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

SUPPLEMENTARY MATERIALS

A. LITERATURE REVIEW

Despite ideology being one of the key predictors of political behavior, its 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 et al. (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 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 or 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 us to infer political orientation even in the absence of any information about the user. Similarly, the only political science study (to my knowledge) that aims at measuring ideology (King, Orlando and Sparks, 2011) uses this type of information. These authors apply a data-reduction technique to the 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 need to solve the potentially severe problem of disambiguation caused by contractions designed to fit the 140-character limit, and are vulnerable to the phenomenon of 'content injection.' As Conover et al. (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.

In conclusion, a critical reading of the literature suggests 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. In order to draw correct inferences, it is also important to indicate the uncertainty of the estimates. Most importantly, none of these studies explores the possibility of placing ordinary citizens and legislators on a common scale or whether this method would generate valid ideology estimates outside of the US context.

B. DATA SOURCES

The lists of Twitter accounts included in the analysis were constructed combining information from different sources. In the US, I have used the NY Times Congress API, complemented with the GovTwit directory. In the UK, I have used lists of political accounts compiled by Tweetminster. In Spain, I have used the Spanish Congress Widget developed by Antonio Gutierrez-Rubi, and the website politweets.es. In Italy and Germany, I used a list of political Twitter users collected by two local experts, to whom I express my gratitude. In the Netherlands, I have used the data set from politiekentwitter.nl.

In the case of the US, this list includes, among others, the Twitter accounts of all Members of Congress, the President, the Democratic and Republican parties, candidates in the 2012 Republican primary election (`@THEHermanCain`, `@GovernorPerry`, `@MittRomney`, `@newtgingrich`, `@timpawlenty`, `@RonPaul`), relevant political figures not in Congress (`@algore`, `@ClintonTweet`, `@SarahPalinUSA`, `@KarlRove`, `@Schwarzenegger`, `@GovMikeHuckabee`), think tanks and civil society group (`@Heritage`, `@HRC`, `@democracynow`, `@OccupyWallSt`), and journalists and media outlets that are frequently classified as liberal (`@nytimes`, `@msnbc`, `@current`, `@KeithOlbermann`, `@maddow`, `@MotherJones`) or conservative (`@limbaugh`, `@glennbeck`, `@FoxNews`). A similar approach was adopted in the other five countries of study. Note that my purpose is not to collect an exhaustive list of all relevant political Twitter accounts, but rather focus on a set of users such that following them is informative about ideology.

C. ADDITIONAL RESULTS

Table 1 summarizes the distribution of demographic characteristics of Twitter users in the U.S., as well as of the population, all online adults, and politically interested Twitter users. Twitter users in the U.S. 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

that of the entire population. (See also Mislove et al., 2011; Parmelee and Bichard, 2011.)

Table 1: Sociodemographic characteristics of Twitter Users in the U.S.

	Population (Census)	Online adults	Twitter users	Pol. interested Twitter users
Average age	46.2	43.2	34.2	40.1
		(42.8, 43.6)	(33.4, 34.9)	(38.5, 41.8)
% female	51.1%	51.0%	50.4%	40.6%
		(49.9, 52.1)	(47.6, 53.2)	(34.4, 46.8)
Income over $50K	45.0%	50.3%	53.0%	54.2%
		(49.1, 51.4)	(50.0, 56.0)	(47.7, 60.7)
% w. college degree	39.1%	44.9%	49.8%	54.6%
		(43.8, 46.0)	(47.0, 52.7)	(48.2, 60.8)
% white	68.1%	69.7%	64.4%	47.9%
		(68.7, 70.8)	(61.7, 67.1)	(41.6, 54.2)
% African-American	11.5%	10.6%	12.3%	29.8%
		(9.9, 11.3)	(10.4, 14.1)	(24.0, 35.5)
Sample size		2,498	324	72

Source: Pew Research Center Poll on Biennial Media Consumption, June 2012, weighted. Descriptive statistics refer to entire U.S. population, according to Pew Research Center estimates based on the 2000 Census (Column 1), 83.3% of adults who use the internet at least occasionally (Column 2), 15.1% of online adults who ever use Twitter (Column 3), and 3% of "politically interested" online adults who use Twitter and read blogs about politics regularly (Column 4). 95% confidence intervals in parentheses.

Figure 1 compares the distribution of ideal points by gender in the sample of Twitter users in the U.S., showing that women tend to be slightly more liberal than men. This result is consistent with what can be found in political surveys. For example, the average ideological placement (in a scale from 1, extremely liberal, to 7, extremely conservative) in the 2008 American National Election Survey was 4.05 for women and 4.24 for men. Gender was estimated using a Naive Bayes classifier (Bird, Klein and Loper, 2009) based on their first name (when available on their profile), relying on a list of common first names by gender in anonymized databases (Betebenner, 2012) as a training dataset. The accuracy of this classifier, computed on a random sample of 500 manually labeled Twitter profiles, is 75.6%. The distribution of users by gender was: 50.2% male (151,497 users), 35.1% female (105,811 users), and 14.7% unknown (44,299 users); which matches the survey marginals

of politically interested Twitter users in Table 1.

