

# Ideological Scaling of Social Media Users: A Dynamic Lexicon Approach

Mickael Temporão<sup>1</sup>, Corentin Vande Kerckhove<sup>2</sup>,  
Clifton van der Linden<sup>3</sup>, Yannick Dufresne<sup>1</sup> and  
Julien M. Hendrickx<sup>2</sup>

<sup>1</sup> Département de science politique, Université Laval, Québec, QC G1V 0A6, Canada. Email: [mickael.temporao.1@ulaval.ca](mailto:mickael.temporao.1@ulaval.ca)

<sup>2</sup> Large Graphs and Networks Group, Université catholique de Louvain, Louvain-la-Neuve, B-1348, Belgium

<sup>3</sup> Department of Political Science, University of Toronto, Toronto, ON M5S 3G3, Canada

## Abstract

Words matter in politics. The rhetoric that political elites employ structures civic discourse. The emergence of social media platforms as a medium of politics has enabled ordinary citizens to express their ideological inclinations by adopting the lexicon of political elites. This avails to researchers a rich new source of data in the study of political ideology. However, existing ideological text-scaling methods fail to produce meaningful inferences when applied to the short, informal style of textual content that is characteristic of social media platforms such as Twitter. This paper introduces the first viable approach to the estimation of individual-level ideological positions derived from social media content. This method allows us to position social media users—be they political elites, parties, or citizens—along a shared ideological dimension. We validate the proposed method by demonstrating correlation with existing measures of ideology across various political contexts and multiple languages. We further demonstrate the ability of ideological estimates to capture derivative signal by predicting out-of-sample, individual-level voting intentions. We posit that social media data can, when properly modeled, better capture derivative signal than discrete scales used in more traditional survey instruments.

*Keywords:* automated content analysis, computational social science, ideal point estimation, ideological scaling, machine learning, social media

## 1 Introduction

Theories of political disagreement often envisage ideology as a single, latent dimension that structures politics on a continuum between left and right (Downs 1957; Hare *et al.* 2015). A primary source in the measurement of ideology has been textual data, with a focus predominantly on official texts such as party manifestos, policy platforms, and proposed legislation (Gabel and Huber 2000; Budge *et al.* 2001; Laver, Benoit, and Garry 2003; Klingemann *et al.* 2006; Slapin and Proksch 2008; Lauderdale and Herzog 2016). While these texts are well-suited for analysis given that they conform to certain norms pertaining to style, quality, and content, they tend to be top-down edicts that reflect the ideological underpinnings of the most dominant political actors within a party. They offer little insight into the range and distribution of ideology among the general population.

*Author's note:* We thank François Gélinau, Thierry Giasson, William Jacoby, Jonathan N. Katz (editor), Gregory Kerr, Michael Lewis-Beck, Alexander Shestopaloff, and two anonymous referees for their helpful comments and discussions. All remaining errors are ours. We also thank the participants of the 3rd Leuven-Montréal Winter School on Elections for their feedback on an earlier draft of this paper. This research was made possible thanks to an allocation of supercomputer resources from *Compute Canada*, specifically the 'Colosse' service administered by *Calcul Québec* at Laval University. The operation of this supercomputer is funded by the Canada Foundation for Innovation (CFI), the ministère de l'Économie, de la science et de l'innovation du Québec (MESI) and the Fonds de recherche du Québec—Nature et technologies (FRQ-NT). We are grateful in particular to Félix-Antoine Fortin for facilitating the computing resources associated with this research. Replication materials are available online on the *Harvard Dataverse* Temporão *et al.* (2018), at doi:10.7910/DVN/0ZCBTB. Supplementary materials for this article are available on the *Political Analysis* Web site.

*Political Analysis* (2018)  
vol. 26:457–473  
DOI: 10.1017/pan.2018.30

Published  
28 August 2018

Corresponding author  
Mickael Temporão

Edited by  
Jonathan N. Katz

© The Author(s) 2018. Published  
by Cambridge University Press  
on behalf of the Society for  
Political Methodology.

By comparison with official party texts, social media data, such as Tweets, tend to be less focused, much shorter, and more informal. Although their structure makes social media data less appealing than official party documents for purposes of textual analysis, their abundance and breadth—both in terms of population and subject matter—arguably permits a more substantive rendering of the ideological landscape as well as a better sense as to how ordinary citizens are situated therein. To date, attempts at applying ideological text-scaling methods to social media data have been limited to users who can be readily associated with a substantial corpus exogenous to a given social media platform (Grimmer 2010; Quinn *et al.* 2010). This has, in most cases, constrained the analysis to political parties, election candidates, and elected officials whose ideological positioning is clearly articulated in official texts.

In this article we develop a method that allows us to estimate the individual-level ideological attributes of both political elites *and* ordinary citizens using the textual content they generate on social media platforms. We first parse the lexicon of political elites in order to create dynamic dictionaries using the *Wordfish* algorithm (Slapin and Proksch 2008). These dictionaries are used to create scales that reflect the ideological dimensions that structure political discourse. We then estimate the position of individual social media users on said scales by using the previously computed dynamic dictionaries to analyze the textual content a user generates.

Ideological estimates are validated using survey data sourced from an uncommonly large and richly detailed sample of ideologically profiled Twitter users gathered using a Voting Advice Application (VAA) called Vote Compass.<sup>1</sup> The application is run during election campaigns and surveys users about their views on a range of policy issues. It generates a rich ideological and sociodemographic profile of its users and, in certain instances, captures their social media identifiers. We use the Vote Compass survey data to estimate a given social media user's ideological position on a multidimensional scale and compare this result with the ideological positions derived for that same user using unsupervised analysis of the content they generate on social media. The high correlation between the two measures suggests convergent validity. As an additional validation step but also to illustrate the usefulness of our approach, we attempt to predict out-of-sample individual-level voting intentions from social media data and contrast the estimates produced with those of models based on survey data from Vote Compass. Interestingly, the predictive power of the ideological estimates can outperform those of surveys when combining textual data with rich network information (Barberá 2015). We thus believe that data generated by social media can be considered in some respects to be even richer than that collected via conventional surveys.

This method represents a unique approach to the measurement of ideology in that it extends the utility of textual analysis to the domain of social media. Current methods for ideological inference from social media data have relied primarily on network analysis (Barberá 2015). We examine the content that social media users generate rather than the connections between them. In order to do so, we develop a technique to identify and extract a political lexicon from the broader social media discourse. The conventional application of textual analysis to the estimation of ideology considers the entire corpus. This is relatively unproblematic when the scaling texts that are concentrated on the factors of interest, as is the case with party manifestos or policy platforms (Grimmer and Stewart 2013). However, social media discourse is much denser and more manifold in terms of subject matter. Using dynamic dictionary files trained on the lexicon of political elites, we parse social media content in such a manner as to distinguish text which is predictive of ideology from the broader discourse.

Our contribution to the measurement of ideology is threefold. First, our model is endogenous—we do not rely on a corpus external to a given social media platform in order to infer the ideological

1 See <http://www.votecompass.com> for details.

position of platform users. As a result, we are able to extend ideological estimation beyond political elites to ordinary citizens, who are normally not associated with the corpora necessary to apply conventional ideological text-scaling methods. Second, our method is eminently scalable. Although we use Twitter data for the case studies presented in this article, the dynamic lexicon approach to ideological inference is platform agnostic—it could be readily applied to most other text-based social media platforms. Moreover, as we demonstrate, its performance is consistent in multiple languages. Third, a dynamic lexicon approach extends the purchase of ideological scaling of political actors in party systems outside the United States, where methods such as DW-NOMINATE (Poole and Rosenthal 1985, 2001) are often limited in terms of their ability to differentiate the ideological variance among legislators due to the effect of party discipline on voting records. We demonstrate the convergent validity of our approach across multiple different political contexts—both national and subnational—including Canada, New Zealand, and Quebec.

