

Boring in a New Way

Estimation and Inference for Political Style at Westminster, 1935–2018*

Arthur Spirling[†] Leslie Huang[‡] Patrick O. Perry[§]

Abstract

We consider the merits of claims that Members of Parliament (MPs) in the UK have become more ‘boring’ over time—that is, less distinctive from one another in terms of their speech and style. We review theory and previous findings in the area, and note their ambiguity in predictions on this matter. We then provide an efficient new measurement model of distinctiveness that extends traditional efforts to statistically characterize the ‘style’ of authors, and apply it to a corpus of *Hansard* speeches from 1935 to 2018. In the aggregate, we find no evidence for the claim of increased boringness. But this hides intriguing covariate effects: at the MP level, panel regression results demonstrate that on average, more senior backbenchers tend to be less interesting in speech terms. We also show, however, that this pattern is changing: in recent times, it is less experienced MPs who speak most distinctively.

Software package, `stylest`, for R available here:
<https://github.com/leslie-huang/stylest>

Word Count: 3910 (excluding Supporting Information)

*Prepared for the Polmeth annual meeting in Provo, UT (2018). First version: March 10, 2018. This version: July 12, 2018. We thank Jack Blumenau for assistance with data. Jennifer Holmes provided helpful feedback on an earlier draft.

[†]Associate Professor of Politics and Data Science, New York University (arthur.spirling@nyu.edu)

[‡]Graduate Student, Politics, New York University (lh1036@nyu.edu)

[§]Assistant Professor of Information, Operations, and Management Sciences, New York University Stern School of Business (pperry@stern.nyu.edu)

1 Introduction

Floccinaucinihilipilification is not a word regularly encountered in the House of Commons; between 1803 and the present day, it was uttered just twice. The first time was in 1947, by the Member of Parliament (MP) for Caithness and Sutherland; subsequently, it was used in 2012 by Jacob Rees-Moog, MP for North Devon. So apparently interesting was the second invoking of the term that it earned Rees-Mogg an interview on the BBC’s current affairs program, *The Daily Politics*. That such an event was deemed newsworthy speaks to a broader, popular concern with modern Westminster politics. In particular, that contemporary MPs are too ‘on message’, sound the same, and are thus ‘boring’—especially compared to the ‘big beasts’ of yesteryear. Thus Cowley and Stuart (2004, 211) (citing Cowley 2002) note that Labour backbenchers of the Blair period were “variously described as sheep, poodles, clones, robots or—most bizarrely of all—daleks.”

The academic assessment of the relative historical distinctiveness of contemporary MPs has been nuanced but its findings point in opposite directions. On the one hand, some scholars note that ‘rebellious’ behavior (on roll calls) is generally increasing—moving the UK away from the traditional Westminster archetype of pliant backbenchers. Furthermore, we observe that MPs increasingly seem to play to constituency preferences, rather than the party line, on at least some votes (Vivyan and Wagner, 2012; Hanretty, Lauderdale and Vivyan, 2017). Relatedly, MPs try to cultivate a personal vote separate to their party (Jackson and Lilleker, 2011), although the size of the effect is debatable (Eggers and Spirling, 2017). But we *also* observe that MPs are more ‘professional’ (Rush, 2001) and career oriented (King, 1981) than in the past. To the extent that being distinctive in one’s parliamentary behavior is likely to be costly to ministerial ambitions (Kam, 2009), and more MPs have such ambitions, we should expect to see less of it, on average, over time.

British legislators have unusual latitude to speak freely in debates (Proksch and Slapin, 2012). Given the above though, our predictions for how they use those opportunities—and how that has changed over time—are ambiguous. That is, there are good reasons to imagine they have become *more* or *less* distinctive, in accordance with the relative weight one places on the forces noted

above. What complicates the picture considerably is that there are likely *within career* effects but these are also ambiguous. On the one hand, we see that more senior MPs are more likely to engage in distinctive or rebellious behavior (Benedetto and Hix, 2007); on the other, we know that the longer they serve, the more members become socialized to its norms and expectations (Rush and Giddings, 2011).

This short paper speaks to these ambiguities—historical and within career—as they pertain to speech behavior. We do this with a dataset of three million speeches (1935–2018), and a new way of estimating how ‘distinctive’ backbench MPs are over-time, in aggregate and individually. Intuitively, MPs are ‘interesting’ if they can be easily identified from their speeches relative to other members; they are ‘boring’ if their words do not mark them out. Ultimately, we find no evidence that MPs are becoming less interesting as a whole. To the extent that service matters, it typically has a negative effect: that is, senior MPs tend to be more ‘boring’, junior ones are more ‘interesting’. Yet this relationship is changing: in particular, since the 1990s, more experienced members are emerging as the most distinctive legislators in the Commons.

