

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

Pablo Barberá

*Wilf Family Department of Politics, New York University, 19 W 4th Street, 2nd Floor,
New York, NY 10012.*

e-mail: pablo.barbera@nyu.edu

Edited by R. Michael Alvarez

Politicians and citizens increasingly engage in political conversations on social media outlets such as Twitter. In this article, I show that the structure of the social networks in which they are embedded can be a source of information about their ideological positions. Under the assumption that social networks are homophilic, I develop a Bayesian Spatial Following model that considers ideology as a latent variable, whose value can be inferred by examining which politics actors each user is following. This method allows us to estimate ideology for more actors than any existing alternative, at any point in time and across many polities. I apply this method to estimate ideal points for a large sample of both elite and mass public Twitter users in the United States and five European countries. The estimated positions of legislators and political parties replicate conventional measures of ideology. The method is also able to successfully classify individuals who state their political preferences publicly and a sample of users matched with their party registration records. To illustrate the potential contribution of these estimates, I examine the extent to which online behavior during the 2012 US presidential election campaign is clustered along ideological lines.

1 Introduction

Measuring politicians' and voters' policy positions is a relevant, yet complex, scientific endeavor. Studies of electoral behavior, government formation, and party competition require systematic information on the placement of key political actors and voters on the relevant policy dimensions. The development of methods to estimate such positions, usually in a single latent dimension characterized as “ideology” (Clinton, Jackman, and Rivers 2004; Poole and Rosenthal 2007; Jessee 2009; Shor, Berry, and McCarty 2010; Bonica 2013b), represents one of the most important methodological contributions to political science in the past two decades.

However, most studies estimate ideal points for legislators only. When the analysis also includes voters, it is done at the expense of strong “bridging” assumptions (Jessee 2009), or only for self-selected population groups (Bonica 2014). There is also little work on cross-national ideological estimation (Lo, Proksch, and Gschwend 2013). Most importantly, given the sparse nature of the data (roll call votes or contributions) and its costly collection (survey data), current measurement methods generate ideal points that are essentially static in the short run. In this article, I show that using Twitter networks as a source of information about policy positions has the potential to solve these difficulties.

Twitter has become one of the most important communication arenas in daily politics. Initially conceived as a Web site to share personal status updates, it now has more than two hundred million monthly active users worldwide,¹ including 18% of all online Americans.² One distinct

Author's note: I would like to thank Jonathan Nagler, Joshua Tucker, Nick Beauchamp, Neal Beck, Ken Benoit, Richard Bonneau, Patrick Egan, Adam Harris, John Jost, Franziska Keller, Michael Laver, Alan Potter, Gaurav Sood, Chris Tausanovitch, Shana Warren, and two anonymous reviewers for helpful comments and discussions. Supplementary materials for this article are available on the *Political Analysis* Web site.

¹Source: Twitter's Official Twitter Account, December 18, 2012. [<https://twitter.com/twitter/status/281051652235087872>].

²Source: The Pew Research Center's Internet and American Life Project, August 2013. [<http://www.pewinternet.org/2013/08/05/72-of-online-adults-are-social-networking-site-users/>].

characteristic of this online social network is the presence of not only ordinary citizens, but also political actors. Virtually every legislator, political party, and candidate in developed democracies has an active Twitter account. Independent of their offline identities, they all interact within the same symbolic framework, using similar language in messages of identical length. Most importantly, they are embedded in a common social network. This opens the possibility of estimating ideological positions of all users on a common scale, which would allow for meaningful comparisons of voters' and legislators' ideal points.

The use of Twitter data presents three additional advantages over other sources of information about preferences. First, the large number of active users on this social networking site can be exploited to estimate highly precise ideal points for politicians, if we consider users as "experts" who are "rating" elites through their decisions of whom to follow. Second, the structure of this network is far from static, which can facilitate the estimation of highly granular dynamic ideal points in real time. Third, it is possible to link Twitter profiles to other data through name identification, which provides interesting ways to examine differences between private and public political behavior.

This series of advantages comes at the expense of one important limitation. Twitter users are not a representative sample of the voting age population. This can represent a difficulty in the context of studies about mass attitudes and behavior, but not for the method I present in this article. Citizens who discuss politics on Twitter are more likely to be educated and politically interested, and that makes them a particularly useful source of information about elites' ideology.

This method relies on the characteristics of the social ties that Twitter users develop with one another and, in particular, with the political actors (politicians, think tanks, news outlets, and others) 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. The decision to follow is considered a costly signal that provides information about Twitter users' perceptions of both their ideological location and that of political accounts. Unlike other studies that estimate political ideology using social media data (Conover et al. 2010; King, Orlando, and Sparks 2011; Boutet et al. 2012),³ I am able to estimate ideal points, with standard errors, on a continuous scale, for all types of active Twitter users, across different countries. To validate the method, I estimate the ideological positions of legislators, political parties, and a large sample of active users in the United States and five European countries. Their estimated ideal points replicate conventional measures of ideology.

This method represents an additional measurement tool that can be used to estimate ideology, an important quantity of interest in political science, for a larger set of political actors and individuals than any other method before. To illustrate a potential use of these estimates, I examine the extent to which online behavior during the 2012 US presidential election campaign is clustered along ideological lines, finding support for the so-called "echo-chamber" theory and high levels of political polarization at the mass level.

2 Ideal Point Estimation Using Twitter Data

2.1 Assumptions

In this article, I 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 between these two sets of users. In order to do so, I develop a Bayesian spatial model of Twitter users' following behavior.

The key assumption of this model is that Twitter users prefer to follow politicians whose positions on the latent ideological dimension are similar to theirs. This assumption is equivalent to that of spatial voting models (see, e.g., Enelow and Hinich 1984). I consider following decisions to be costly signals about users' perceptions of both their ideological location and that of political accounts. Such cost can take two forms. If the content of the messages users are exposed to as a

³See Section A of the supplementary materials for a summary of previous studies measuring ideology using social media.

result of their following decisions challenges their political views, it can create cognitive dissonance. Second, given the fast-paced nature of Twitter, it also creates opportunity costs, since it reduces the likelihood of being exposed to other messages, assuming the amount of time a user spends on Twitter is constant. In other words, these decisions provide information about how social media users decide to allocate a scarce resource—their attention. Although obviously less costly than campaign contributions or votes in a legislature, the assumption behind this model is similar in nature to that justifying how donations and roll call votes can be scaled onto a latent ideological dimension (Poole and Rosenthal 2007; Bonica 2014).

Two additional arguments support the notion that following decisions can be informative about ideology. First, 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 homogeneous networks with regard to many sociodemographic and behavioral traits. Multiple studies have observed patterns of homophilic segregation consistent with these models in networks of interactions between Twitter users (Wu et al. 2011; Conover et al. 2012).