The potential applications for this method in electoral behavior studies and public opinion research are vast. Perhaps most notably, it offers a cost effective alternative to panel studies as a means of estimating the ideological position of both a given social media user and the ideological distribution within a population of interest.

## 2 Deriving Ideological Scales from Social Media Text

Existing text-based ideological scaling methods require large corpora of formal and politically oriented texts in order to scale a valid ideological dimension. These methods either follow a supervised approach, such as the *Wordscore* method (Laver, Benoit, and Garry 2003), or an unsupervised approach, such as the *Wordfish* method (Slapin and Proksch 2008). The former approach uses the guidance of experts to choose and position of reference texts that define the ideological space in order to position actors in that space (Lowe 2008). The latter approach estimates the quantities of interest, without requiring any human judgment, by relying on the assumption of ideological dominance in the texts under scrutiny (Monroe and Maeda 2004). Such constraints limit the texts that can be scaled to those that are specifically designed to articulate ideological positions, such as political manifestos and policy platforms. This typically narrows the focus of any analysis to the political elites who produce these types of texts.

The user-generated text on social media platforms is, by contrast, characterized by its brevity, informality, and broad range of subject matter. These dynamics tend to be a function of platform design and user culture, as even political elites adapt their communication style to the parameters of a particular social media platform. Indeed, politicians who are able to demonstrate a fluency in the lexicon of social media are often rewarded with out-sized recognition for their efforts. For example, Hillary Clinton's most liked and re-Tweeted Tweet of the 2016 US Presidential election campaign referenced the popular phrase "Delete your account" in response to a Tweet from Donald Trump.<sup>2</sup> Despite notable outliers, political elites generally exhibit common patterns in terms of how they use social media. They tend to be most active during election campaigns (Larsson 2016) and generally use social media for purposes of broadcasting rather than dialog (Grant, Moon, and Busby Grant 2010; Small 2010). In this sense, they are typically using social media as an extension of their campaign messaging (Straus *et al.* 2013; Evans, Cordova, and Sipole 2014; Vergeer 2015).

As such, the discourse of political elites—even on social media—is distinct and thus distinguishable from that of the general public. If we assume that the social media content generated by political elites has a more ideological focus relative to the general user base of a given platform, it may be leveraged as a means to detect ideological signal in the noise of social media chatter. Attempts to measure ideology based on the overall content generated on social

2 Clinton, Hillary. (@HillaryClinton) 9 June 2016, 2:27 PM EST. Tweet: <https://twitter.com/HillaryClinton/status/740973710593654784>.

media platforms have generally produced estimates that are highly biased by content unrelated to ideology (Grimmer and Stewart 2013; Lauderdale and Herzog 2016). However, preliminary evidence suggests that political elites, as a subset of social media users, may be more reliably scaled based on the content they generate (see Ceron 2017).

We use the lexicon of political elites to dynamically parse social media content so as to extract recurrent, comparable, and reliable ideological positions for a broader population of social media users—both political elites and ordinary citizens. This dynamic lexicon approach proceeds from the assumption that, on average, political elites tend to discuss a narrower range of subjects than do ordinary citizens (at least in public), and that these subjects tend to be more politically oriented. The texts generated by political elites on social media—as a subset of the general population—can in this sense define ideological space as has been demonstrated with earlier text-scaling approaches (Monroe, Colaresi, and Quinn 2008; Grimmer and Stewart 2013; Ceron 2017).

In order to apply the dynamic approach, the texts generated by political elites on their social media profiles are first collected in order to construct a lexicon of political terms that structure the dimensions of ideological discourse. We restrict the collection of social media content to defined time periods, specifically election campaigns. This is due in part to evidence that political elites are most active on social media during election campaigns (Barberá 2015; Larsson 2016), but also because election campaigns force ideological differences into sharper focus. To maintain a standardized operationalization of ‘political elites’ for comparative purposes, we limit the selection of social media users to politicians, which we define as incumbents or challengers who are candidates for election in a given campaign. To control for the effect of word use variation introduced by different languages, only texts written in the dominant language of the election discourse are included. In order to ensure that this method is scalable across contexts and not constrained by the subjective or arbitrary judgment of individual researchers, we rely on an unsupervised approach—specifically, the *Wordfish* algorithm—to parse the given texts and identify terms that are indicative of certain ideological positions. Though this method varies from the more established network approach, whereby ideological estimates are derived based on network connections such as Twitter followers (Barberá 2015), it will by design incorporate a network effect dynamic as social media as the sharing of texts is more likely to occur between social media users who share a connection. The practice of sharing on social media platforms thus amplifies the occurrence of certain terms in the dynamic lexicon approach. Finally, the lexicon extracted from the discourse of politicians is scaled and then compared with the texts generated by other social media users so as to estimate their respective ideological position.

## 2.1 Calibrating a dynamic lexicon of ideology

Ideology is generally conceived of in multidimensional terms, as represented by a  $d$ -dimensional vector in  $\mathbb{R}^d$ . Extracting these dimensions is referred as ideological scaling. Intuitively, two members who share similar political views should be positioned nearby one another in ideological space, which is generally defined in terms of Euclidean distance.

Text-scaling approaches assume that the frequency of expression of specific terms made by a given actor can be used to estimate said actor’s position in an ideological space. Initially Monroe and Maeda (2004) proposed a scaling model based on item response theory that was later extended by Slapin and Proksch (2008) to measure political positions from written documents. Their model considers that the use of words can be modeled by a Poisson process. The process first builds a matrix  $Z$  (also called a term-document matrix or TDM) where each element  $z_{jw}$  corresponds to the number of times the word  $w$  is present in party  $j$  manifesto. Then, the expected value of  $z_{jw}$  is expressed as a function of the ideologies:

$$Pr(z_{jw} = k) = f(k; \lambda_{jw}) \quad \text{with } \lambda_{jw} = \exp(s_j + p_w + \theta_w x_j), \quad (1)$$

where  $f(k; \lambda_{jw})$  represents the Poisson probability mass function. The model introduces two factors  $\rho_w$  and  $s_j$  to deal with the more frequent words and more active political candidates. The  $\theta_w$  factor takes into account the impact of words in the author's position and  $x_j$  is the estimate of the author's position. Estimators of the unknown parameters are derived using a maximum likelihood estimation. An iterative approach (using expectation–maximization) is implemented in the *Wordfish* algorithm (Slapin and Proksch 2008).

This dynamic lexicon creation process assumes that the textual content is politically dominant. This is often the case for political candidates' tweets, especially during an election campaign. We first record all the sequences of  $N$  adjacent words,  $N$ -gram, in the political candidates' tweets, for particular values of  $N = 1, 2$  or  $3$ . We take into account co-occurrences with other words to more readily distinguish between semantic groupings (Brown *et al.* 1992; Jurafsky and Martin 2009) and to better interpret sarcasm (Lukin and Walker 2013). Stop words are discarded and stemming is performed to improve the extraction accuracy. We handle nonpolitical terms by introducing a threshold parameter  $\beta$ . We discard  $N$ -grams that are not shared by at least  $\beta$  percent of the set of political candidates.<sup>3</sup> We then build a TDM ( $Z^{elit}$ ) by matching the  $N$ -grams' occurrences to the corresponding politicians. Finally, we use the Poisson model described in equation (1) to generate the dynamic lexicon of political terms, specifically we estimate  $\hat{\theta}_w^{elit}$  and  $\hat{\rho}_w^{elit}$ .

