Measuring Agenda-Setting Influence from Legislative Speech

Jack Blumenau

University College London

Abstract

Assessing which actors are influential in political debate is important for understanding the mechanisms behind legislative decision-making. Conceiving of 'influence' as a speaker's ability to mould discussion of an issue towards their own framing, I propose a measurement strategy which infers influence by modelling each speech in a debate as a function of the speeches that preceded it. Intuitively, an influential speech is one that is highly predictive of other speeches that occur later in the debate, and influential legislators are those who deliver influential speeches. I apply this method to debates in the UK House of Commons from 1979 to 2018, and compare the measure to potential alternatives in a series of validation tests. I demonstrate the value of this approach by using the measure to address important questions in legislative politics.

This version: April 2018. I thank Ben Lauderdale and Arthur Spirling for helpful discussions.

Introduction

Who is influential in political debate? The speeches that politicians make can be viewed as purposive attempts to alter the perceived issue space in such a way as to change “the interpretation of what is at stake” (Shepsle, 2003, 309) and thereby shift debate outcomes (Riker, 1986). However, we have few empirical measures for determining which actors are successful in this regard. As debate is a central feature in the work of all parliaments, understanding which legislators are most able to shape the debate agenda is crucial for addressing questions of intra-party politics, government accountability, and legislative behaviour. This paper presents an approach for measuring relative levels of agenda-setting influence amongst legislators, which I apply to speeches made by Members of Parliament (MPs) in the UK House of Commons between 1979 and 2018.

Debaters may acknowledge influential speakers by explicitly referring to important speeches, or by implicitly alluding to the ideas and arguments that such speeches express. In the legislative context, explicit references are rare, but I argue that implicit endorsements of influential speakers can be revealed by studying the vocabulary used by participants in political discourse. Intuitively, influential MPs will use language that is adopted and discussed by other MPs in subsequent speeches. When another MP adopts your framing of an issue, they are implicitly indicating that your perspective is important. Influential MPs therefore literally ‘shape the debate’, as the framing they use alters the vocabulary used by other MPs in subsequent discussion.

I capture this intuition by modelling the text of each speech in a debate as a function of the texts of the speeches that preceded it. At the speech-level, I use a penalised regression framework to model the words used in each speech, where the words used by all previous speeches are explanatory variables. The model implies that earlier speeches must be strongly predictive of future speeches, conditional on all other speeches in the debate, in order to be counted as influential. At the debate-level, I rank the influence of specific MPs by accounting for not only the *number* of speakers an MP influences, but also the *importance* of those influenced speakers.

Like other unsupervised measurement approaches, these estimates of MP influence are based on empirical assumptions that might not hold in practice, and therefore validation is essential. I show that, at the debate-level, these scores predict whether an MP's speeches in a particular debate are quoted by the media in coverage of that debate, as well as how often MPs explicitly refer to one another in the debate. At the individual-level, inferred influence maps closely with changes in an MPs' institutional agenda-setting power, and correlates highly with coverage of MPs in national newspapers (Ban et al., 2018). Furthermore, the approach described here outperforms a variety of competing measures across these validation tasks (Fader et al., 2007; Eggers and Spirling, 2016).

I showcase the value of these estimates by applying them to study important questions in the parliamentary politics literature. First, focussing on the effects of parliamentary tenure on legislative behaviour (Benedetto and Hix, 2007; Besley and Case, 1995; Fouirnaies and Hall, 2018), I trace how the influence of MPs in the Commons varies over the course of their parliamentary careers. I show that while the most experienced and least experienced MPs deliver speeches of roughly equal length, more experienced MPs are substantially more influential in the cut-and-thrust of political debate.

Second, I compare MP influence in debates that are associated with votes in which the MP toes the party line to debates where the MP defects from the party. MPs from governing parties make more speeches when they rebel from the party than when they are loyal (Slapin et al., 2018), and I find that, in addition, rebel MPs from the governing party are also more influential in debates where they defect. This emphasises the potential damage that even non-decisive roll-call defections can cause for governing parties: rebel MPs are more important focal points in debate than are loyal MPs.

The paper is organised as follows. Next, I provide the intuition behind the measure, and comment on what it can – and cannot – be expected to capture. I then describe the data and measurement strategy. Following this, I outline alternative influence measures, and compare these measures in a series of validation tests. The penultimate section considers applications, and a final section concludes.

Agenda-setting influence in parliamentary debate

For two days in early 2017, MPs in the House of Commons debated the Government's proposal to initiate the process of leaving the European Union: the start of what has become known as the "Article 50" process.[1] A key speech in the debate was given by Kenneth Clarke, a member of the Conservative Party who has long-argued in favour of the UK's membership of the European Union. Clarke's speech provided both an economic and an ideological foundation for voting against the Government's motion. Drawing on the classic Burkean view of representation, Clarke argued that MPs have a responsibility to follow their individual consciences in deciding how to vote rather than simply following the orders of their constituents. These themes of individual responsibility and conscience voting reverberated throughout the ensuing debate, and were addressed by 'leave' and 'remain' MPs alike. Although Clarke was not a member of the government at this point, and so had no formal agenda-setting powers, his speech undoubtedly shaped the course of the debate. The speech was widely acknowledged in the media as an important and influential contribution,[2] and the speech was referred to both implicitly and explicitly by other MPs throughout the two days of debate. For example, one Conservative MP – Robert Courts – characterised Clarke's contribution as a "masterclass display of oratory and expertise".[3]

Clarke's speech is an example of the type of agenda-setting influence that I seek to measure in this paper. I conceive of 'influence' as the degree to which the contributions an MP makes to parliamentary debate shape the subsequent contributions by other MPs. During debate MPs will raise different issues, some of which will be picked up and developed

[1]Article 50 of the Lisbon Treaty sets out the procedure by which a member state can leave the EU.

[2]Clarke's speech was covered by many news outlets, from all sides of the political spectrum, and described as "barnstorming" (Mirror, Link) "blistering" (New Statesman, Link), "memorable" (Business Insider, Link), and "soberly compelling" (Telegraph, Link). The Guardian described the speech as "a prophet crying in the wilderness", and The Telegraph suggested that Clarke had become a de facto "Leader of the Opposition" (Labour, the official opposition party, were split on their approach to Article 50.) (Link)

[3]Similarly, Nicky Morgan, a Conservative MP, said "I want to pay tribute to my right honourable and learned friend Mr Clarke...for his wonderful speech. Boy, does he show us how it is all done." Clarke also drew support from Labour party MPs such as Clive Efford who said "I find myself in the invidious position of agreeing with virtually everything that was said by Mr Clarke." Additionally, those on the 'leave' side of the argument also credited the relevance of Clarke's speech. For example, Conservative MP Nigel Adams said that "Mr Clarke is always eloquent and impassioned. Occasionally he is wrong, but it was great to hear from him." All quotations in House of Commons (2017)

by other members in subsequent speeches while others will be ignored. Having another MP pick up on your framing of an issue is a way of controlling how the debate proceeds: it means that other people are taking up your perspective, whether they agree with it or not. The underlying measurement assumption, then, is that the influence of a speaker can be inferred from the degree to which that speaker's words are echoed by subsequent speakers in the same debate. When the language in one speech is adopted by other MPs, this indicates that the speech is influential in the debate. Speakers who make influential speeches are themselves considered to be influential.

This definition of 'influence' does not imply that a speech shifts the *preferences* of other MPs, at least not directly. Instead, I evaluate the degree to which an MP's speeches determine *how* an issue is discussed by the parliament as a whole. Agenda-setting influence is therefore a quality that resides in a speaker's ability to mould discussion of a given issue towards their own framing. One can imagine that industrious agenda-setters may shift policy by seeking "to frame things in the light most favourable to their cause." (Shepsle, 2003, 313) and that strategic framing of an issue can shift the "evaluation of an activity by changing the way it is perceived." (Finlayson and Martin, 2008, 452) Indeed, the idea that reframing issues in a different light can allow agenda-setters to nudge policy towards their ideal has a long history in political science (Riker, 1986, 1996), but I do not evaluate such claims directly. Rather, I suggest a method for measuring the relative agenda-setting effectiveness of different political actors in the context of legislative debate.

Political scientists have focussed heavily on two main quantities of interest when using legislative speech as a source of data: first, for scaling the positions of political actors (Laver, Benoit and Garry, 2003; Slapin and Proksch, 2008; Lauderdale and Herzog, 2016); and second, for measuring the priorities that politicians place on different topics and issues (Quinn et al., 2010; Hopkins and King, 2010; Grimmer, 2010; Roberts et al., 2014). For both scaling and topic models, texts are clustered in a latent space according to the dominant sources of variation (ideological position, topic, etc) in word-use across documents. By contrast, my method traces the diffusion of word-use throughout a debate, and characterises each text by the degree to which it predicts the words used in future texts.[4]

The central problem in estimating patterns of influence from the words used in sequences of texts is that it is difficult to distinguish instances where one text influences another from instances where two texts simply share words from a common topic. I address this issue in a number of ways. First, similar to recent approaches (Laver and Benoit, 2002; Herzog and Benoit, 2015; Lauderdale and Herzog, 2016), I use legislative *debates* as a natural way to group legislative *speeches*. Conditioning on debate means that much of the topic-driven variation in word-use across debates is held constant. Second, while both scaling and topic models treat debates as a set of unordered texts, I make speech order an important organising feature in the modelling process, using it to reflect the intuition that earlier speeches can influence later speeches, but not vice versa. Finally, I use a penalised regression framework for modelling the flow of word-use between speeches, which means that the words used in one speech must be highly predictive of the words used in subsequent speeches in order for the speech to be considered influential in the debate.

My approach is subject to a number of assumptions. Most importantly, the method assumes that debate-level patterns of word-use can reveal something useful about relative levels of agenda-setting influence between MPs. To simplify, I argue that if the words used in speech A are highly predictive of the words used in speech B, then this indicates that speech A influences speech B. Further, I assume that a speech that influences many other speeches is influential, and speakers who deliver many such influencing speeches should also be considered influential. Finally, I assume that speeches within a given debate can influence one another, but speeches in different debates cannot, and that earlier speeches can influence later ones, but not vice versa. These assumptions may, of course, be wrong, and so validation of the resulting measures is essential. In the next section I outline the measurement approach, and in subsequent sections I validate this approach at both the debate-level and individual-level.