However, Twitter is not only an online social network—it is also a news media (Kwak et al. 2010). From this perspective, we should also consider the existing literature on the selective exposure theory (Lazarsfeld, Berelson, and Gaudet 1944; Bryant and Miron 2004; Stroud 2008; Iyengar and Hahn 2009) that argues that individuals exhibit a preference for opinion-reinforcing political information and that they systematically avoid opinion challenges. Given the dynamic nature of social media, its large size, 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.

2.2 The Statistical Model

The statistical model I employ is similar in nature to latent space models applied to social networks (Hoff, Raftery, and Handcock 2002), item-response theory models (see, e.g., Linden and Hambleton 1997), and other methods that scale roll call votes or campaign contributions into latent political dimensions (Clinton, Jackman, and Rivers 2004; Poole and Rosenthal 2007; Bonica 2014), but adapted to allow the estimation of ideal points for hundreds of thousands of individuals.

Suppose that each Twitter user $i \in \{1, \dots, n\}$ is presented with a choice between following or not following another target user $j \in \{1, \dots, m\}$, where j is a political actor who has a Twitter account.⁴ 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 dimension⁵ 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 core model, I add two additional parameters, α_j and β_i . The former measures the popularity of user j . This parameter accounts for the fact that some political accounts are more likely to be followed, due to the higher profile of the politicians behind them (e.g., we would expect the probability of following @BarackObama to be higher than the probability of following a random member of the US Congress) or for other reasons (politicians who “tweet” more often are more likely to be highly visible and therefore also to have more followers). The latter measures the level of political interest of each user i . Similarly, this parameter accounts for the differences in

⁴If 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 article, 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 .

⁵I 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 generalized to multiple dimensions.

the number of political accounts each user i decides to follow, which could be related to the overall number of Twitter users they follow, or their overall level of interest in politics.

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

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

Given that none of these parameters is directly observed, the statistical problem here is the inference of $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)'$, $\boldsymbol{\phi} = (\phi_1, \dots, \phi_m)'$, $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m)'$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_n)'$, 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 this model is as follows:

$$p(\mathbf{y} | \boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma) = \prod_{i=1}^n \prod_{j=1}^m \text{logit}^{-1}(\pi_{ij})^{y_{ij}} (1 - \text{logit}^{-1}(\pi_{ij}))^{1-y_{ij}}, \quad (2)$$

where $\pi_{ij} = \alpha_j + \beta_i - \gamma \|\theta_i - \phi_j\|^2$.

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. However, samples from the posterior density of each parameter in the model can be obtained using Markov Chain Monte Carlo methods. To improve the efficiency of this procedure, I use a Hamiltonian Monte Carlo algorithm (Gelman et al. 2013) and employ a hierarchical setup that considers each of the four sets of parameters as draws from four common population distributions: $\alpha_j \sim N(\mu_\alpha, \sigma_\alpha)$, $\beta_j \sim N(\mu_\beta, \sigma_\beta)$, $\theta_i \sim N(\mu_\theta, \sigma_\theta)$, and $\phi_j \sim N(\mu_\phi, \sigma_\phi)$. The full joint posterior distribution is thus

$$p(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma | \mathbf{y}) \propto p(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma, \boldsymbol{\mu}, \boldsymbol{\sigma}) \prod_{i=1}^n \prod_{j=1}^m \text{logit}^{-1}(\pi_{ij})^{y_{ij}} (1 - \text{logit}^{-1}(\pi_{ij}))^{1-y_{ij}} \prod_{j=1}^m N(\alpha_j | \mu_\alpha, \sigma_\alpha) \prod_{i=1}^n N(\beta_i | \mu_\beta, \sigma_\beta) \prod_{i=1}^n N(\theta_i | \mu_\theta, \sigma_\theta) \prod_{j=1}^m N(\phi_j | \mu_\phi, \sigma_\phi). \quad (3)$$

2.3 Identification

The model described by equation (1) 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 nonzero constant, with γ divided by its square, leaving the model predictions unchanged. These indeterminacies, which are common to item-response theory models, are sometimes called “additive aliasing” and “scaling invariance” (see, e.g., Bafumi et al. 2005; Londregan 1999).

Existing studies on ideal point estimation employ two different strategies to identify the model. One is to apply two linearly independent restrictions on the ideal point parameters, $\boldsymbol{\theta}$ or $\boldsymbol{\phi}$ in this case. In particular, the usual procedure is to constrain the ideal points of two legislators (liberal and conservative) at arbitrary positions, such as -1 and $+1$ (see, e.g., Londregan 1999; Clinton, Jackman, and Rivers 2004). An alternative is to apply a unit variance restriction on the set of ideal points, which in the multilevel setting would be equivalent to giving the θ_i s or ϕ_j s an informative $N(0, 1)$ prior distribution (Gelman and Hill 2007, 318). However, note that this second approach does not solve the problem of “reflection invariance”: the resulting scale can be reversed (flipped left-to-right) without changing the prediction of the models. As Jackman (2001) shows, choosing starting values that are consistent with the expected direction of the scale (e.g., liberals with -1 and conservatives with $+1$) is sufficient to ensure global identification in most cases. As I show in the supplementary materials, either of these strategies can be applied in this case to resolve the indeterminacies.

2.4 MCMC Algorithm

To improve the efficiency of the estimation procedure, I divide it into two stages. First, I use a No-U-Turn sampler, a variant of Hamiltonian Monte Carlo sampling algorithms (Gelman et al. 2013), to estimate the parameters indexed by j , using a random sample of ten thousand i users who follow at least ten j users. In the second stage, I use a random-walk Metropolis–Hastings algorithm (Metropolis et al. 1953) to estimate all parameters indexed by i . Note that each of these parameters can be estimated individually because I assume local independence, conditional on the j parameters, and therefore multi-core processors can be used to run multiple samplers simultaneously and dramatically increase computation speed.⁶

The first stage is implemented using the Stan modeling language (Stan Development Team 2012), whereas the second stage is implemented using R. I use flat priors on all parameters, with the exception of μ_θ , σ_θ , and μ_α , which are fixed to 0, 1, and 0, respectively, for identification purposes. The samplers in both stages are run using two chains with as many iterations as necessary to ensure that all ideology parameters have an effective number of simulation draws (Gelman and Rubin 1992) of at least 200. Each chain is initiated with random draws from a multivariate normal distribution for θ and γ , the logarithm of the number of followers of user j or number of friends of user i for α and β (to speed up convergence), and values of zero for ϕ , with the exception of those who belong to a party, -1 for left-wing politicians and $+1$ for right-wing politicians. This model-fitting strategy appears to be quite robust, and my results are largely insensitive to the choice of priors and initial values.⁷

2.5 Discussion

A key challenge in implementing this method is the choice of the m target Twitter users who are political elites: the set of users with “discriminatory” predictive power such that the decision to follow them (or not) provides information about an individual’s ideology. Following Conover et al. (2010), we could analyze the entire Twitter network and let the different clusters emerge naturally. However, “homophilic” networks can be based not only on political traits, but also on other personal characteristics. Instead, the approach I use is to select a limited number of target users that includes politicians, think tanks, and news outlets with a clear ideological profile that span the full range of the ideological spectrum. The set of users that are considered will determine the interpretation of the latent scale where ideal points are located and, for this reason, it is important to include identifiable figures with extreme ideological positions, beyond just partisans.⁸

3 Data

The estimation method I propose in this article can be applied to any country where a high number of citizens are discussing politics on Twitter.⁹ However, in order to test the validity of the estimated parameters, I will focus on six countries where high-quality ideology measures are available for a subset of all Twitter users: the US, the UK, Spain, Germany, Italy, and the Netherlands.

