Birds of the Same Feather Tweet Together. Bayesian Ideal Point Estimation Using Twitter Data

Pablo Barberá
pablo.barbera@nyu.edu

New York University

Abstract

Parties, candidates, and voters are becoming increasingly engaged in political conversations through the micro-blogging platform Twitter. In this paper I explore whether the structure of the social networks in which they are embedded has the potential to become a source of information about policy positions. Under the assumption that social networks are homophilic (McPherson, Smith-Lovin, and Cook, 2001), this is, the propensity of users to cluster along partisan lines, I develop a Bayesian Spatial Following model that scales Twitter users along a common ideological dimension based on who they follow. I apply this network-based method to estimate ideal points for Twitter users in the US, the UK, Spain, and the Netherlands. The resulting positions of the party accounts on Twitter are highly correlated with offline measures based on their voting records and their manifestos. Similarly, this method is able to successfully classify individuals who state their political orientation publicly, and a sample of users from the state of Ohio whose Twitter accounts are matched with their voter registration history. Finally, I introduce three applications that rely on Twitter-based ideal points: first, I estimate the posterior ideal point of the weighted average Twitter user in each state; second, I examine the extent to which online behavior is clustered along ideological lines; and third, I use sentiment analysis to examine the polarization of public opinion about presidential candidates.

1 Introduction

The micro-blogging service Twitter has become one of the most important communication arenas in daily politics. Despite being initially conceived as a website to share personal status updates, it has now become a massive phenomenon, with 200 million montly active users worldwide[1], including 15% of all online Americans[2]. All these users engage in a permanent interaction, 140 characters at a time, exchanging opinions and debating about news events in real-time. The content and structure of this conversation, easily accessible through the Twitter API, represents a unique opportunity for researchers interested in the study of elections and public opinion.

One distinct characteristic of this online social network is the presence of not only ordinary citizens, but also public officials, political parties, and candidates. A vast amount of information is publicly available about each of them: the content of their messages, who they decide to follow, and how they interact with other users. The purpose of this paper is to explore to what extent this new source of data can be used to estimate reliable policy positions for both types of actors.

Measuring parties' and voters' policy positions is a relevant, yet complex, scientific endeavor. Studies of government formation and stability (Strom, 1990; Laver and Shepsle, 1996; King, Alt, Burns, and Laver, 1990), political competition (Inglehart, 1990; Franklin, 2004; Dow, 2001; Adams, Merrill, and Grofman, 2005), policy outcomes (Blais, Blake, and Dion, 1993; Kedar, 2005), and institutional reform (Sartori, 1994; Reynolds, 2002) require systematic information on the placement of key political actors on the relevant policy dimensions. Empirical tests of spatial voting models (Downs, 1957; Stokes, 1963; Lau and Redlawsk, 1997; Jessee, 2009) also rely on measures of citizens' positions on these dimensions. This type of information is necessary not only to test whether they are accurate predictors of vote choice, but also in order to advance our knowledge of electoral behavior and public opinion.

A particularly promising aspect of Twitter is that ordinary users and politicians interact within the same symbolic framework. They use the same type of language, in messages of identical length, which feature the same external references and very frequently even the same content - as a result of the use of hashtags and "re-tweets" –, and most importantly, they are embedded in a common social network. This opens the possibility of estimating ideological positions of both political types of actors on the same scale, which could help overcome a significant limitation of the existing methods: the lack of a common ideological scale for legislators and voters (Shor, Berry, and McCarty, 2010).

[1]Source: Twitter's Official Twitter Account, December 18, 2012. [link]
[2]Source: The Pew Research Center's Internet & American Life Project, February 2012. [link]

The method I propose relies on the characteristics of the social ties that Twitter users develop with each other and, in particular, with the political actors (politicians, think tanks, news outlets...) they decide to follow. I argue that valid policy positions for ordinary users *and* political actors can be inferred from the structure of the following links across these two sets of Twitter users. Following decisions are considered costless signals that provide information about each users' perceptions of both his/her ideological location and that of the political accounts. This is thus a 'cheap following' model.

My argument hinges on the assumption that Twitter users prefer to follow other accounts whose ideology is similar to theirs. This is not a strong assumption for two reasons. First, it is a well-established finding that social networks are homophilic (McPherson, Smith-Lovin, and Cook, 2001) and thus individuals tend to relate and interact more often with those of similar traits. Second, given that Twitter is also a news media (Kwak, Lee, Park, and Moon, 2010), this pattern is reinforced by "selective exposure" (Bryant and Miron, 2004) to sources of information biased in the same direction as each user. Drawing an analogy with offline behavior, this argument would be equivalent to using the choice of sources of political information a voter makes as a proxy for his/her political preference.

More specifically, I implement a spatial following model that considers ideology as a latent variable, whose value can be inferred by examining which political actors each user is following. Whether a user i decides to follow a political account j is modeled as a function of the Euclidean distance between them on the latent ideological dimension, as well as two other parameters measuring the popularity of user j and how interested in politics user i is. Unlike similar studies (Conover, Gonçalves, Ratkiewicz, Flammini, and Menczer, 2010; King, Orlando, and Sparks, 2011; Boutet, Kim, Yoneki, et al., 2012), the method I propose allows us to estimate policy positions, with standard errors, on a multidimensional scale, for all types of Twitter users, and across different countries.

To illustrate the method, I generate ideal point estimates for a large sample of active users in the US and the three European countries with the highest number of Twitter accounts – the United Kingdom, Spain, and the Netherlands. In all four countries, parties and individual candidates have a very visible presence in Twitter, and engage in frequent conversations between each other and with ordinary citizens – which is a clear sign of the increasing use of this social network as a tool for political debate. At the same time, the variation in the size of their party systems allows me to examine whether the positions of the different parties on the resulting scale are congruent with their location on the left-right axis.

My results show that this method generates valid ideology estimates for both politicians and citizens. In the US, the resulting ideal point estimates for the members of the House and Senate are highly correlated with measures that rely on their roll-call votes. Similarly, most individuals who self-identify as "liberal", "independent", and "conserative" on their Twitter profiles are successfully classified on the left, center, and right of the resulting ideological scale. In order to further validate the method, I also match a sample of Twitter accounts from the state of Ohio with their voter registration records, based on their full name and county, finding that Twitter-based ideal points are good predictors of party registration. Finally, in the UK, Spain, and the Netherlands, the method I propose is able to cluster members of the same political party on similar locations of the latent dimension, and their positions are congruent with other measures based on manifestos or surveys.

I also provide three applications in Political Science where the method I propose can make a substantive contribution. First, I estimate the location of the average Twitter user in each state in the US. Second, I examine the extent to which online behavior during the 2012 US presidential election campaign is clustered along ideological lines. Finally, I examine the predictive power of ideology on the evaluation of presidential candidates – constructed using sentiment analysis (Pak and Paroubek, 2010). While my analysis does not provide complete certainty about the validity of this approach, it does suggest the potential of this procedure to accurately scale politicians and citizens along a meaningful axis.

The rest of this article proceeds as follows. Section 2 examines the existing literature on the use of Twitter data in the Social Sciences and discusses the opportunities and challenges in this field. Section 3 presents the ideal point estimation method I propose in the context of the different alternatives available. Section 4 describes the data, together with some basic summary statistics. Results of my analysis are shown in section 5, and three applications that use the resulting measures are presented in section 6. The article concludes in section 7 with a summary of my main findings and a list of possible paths for future research.

2 Background. Wading into the (Political) Tweet Stream.

The increase in the use of social media has led many social scientists to examine whether specific patterns in the stream of tweets might be able to predict real-world outcomes. Asur and Huberman (2010) show how a simple model measuring chatter from Twitter about movies predicts box-office revenues, outperforming market-based predictors[3]. Applying a similar method, Lampos, De Bie, and Cristianini (2010) are able to accurately

[3]However, a recent study conducted by Wong, Sen, and Chiang (2012) qualifies this conclusion. Using a machine learning algorithm that predicts that "positiveness" of the tweets about movie, these authors show that scores computed from Twitter reviews do not necessarily translate into predictable box-office performance, although it can be indicative of financial success.

track the prevalence of Influenza-like illnesses in several regions of the United Kingdom. Paul and Dredze (2011) extend this analysis into a broader range of illnesses in the United States, opening a whole new agenda in the field of public health research. In two highly publicized articles, Golder and Macy (2011) and Dodds, Harris, Kloumann, Bliss, and Danforth (2011) study the temporal patterns of happiness of millions of people in real time based on their *tweets*. An innovative study conducted by Hannak, Anderson, Barrett, Lehmann, Mislove, and Riedewald (2012) builds up on this research to show how weather affects aggregated sentiment. Measurements of collective mood states derived from Twitter feeds had already been found to be correlated to stock market indexes by Bollen, Mao, and Zeng (2011). Finally, exploiting the geographic information that Twitter users provide has allowed researchers to estimate the epicenter of earthquakes in Japan (Sakaki, Okazaki, and Matsuo, 2010) or even regional support for "American Idol" contestants (Ciulla, Mocanu, Baronchelli, Gonçalves, Perra, and Vespignani, 2012).

Given the accuracy of these predictions, and the consolidation of Twitter as a source of political information, a battlefield for campaigning, and a public forum of political expression, some researchers have wondered whether "tweets" validly mirror offline public opinion. "Can we analyze publicly available data to infer population attitudes in the same manner that public opinion pollsters query a population?" (O'Connor, Balasubramanyan, Routledge, and Smith, 2010, p.122). Were this approach to be successful, its advantages would be obvious: Twitter provides (relatively) easy and free access to millions of public messages in real-time and from most countries around the world. Inferring public opinion from social media messages is challenging, but also potentially very rewarding, given the wealth of this information.

The first studies of this kind, in the context of the German legislative elections of 2009 (Tumasjan, Sprenger, Sandner, and Welpe, 2010) and the first two years of the Obama presidency in the United States (Cummings, Oh, and Wang, 2010; O'Connor, Balasubramanyan, Routledge, and Smith, 2010) gave reasons to be optimistic. Tumasjan, Sprenger, Sandner, and Welpe (2010) found that "the mere number of messages [mentioning each German political party] reflect[ed] the election result and even [came] close to traditional electoral polls". O'Connor, Balasubramanyan, Routledge, and Smith (2010), on the other hand, showed that "a relatively simple sentiment detector based on Twitter data replicate[d] presidential job approval polls. The results highlight the potential of text streams as a substitute for traditional polling". Similar studies conducted in Singapore (Choy, Cheong, Laik, and Shung, 2011; Skoric, Poor, Achananuparp, Lim, and Jiang, 2012), United Kingdom (Lampos, 2012), Netherlands (Sang and Bos, 2012), and Spain (Congosto, Fernández, and Moro Egido, 2011) have also found that, during the elections, "the Twittersphere represents a rich source of data for gauging public opinion and that the frequency of tweets mentioning names of political parties, political candidates and contested constituencies could be used to make predictions about the share of votes at the national level" (Skoric, Poor, Achananuparp, Lim, and Jiang, 2012, p.2583). With these results, Cummings, Oh, and Wang (2010) even wondered "who needs polls?".

The response to this set of papers arrived in two recent research articles by Metaxas, Mustafaraj, and Gayo-Avello (2011) and Gayo-Avello (2012). These authors warn against "turning social media into another 'Literary Digest' poll" and claim that the "predictive power of Twitter regarding elections has been greatly exaggerated". They illustrate their concerns with analyses of several Senate races in the two most recent US Congressional elections, and find that electoral predictions applying similar methods as those used by the previous authors do not perform better than chance. These authors also criticize the previous results in this literature. For example, they point out that if Tumasjan, Sprenger, Sandner, and Welpe (2010) had not restricted their analysis to parties with parliamentary representation, the Pirate Party would have won the 2009 German elections – it was the party with the highest number of mentions in Twitter (Jungherr, Jurgens, and Schoen, 2011). In their view, an accurate prediction can only come through "correctly identifying likely voters and getting an un-biased representative sample of them". Self-selection biases, overrepresentation of younger, more educated citizens on Twitter, and the simplistic assumptions of the existing sentiment analysis techniques are the three most important methodological challenges to overcome.

This set of research papers evidenced some of the challenges that the use of Twitter data presents. The average internet user is younger, more interested in politics, and comes from a higher socioeconomic background than the average citizen, which raises concerns about external validity (Mislove, Lehmann, Ahn, Onnela, and Rosenquist, 2011; Gong, 2011). Furthermore, the voice of political minorities tend to be underrepresented in the public debate on Twitter, and differences in party strategies regarding their presence in social media can also bias any measure of public opinion that relies on the number and content of tweets. It is therefore necessary to obtain more background information about each individual user, so that it is possible to stratify them and weight public opinion estimates.

Extracting socioeconomic information in Twitter is a difficult task, because users are not even asked to provide their age or gender, as it is the case in other social networks. However, developing techniques to estimate social media users' individual attributes serves three important purposes. First, this type of information improves our understanding of the profile of who participates in online social networks and, most importantly, how representative of the entire population is a random sample of social media users. Second, individual-level data can be very valuable in the process of generating reliable public opinion estimates. For example, if we are interested in studying the Republican primary election, it would allow us to sample only supporters of this party and avoid simplifying assumptions (see King, Orlando, and Sparks, 2012). Future studies that aim at capturing public opinion trends using language processing techniques could use these data to stratify and weigh their estimates. Finally, ideology and other personal traits could be particularly useful as a covariate in future studies about online and offline political behavior. For instance, given the use of Twitter as a coordination mechanism in an era of new types of social protest, this variable might be useful to study how political action spreads across partisan networks. It could also prove to be useful in the study of party competition and electoral behavior, for it might provide ideal point estimates for legislators and ordinary citizens within different regions or states, which would improve the existing empirical tests of spatial voting models. These reasons justify the relevance of the method I present in this paper.

3 Ideal Point Estimation Using Twitter Data

3.1 Previous Studies

There is a limited but increasing literature on the measurement of users' attributes in social media, particularly in the field of computer science. Despite ideology[4] being one of the key predictors of political behavior online, their measurement through social media data has only been examined in a handful of studies.

These studies have relied on three different sources of information to infer Twitter users' ideology. First, Conover, Gonçalves, Ratkiewicz, Flammini, and Menczer (2010) focus on the structure of the conversation on Twitter: who replies to whom, and who retweets whose messages. Using a community detection algorithm, they find two segregated political communities in the US, which they identify as democrats and republicans. Second, Boutet, Kim, Yoneki, et al. (2012) argue that the number of tweets referring to a British political party sent by each user before the 2010 elections are a good predictor of his/her party identification. However, Pennacchiotti and Popescu (2011) and Al Zamal, Liu, and Ruths (2012) have found that the inference accuracy of these two sources of information is outperformed by a machine learning algorithm based on a user's social network properties. In particular, their results show that the network of friends (who each individual follows on Twitter) allows to infer political orientation even in the absence of any information about the user. Similarly, the only (to my knowledge) political science study that aims at measuring ideology (King, Orlando, and Sparks, 2011) uses this type of information. These authors apply a data-reduction technique to the

[4]Ideology is defined here as the main policy dimension that articulates political competition: "a line whose left end is understood to reflect an extremely liberal position and whose right end corresponds to extreme conservatism" (Bafumi, Gelman, Park, and Kaplan, 2005, p.171). Each individual's ideal point or policy preference corresponds to their position on this scale.

complete network of followers of the U.S. Congress, and find that their estimates of the ideology of its members are highly correlated with estimates based on roll-call votes.