## 2.2 Scaling ideological positions for social media users

Text-scaling methods fail to retrieve a valid ideological dimension when the subject matter of the texts in question is broader than politics (Grimmer and Stewart 2013). It is reasonable to assume that politicians' user-generated content will be concentrated around political topics and themes. This assumption does not hold, however, for the average social media user. Ergo, an analysis of the content generated by average users—the political centrality of which is variable—does not readily lend itself to the identification of valid ideological positions.

In order to detect the ideological signal in the textual data of average social media users, a dynamic lexicon approach analyzes user-generated content through the lens of the content generated by politicians on the same social media platform. This approach is inspired by the concept of *transfer learning*, that is, estimating certain parameters using one dataset and then using those parameter values to make inferences on another dataset (Do and Ng 2005). When generating ideological estimates for politicians, the Poisson model (1) associates a weight estimator ( $\hat{\theta}^{elit}$ ) and a popularity estimator ( $\hat{\rho}^{elit}$ ) to each term present in the derived dynamic lexicon. One can then estimate a given user's ideology by taking into account the precomputed values ( $\hat{\theta}^{elit}, \hat{\rho}^{elit}$ ) in the Poisson model fit to the matrix  $Z^{elit}$  when handling the citizens' social media content. This entails an adapted TDM  $Z^{cit} \in \mathbb{R}^{n \times q}$ . The TDM is built by matching the  $q$  terms existing in the lexicon of political terms to a given social media user. The matrix entries  $z_{iw}$  count the number of instances in which each term (identified by  $w = 1 \dots q$ ) appears in the aggregated content of the user. The estimates are obtained by solving for the convex optimization problem (2):

The log likelihood function of a Poisson distribution is given by

$$l(\lambda) = \sum_{i=1}^n \sum_{w=1}^q (z_{iw} \ln(\lambda_{iw}) - \lambda_{iw} - \ln(z_{iw}!))$$

where  $\lambda$  represents a matrix, here the matrix of event rates  $\lambda_{iw} = \exp(s_i + \hat{\rho}_w^{elit} + x_i \hat{\theta}_w^{elit})$ . We optimize the log likelihood over the variables  $s_i$  and  $x_i$  conditional on  $\hat{\theta}_w^{elit}$  and  $\hat{\rho}_w^{elit}$ . That is, the parameters  $\hat{\theta}^{elit}$  and  $\hat{\rho}^{elit}$  are now precomputed in the political candidates' dynamic lexicon

3 For more information regarding the parameter selection for the dynamic lexicon of political terms see Supplementary Material.

estimation process. This leads to the maximization problem

$$\text{maximize}_{s_i, x_i} \sum_{i=1}^n \sum_{w=1}^q (z_{iw}(s_i + x_i \hat{\theta}_w^{elit}) + z_{iw} \hat{\rho}_w^{elit} - \exp(\hat{\rho}_w^{elit}) \exp(s_i + x_i \hat{\theta}_w^{elit}) - \ln(z_{iw}!))$$

where both terms  $z_{iw} \hat{\rho}_w^{elit}$  and  $\ln(z_{iw}!)$  do not depend on the instances  $x_i$  and  $s_i$ . Suppressing these terms does not alter the optimal solution and leads to the convex optimization problem (2).

$$\text{maximize}_{s_i, x_i} \sum_{i=1}^n \sum_{w=1}^q (z_{iw}(s_i + x_i \hat{\theta}_w^{elit}) - \alpha_w \exp(s_i + x_i \hat{\theta}_w^{elit})), \tag{2}$$

where  $x_i$  denotes the ideology of citizen  $i$  and  $s_i$  depicts their publication activity. The constant  $\alpha_w = \exp(\hat{\rho}_w^{elit})$  takes into account the  $N$ -gram popularity vectors estimated from political candidates' lexicon. The optimization problem is derived by maximizing the likelihood function of Equation 2 over two sets of variables ( $x_i$  and  $s_i$ ) instead of the four initial vectors.

### 3 Data

To demonstrate the dynamic lexicon approach, we apply the method to the Twitter platform.<sup>4</sup> We note, however, that the approach is designed to be applied across text-based social media platforms.

In order to test the dynamic lexicon approach we rely on two sources of data. The first, of course, is Twitter data itself. The second is from a VAA called Vote Compass. We use data from three recent elections as case studies: the 2014 New Zealand general election, the 2014 Quebec provincial election, and the 2015 Canadian federal election (see Section 4.1 for details on case study selection).

As language explains most of the variation in the terms used on social media, we restrict the analysis to English-speaking users only for New Zealand 2014 and Canada 2015, and to French-speaking users for Quebec 2014.

#### 3.1 Vote Compass

In order to validate the ideological estimates derived from the dynamic lexicon approach, we rely on the survey data generated by a large-scale VAA with active instances in each of the election campaigns we draw on as case studies in this article.

Vote Compass is an online application that surveys users' views on a variety of public policy issues germane to a given election campaign, and then offers users an estimation of their position in the political landscape and, by extension, their alignment with each of the political parties contesting said campaign. The application is wildly popular, drawing millions of users worldwide, and is generally run in partnership with major media organizations in the jurisdiction where an election is being held.

In addition to attitudes to policy issues, Vote Compass also collects sociodemographic attributes and ideological self-placement on a scale of 0 to 10, representing left to right. Users also have the option to self-select into a research panel which associates their respective Twitter accounts with their responses to the survey. A total of 62,430 Twitter users were recruited through Vote Compass ( $n = 11,344$  for Quebec 2014,  $n = 8,452$  for New Zealand 2014, and  $n = 42,634$  for Canada 2015).

4 Replication materials are available online on the *Harvard Dataverse* Temporão et al. (2018), at doi:10.7910/DVN/0ZCBTB.

### 3.2 Twitter

The user-generated content during the campaign period for accounts associated with verifiable candidates for election is collected using the Twitter REST API. These data form the corpus of terms that serve as the lexicon that constitute the ideological discourse of a given campaign. A content filter with threshold ( $\beta$ ) excludes candidates whose tweets do not include at least 5% of these political terms by word count. A second network filter eliminates candidates who have fewer than 25 followers who we can validate externally using Vote Compass data (see Section 3.1). A third filter excludes candidates from political parties that have fewer than three politicians in the sample once the first two filters have been applied. Excluding candidates from such minor parties from the analysis prevents unreliable conclusions resulting from small sample sizes for classification.

Across the three election campaigns served as case studies for our analysis, we identify 297 candidates with public Twitter accounts in the 2014 Quebec election, 131 candidates in the 2014 New Zealand general election, and 759 candidates in the 2015 Canadian federal election. The number of active candidates in our sample once filtering has been applied is as follows:  $m = 106$  for Quebec 2014,  $m = 56$  for New Zealand 2014 and  $m = 120$  for Canada 2015. Citizens consist of potentially eligible voters that are active on Twitter during the electoral campaigns under scrutiny. As with candidates, we only consider citizens with at least 25 political bigrams<sup>5</sup> in the dynamic lexicon within each context, and who follow a minimum of 3 active candidates. The citizens' sample sizes after applying these filtering conditions are  $n = 796$  (Quebec 2014),  $n = 123$  (New Zealand 2014) and  $n = 1,717$  (Canada 2015).

## 4 Results and Validation

When relying on unsupervised ideological scaling methods, validation is essential to ensure that the estimates correspond to ideological positions. As any single test is generally not sufficient to establish the validity of a measurement, we take several different approaches to validity testing.