⁶Samples 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. In all the examples in this article, the correlation between these two sets of estimates is $\rho = .999$.

⁷See Section D in the supplementary materials for the code to estimate the model in Stan, as well as results of a battery of tests that assess model fit.

⁸Note that the model is agnostic regarding the interpretation of the latent dimension, which will depend on the set of m political actors that are considered. As I show in Section 4.1, the results from multiparty systems clearly show that this dimension overlaps with the left-right scale. In the United States, where partisanship and ideology are highly correlated, it is not as clear. However, the fact that Twitter-based ideal points are highly correlated with DW-NOMINATE scores (commonly thought to capture legislators’ ideology), and that state-level estimates are better predictors of survey-based measures of ideology rather than partisanship, also suggest that the estimated dimension is the liberal-conservative scale.

⁹Estimating ideal points using data from different countries simultaneously is more complex, given the high intra-country locality effect (Gonzalez et al. 2011), which limits the number of Twitter users who could serve as “bridges” across countries in the estimation.

Furthermore, the increasing complexity of the party system in each of these countries will show how the method performs as the number of parties increases.

For each of these countries, I identified a set of political actors with visible profiles on Twitter: (1) all political representatives in national-level institutions; (2) political parties with accounts on Twitter; and (3) media outlets and journalists who tweet about politics. I considered only political Twitter users with more than five thousand (United States) or two thousand (UK, Spain, Italy, Germany, the Netherlands) followers. This represents a total of $m = 318$ target users in the United States, $m = 244$ in the UK, $m = 298$ in Spain, $m = 214$ in Italy, $m = 273$ in Germany, and $m = 118$ in the Netherlands.¹⁰

Next, using the Twitter REST API, I obtained the entire list of followers for all m users in each country, resulting in an entire universe of Twitter users following at least one politician of $n = 32,919,418$ in the United States, $n = 2,647,413$ in the UK, $n = 1,059,890$ in Spain, $n = 1,119,763$ in Italy, $n = 1,559,311$ in Germany, and $n = 856,201$ in the Netherlands.¹¹ However, an extremely high proportion of these users are either inactive, are spam bots, or reside in different countries. 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 fewer than one hundred tweets; (2) have not sent one tweet in the past six months; (3) have less than twenty-five followers; (4) are located outside the borders of the country of interest; and (5) follow less than three political Twitter accounts.¹² The final sample size is $n = 301,537$ users in the United States, $n = 135,015$ in the UK, $n = 123,846$ in Spain, $n = 150,143$ in Italy, $n = 49,142$ in Germany, and $n = 96,624$ in the Netherlands.¹³

This is a highly self-selected sample because Twitter users are not a representative sample of the population.¹⁴ In addition, the inferences I make based on our sample would not represent the full set of Twitter users, as I am selecting only those users who follow three or more political accounts. However, this should not affect the inference of *politicians'* ideal points, since these users can indeed be considered as "authoritative" when it comes to politics. Precisely because they are more likely to be knowledgeable and interested in politics than the average citizen, examining their online behavior can be highly informative about policy positions. This procedure is roughly analogous to an expert survey with many respondents where each respondent provides a small amount of information that, when aggregated, results in highly accurate policy estimates.

4 Results and Validation

In this section, I provide a summary of the ideology estimates for the six countries included in my study. 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.

My analysis is divided into three parts, with each of them providing a different type of evidence to the validation. The first part shows that Twitter-based ideal points replicate existing measures of ideology for elites (legislators and political parties) in six different countries. Then, I validate mass ideology at the aggregate level by examining groups of Twitter users by self-identified ideology and state of residence. Here, I am also able to replicate previous findings in the literature about elite and

¹⁰See the supplementary materials for additional details on the data collection. Full replication files are available in Barberá (2014).

¹¹As of November 2012 in the United States, Spain, the Netherlands, and the UK; February 2013 in Italy; August 2013 in Germany.

¹²In the US sample, I further restricted the sample to accounts that tweeted at least three times mentioning "Obama" or "Romney" during the three months before the 2012 election, in order to include only users who tweet frequently about politics.

¹³Note that the sample selection process requires identifying the specific country in which each user is located. This information was inferred from the "time zone" and "location" fields in the user profile, which was 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 United States (71%).

¹⁴Table 1 in the supplementary materials shows that Twitter users in the United States tend to be younger and to have a higher income level than the average citizen, and their educational background and racial composition are different from that of the entire population.

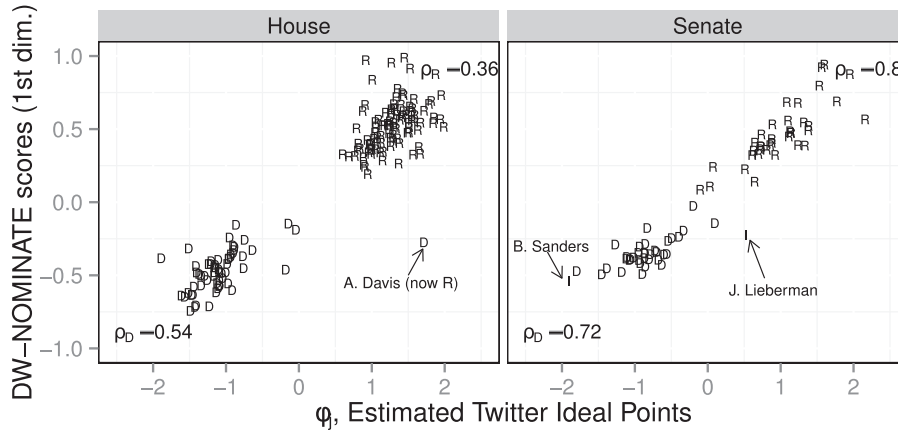


Fig. 1 Ideal point estimates for members of US Congress.

mass polarization. Finally, I also validate mass ideology at the individual level using campaign contribution records and information about voters' party registration history.

4.1 Replication of Legislators' and Parties' Ideal Points

The first set of results I focus on are those from the United States. Figure 1 compares ϕ_j , the ideal point estimates, of 231 members of the 112th US Congress¹⁵ based on their Twitter network of followers (y -axis) with their DW-NOMINATE scores,¹⁶ based on their roll call voting records (Poole and Rosenthal 2007), on the x -axis. Each letter corresponds 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.