Figure 1: Distribution of Ideal Point Estimates, by Gender

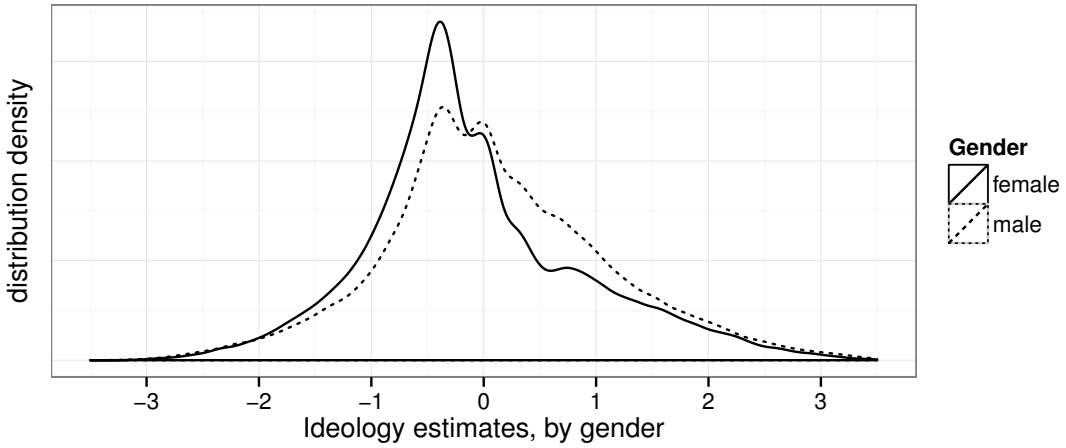


Figure 2 displays the ideology of the median Twitter user in each state, where the shade of the color indicates the quartile of the distribution.

Figure 2: Ideal Point of the Average Twitter User in the Continental US, by State

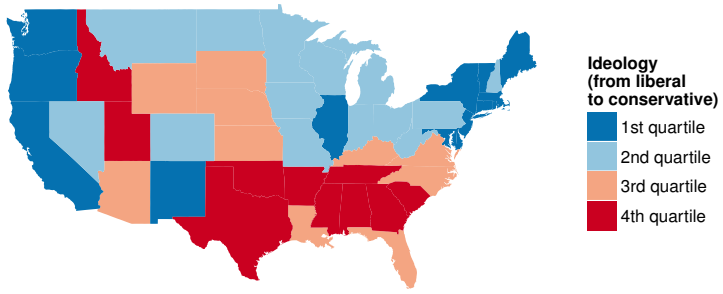
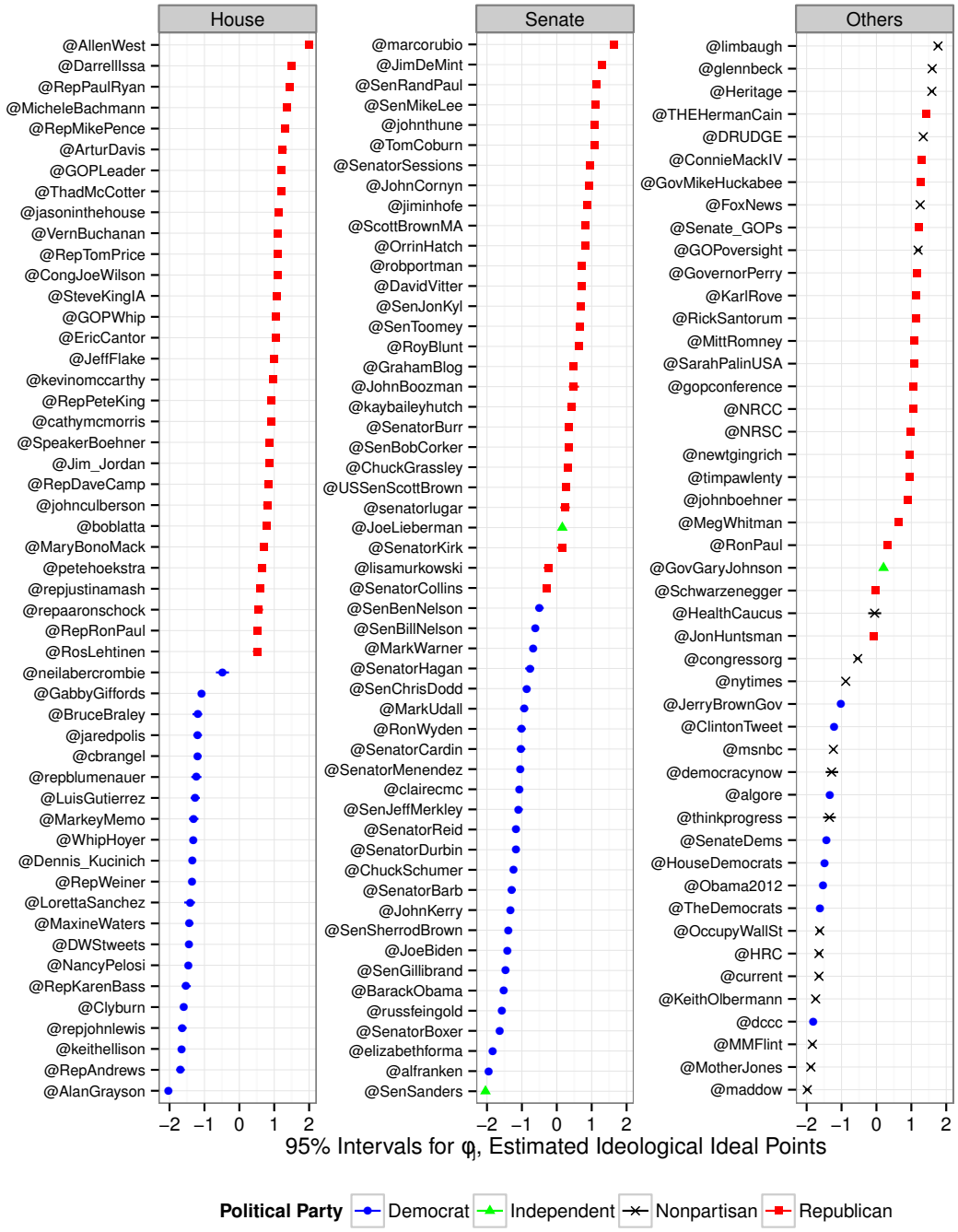


Figure 3 displays the estimated ideal points for the set of m key political actors in the US with 10,000 or more followers.