[4]The two other studies in political science (Fader et al., 2007; Eggers and Spirling, 2016) that focus on the measurement influence in political debate are discussed in greater detail below.

Modelling 'influence' in ordered sequences of texts

I work with data from all debates in the UK House of Commons between 3rd May 1979 and 1st March 2018. Data is retrieved from `TheyWorkForYou.com`, a public website that catalogues parliamentary proceedings. A *speech* is each uninterrupted utterance by an MP and a *debate* constitutes a number of speeches on the same topic on the same day.[5] The full corpus includes 60,401 debates and just over 1.5 million speeches.

Penalised text regression

I characterise a debate as an ordered sequence of speeches (indexed as $s \in 1, ..., S$), where each speech is represented by a vector w_s of length V where V is the number of unique words in the debate.[6] Each entry (w_{sv}) in such a vector is the number of times word v appears in speech s divided by the total number of words in speech s. Each speech is therefore represented as a vector of word proportions which sum to 1, where larger entries correspond to words that are used more frequently within the speech. The goal is to model the word-proportions of a given speech as a function of the word-proportions of all previous speeches that occurred in the debate.

Consider the relationship between speech i and speech j, where j occurs after i in debate. A simple approach would be to regress w_j on w_i, and record i as influencing j if the coefficient from such a model is greater than 0. A regression coefficient of 1 would indicate that the word proportions of speech i were identical to those of speech j. However, establishing whether speech i *influences* speech j is complicated by the fact that i can be just one of a sequence of prior texts $(1, ..., i, ..., j-1)$ that might influence j. That is, in a set of ordered texts, there are $j-1$ possible influencers of text j.

Imagine we have a sequence of three texts where $h < i < j$ and where we are interested in predicting the word use of j. Although j may share many of the same words as i, both may also share words with h. If this is the case, then any relationship between i and j

[5] In practice, debate identifiers are given in the `TheyWorkForYou` data.

[6] In terms of preprocessing, I convert all text to lower case, remove stop-words and punctuation, and then restrict attention to the 30,000 most common words in the corpus as a whole.

may disappear once we condition on h. This is just the usual concern that any bivariate relationship between i and j may be confounded by h. In substantive terms, j may use similar words to i, but both may really be influenced by the truly agenda-setting speech h. As debates will include words common to many texts, estimating the bivariate relationship between speech pairs will generally overstate the true relationship between speeches.

To overcome this problem, we can model w_j as a function of the words of both of the previous two texts:

$$w_j = \beta_i w_i + \beta_h w_h \tag{1}$$

where β_i can be viewed as the influence of i on j conditional on h, and β_h the influence of h on j conditional on i. We can generalise this approach by modelling a given speech as a function of all previous speeches in the debate:

$$w_j = \sum_{s=1}^{j-1} \beta_s w_s \tag{2}$$

Estimating equation 2 for speech j results in a vector of β coefficients, one for every speech that occurred prior to j ($s \in 1, ..., j-1$). Each β_s represents the influence of speech s on speech j, conditional on all other speeches in the sequence. At the debate-level, we therefore estimate equation 2 for each speech (except the first) in the debate, storing the resulting coefficients in a speech-by-speech matrix, D. The typical element of this matrix – $D_{i,j}$ – is the coefficient for the ith speech from the jth regression model and represents the degree to which the word use in speech i is predictive of the word use in speech j, conditional on all other speeches that occurred prior to speech j. Equation 2 is defined only for $s < j$, meaning that earlier speeches can influence later ones but not vice versa, and so values in the upper triangle of D will be missing. Further, when the ith speech gets an estimated coefficient of 0 from the jth model, the corresponding element of D will be 0. Finally, as some speakers give multiple speeches in a debate, but we do not want a speaker to be able to 'influence' herself, the relevant entries in D are also coded as missing.[7]

[7]Note, however, that all prior speeches made by the same speaker of speech j *are* included in the

Each speech is modelled as a function of all the speeches that preceded it, and so as the debate grows in length there are very many possible sources of influence for each speech. For example, when $j = 100$, there are 99 prior speech vectors included in the model. We would like it to be 'hard' for one speech to influence another, in the sense that speech i should be highly predictive of speech j – conditional on all other speeches – in order for us to count i as having influenced j. I therefore estimate equation 2 with the LASSO method proposed by Tibshirani (1996). The LASSO estimates of the β parameters in are obtained by minimizing the sum of the squared residuals subject to an L_1 norm:

$$\hat{\beta}^{lasso} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^{N} \left(w_{j,i} - \sum_{s=1}^{j-1} \beta_s w_{s,i} \right)^2 + \lambda \sum_{s=1}^{j-1} |\beta_s| \right\} \tag{3}$$

where λ is a penalty parameter that controls the amount of shrinkage in the regression. A key advantage of the LASSO approach is that it will induce sparsity in the resulting coefficient estimates, with many of the βs shrunk exactly to zero. The larger λ becomes, the greater the amount of shrinkage, and thus the greater the number of β coefficients that will shrink to zero. Intuitively, this implies that the larger the λ, the stronger the (conditional) relationship between two speeches needs to be before we count one speech as influencing another.[8] Estimating equation 3 for each speech results in a sparse D matrix for the debate, where most speech-pair coefficients are equal to zero.[9]

relevant regression for speech j. As word use is likely to be correlated between speeches made by the same member, excluding these speeches from equation 2 would likely lead to an overestimate of the influence of speeches by other members. It is only when constructing the speech-by-speech coefficient matrix D that self-referencing speeches are excluded.

[8] I select λ using K-fold cross-validation for each regression. A more nuanced alternative approach would be to allow for different penalty factors to apply to each coefficient in the model, in the spirit of Zou (2006):

$$\hat{\beta}^{lasso} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^{N} \left(w_{j,i} - \sum_{s=1}^{j-1} \beta_s w_{s,i} \right)^2 + \lambda_n \sum_{s=1}^{j-1} \lambda_{n,s} |\beta_s| \right\} \tag{4}$$

where λ_n is a grand penalty parameter, and $\lambda_{n,s}, s = 1, ..., j-1$ are specific penalty parameters associated with each prior speech in the debate. Imposing the constraint that $\lambda_{n,j-1} < \lambda_{n,j-2}$, for example, would make it 'harder' for the speech two positions away from j to influence j than the speech that occurred immediately prior to j. This approach is attractive intuitively, as it would suggest that influence is mostly 'local' in debates, but selecting appropriate $\lambda_{n,s}$ values in a data-driven way is non-trivial (Bergersen, Glad and Lyng, 2011). I intend to explore a model similar to 4 in future iterations of this work.

[9] In the full set of debates described below, the D matrices are 71% sparse on average.

Aggregating speech-level scores and ranking speakers

The D matrix gives a complete set of speech-by-speech influence scores in a given debate. However, we are ultimately interested in ranking the influence of *speakers* rather than characterising the flow of influence between individual *speeches*. There are two related issues that need to be addressed in the aggregation stage.

First, as some MPs deliver many speeches in a given debate, we require a method to aggregate the entries in D to the speaker level. To do so, I create a new matrix, $\tilde{D}$, with typical entry $\tilde{D}_{i,j}$, which is simply the sum of the LASSO coefficients of speeches by speaker i for speeches by speaker j. For instance, if speaker i gave two speeches, and speaker j gave only one speech, then $\tilde{D}_{i,j}$ would be the sum of the relevant two coefficients from the LASSO regression for speaker j's speech. This speaker-by-speaker matrix therefore represents the degree to which any given speaker (rows) influences any other speaker (columns) in a debate.[10] Note that $\tilde{D}$ is not symmetric ($\tilde{D}_{i,j} \neq \tilde{D}_{j,i}$), and thus can be viewed as a directed network, where speakers are the nodes, and the edges are described by the cell values.

Second, these speaker-by-speaker scores need to be aggregated in a way that reveals a sensible ranking of influential speakers in the debate. Consider the elements of $\tilde{D}$ in a simple (imaginary) example depicted in figure 1. In this debate, there are five speakers, four of whom are backbenchers (B1 to B4) and the fifth is the prime minister (PM).[11] Here, B1 (top row) influences B2, B3 and B4. By contrast, B2 only influences the Prime Minister, B3 only influences B4, and B4 only influences B2. The Prime Minister is clearly influential in the debate, as she influences each of the backbenchers.

Given $\tilde{D}$, what is an appropriate way of ranking the influence of these speakers? The most straightforward approach would be to simply count the number of speakers that a given individual influenced. Here, the Prime Minister being would be most influential, with

[10] One potential problem with this aggregation approach is that it is likely to reward MPs who make many speeches relative to those who make few speeches. This is a reasonable concern, but only to the extent that those additional speeches have non-zero LASSO coefficients. That most speech-pairs have coefficients of zero mitigates this problem to some degree. Furthermore, in the validation checks below I compare the LASSO approach to a simpler approach where speakers who deliver many speeches (or long speeches) as being more influential. The LASSO approach clearly outperforms these simpler measures.

[11] In this example, for clarity, I use a binary variable to indicate whether one speaker influenced another.

$$
\begin{array}{c}
 \\
B1 \\
B2 \\
B3 \\
PM \\
B4
\end{array}
\begin{array}{c}
\begin{array}{ccccc} B1 & B2 & B3 & PM & B4 \end{array} \\
\begin{bmatrix}
0 & 1 & 1 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
1 & 1 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 0
\end{bmatrix}
\end{array}
\quad
\begin{array}{cc}
\text{Degree Rank} & \text{PageRank} \\
2 & 3 \\
3 & 2 \\
3 & 5 \\
1 & 1 \\
3 & 4
\end{array}
$$

Figure 1: $\tilde{D}$ example

B1 second, and B2, B3 and B4 tied in third place ('Degree Rank' in figure 1).