As we can see, the estimated ideal points are clustered into two different groups that align almost perfectly with party membership. The correlation between Twitter- and roll-call-based ideal points is $\rho = 0.941$ in the House and $\rho = 0.954$ 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: $\rho = 0.546$ for Republicans, $\rho = 0.610$ for Democrats.¹⁷

Figure 2 displays ideal points for a set of key political actors. (See supplementary materials for an expanded version of this plot.) As was the case with members of Congress, the resulting estimates show a clear division across members of each party, and the ideal points for all nonpartisan actors have face validity. Note also that their positions within each cluster are what we would expect based on anecdotal evidence. For example, Jon Huntsman appears among the most liberal Twitter accounts in the Republican Party, whereas Rush Limbaugh and Glenn Beck are in the group of the most conservative nonpartisan Twitter accounts. On the left side of the ideological dimension we find Michael Moore, Rachel Maddow, and the Human Rights Campaign (HRC) among the most liberal Twitter accounts.

Turning to the results in the five European countries I consider, Fig. 3 shows that Twitter-based estimates of the locations of political elites are congruent with other measures based on surveys of experts. The left panel displays the Twitter-based ideal points for members of the most important political parties in the UK, Spain, the Netherlands, Germany, and Italy. Each dot represents the

¹⁵As explained in Section 3, only Twitter accounts with more than five thousand followers as of November 2012 are included in the sample.

¹⁶Source: voteview.com.

¹⁷These results are essentially identical if I compare Twitter-based estimates with other ideal points based on voting records, such as those estimated by Jackman (2012) using an item-response theory scaling method (Clinton, Jackman, and Rivers 2004): $\rho = .962$ in the House, $\rho = .936$ in the Senate, $\rho = .542$ for Republicans, and $\rho = .675$ for Democrats.

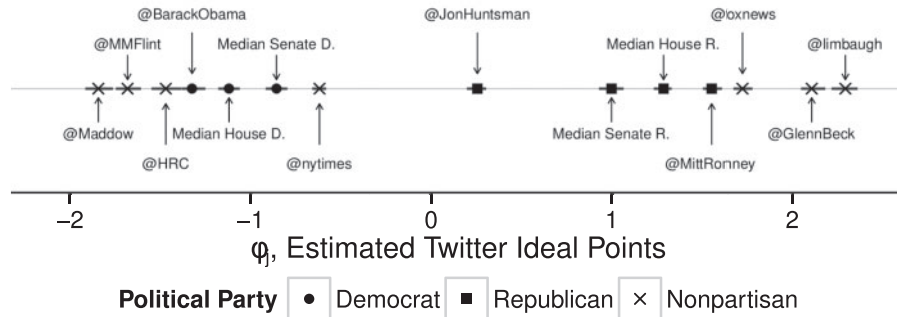


Fig. 2 Ideal point estimates for key political actors in the United States.

position of an individual that belongs to each party: a public official (e.g., the president or a member of the Cabinet), a candidate, or the main Twitter account of the party. The vertical lines indicate the location of the median user within each party. To facilitate the comparison, the right panel shows estimates for the same parties according to the 2010 Chapel Hill Expert Survey (Bakker et al. 2012).

As in the case of the United States, this set of results shows that my estimation method is able to classify accounts according to the party to which they belong. With few exceptions, all Twitter accounts from the same party are clustered together, and parties with similar ideologies are located in similar positions on the latent ideological dimension. Furthermore, the order of the parties seems to be similar to that reported by different studies based on expert surveys for the “left-right” dimension.

The results show the lower degree of accuracy in the UK: both the Labor Party and the Liberal Democrats are located to the left of the Conservative Party on average, but the latter is classified as right-wing, almost overlapping with the conservatives, perhaps indicating their status as coalition partners. In Spain, the Socialist Party (PSOE) is located to the left of the Conservative Party (PP), with the recently created Center Party (UPyD) between them, and the Communist Party to the left of all three, although overlapping with the PSOE. In the Netherlands, all parties are accurately scaled, with the exception of the Socialist Party (SP) and the right-wing Party for Freedom (PVV), which appear slightly more centrist in my Twitter estimates than in the expert surveys. In Germany, the locations of the three right-wing parties (CDU, CSU, and FDP) on the ideological dimension are indistinguishable, which is consistent with the experts’ assessments, but Die Linke appears to be more centrist. In Italy, the two largest parties are also correctly classified on the left (PD) and right (PDL) of the ideological spectrum.

It is 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 homogeneous scale across countries. However, it would be possible to overcome this difficulty under certain assumptions. For example, following Lo, Proksch, and Gschwend (2013), we could exploit the membership of national parties in political groups inside the European Parliament. Alternatively, it may be possible to use Twitter users who follow accounts from more than two countries as “bridges” across ideological scales. This procedure would be similar to the method developed by Bonica (2014) to scale legislators across different states using campaign contributions from donors who contribute to candidates in multiple states.

It is also necessary to explore why this method performs differently across countries, particularly when estimating ideal points for parties on the extremes of the distributions. One possibility is that the emerging scale collapses different dimensions. As it is formulated in this article, the latent dimension that emerges from the model is the one that captures the largest share of the variation in users’ following decisions. This is probably the reason why the model performs well in the US case, where economic and social left-right positions are highly correlated and policy preferences are mostly unidimensional (Poole and Rosenthal 2007). But that is not the case in most European countries, where the policy space is usually conceived as two-dimensional (see, e.g., Benoit and

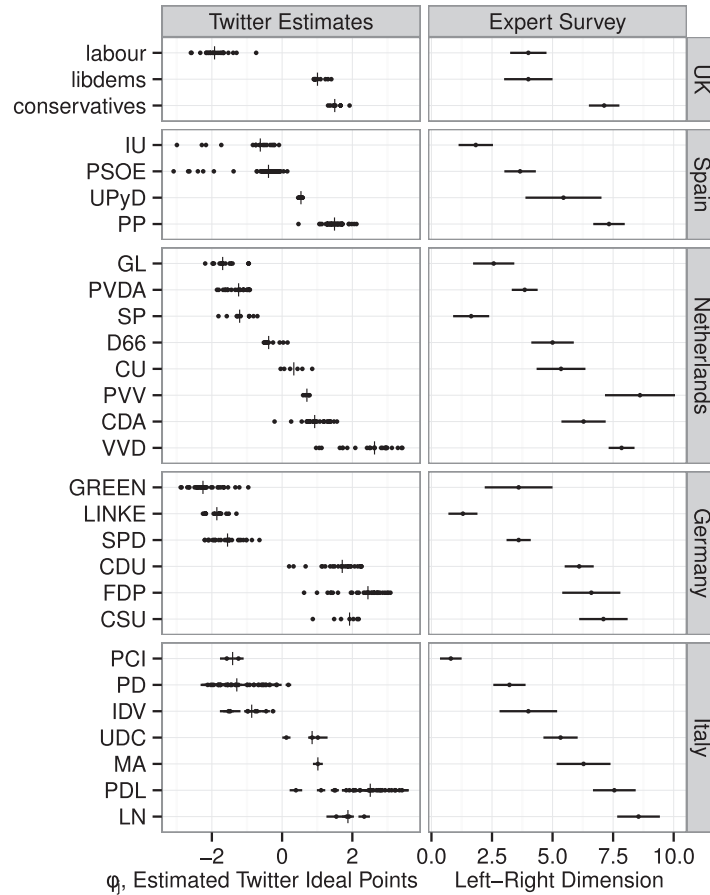


Fig. 3 Ideological location of parties in five European countries.