Figure 4 shows the distribution of ideal points for a sample of Twitter users who "self-reported" their vote for Obama ($N = 2539$) or Romney ($N = 1601$) on election day. To construct this dataset, I captured all tweets mentioning the word "vote" and either "obama" or "romney" and then applied a simple classification scheme to select only tweets

Figure 3: Estimated Ideal Points for Key Political Actors with 10,000 or more followers



where it was openly stated that the user had cast a vote for one of the two candidates.[1] As expected, ideology is an excellent predictor of vote choice, which provides additional evidence in support of the external validity of these ideal point estimates.

Figure 4: Distribution of Users' Ideal Points, by Self-Reported Votes

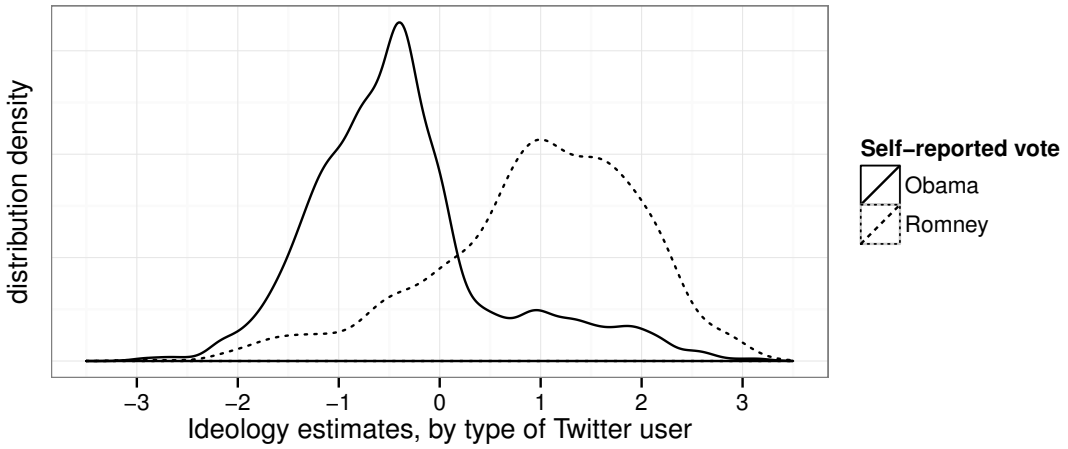


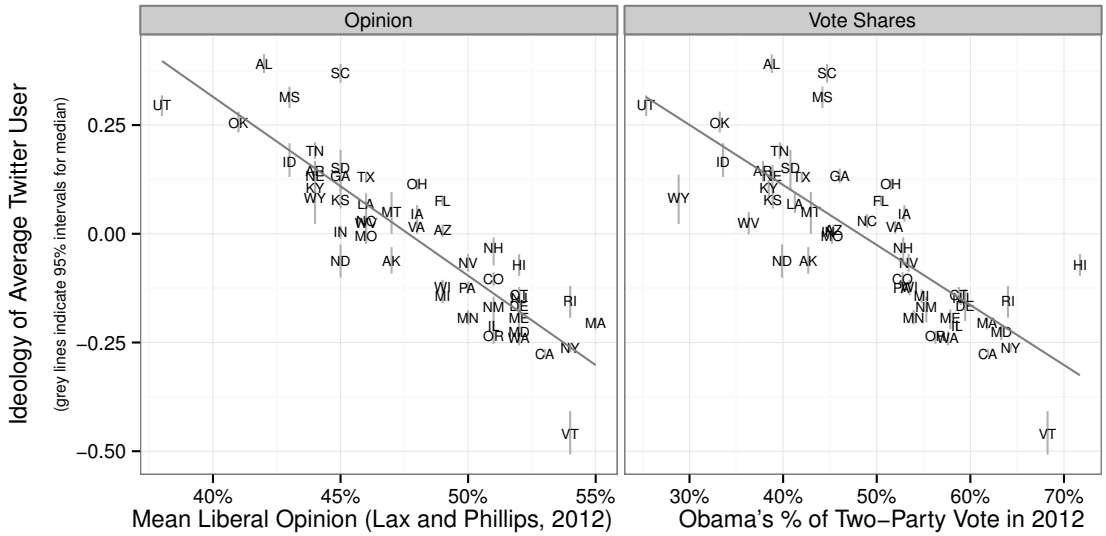
Figure 5 shows that the estimated ideal points for the median Twitter user in each state are highly correlated ($\rho = .880$) with the proportion of citizens in each state that hold liberal opinions across different issues (Lax and Phillips, 2012). Ideology by state is also a good predictor of the proportion of the two-party vote that went for Obama in 2012, as shown on the right side of the figure, but the magnitude of the correlation coefficient is smaller ($\rho = -.792$), which suggests that the meaning of the emerging dimension in my estimation is closer to ideology than to partisanship.[2]

The left panel of Figure 6 displays the distribution of Twitter-based ideal points for each group of contributors (Bonica, 2014), classified into three categories: those who donate

[1] For example, in the case of Obama I selected those tweets that mentioned "I just voted for (president, pres, Barack) Obama", "I am voting for Obama", "my vote goes to obama", "proud to vote Obama", and different variations of this pattern, while excluding those that mentioned "didn't vote for Obama", "never vote for Obama", etc.