However, this simple ranking counts all speaker-by-speaker influence scores equally. This method seems problematic, as it misses an important feature of this debate: although B2, B3 and B4 all only influence one other speaker, and thus receive the same ranking, B2 is the only speaker to influence the Prime Minister who is herself very influential. It seems clear that B2 should be counted as more influential than B3 or B4, as her speeches influence the Prime Minister who in turn influences other backbenchers, whereas B3 and B4 only influence other low-influence speakers. So, in addition to the *quantity* of people an MP influences, the *quality* of those influenced people should matter when we rank MPs.

The upshot is that we may prefer an aggregation approach that gives additional weight to speakers who influence other influential speakers. Ranking problems of this type have been well studied in the computer science literature (Page et al., 1999; Kleinberg, 1999).[12] One such method that captures this idea is to treat $\tilde{D}$ as a network, where each speaker is a node and the entries of $\tilde{D}$ represent the (weighted) edges between speakers (Mihalcea, 2004; Erkan and Radev, 2004; Fader et al., 2007). Given this graph, then an intuitive way of formulating this idea is to imagine that the influence score for each speaker is a function of the influence scores of the speakers that she influences:

$$p(i) = \sum_{j \in adj(i)} \frac{\tilde{D}_{i,j}}{\sum_{k \in adj(j)} \tilde{D}_{k,j}} p(j) \tag{5}$$

where $p(i)$ is the influence of speaker i and $adj(i)$ is the set of speakers that i influences.

[12] In related work, (Fader et al., 2007) use a similar approach to determine central speakers in the US Senate, and I discuss the differences between their work and my own in more detail below.

This formulation emphasises that the influence of speaker i is determined by the degree to which she influences other speakers ($adj(i)$ and $\tilde{D}_{i,j}$), and, crucially, by the influence of those speakers that are them selves influenced by i ($p(j)$). Erkan and Radev (2004) show that the vector of speaker-level influence scores, P, can be calculated as the left eigenvector of the row-normalised speaker-matrix $\tilde{D}$ via the *PageRank* algorithm (Page et al., 1999), which I employ here.[13] Because the absolute values of P are sensitive to debate-length, I normalise the debate-level influence scores to the unit interval throughout the paper in order to be able to make comparisons across debates.[14]

Applying *PageRank* to the example in figure 1 results in a subtly different ranking of speakers. In this example, the new ranking is: $PM = 1, B2 = 2, B1 = 3, B4 = 4, B3 = 5$. Importantly, while B2 was ranked joint last using the simple method, using *PageRank* elevates B2 to the second most influential speaker in this debate. This is because B2's influence score is driven up by her influence over the Prime Minister, who is the most influential member of the debate. The Prime Minister remains the most influential member, and B1 moves from 2nd to 3rd.

Although I present more formal validity checks below, it is useful to briefly consider the estimates that the method recovers for some individual debates. Figure 2 visualises the normalised scores from equation 5 for two debates for the most recent parliamentary term. In each plot, grey points represent individual speeches (scaled proportionally to speech length), and speeches made by the same MP are linked by a horizontal line. MP names are scaled proportionally to the influence score of the MP. The x-axis gives the position of the speech in the debate, and the y-axis gives the influence score of the speaker.

The left-hand plot presents the influence scores for the debate on Article 50 described previously. The most influential speakers are David Davis – the Secretary of State for the Department of Exiting the EU; Keir Starmer – the relevant Shadow Secretary of State; and Alex Salmond – the SNP's Europe Spokesman. That the measure identifies these

[13] Mihalcea (2004) shows that either the Kleinberg (1999) *HITS* algorithm or *PageRank* can be used to calculate P and that both perform well in approximating human judgements. *PageRank* can also be viewed as an approximation of the Bradley-Terry model for paired comparisons.

[14] $p(i)' = \frac{p(i)-min(P)}{max(P)-min(P)}$, where P is the vector given by the *PageRank* algorithm for that debate.

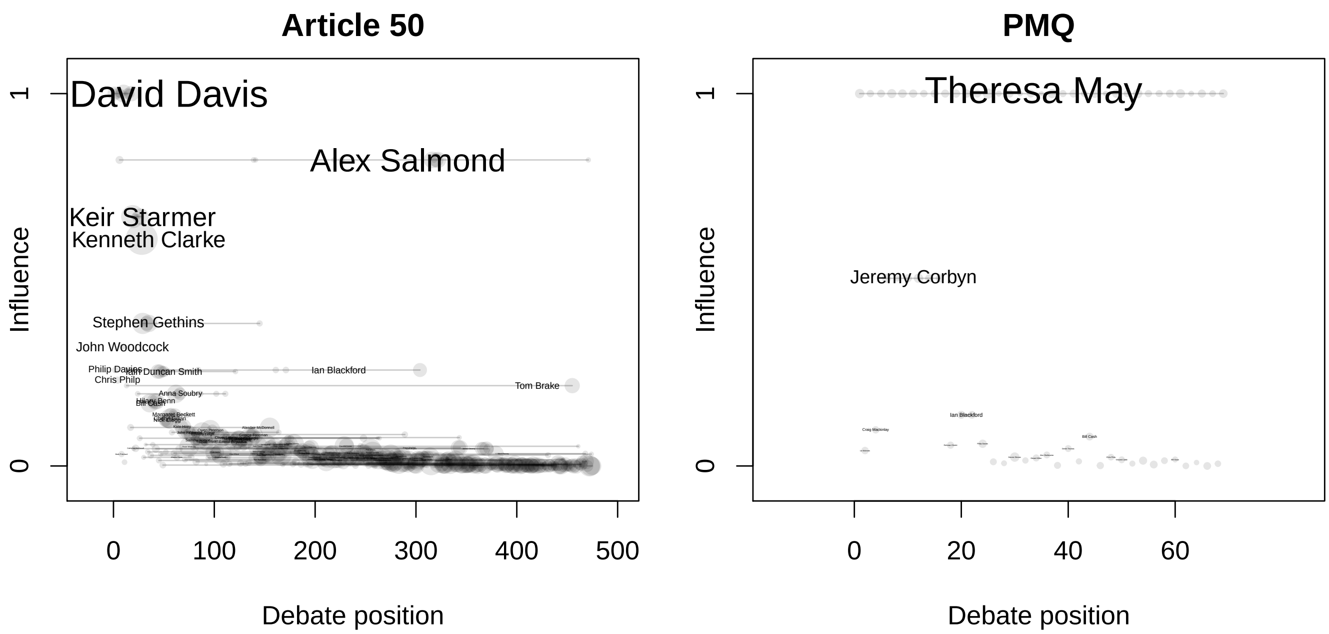

Figure 2: *Influence scores in two example debates*
The plots visualise the influence scores as defined in equation 5 for two example debates. The left-hand plot shows the scores for the Article 50 debate, and the right-hand plot shows scores for Prime Minister's Question Time (PMQ).

institutionally powerful actors is a encouraging, but it also captures the influence of MPs who do not hold such positions but who we would nonetheless expect to be influential. Most notably, it is reassuring that Kenneth Clarke – whose speech was described above – appears prominently in the plot, demonstrating almost equal influence with Keir Starmer.

The right-hand plot presents scores from a Prime Minister's Question Time (PMQs) debate, a weekly event in the parliamentary calendar when the Prime Minister is questioned by MPs of all parties. Theresa May, the current Prime Minister, is the most influential MP in this debate, followed by Jeremy Corbyn who is the current Leader of the Opposition. All other MPs in this debate receive fairly low influence scores. Again, this corresponds to intuitive notions of influence in such debates: the Prime Minister has opportunities throughout the debate to communicate the government's agenda, and the Leader of the Opposition has procedural privileges which allow him to repeatedly question the Prime Minister. Taken together the plots indicate, at least in these debates, that the measurement strategy described above provides plausible estimates of MP influence in the Commons.

Aggregating and ranking influence across debates

The previous subsection described a method for aggregating from a speech-by-speech influence matrix, D, to a speaker-by-speaker matrix, $\tilde{D}$, for a given debate, and then a ranking method for producing influence scores for each MP. While single debates provide a natural grouping for making comparisons between parliamentary speeches (Lauderdale and Herzog, 2016), this general approach can also be applied to any arbitrary grouping of debates.

For instance, say we were interested in ranking MPs across K debates in a particular policy area. To do so, we could collect the specific debates of interest, and then estimate debate-specific speech-by-speech influence matrices using the LASSO approach described in equation 3, repeating the process for each debate that matched our selection criteria. This would result in K separate D_k matrices, which could be combined to form a single $\tilde{D}$ speaker-by-speaker matrix, where the typical element $\tilde{D}_{i,j}$ is the sum of the LASSO coefficients relating to speeches by speaker i for speeches by speaker j, *across all debates*. $\tilde{D}$ will therefore capture the patterns of influence between all MPs who participated in a given set of debates, and applying equation 5 to $\tilde{D}$ would produce a ranking of MP influence across these debates.[15] The selection of which debates to use could be driven by a researcher's particular area of interest, including ranking MPs in particular policy areas, time periods, or in different types of debate. I use this more general approach to ranking MPs in some of what follows.

Alternative strategies

In this section I briefly outline some alternative approaches to measuring influence. These alternatives measures aim to capture slightly different concepts, and so comparisons across methods will not reveal which is 'better' in any general sense. Nevertheless, it will be informative to use these alternatives as baselines for the validity tests that follow.

[15]It is worth noting that this approach maintains a key strength of the strategy outlined above, as the individual speech-influence scores are still determined at the debate-level. While I calculate D at the debate-level, $\tilde{D}_{i,j}$ represents the overall degree to which one MP influences another over all debates.

MavenRank

The work most closely related to my own is that of Fader et al. (2007), who also analyse the rhetorical influence of US Senators in legislative debate with an approach they title `MavenRank`. Fader et al. first concatenate all speeches made by the same member on the same topic, and then measure the cosine-similarity between the (tf-idf weighted) word vectors for each speaker. This results in a similar speaker-by-speaker matrix as described above ($\tilde{D}$), with the difference that the cells give the similarity between the speech documents rather than the coefficients from LASSO regressions. With this matrix in hand, they apply *PageRank* to produce individual-level influence scores.