From a theoretical perspective, the use of network properties to measure ideology has several advantages in comparison to the alternatives. Text-based measures are vulnerable to the phenomenon of 'content injection'. As Conover, Gonçalves, Ratkiewicz, Flammini, and Menczer (2010) show, hashtags are often used incorrectly for political reasons: "politically-motivated individuals often annotate content with hashtags whose primary audience would not likely choose to see such information ahead of time". This reduces the efficiency of this measure and results in bias if content injection is more frequent among one side of the political spectrum. Similarly, conversation analysis is sensitive to two common situations: the use of 'retweets' for ironic purposes, and '@-replies' whose purpose is to criticize or debate with another user. As a result, it is hard to characterize the emerging communities, and whether this divide overlaps with the ideological composition of the electorate, or even if it is stable over time.

In conclusion, a critical reading of the literature suggest the need to develop new, network-based measures of political orientation. It is also necessary to improve the existing statistical methods that have been applied. Pennacchiotti and Popescu (2011) and Al Zamal, Liu, and Ruths (2012) focus only on classifying users, but most Political Science applications require a continuous measure of ideology, for it is considered a latent variable that is scaled on a single dimension. In order to draw correct inferences, it is also important to indicate the uncertainty of the estimates. King, Orlando, and Sparks (2011), for example, do not provide standard errors for their measures of members of Congress' ideology. Without these, it is not possible to make inferences about their rank-ordering. Similarly, these authors do not explore the possibility of placing ordinary citizens and legislators on a common scale. The main contribution of this paper is thus to implement a method to provide reliable and valid estimates (and standard errors) of Twitter users' ideology on a continuous scale.

3.2 Implementing a Bayesian Spatial Following Model of Ideology

The purpose of this paper is to demonstrate that valid ideal point estimates of individual Twitter users *and* political actors with a Twitter account can be derived from the structure of the 'following' links across these two sets of users. In order to do so, I develop a Bayesian spatial model of Twitter users' following behavior. Ideology is defined as a position in a latent ideological space (Poole and Rosenthal, 1997, 2007), and individual estimates are derived on the basis of the observed 'following' decisions, under the assumption that Twitter users are instrumentally rational.

The key assumption of this model is that Twitter users prefer to follow politicians whose position on the latent ideological dimension are similar to theirs. There exists broad theoretical and empirical support for this notion. Firstly, the vast body of research about homophily in personal interactions can easily be extended to online social networks such as Twitter. As McPherson, Smith-Lovin, and Cook (2001) theorize, individuals tend to be embedded in homogenous networks with regard to many sociodemographic and behavioral traits, because of a dual process of creation and dissolution of social ties due to shared geographical, organizational, and symbolic spaces. A very small individual preference for interactions with similar people is required for segregated communities to emerge at the aggregate level (Schelling, 1978). Different studies have provided solid evidence of the existence of these patterns in Twitter (Gayo-Avello, 2010; Wu, Hofman, Mason, and Watts, 2011; Conover, Gonçalves, Flammini, and Menczer, 2012).

However, Twitter is not only an online social network – it is also a news media (Kwak, Lee, Park, and Moon, 2010). From this perspective, we can also rely on the existing literature on the selective exposure theory (Lazarsfeld, Berelson, and Gaudet, 1944; Bryant and Miron, 2004) to argue that Twitter users exhibit a preference for opinion-reinforcing political information and that they systematically avoid opinion challenges. Given the fast-paced nature of this social network and individuals' finite ability to process incoming information (Oken Hodas and Lerman, 2012), we should expect Twitter users to maximize the value of their online experience by choosing to follow political actors who can provide information that can be of higher value to them (Chan and Suen, 2008).

The theoretical model I employ relies on similar assumptions as the Euclidean spatial voting model (Enelow and Hinich, 1984). Suppose that each Twitter user $i \in \{1, \ldots, n\}$ is presented with a choice between following or not following another target user $j \in \{1, \ldots, m\}$, where j is a political actor who has a Twitter account[5]. Let $y_{ij} = 1$ if user i decides to follow user j, and $y_{ij} = 0$ otherwise. For the reasons explained above, I expect this decision to be a function of the squared Euclidean distance in the latent ideological dimension[6] between user i and j: $-\gamma||\theta_i - \phi_j||^2$, where $\theta_i \in \mathbb{R}$ is the ideal point of Twitter user i, $\phi_j \in \mathbb{R}$ is the ideal point of Twitter user j, and γ is a normalizing constant.

To this basic setup, I add two extra parameters, α_j and β_i. The former measures the popularity or "indegree" of user j. This parameter accounts for the fact that some political accounts are more likely to be followed, because of the role of the politicians

[5]If we considered not only politicians, but the entire Twitter network, then $n = m$. In that case, the model would still yield valid estimates, but the estimation would be computationally intractable and inefficient and, as I argue below, the resulting latent dimension might not be ideology. In this paper I show that it is possible to obtain valid ideal point estimates choosing a small m whose characteristics make 'following' decisions informative about the ideology of users i and j.

[6]I assume that ideology is unidimensional, which is a fairly standard assumption in the literature (e.g., see Poole and Rosenthal, 1997, 2007). However, the model I estimate could be easily generalized to multiple dimensions.

behind it (for example, we would expect the probability of following `@BarackObama` to be higher than the equivalent for a member of Congress, all else equal) or other reasons (politicians who 'tweet' more often are more likely to be highly visible and therefore also to have more followers, all else equal). The latter measures the level of political interest or "outdegree" of users i. Similarly, this parameter accounts for the differences in the number of political accounts each user i decides to follow, which could be due to the overall number of Twitter users he/she follows, or how interested in politics he/she is.

The probability that user i follows a political account j is thus formulated as a logit model:

$$P(y_{ij} = 1|\alpha_j, \beta_i, \gamma, \theta_i, \phi_j) = \text{logit}^{-1}\left(\alpha_j + \beta_i - \gamma||\theta_i - \phi_j||^2\right) \tag{1}$$

Estimation and inference for this type of model is not trivial. Maximum-likelihood estimation methods are usually intractable given the large number of parameters involved. For this reason, I implement a Bayesian method, where the posterior density of the five sets of parameters is explored via Markov Chain Monte Carlo methods and, more specifically, a Hamiltonian Monte Carlo algorithm. A detailed explanation of the model and the estimation algorithm is provided in section B.1 of the Appendix.

This approach presents three important advantages. Firstly, it allows the researcher to incorporate prior information into the estimation, based on observed data or relevant new information, but it is also possible to choose priors that reflect complete ignorance. Secondly, it provides proper estimates of the stochastic error at a low computational cost. As I show in Section 6, samples from the posterior distribution of the ideology parameters can be easily combined with simulated values of covariates to propagate uncertainty and obtain more accurate standard errors. Finally, although this option is not explored in the paper, this approach provides enough flexibility to easily extend the model to a hierarchical setup, where the estimated parameters are explicitly sampled as a function of covariates.

An important challenge regarding the implementation of method I propose is the choice of m, this is, those Twitter users with such "discriminatory" predictive power that the decision to follow them or not can provide information about an individual's ideology. Following Conover, Gonçalves, Ratkiewicz, Flammini, and Menczer (2010), we could analyze the entire networks of friendships in Twitter and let the different clusters emerge naturally, this is, without pre-imposing any structure or any reference point. However, this decision can violate the independence assumptions, as "homophilic" networks emerge based not only on political traits, but also as a result of similarities in other dimensions. Instead, the approach I suggest is to select a limited number of target users that includes politicians, think tanks, news outlets with a clear ideological profile, etc. Considering only those accounts that can be more informative about individuals' ideal points will ensure that the estimation is efficient, and that the latent dimension in which we are locating the ideal points is political ideology.

4 Data

The estimation method I propose in this paper can be applied to any country where a high number of citizens are discussing politics on Twitter[7]. However, in order to test the validity of the estimated parameters, I will focus on four countries where high-quality ideology measures are available for at least a subset of all Twitter users: the US, the UK, Spain, and the Netherlands. Furthermore, the increasing complexity of the party system in each of these countries will show how the method performs where more than two parties are present on Twitter.

For each of these countries, I identified a list of political representatives in national-level institutions, parties, and individuals with a highly political profile who are active on Twitter[8]. This represents a total of $m = 548$ target users in the US, $m = 244$ in the UK, $m = 298$ in Spain, and $m = 118$ in the Netherlands.

Next, using the Twitter REST API, I obtained the entire list of followers (as of November 4th, 2012) for all m users in each country, resulting in a entire universe of Twitter users following at least one politician of $n = 32{,}919{,}418$ in the US, $n = 2{,}647{,}413$ in the UK, $n = 1{,}059{,}890$ in Spain, and $n = 856{,}201$ in the Netherlands. However, an extremely high proportion of these users are either inactive, spam bots or reside in different countries[9]. To overcome this problem, I extracted the available personal attributes from each user's profile, and discarded from the sample those who 1) have sent less than 100 tweets, 2) have not sent one tweet in the past six months, 3) have less than 25 followers, 4) are located outside the borders of the country of interest, and 5) follow less than three political Twitter accounts. The final sample size is $n = 473{,}640$ users in the US[10], $n = 135{,}015$ in the UK, $n = 123{,}846$ in Spain, and $n = 96{,}624$ in the Netherlands.

[7] Estimating ideal points using data from different countries simultaneously is more complex, given the high intra-country locality effect (Gonzalez, Cuevas, Cuevas, and Guerrero, 2011).

[8] These lists combine information from different sources. In the US, I have used the NY Times Congress API, the Sunlight Labs Congress API and the GovTwit directory. In the UK, I have used the Twitter lists compiled by Tweetminster. In Spain, I have used the Spanish Congress Widget developed by Antonio Gutierrez-Rubi, and the website politweets.es. In the Netherlands, I have used the data set of politiekentwitter.nl. I considered only political Twitter users with more than 5,000 (US) or 2,000 (UK, Spain, Netherlands) followers.

[9] For example, in my analysis I found that only around 56% of Barack Obama's 23 million followers as of December 2012 are located in the United States.

[10] The actual sample size in the US is 301,537 users, which is the number of Twitter accounts who tweeted at least therefore times mentioning 'Obama' or 'Romney' and can thus be included in the analysis in section 6.

Of course, this is a highly self-selected sample. Twitter users are not a representative sample of the population: they tend to be younger and to have a higher income level than the average citizen, and their educational background and racial composition is different than of the entire country (Mislove, Lehmann, Ahn, Onnela, and Rosenquist, 2011; Parmelee and Bichard, 2011). In the context of this paper, the inferences I make based on my sample won't even be valid for the entire universe of Twitter users, since I am only selecting those who follow a certain number of political accounts. However, this should not affect the inference of *individual* ideal points, because these users can indeed be considered as "authoritative" when it comes to politics. Even if they are more knowledgeable and interested in politics than the average citizen, this set of users are at the core of most exchanges about current events that take place over Twitter, and play an important role in setting public opinion. As a consequence, examining their online behavior can be highly informative about how information spreads through social networking sites. Furthermore, if we desire to extend our results to the entire population, we can estimate individual characteristics of each user and then use these to calibrate and weight our sample.

The sample selection process requires identifying the specific country from where each user tweets. This information was extracted from the "location" field in the user profile, and structured using the Yahoo geolocation API[11]. Interestingly, the geographical distribution of Twitter users resembles that of the general population. For example, in Figure 1 I show how the percentage of Twitter accounts located in each state approximates the proportion of US citizens living in that state.

In the case of the US, I also imputed the gender of all Twitter users in my sample. This variable was estimated using the full name of each Twitter user (when it was provided), and applying a probabilistic model that relies on the list of most common first names by gender available in the US Social Security Administration, and other anonymized databases where gender is known, available in the RandomNames `R` package. The table below provides descriptive statistics for this variable. A more detailed explanation of the imputation method I use can be found in Appendix C.

Two of the substantive applications I present in section 6 analyze the structure and content of the political conversation in Twitter during the 2012 US Presidential election campaign. The data I use consists of all public tweets that mentioned 'Obama' or

[11] Note that location information is only available for the subset of users who decide to provide it on their profiles, which is around 70% (Hale, Gaffney, and Graham, 2012). Even if they decide to provide such information, it is highly unstructured and sometimes states only the country but not the region or city, or it refers to imaginary places (Hecht, Hong, Suh, and Chi, 2011). However, in combination with the information about the time zone in each user's profile, it is sufficient to identify the country of residence in 90% of the cases. This proportion is lower when we consider more specific geographical levels, such as state in the US (71%), country in the UK (67%), or province in Spain (61%) and the Netherlands (64%).

Figure 1: Distribution of Twitter Users in the Continental US, by State

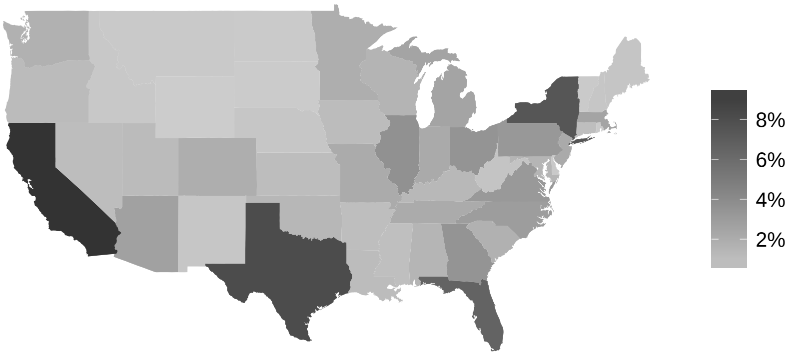


Table 1: Distribution of Twitter Users, by Gender

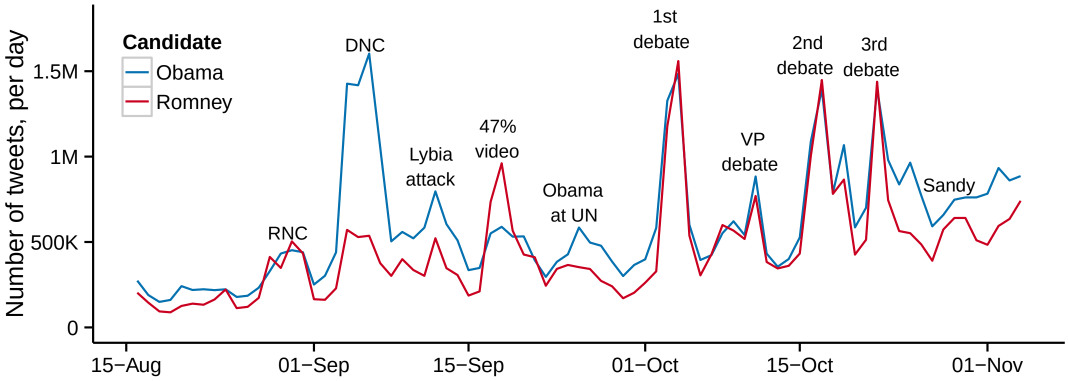
	Male	Female	Unknown	Total
Number of Twitter Users	151,497 (50.2%)	105,811 (35.1%)	44,229 (14.7%)	301,537 (100%)

'Romney' from August 15th to November 4th. These messages total over 75 million tweets (15 million of which were published by the 301,537 users in my sample – an average of 50 tweets per person) and were collected using the Twitter Streaming API and stored on a MySQL server. (All relevant source code is available upon request.) Figure 2 plots the evolution in the daily number of tweets sent over the course of the electoral campaign. As expected, this metric peaks during significant political events, such as the party conventions or the three presidential debates.