[2] Additional evidence in support of this conclusion is that the correlation of the state-level Twitter-based estimates with measures of "Republican advantage" (difference between proportion of self-identified Republicans and Democrats in each state) according to Gallup is even lower: $\rho = -0.712$. Furthermore, in an OLS regression of vote share for Obama on by state on "Republican advantage" and Twitter-based ideology estimates, both coefficients are significantly different from zero at the 1% level.

Figure 5: Twitter-Based Ideal Points, by State



to Democratic candidates only, to Republican candidates only, or to both. As expected, individuals in the first (second) group are systematically placed to the left (right) of the average voter, and Twitter users who donated to both parties have centrist positions. The panel on the right compares ideal points estimated using Twitter networks and contribution records, showing that both measures are highly correlated (Pearson's $\rho = 0.80$). Note, however, that the correlations within each quadrant of Figure 6 are positive but low: $\rho = .164$ for the bottom-left quadrant and $\rho = .100$ for the bottom-right quadrant.

Figure 7 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.

Figure 8 replicates the analysis in Section 5 using "mentions" instead of retweets.

Figure 6: Ideal Point Estimates and Campaign Contributions

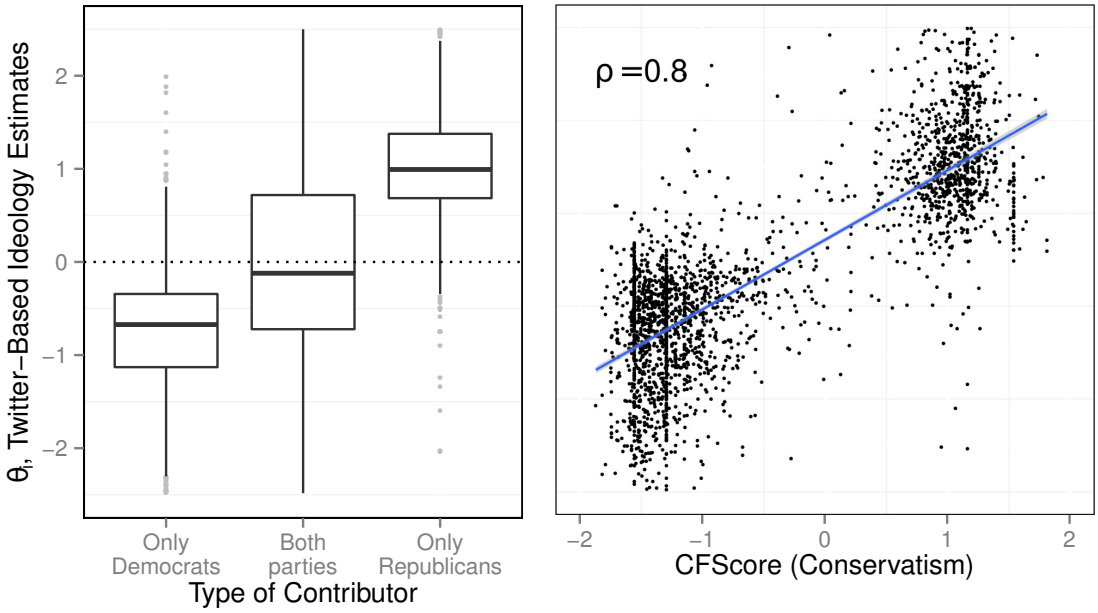


Figure 7: Evolution of mentions to Obama and Romney on Twitter

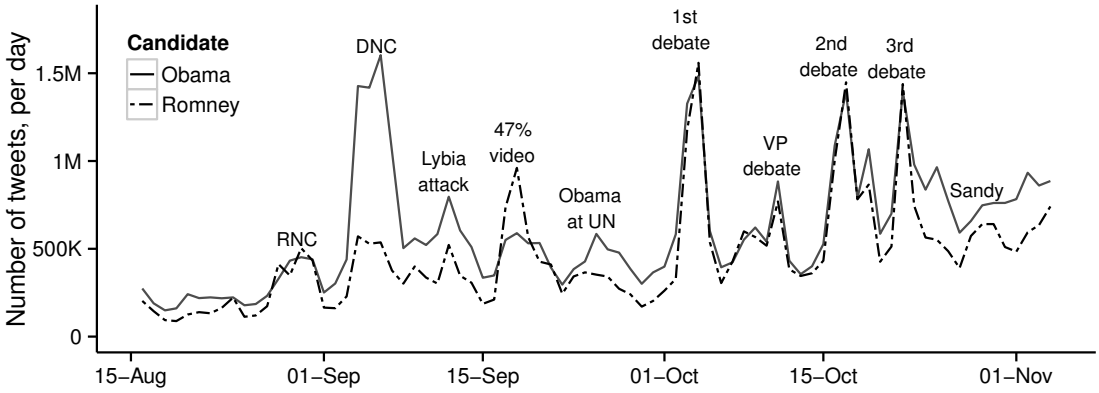
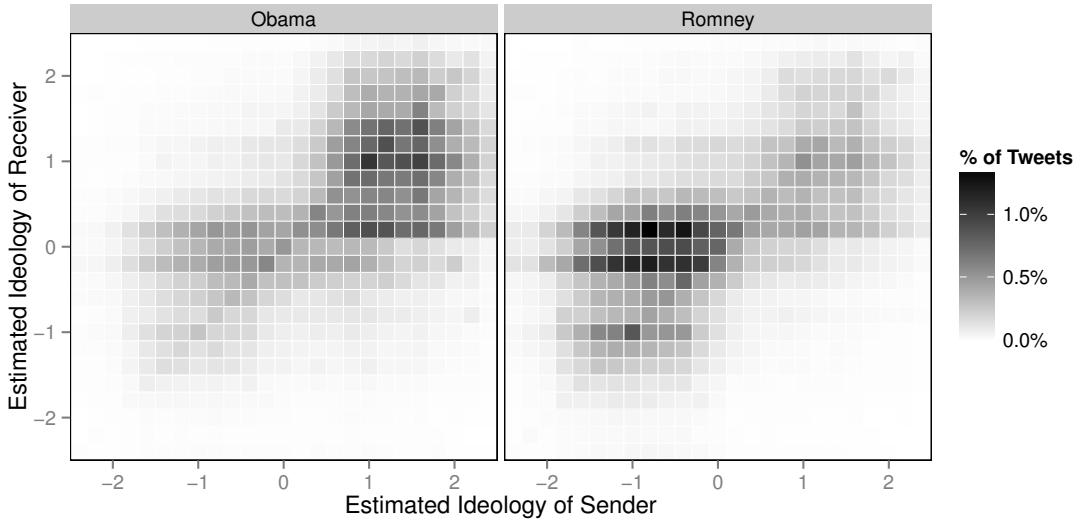


