Examining Political Trustworthiness through Text-Based Measures of Ideology

Eitan Sapiro-Gheiler

Department of Economics, Princeton University
Julis Romo Rabinowitz Building, Princeton, NJ 08544
Phone: 612-321-1849, Email: eitans@princeton.edu

Abstract

This work shows the value of word-level statistical data from the US Congressional Record for studying the ideological positions and dynamic behavior of senators. Using classification techniques from machine learning, we predict senators' party with near-perfect accuracy. We also develop text-based ideology scores to embed a politician's ideological position in a one-dimensional policy space. Using these scores, we find that speech that diverges from voting positions may result in higher vote totals. To explain this behavior, we show that politicians use speech to move closer to their party's average position. These results not only provide empirical support for political economy models of commitment, but also add to the growing literature of machine-learning-based text analysis in social science contexts.

Introduction and Motivation

It is commonly believed that politicians lie, and do it often. After all, speeches and votes represent two different levels of commitment. While speeches are public, they are nonbinding; on the other hand, votes represent recorded positions that influence on whether policy proposals are approved or not. However, existing political science literature on promise-keeping is more mixed than that adage would suggest. Political parties are surprisingly trustworthy, as discussed for example by (Pétry and Collette 2009). Early work in this area used party platforms to manually derive a list of promises and examined policy implementations to see whether they line up, as in (Budge and Hofferbert 1990). The development of automatic text analysis and classification techniques from machine learning provides a new set of tools with which to address these questions. Recently, (Gentzkow, Shapiro, and Taddy 2016) and (Laurendale and Herzog 2016) have used these modern methods to quantify political polarization by extracting features from speeches in the US Congress, a much larger and more personalized body of text than party platforms.

The work presented here builds on these contributions in two ways. First, using methods from data classification on political texts to predict political party simulates a simplified version of voting—the determination of a political candidate's ideological position from available information.

While party-blind elections are uncommon, they do exist—as described by e.g., (Bonneau and Cann 2013)—so this machine-learning-inspired simulation of social behavior is more realistic than it seems on face. Furthermore, party affiliation is an objective ground-truth label, while voting-based measures of ideology, which this work uses to discuss trustworthiness, are constructed from data and thus a step away from the reality of political behavior. Second, this work develops "scores" that embed ideology in a one-dimensional space, similar to those based on voting behavior.[1] We use these scores to examine the gap between text-based and voting-based measures of political position, thus providing quantitative analysis of both the magnitude and causes of political untrustworthiness.

Methods and Data

Using word-count approaches (bag-of-words), we employ four classical classification methods: decision tree (DT), naïve Bayes (NB), support vector machine (SVM), and lasso-penalty regression (LR). This paper studies approximately 1000 senators—all those serving between 1995-2015.[2] All classification methods use 10-fold cross-validation. We train each method using all ten sessions of Congress and test its performance on each individual session of Congress. We also train classifiers for each method using only the data from each individual session of Congress, each of which is then tested not only on its own session, but also on true "out-of-sample" data by applying it to the other sessions of Congress and the combined 10-session sample.

Because of LR's excellent performance across the full sample and because it provides a wide range of probabilities of party membership, we use it to construct one-dimensional ideology scores that are directly comparable to DW-NOMINATE scores, the standard such measure based on voting records. In other words, we provide a novel text-based embedding of political ideology. These scores allow more fine-grained and quantitative approaches to measuring ideology than existing political-text-analysis literature,

---

[1]The most well-known example of this are Poole-Rosenthal DW-NOMINATE scores, described in (Poole and Rosenthal 1985).

[2]Data is drawn from a preprocessed version of the US Congressional Record used in (Laurendale and Herzog 2016). We perform further text preprocessing using standard tools in Matlab.

which has focused on partisanship and polarization over individual behavior. We use them to offer an initial explanation for the dynamics of the gap between text-based and voting-based measures, i.e., an explanation for the degree of truthfulness shown by particular senators at different points in time. Specifically, we use linear regression models to test hypotheses around adherence to potential priors, incumbency, measures of electoral success, and district demographics.