There are two main differences between this approach and my own. First, the central intuition of the measurement strategy here is that *debate-level* patterns of word use are revealing of influence between politicians. In MavenRank, by contrast, debate-groupings are ignored, and so too is speech order as all texts by the same speaker within the same topic (across many debates) are concatenated into a single document before being analysed.

Second, because the cosine-similarity relation is symmetric, so to is the speaker-by-speaker matrix in the MavenRank approach. By contrast, the LASSO approach implies an asymmetric $\tilde{D}$ matrix. A consequence of this is that graph analysed in the ranking step in `MavenRank` is undirected, whereas the graph I investigate is directed. Although both directed and undirected networks can be ranked via the *PageRank* algorithm, the resulting influence scores are very different. In addition, the LASSO approach takes account of the possibility that word similarities between speeches might be confounded by more general patterns of word use in debate, something that is ignored in the MavenRank framework.

Burstiness

Eggers and Spirling (2016) introduce a different approach to identifying important agenda-setting speakers in parliamentary debates in the 19th century House of Commons. Taking inspiration from the computer science literature on document 'streams' (Kleinberg, 2003), Eggers and Spirling model parliamentary speeches using a measure that considers the `burstiness` of word use over time. The intuition behind their approach is to detect spikes in the use of a word, and to attribute importance to the MPs who begin such spikes, viewing this as an indicator of their latent agenda-setting ability.

Seniority and Speechiness

Beyond these text-based approaches, I also consider whether agenda-setting influence can be captured by simpler measures. First, parliamentary actors with formal institutional powers to set the legislative agenda should also be more influential in the course of legislative debate. Government cabinet ministers, as well as members of the opposition 'shadow' cabinet, have institutional privileges when it comes to suggesting legislation and proposing debates in parliament. These actors speak more often than other members, and deliver longer speeches on average. A simple baseline measure is therefore a binary indicator for `Frontbench` MPs, coded as 1 if an MP held a government or shadow cabinet ministerial position at the time of a given debate. Similarly, I also evaluate whether the number of words (`N words`) or the number of speeches (`N speeches`) are predictive of debate influence.

Validating debate influence

Validation is particularly important for methods that use automated text analysis (Grimmer and Stewart, 2013). In this section, I use a variety of validation tasks to evaluate the LASSO strategy, and to make comparisons with other approaches.

Debate-level validity

The primary measurement goal is to identify which MPs are influential in shaping the agenda in individual parliamentary debates. To validate the measure at the debate level, I draw on two sources of data which are proxies for debate-level agenda-setting influence.

First, parliamentary politics in the UK are reported on extensively in the national media, and parliamentary debates have played a central role in that reporting since at least the 19th century (Erdman, 1960). One useful source of parliamentary coverage is the BBC's Today in Parliament (TiP) radio programme which includes short excerpts from the most salient speeches to be made on the floor of the Commons each day. The speeches selected for broadcast are drawn from a wide variety of debates, including key debates on legislation, ministerial question time, and debates selected by the opposition parties. Although the BBC is required to pursue 'impartial' coverage of different parties in the speeches it selects, there is generally a wide mix of speeches from different parties and from both frontbench and backbench MPs.[16] This is a valuable source of validation data, as it provides a natural coding – was a speaker featured in the radio programme or not – of the relative importance of MPs in a given debate. Although I show that the LASSO approach also correlates highly with more general media coverage in the next section, this data allows a more fine-grained evaluation of the measure at the debate level.

I collect data on the coverage of parliamentary debates in TiP for 17 days in January and February 2018. The data comes from 78 debates held across this period, and includes 1749 speaker-debate observations. For each speaker in each debate, I code whether the MP was featured in the TiP programme or not, and use this as the dependent variable in a series of logistic regressions, where for each regression the predictor variable is the influence scores from one of the methods described above. To make the scores comparable across debates, I scale each score within each debate to the unit interval. The intuition is that if the scores accurately capture the idea of speaker influence, they should be predictive of whether a given speaker was subsequently covered in the TiP programme. I evaluate the predictive power of each measure by using K-fold cross-validation for each regression, and the average cross-validation error for each measure is given in figure 3.

Second, I also compare the influence scores to the number of times an MP is directly mentioned by other MPs in during a debate. MPs follow strict conventions when commenting on other members, referring to the constituency that the member represents rather than their name (i.e. MPs will address the "Member for Hornsey and Wood Green" or the "Honourable Member for Enfield North", etc). Constituency names are unique to each MP, and so I search the entire corpus of debate texts between 1979 and 2018 for direct mentions of each backbench MP using their constituency identifiers, and normalise the number of

[16] In the sample here, 27% of broadcast speeches were given by frontbench MPs and 73% by backbenchers.

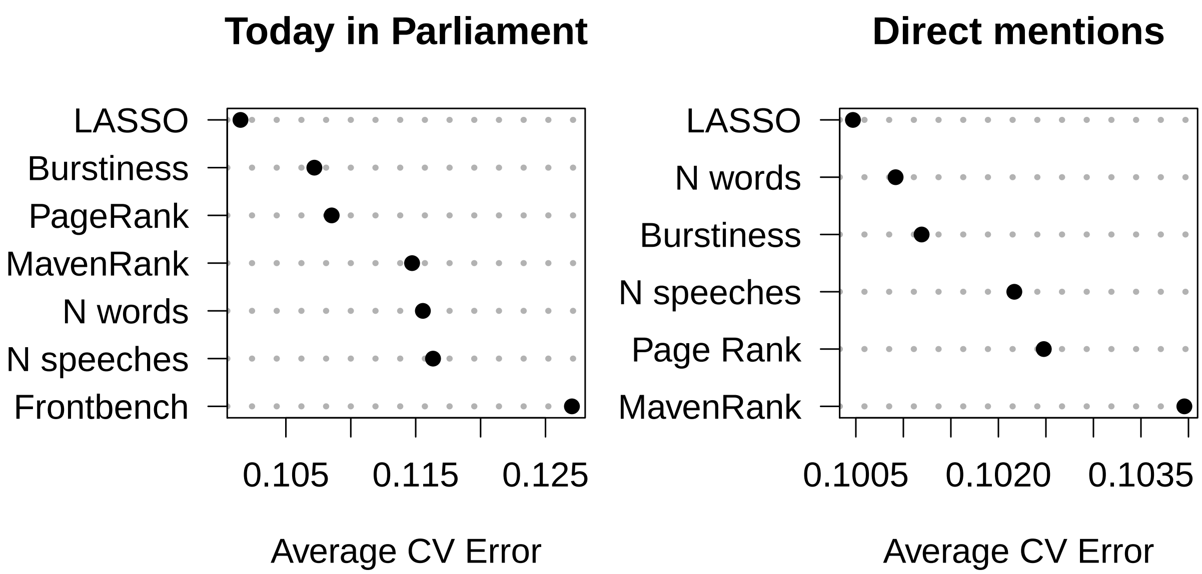

Figure 3: *Debate-level validity*
The plots present cross-validation errors of different 'influence' measures for (left) coverage of MPs in the Today in Parliament programme, and (right) the normalised number of times a back-bench MP was directly mentioned by name in debate. Left panel is based on logistic regressions of 1749 observations in 78 debates in January and February 2018. Right panel is based on linear regressions of nearly 600,000 observations from 1979 to 2018.

mentions of MPs within each debate to the unit interval. It seems clear that an MP who is directly mentioned by many other MPs in their speeches is playing an important role in the debate and the LASSO influence score should correlate positively with the number of direct mentions that an MP receives. I use the normalised counts of explicit references to each MP as a dependent variable in a series of linear regressions, again with each of the unit-normalised influence scores as the sole predictor variable. The average cross-validation error from these models are presented in the right panel of figure 3.[17]

The main message from both of these validity checks is clear: at the debate-level, the LASSO approach outperforms all alternative measures discussed above. The predictive error across these two very different validation strategies is smaller for the LASSO model

[17]The convention for referring to government and opposition frontbench MPs is somewhat different, and does not allow me to count direct references for those MPs, and so in this analysis I subset the data just to backbench MPs. As a consequence the `Frontbench` dummy variable is not considered for this task.

than for other models. In addition, while the relative performance of the LASSO approach is superior in these tests, it also performs well in absolute terms. For the TiP data, the average correlation between the LASSO based influence scores and coverage in TiP is 0.55. The correlation with the number of direct mentions is somewhat lower, but still positive at 0.24. Overall, these validity checks suggest that the LASSO approach identifies important speakers in the cut-and-thrust of individual debates.

Individual-level validity

In this section I evaluate the validity of the LASSO influence measure at the level of the individual MP. One basic check on the face validity of this measure is to evaluate the degree to which actors who hold positions of institutional agenda-setting power are also marked by high levels of influence in parliamentary debate. In the House of Commons, government cabinet ministers control the direction and implementation of government policy, and are dominant in the political process at Westminster. In the context of floor debates, government ministers speak frequently to propose legislation, setting the agenda for the debate that follows. If the LASSO measure is valid, it should clearly distinguish government ministers as being more influential than other MPs.

Table S1 in the appendix presents results from a model where I regress the LASSO influence score from equation 5 for each MP in each debate from 1979 to the present on a binary indicator for whether the MP in question was a current cabinet minister. I also include an additional explanatory variable for *shadow* cabinet ministers, who also have (more limited) agenda-setting powers. The LASSO measure is strongly associated with institutional status: cabinet ministers are more than three times as influential as backbench MPs, and nearly twice as influential as shadow cabinet ministers. While reassuring, simple comparisons between frontbench and backbench MPs may reflect more systematic differences between these groups, rather than changes in influence that result from institutional position. For example, MPs who demonstrate their aptitude as debaters may be more likely to be appointed as cabinet ministers. To control for potential baseline differences, I instead focus on within-MP changes in influence over the course of an MP's career.