Finally, in order to improve the validation of the ideal points I estimate in the US, a sample of Twiter users from the state of Ohio were matched with their voting registration records. This choice was based on it being considered one of the prime 'battleground' states, and also for reasons of data availability. (The entire voter file is available online at the Ohio Secretary of State website.) A total of 2,462 Twitter users were matched with their individual records, based on perfect matches of their reported full name and county of residence, which represents over 12% of the 20,153 Twitter users from Ohio in the full sample[12]. Again, this subset cannot be considered representative for any

[12]This proportion is comparatively not too small, particularly if we consider that most Twitter users do not provide their real full name. The other existing study that performed a similar analysis (Bond, Fariss, Jones, Kramer, Marlow, Settle, and Fowler, 2012) was only able to successfully match 1 in 3 Facebook users to voter records.

Figure 2: Evolution of mentions to Obama and Romney in Twitter



population of interest, and will only be used to examine whether the resulting ideology estimates are good predictors of the party under which they are registered.

5 Results

In this section I provide a summary of the ideology estimates for the four countries included in my study, and the oversampled set of Twitter users in the state of Ohio. To validate the method, I will use different sources of external information to assess whether this procedure is able to correctly classify and scale Twitter users on the left or right side of the ideological dimension.

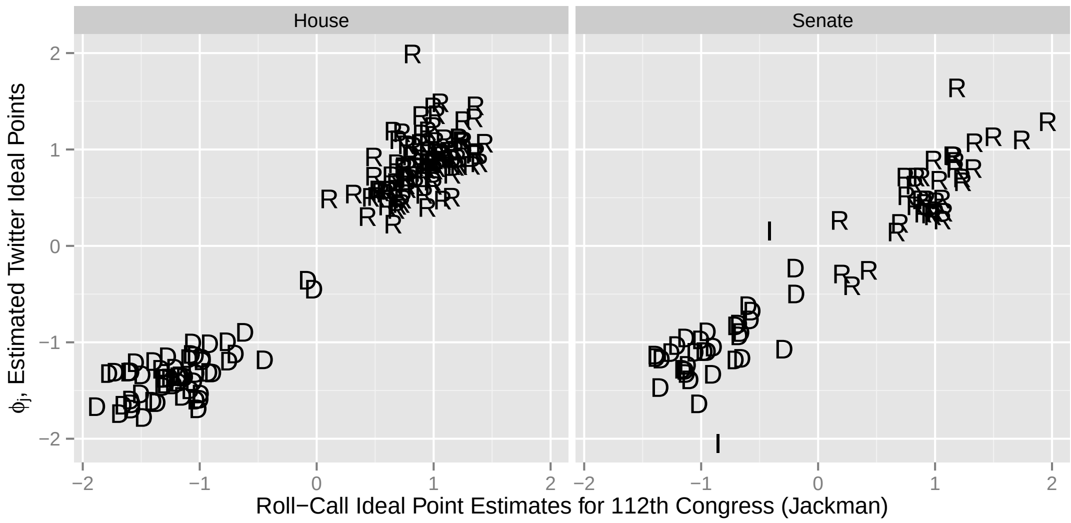
5.1 United States

The first set of results I focus on are those from the United States. Figure 3 compares ϕ_j, the ideal point estimates, of 219 members of the 112th U.S. Congress[13] based on their Twitter network of followers (y axis) and on their roll-call voting records[14] (x axis). Each letter correspond to a different member of congress, where D stands for democrats and R stands for republicans, and the two panels split the sample according to the chamber of Congress to which they were elected.

[13] Only members of congress whose Twitter acounts have more than 5,000 followers are included in the sample.

[14] The source for the ideal points based on voting records is Jackman (2012), and were estimated using the item-response model described in Clinton, Jackman, and Rivers (2004).

Figure 3: Comparing Ideal Points Based on Roll-Call Records and Based on Twitter Network of Followers in the U.S. Congress



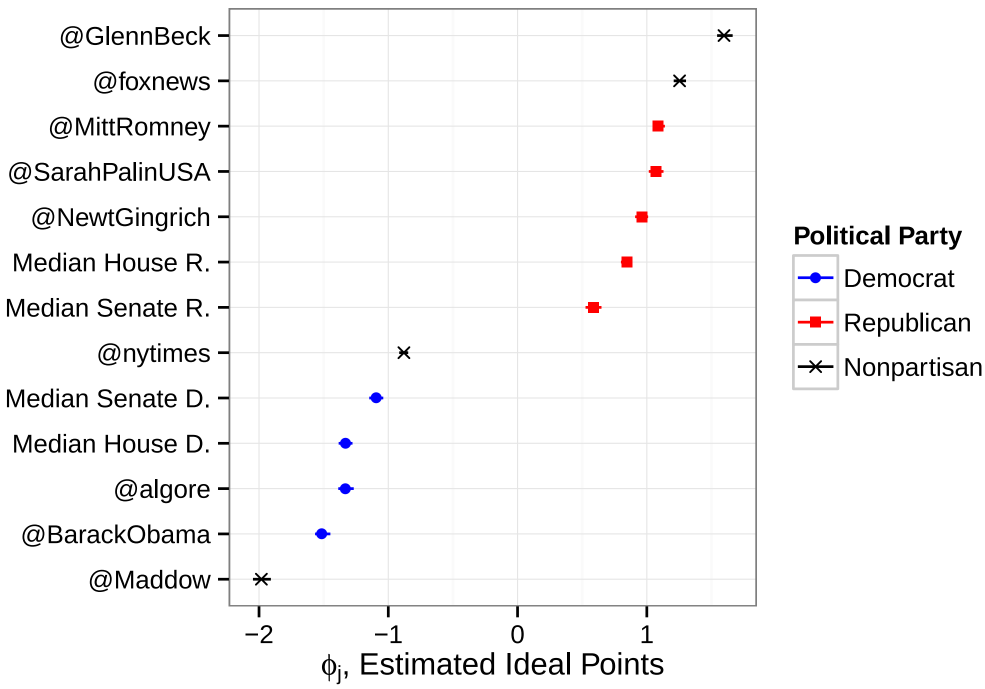
As we can see, the estimated ideal points are clustered in two different groups, which align almost perfectly with party membership. The correlation between Twitter- and roll-call-based ideal points is $\rho = .966$ in the House and $\rho = .950$ in the Senate. Furthermore, if we examine the most extreme legislators, we find that their Twitter-based estimates also position them among those with the highest and lowest values on the ideological scale. Within-party correlations are also relatively high: $\rho = .538$ for republicans, $\rho = .749$ for democrats.

While the method I propose is able to classify each member of congress by party – note that party information was not included at any point in the estimation –, its success at scaling them along the ideological dimension is still limited. Twitter-based ideal points cannot be considered an alternative to estimates based on roll-call data, but their level of accuracy shows that they can be informative about their ideology, at a low cost.

Ideal points for a wider set of political actors are plotted in Figure 4. (See Appendix A for an expanded version of this plot.) As it was the case with members of congress, the resulting estimates show a clear division across members of each party, and all non-partisan actors whose ideology is known are correctly classified. Note also their positions within each cluster are also what we would expect based on anecdotal evidence. For example, Schwarzenegger and Jon Huntsman are classified as the two most liberal

Twitter accounts in the Republican Party, while Limbaugh and Glenn Beck are the most conservative among the nonpartisan Twitter accounts. On the left side of the ideological dimension we find Olbermann, Michael Moore, Rachel Maddow or the HRC as the most liberal Twitter acounts.

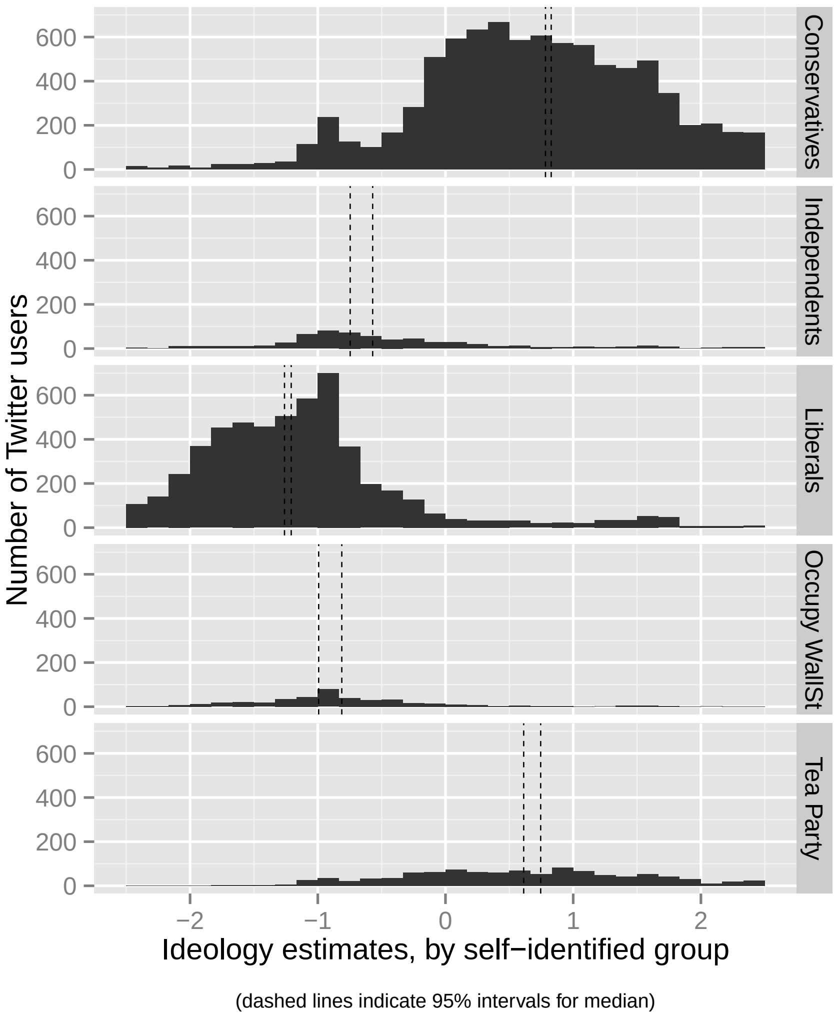
Figure 4: Estimated Ideal Points for Key Political Actors



Now I turn to assess whether the estimated ideal points for ordinary citizens are also valid. In Figure 5 I plot the distribution of the ideology estimates for different groups of individuals. Here I exploit the fact that many Twitter users define themselves politically in their profiles[15]. Using this information, I extracted five subsets of accounts, according to whether they mention specific keywords on their profiles: conservatives ("conservative", "GOP", "republican"), independents ("independent"), liberals ("liberal", "progressive", "democrat"), Tea Party members ("tea party", "constitution"), and Occupy Wall Street members ("occupy", "ows"). The dashed line in each panel indicates the ideology of the median member of each subset.

[15]Three different examples of profiles that can be used to identify ideology would be: "Student of History and Politics. Christian and **Conservative**. Southern and Saved [...]", "Idaho native. Oregon **democrat**. Fly Fisherwoman. Political Nerd [...]", and "reader, citizen patriot, concerned, recently changed registration to **Independent**".

Figure 5: Distribution of Ideal Point Estimates, by Self-Identified Political Group



The distribution of ideal points for each group closely resembles what one would expect: conservatives are located to the right of independents, and independents are located to the right of liberals. Similarly, supporters of the Occupy Wall Street movement tend to be more liberal than Tea Party members, although they appear to be more centrist than the median conservative or liberal Twitter user, which suggests that this scale could be capturing the intensity of party support rather than strictly ideology. While classification is not completely perfect, this plot shows that the estimation is able to distinguish and scale Twitter accounts according to the policy position of who updates them.

Figures 19 and 20 in Appendix A show the distribution of ideal points by two other important individual characteristics: gender and language[16]. These figures validate what can be found in political surveys: women tend to be more liberal than men, and Spanish-speaking citizens are more liberal than the average English-speaking voter.

5.2 Ohio

To further validate the ideal point estimates I introduced in the previous section, now I turn to examine the results from the sample of 2,360 Twitter users from Ohio whose names were matched with the voter file.

In Figure 6, I plot the distribution of the ideology estimates across three different groups of voters, based in their most recent party registration in the period 2008–2012: not registered, registered as democratic, registered as republican. (Note that this variable is available because being registered with a party is a necessary condition in order to vote in the primary elections.)

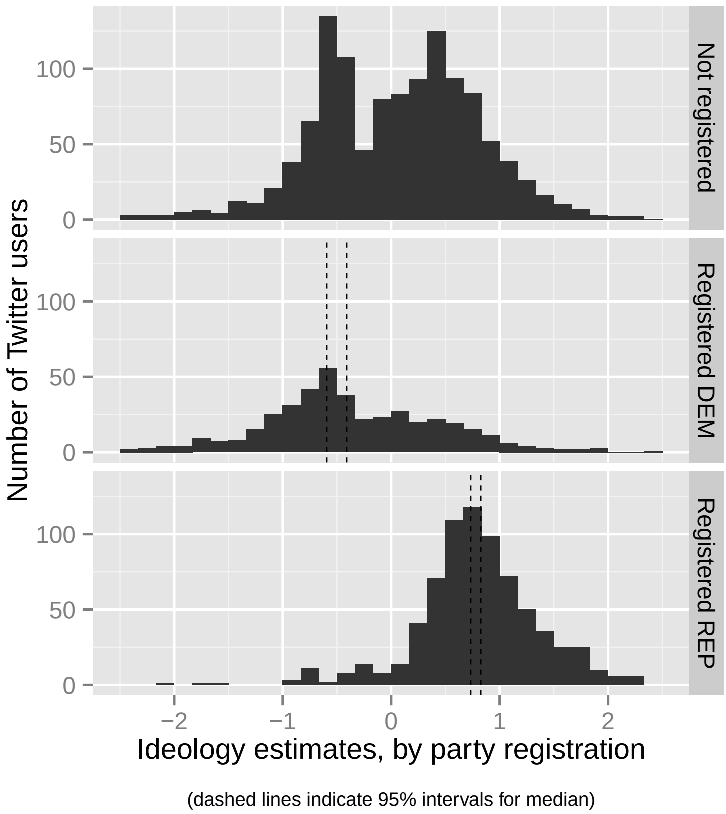
The evidence I present provides additional support for the external validity of my method: the average registered republican is located to the right of the average democrat, and this difference is large and statistically significant. If we consider the distribution of ideal points across these two groups, we see that most individuals are correctly classified to the left or right based on their party registration, with some overlap, particularly in the case of liberal voters. On the contrary, the ideal points of citizens who are not registered with any party have a distribution very similar to the overall US population (see Figure 19).

Having access to the voter file also allows the use of two additional sources of information about Twitter user: their year of birth and their registration address. This second variable, in turn, can be used to approximate each Twitter user's income based on their home values. (I obtain these estimates from the Yahoo! Homes website, which aggregates home valuations from several websites, such as Zillow or eppraisal. These values were available for around 78% of the residential addresses in my dataset.)

In Figure 7 I examine how each of these variables correlate with ideology. Each dot is labeled according to the party of registration for each voter. The left panel shows that age and estimated ideology are positively correlated, such that older Twitter users tend to be more conservative. This relationship, albeit being statistically significant at the 5% level, has a small magnitude. Not surprisingly, I also find that the median matched

[16] Language refers to the language in which users have set up their interface when they access Twitter. Only English and Spanish are considered, given that the sample size for other languages is very small.