Results

Classifiers trained on the full data set have party-prediction accuracy rates of 0.745 (DT), 0.728 (NB), 0.900 (SVM), and 0.983 (LR). Classifiers trained on only one session have average success rates across all sessions of 0.651 (DT), 0.729 (NB), 0.761 (SVM), and 0.678 (LR). Predicting DW-NOMINATE scores and then classifying based on those predictions notably improves the LR result. Full-data-trained classifiers have accuracy rates of 0.723 (DT), 0.701 (SVM), and 0.998 (LR) while the average for single-session-trained classifiers is 0.639 (DT), 0.757 (SVM), and 0.998 (LR); we omit NB because it sorts into discrete classes. LR usually outperforms the other methods in sessions close to the training session, while NB and SVM have least variance in their success rate. However, all classifiers show a decrease in performance on sessions at increased temporal distance from their training session. Both of these properties are preserved when predicting DW-NOMINATE scores as an intermediate step. This indicates, as observed by, e.g., (Jensen et al. 2012), that Congressional vocabulary shifts over time. Thus, voters who do not keep up with political developments are likely to lose the ability to perform even basic partisan sorting after only one or two sessions of Congress.

Regarding the hypotheses about differences between political speech and voting records, this work also provides initial evidence to believe that speech which diverges from policy positions may be beneficial, yielding larger vote totals, but mixed statistical significance results temper this claim. As an explanation, we show that politicians are likely aware of the "party line," using their speech to move closer to voters' priors than their voting record allows. We regress the difference between each senator's text-based and DW-NOMINATE scores on an indicator variable for whether the text-based score is closer to party-level measures of ideology than to the senator's DW-NOMINATE score. The resulting coefficient is positive, consistent in magnitude, and significant at the 1% level for several choices of party-level ideology, including current and recent-past party-mean or party-median DW-NOMINATE scores. It also passes a placebo test using non-viable but correlated priors. A similar regression shows that incumbents have larger gaps between speech and votes after controlling for state-level characteristics directly or with a fixed-effects model, but there is room for future work to determine whether this behavior is a cause or result of incumbents' persistent advantages.

Further Work

There are of course limitations to this analysis and directions for future work. We do not use neural networks, due in large part to the volume of data needed. However, refinements of the methods implemented here can improve upon the already-successful classification and ideology embedding. With respect to the text-based ideology scores, there is an opportunity to develop domain-specific embeddings by using positions and speech on issues of interest, such as trade, immigration policy, or taxation. To provide a larger corpus for neural network approaches or a more focused one for issue-specific scores, web-scraping and other techniques would allow access to a wider range of speeches and public statements. There are a variety of non-word-count-based methods of text analysis, such as word embedding or natural language processing, that incorporate more complex relationships between words and phrases or even include meaning as a component of the analysis. These approaches would enrich both the predictive power and the verisimilitude of these models as they relate to the task faced by voters. Theoretically, this work would provide ground for testing the predictions of signaling and Bayesian persuasion models that analyze political postures with and without commitment. The realm of text analysis is a growing field with great potential for applications in political science, economics, and beyond. This work is simply a first step towards taking full advantage of these rich new tools and datasets.

Acknowledgments

References

Bonneau, C. W., and Cann, D. M. 2013. Party identification and vote choice in partisan and nonpartisan elections. *Political Behavior* 37(1):43–66.

Budge, I., and Hofferbert, R. 1990. Mandates and policy outputs: U.S. party platforms and federal expenditures. *American Political Science Review* 84(1):111–131.

Gentzkow, M.; Shapiro, J.; and Taddy, M. 2016. Measuring polarization in high-dimensional data: Method and application to congressional speech. *NBER Working Paper No. w22423*.

Jensen, J.; Kaplan, E.; Naidu, S.; and Wilse-Samson, L. 2012. Political polarization and the dynamics of political language: Evidence from 130 years of partisan speech. *Brookings Papers on Economic Activity* 43(2):1–81.

Laurendale, B., and Herzog, A. 2016. Measuring political positions from legislative speech. *Political Analysis* 24(3):374–394.

Pétry, F., and Collette, B. 2009. Measuring how political parties keep their promises: A positive perspective from political science. In Imbeau, L. M., ed., *Do They Walk Like They Talk? Speech and Action in Policy Processes*. New York: Springer. 694–702.

Poole, K., and Rosenthal, H. 1985. A spatial model for legislative roll call analysis. *American Journal of Political Science* 29(2):357–384.