.

We now turn to a preliminary analysis of the effect of candidates' engagement on their likelihood of being elected to the European Parliament. As we discussed earlier in the paper, candidates face a trade-off between what is democratically desirable (dialogue) and what is more likely to result in electoral benefits (marketing). Our results provide preliminary evidence that dialogue with citizens indeed has negative electoral consequences. Table 9 shows the estimates from three logistic regressions

of whether each candidate was elected on the proportion of engaging tweets they sent. We find that a one-standard deviation increase in engagement (18 percentage points) is associated with a decrease in the probability of being elected of 2.4 percentage points for the average candidate in our sample. This quantity of interest was computed with the coefficients of Model 2, where we control for potential confounders of this relationship, such as measures of popularity (count of followers), activity on Twitter (number of tweets sent), and electability (position in list and national vote share). However, this result appears to be driven by the results in Greece – when we estimate the marginal effect of sending engaging tweets in each country, we find that it only has a statistically significant effect in that country.

Table 9: Logit regressions of being elected on Twitter variables

	Model 1	Model 2	Model 3
% Engaging tweets	−4.54*	−3.20*	−5.99
	(0.97)	(1.38)	(3.17)
Number of tweets sent	−0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)
log(count of followers)	1.05*	0.64*	0.64*
	(0.11)	(0.13)	(0.13)
Greece (dummy)	−2.60*	−0.55	0.60
	(0.53)	(0.75)	(1.40)
Spain (dummy)	−0.98*	−0.25	−1.56
	(0.37)	(0.65)	(1.77)
UK (dummy)	−0.01	1.33*	−0.33
	(0.38)	(0.68)	(1.40)
Candidate is incumbent		1.69*	1.68*
		(0.45)	(0.46)
Viability: Safe		2.75*	2.82*
		(0.70)	(0.75)
Viability: Unpromising		−2.62*	−2.83*
		(0.48)	(0.51)
Vote share (national)		−0.00	−0.00
		(0.02)	(0.02)
Prime minister (national)		0.40	0.26
		(0.66)	(0.69)
Engaging × Greece			−7.38
			(6.29)
Engaging × Spain			3.87
			(4.33)
Engaging × UK			4.71
			(3.63)
Intercept	−6.45*	−4.26*	−3.35*
	(0.77)	(1.21)	(1.39)
N	617	617	617
$\log L$	−214.64	−84.78	−72.86

Standard errors in parentheses

* indicates significance at $p < 0.05$

Discussion

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 is risky: it can attract the vitriol of citizens who, protected by the apparent anonymity of the platform, harass or attack the candidate, potentially destroying her reputation. From this perspective, 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, which suggests that candidates with more engaging messages are also more exposed to criticism and harassment, and that they respond to such exposure by increasing the number of engaging messages they send, which could in turn decrease their chances of winning a seat.

We have noted two important shortcomings in our analysis. 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 how candidates' behavior evolves during the campaign partially addresses this concern, although an experimental setup would be necessary to achieve causal identification. And second, it is possible that candidates who expect a lower likelihood of winning a seat send more engaging tweets under the assumption that such behavior would improve their chances. One potential way to disentangle this effect would be to examine whether engaging tweets could be reducing candidates' likelihood of getting elected because they generally have a negative tone

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.

Barber, B. (2004). *Strong democracy: Participatory politics for a new age*. University of California Press, Berkeley.

Barberá, P. (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using twitter data. *Political Analysis*, 23(1):76–91.

Barberá, P., Bonneau, R., Egan, P., Jost, J., Nagler, J., and Tucker, J. (2014). Leaders or Followers? Measuring Political Responsiveness in the US Congress Using Social Media Data. In *Annual Meeting of the American Political Science Association*.

Barberá, P., Boydstun, A., Linn, S., and Nagler, J. (2014). Economic conditions, economic perceptions, and media coverage of the united states economy. Paper presented at the 2014 EPSA Conference.

Barberá, P., Popa, S. A., and 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. In *MPSA Conference 2015*.

Bartlett, J. (2015). Which party leader gets the most abuse on Twitter?

BBC (2014). Twitter troll: What I said was utterly appalling and disgusting.

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.

Chadwick, A. (2006). *Internet Politics: State, Citizens and New Communication Technologies*. Oxford University Press, Oxford.

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

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.

Fraser, N. (1992). Rethinking the Public Sphere: A Conrtibution to the Critique of Actually Existing Democracy. In Calhoun, C., editor, *Habermas and the Public Sphere*, pages 109–142. MIT Press, Cambridge.

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 âĂŸCitizen-initiatedâĂŹ 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. and Barberá, P. (2013). The Dynamics of Information Diffusion in the Turkish Protests.

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.

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). PolandâĂŹs 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. (2015). Going Public with a Private Profile? Analyzing the Online Strategies of 2014 European Parliament Election Candidates. In *73rd MPSA Conference*, Chicago.

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.

Lyotard, F. (1984). *The Postmodern Condition*. University of Minnesota Press, Minneapolis.

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).

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.

Prior, M. (2007). *Post-broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections*. Cambridge University Press, Cambridge.

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.

Schudson, M. (1997). Why Conversation is Not the Soul of Decmoracy. *Critical Studies in Mass Communication*, 14(4):1–13.

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). OnâĂŘline 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.

Theocharis, Y., Lowe, W., van Deth, J., and García 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(22):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. (2014). How (Not) to Talk on Twitter: Effects of PoliticiansâĂŹ Tweets on the Whole Twitter Environment. In *Annual Meeting of the Association for Education in Journalism and Mass Communication*, Montreal.