2 Competing Pressures in Westminster Systems

The ‘textbook’ account of Westminster systems notes that they have a strong executive, to which backbenchers on both government and opposition sides are supplicant (Kam, 2009; Lijphart, 2012). Exactly what drives MP decisions on roll calls, speeches and other behaviors is debatable, but there is a broad consensus that at least two factors matter. First, is ministerial ambition; since backbenchers want to become frontbenchers (see Rush and Childs, 2004, for discussion of survey evidence), and the party leadership controls access to this resource, MPs tend to toe the party line (Rush, 2001). Evidence of this mechanism comes in two forms. For one, it is clear that ‘rebellious’—i.e. not following the whips’ instructions—is not beneficial to one’s career (Cowley, 2002). Relatedly, scholars have found that ‘rebels’ are those passed over for promotion, or unlikely to ever be ministers for other reasons (Benedetto and Hix, 2007). Of course, career incentives are

not the only ones at play in ensuring compliance with leadership commands. Thus, Crowe (1986) and Rush and Giddings (2011) present evidence of ‘socialization’ by which incoming MPs learn the ropes in terms of both Commons procedure and with respect to more general expectations of behavior. Beyond the carrot of cabinet office, there is some evidence that communication of preferences is a two-way street: party leaders listen privately to the policy concerns that their troops have (Norton, 1999; Cowley and Childs, 2003) and are otherwise constrained by backbench opinions in terms of central personnel choices (Kam et al., 2010).

Unlike in the US (e.g. Ansolabehere, Snyder and Stewart, 2001; Canes-Wrone, Brady and Cogan, 2002), the forces of political life in Westminster systems are thought to mitigate against notions of ‘dyadic representation’ (Weissberg, 1978)—that is, the notion that MPs might seek to legislate in a way that reflects political preferences of their constituents. In the United Kingdom, where politics become increasingly ‘national’ after the Great Reform Act (Cox, 1987) scholars have historically emphasized the importance of decidedly non-local factors in deciding voter minds when selecting a local representative. These include (national) party identification in conjunction with class (e.g. Butler and Stokes, 1969; Heath et al., 1991) and perceptions of leader competence (e.g. Green and Hobolt, 2008). In line with this skepticism about the effects of local connections, while MPs have long done service in their constituencies (Cain, Ferejohn and Fiorina, 1987; Norton and Wood, 1993; Searing, 1994), evidence that it matters for re-election performance has not been readily forthcoming (Gaines, 1998; Eggers and Spirling, 2017).

With the above in mind, it is surprising that there is any reason at all for MPs to seek to be different to one another. But this ignores recent trends in voter and politician behavior. Scholars note that citizens in advanced democracies are generally less attached to major parties than they used to be, and Britain is no exception (Clarke and Stewart, 1998). Perhaps because of this (or perhaps despite it), MPs became more rebellious from the 1960s onwards (Norton, 1975, 1980). Modern MPs, especially Labour ones, disobey the whip more often, and on more important matters, than is commonly realized (Cowley, 2002). And they rebel strategically, at least as regards their status as government or opposition at a given time (Slapin et al., 2017). Furthermore, there

is evidence that MPs do, in fact, attempt to respond to constituent opinion on important issues (Hanretty, Lauderdale and Vivyan, 2017) and that voting particular ways at Westminster may have a positive, if weak, effect on their popularity at home (Pattie, Fieldhouse and Johnston, 1994). It is also apparently the case that rebellion itself is rewarded (Vivyan and Wagner, 2012), and those who defy their leadership are more recognizable in surveys than those who don't (Kam, 2009). In keeping with these self-promotional efforts, MPs seek to actively manage their personal 'brand' via social media (Jackson and Lilleker, 2011).

Matching these behavioral changes, and possibly a cause of the same, are changes to the sociological make-up of the House of Commons. Since the mid-1850s, parliament has become more 'professional' (Rush, 2001). Members are increasingly "career" politicians (in the sense of King, 1981) and view their positions as full time jobs. Since the 1960s, MPs are increasingly drawn from the university educated middle-classes and whatever distinctions traditionally existed between Labour and Conservative backgrounds are now much weaker (Norris and Lovenduski, 1995; Heath, 2015). At least part of this trend is the increasing tendency to draw MPs from the ranks of 'special advisors' and other professional party workers (Shaw, 2001).