Laver 2006).¹⁸ In these cases, a model that is able to locate users' and politicians' positions on two different dimensions would be more appropriate, although it would also imply additional identifying assumptions (Jackman 2001).

4.2 Validation of Mass Ideology at the Aggregate Level

Now I assess whether the estimated ideal points for ordinary citizens are also valid. The left panel of Fig. 4 compares the distribution of ideal points for the two types of Twitter users in the sample—political actors and ordinary citizens. The pattern that emerges is consistent with previous findings in the literature on elite and mass political polarization in the United States (Fiorina and Abrams 2008). In fact, this figure is an almost exact replication of Fig. 5 in Bafumi and Herron (2010).¹⁹ Four features stand out: both distributions are bimodal, liberal citizens represent a majority of the sample population, political actors are more polarized than mass voters, and the distribution of mass ideology has longer tails.

The right panel displays the distribution of ideology estimates for different groups of individuals. Here, I exploit the fact that many Twitter users define themselves politically in their profiles. Using

¹⁸Note that this should not affect the estimation as long as these two dimensions are orthogonal to each other, in which case the estimated latent space would correspond to the dimension that explains a larger proportion of the variance.

¹⁹Bafumi and Herron (2010) estimate ideology for the mass public using the 2006 Cooperative Congressional Election Study ($N=33,000$), which included questions asking respondents to take positions on roll call votes as if they were members of Congress. These responses were then used to estimate ideology using an item-response theory model (Jessee 2009).

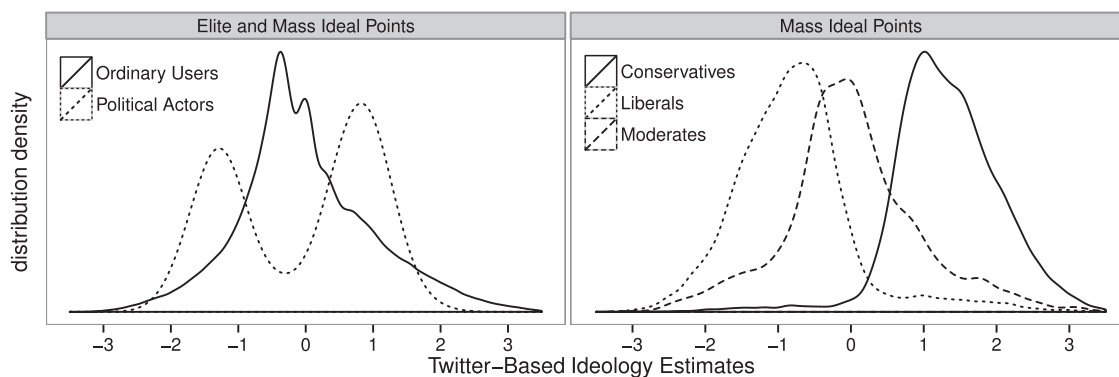


Fig. 4 Distribution of political actors and ordinary Twitter users' ideal points.

this information, I defined three subsets of accounts, according to whether they mention specific keywords on their profiles: conservatives (“conservative,” “GOP,” “Republican”), moderates (“independent,” “moderate”), and liberals (“liberal,” “progressive,” “Democrat”). 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, with some overlap.

I also compared the estimated ideal points for the median Twitter user in each state with the proportion of citizens in each state that hold liberal opinions across different issues, as estimated by Lax and Phillips (2012), combining surveys and demographic indicators. The correlation between these two variables is $\rho = 0.876$, and it is higher than the correlation with the proportion of the two-party vote for Obama in 2012 ($\rho = 0.785$), which suggests that the meaning of the emerging dimension in my estimation is closer to ideology than to partisanship.

4.3 Validation of Mass Ideology at the Individual Level

Finally, in order to further examine the validity of the ideal points I estimate in the United States, I matched a sample of Twitter users with their campaign contribution records (Bonica 2013b) and, for those in Ohio, with their voter registration records.²⁰ A total of 2707 and 2462 Twitter users were successfully matched in each case, based on perfect nonduplicated matches of first and last name, and zip code or county. Again, it is not critical that this subset cannot be considered representative for any population of interest, as it is used only to examine whether the resulting ideology estimates are good predictors of campaign contributions and the party under which they are registered.

I find that Twitter-based ideology estimates are good predictors of campaign contributions. In all, 90% of the 1058 matched Twitter users on the right of the median voter donated to Republican candidates, whereas 98% of the 1649 matched users on the left donated to Democratic candidates. The correlation between ideal points estimated using Twitter networks and contribution records (Bonica 2014) is $\rho = 0.80$.

Figure 5 compares the estimated ideal points with partisan registration for the sample of Twitter users matched to the Ohio voter registration file. The panel on the left shows the results for the most recent available data (2012). I compare the distribution of estimated ideal points for registered Democrats to the distribution of estimated ideal points for registered Republicans. The mean ideal point for registered Republicans is higher (more conservative) than the mean estimated ideal point for registered Democrats. And in fact most registered Republicans have estimated ideal points to the right of the estimated ideal points of most registered Democrats.

Since each voter's registration history is available since 2000, I can also examine whether the most conservative voters in Ohio tend to consistently register as Republicans, and vice versa. The

²⁰The entire voter file is available online at the Ohio Secretary of State's Web site, www.sos.state.oh.us.