Figure 6: Distribution of Ideal Point Estimates, by Party of Registration

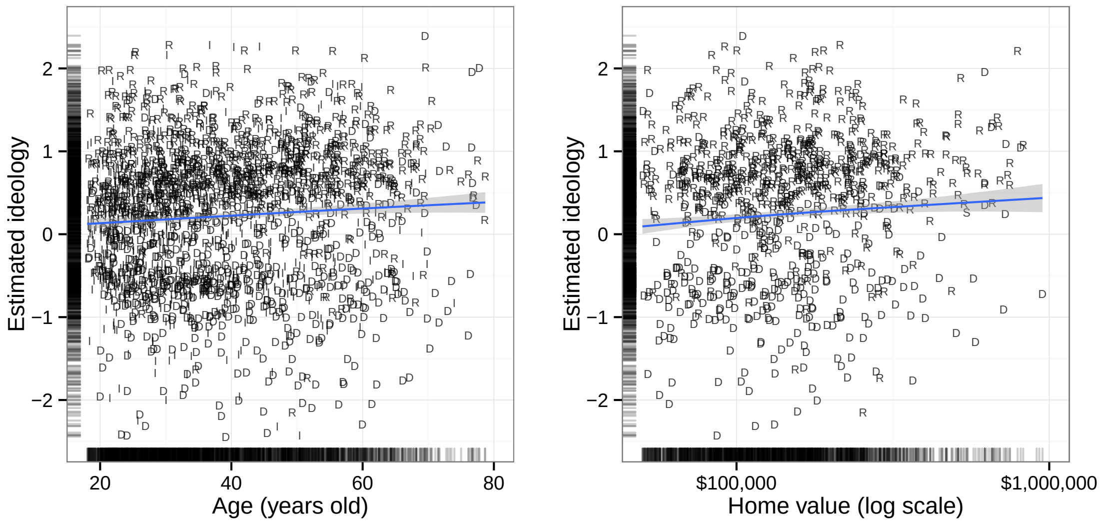


voter is young, around 35 years old. Home values and ideology of their residents are slightly more correlated, as shown in the right panel, with owners or tenants of more expensive houses being more conservative on average. However, the size of this effect is still small. After controlling for age, a change of one unit on the log scale (from \$100,000 to \$1,000,000) is associated with an expected change of 0.15 units on the ideological scale. (This effect is significant at the 0.1% level.)

5.3 UK, Spain, and Netherlands

The next three figures refer to the other countries I included in the study, and display credible intervals for the location of the average follower of the party and leader accounts for each of the four largest national parties. In each figure, I also show the ideological locations of each party on the left-right scale, estimated using expert surveys (Benoit and

Figure 7: Age and Home Values Are Correlated with Estimated Ideology



Laver, 2006). (Ideology estimates based on party manifestos (Budge, 2001) are similar, and therefore omitted for reasons of space.)

As in case of the US, these results show that my estimation method is able to classify accounts according to the party to which they belong. With few exceptions, all pairs of Twitter accounts from the same party are clustered together. Furthermore, the order of the parties seems to be similar to that reported by different studies based on expert surveys (Benoit and Laver, 2006) for the "taxes vs. spending" dimension. In Spain, the Communist party (IU) is placed to the left of the Socialist party (PSOE), with the Conservative party (PP) on the right, and the recently-created Center party (UPyD) located inbetween. In the Netherlands, the Labour party (PvDA) and the Socialist party (SP) are located to the left of the Christian-democratic party (CDA) and the Liberal party (VVD). In contrast, classification is less accurate in the UK: the Labour Party is to the left of the Conservative Party, but the Liberal-Democrats and the Scottish SNP are located in the ideological center[17].

However, note that the distances between parties do not match in some cases what we would expect based on an analysis of their platforms. For example, in Spain the Communist party (IU) is significantly more left-wing than the Socialist party (PSOE), but in my analysis I found them to have nearly identical positions. Finally, it is also

[17] One possible explanation for the poor results could be that party competition in the UK is multidimensional, which the model does not capture, since it assumes a single ideological dimension.

Figure 8: Ideological Location of Parties in Spain

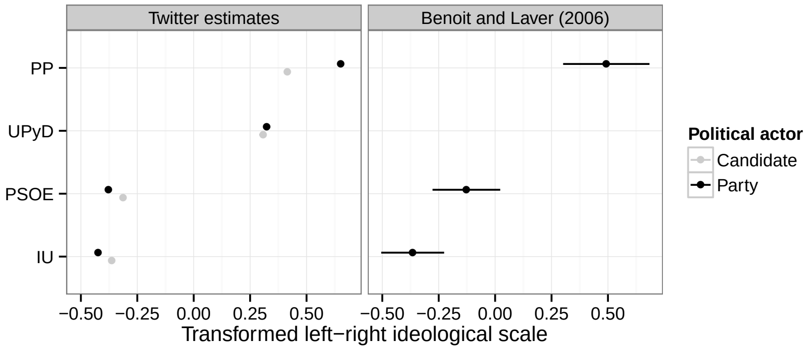
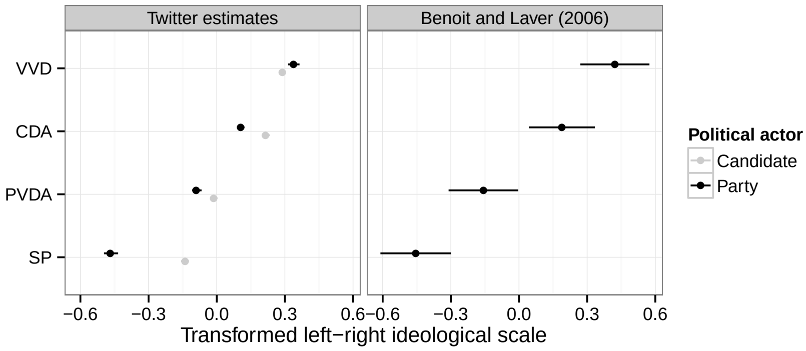


Figure 9: Ideological Location of Parties in the Netherlands

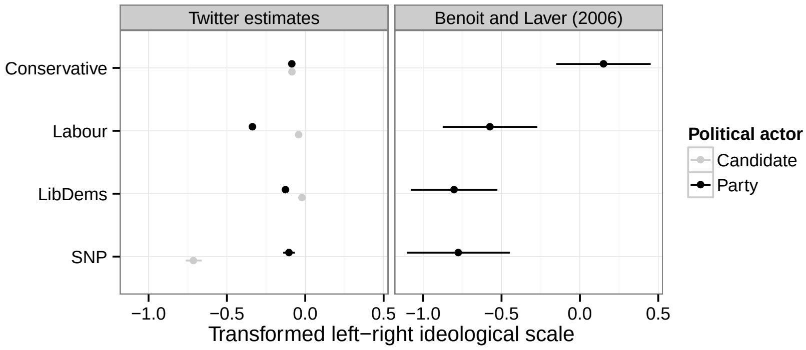


important to note that the results for different countries are not directly comparable, as the estimation was performed independently, and the resulting dimension does not have a homogenous scale across countries.

6 Applications

In this section I present three potential applications of the ideal point estimates obtained using Twitter data, which illustrate different research questions that could be answered using my method. First, I exploit the convenience of MCMC methods to estimate the ideal point of the average Twitter user in each state via simulation. These results help us understand how the sample of Twitter users differs from the entire population. Second, I

Figure 10: Ideological Location of Parties in the United Kingdom



address a heated debate in the literature: does social media increase polarization? Using the 2012 US presidential election campaign as a case study, I find that public exchanges on Twitter take place predominantly among users with similar viewpoints. And third, I apply a sentiment analysis technique to measure how the polarity of the tweets sent by the panel of Twitter users I consider in this paper is a good predictor of the overall evaluation of the two candidates.

6.1 Estimating the Location of the Average Voter

One of the advantages of Bayesian computation is that it allows us to easily combine uncertainty at each state of the estimation when obtaining the density of the posterior distribution of the parameters of interest. I illustrate this approach by estimating the ideology of the median Twitter user in each state, after stratifying by gender, and therefore combining the uncertainty about the ideal points for each individual and their imputed gender (see Appendix C for more details on the imputation method).

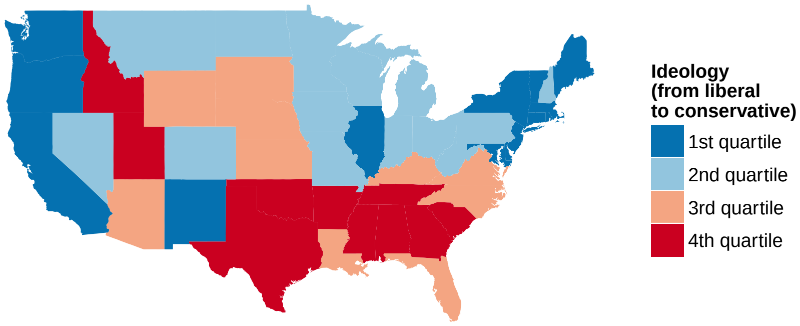
The results of this method are shown in in Figure 11, where the shade of the color indicates the quartile of the distribution of posterior ideology for the (weighted) average Twitter user in each state of the US ($\bar{\theta}_j$). Darker colors correspond to 'blue Twitter states' and lighter colors correspond to 'red Twitter states'.

The estimate for each state j has been computed as follows:

$$\bar{\theta}_j | y_j = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{N_j} \sum_{i=1}^{N_j} w_{ij} \bar{\theta}_{ijt}$$

where T is the number of iterations of the MCMC chain, N_j is the number of Twitter users in each state, $\bar{\theta}_{it}$ is the average of the posterior distribution of the ideal points for user i in state j in iteration t, and w_i is the post-stratification weight given to each user based on their gender. This latter variable is thus computed as the quotient of the proportion of users in the population for each gender group (using data from the 2010 US Census) over the proportion of the same group in the sample of Twitter users.

Figure 11: Ideal Point of the Average Twitter User in the Continental US, by State

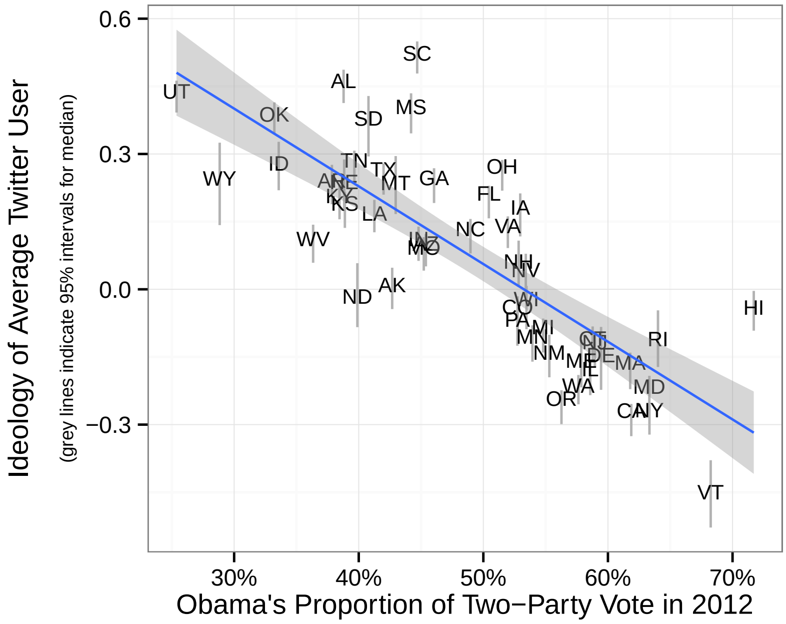


Despite Twitter users being a highly self-selected sample of the population, this figure nonetheless presents a close resemblance to ideology estimates based on surveys (Holbrook-Provow and Poe, 1987; Carsey and Harden, 2010). Furthermore, state ideal point estimates are highly correlated with electoral results. As I show in Figure 12, the states where the average Twitter user is more liberal supported Barack Obama in greater proportion in 2012. The correlation between these two variables is $\rho = .81$. (This correlation increases to $\rho = .83$ if we exclude Hawaii.)

6.2 Social Media and Political Polarization: Echo Chamber or Pluralist Debate?

A recurring theme in the literature on internet and politics is how the increasing amount and heterogeneity of political information citizens have access to affects their political views (Farrell, 2012). Several authors argue that, as a result of this transformation, individuals are being increasingly exposed to only information that reinforces their existing views, thus avoiding challenging opinions (Sunstein, 2001; Garrett, 2009). This generates a so-called echo-chamber environment (Adamic and Glance, 2005) that fosters social extremism and political polarization. Given that a substantial proportion of citizens now

Figure 12: Twitter-Based Ideal Points and Obama Vote Share in 2008, by State



rely mostly on the internet to gather political information[18], the policy implications of this issue are obvious: how individuals gather political information affects the quality of political representation, the policy-making process, and the stability of the democratic system (Mutz, 2002).

In the specific context of Twitter, this issue is also relevant because the extent to which users' behavior on this platform is polarized remains an open debate. On one hand, Conover, Gonçalves, Ratkiewicz, Flammini, and Menczer (2010); Conover, Ratkiewicz, Francisco, Goncalves, Flammini, and Menczer (2011); Conover, Gonçalves, Flammini, and Menczer (2012) find high levels of clustering along party lines: "the network of political retweets exhibits a highly segregated partisan structure, with extremely limited connectivity between left- and right-leaning users" (Conover, Ratkiewicz, Francisco, Goncalves, Flammini, and Menczer, 2011, p.89). Yardi and Boyd (2010) and Gruzd (2012) qualify this conclusion. While they also find that Twitter users tend to cluster around shared political views, their results show that open cross-ideological exchanges are very frequent, and individuals are exposed to broader viewpoints. Similarly, when examining other types of behavior on Twitter, Conover, Ratkiewicz, Francisco, Goncalves,

[18] According to a survey conducted by the Pew Research Center in 2001, 31% of U.S. adults rely most on internet for political information.

Flammini, and Menczer (2011) also find that user-to-user interactions via "@-replies" between ideologically-opposed individuals take place at a higher rate compared to the network of retweets.

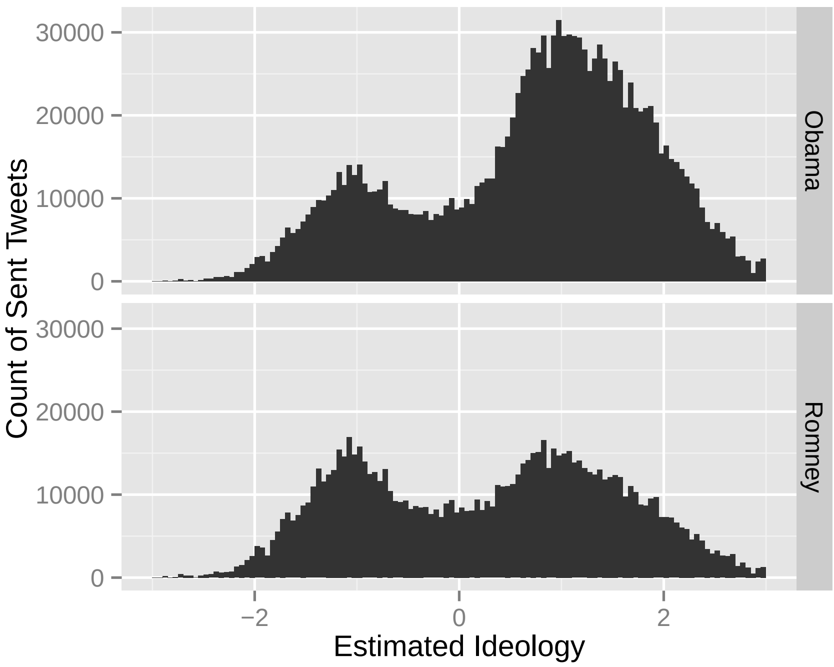
One possible reason for the variability in the results of these studies is that they rely on similar sources of information in their measurement of both ideology and polarization. For example, Conover, Gonçalves, Ratkiewicz, Flammini, and Menczer (2010); Conover, Ratkiewicz, Francisco, Goncalves, Flammini, and Menczer (2011); Conover, Gonçalves, Flammini, and Menczer (2012) use network clustering algorithms to classify users by their tweeting behavior, and then see to what extent users that belong to the same group interact with each other. The problem with this approach is that these algorithms are trained precisely to maximize the distance between individuals across different communities, and are thus biased towards finding polarized networks.