We begin with a focus on the face validity (Dinas and Gemenis 2010) of the estimates produced by the dynamic lexicon approach, drawing on an overview of the ideological landscapes in each of our three case studies. Our objective is to develop an intuitive rendering of the ideological positioning of the major parties in each case, so that we have a frame of reference by which to compare the estimates generated by the dynamic lexicon approach. We then test the convergent validity of the calibration process for the dynamic lexicon of political terms. Here we focus on the subset of political candidates, whose user-generated content defines the dynamic lexicon, by comparing the positions derived from this approach to an ideological scaling approach based on network information (Barberá 2015; Bond and Messing 2015). Finally, we extend this approach to individual citizens in our sample and compare their estimated positions to a reference position derived from survey data collected from Vote Compass.

### 4.1 Mapping ideological landscapes

The three case studies selected to demonstrate the dynamic lexicon approach are the 2015 Canadian federal election, 2014 New Zealand general election, and the 2014 Quebec provincial election. The intuitive ideological placement of the parties contesting these elections offers the opportunity to compare the estimates produced by the dynamic lexicon approach for both an external validity test and robustness check of the method.

Canada is a federal state, and elections take place in the context of a first-past-the-post electoral system (Kam 2009; Godbout and Høyland 2011). Four federal parties contest the nationwide vote, with the sovereigntist Bloc Québécois running candidates in the province of Quebec. The *New Democratic party (NDP)* and the *Green Party of Canada (GPC)* typically vie for the position of most

<sup>5</sup> For more details on how the variation of the bigram filtering criteria affects the sample size and the correlations with the reference ideology within each context, see Supplementary Materials.

left-leaning in federal politics, depending on the policy context, with the *Liberal Party of Canada (LPC)* adopting a generally center-left position and the *Conservative Party of Canada (CPC)* situated on the right of the ideological spectrum. Only the Liberal and Conservative parties have ever formed government, with the Conservatives having won, at minimum, a plurality of seats in the Canadian Parliament from 2006 to 2015, when the Liberals swept the federal election and formed a majority government.

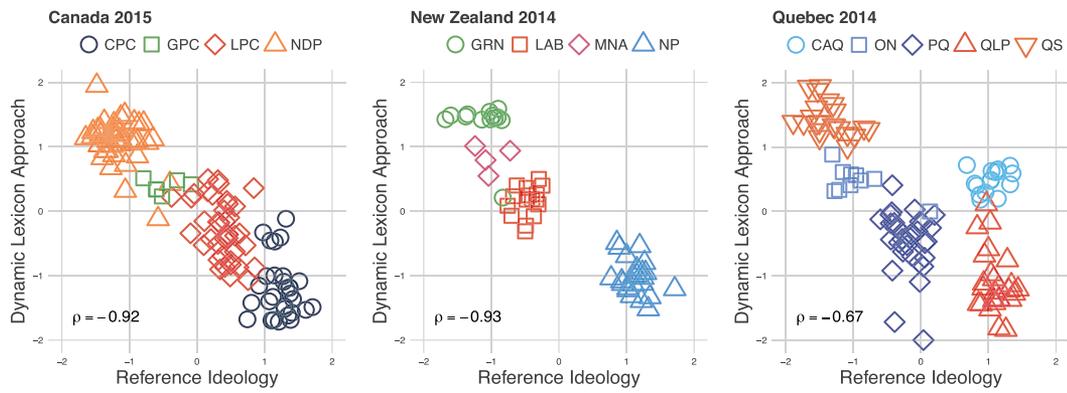
New Zealand is a unitary state, and its elections have been run under a mixed-member proportional framework since 1996. Although more than a dozen parties contest general elections, seven of which had at least one seat in Parliament following the 2014 general election. Of these, the filtering process (see Section 3.2) results in the inclusion of candidates from four parties. Of these the *Green Party of Aotearoa New Zealand (GRN)* is generally considered the most left-leaning. While the Mana Party has adopted certain radical positions, its coalition with the Internet Party during the 2014 campaign was perceived to have a slightly moderating effect on its ideological positioning, placing *Internet Mana (MNA)* slightly to the right of the Greens. The *New Zealand Labour Party (LAB)* is generally considered a center-left party and the *New Zealand National Party (NP)* on the center-right.

The case of the 2014 Quebec provincial election offers two unique tests of the dynamic lexicon approach. First, it extends the approach to the subnational level by way of its application in the context of a provincial election campaign. Second, it tests the scalability of an unsupervised model to contexts where English is not the dominant language. Quebec has a majority francophone population and the sovereigntist sentiment in the province gives rise to a unique ideological landscape within Canada wherein nationalism and identity constitute an independent dimension orthogonal to the more socioeconomic left–right spectrum. Toward the nationalist end of this spectrum, three parties advocate for Quebec independence: *Option Nationale (ON)*, the *Parti Québécois (PQ)*, and *Québec Solidaire (QS)*. The least nationalist party and that which is most opposed to Quebec independence is the *Quebec Liberal Party (QLP)*, the party that won office in 2014, putting an end to a short-lived PQ minority government. The *Coalition Avenir Québec (CAQ)* positions itself in between these two poles by advocating more autonomy for the French-speaking province while remaining a member of the Canadian federation. But the socioeconomic left–right ideological dimension also structures Quebec party competition, following roughly the same order as the identity dimension with most nationalist parties leaning generally more on the left side of the spectrum relative to federalist parties. Notable exceptions include, for example, attitudes toward religious accommodation, where the nationalist PQ takes more right-leaning positions than other parties.

## 4.2 Validating ideological estimates for election candidates

In order to assert that the dynamic lexicon approach properly classifies social media users' ideologies, we must first convincingly position election candidates within the ideological landscape. As candidates' lexicon is used to generate the ideological scales upon which other social media users are positioned, the ideological estimates of candidates serve as the reference point for those of all other users.

To validate the ideological estimates ascribed to candidates by the dynamic lexicon approach, we adapt the roll-call scaling method—an established standard in the estimation of elected officials' ideological positions (Poole and Rosenthal 2000; Clinton, Jackman, and Rivers 2004; Carroll *et al.* 2009). In doing so, however, we face two immediate constraints. Roll-call data are only available for elected officials, not political challengers. While this is less problematic in a two-party system given that both parties are likely to have substantial representation in government, it becomes more problematic in a multi-party, Westminster system (Spirling and McLean 2007; Grimmer and Stewart 2013). Not only does the presence of smaller parties increase the error



**Figure 1. Political candidates—Comparison of estimated positions for the reference method (network scaling approach) and the dynamic lexicon approach.** The x-axis shows the standardized position of political candidates on the unidimensional latent space derived from network data (Barberá 2015). The y-axis shows the standardized position of political candidates on the unidimensional latent space derived from the dynamic lexicon approach. Pearson correlations ( $\rho$ ) are all statistically significant ( $p$ -value < 0.05).

associated with ideological estimates, but a political culture that emphasizes party discipline makes it difficult to distinguish party ideology from candidate ideology. While it is reasonable to assume that most candidates' ideologies should be generally more aligned with their own party than with any other party, some degree of variation can be reasonably expected but would be difficult to detect using a DW-NOMINATE approach.

To address this limitation, one can estimate ideological positions based on network information, such as the existing connections between citizens and political candidates on social media (Barberá 2015; Bond and Messing 2015; Rivero 2015).