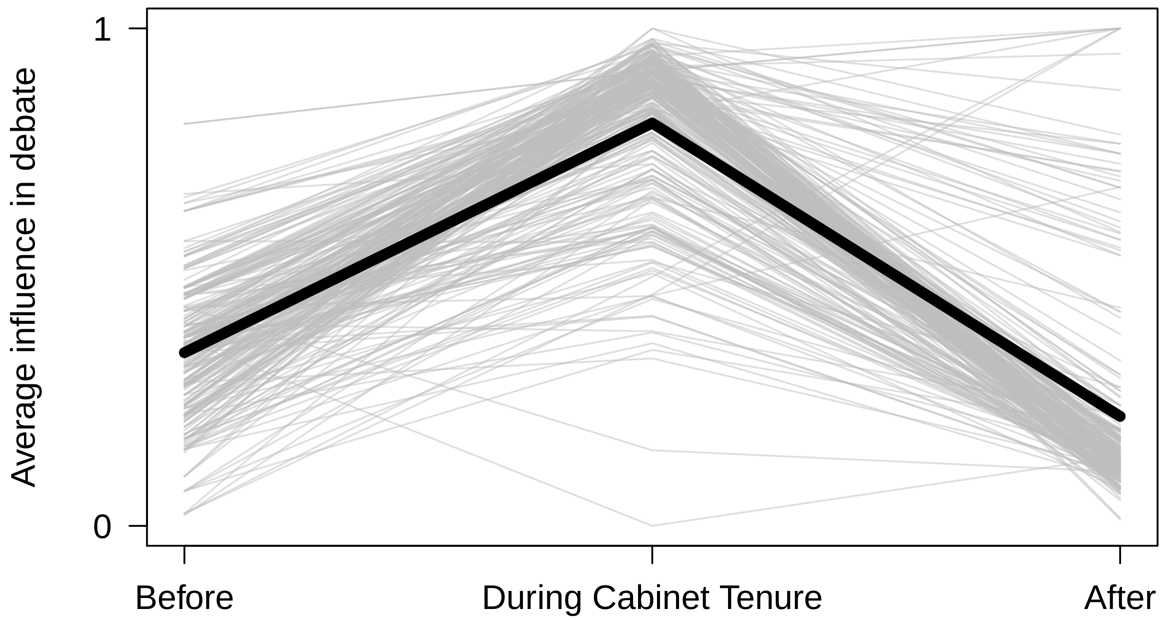

Figure 4: *Cabinet member debate influence before, during and after cabinet tenure.* MP influence increases significantly after being appointed to cabinet ministry positions. Grey lines represent individual MPs and the thick black line represents the average across all MPs who held a cabinet position between 1979 and 2018.

Figure 4 illustrates within-MP changes in influence for those MPs who held a cabinet position at some point during their time in parliament. For each cabinet-serving MP, I calculate the average LASSO score across all debates in which they spoke for three time periods: before, during, and after their tenure in the cabinet. I plot the influence trajectory for each MP (grey lines) and the average across all MPs (black line). The plot clearly reveals that MPs become more influential in debate when they are promoted to positions in the cabinet, and become less influential once they leave office. The average pre-cabinet influence score is 0.34, increasing to 0.81 during cabinet tenure, before declining to 0.22 in an MP's post-cabinet career. The difference between the pre- and post-cabinet scores might be accounted for by the fact that most cabinet ministers hold junior government positions before they are promoted, but return to the backbenches (or leave parliament altogether) after they leave the cabinet. Regardless, figure 4 clearly demonstrates that the LASSO scores capture important within-MP changes in agenda-setting power.

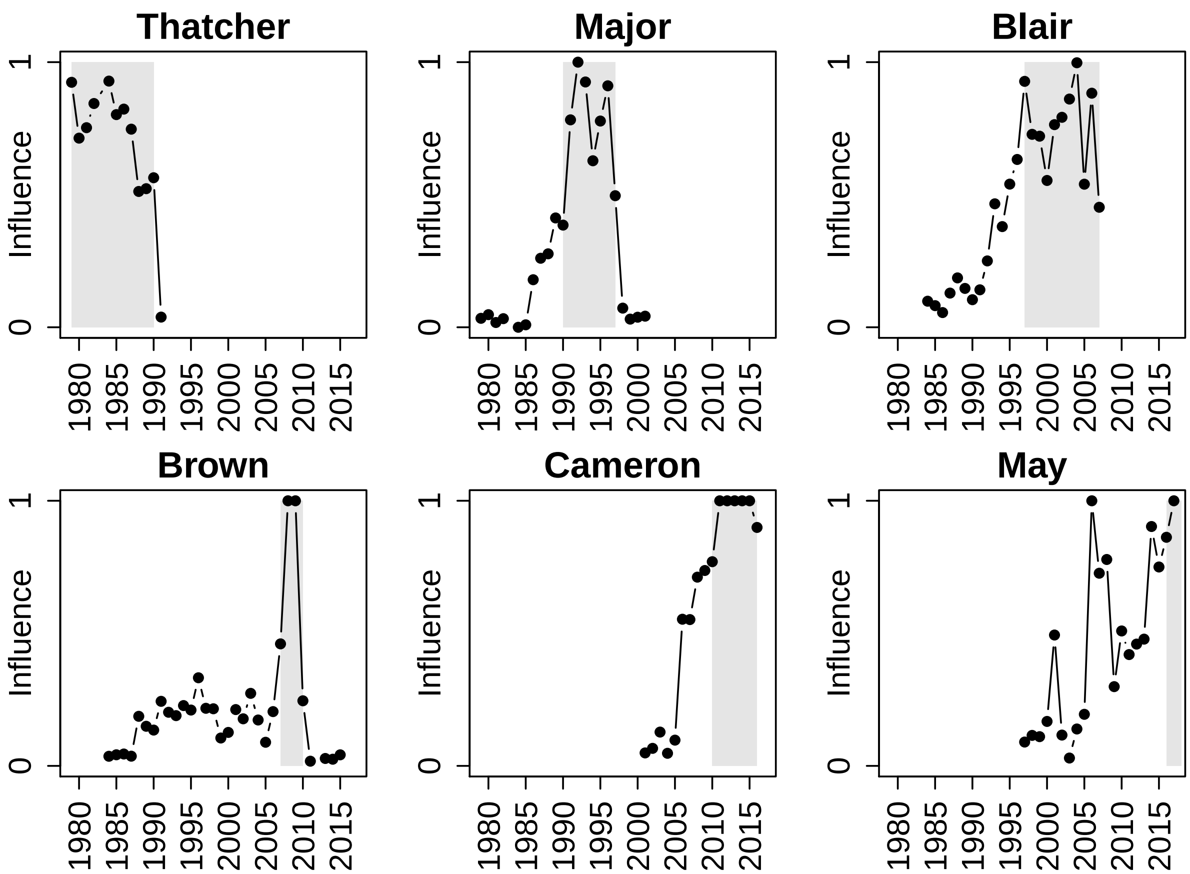

Figure 5: *Debate influence of the prime minister, by year.*

We can also chart the influence of some well-known MPs over time, and to see whether the change in their influence over time is consistent with prior expectations. Here I aggregate the speech-by-speech matrices for all debates within a given calendar year to a single speaker-by-speaker matrix for that year. I then apply *PageRank* to these yearly $\tilde{D}$ matrices to produce a ranking of MPs within each year.[18]

Figure 5 plots the resulting estimates for the six MPs who held the position of Prime Minister during this time period (grey-shaded areas in the plot indicate the periods in which the relevant Prime Minister was in office). It is clear that these six MPs were substantially more influential as Prime Minister than they were either before or after their tenure. In addition, the LASSO measure also seems to capture other salient details of career

[18]As with the debate-level scores, I normalise these yearly influence scores to the unit interval.

progression. For example, we accurately identify periods in which MPs held the position of Leader of the Opposition: David Cameron (bottom-middle) becomes more influential around 2005, and the same can be said for Tony Blair (top-right) around 1994. Similarly, Gordon Brown (bottom-left) becomes more influential after his appointment to the shadow cabinet in 1987 than he was before that date.

A more vigorous validity check of these influence scores at the individual level requires making comparisons between the LASSO method and the alternative strategies detailed above. Unlike the debate-level validation exercise, here we require some measure of the overall influence of individual MPs at a given point in time. Ban et al. (2018) argue that relative levels of newspaper coverage of different political actors can be used to measure the relative power of those actors over time. In particular, they suggest that the following metric can be used to measure the relative power of different actors:

$$\text{Relative coverage of MP}_{it} = \frac{\text{Newspaper Mentions of MP}_{it}}{\sum_j^N \text{Newspaper Mentions of MP}_{it}} \tag{6}$$

I collect all newspaper articles from The Guardian newspaper from 2010 to 2017 which are related to politics, and search these texts for the names of the MPs in parliament during each year.[19] For each MP i in each year t I calculate equation 6. As Ban et al. suggest that their measure should be treated as ordinal rather than cardinal (p. 5), I then rank the newspaper coverage scores within each year and rescale the MP rankings to the unit interval. I use these scores as the dependent variable in a series of linear regressions, where the sole predictor in each regression is one of the influence scores described above.[20] In

[19]The Guardian provides a publicly available API for searching the newspaper's archives, and 'tags' articles as being relevant to different domains. I downloaded all articles with the 'politics' tag, which returns 19844 articles. I search each article for the name of each MP in parliament during this time period, using multiple search strings for MPs whose names are often reported differently in the media than they are in parliamentary proceedings (for example, Edward Balls becomes Ed Balls, and so on). The Guardian is a major national broadsheet newspaper in the UK, and has a left-of-centre political orientation. Ideally it would be helpful to replicate this analysis using other news sources, but fully searchable texts are not available for other publications.

[20]In order to make fair comparisons, I 1) calculate `MavenRank` by concatenating all speeches by a given MP in a given year, and then construct cosine-similarity matrices at the year level before applying *PageRank*; 2) calculate `Burstiness` by summing the burstiness score of each MP across all debates in each year; 3) calculate `N speeches` and `N words` as the sum of the number of speeches and number of words by an MP in a given year; 4) code an MP as being a member of the `frontbench` if that MP held a cabinet or shadow cabinet position at any point during the relevant year. All scores are unit-normalised for each year.