As a second application of the estimation method I propose in this paper, I replicate the analysis of this set of studies. In contrast with their approach, I use two completely different sources of data to measure ideology and users' behavior of U.S. Twitter accounts. As it was explained in section 3, my ideal point estimates are based on the 'following' connections established between users. In parallel, I have captured all tweets mentioning any of the two presidential candidates ("Obama" or "Romney") from August 15th, 2012, to November 6th, 2012, and selected those (nearly 20%) that were sent by Twitter users in my sample of over 300,000 accounts. This dataset will allow me to measure to what extent political conversations on Twitter are polarized along ideological lines.

I show results of my analysis in Figures 13, 14 and 15. The first figure plots the number of tweets published in the interval of study by users along the latent ideological dimension (in bins of width 0.05). The top panel refers to tweets that mention "Obama", while the bottom panel refers to tweets mentioning "Romney". The pattern that emerges yields two results. First, I find that the conversation in Twitter is dominated by individuals with extreme views. Despite the fact that (by construction) ideology has a unit variance distribution, we find that the distribution of the number of tweets is highly bimodal, with the modes at approximately -1 and $+1$ – this is, one standard deviation away from the average Twitter user. Second, I find a very distinct pattern in the tweets mentioning President Obama: conservative Twitter users sent a substantively higher proportion of tweets than their liberal counterparts. This finding is consistent with the results of the analysis by Conover, Gonçalves, Flammini, and Menczer (2012), who also discovered that right-leaning Twitter users exhibit greater levels of political activity.

Figures 14 and 15 provide additional evidence of the previous finding. Here, I use a heat plot to visualize the structure of the two most common types of interactions in

Figure 13: Number of Tweets Mentioning Presidential Candidates, by Ideal Point Bin



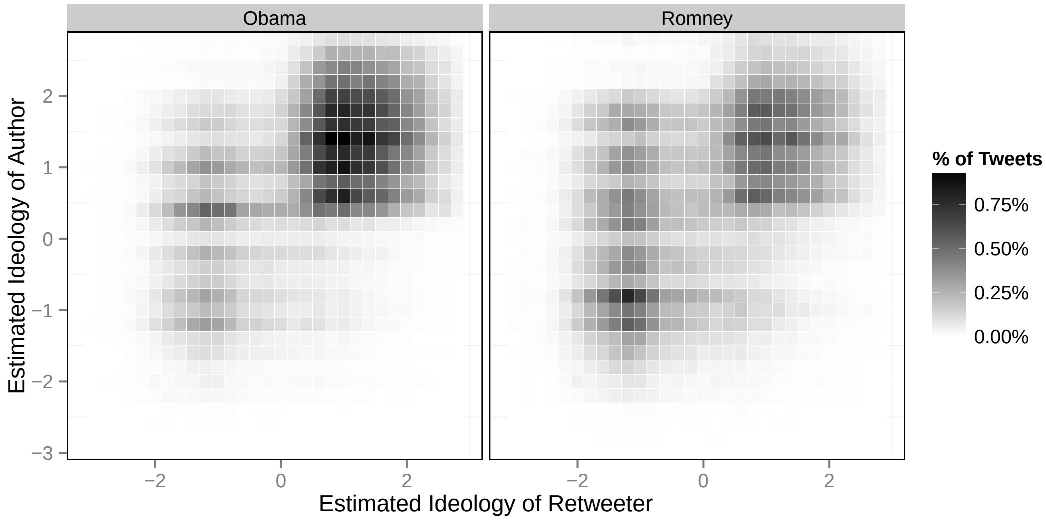
Twitter: retweets and mentions[19]. The color of each cell (of size 0.2×0.2) represents the proportion of tweets in the sample that were retweets/mentions of users with ideal point X to users with ideal point Y[20]. Therefore, if we were to find perfect polarization (this is, users interacting only with those of identical ideology), we would find a pattern that would resemble a line with slope one.

Both figures show very similar results. On one hand, right-leaning users appear to be more engaged in conversations, but their interactions tend to take place only among them. If we focus on the right-hand side of each panel, we can see that very few interactions originating from conservative users are addressed to liberal users. On the other hand, liberal users tend to engage in conversations all along the ideological spectrum. To

[19] A retweet consists on re-posting another user's content with an indication of its original author. It is used whenever the 'retweeter' wants to publicize the content of the original post, but it is not necessarily a sign of endorsement. In politics, candidates often encourage their followers to retweet their messages. A mention consists on including in a tweet the handle of another user (e.g. "@BarackObama"), so that the user that is mentioned can easily find the tweet. It is therefore an indication of a conversation between two Twitter users.

[20] For example, in the left panel of Figure 14 we can see that around 1% of all retweets mentioning Obama had an original author a Twitter user whose ideal point was in the interval between 1 and 1.2, and were retweeted by Twitter users in the same interval.

Figure 14: Political Polarization in Retweets Mentioning Presidential Candidates

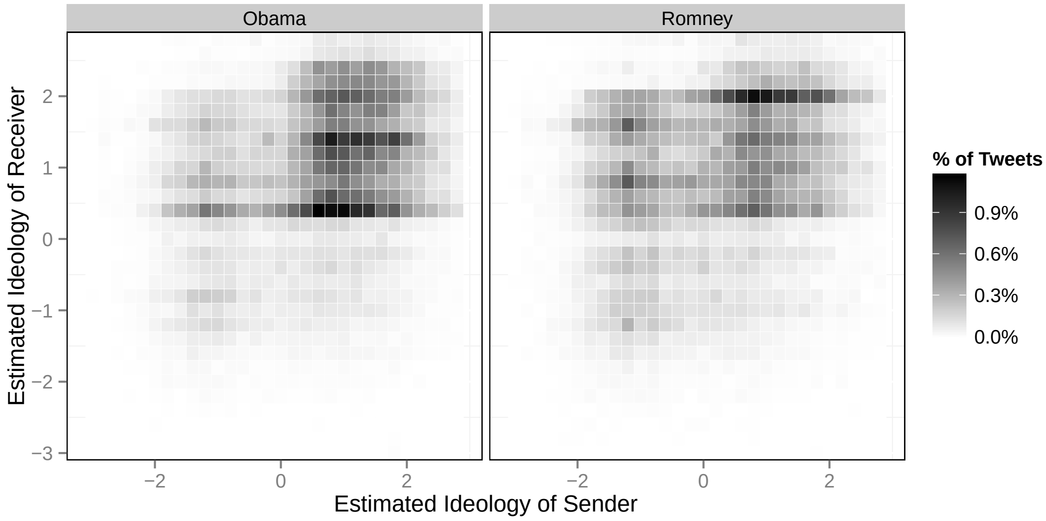


sum up, the picture that emerges points towards a certain degree of polarization, which is driven predominantly by conservative Twitter users. The strength of their ties has important implications during the electoral campaign. As Conover, Gonçalves, Flammini, and Menczer (2012) argue, the topology of the network of right-leaning Twitter users facilitate the rapid and broad dissemination of political information. In a context in which interactions taking place through this platforms are increasingly covered in the traditional media, the cohesiveness of this group of users has the potential to manipulate the public agenda.

6.3 Can We Measure Public Opinion Using Twitter Data?

The third application introduces an additional methodological tool that will be used to study the content of the tweets sent by ordinary citizens with different ideal point estimates: a sentiment analysis technique (Pak and Paroubek, 2010). This method consists of two steps. First, each tweet is decomposed in words, and each word is classified as positive, negative or neutral using a dictionary of positive and negative words (Riloff and Wiebe, 2003). The composition of each tweet determines whether the entire message is classified as positive or negative. Secondly, for each user's tweets (or a subset), I compute a daily "thermometer" or "affective score", which ranges from 0 (negative) to 100 (positive), measuring the proportion of positive tweets over the sum of positive and negative tweets. These individual scores are averaged using weights to

Figure 15: Ideological Polarization in Conversations Mentioning Presidential Candidates



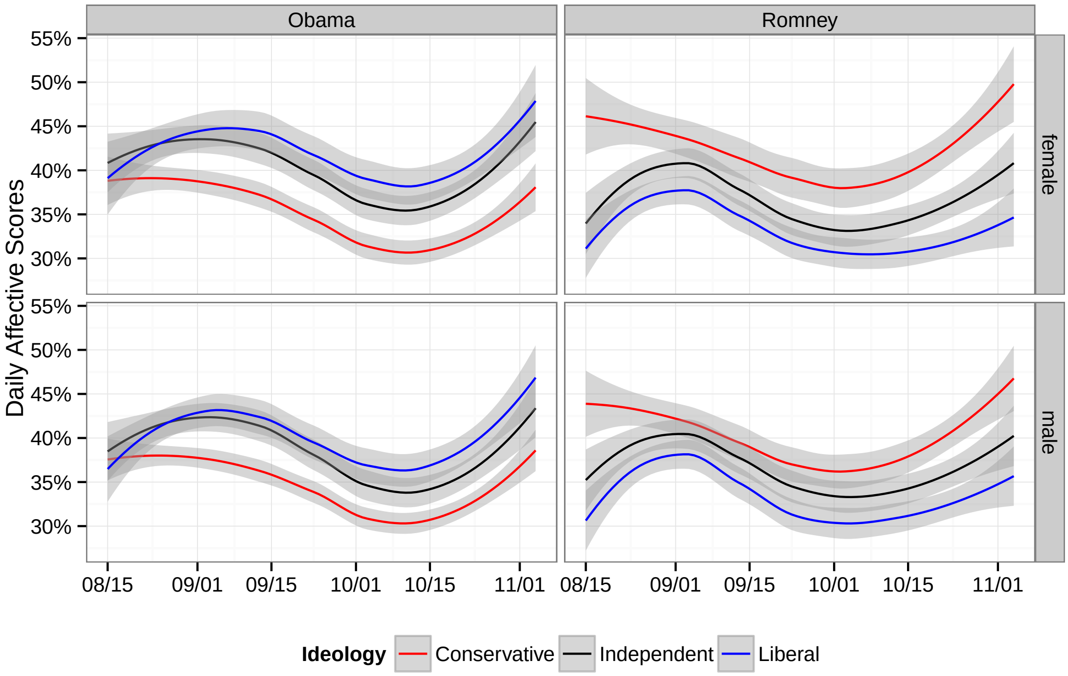
stratify by gender and location in order to compute daily aggregate scores. When we restrict the estimation to only those tweets mentioning a particular political actor, it is possible to construct an approximation of how that actor is evaluated by the panel of Twitter users I consider.

This technique is currently being used by companies that perform analysis of social media, and different academic papers have also applied similar methods to measure public opinion (see for example O'Connor, Balasubramanyan, Routledge, and Smith, 2010; Bollen, Pepe, and Mao, 2010; Rao and Srivastava, 2012). Even Twitter as a company has released a similar tool (the 'Twindex') to track the emotional content of tweets about Obama and Romney.

My approach improves the existing methods in three different ways. First, by tracking a panel of Twitter users carefully chosen to be 'opinion leaders' in politics, I make sure that the specific users included in the sample at each time are the same, and also that their tweets have the potential to be influential and affect offline public opinion. Second, I introduce weights by socioeconomic characteristics, in order to partially correct for the lack of representativeness of the sample of Twitter users. Finally, by splitting the sample across ideological groups, I am able to examine with greater accuracy public opinion trends and, in particular, how independent or centrist Twitter users change their evaluation of Obama and Romney over time.

I show preliminary results of this analysis in Figures 21 and 22 in Appendix A. Both plots show that the affective score is a noisy measure, with what appears to be lot of random variation from day to day. For this reason, in the text of the article I focus on Figures 16 and 17, where I apply a non-parametric loess smooth so that the underlying patterns are easier to observe. Despite the noise, both plots show very clearly that the average affective scores for Obama are higher among liberal Twitter users (those with $\hat{\theta} < -0.33$), while Romney is evaluated more favorably by conservative Twitter users (with $\hat{\theta} > 0.33$). Centrist (or independent) users tend to express opinions about Obama and Romney that are similar to those of liberal users, which is perhaps not surprising given the outcome of the election. These differences are consistent over time.

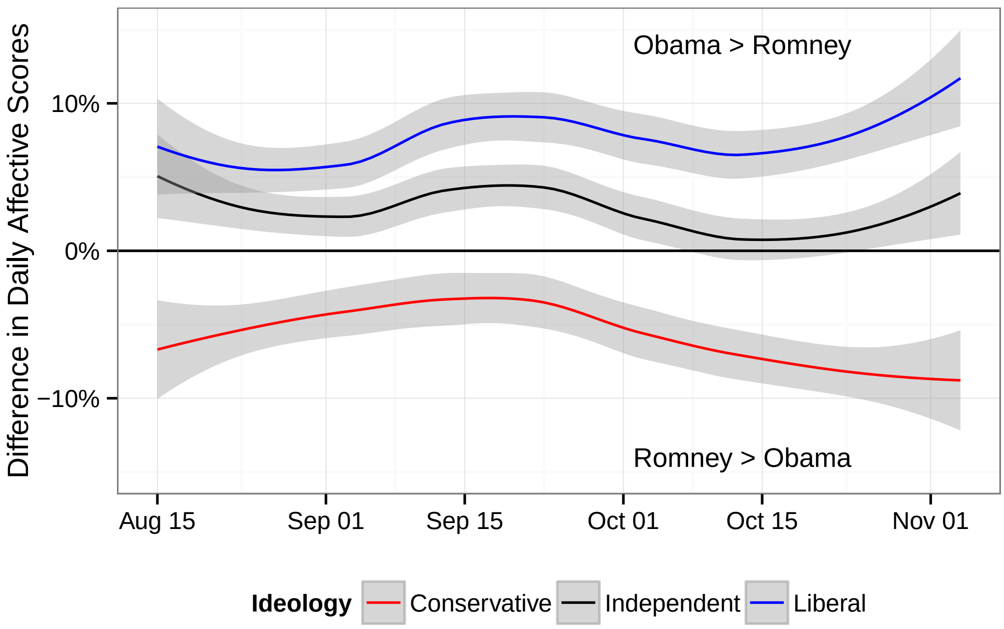
Figure 16: Affective Scores of Tweets Mentioning Obama or Romney, by Gender and Ideology of Twitter Users (with Loess smoothing)



On the other hand, when distinguishing Twiter users by gender, I find that women tend to have a better evaluation of Obama than men, and the opposite for Romney, although these differences are not as large as those related to ideology. This difference makes it thus necessary to weight by gender when computing an average score, as I do in Figure 17, where female Twitter users are weighted so that they represent a similar proportion to that of male Twitter users.

The evolution of the affective scores over time also seem to respond to relevant cam-

Figure 17: Daily Differences in Weighted Affective Scores of Tweets Mentioning Obama or Romney, by Ideology of Twitter Users (with Loess smoothing)



paign events. Both candidates improved their evaluations during their party conventions (late August for Romney and early September for Obama). These 'convention bumps' are visible across all ideological groups. Affective scores also decrease significantly for both candidates right after September 11, particularly for Mitt Romney, perhaps due to their reaction to attack on the American embassy in Lybia. Finally, Romney does not appear to see his evaluation worsen when his "47% remarks" were released, on September 17; but Obama does improve his affective score from this moment on. Finally, it is interesting to note how the gap between Romney and Obama widens after the first debate, but only among conservative Twitter users[21], and how opinions about candidates become even more polarized by ideology during the last two weeks of the campaign.