The scaling process relies on the assumption that these virtual connections are homophilic in nature, that is that citizens tend to follow political candidates whose ideologies lie close to their own. We refer to this approach hereafter as the network scaling approach (Barberá 2015). We validate the results of the dynamic lexicon approach by comparing how it positions election candidates *vis-à-vis* the estimates produced by network scaling.

The results demonstrate that, compared to the ideological estimates generated by a network scaling approach, the dynamic lexicon approach provides highly convergent results (see Figure 1). It is apparent that the point estimates showing each political candidates' positions measured using two fundamentally different scaling methods are strongly correlated. The correlations are particularly strong in the cases of the 2015 Canadian and the 2014 New Zealand elections. The linear relationship between the estimates indicates that they seem to capture a similar ideological dimension, which implicitly confirms the validity of both methods. Further examination suggests that these methods can identify clusters of candidates belonging to the same political party, even though within-party correlations are weak. This supports the premise that candidates are generally more aligned in ideological terms with candidates from their own party than with those from other parties. The positioning of candidates from each party accords with conventional wisdom as to how said parties are generally situated within their respective political landscapes. It also aligns with estimates derived by aggregating public articulations of public policy positions. This suggests that the dynamic lexicon approach is able to scale ideological positions at the individual level for political candidates and, by extension, render a valid ideological landscape from social media data.

Figure 1 compares the estimates generated by the dynamic lexicon approach with those of the network scaling approach. In the context of the 2015 Canadian election, the clustering of candidates generated by both methods depicts a similar ideological landscape. The candidates

cluster by party association and the party positioning exhibits face validity given the Canadian context.

The 2014 New Zealand context represented in Figure 1 shows once again that both the network scaling and dynamic lexicon approaches produce a consistent ideological landscape. The positioning of the party candidates also demonstrates face validity.

We find a slightly divergent result in the 2014 Quebec context. Unlike in Canada or New Zealand, the dynamic lexicon approach produces a different ideological landscape than the network scaling approach. For instance, the network scaling estimates, when taken alone, cannot differentiate the QLP from the CAQ. Although, upon examination of the dynamic lexicon approach estimates, we are able to clearly separate the CAQ from the QLP; although it then becomes difficult to distinguish the CAQ from ON. These results might suggest that these two methods, when taken together, can complement each other when capturing the ideological landscape. That said, the two methods identify relative clusters that are consistent with the parties' positions as one would intuitively expect. The results may also indicate that more than one dimension is being needed to capture nuances in particular ideological landscapes.

### 4.3 Validating ideological estimates for social media users

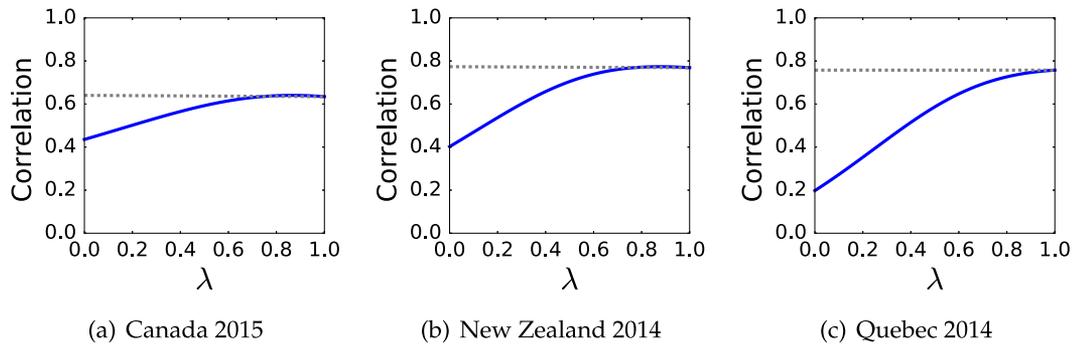
Having demonstrated that the dynamic lexicon approach can derive valid ideological dimensions from the content that political elites generate on social media, and that it can position those political elites in ideological space in ways that correlated highly with more established methods, we now examine its ability to classify the individual ideologies of the broader population of social media users.

To do so, we compare the dynamic lexicon approach to a basic *Wordfish* approach, in which we ignore the precomputed values and instead simply apply the algorithm. Two different baseline strategies are proposed, depending on whether or not we consider the use of a lexicon of political terms to create the TDM for citizens. In the first strategy (*Wordfish-all*), users' TDM is generated keeping all the  $N$ -grams (i.e., following the same building process as for elites' TDM  $Z^{elit}$ ). The second strategy (*Wordfish-political*) considers the adapted TDM  $Z^{cit}$  built from the dynamic lexicon of political terms. The reference ideological position for each users is derived from the survey data collected by Vote Compass.<sup>6</sup> We use an expectation–maximization algorithm to solve the maximum likelihood problem that generates these estimates (Barber 2012). We consider a social media extraction to be effective if two social media users with similar ideologies according to the Vote Compass survey data are also situated in close ideological proximity to one another using the dynamic lexicon approach. The *quality* of estimates is then defined as the Pearson correlation coefficient ( $\rho$ ) between the vector of textual estimates and the vector of reference ideologies.

Table 1 demonstrates the performance of the dynamic lexicon approach compared with the two baseline strategies. It clearly outperforms both baselines, especially for Quebec 2014, however, we observe weak to moderate correlations for this approach, which suggest that the dimension extracted by this method differs for the most part from that of the reference ideology. When we look solely at the politically interested citizens<sup>7</sup> individuals in our sample, the correlations remain similar even though one would expect the correlations to be higher for these more politically interested individuals as they are likely to demonstrate greater ideological consistency (Converse 1964). This can be partly explained by the selection bias inherent to the user base of VAA, wherein the average VAA user is more politically interested *vis-à-vis* the overall population. We furthermore

<sup>6</sup> See Section 3 for details.

<sup>7</sup> The subset of users reflected in parenthesis is based on filtering respondents that answered “*very interested*” to a political interest question asked in the Vote Compass survey. The question is as following: “*Generally speaking, how interested are you in politics?*” Users are offered the choice of one of the following four options: (1) *not interested at all*; (2) *not very interested*; (3) *somewhat interested*; (4) *very interested*.



**Figure 2. Citizens—Assessing linear combinations of textual and network ideologies.** The x-axis displays the parameter  $\lambda$  (3). The textual ideology corresponds to the case  $\lambda = 0$ . The network ideology corresponds to the case  $\lambda = 1$ . The y-axis displays Pearson correlations ( $p$ -value  $< 0.05$ ) between the linear combination and the reference ideologies. The dashed line depicts the correlation for  $\lambda = 1$ .

**Table 1. Citizens—Assessment of the dynamic lexicon approach for citizens.** Results are expressed for a bigram dictionary. The values indicate the Pearson correlations ( $p$ -value  $< 0.05$ ) between the ideologies extracted from Twitter data (text and network) and the reference ideologies based on policy positions. The values in parenthesis indicate the Pearson correlations for the subset of politically interested citizens.

Method	2015 Canada	2014 New Zealand	2014 Quebec
Dynamic lexicon	0.43 (0.44)	0.40 (0.38)	0.21 (0.22)
– Wordfish-all	0.29 (0.27)	0.34 (0.29)	0.09 (0.06)
– Wordfish-political	0.38 (0.39)	0.39 (0.42)	0.09 (0.11)
Network	0.63 (0.64)	0.76 (0.77)	0.75 (0.75)

tested the network scaling approach that was used to validate candidate positions (Barberá 2015) on the broader population of social media users. The results of the network scaling approach are strongly correlated with the reference ideology ( $\rho > 0.6$ ). This result suggests that, by looking at one’s connections on social media (in this case, Twitter followers), we are able to derive meaningful information about one’s ideological position.