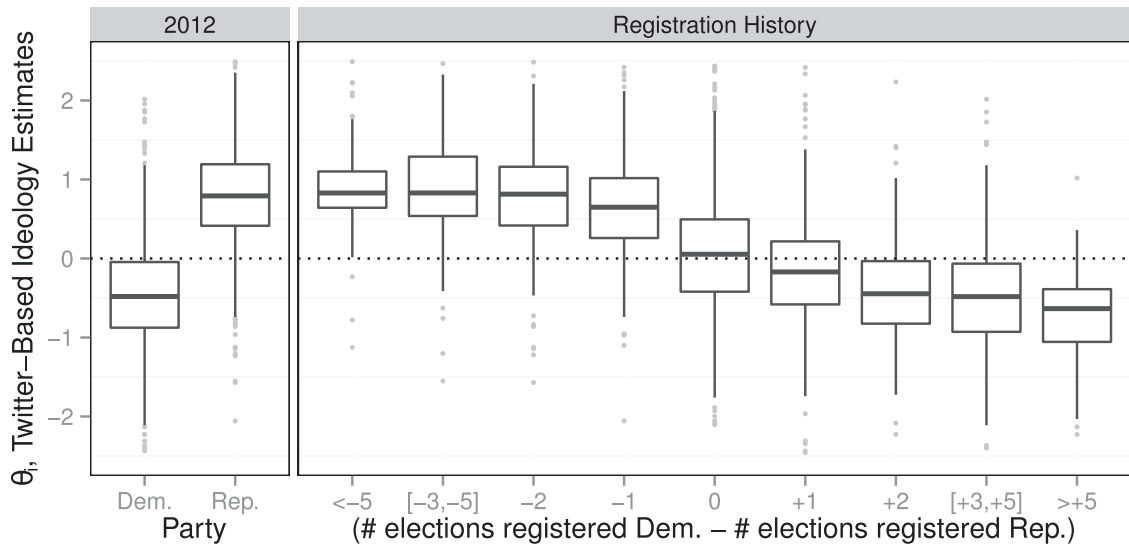


Fig. 5 Ideal point estimates and party registration history.

panel on the right, which classifies voters according to their registration history (the number of times they registered as Democrat minus the number of times they registered as Republican), indicates that this is indeed the case. Users who registered at least five times more often as Democrats than Republicans since 2000 have substantially more liberal ideal points than users who registered more often as Republican than as Democrats.

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

A recurring theme in the literature on the Internet and politics is how the increasing amount and heterogeneity of political information citizens have access to affects their political views (Farrell 2012). 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; Iyengar and Hahn 2009). This generates a so-called echo-chamber environment that fosters social extremism and political polarization. Given that a substantial proportion of citizens now rely mostly on the Internet to gather political information,²¹ 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 et al. (2010, 2011) 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 et al. 2011, 89). Yardi and Boyd (2010) and Gruzd and Roy (2014) qualify this conclusion. Their results show that open cross-ideological exchanges are frequent, and individuals are exposed to broader viewpoints.

As a substantive application of the estimation method I propose in this article, I re-examine the question of whether political discussion is polarized. By using the method developed in this article, I am able to compute ideal points for a far larger sample of users than any of the previous studies, which now includes most active Twitter users in the United States, and with a much higher degree of precision. In order to do so, I collected all public tweets that mentioned "Obama" or "Romney" from August 15 to Election Day. These messages total over 65 million tweets sent by 8.8 million

²¹According to a survey conducted by the Pew Research Center in 2011, 31% of adults in the United States rely mainly on the Internet for political information.

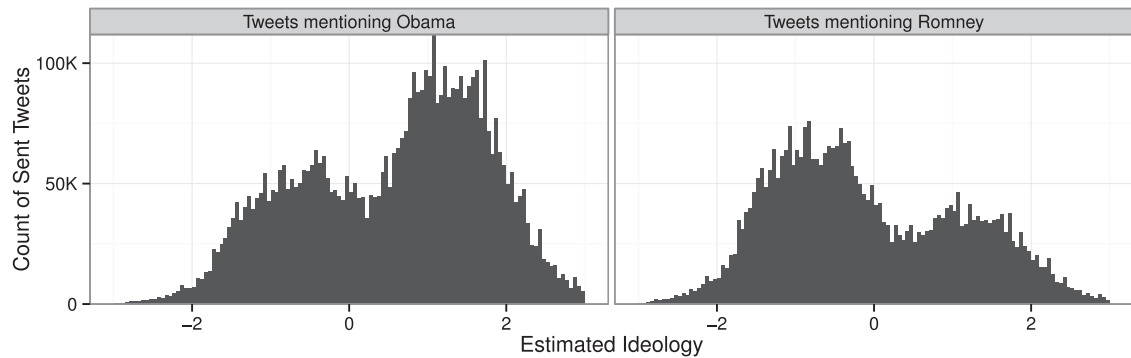


Fig. 6 Number of tweets mentioning presidential candidates, by ideal point bin.

users and were collected using the Twitter Streaming API and the `streamR` package for R (Barberá 2013).

In the analysis below, I focus on the 15 million tweets that were sent by the 301,537 individuals in my sample of politically interested users.²² Figures 6 and 7 display my main results. 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 left panel refers to tweets that mention “Obama,” whereas the right panel refers to tweets mentioning “Romney.”

I find that the conversation on Twitter is dominated by individuals with strong political views: the distribution of the number of tweets is highly bimodal, with the modes at approximately -1 and $+1$ —that is, one standard deviation away from the average Twitter user. Moderate individuals are less active than those with more extreme ideological positions, even if they are a majority of the population: users with ideal points between -0.5 and 0.5 represent almost 40% of the sample considered in the study, but they sent only 21% of the tweets included in the collection. A distinct pattern emerges in the tweets mentioning President Obama: conservative Twitter users sent a substantially higher proportion of tweets than their liberal counterparts. This result suggests that most tweets sent during this period were negative. Overall, conservative users are more active: users with ideal points between 0.5 and 1.5 sent an average of 118 tweets, whereas those in the interval between -1.5 and -0.5 sent 82 tweets on average.

Figure 7 provides evidence for the existence of an “echo-chamber” environment on Twitter. Here, I use a heatmap to visualize the structure of the most common type of interactions on Twitter: retweets.²³ The shading of each cell (of size 0.2×0.2) represents the proportion of retweets in the sample that were originally sent by users with ideal point X and subsequently retweeted by users with ideal point Y . Therefore, if we were to find perfect polarization (i.e., users interacting only with those of identical ideology), we would find a pattern that would resemble a line with a slope coefficient equal to one.

Figure 7 shows that cross-ideological interactions are rare, since 85% of retweet interactions take place among Twitter users with similar ideological positions. After dividing each panel into four parts, according to whether they are to the left or right of the average Twitter user (with ideal point 0), I find that 20% of the retweets mentioning Obama (and 49% of the retweets mentioning Romney) are located in the bottom-left quadrant, whereas 69% of the retweets mentioning Obama (and 31% of the retweets mentioning Romney) are located in the top-right quadrant. Political polarization is particularly intense among right-leaning Twitter users, as indicated by the darker shading. Although liberal users also present this pattern, they tend to engage more often in conversations all along the ideological spectrum.²⁴ To sum up, the picture that emerges

²²Note that although this sample of users includes 3.5% of all unique users in the collection of tweets, they sent 23% of all tweets, and thus are a good representation of the political discussion that takes place on Twitter.

²³See Fig. 8 in the supplementary materials for a replication of this analysis using mentions instead of retweets.

²⁴For example, of all the retweets by liberal Twitter users (with ideal points lower than -0.5 in the latent scale), 18% were originally sent by moderate and conservative users (with ideal points higher than -0.5). In the case of conservatives (with ideal points 0.5 or higher), only 7% of their retweets were tweets from liberal or moderate users.