[21] However, note that their absolute affective scores decrease in both cases. One explanation for this could be many Twitter users do not join the conservation about the candidates until the debates. Since these individuals are less interested in politics, they might also be less likely to express positive opinions about Obama or Romney. Even if I focus only on tweets sent by a 'panel' of users, the problem of having different individuals 'tweeting' about the candidates at different times represents an additional source of bias, that may be solved by assigning prior probabilities to their opinions, so that we can estimate the polarity of their tweets even when they didn't send them.

7 Conclusions

Millions of people are writing personal messages on Twitter everyday. Many of these "tweets" are either irrelevant personal experiences, replication of existing information or simply spam. However, given the number and heterogeneity of users, some valuable data can be extracted from this source. Recently, some scholars have started to examine whether specific patterns in the stream of tweets might be able to predict consumer behavior. But the literature on the measurement of public opinion using Twitter data is still underdeveloped.

One of the main reasons is the lack of certainty about any inference that we might draw from this data. Twitter users are younger, more interested in politics and have higher incomes than the average citizen. It is therefore necessary to know more about the distribution of key socioeconomic and political factors among Twitter users in order to be able to infer valid estimates from this data.

That was the motivation behind this paper. Addressing these concerns, I have proposed a new measure of ideology in Twitter that might be used to weight estimates of public opinion in future studies. In contrast with the existing content-based measures, I have argued that a more promising approach is to study the 'following' links between ordinary Twitter users and political actors with a strong presence on this platform. I have applied this measure in four different countries: United States, United Kingdom, Spain, and the Netherlands; and in a sample of voters from Ohio. My results show that this method successfully classifies most political actors and ordinary citizens according to their political orientation, with the locations along the ideological scale resembling positions estimated using roll-call voting, party manifestos, and expert surveys. While further research is needed to validate the method, and to interpret the meaning of the emerging scale, my results illustrate the unexplored potential of Twitter data to generate new estimates that could prove useful in political science.

I have presented three examples of such applications in this paper. First, I have estimated the ideological ideal point for the average Twitter user in each state. By using the proper weighting to partially correct for biases due to the inherent self-selective nature of Twitter, these types of measures could be used to test spatial voting models. In particular, one of the most promising possibilities of this type of analysis is that it would allow meaningful comparisons across citizens and politicians positions on a similar ideological scale.

Secondly, I have addressed a heated debate in the literature: do social media increase polarization? Using the 2012 US presidential election campaign as case of study, I find that public exchanges in Twitter take place predominantly among users with similar viewpoints, and also that right-leaning users form a cluster of highly-motivated individuals, who dominate public conversations on Twitter.

Finally, I have used sentiment analysis to measure public opinion about politicians using the content of "tweets" who mention them. Despite the simplicity of my approach, it has shown the potential of using Twitter as a source of information about current events in real time and from a "micro" perspective. While many other studies (and media outlets) are currently using this type of technique, they do not distinguish by the individual characteristics of who sends the tweets. Similar to how micro-level covariates can be very informative in the analysis of survey data, examining sentiment across different ideological groups of Twitter users can improve the analysis of public opinion estimates based on social media data.

References

Adamic, L., and N. Glance (2005): "The political blogosphere and the 2004 US election: divided they blog," in *Proceedings of the 3rd international workshop on Link discovery*, pp. 36–43. ACM.

Adams, J., S. Merrill, and B. Grofman (2005): *A unified theory of party competition: A cross-national analysis integrating spatial and behavioral factors*. Cambridge Univ Pr.

Al Zamal, F., W. Liu, and D. Ruths (2012): "Homophily and latent attribute inference: Inferring latent attributes of twitter users from neighbors," in *Proceedings of the International Conference on Weblogs and Social Media*.

Ambekar, A., C. Ward, J. Mohammed, S. Male, and S. Skiena (2009): "Name-ethnicity classification from open sources," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '09, pp. 49–58, New York, NY, USA. ACM.

Ansolabehere, S., and E. Hersh (2012): "Validation: What Big Data Reveal About Survey Misreporting and the Real Electorate," *Political Analysis*.

Asur, S., and B. Huberman (2010): "Predicting the future with social media," in *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on*, vol. 1, pp. 492–499. IEEE.

Bafumi, J., A. Gelman, D. Park, and N. Kaplan (2005): "Practical issues in implementing and understanding Bayesian ideal point estimation," *Political Analysis*, 13(2), 171–187.

Benoit, K., and M. Laver (2006): *Party policy in modern democracies*, vol. 19. Taylor & Francis.

Blais, A., D. Blake, and S. Dion (1993): "Do parties make a difference? Parties and the size of government in liberal democracies," *American Journal of Political Science*, pp. 40–62.

Bollen, J., H. Mao, and X. Zeng (2011): "Twitter mood predicts the stock market," *Journal of Computational Science*.

Bollen, J., A. Pepe, and H. Mao (2010): "Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena," in *Proc. of WWW 2010 Conference*.

Bond, R., C. Fariss, J. Jones, A. Kramer, C. Marlow, J. Settle, and J. Fowler (2012): "A 61-million-person experiment in social influence and political mobilization," *Nature*, 489(7415), 295–298.

Boutet, A., H. Kim, E. Yoneki, et al. (2012): "Whats in Your Tweets? I Know Who You Supported in the UK 2010 General Election," in *Sixth International AAAI Conference on Weblogs and Social Media*.

Bradley, A. (1997): "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern recognition*, 30(7), 1145–1159.

Brier, G. (1950): "Verification of forecasts expressed in terms of probability," *Monthly weather review*, 78(1), 1–3.

Bryant, J., and D. Miron (2004): "Theory and research in mass communication," *Journal of communication*, 54(4), 662–704.

Budge, I. (2001): *Mapping policy preferences: estimates for parties, electors, and governments, 1945-1998*. Oxford University Press, USA.

Carsey, T., and J. Harden (2010): "New measures of partisanship, ideology, and policy mood in the American states," *State Politics & Policy Quarterly*, 10(2), 136–156.

Chan, J., and W. Suen (2008): "A spatial theory of news consumption and electoral competition," *Review of Economic Studies*, 75(3), 699–728.

Chang, J., I. Rosenn, L. Backstrom, and C. Marlow (2010): "ePluribus: Ethnicity on Social Networks," in *Proceedings of the Fourth International Conference on Weblogs and Social Media (ICWSM-10)*, Washington DC. AAAI Press, AAAI Press.

Choy, M., M. Cheong, M. Laik, and K. Shung (2011): "A sentiment analysis of Singapore Presidential Election 2011 using Twitter data with census correction," *Arxiv preprint arXiv:1108.5520*.

Ciulla, F., D. Mocanu, A. Baronchelli, B. Gonçalves, N. Perra, and A. Vespignani (2012): "Beating the news using Social Media: the case study of American Idol," *Arxiv preprint arXiv:1205.4467*.

Clinton, J., S. Jackman, and D. Rivers (2004): "The statistical analysis of roll call data," *American Political Science Review*, 98(2), 355–370.

Congosto, M., M. Fernández, and E. Moro Egido (2011): "TWITTER Y POLÍTICA: INFORMACIÓN, OPINIÓN Y PREDICCIÓN?," *Cuadernos de Comunicación Evoca*, (4).

Conover, M., B. Gonçalves, A. Flammini, and F. Menczer (2012): "Partisan Asymmetries in Online Political Activity," .

Conover, M., B. Gonçalves, J. Ratkiewicz, A. Flammini, and F. Menczer (2010): "Predicting the political alignment of twitter users," Unpublished manuscript.

Conover, M., J. Ratkiewicz, M. Francisco, B. Goncalves, A. Flammini, and F. Menczer (2011): "Political polarization on twitter," in *Proc. 5th Intl. Conference on Weblogs and Social Media*.

Cummings, D., H. Oh, and N. Wang (2010): "Who Needs Polls? Gauging Public Opinion from Twitter Data," Unpublished manuscript.

Dodds, P., K. Harris, I. Kloumann, C. Bliss, and C. Danforth (2011): "Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter," *Arxiv preprint arXiv:1101.5120*.

Dow, J. K. (2001): “A comparative spatial analysis of majoritarian and proportional elections,” *Electoral Studies*, 20(1), 109 – 125.

Downs, A. (1957): “An economic theory of political action in a democracy,” *The Journal of Political Economy*, 65(2), 135–150.

Enelow, J., and M. Hinich (1984): *The spatial theory of voting: An introduction.* Cambridge Univ Pr.

Farrell, H. (2012): “The Internet’s Consequences for Politics,” *Annual Review of Political Science, forthcoming.*

Franklin, M. (2004): *Voter turnout and the dynamics of electoral competition in established democracies since 1945.* Cambridge University Press.

Garrett, R. (2009): “Politically motivated reinforcement seeking: Reframing the selective exposure debate,” *Journal of Communication*, 59(4), 676–699.

Gayo-Avello, D. (2010): “All liaisons are dangerous when all your friends are known to us,” *Arxiv preprint arXiv:1012.5913.*

——— (2012): “I Wanted to Predict Elections with Twitter and all I got was this Lousy Paper A Balanced Survey on Election Prediction using Twitter Data,” *Arxiv preprint arXiv:1204.6441.*

Gelman, A., J. Carlin, H. Stern, and D. Rubin (2013): *Bayesian data analysis, 3rd Edition.* Chapman & Hall/CRC.

Gelman, A., J. Hill, and E. Corporation (2007): *Data analysis using regression and multilevel/hierarchical models.* Cambridge University Press.

Gelman, A., and D. Rubin (1992): “Inference from iterative simulation using multiple sequences,” *Statistical science*, 7(4), 457–472.

Golder, S., and M. Macy (2011): “Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures,” *Science*, 333(6051), 1878–1881.

Gong, A. (2011): “An automated snowball census of the political web,” .

Gonzalez, R., R. Cuevas, A. Cuevas, and C. Guerrero (2011): “Where are my followers? Understanding the Locality Effect in Twitter,” *Arxiv preprint arXiv:1105.3682.*

Gruzd, A. (2012): “Investigating Political Polarization on Twitter: A Canadian Perspective,” *unpublished manuscript.*

Hale, S., D. Gaffney, and M. Graham (2012): “Where in the world are you? Geolocation and language identification in Twitter,” Discussion paper, Working paper.

Hannak, A., E. Anderson, L. Barrett, S. Lehmann, A. Mislove, and M. Riedewald (2012): “Tweetinin the Rain: Exploring societal-scale effects of weather on mood,” .

Hecht, B., L. Hong, B. Suh, and E. Chi (2011): “Tweets from justin bieber’s heart: the dynamics of the location field in user profiles,” in *Proceedings of the 2011 annual conference on Human factors in computing systems*, pp. 237–246. ACM.

Holbrook-Provow, T., and S. Poe (1987): "Measuring state political ideology," *American Politics Research*, 15(3), 399–416.

Inglehart, R. (1990): *Culture shift in advanced industrial society*. Princeton Univ Press.

Jackman, S. (2012): "Estimates of Members' Preferences, 112th U.S. House and Senate," Retrieved on August 3rd, 2012.

Jessee, S. (2009): "Spatial voting in the 2004 presidential election," *American Political Science Review*, 103(01), 59–81.

Jungherr, A., P. Jurgens, and H. Schoen (2011): "Why the pirate party won the german election of 2009 or the trouble with predictions: A response to "predicting elections with twitter: What 140 characters reveal about political sentiment"," Social Science Computer Review.

Kedar, O. (2005): "When moderate voters prefer extreme parties: Policy balancingin parliamentary elections," *American Political Science Review*, 99(02), 185–199.

King, A., F. Orlando, and D. Sparks (2011): "Ideological Extremity and Primary Success: A Social Network Approach," *Paper presented at the 2011 MPSA Conference.*

——— (2012): "The Social Primary: Measuring Momentum in the 2012 Republican Primary Campaign," *Unpublished Manuscript*.

King, G., J. Alt, N. Burns, and M. Laver (1990): "A unified model of cabinet dissolution in parliamentary democracies," *American Journal of Political Science*, pp. 846–871.

Kwak, H., C. Lee, H. Park, and S. Moon (2010): "What is Twitter, a social network or a news media?," in *Proceedings of the 19th international conference on World wide web*, pp. 591–600. ACM.

Lampos, V. (2012): "On voting intentions inference from Twitter content: a case study on UK 2010 General Election," *Arxiv preprint arXiv:1204.0423*.

Lampos, V., T. De Bie, and N. Cristianini (2010): "Flu detector-tracking epidemics on twitter," *Machine Learning and Knowledge Discovery in Databases*, pp. 599–602.

Lau, R., and D. Redlawsk (1997): "Voting correctly," *American Political Science Review*, pp. 585–598.

Laver, M., and K. Shepsle (1996): *Making and breaking governments: Cabinets and legislatures in parliamentary democracies*. Cambridge Univ Pr.

Lazarsfeld, P., B. Berelson, and H. Gaudet (1944): *The peoples choice: How the voter makes up his mind in a presidential election*. New York: Duell, Sloan and Pearce.

Lewis-Beck, M., H. Norpoth, W. Jacoby, and H. Weisberg (2008): *The American voter revisited*. University of Michigan Press.

McPherson, M., L. Smith-Lovin, and J. Cook (2001): "Birds of a feather: Homophily in social networks," *Annual review of sociology*, pp. 415–444.

Metaxas, P., E. Mustafaraj, and D. Gayo-Avello (2011): "How (Not) To Predict Elections," Unpublished manuscript.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller (1953): "Equation of state calculations by fast computing machines," *The journal of chemical physics*, 21, 1087.

Mislove, A., S. Lehmann, Y. Ahn, J. Onnela, and J. Rosenquist (2011): "Understanding the Demographics of Twitter Users," in *Fifth International AAAI Conference on Weblogs and Social Media*.

Mutz, D. (2002): "Cross-cutting social networks: Testing democratic theory in practice," *American Political Science Review*, 96(1), 111–126.

O'Connor, B., R. Balasubramanyan, B. Routledge, and N. Smith (2010): "From tweets to polls: Linking text sentiment to public opinion time series," in *Proceedings of the International AAAI Conference on Weblogs and Social Media*, pp. 122–129.

Oken Hodas, N., and K. Lerman (2012): "How Visibility and Divided Attention Constrain Social Contagion," .

Pak, A., and P. Paroubek (2010): "Twitter as a corpus for sentiment analysis and opinion mining," *Proceedings of LREC 2010*.

Parmelee, J., and S. Bichard (2011): *Politics and the Twitter Revolution: How Tweets Influence the Relationship Between Political Leaders and the Public*. Lexington Books.

Paul, M., and M. Dredze (2011): "You are what you Tweet: Analyzing Twitter for public health," in *Barcelona, Spain: 5th International AAAI Conference on Weblogs and Social Media (ICWSM 2011)*.

Pennacchiotti, M., and A. Popescu (2011): "A machine learning approach to twitter user classification," in *Fifth International AAAI Conference on Weblogs and Social Media*.

Poole, K., and H. Rosenthal (1997): *Congress: A political-economic history of roll call voting*. Oxford University Press, USA.

——— (2007): *Ideology and Congress*. Transaction Pub, 2nd edition edn.

Rao, T., and S. Srivastava (2012): "Twitter Sentiment Analysis: How To Hedge Your Bets In The Stock Markets," *arXiv preprint arXiv:1212.1107*.

Reynolds, A. (2002): *The architecture of democracy: Constitutional design, conflict management, and democracy*. Oxford University Press.