Figure 8: Ideological Polarization in Conversations Mentioning Presidential Candidates



D. TECHNICAL NOTES: ESTIMATION OF THE BAYESIAN SPATIAL FOLLOWING MODEL

D.1. *Code*

Table 2 displays the stan code to fit the statistical model introduced in Section 2. The code that implements the second stage of the estimation procedure, as well as the scripts to collect and process the Twitter data, will be made available online upon publication.

Table 2: STAN Code for Spatial Following Model

```
data {
  int<lower=1> J; // number of twitter users
  int<lower=1> K; // number of elite twitter accounts
  int<lower=1> N; // N = J x K
  int<lower=1,upper=J> jj[N]; // twitter user for observation n
  int<lower=1,upper=K> kk[N]; // elite account for observation n
  int<lower=0,upper=1> y[N]; // dummy if user i follows elite j
}
parameters {
  vector[K] alpha; // popularity parameters
  vector[K] phi; // ideology of elite j
  vector[J] theta; // ideology of user i
  vector[J] beta; // pol. interest parameters
  real mu_beta;
  real<lower=0.1> sigma_beta;
  real mu_phi;
  real<lower=0.1> sigma_phi;
  real<lower=0.1> sigma_alpha;
  real gamma;
}
model {
  alpha ~ normal(0, sigma_alpha);
  beta ~ normal(mu_beta, sigma_beta);
  phi ~ normal(mu_phi, sigma_phi);
  theta ~ normal(0, 1);
  for (n in 1:N)
    y[n] ~ bernoulli_logit( alpha[kk[n]] + beta[jj[n]] -
      gamma * square( theta[jj[n]] - phi[kk[n]] ) );
}
```

D.2. *Identification (Continued)*

To illustrate how global identification of the latent parameters is achieved, consider the estimated probability that the average individual in the U.S. sample ($\theta_i = 0$ and $\beta_i = -1.16$) follows Barack Obama ($\phi_j = -1.51$ and $\alpha_j = 3.51$),

$$\begin{aligned} P(y_{ij} = 1) &= \text{logit}^{-1}\left(\alpha_j + \beta_i - \gamma(\theta_i - \phi_j)^2\right) \\ &= \text{logit}^{-1}\left(3.51 - 1.16 - 0.93 \times (0 + 1.51)^2\right) \\ &= 0.56, \end{aligned}$$

which roughly corresponds with the observed proportion of users in the sample that follow him (191,986 of 301,537; 63%).

This equation has three indeterminacies. First, a constant k can be added to α_j and then subtracted from β_i leaving the predicted probability unchanged. Second, the same occurs when we add k to both θ_i and ϕ_j. This is usually referred to as "additive aliasing" (Bafumi et al., 2005), and implies that the latent scale on which these parameters are located can be "shifted" right or left without affecting the likelihood. A third type of indeterminacy is "multiplicative aliasing": ϕ_j and θ_i can be multiplied by any non-zero constant and γ divided by its square without changing the predicted probability. In other words, any change in how "stretched" the latent scale is can be offset by changes in γ. Equations 4 to 6 below illustrate each of these three indeterminacies.

$$\begin{aligned} P(y_{ij} = 1) &= \text{logit}^{-1}\left(\alpha_j + \beta_i - \gamma(\theta_i - \phi_j)^2\right) \\ &= \text{logit}^{-1}\left((\alpha_j + k) + (\beta_i - k) - \gamma(\theta_i - \phi_j)^2\right) \quad (1) \\ &= \text{logit}^{-1}\left(\alpha_j + \beta_i - \gamma((\theta_i + k) - (\phi_j + k))^2\right) \quad (2) \\ &= \text{logit}^{-1}\left(\alpha_j + \beta_i - \left(\frac{\gamma}{k^2}\right) \times ((\theta_i - \phi_j) \times k)^2\right) \quad (3) \\ &= 0.56 \end{aligned}$$

These equations show that without imposing any constraints, there is not a unique solution to the model, and therefore the Bayesian sampler will not converge to the posterior distribution of the parameters. As I discussed in Section 2.4, the model can be identified applying different restrictions. Table 3 shows the two most common approaches in the