Vergeer, M. and Hermans, L. (2013). Campaigning on Twitter: Microblogging and Online Social Networking 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*.

A Coding instructions

Social Media and 2014 EU Election Project

In this job, you will be presented with tweets about the 2014 European elections. You will need to classify each tweet into the following series of categories:

1. Polite Vs. Impolite

- **Polite** (a tweet that adheres to politeness standards, i.e. it is written in a well-mannered and non-offensive way)
  - *@paulmasonews why doesnt #EU take a longer term view? Doesnt #Germany remember their 1940s bailout allowing recovery & growth? #Greece*
- **Impolite** (an ill-mannered, disrespectful tweet that may contains offensive language. This includes: threatening one's rights (freedom to speak, life preferences), assigning stereotypes or hate speech ("nigger", "faggot"), name-calling ("weirdo", "traitor", "idiot"), aspersion ("liar", "traitor"), pejorative speak or vulgarity, sarcasm, ALL CAPS, incendiary, obscene, humiliating.
  - *@Nigel_Farage How's your dirty European non British dirty bitch of a wife? Is she ok? Can't imagine what it's like for you.*
  - *@SLATUKIP – "@DavidCoburnUKip Oh shut up David. You're a bore. @marley68xx"*

2. Communication Style

- **Broadcasting** (a statement or an expression of opinion)
  - *@PaulBrannenNE – "Labour's freepost election address dropping through letter boxes across the North East this week."*
- **Engaging: directed to someone else/another user** (a direct response)
  - *@GreenJeanMEP – "@klebudd Thank you Katie. We aimed for a positive campaign #VoteGreen2014"*

3. Political content (other categories omitted)

- **Morality and democracy** (tweets that make reference to one of the following topics: freedom and human right, traditional morality, law and order, social harmony, freedom and human rights , democracy, constitutionalism)
  - *@NATOWales but what about the defense of democracy and freedom of speech???*
  - *@Magee__ That was dropped. He was then arrested for the content of the speech.*

B Summary statistics

Table 10: Summary statistics: Germany

Variable	Mean	Std. Dev.	Min.	Max.	N
% engaging tweets sent	0.26	0.17	0.01	0.92	95
% impolite tweets received	0.06	0.03	0	0.2	93
% tweets about morality/democ. received	0.09	0.1	0.01	0.88	93
Incumbent candidate (dummy)	0.28	0.45	0	1	123
Position in party list	17.65	21.32	1	95	123
Candidate is male (dummy)	0.68	0.47	0	1	123
Tweets sent by candidate	109.47	201.45	0	979	123
Tweets received by candidate	557.48	3151.3	0	33466	123
Number of followers	3482.31	15298.34	1	155193	108
Ideology of candidate	4.75	1.27	-0.03	6.26	67
EU position of candidate	6.46	0.85	4.48	7.45	67
National vote share	14.49	12.42	0	34.1	123
National incumbent party (dummy)	0.39	0.49	0	1	123

Table 11: Summary statistics: Spain

Variable	Mean	Std. Dev.	Min.	Max.	N
% engaging tweets sent	0.45	0.11	0.22	0.84	211
% impolite tweets received	0.04	0.04	0	0.28	211
% tweets about morality/democ. received	0.1	0.06	0	0.43	211
Incumbent candidate (dummy)	0.07	0.26	0	1	225
Position in party list	18.81	14.42	1	54	225
Candidate is male (dummy)	0.6	0.49	0	1	225
Tweets sent by candidate	269.64	385.93	0	2655	225
Tweets received by candidate	1718.62	8342.45	0	99324	225
Number of followers	8452.71	61523.6	10	866563	205
Ideology of candidate	4.57	1.19	1.6	6.52	175
EU position of candidate	6.01	0.24	5.46	6.41	175
National vote share	8.4	11.84	0	41.9	225
National incumbent party (dummy)	0.05	0.22	0	1	225

Table 12: Summary statistics: Greece

Variable	Mean	Std. Dev.	Min.	Max.	N
% engaging tweets sent	0.19	0.11	0.03	0.58	79
% impolite tweets received	0.18	0.11	0	0.52	70
% tweets about morality/democ. received	0.04	0.04	0	0.28	70
Incumbent candidate (dummy)	0.08	0.27	0	1	99
Position in party list	20.75	13.07	1	42	99
Candidate is male (dummy)	0.66	0.48	0	1	99
Tweets sent by candidate	58.64	110.87	0	839	99
Tweets received by candidate	93.45	260.79	0	1692	99
Number of followers	2056.33	4797.74	3	37314	90
Ideology of candidate	4.63	1.91	-0.29	6.9	53
EU position of candidate	6.66	0.05	6.49	6.74	53
National vote share	15.18	11.04	0	29.7	99
National incumbent party (dummy)	0.4	0.49	0	1	99

Table 13: Summary statistics: UK

Variable	Mean	Std. Dev.	Min.	Max.	N
% engaging tweets sent	0.53	0.14	0.04	0.92	271
% impolite tweets received	0.05	0.03	0	0.23	266
% tweets about morality/democ. received	0.06	0.04	0	0.27	266
Incumbent candidate (dummy)	0.16	0.37	0	1	303
Position in party list	3.34	2.16	1	10	298
Candidate is male (dummy)					0
Tweets sent by candidate	169.44	330.07	0	3720	304
Tweets received by candidate	657.61	3078.34	0	48781	304
Number of followers	3119.31	13093.55	0	191616	264
Ideology of candidate	5.19	1.04	4.24	8.18	176
EU position of candidate	5.18	0.58	3.75	6.14	176
National vote share	15.61	14.33	0	36.1	304
National incumbent party (dummy)	0.32	0.47	0	1	304