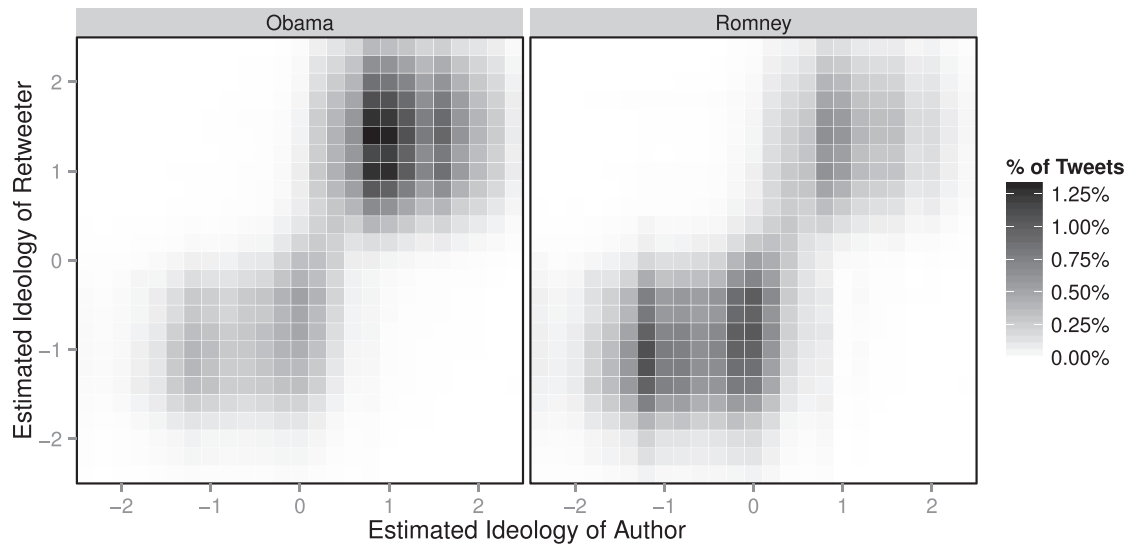


Fig. 7 Political polarization in retweets mentioning presidential candidates.

points toward a high degree of polarization, which is driven predominantly by conservative Twitter users.²⁵ The strength of their ties has important implications during the electoral campaign. As Conover et al. (2012) argue, the topology of the network of right-leaning Twitter users facilitates the rapid and broad dissemination of political information, and increases their potential to manipulate the public agenda.

6 Conclusions

Millions of people are writing personal messages on Twitter every day. Many of these “tweets” are either politically irrelevant personal updates, replication of existing information, or simply spam. However, given the number and heterogeneity of users, some valuable data can be extracted from this source. That was the motivation behind this article.

I have proposed a new method to estimate ideology for a large number of political elites and ordinary citizens that relies on the “following” links created across them. I have applied this measure in six different countries: United States, United Kingdom, Spain, Italy, Germany, and the Netherlands. My results show that this method successfully classifies political actors and ordinary citizens according to their political orientation, with the locations along the ideological scale being verified by comparisons to positions estimated using roll call voting, party manifestos, and expert surveys. I have also validated the results by comparing the estimated ideology of a subset of individual Twitter users with their official party registration records.

These results highlight the unexplored potential of Twitter data to generate ideology estimates that could prove useful in exploring political behavior. To illustrate this possibility, I have presented an application that relies on such estimates. Using the 2012 US presidential election campaign as a case study, I have shown that public exchanges on 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.

This application is just one of many research questions that could be examined exploiting the opportunities of this new source of information. For instance, the analysis in Section 5 could easily

²⁵It may be argued that this conclusion is driven at least in part by the key assumption behind the method to estimate ideology, since users are more likely to retweet someone they follow. However, note that retweets between users who do not follow each other are indeed frequent. For example, of the 460,000 users who retweeted at least one of Barack Obama’s tweets between August 15, 2012, and November 6, 2012, 152,000 (33%) did not follow him as of November 5, 2012. In fact, 21.2 of his 21.5 million followers at the time did not retweet any of his tweets during this period.

be extended to other countries to examine the determinants of elite and mass political polarization from a comparative perspective. Relying on Twitter users who follow accounts in more than one country, or under other possible “bridging” assumptions, such as shared party groups in the European Parliament, it would also be possible to make meaningful comparisons about the ideological positions of parties and voters in different countries, which could increase the interest of such comparative studies.

This measurement tool also represents a relevant contribution to the emerging literature on social media and politics. As citizens spend a greater proportion of their time online, most of their political behavior also takes place through social networking platforms. Being able to estimate covariates of interest represents an important step in the study of such behavior, since most existing work relies on crude measures of political preferences.

In this article, I estimated ideal points for a specific time period, but the granularity of social media data could be exploited to develop dynamic ideology estimates (Martin and Quinn 2002). This new source of information has the potential to generate much value added to explanatory models of political behavior at a low cost. For example, it could enable the study of short-term changes in political ideology in response to campaign events without the need to field panel surveys, which usually require weeks of preparation. It also opens up the possibility of conducting future research that could be done only with this kind of social media data, such as the study of how ideological coalitions emerge and evolve on rapidly changing issues.

The statistical model as formulated in this article can be used to estimate the ideology only of individuals who follow at least one political account. However, a third stage in the estimation could be added to compute ideal points for *all* Twitter users based on which politically interested users they choose to follow (even if they do not follow any political account), which in practice would allow us to estimate dynamic ideal points for millions of individuals across multiple countries. In all, 4.3 million US citizens made a political contribution in the 2012 electoral cycle (Bonica 2013a), but 15% of online adults (almost forty million citizens) are active on Twitter.²⁶ Furthermore, as I have shown in this article, it is possible to identify the location of most Twitter users, as well as their gender, and to track their behavior over time, which could help recover the representativeness of this sample of individuals.

A final advantage of Twitter as a source of data is the fact that all types of actors are present on this platform. One important reason why the work by Jessee (2009), Shor, Berry, and McCarty (2010), and Bonica (2013b) represents an important contribution to the measurement literature is that ideology estimates for citizens, state legislators, political action committees (PACs), and political candidates were not previously available. Of course, Twitter users are a self-selected minority of the population, but this minority includes virtually all political parties and candidates, public officials, interest groups, media outlets, journalists, and many others for whom we do not necessarily have good measures of ideology. Exploiting this new set of ideal points can represent a significant contribution to studies on a wide array of political phenomena, ranging from policy processes to media effects.

Funding

The present work has been supported by the National Science Foundation (Award #1248077) and the “La Caixa” Fellowship Program.

References

- Bafumi, J., A. Gelman, D. K. Park, and N. Kaplan. 2005. Practical issues in implementing and understanding Bayesian ideal point estimation. *Political Analysis* 13(2):171–87.
- Bafumi, Joseph, and Michael C. Herron. 2010. Leapfrog representation and extremism: A study of American voters and their members in Congress. *American Political Science Review* 104(3):519–42.