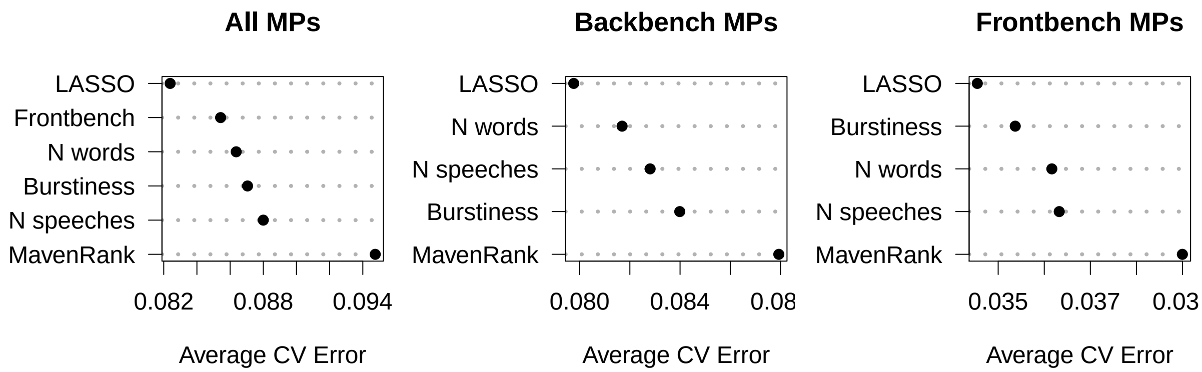

Figure 6: *Debate influence and relative media coverage.*
Plots present cross-validation errors of different 'influence' measures for predicting the relative coverage of MPs (equation 6) in The Guardian newspaper (2010 to 2017).

order to evaluate performance for different MP-types, I run one set of regressions for all MPs, one set for only frontbench MPs, and one set for only backbench MPs. I again use K-fold cross-validation for evaluation, and present average errors in figure 6.

The results in figure 6 suggest that at the individual-level, the LASSO approach outperforms all alternative measures of influence discussed in the paper. In the left-hand panel, the LASSO scores are the only ones to outperform the `Frontbench` indicator in predicting newspaper coverage. This is a stern test, as frontbench MPs are significantly more visible in the press than are backbench MPs. Furthermore, even conditional on frontbench status (right-hand panel), the LASSO method is more predictive of media coverage than either of the text-based `MavenRank` or `Burstiness` approaches, or either of the measures of speech quantity. The same is true for back-bench MPs, where the speech quantity measures are somewhat better at predicting media coverage than `MavenRank` or `Burstiness`, but are outperformed by the LASSO measure. Again, the absolute performance of the LASSO method is also reasonably strong: across all MPs and all years, the correlation between an MP's influence score and the newspaper-based ranking of the MP is 0.42.[21]

[21]The equivalent correlations for frontbench MPs and backbench MPs are 0.38 and 33, respectively.

Applications

The validation tasks in the previous section suggest that the influence scores from the LASSO-based method can be usefully employed to describe the relative influence of MPs in parliamentary debate. In this section, I demonstrate the value of these scores by applying them to three questions in legislative politics. These examples are necessarily brief, but they reveal some potential uses of these scores for future research.

Influence and parliamentary tenure

A key question in legislative politics is how the behaviour of MPs changes over the course of their legislative careers. Several studies demonstrate that legislative behaviour varies over the career cycle. For example, MPs in the House of Commons are more likely to defect from their party leadership towards the end of their tenure (Benedetto and Hix, 2007). Similarly, there is a wealth of literature that suggests politicians who are serving their final term in office exert less effort than those serving prior terms (Besley and Case, 1995; Wright, 2007; Ferraz and Finan, 2011; Alt, Bueno de Mesquita and Rose, 2011; Fouirnaies and Hall, 2018). In this context, we can ask a related question: How does the *influence* of an MP vary over their tenure in parliament?

To evaluate the effects of parliamentary tenure on MP debate influence, I estimate a generalised additive model (GAM), where the influence score of each MP in each debate is regressed on a smooth function of the number of days that the MP has served in parliament up until that point (`tenure`). To account for the fact that MPs are also more likely to be promoted to institutionally powerful positions the longer they have served in parliament, I control for whether an MP held a ministerial or party leadership position at the time of a given debate, and whether the MP was the current chair of a parliamentary committee. I also control for whether the MP was a member of the current government party.

Figure 7 presents the fitted values for this regression over the range of the `tenure` variable in the data. As the plot makes clear, agenda-setting influence increases dramatically over the first 10 years of an MP's career, and then reduces slightly thereafter. The rapid

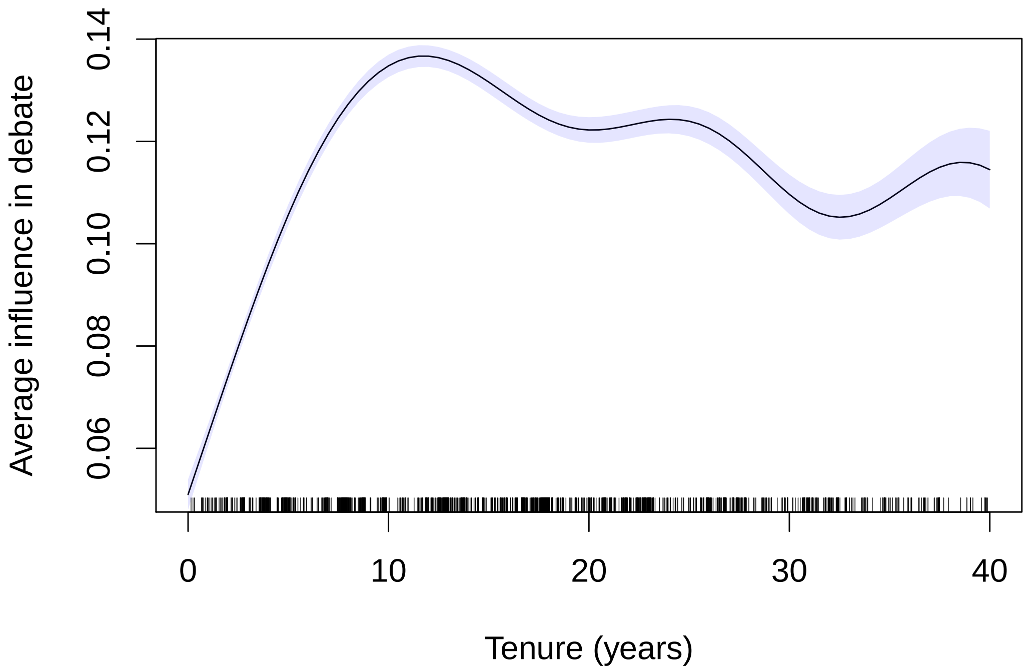

Figure 7: *Debate influence and parliamentary tenure*
The figure shows the average predicted level of debate influence by an MP's tenure in parliament. Predicted values are from a generalised additive model, controlling for whether an MP holds a ministerial, committee, or party leadership position, and for whether the MP is a member of a governing party. Ticks at the bottom of the plot represent the maximum tenure for each MP in the sample.

increase in influence over the beginning of an MP's career is interesting, if not surprising: MPs gain leverage in parliamentary debates as they become more experienced in the Commons. Similarly, although influence does appear to wane in later periods of an MP's tenure, the decline is limited, suggesting that more senior figures retain much of the sway that they hold over their colleagues in their later years.

It is also notable that there is a very different relationship between an MP's tenure and the amount that they speak in debates over time. Figure S1 in the appendix replicates this analysis with the average number of words in debate as the outcome variable. While the figure shows a similar increase in the number of words spoken by MPs over the first ten years of their career, it then depicts a dramatic decline thereafter. This implies that while the speeches of the most experienced MPs are of comparable length as the speeches the least experienced, more experienced MPs exert significantly more influence than their newer colleagues. This demonstrates that the influence measure allows us to capture nuances about political behaviour that would not be possible by looking at speech quantity alone.

Influence and party discipline

In recent work, Slapin et al. (2018) argue that patterns of rebellion in roll-call votes in the Commons can in part be explained by the incentives that MPs face to pander to their constituents. They argue that ideologically extreme MPs vote against their own party in order to telegraph to constituents their opposition to party policy. Moreover, when MPs cast rebellious votes, they are substantially more likely to pair that vote with a speech in the relevant debate in parliament – something that is true particularly for government party MPs. An open question, however, is whether these MPs play an important role in the debates in which they rebel. That is, we can also investigate whether those defecting MPs are more likely to be influential in steering the parliamentary discussion than non-defecting MPs. One potential reason why defecting government MPs may gain more influence in debates is that such defections – and the associated speeches – send a clear signal to the opposition of a key line of attack for government policy that is under discussion. Similarly, opposition parties may want to draw attention to internal divisions within the governing party. Accordingly, we might expect MPs from the governing party who defect in roll-call votes will become focal points in relevant debates.

To test this hypothesis, I collect data on all votes cast in the Commons between 2001 and 2018, and link these votes (known as 'divisions') to the relevant debates.[22] For each MP, on each division, I code whether the MP 'defected' (cast a vote against the majority of their party).[23] Not all debates are followed by recorded votes, and not all MPs who vote in a division participate in the relevant debate, and so the data here is from 2143 debates, and

[22] Roll-call votes are collected from `publicwhip.org.uk`

[23] This measure is likely to be reliable for large parties, but less so for small parties, and so I only include observations from the Conservatives, Labour and the Liberal Democrats here.

includes 42,486 speaker-vote observations. I then estimate linear regressions of the form:

$$\begin{aligned}\text{Influence}_{iv} = \alpha + \beta_1\text{Defect}_{iv} + \beta_2\text{Governing Party}_{iv} + \beta_3\text{Defect * Governing Party}_{iv} \\ +\beta_4\text{Cabinet Minister}_{iv} + \beta_5\text{Shadow Minister}_{iv} + \beta_6\text{Committee Chair}_{iv} \\ +\delta_v + \lambda_i + \epsilon_{iv}\end{aligned} \tag{7}$$

where Influence_{iv} is the debate-level influence score (equation 5) of MP i on debate-vote v, Defect_{iv} is whether an MP defected from their party-majority on a given vote, $\text{Governing Party}_{iv}$ indicates whether an MP was a member of the current governing party, and $\text{Defect}_{iv} * \text{Governing Party}_{iv}$ captures the interaction of interest. I also control for cabinet, shadow cabinet, and committee chair status, and in some specifications I include debate- and individual fixed-effects (δ_v and λ_i, respectively). The expectation is that MPs who defect from the party line will be more influential in the course of relevant debates than non-defecting MPs, and that this effect will be particularly pronounced for government party MPs. Accordingly, the crucial quantity of interest is the sum of β_1 and β_3, which captures the effect of defection on debate influence amongst government party MPs, and which I expect to be positive.