As the information captured by these methods seems to differ, we consider combining network and textual ideologies into a single estimate as a means of enhancing the measurement. A classical method of merging estimators is to consider the set of estimators  $\hat{\mathbf{x}}_\lambda$  generated by convex combinations (Struppeck 2014). This family of ideological estimators is described by:

$$\hat{\mathbf{x}}_\lambda = \lambda * \hat{\mathbf{x}}^{net} + (1 - \lambda) * \hat{\mathbf{x}}^{txt}, \quad \lambda \in [0, 1]. \tag{3}$$

We compare the efficiency of combining ideologies by analyzing the correlation of the new estimates  $\hat{\mathbf{x}}_\lambda$  with the reference ideologies according to multiple values of  $\lambda$  (Figure 2). The optimal *quality* for Canada 2015, New Zealand 2014 and Quebec 2014 is respectively reached for  $\lambda = 0.86$ ,  $\lambda = 0.87$  and  $\lambda = 1$ . None of these combinations leads to enhancements of the network performance by more than 1%. This suggests that combining estimates does not lead to a significant improvement.

The moderate correlations observed in the dynamic lexicon approach could be attributable to a variety of factors. Ordinary citizens may exhibit less ideological consistency than do political elites, which could partly explain the lower correlations between the estimates (Converse 1964). However, the results emphasized in Table 1 are not sufficient to independently validate said hypothesis. Another explanation could be that the periods of time studied are too short to capture a concept as substantive as ideology. Increasing the time frame could plausibly result

in improved estimates. The presence of nonpolitical bigrams generates noisy estimates that explain the lower correlations observed for the textual estimates. The derived positions could also contain some other relevant information about individuals, but do not capture the same ideological dimension that captured by the reference ideology. Nevertheless, these two types of information could be complementary and could result in improved performance when trying to predict phenomena related to political ideologies, such as voting behavior.

## 5 Validating Ideological Estimates using Voting Intention

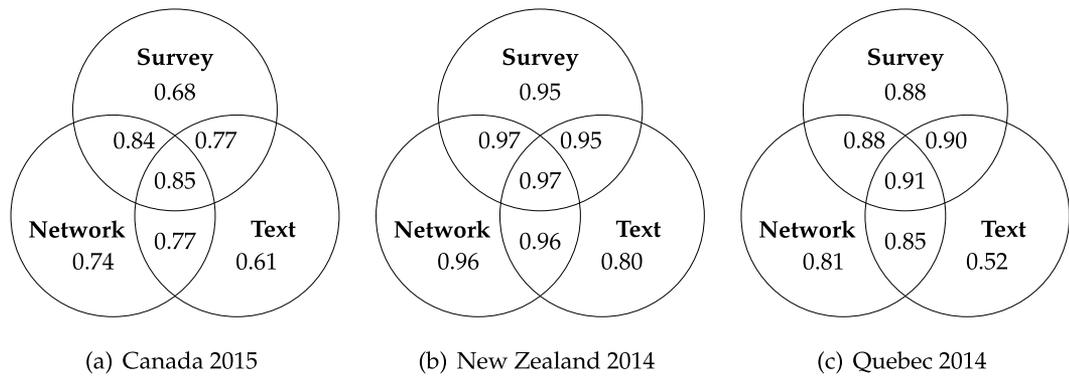
The notion that social media data may contain other useful information about its users provides an additional opportunity for validation. Scholarly interest in ideological positioning is in part motivated by the possibility to predict voting behavior (Downs 1957; Jessee 2009). Yet, predicting voting behavior remains a challenging empirical task. Contrary to convergent validity, which explores the correlations between multiple indicators of the same concept, predictive validity compares different phenomena that are linked together by an explanatory relation (Adcock 2001). The predictive validity of the ideological positions derived from the dynamic lexicon approach, that is, the ability to predict individual-level voting intentions from the ideological estimates the method produces, can serve as an additional means of validation of the approach.

To classify citizens by their voting intentions, we require two supplementary filtering criteria. The analysis is restricted to members who reported one of the parties that pass the filtering criteria in their voting intentions. Also, training sets with highly unbalanced class sizes generate important biases (Huang and Du 2005). The classification analysis is therefore performed on parties with at least 20 voters. This leads us to a total of 8 major parties: Canada's Conservative (CPC), Liberal (LPC), and New Democratic (NDP) parties, New Zealand's National (NP) and Labour (LAB) parties, as well as Quebec's Parti Quebecois (PQ), Liberal Party (QLP) and Coalition Avenir Quebec (CAQ). The resulting sample sizes are  $n = 796$  for Quebec 2014,  $n = 123$  for New Zealand 2014 and  $n = 1,717$  for Canada 2015. We perform a machine learning classification task to investigate the ability of the ideological estimates to correctly classify party affiliation (for elites) and predict vote intentions (for citizens). The machine learning approach is preferred to a descriptive statistical study since we focus on predicting an outcome rather than providing an explanatory model.<sup>8</sup>

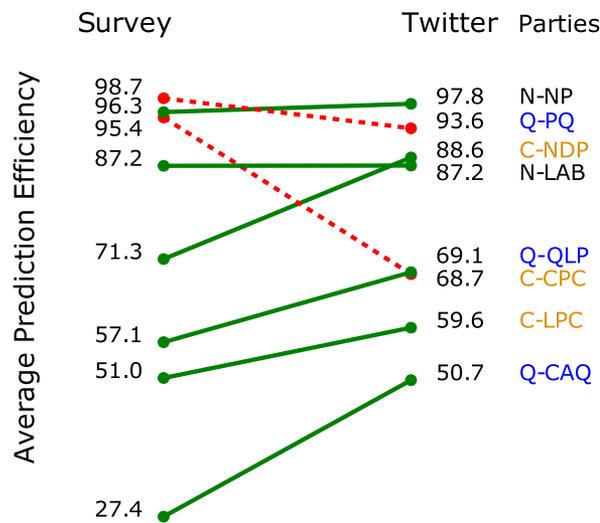
New data sources combined with novel scaling methods can allow the estimation of ideological positions that can predict votes without the cost of designing and administering complex and costly surveys. Furthermore, each different source of information available has the potential to complement others and improve the accuracy of a classification task (Jurafsky and Martin 2009). In order to investigate the complementarity of these methods, we evaluate the power of each of the two methods taken individually, but also in combination, to predict individual-level voting intentions.

The Venn diagrams (Figure 3) illustrate the average prediction efficiencies based on the dynamic lexicon approach (text), network scaling approach (network), and Vote Compass survey-based ideological positions (survey). This allows us to illustrate the quality of the predictions for each approach taken individually and for any combination between these. The metrics displayed are an evaluation of the quality of the predictions based on the AUC of precision and recall curves (PR) for each party within each context. These metrics evaluate the proportion of time the trained algorithm guesses the voting intentions of an individual correctly. The advantage of this metric is that it is less affected by sample balance than simply using the prediction accuracy. A cross-validation process handles over-fitting so that the results of the combinations of these estimates are nontrivial and actually result from new information captured.

8 For more information regarding for the setting of the machine learning classification, see Supplementary Materials.



**Figure 3. Venn diagram illustrating the complementarity between the ideology estimates to predicting voting intentions of citizens.** The metric values inside each set represent the prediction’s efficiency measured by the area under curve (AUC).



**Figure 4. Comparison at the party level of citizens’ Twitter and Survey prediction efficiencies.** The values displayed correspond to the area under curve of the precision and recall curves related to each party.