Riloff, E., and J. Wiebe (2003): "Learning extraction patterns for subjective expressions," in *Proceedings of the 2003 conference on Empirical methods in natural language processing*, pp. 105–112. Association for Computational Linguistics.

Sakaki, T., M. Okazaki, and Y. Matsuo (2010): "Earthquake shakes Twitter users: real-time event detection by social sensors," in *Proceedings of the 19th international* *conference on World wide web*, pp. 851–860. ACM.

Sang, E., and J. Bos (2012): "Predicting the 2011 Dutch Senate Election Results with Twitter," *EACL 2012*, p. 53.

Sartori, G. (1994): *Comparative constitutional engineering: an inquiry into structures, incentives, and outcomes*. NYU Press.

Schelling, T. (1978): *Micromotives and macrobehavior*. WW Norton & Company.

Shor, B., C. Berry, and N. McCarty (2010): "A Bridge to Somewhere: Mapping State and Congressional Ideology on a Cross-institutional Common Space," *Legislative Studies Quarterly*, 35(3), 417–448.

Skoric, M., N. Poor, P. Achananuparp, E. Lim, and J. Jiang (2012): "Tweets and Votes: A Study of the 2011 Singapore General Election," in *2012 45th Hawaii International Conference on System Sciences*, pp. 2583–2591. IEEE.

Stokes, D. (1963): "Spatial models of party competition," *The American Political Science Review*, 57(2), 368–377.

Strom, K. (1990): *Minority government and majority rule*. Cambridge Univ Press.

Sunstein, C. (2001): *Republic. com*. Princeton University Press.

Team, S. D. (2012): "Stan Modeling Language: User's Guide and Reference Manual. Version 1.0.," .

Tumasjan, A., T. Sprenger, P. Sandner, and I. Welpe (2010): "Predicting elections with twitter: What 140 characters reveal about political sentiment," in *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, pp. 178–185.

Wong, F., S. Sen, and M. Chiang (2012): "Why Watching Movie Tweets Won't Tell the Whole Story?," *Arxiv preprint arXiv:1203.4642*.

Wu, S., J. Hofman, W. Mason, and D. Watts (2011): "Who says what to whom on twitter," in *Proceedings of the 20th international conference on World wide web*, pp. 705–714. ACM.

Yardi, S., and D. Boyd (2010): "Dynamic debates: An analysis of group polarization over time on twitter," *Bulletin of Science, Technology & Society*, 30(5), 316–327.

A Appendix. Additional Results

Figure 18: Estimated Ideal Points for Key Political Actors with 10,000 or more followers

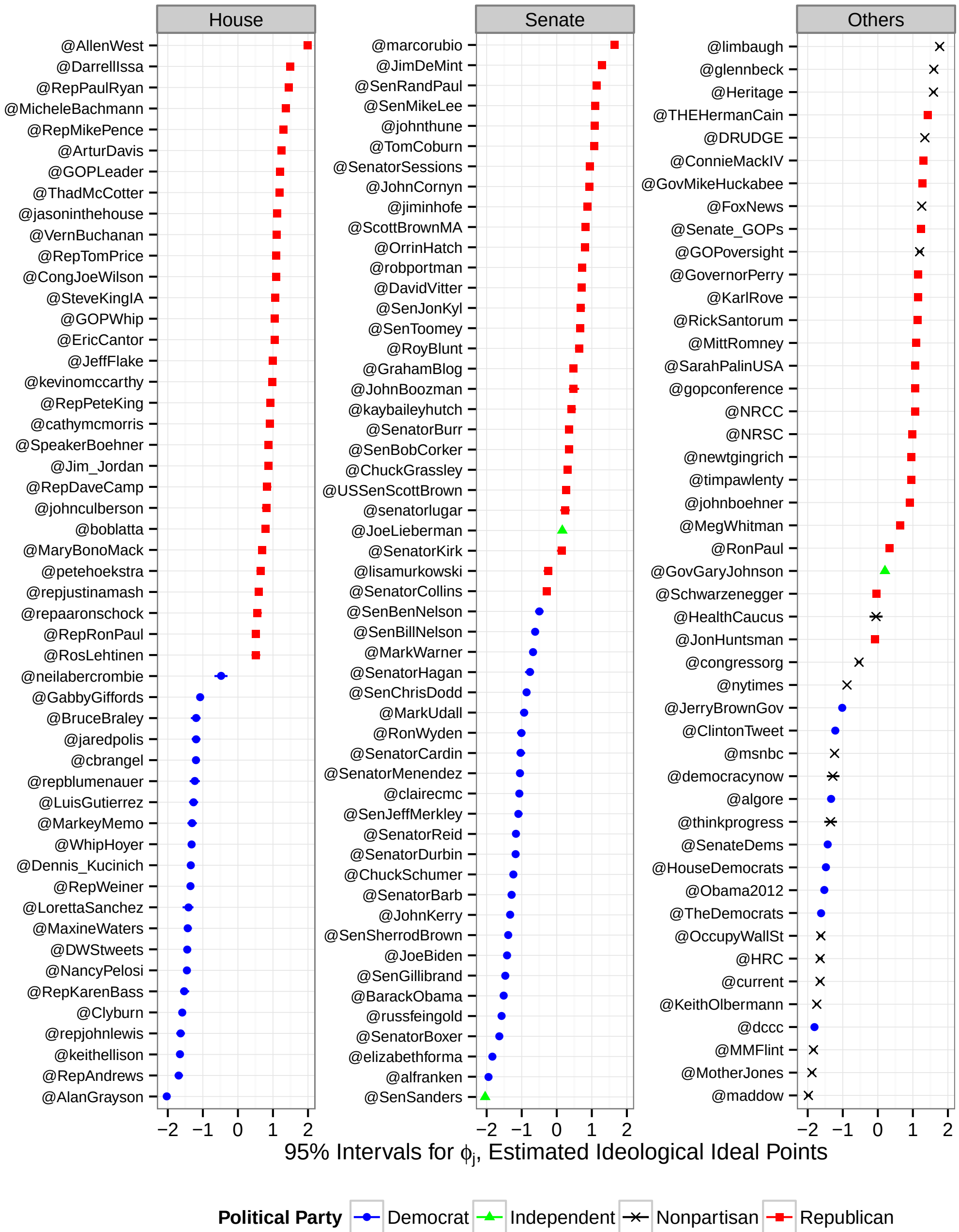


Figure 19: Distribution of Ideal Point Estimates, by Gender

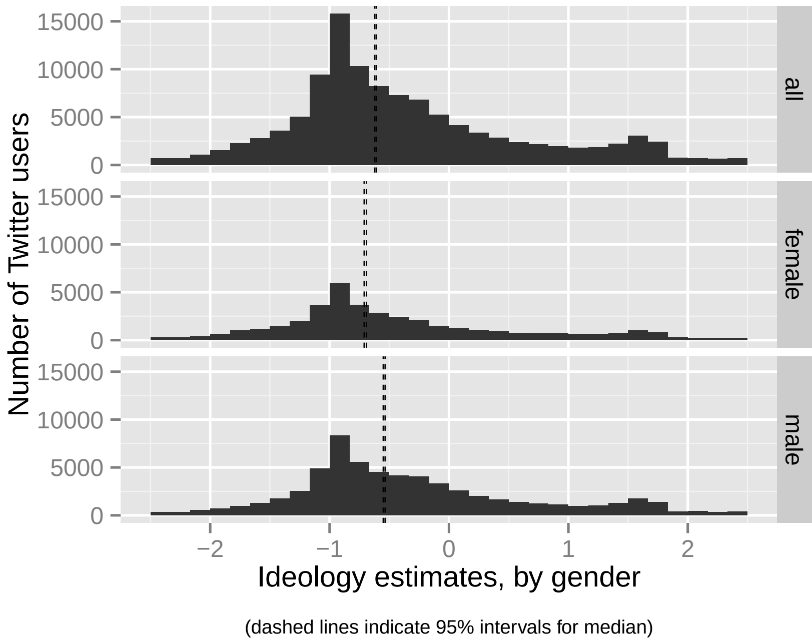


Figure 20: Distribution of Ideal Point Estimates, by Language

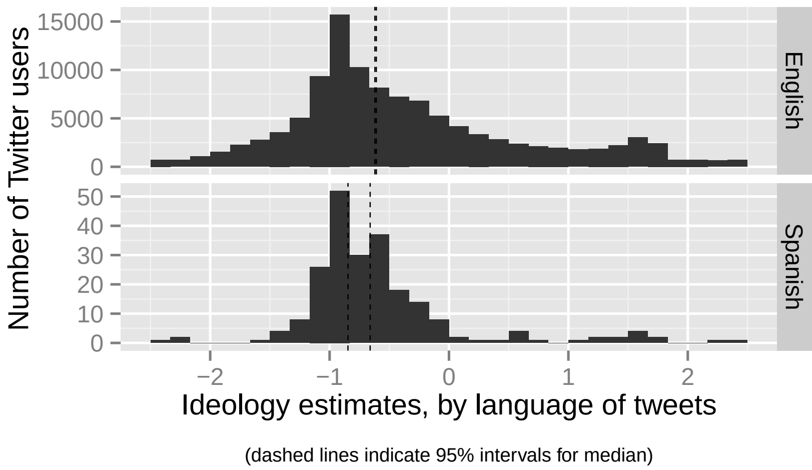


Figure 21: Affective Scores of Tweets Mentioning Obama or Romney, by Gender and Ideology of Twitter Users

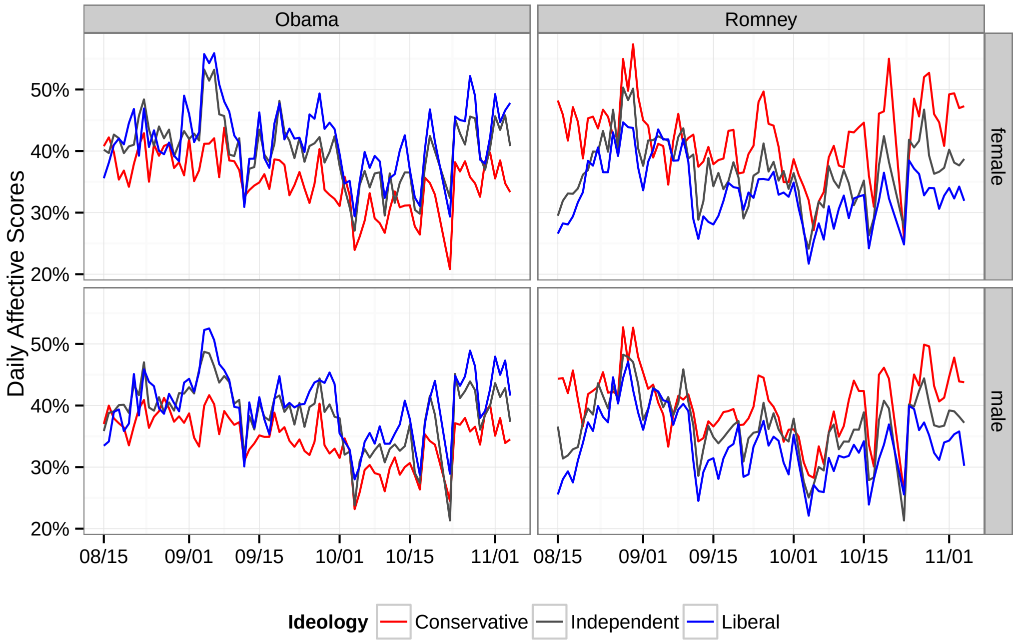
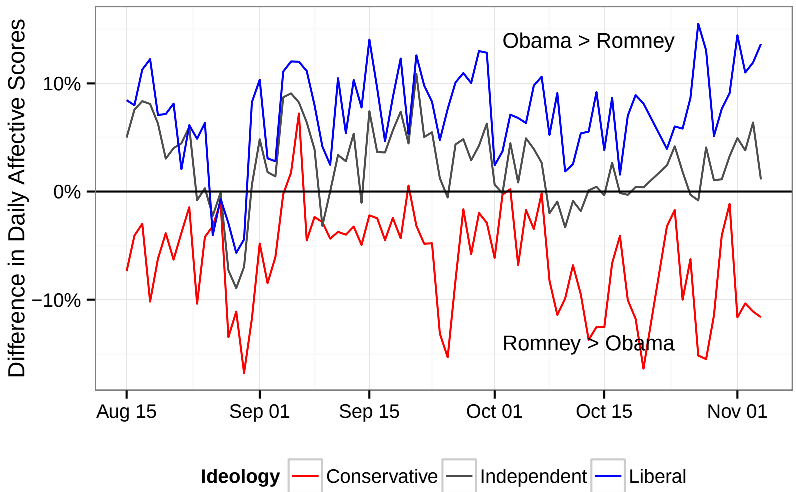


Figure 22: Daily Differences in Weighted Affective Scores of Tweets Mentioning Obama or Romney, by Ideology of Twitter Users



B Technical Appendix

B.1 Estimation of the Bayesian Spatial Following Model

B.1.1 The Model

The statistical model I assume to explain the decision of following a political account in Twitter is:

$$P(y_{ij} = 1|\alpha_j, \beta_i, \gamma, \theta_i, \phi_j) = \text{logit}^{-1}\left(\alpha_j + \beta_i - \gamma||\theta_i - \phi_j||^2\right) \tag{2}$$

where y_{ij} equals 1 when user $i \in \{1, \ldots, n\}$ decides to follow political actors $j \in \{1, \ldots, m\}$, θ_i is the ideal point of user i on the ideological dimension; ϕ_j is the ideal point of user j; α_j and β_i are the indegree and outdegree parameters of users i and j, respectively; and γ is a normalizing constant.

Given that none of these parameters is observed, the statistical problem here is inference for $\theta = (\theta_i, \ldots, \theta_n)'$, $\phi = (\phi_j, \ldots, \phi_m)'$, $\alpha = (\alpha_j, \ldots, \alpha_m)'$, $\beta = (\beta_i, \ldots, \beta_n)'$, and γ. Assuming local independence (individual decisions to follow are independent across users n and m, conditional on the estimated parameters), the likelihood function to maximize is the following:

$$p(\mathbf{y}|\theta, \phi, \alpha, \beta, \gamma) = \prod_{i=1}^{n}\prod_{j=1}^{m} \text{logit}^{-1}(\pi_{ij})^{y_{ij}}(1 - \text{logit}^{-1}(\pi_{ij}))^{1-y_{ij}} \tag{3}$$
$$\text{where } \pi_{ij} = \alpha_j + \beta_i - \gamma||\theta_i - \phi_j||^2$$

The complexity of this equation makes direct estimation using maximum likelihood highly intractable. However, samples from the posterior density of each parameter in the model can be obtained using Markov-Chain Monte Carlo methods. More specifically, to make computation more efficient, I employ a hierarchical setup that considers each of the four set of parameters as drawn from common population distributions whose hyperparameters are also estimated:

$$\alpha_j \sim \text{N}(\mu_\alpha, \sigma_\alpha) \qquad \beta_j \sim \text{N}(\mu_\beta, \sigma_\beta)$$
$$\theta_i \sim \text{N}(\mu_\theta, \sigma_\theta) \qquad \phi_j \sim \text{N}(\mu_\phi, \sigma_\phi)$$

The full joint posterior distribution is thus:

$$
\begin{aligned}
p(\theta, \phi, \alpha, \beta, \gamma|\mathbf{y}) \propto\ & p(\theta, \phi, \alpha, \beta, \gamma, \mu, \sigma) \\
& \prod_{i=1}^{n}\prod_{j=1}^{m} \text{logit}^{-1}(\pi_{ij})^{y_{ij}}(1-\text{logit}^{-1}(\pi_{ij}))^{1-y_{ij}} \\
& \prod_{j=1}^{m} \text{N}(\alpha_j|\mu_\alpha, \sigma_\alpha) \prod_{i=1}^{n} \text{N}(\beta_i|\mu_\beta, \sigma_\beta) \\
& \prod_{i=1}^{n} \text{N}(\theta_i|\mu_\theta, \sigma_\theta) \prod_{j=1}^{m} \text{N}(\phi_j|\mu_\phi, \sigma_\phi)
\end{aligned}
\tag{4}
$$

B.1.2 Identification

The increasing number of parameters that this method estimates[22] comes at a cost. As it stands in equation 1, the model is unidentified: any constant can be added to all the parameters θ_i and ϕ_j without changing the predictions of the model; and similarly θ_i or ϕ_j can be multiplied by any non-zero constant, with γ divided by the same constant, or α_j or β_i divided by its square root. These problems are sometimes called "additive aliasing" and "scaling invariance" (see e.g. Bafumi, Gelman, Park, and Kaplan, 2005). In order to solve these two issues and identify the equation, I fix the mean of σ_α at one and apply a unit variance restriction on θ. In the multilevel setting, this is equivalent to giving the θ_i's an informative $\text{N}(0, 1)$ prior distribution (Gelman, Hill, and Corporation, 2007, p.318).