The estimates of these regressions, given in table 1, tell a consistent story: governing-party MPs who vote against the government whip in roll-call votes are substantially more influential in the corresponding parliamentary debates than are MPs who remain loyal to the government. Based on the most conservative model (model 3), the effect of defection on the influence of opposition party MPs is not distinguishable from zero ($\beta_1 = -0.007$, $t = -0.778$) but there is a large and significant effect for governing party MPs. The marginal effect ($\beta_1 + \beta_3$) implies that, for government MPs, defecting from the party line increases an MP's influence score by 0.02, or approximately 10% over the baseline influence level for all MPs in debate. These results suggest that not only do government party rebels speak more often on rebellious votes, but that the speeches they give play a more important role in debate than do the speeches of more loyal party members.[24]

[24] Table S2 in the appendix replicates this analysis, but using the proportion of words in debate as the dependent variable. These effects disappear when considering speech length, again suggesting that the influence measure offers something different from a simple analysis of participation in debate.

	Influence		
	(1)	(2)	(3)
Defection	−0.015	−0.015	−0.007
	(0.009)	(0.009)	(0.009)
Governing Party	−0.011	0.007	0.009
	(0.002)	(0.003)	(0.003)
Defection * Governing Party	0.035	0.041	0.027
	(0.010)	(0.010)	(0.010)
Constant	0.108	0.045	0.021
	(0.007)	(0.025)	(0.041)
Controls	Yes	Yes	Yes
MP fixed-effects	No	Yes	Yes
Debate fixed-effects	No	No	Yes
Observations	42,486	42,486	42,486
R^2	0.131	0.248	0.365

Table 1: *Influence and roll-call defections*
Linear regression estimates for the relationship between debate influence and defections in roll-call votes based on 2143 debates and votes from 2001 to 2018. All models control for whether an MP held a cabinet, shadow cabinet, or committee chair position. Model 2 includes MP fixed-effects, and model 3 includes MP and debate fixed-effects.

Who influences whom?

Our primary quantities of interest thus far have been the MP-level influence scores defined in equation 5. However, an additional advantage to the measurement strategy is that, over and above simply ranking MPs by their relative influence scores, we are able to ask questions about which MPs influence which other MPs. Consider the speaker-by-speaker matrix $\tilde{D}$ defined above, which I construct here by combining the debate-specific D matrices across all debates in two periods: one for the 2010-2015 parliament, and one for the current parliament. A column of this matrix indicates the degree to which the speeches by a given MP were influenced by all other MPs during the time period. We can therefore measure 'who typically influences whom' by examining the ranking of MPs in each of these vectors.

To demonstrate the value of this characterisation of debate, figure 8 depicts the names of the MPs who most strongly influence the speeches of important cabinet ministers in these two parliamentary terms. In the centre of each panel is the name of the relevant cabinet minister, and the names surrounding that minister are sized proportionally to the influence that the relevant MP exerts over the relevant minister. The solid lines indicate the influence exerted on the minister by the MP, and the dashed lines indicate the influence exerted in the opposite direction: from the minister to the MP. The top row of the plot relates to the 2010-2015 coalition government and the bottom row to the current Conservative minority government.

For example, the top-left panel gives the names of the MPs who exerted most influence over the speeches of David Cameron – the Prime Minister in the 2010-2015 parliament. The most influential MP here is Ed Miliband, then Leader of the Opposition, followed by Harriet Harman who was Miliband's deputy at the time. Similarly, George Osborn (then Chancellor) is most influenced by Edward Balls (then Shadow Chancellor) and William Hague (then Foreign Secretary) by Douglas Alexander (then Shadow Foreign Secratery). Similar relationships can be seen in the bottom panels, as the most important influencer of each cabinet minister in each case is the relevant shadow minister.

These figures provide an interesting view of parliamentary politics in the UK as they appear to demonstrate that far from being a hollow 'echo-chamber' in which party politics drives all political discussion, cabinet ministers are in fact primarily responsive to their opposition counterparts. This is encouraging as a key normative foundation of the Westminster system is that the government-opposition dynamics encourage high-quality scrutiny of the executive, which primarily operates through parliamentary debate and questioning (Franklin and Norton, 1993). In general, characterising debates in this way allows for a rich representation of parliamentary debate that could be used to pursue a variety of interesting research questions. For example, future work could examine whether and how these dynamics have changed over time, and the degree to which they vary by policy area.

Discussion and future work

Speechmaking is omnipresent in politics, and both political scientists and computer scientists have made great progress in recent years in tackling problems related to the measure-

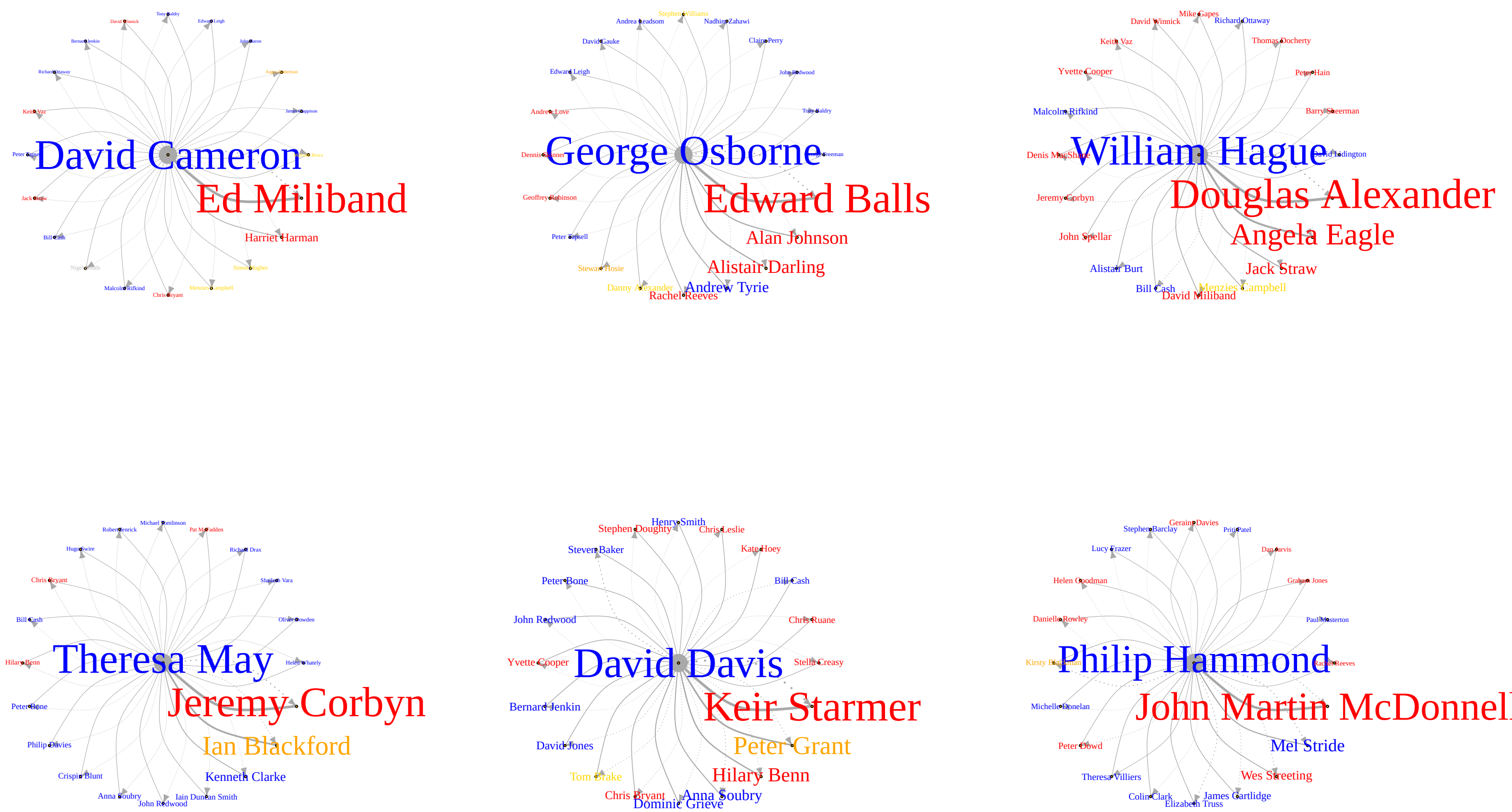

Figure 8: Plots present the MPs (on the outside of the star) whose language is most commonly adopted by a given cabinet minister (on the inside of the star). MP names are sized proportionally to their 'influence' on the relevant cabinet minister, and coloured by party. Solid lines indicate the influence running from the MP to the minister's speeches, and dashed lines indicate influence running in the opposite direction. The top row relates to the 2010-2015 coalition government, and the bottom row relates to the current Conservative minority government.

ment of ideological position-taking and topical-attention. This paper outlines an approach for measuring an alternative quantity of interest – agenda-setting influence in debate – which could be applied to many different legislative settings. The validation tasks suggest that this approach captures reasonable patterns of parliamentary influence, and the applications suggest that such a measure could be profitably applied to important questions in legislative politics.

There are two main methodological issues that deserve attention in future work. First, there is uncertainty inherent in the LASSO measure at the speech-level, and also uncertainty in the aggregation of speech-level scores to the individual MP-level. While both sources of uncertainty need to be incorporated into the current strategy, the former is likely to be small, for the same reasons that uncertainty estimates from typical word-scaling methods is also small (Lowe and Benoit, 2011). The latter is more concerning, and one potential approach would be to treat the speech-level LASSO scores as data to inform a more principled hierarchical model at the speaker-level, as in Lauderdale and Herzog (2016). Second, the framework I employ requires selecting an appropriate penalty parameter, λ, for the speech-level regressions. Selecting separate penalties for each regression implies that it will be 'harder' to influence some speeches than others within the same debate. A more suitable alternative would be a joint model of all speeches in a debate with a common λ across speeches, something that is certainly possible but more computationally burdensome. I intend to address both of these issues in future iterations of this work.