To summarize, MPs are under competing pressures. On the one hand, they need to behave to be accepted by their peers and promoted. On the other, this may mean they miss out on opportunities to win over local constituents. And if that happens, they may cease to be MPs at all. In that sense, professionalisation cuts both ways: optimistically, it implies MPs have better 'streetwise' knowledge of the political environment in which they operate, perhaps facilitating more independent thought and action. More pessimistically, it may imply less diversity in terms of ideology and style, twinned with more dependence on a party apparatus through on which they have always depended for advancement. To know how these forces play out, we need a valid measure of speech 'style' for MPs.

3 Measuring Style

To measure ‘interestingness’, we rely on the basic principle of ‘stylometry’—that authors have idiosyncratic markers in the documents they produce. The typical goal in that literature is detecting the most likely author of a given text of uncertain origin, by way of the candidate authors’ known preferences in word use. In terms of applications to politics, the work of Mosteller and Wallace (1963; 1964) is well-known. Their challenge was to identify the most likely author—Madison or Hamilton—of twelve ‘disputed’ *Federalist Papers* papers. They found overwhelming evidence for Madison. Using similar methods, Airoidi et al. (2006) and Airoidi, Fienberg and Skinner (2007) investigate the empirical claim that Ronald Reagan may not have authored a series of radio addresses he gave in the 1970s.

Our strategy here is broadly similar in that we care about the distinctiveness of one speaker/author (MP) relative to another. But in detail it differs markedly. In particular, our interest is obviously not in identifying origins of texts: we have labels for all our data (in terms of the MP giving the speech). Instead, authors/speakers are our focus, and the extent to which any linguistic features exist that mark them apart from one another.

3.1 Our Conception of Style

To see the intuition, note that in the case of Mosteller and Wallace what is of interest is the comparison of two probabilities for an unlabeled text. The model implies some probability the paper was written by Madison given the counts of the words (\mathbf{w}) it contains, $\text{Pr}(\text{Madison}|\mathbf{w})$. That is then compared with the probability the essay is, conditioned on its contents, from Hamilton, $\text{Pr}(\text{Hamilton}|\mathbf{w})$. Subject to some mathematical housekeeping, the larger of these probabilities then yields the author prediction.

But in our application, we care about how different Madison is from Hamilton *in general* for all the (labeled) data. To fix ideas, suppose we have an essay we know Madison wrote. Given the words it contains, our model can still provide us with an *estimate* of the probability Madison

penned it: while we know the true probability to be one, the model will give us an in-sample prediction (which is not one). We can do the same thing for Hamilton for that given Madison speech: obtain a model probability that Hamilton wrote it (though we know the true probability to be zero) by plugging in Hamilton’s word-use tendencies from the essays which we know *he* wrote. Suppose that we do this for every one of Madison’s known texts. If, in general, the predicted probability that Madison wrote them is much higher than the predicted probability Hamilton wrote them, we have *prima facie* evidence that the authors differ in style terms. That is, the model is finding features that enables it to distinguish one author from the other for the Madison documents. But if the Madison essays have predicted probabilities that are always similar for Madison and Hamilton, then the opposite lesson applies: the model simply cannot distinguish between the two men as most plausible authors for the selection of texts. To summarize, and in keeping with our simple statement of the problem, we are interested in the ‘distinctiveness of Madison’ \mathbb{D}_M . We now generalize this to an arbitrary number of essays (speeches) and authors (MPs).

3.2 Our Formulation of Style

In Supporting Information A we give a mathematical derivation of our style model; for now, we give a non-technical intuition that builds on the logic laid out above.

1. For a given session of parliament, each member’s speeches are summarized as a vector with word counts for every one of the tokens spoken by any MP in that session. These counts are converted into *probabilities* by simply dividing them by the length of the vectors (which are the size of the vocabulary used in a given parliamentary session). For convenience, those probabilities are then logged, to produce a vector η_s where s denotes a specific MP.
2. To compare MP s with MP t as a pair for a given word in a random speech i that one of them gave, we subtract the value of η_s from η_t for that word in the speech (multiplied by the number of times that word is used in the speech). If this number is ‘large’ then t is *distinctive* relative to s in her use of this word. We can take the average of this quantity over

all the words in the vocabulary to get a more general measure of the distinctiveness of t .

3. For the random speech i , we generalize this measure to all members—that is, we produce a number that compares t to everyone else pairwise—by taking the average of the quantity in (2) *over all speakers*. The second generalization is *over all speeches* given by t which essentially requires summing the (average) distinctiveness per word and dividing out by the number of speeches t gave.

This quantity has a Bayesian interpretation, but in essence represents the evidence that a given speaker (relative to all others) produced the words she spoke—averaged over all words, all speeches and all possible pairwise comparisons to other members. It comes with several caveats, discussed in Supporting Information B. We take this model to data.