Table 3: Identifying restrictions

Indeterminacy	Approach 1	Approach 2
Additive aliasing on α_j and β_i	Fix $\alpha'_j = 0$ or $\beta'_i = 0$	Fix $\mu_\alpha = 0$ or $\mu_\beta = 0$
Additive aliasing on ϕ_j and θ_i	Fix $\phi'_j = +1$ or $\theta'_i = +1$	Fix $\mu_\phi = 0$ or $\mu_\theta = 0$
Multiplicative aliasing on ϕ_j and θ_i	Fix $\phi''_j = -1$ or $\theta''_i = -1$	Fix $\sigma_\phi = 1$ or $\sigma_\theta = 1$

literature on scaling. One is to fix a subset of parameters at specific values, generally one ideal point at -1 for a liberal legislator and another at $+1$ for a conservative legislator. In the model I use here, since I'm computing more parameters than the standard item-response theory model, I would also need to fix one α_j or β_j.

A second approach, which is the one I use in this paper, is to fix the hyperparameters of the prior distributions of the latent parameters. In particular, I choose to give an informative prior distribution to the users' ideal point estimates, so that they have mean zero and standard deviation one, which facilitates the interpretation of the results. However, note that this set of restrictions achieves local identification but not global identification: all ideal points can be multiplied by -1 leaving the likelihood unchanged. In practice, this implies that the likelihood and posterior distribution are bimodal, and each individual chain may converge to a different mode. This could be solved after the estimation has ended, by multiplying the values sampled for $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ in each chain by -1 whenever the resulting scale is not in the desired direction. An alternative solution is to choose starting values for a set of ideal points that are consistent with the expected direction (liberals on the left, conservatives on the right), which has the advantage of speeding up convergence. This is the approach I implement in this paper. In particular, I set the starting values for ϕ_j to $+1$ for Republican legislators and to -1 for Democratic legislators. One advantage of this strategy is that it allows me to easily compute the percentile in the population of users' ideal points to which a given politicians' ideology estimate corresponds. For example, a politician with an estimated ideal point of -2 would be among the top 2.5% most liberal individuals.

To demonstrate that either set of restrictions identifies the model, I simulated data for 1,000 individuals and 100 political actors under the data generating process in equation 1, assuming that the distribution of ideal points is unimodal for individuals and bimodal for legislators (see Table 4). Then, I estimated the model under each of the two sets of identifying constraints (see Table 2 and 5), running two chains of 1,000 iterations with

a warmup period of 200 iterations. In both cases, the two chains converged to the same posterior distribution ($\hat{R}$ was below 1.10 for all parameters in the model), and the posterior estimates for the ideology parameters were indistinguishable from their true value ($\rho = .99$ in both cases), as shown in Figure 9.

Table 4: R Code to Simulate Data for Testing Purposes

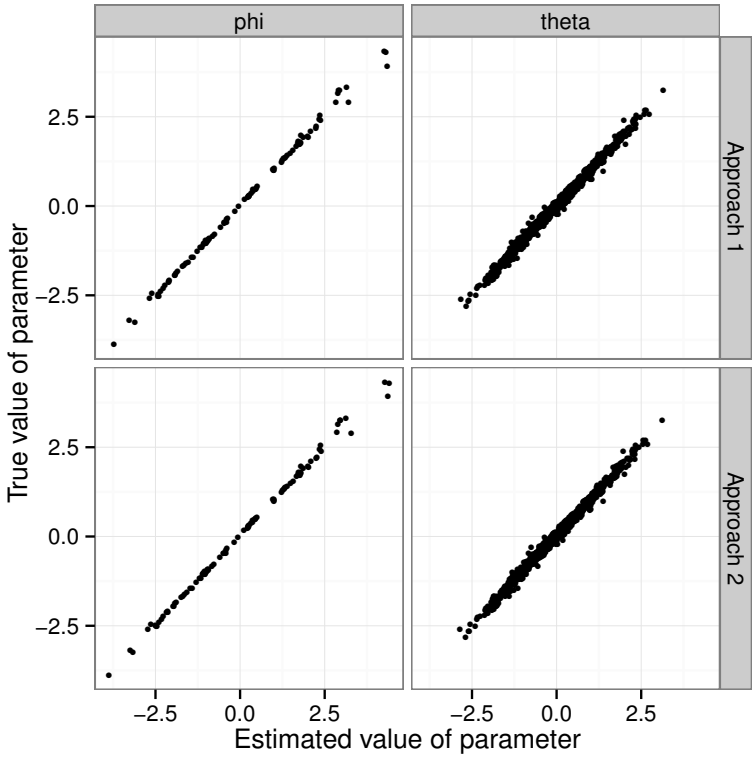
```
simulate.data <- function(J, K){
    # J = number of twitter users
    # K = number of elite twitter accounts
    theta <- rnorm(J, 0, 1) # ideology of users
    # ideology of elites (from bimodal distribution)
    phi <- c(-1, 1, c(rnorm(K/2-1, 1.50, 1), rnorm(K/2-1, -1.50, 1)))
    gamma <- 0.8 # normalizing constant
    alpha <- c(0, rnorm(K-1, 0, .5)) # popularity parameters
    beta <- rnorm(J, 1, .5) # pol. interest parameters
    jj <- rep(1:J, times=K) # twitter user for observation n
    kk <- rep(1:K, each=J) # elite account for observation n
    N <- J * K
    y <- rep(NA, N) # data

    for (n in 1:N){ # computing p_ij
        y[n] <- plogis( alpha[kk[n]] + beta[jj[n]] -
            gamma * (theta[jj[n]] - phi[kk[n]])^2 + rnorm(1, 0, 0.5))
    }
    y <- ifelse(y>0.50, 1, 0) # turning p_ij into 1,0
    return(list(data=list(J=J, K=K, N=N, jj=jj, kk=kk, y=c(y)),
        pars=list(alpha=alpha, beta=beta, gamma=gamma, phi=phi, theta=theta)))
}
```