An additional difficulty is reflection invariance: the resulting scale can be reversed (flipped left-to-right) without changing the prediction of the models. This is a problem for interpretation, but not for estimation, and can be easily solved by specifying starting values that are consistent with the expected direction of the scale (e.g. setting θ_i for a liberal Twitter user as -1), or by multiplying the estimated parameters by -1 after the chain has converged if the scale is not in the proper left-right direction.

B.1.3 MCMC algorithm

To improve the efficiency of the estimation procedure, I divide it in two stages. First, I use a No-U-Turn sampler, a variant of Hamiltonian Monte Carlo sampling algorithms (Gelman, Carlin, Stern, and Rubin, 2013), to estimate the parameters indexed by j. This allows for very efficient sampling from the posterior distribution – 1,000 iterations from 2 chains with overdispersed starting values is enough to obtain effective number

[22] In the case of the US, for example, I am estimating a total of over 600,000 parameters.

of simulation draws over 500. However, the current implementation of this algorithm in the Stan modeling language (Team, 2012) is still not prepared for large datasets, and therefore I use only a random sample of 50,000 i users. This limitation does not affect the estimation of the parameters indexed by j, whose posterior distribution remains essentially constant once the size of this random sample is higher than 10,000 i users.

In the second stage, I use a Metropolis algorithm (Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller, 1953) with a uniform jumping distribution to estimate all parameters indexed by i. The only difference with respect to the original formulation of this sampling algorithm is that, in each iteration, a new value of the j parameters is drawn from their posterior distributions (estimated in the first stage), in order to account for our uncertainty about their true value. This approach seems less efficient than the one used in the first stage – over 2,000 iterations and 2 chains are now necessary to obtain effective number of simulation draws around 200. However, note that each of the i parameters can be estimated individually because we assume local independence[23], and therefore multi-core processors can be used to run multiple samplers simultaneously and dramatically increase computation speed[24].

The first stage is implemented using the Stan modeling language, while the second stage is implemented using `R`. I use flat priors on all parameters, with the exception of μ_θ, σ_θ and σ_α, which are fixed to 0, 1, and 1 respectively for identification purposes. The samplers in both stages are run using two chains with as many iterations as necessary to obtain effective number of simulation draws (Gelman and Rubin, 1992) over 200, which is enough to estimate the parameter means with a precision of two decimal digits. Each chain is initiated with random draws from a multivariate normal distribution for ϕ and γ, the logarithm of the "indegree" of user j or "outdegree" of user i for α and β (to speed up convergence), and values zero for θ, with the exception of those who belong to a party, -1 for left-wing politicians and $+1$ for right-wing politicians. This choice is only necessary in order to interpret the results of the model (see section above). The results appear to be insensitive to the choice of priors and initial values. All relevant code is available upon request.

B.2 Convergence Diagnostics and Model Fit

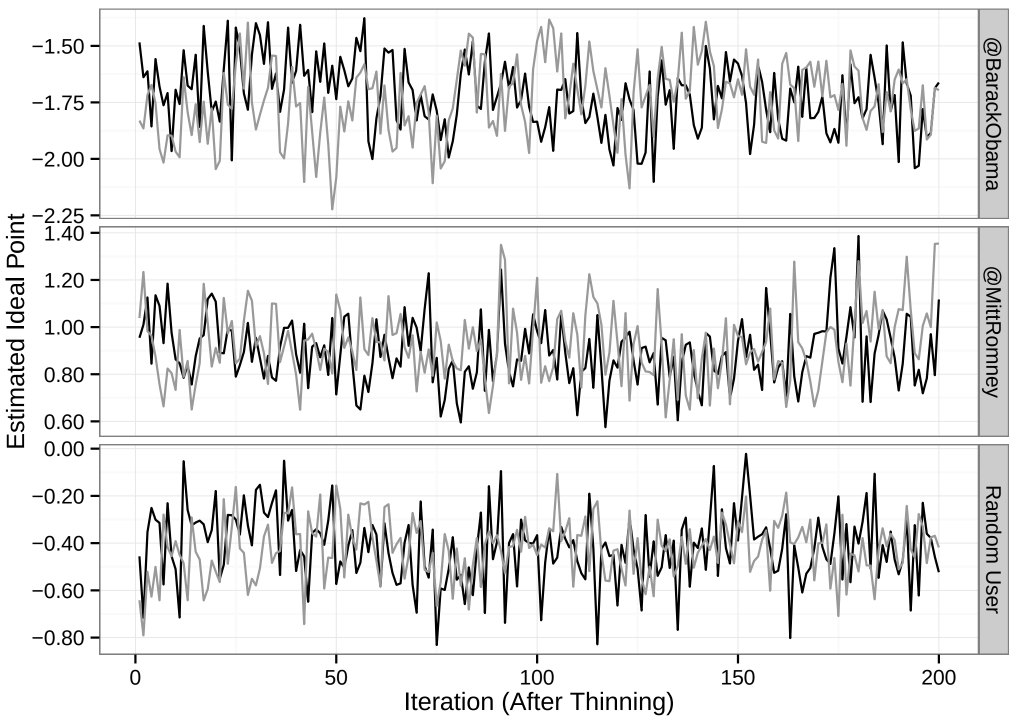
Despite the relatively low number of iterations, visual analysis of the trace plots, estimation of the $\hat{R}$ diagnostics, and effective number of simulation draws show high level of convergence in the Markov Chains. Figure 23 shows that each of the two chains used

[23] Note also that it is the independence assumption what allows us to divide the estimation method in these two stages.

[24] Samples from the i parameters in the second stage can be compared with those obtained for the random sample in the first stage to ensure that there were no errors in the estimation.

to estimate the ideology of Barack Obama, Mitt Romney and a random i user have converged to stationary distributions. Similarly, all $\hat{R}$ values are below 1.05, which is considered the rule-of-thumb to monitor convergence of multiple chains, and the effective number of simulation draws is over 200 for all parameters. The results of running Geweke and Heidelberg diagnostics also indicate that the distribution of the chains is stationary.

Figure 23: Trace Plots. Iterative History of the MCMC Algoritm



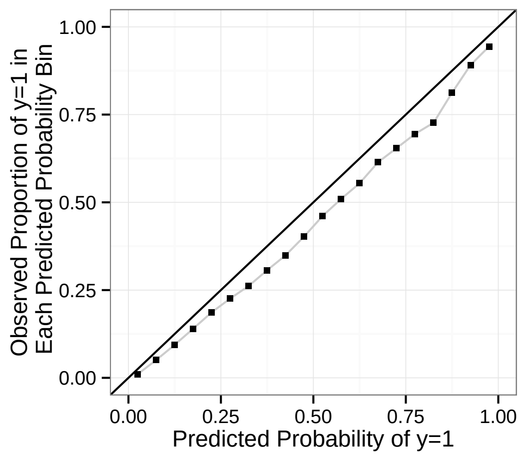
The results of a battery of predictive checks for binary dependent variables are shown in Table 2. All of them show that the fit of the model is adequate: despite the sparsity of the ‘following’ matrix (less than 8% of values are 1’s), the model’s predictions improve the baseline (predicting all y_{ij} as zeros), which suggests that Twitter users’ following decisions are indeed guided by ideological concerns. Besides the widely known Pearson’s ρ correlation and the proportion of correctly predicted values, Table 2 also shows the AUC and Brier Scores. The former measures the probability that a randomly selected $y_{ij} = 1$ has a higher predicted probability than a randomly selected $y_{ij} = 0$ and ranges from 0.5 to 1, with higher values indicating better predictions (Bradley, 1997). The latter is the mean squared difference between predicted probabilities and actual values of y_{ij} (Brier, 1950).

Table 2: Model Fit Statistics.

Statistic	**Value**
Pearson's ρ Correlation	0.592
Proportion Correctly Predicted	0.940
PCP in Baseline (all $y_{ij} = 0$)	0.924
AUC Score	0.915
Brier Score	0.046
Brier Score in Baseline (all $y_{ij} = 0$)	0.076

A visual analysis of the model fit is also shown in Figure 24, which displays a calibration plot where the predicted probabilities of $y_{ij} = 1$, ordered and divided into 20 equally sized bins (x-axis), are compared with the observed proportion of $y_{ij} = 1$ in each bin. This plot also confirms the good fit of the model, given that the relationship between observed and predicted values is very close to a 45-degree line (in dark color).

Figure 24: Model Fit. Comparing Observed and Predicted Proportions of $y_{ij} = 1$



C Imputing Twitter Users' Gender and Race

C.1 Description

Gender and race are two of the most important predictors of electoral behavior and individual political attitudes in the United States (see e.g. Lewis-Beck, Norpoth, Jacoby, and Weisberg, 2008). In this section I present a method to impute these two variables for each Twitter user based on their self-reported full name. This method uses the frequencies of appearance of first (and last) names by gender (and race) in anonymized databases as the basis for a model that estimates the probability that a given user belongs to each gender or race group. This approach has been found to yield highly accurate estimates by different studies (Ambekar, Ward, Mohammed, Male, and Skiena, 2009; Chang, Rosenn, Backstrom, and Marlow, 2010; Mislove, Lehmann, Ahn, Onnela, and Rosenquist, 2011). Imputing gender based on first names is also a common procedure when matching and validating voting records (Ansolabehere and Hersh, 2012, p.6).

In short, the method has three different steps. First, each Twitter user's full name is pre-processed and split in first, middle, and last name; and users who report their gender or race in the description are pre-classified. Second, Twitter users' names are merged with a database of common first and last names using a fuzzy string matching method. Finally, I use Bayes' rule to compute the probability that each user belongs to a given gender and race group, and use this vector of probabilities to impute gender and race.

In detail:

1. **Pre-processing.** Besides their screen names, Twitter users are also asked to provide their full name. Most of them accept to do so. Even if they choose pseudonyms, it is probably the case that the associated gender and race of their alternative names is the same as their own. The first step in the analysis is thus to split their reported names in first, middle (when available), and last name. In addition, before proceeding to the following step, I also classify those users who report their gender in their description. Individuals who define themselves as "mother", "woman", "girl", "wife", "grandmother"... are classified as women, whereas those who define themselves as "guy", "man", "dude", "dad", "father", "husband"... are classified as men. Around 18% of the users in the U.S. sample include any of these words in their description, and can therefore be classified with nearly perfect accuracy.

2. **Partial String Matching.** The second step is to match first, middle, and last names with their equivalent in a database that reports their use frequency among

different gender and race groups. In this paper, for reasons of convenience, I use the anonymized database available in the RandomNames `R` package. I then compute the generalized Levenshtein distance between each name in my dataset and each name in this database. This statistic measures the distance between two strings by computing the minimal number of insertions, deletions, and substitutions needed to transform one string into the other. This score is then divided by the number of characters in each name to obtain a score that ranges from 0 (perfect match) to the number of characters in the name (completely imperfect match). Matches whose score is above 0.25 are discarded. (This is threshold is arbitrary, but it seems to maximize precision in this application.) Around 80% of Twitter users have at least one of their names matched to a name in the `randomNames` database. Note that the partial string matching step is necessary because some names are misspelled or abbreviated.

3. **Computing Vector of Probabilities.** Once the names have been matched, how frequent they are across gender and race groups is known, and the probability that a given user belongs to each of these groups can be easily computed using Bayes rule. The probability that a given Twitter user belongs to a category j is:

$$P(C_j|fullname_i) = \frac{1}{K_i} \sum_{k=1}^{K_i} \frac{P(name_{ik}|C_j)P(C_j)}{\sum_{j=1}^{J} P(name_{ik}|C_j)P(C_j)}$$

where $name_{ik}$ is each of the K_i names of user i that have been matched with the database, $P(name_{ik}|C_j)$ is the probability that an individual in class j is named with $name_{ik}$, and $P(C_j)$ is the prior probability of belonging to class j. Note that each matched name is considered equally informative, since we assign equal weight to each match. Prior probabilities are obtained from the approximate distribution of gender and race in the US according to the 2010 census: 49% male, 51% female; 0.9% American Indian or Native Alaskan, 4.8% Asian or Pacific Islander, 12.4% Black (not Hispanic), 16.7% Hispanic, 65.2% White (not Hispanic).

This procedure generates, for each Twitter user and characteristic, a vector of probabilities indicating how likely it is that he/she belong to each group. These vectors are used to sample from a binomial (or multinomial) distribution and thus impute a value for the race and gender variables for each Twitter user. The resulting distribution of both variables is reported in Table 1.

C.2 Validation

In this section I present the results of a preliminary validation analysis of the method I employ to impute gender and race. A random sample of 500 users was drawn from the universe of Twitter users in the US I consider in this paper. Their gender and race was labelled manually by a third person, based on their profile pictures, description, and most recent tweets. This data is considered "ground truth" and is compared to the imputed race and gender for each user in the following two confusion matrices.

Results are mixed. On one hand, gender is correctly imputed for 75.6% of the users in the sample, as reported in Table 3. This level of accuracy is high given the relatively low level of information that the method requires. In comparison, Al Zamal, Liu, and Ruths (2012) achieve a maximum accuracy of 80.2% using a complex machine learning model whose features are the complete text of each user's latest tweets, their network of friends and followers, their retweeting tendency, and other information. Note that this result is also significantly better than assigning gender at random – in which case the level of accuracy would be around 50% by construction.

Table 3: Confusion Matrix. Gender

		Imputed Gender		
		Male	Female	Unknown
Labelled Gender	Male	**206**	22	13
	Female	14	**117**	18
	Unknown	37	18	**55**

However, results are considerably worse in the case of race. As reported in Table 4, only 49.6% of users are correctly classified. This level of precision is better than what a random classifier would yield (around 15%), but not the level of precision that would result from classifying all users as "white" (54.4%). An analysis of the confusion table also shows that the method's performance is particularly poor at classifying African-American and Hispanic Twitter users.

Further work is necessary to understand why the method fails, and whether it can still be applied to impute Twitter users' race.

Table 4: Confusion Matrix. Race

		Imputed Race					
		Am.Ind.	Asian	Afr.Am.	Hisp.	White	Unkn.
Labelled Race	American Indian	**0**	0	0	1	1	0
	Asian	0	**1**	0	2	3	0
	African-American	0	3	**4**	2	14	7
	Hispanic	0	2	0	**10**	10	2
	White	2	25	22	32	**164**	27
	Unknown	3	18	16	13	47	**69**