²⁶Source: The Pew Research Center’s Internet and American Life Project, August 2013.

- Bakker, Ryan, Catherine de Vries, Erica Edwards, Liesbet Hooghe, Seth Jolly, Gary Marks, Jonathan Polk, Jan Rovny, Marco Steenbergen, and Milada Anna Vachudova. 2012. Measuring party positions in Europe: The Chapel Hill expert survey trend file, 1999–2010. *Party Politics*.
- Barberá, P. 2013. *streamR: Access to Twitter Streaming API via R*. R package available on CRAN. <http://CRAN.R-project.org/package=streamR>.
- . 2014. *Replication data for: Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data*. Dataverse Network, <http://dx.doi.org/10.7910/DVN/26589>.
- Benoit, Kenneth, and Michael Laver. 2006. *Party policy in modern democracies*. London: Routledge.
- Bonica, Adam. 2013a. *Database on ideology, money in politics, and elections*. <http://data.stanford.edu/dime>.
- . 2013b. Ideology and interests in the political marketplace. *American Journal of Political Science* 57(2):294–311.
- . 2014. Mapping the ideological marketplace. *American Journal of Political Science* 58:367–86.
- Boutet, A., H. Kim, E. Yoneki. 2012. What's in your Tweets? I know who you supported in the UK 2010 General Election. *Proceedings of Sixth International AAAI Conference on Weblogs and Social Media*. Palo Alto, California.
- Bryant, J., and D. Miron. 2004. Theory and research in mass communication. *Journal of Communication* 54(4):662–704.
- Clinton, J., S. Jackman, and D. Rivers. 2004. The statistical analysis of roll call data. *American Political Science Review* 98(2):355–70.
- Conover, M., B. Gonçalves, J. Ratkiewicz, A. Flammini, and F. Menczer. 2010. Predicting the political alignment of Twitter users. *Proceedings of 3rd International Conference on Social Computing*. Bethesda, Maryland.
- Conover, Michael D., Bruno Gonçalves, Alessandro Flammini, and Filippo Menczer. 2012. Partisan asymmetries in online political activity. *EPJ Data Science* 1(1):1–19.
- Conover, M. D., J. Ratkiewicz, M. Francisco, B. Goncalves, A. Flammini, and F. Menczer. 2011. Political polarization on Twitter. *Proceedings of 5th International Conference on Weblogs and Social Media*. Menlo Park, California.
- Enelow, J. M., and M. J. Hinich. 1984. *The spatial theory of voting: An introduction*. New York, New York: Cambridge University Press.
- Farrell, H. 2012. The Internet's consequences for politics. *Annual Review of Political Science* 15:35–52.
- Fiorina, Morris P., and Samuel J. Abrams. 2008. Political polarization in the American public. *Annual Review of Political Science* 11:563–88.
- Garrett, R. K. 2009. Politically motivated reinforcement seeking: Reframing the selective exposure debate. *Journal of Communication* 59(4):676–99.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2013. *Bayesian data analysis*. 3rd ed. Boca Raton, Florida: Chapman & Hall/CRC.
- Gelman, A., and J. Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*. New York, New York: Cambridge University Press.
- Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. *Statistical Science* 7(4):457–72.
- 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., and J. Roy. 2014. Investigating political polarization on Twitter: A Canadian perspective. *Policy & Internet* 6:28–45.
- Hoff, Peter D., Adrian E. Raftery, and Mark S. Handcock. 2002. Latent space approaches to social network analysis. *Journal of the American Statistical Association* 97(460):1090–98.
- Iyengar, Shanto, and Kyu S. Hahn. 2009. Red media, blue media: Evidence of ideological selectivity in media use. *Journal of Communication* 59(1):19–39.
- Jackman, Simon. 2001. Multidimensional analysis of roll call data via Bayesian simulation: Identification, estimation, inference, and model checking. *Political Analysis* 9(3):227–41.
- . 2012. *Estimates of Members' Preferences, 112th U.S. House and Senate*. Retrieved on August 3, 2012, from <http://jackman.stanford.edu/blog/>.
- Jessee, S. A. 2009. Spatial voting in the 2004 presidential election. *American Political Science Review* 103(1):59–81.
- King, A., F. Orlando, and D. B. Sparks. 2011. Ideological extremity and primary success: A social network approach. Chicago, Illinois: Paper presented at the 2011 MPSA Conference.
- Kwak, H., C. Lee, H. Park, and S. Moon. 2010. What is Twitter, a social network or a news media? *Proceedings of 19th International Conference on World Wide Web*. ACM, 591–600.
- Lax, Jeffrey, and Justin Phillips. 2012. The democratic deficit in the States. *American Journal of Political Science* 56(1):148–66.
- Lazarsfeld, P. F., B. Berelson, and H. Gaudet. 1944. *The people's choice: How the voter makes up his mind in a presidential election*. New York: Duell, Sloan and Pearce.
- Linden, W. J., and R. K. Hambleton. 1997. *Handbook of modern item response theory*. New York: Springer.
- Lo, James, Sven-Oliver Proksch, and Thomas Gschwend. 2013. A common left-right scale for voters and parties in Europe. *Political Analysis*. Forthcoming.
- Londregan, John. 1999. Estimating legislators' preferred points. *Political Analysis* 8(1):35–56.
- Martin, Andrew D., and Kevin M. Quinn. 2002. Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999. *Political Analysis* 10(2):134–53.
- McPherson, M., L. Smith-Lovin, and J. M. Cook. 2001. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27:415–44.

- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. *Journal of Chemical Physics* 21:1087–92.
- Mutz, D. C. 2002. Cross-cutting social networks: Testing democratic theory in practice. *American Political Science Review* 96(1):111–26.
- Oken Hodas, N., and K. Lerman. 2012. How visibility and divided attention constrain social contagion. College Park, Maryland: Proceedings of 2012 International Conference on Social Computing.
- Poole, K. T., and H. Rosenthal. 1997. *Congress: A political-economic history of roll call voting*. New York: Oxford University Press.
- . 2007. *Ideology and Congress*. 2nd ed. New Brunswick, New Jersey: Transaction Pub.
- 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):417–48.
- Stan Development Team. 2012. Stan modeling language: User's guide and reference manual. Version 1.0.
- Stroud, Natalie Jomini. 2008. Media use and political predispositions: Revisiting the concept of selective exposure. *Political Behavior* 30(3):341–66.
- Sunstein, C. 2001. *Republic.com*. Princeton, New Jersey: Princeton University Press.
- Wu, S., J. M. Hofman, W. A. Mason, and D. J. Watts. 2011. Who says what to whom on Twitter. Proceedings of 20th International Conference on Worldwide Web. ACM, 705–14.
- Yardi, S., and D. Boyd. 2010. Dynamic debates: An analysis of group polarization over time on Twitter. *Bulletin of Science, Technology & Society* 30(5):316–27.