When we examine the AUC for each source in Figure 3, we can see that the predictions based on ideological estimates derived from a network scaling approach can outperform the predictions based on ideological estimates from Vote Compass survey data. This is the case for Canada and for New Zealand. In the case of Quebec, the ideological estimates from the Vote Compass survey data outperform the alternative approaches. Any combination of these estimates improves the quality of the prediction, with the exception of New Zealand, where it is clear that the estimates from the dynamic lexicon approach do not add any information to network scaling or survey-based estimates. This result is consistent to what is shown for elites in Figure 1. These findings indicate a cumulative improvement in the ability to predict voter intentions by combining different types of social media data, specifically ideological estimates derived from a combination of network and textual data. These combinations can outperform, or at least approach, established survey-based estimates without the burden of having to design or administer a survey.

To further examine the effect of combining different types of social media data, we illustrate the performance of our classifier when it combines a network scaling approach with the dynamic lexicon approach to predict individual voting intentions. We compare this combination to a classification based solely on survey data. Figure 4 shows the comparison of the two models

by illustrating their performance at the party level. The values displayed correspond to the AUC of PR curves related to each party. The solid lines highlight the parties where Twitter estimates perform better than the conventional survey estimates. The dashed lines indicate when Twitter information leads to higher prediction errors. We can see in Figure 4 that Twitter-based predictions of voting intentions generally outperform survey predictions. Even though the evidence does not allow us to generalize this pattern, it is worth noting that the dashed lines are related to the Parti Québécois (PQ) and the Conservative Party of Canada (CPC), two incumbent parties that failed to be re-elected in the cases under examination. Besides having that in common, these two parties also were the least active on social media relative to other parties. That is, we have much less data for those parties where the predictions are less accurate than those derived from survey data. There is also a noticeably higher efficiency rate for the social media model in the cases of for smaller parties such as the CAQ and NDP. This could be explained in part by the fact that said smaller parties have a more pervasive online presence in terms of quantity of Tweets published by their respective political candidates. Parties with smaller campaign budgets tend to rely more extensively on social media outreach than do parties with more substantial war chests (Haynes and Pitts 2009). They tend to publish larger quantities of information on social media in order to try to reach a broader audience at a lower cost. The additional data available for smaller parties can increase the efficiency of the classifiers in terms of identifying individuals intending to vote for said parties. Indeed, the more information we have to train the classifiers for these parties, the fewer classification errors we should observe. This is not the case with the Vote Compass survey data as every user answers questions related to specific policy issues relevant to the political campaign. On social media, however, the unsupervised approaches have to identify signals among large quantities of relevant and irrelevant information relative to the specific parties which require larger amounts of data in order to develop efficient estimates.

## 6 Conclusion and Discussion

In this article we introduce a dynamic lexicon approach to ideological scaling of social media users based on the textual content they generate on a given platform. This approach extends the capacity for ideological estimation from political elites to ordinary citizens by foregoing the requirement of a verbose and formal corpus of ideological texts exogenous to the platform. Instead, a given social media user's ideology can be estimated within a given ideological discourse, as defined by the lexicon of political elites and using only data endogenous to the platform. The findings from a series of validation tests indicate that a dynamic lexicon approach can extract ideological dimensions which demonstrate convergent validity in that they are correlated with other measures of ideology such as network scaling and survey-based approaches. It also exhibits predictive validity in terms of predicting individual-level voting intentions. Although we find that network scaling performs better than the dynamic lexicon approach in terms of predicting individuals' voting intentions, the combination of the two methods into a single model generally outperforms predictions of individual-level voting intentions extrapolated from survey-based measures of ideological self-placement.

The implications of these findings, should they withstand further empirical scrutiny, are significant for researchers in the fields of electoral behavior and public opinion research, who could effectively measure ideological dynamics with less data, lower cost, but greater accuracy than conventional survey-based methods. Moreover, in the context of its utility for predictive purposes, the dynamic lexicon approach not only serves to boost the accuracy of network scaling. Network estimates may demonstrate more predictive power but they are mostly static in the short term, computationally intensive to derive, and therefore not scalable to large pools of social media users. By contrast, text-based estimates have the potential to allow for real-time analysis of ideological volatility.

In interrogating the robustness of the dynamic lexicon approach, certain limitations of the method were surfaced. Though successful at scaling individual-level ideologies and predicting vote intentions, the approach comes at the cost of high filtering criteria. These criteria directly affect the number of users for whom an ideological estimation can be modeled. For example, the constraints in terms of content generation and time frames limited the analysis undertaken in this article to Twitter users who posted messages during elections campaigns included as case studies, i.e. the 2015 Canadian federal election, the 2014 New Zealand general election, and the 2014 Quebec provincial election.

These limitations are far from insurmountable and they open up opportunities for future research. A longer-term collection of Twitter data would provide an indication as to whether elections are particularly effective moments for detecting signal within the elite lexicon used to define the ideological landscape. Moreover, the platform scalability of the dynamic lexicon approach, while theoretically plausible, requires empirical testing. Does the approach produce comparable ideological estimates when applied, for example, to user-generated content derived from Facebook? Further testing of the linguistic scalability of the dynamic lexicon approach is also necessary, particularly given the variance in the convergent validity of the Quebec case—the only non-English case—*vis-à-vis* Canada and New Zealand. The more likely explanation of this variance, however, is the structure of the Quebec ideological landscape, which begs the question of whether multidimensional scaling of ideology would produce more accurate estimates in contexts such as Quebec. Finally, future research to expand classifiers—previously trained on previous, known contexts—may provide additional validation as to the capacity of this method to predict individual voting intentions for specific elections.

Avenues for future research notwithstanding, this research stands on its own as a novel contribution to experimental methods for deriving valid inferences from social media data. Political scientists often require recurrent, comparable, and valid estimates of ideological positions of political actors in order to develop and test theories. Acquiring such estimates through surveys is often prohibitive in terms of cost (Slapin and Proksch 2008; Bond and Messing 2015). A dynamic lexicon approach permits the study of a wide range of actors, including some for which reliable measures of ideology may not be otherwise available. Its scalability and multilingual capacity also suggest that the dynamic lexicon approach could be used to study political contexts where it is impractical to conduct conventional surveys, but where there is ample uptake of social media platforms. This approach could be used to offer a more refined analysis across a broad array of salient topics in the study of politics, such as representation, polarization, and populism. Moreover, as our estimates of vote intention demonstrate, the signal contained within social media data enables inferences beyond a particular variable of interest. We demonstrate that ideological estimates generated from social media data can, when properly modeled, better predict individual-level voting intentions than ideological estimates derived from discrete scales. This suggests that the rich information contained with social media data may provide additive signal.

The results from this initial interrogation of the dynamic lexicon approach are sufficiently promising to warrant further investigation into the promise and potential of this method.

### Supplementary material

For supplementary material accompanying this paper, please visit <https://doi.org/10.1017/pan.2018.30>.