References

Alt, James, Ethan Bueno de Mesquita and Shanna Rose. 2011. "Disentangling accountability and competence in elections: evidence from US term limits." *The Journal of Politics* 73(1):171–186.

Ban, Pamela, Alexander Fouirnaies, Andrew B Hall and James M Snyder. 2018. "How Newspapers Reveal Political Power." *Political Science Research and Methods* pp. 1–18.

Benedetto, Giacomo and Simon Hix. 2007. "The rejected, the ejected, and the dejected: Explaining government rebels in the 2001-2005 British House of Commons." *Comparative Political Studies* 40(7):755–781.

Bergersen, Linn Cecilie, Ingrid K Glad and Heidi Lyng. 2011. "Weighted lasso with data integration." *Statistical applications in genetics and molecular biology* 10(1).

Besley, Timothy and Anne Case. 1995. "Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits." *The Quarterly Journal of Economics* 110(3):769–798.

Eggers, Andrew C and Arthur Spirling. 2016. "The shadow cabinet in Westminster systems: modeling opposition agenda setting in the House of Commons, 1832–1915." *British Journal of Political Science* pp. 1–25.

Erdman, David V. 1960. "Coleridge in Lilliput: The quality of parliamentary reporting in 1800." *Communications Monographs* 27(1):33–62.

Erkan, Günes and Dragomir R Radev. 2004. "LexRank: Graph-based lexical centrality as salience in text summarization." *Journal of Artificial Intelligence Research* pp. 457–479.

Fader, Anthony, Dragomir R Radev, Michael H Crespin, Burt L Monroe, Kevin M Quinn and Michael Colaresi. 2007. MavenRank: Identifying influential members of the US senate using lexical centrality. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL).*

Ferraz, Claudio and Frederico Finan. 2011. "Electoral accountability and corruption: Evidence from the audits of local governments." *American Economic Review* 101(4):1274–1311.

Finlayson, Alan and James Martin. 2008. "'It ain't what you say...': British political studies and the analysis of speech and rhetoric." *British Politics* 3(4):445–464.

Fouirnaies, Alexander and Andrew B Hall. 2018. "How Do Electoral Incentives Affect Legislator Behavior?" *Working Paper* .

Franklin, Mark N and Philip Norton. 1993. *Parliamentary questions: for the study of Parliament group.* Oxford University Press, USA.

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.

Herzog, Alexander and Kenneth Benoit. 2015. “The most unkindest cuts: speaker selection and expressed government dissent during economic crisis.” *The Journal of Politics* 77(4):1157–1175.

Hopkins, Daniel J and Gary King. 2010. “A method of automated nonparametric content analysis for social science.” *American Journal of Political Science* 54(1):229–247.

House of Commons. 2017. “European Union (Notification of Withdrawal) Bill, Debate.”. **URL:** *https://www.theyworkforyou.com/debates/?id=2017-01-31c.817.2*

Kleinberg, Jon. 2003. “Bursty and hierarchical structure in streams.” *Data Mining and Knowledge Discovery* 7(4):373–397.

Kleinberg, Jon M. 1999. “Authoritative sources in a hyperlinked environment.” *Journal of the ACM (JACM)* 46(5):604–632.

Lauderdale, Benjamin E and Alexander Herzog. 2016. “Measuring political positions from legislative speech.” *Political Analysis* 24(3):374–394.

Laver, Michael and Kenneth Benoit. 2002. “Locating TDs in policy spaces: the computational text analysis of Dáil speeches.” *Irish Political Studies* 17(1):59–73.

Laver, Michael, Kenneth Benoit and John Garry. 2003. “Extracting policy positions from political texts using words as data.” *American Political Science Review* 97(2):311–331.

Lowe, Will and Kenneth Benoit. 2011. Estimating uncertainty in quantitative text analysis. In *Annual Conference of the Midwest Political Science Association*. Vol. 31.

Mihalcea, Rada. 2004. Graph-based ranking algorithms for sentence extraction, applied to text summarization. In *Proceedings of the ACL 2004 on Interactive poster and demonstration sessions*. Association for Computational Linguistics p. 20.

Page, Lawrence, Sergey Brin, Rajeev Motwani and Terry Winograd. 1999. “The PageRank citation ranking: bringing order to the Web.”.

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.

Riker, William H. 1986. *The art of political manipulation*. Vol. 587 Yale University Press.

Riker, William H. 1996. *The strategy of rhetoric: Campaigning for the American Constitution*. Yale University Press.

Roberts, Margaret E, Brandon M Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson and David G Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science* 58(4):1064–1082.

Shepsle, Kennet A. 2003. "Losers in politics (and how they sometimes become winners): William Riker's heresthetic." *Perspectives on politics* 1(2):307–315.

Slapin, Jonathan B, Justin H Kirkland, Joseph A Lazzaro, Patrick A Leslie and TOM O'GRADY. 2018. "Ideology, Grandstanding, and Strategic Party Disloyalty in the British Parliament." *American Political Science Review* 112(1):15–30.

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.

Tibshirani, Robert. 1996. "Regression shrinkage and selection via the lasso." *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 267–288.

Wright, Gerald C. 2007. "Do term limits affect legislative roll call voting? Representation, polarization, and participation." *State Politics & Policy Quarterly* 7(3):256–280.

Zou, Hui. 2006. "The adaptive lasso and its oracle properties." *Journal of the American statistical association* 101(476):1418–1429.

Appendix

	Influence
Cabinet Minister	0.619
	(0.002)
Shadow Cabinet Minister	0.195
	(0.002)
Constant	0.183
	(0.0004)
Observations	637,713
R^2	0.144

Table S1: *Debate influence of cabinet, shadow cabinet, and backbench MPs.*
The table presents results from a linear regression of the LASSO-based influence scores on binary indicators for whether an MP is a cabinet minister or a shadow cabinet minister during the relevant debate. The constant represents MPs who do not hold either cabinet or shadow cabinet positions. Shadow cabinet ministers are about twice as influential as backbenchers, and cabinet ministers are about 3 times as influential.

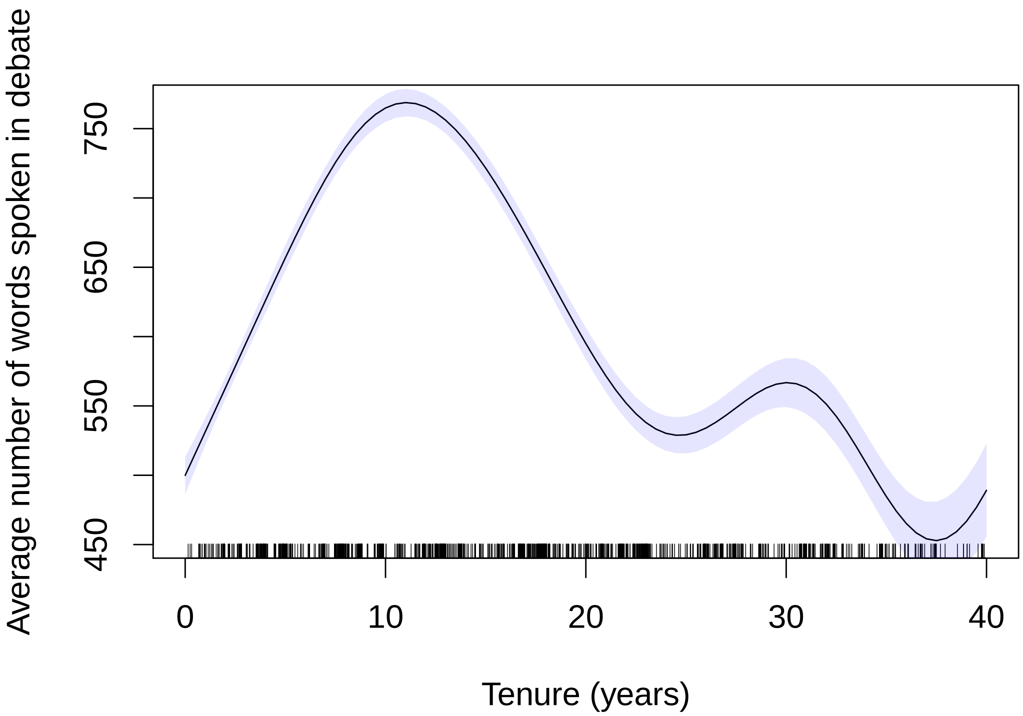

Figure S1: *Average number of words spoken by MPs in debate by parliamentary tenure* The figure show the average number of words spoken in debate by an MP's tenure in parliament. While MPs speak more and become more influential over the first ten years of their career, the number of words declines dramatically later in the course of a career while the influence of the MP does not (figure 7). Predicted values are from a generalised additive model, controlling for whether an MP holds a ministerial, committee, or party leadership position, and for whether the MP is a member of a governing party.

	Proportion of words		
	(1)	(2)	(3)
Defection	−0.001	0.001	0.008
	(0.003)	(0.003)	(0.003)
Governing Party	−0.002	0.004	0.005
	(0.001)	(0.001)	(0.001)
Defection * Governing Party	0.003	0.007	−0.002
	(0.003)	(0.003)	(0.003)
Constant	0.041	0.013	0.007
	(0.002)	(0.008)	(0.012)
Controls	Yes	Yes	Yes
MP fixed-effects	No	Yes	Yes
Debate fixed-effects	No	No	Yes
Observations	42,487	42,487	42,487
R^2	0.069	0.248	0.455

Table S2: *Proportion of words in debate and roll-call defections*
Linear regression estimates for the relationship between the proportion of words spoken and defections in roll-call votes based on 2143 debates and votes from 2001 to 2018. All models control for whether an MP held a cabinet, shadow cabinet, or committee chair position. Model 2 includes MP fixed-effects, and model 3 includes MP and debate fixed-effects.