Table 5: STAN Code for Model With Different Identifying Restrictions (excerpt)

```
...
model {
  phi[1] ~ normal(-1, 0.01);
  phi[2] ~ normal(+1, 0.01);
  for (k in 3:K)
    phi[k] ~ normal(mu_phi, sigma_phi);
  alpha[1] ~ normal(0, 0.01);
  for (k in 2:K)
    alpha[k] ~ normal(mu_alpha, sigma_alpha);
  beta ~ normal(mu_beta, sigma_beta);
  theta ~ normal(mu_theta, sigma_theta);
  for (n in 1:N)
    y[n] ~ bernoulli_logit( alpha[kk[n]] + beta[jj[n]] -
      gamma * square( theta[jj[n]] - phi[kk[n]] ) );
}
```

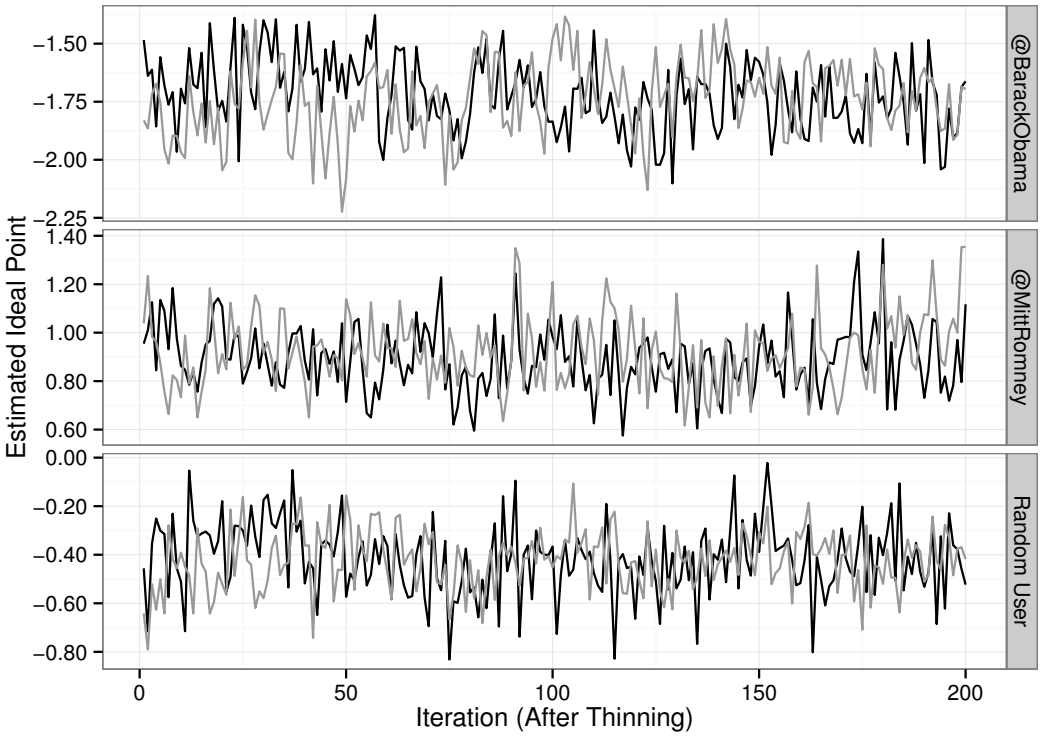
Figure 9: Comparing True Value of Parameters with their Estimates to Prove Identification



D.3. *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 10 shows that each of the two chains used 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 1.10 (consistent with robust convergence of multiple chains) and the effective number of simulation draws is over 200 for all ideology parameters – and in most cases around 400. The results of running Geweke and Heidelberg diagnostics also indicate that the distribution of the chains is stationary.

Figure 10: Trace Plots. Iterative History of the MCMC Algorithm



The results of a battery of predictive checks for binary dependent variables are shown in Table 6. All of them show that the fit of the model is adequate: despite the sparsity of the 'following' matrix (less than 3% of values are 1's), the model's predictions improve the

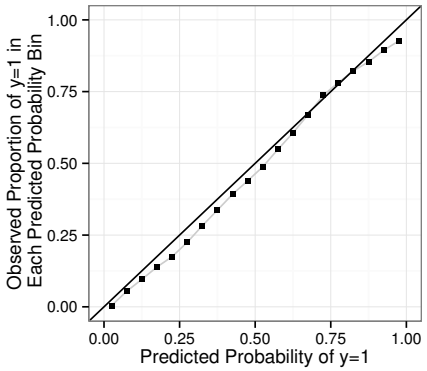
baseline (predicting all y_{ij} as zeros), which suggests that Twitter users' following decisions are indeed guided by ideological concerns. In addition to the widely known Pearson's ρ correlation coefficient and the proportion of correctly predicted values, Table 6 also shows the AUC and Brier Scores. 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), with lower values indicating better predictions.

Table 6: Model Fit Statistics.

Statistic	**Value**
Pearson's ρ Correlation	0.589
Proportion Correctly Predicted	0.977
PCP in Baseline (all $y_{ij} = 0$)	0.971
AUC Score	0.954
Brier Score	0.018
Brier Score in Baseline (all $y_{ij} = 0$)	0.029

A visual analysis of the model fit is also shown in Figure 11, 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 close to a 45-degree line (in dark color).