## References

- Adcock, Robert. 2001. Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Association* 95(03):529–546.
- Barber, David (ed.) 2012. Factor analysis. In *Bayesian reasoning and machine learning*. Cambridge: Cambridge University Press, pp. 462–478.
- Barberá, Pablo. 2015. Birds of the same feather tweet together: Bayesian ideal point estimation using twitter data. *Political Analysis* 23(1):76–91.
- Bond, Robert, and Solomon Messing. 2015. Quantifying social media's political space: Estimating ideology from publicly revealed preferences on facebook. *American Political Science Review* 109(01):62–78.
- Brown, Peter F., Peter V. Desouza, Robert L. Mercer, Vincent J. Della Pietra, and Jenifer C. Lai. 1992. Class-based n-gram models of natural language. *Computational Linguistics* 18(4):467–479.
- Budge, Ian, Hans-Dieter Klingemann, Andrea Volkens, Judith L. Bara, and Eric Tanenbaum. 2001. *Mapping policy preferences: Estimates for parties, electors, and governments, 1945–1998*, vol. 1. Oxford: Oxford University Press on Demand.
- Carroll, Royce, Jeffrey B. Lewis, James Lo, Keith T. Poole, and Howard Rosenthal. 2009. Measuring bias and uncertainty in DW-NOMINATE ideal point estimates via the parametric bootstrap. *Political Analysis* 17(3):261–275.
- Ceron, Andrea. 2017. Intra-party politics in 140 characters. *Party Politics* 23(1):7–17.
- Clinton, Joshua, Simon Jackman, and Douglas Rivers. 2004. The statistical analysis of roll call data. *American Political Science Review* 98(02):355–370.
- Converse, Philip E. 1964. The nature of belief systems in mass publics. In *Ideology and Discontent*, ed. David Apter. New York: Free Press of Glencoe.
- Dinas, Elias, and Kostas Gemenis. 2010. Measuring parties' ideological positions with manifesto data a critical evaluation of the competing methods. *Party Politics* 16(4):427–450.
- Do, Chuong, and Andrew Y. Ng. 2005. Transfer learning for text classification. In *Advances in Neural Information Processing Systems 18*, ed. Y. Weiss, B. Schölkopf, and J. C. Platt. Cambridge, MA: MIT Press, pp. 299–306.
- Downs, Anthony. 1957. *An economic theory of democracy*. New York: Addison Wesley.
- Evans, Heather K., Victoria Cordova, and Savannah Sipole. 2014. Twitter style: An analysis of how house candidates used twitter in their 2012 campaigns. *PS: Political Science & Politics* 47(2):454–462.
- Gabel, Matthew J., and John D. Huber. 2000. Putting parties in their place: Inferring party left-right ideological positions from party manifestos data. *American Journal of Political Science* 44(1):94–103.
- Godbout, Jean-François, and Bjørn Høyland. 2011. Legislative voting in the Canadian Parliament. *Canadian Journal of Political Science* 44(2):367–388.
- Grant, Will J., Brenda Moon, and Janie Busby Grant. 2010. Digital dialogue? Australian politicians' use of the social network tool twitter. *Australian Journal of Political Science* 45(4):579–604.
- Grimmer, Justin. 2010. A Bayesian hierarchical topic model for political texts: Measuring expressed agendas in senate press releases. *Political Analysis* 18(1):1–35.
- Grimmer, Justin, and Brandon M. Stewart. 2013. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis* 21(3):267–297.
- Hare, Christopher, David A. Armstrong, Ryan Bakker, Royce Carroll, and Keith T. Poole. 2015. Using Bayesian Aldrich–McKelvey scaling to study citizens' ideological preferences and perceptions. *American Journal of Political Science* 59(3):759–774.
- Haynes, Audrey A., and Brian Pitts. 2009. Making an impression: New media in the 2008 presidential nomination campaigns. *PS: Political Science & Politics* 42(01):53–58.
- Huang, Yi-Min, and Shu-Xin Du. 2005. Weighted support vector machine for classification with uneven training class sizes. In *2005 international conference on machine learning and cybernetics*, vol. 7. Guangzhou, China: Institute of Electrical and Electronics Engineers, pp. 4365–4369.
- Jessee, Stephen A. 2009. Spatial voting in the 2004 presidential election. *American Political Science Review* 103(01):59–81.
- Jurafsky, Daniel, and James H. Martin. 2009. *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice-Hall.
- Kam, Christopher J. 2009. *Party discipline and parliamentary politics*. Cambridge: Cambridge University Press.
- Klingemann, Hans-Dieter, Andrea Volkens, Judith L. Bara, Ian Budge, and Michael D. McDonald. 2006. *Mapping policy preferences II: Estimates for parties, electors, and governments in Eastern Europe, European Union, and OECD 1990–2003*, vol. 2, Oxford University Press on Demand.
- Larsson, Anders Olof. 2016. Online, all the time? A quantitative assessment of the permanent campaign on facebook. *New Media & Society* 18(2):274–292.
- Lauderdale, Benjamin E., and Alexander Herzog. 2016. Measuring political positions from legislative speech. *Political Analysis* 24(3):374–394.

- Laver, Michael, Kenneth Benoit, and John Garry. 2003. Extracting policy positions from political texts using words as data. *American Political Science Review* 97(02):311–331.
- Lowe, Will. 2008. Understanding wordscores. *Political Analysis* 16(4):356–371.
- Lukin, Stephanie, and Marilyn Walker. 2013. Really? Well. Apparently bootstrapping improves the performance of sarcasm and nastiness classifiers for online dialogue. In *Proceedings of the workshop on language analysis in social media*. Atlanta, GA: Association for Computational Linguistics, pp. 30–40.
- Monroe, Burt L., and Ko Maeda. 2004. Talk's cheap: Text-based estimation of rhetorical ideal-points. In *Annual Meeting of the Society for Political Methodology*. Cambridge: Cambridge University Press, pp. 29–31.
- Monroe, Burt L., Michael P. Colaresi, and Kevin M. Quinn. 2008. Fightin' words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis* 16(4):372–403.
- Poole, Keith T., and Howard Rosenthal. 1985. A spatial model for legislative roll call analysis. *American Journal of Political Science* 29(2):357–384.
- Poole, Keith T., and Howard Rosenthal. 2000. *Congress: A political-economic history of roll call voting*. Oxford: Oxford University Press.
- Poole, Keith T., and Howard Rosenthal. 2001. D-nominate after 10 years: A comparative update to congress: A political-economic history of roll-call voting. *Legislative Studies Quarterly* 26(1):5–29.
- Quinn, Kevin M., Burt L. Monroe, Michael Colaresi, Michael H. Crespin, and Dragomir R. Radev. 2010. How to analyze political attention with minimal assumptions and costs. *American Journal of Political Science* 54(1):209–228.
- Rivero, Gonzalo. 2015. Preaching to the choir. The offline determinants of following members of the US congress on twitter.
- Slapin, Jonathan B., and Sven-Oliver Proksch. 2008. A scaling model for estimating time-series party positions from texts. *American Journal of Political Science* 52(3):705–722.
- Small, Tamara. 2010. Canadian politics in 140 characters: Party politics in the twitterverse. *Canadian Parliamentary Review* 33(3):39–45.
- Spirling, Arthur, and Iain McLean. 2007. UK OC OK? Interpreting optimal classification scores for the UK House of Commons. *Political Analysis* 15(1):85–96.
- Straus, Jacob R., Matthew Eric Glassman, Colleen J. Shogan, and Susan Navarro Smelcer. 2013. Communicating in 140 characters or less: Congressional adoption of twitter in the 111th congress. *PS: Political Science & Politics* 46(1):60–66.
- Struppeck, Thomas. 2014. Combining estimates. *Casualty Actuarial Society E-Forum* 2:1–14.
- Temporão, Mickael, Corentin Vande Kerckhove, Clifton van der Linden, Yannick Dufresne, and Julien M. Hendrickx. 2018. Replication data for: Ideological scaling of social media users. A dynamic lexicon approach, <https://doi.org/10.7910/DVN/0ZCBTB>, Harvard Dataverse, V1.
- Vergeer, Maurice. 2015. Twitter and political campaigning. *Sociology Compass* 9(9):745–760.