Figure 11: Model Fit. Comparing Observed and Predicted Proportions of $y_{ij} = 1$



D.4. *Estimating the Model with Covariates*

The decision to follow a political actor on Twitter may respond to a variety of reasons beyond ideological proximity. As it is formulated in equation 1, the model already incorporates two important factors: politicians' popularity (α_j) and users' interest in politics (β_i). However, it is likely that other reasons also explain how Twitter users decide who to follow. One of these is geographic distance. Gonzalez et al. (2011) show that around 50% of all "following" links take place between users less than 1,000 km apart. This is particularly relevant for Members of the U.S. Congress. An analysis of the data I use in this paper shows than an average of 21% of the followers of each individual legislator are Twitter users from their state. This may represent a problem for the estimation of the ideal point parameters if the effect of geographic distance is not orthogonal to ideology.

In order to explore whether that's the case, here I report the results of running the model incorporating geographic distance as an additional covariate. As shown in equation 4, I add an additional indicator variable, s_{ij} that takes value 1 whenever user i and political actor j are located in the same state, and 0 otherwise. The effect of geographic distance on the probability of establishing a following link is therefore δ, which I expect to be positive.

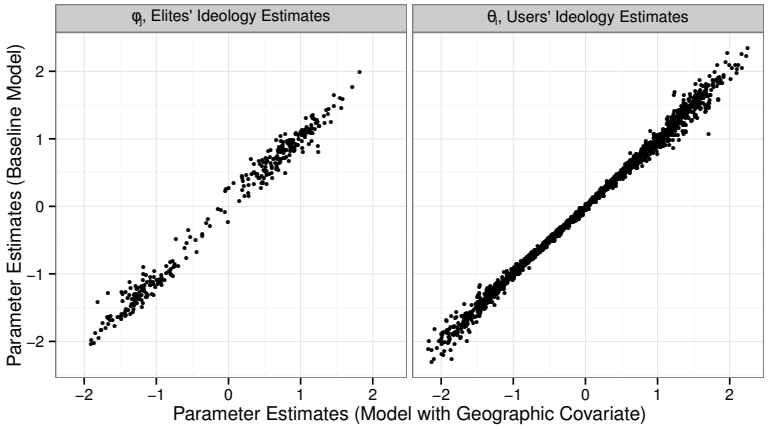
$$P(y_{ij} = 1) = \text{logit}^{-1}\left(\alpha_j + \beta_i - \gamma(\theta_i - \phi_j)^2 + \delta s_{ij}\right) \tag{4}$$

Geographic distance does have a large effect on following decisions ($\delta = 1.24$). The effect of being located in the same state as a Member of Congress on the probability of following him or her is equivalent to decreasing ideological distance by approximately one standard deviation.[3] For example, the model predicts that the probability that the average U.S. Twitter user ($\theta_i = 0$, $\beta_i = -2.29$) follows Barbara Boxer ($\theta_j = -1.68$, $\alpha_j = 0.60$) is 1.1%. If that user was located in California, then the probability would increase to 4%.

Despite the importance of geographic distance, I find that the ideology estimates for users and elites remain essentially unchanged after controlling for this effect. Figure 12 compares both sets of parameters across the baseline model and that in equation 4, estimated with a random sample of 2,000 users with an identifiable geographic location. I find that users' ideal point estimates are indistinguishable across models (Pearson's $\rho = 0.997$). There's slightly more variation in the case of elites' ideology estimates, partly due to the smaller sample size, but they are still highly correlated ($\rho = 0.992$).

[3]Note that $\gamma = 0.96$, and therefore the equivalent effect of increasing s_{ij} by one unit is $\sqrt{\delta/\gamma} = 1.13$,

Figure 12: Comparing Parameter Estimates Across Different Model Specifications



References

Al Zamal, F., W. Liu and D. Ruths. 2012. Homophily and latent attribute inference: Inferring latent attributes of twitter users from neighbors. In *Proc. 6th International Conference on Weblogs and Social Media*.

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):171–187.

Betebenner, Damian W. 2012. "randomNames: Function for creating gender and ethnicity correct random names." R package available on CRAN.
URL: *http://cran.r-project.org/web/packages/randomNames/index.html*

Bird, Steven, Ewan Klein and Edward Loper. 2009. *Natural language processing with Python*. O'reilly.

Bonica, Adam. 2014. "Mapping the Ideological Marketplace." *American Journal of Political Science* 58:367–386.

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 *Proc. Sixth International AAAI Conference on Weblogs and Social Media*.

holding all else equal.

Bradley, A.P. 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.W. 1950. "Verification of forecasts expressed in terms of probability." *Monthly weather review* 78(1):1–3.

Conover, M., B. Gonçalves, J. Ratkiewicz, A. Flammini and F. Menczer. 2010. Predicting the political alignment of twitter users. In *Proc. 3rd Intl. Conference on Social Computing*.

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* .

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

Lax, Jeffrey and Justin Phillips. 2012. "The democratic deficit in the states." *American Journal of Political Science* 56(1):148–166.

Mislove, A., S. Lehmann, Y.Y. Ahn, J.P. Onnela and J.N. Rosenquist. 2011. Understanding the Demographics of Twitter Users. In *Proc. 5th International Conference on Weblogs and Social Media*.

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

Pennacchiotti, M. and A.M. Popescu. 2011. A machine learning approach to twitter user classification. In *Proc. 5th International Conference on Weblogs and Social Media*.