4 Data

Our data supplements Rheault et al. (2016) with data to the present day, and consists of approximately three million speeches in the House of Commons (1935–2018). There is meta-data pertaining to party membership and ministerial position. We calculate the *experience* of a given MP as being the number of sessions they have served in parliament since their first speech.¹ We also introduce a variable that records whether an MP has ever been *demoted* from ministerial office.

For the purposes of this paper, we are only interested in MPs that actually speak (at least ten times, speeches of at least three words) in a given session and in particular we are focussed on government *backbenchers*. That is, our analysis of style pertains to MPs that are in the governing party (the party of the Prime Minister), but who do not hold ministerial office. This allows us to compare like-with-like over time in terms of the incentives MPs have, even though power shifts across parties.

¹We augment the data with information on when—what session—the MPs speaking in the first period of the data entered parliament.

In Supporting Information C we give more details regarding the averages and ranges of (average) values of our data by session. A take-away is that the (average) means and medians are generally very close, implying little skew in these variables.

5 Results: Validation, Aggregates and Effects of Service

Our measure of distinctiveness borrows from a long tradition in social science in statistics—but is it any good? To assess this, we produced a top-20 ‘most interesting’ and bottom-20 ‘most boring’ list for the parliamentary sessions either side of Blair’s election landslide in 1997. This has the advantage of being a period in which (a) control of the Commons switched (from Conservative to Labour), meaning we have variation in the party of the backbenchers in question and (b) we have several academic accounts which help ground our understanding of MP behavior (Cowley, 2002; Spirling and Quinn, 2010; Kellermann, 2012).

We use a *convergent validity* approach: we compare our measure to another (computed independently) and show they are related in the expected way. In particular, we look at how often MPs were mentioned in the *Times* politics news section for the relevant sessions. Fuller details are in Supporting Information D, and obviously there are caveats to our comparisons. Nonetheless, we find that newspaper reports agree with our assessment: our more interesting MPs are discussed more, and our more boring ones are discussed less (and this difference is statistically significant).

We also investigated the validity of the most influential tokens (from our measure), and the those results are in Supporting Information E. Broadly, we are satisfied with what we found.

5.1 Time Series: Distinctiveness is Not Decreasing

For every session in the aggregate, it is trivial to produce an average (median or mean) distinctiveness score, along with a variance. In Figure 1 we do exactly that. Two observations are immediate. First, as demonstrated in the top panel, the median distinctiveness of MPs doesn’t appear to decrease over time. At the very least, the points and the solid [red] lowess of the same seem fairly

stable and certainly not moving downwards. This is also true of the mean (shown via the broken [blue] lowess curve). Indeed, if anything, these averages appear to be *increasing*. For completeness, we also plot the tails of the distribution of scores. These are the broken lines at the top and bottom of the upper plot. While they show some variance, there is no obvious trend in the extremes.

Nor does it seem to be the case that MPs are, on average, becoming more or less spread out in terms of their distinctiveness. This can be seen in the lower plot which reports the variance of the scores over time. The spread decreases until around 1975 (note [red] lowess), before rising again to reach a level approaching the beginning of the data. Plots can be misleading, but formal statistical examinations suggest initial impressions are correct. Standard (monotonic) tests—Cox-Stuart (Cox and Stuart, 1955) and Mann-Kendall (Mann, 1945; Kendall, 1975) (both $p > 0.05$)—reveal no presence of a ‘trend’ in the medians. To summarize then, the average MP has not become more or less distinctive in terms of style over the past 80 years.

Possibly, assessing ‘average’ MPs is unfair in that the ‘real’ action of decline is at the top end of the distribution: that is, it is the maverick outliers who have disappeared. But the evidence for this is equivocal: while a Mann-Kendall test suggests the 90th, 95th and 97.5th percentiles are trending down, Cox-Stuart tests cannot reject the null. Interestingly, by both tests ($p < 0.01$), there is a trend in the lowest percentile we checked—the 2.5th—but it is upwards, not downwards.

5.2 Model Inference: Serving Longer Makes You More Boring

Of course, aggregates can oversimplify: it’s possible that over time the relationship between being distinctive and other MP features has changed in such a way as to disguise something more profound. To look at this possibility, we begin with panel regressions. Here, the cross-section time series is MP-by-session, and is unbalanced since members only serve for a limited number of years. As noted above, we have two covariates (in addition to the fixed effects) for predicting distinctiveness: experience (in session terms) and whether the MP has ever been demoted from government front-bench responsibilities.

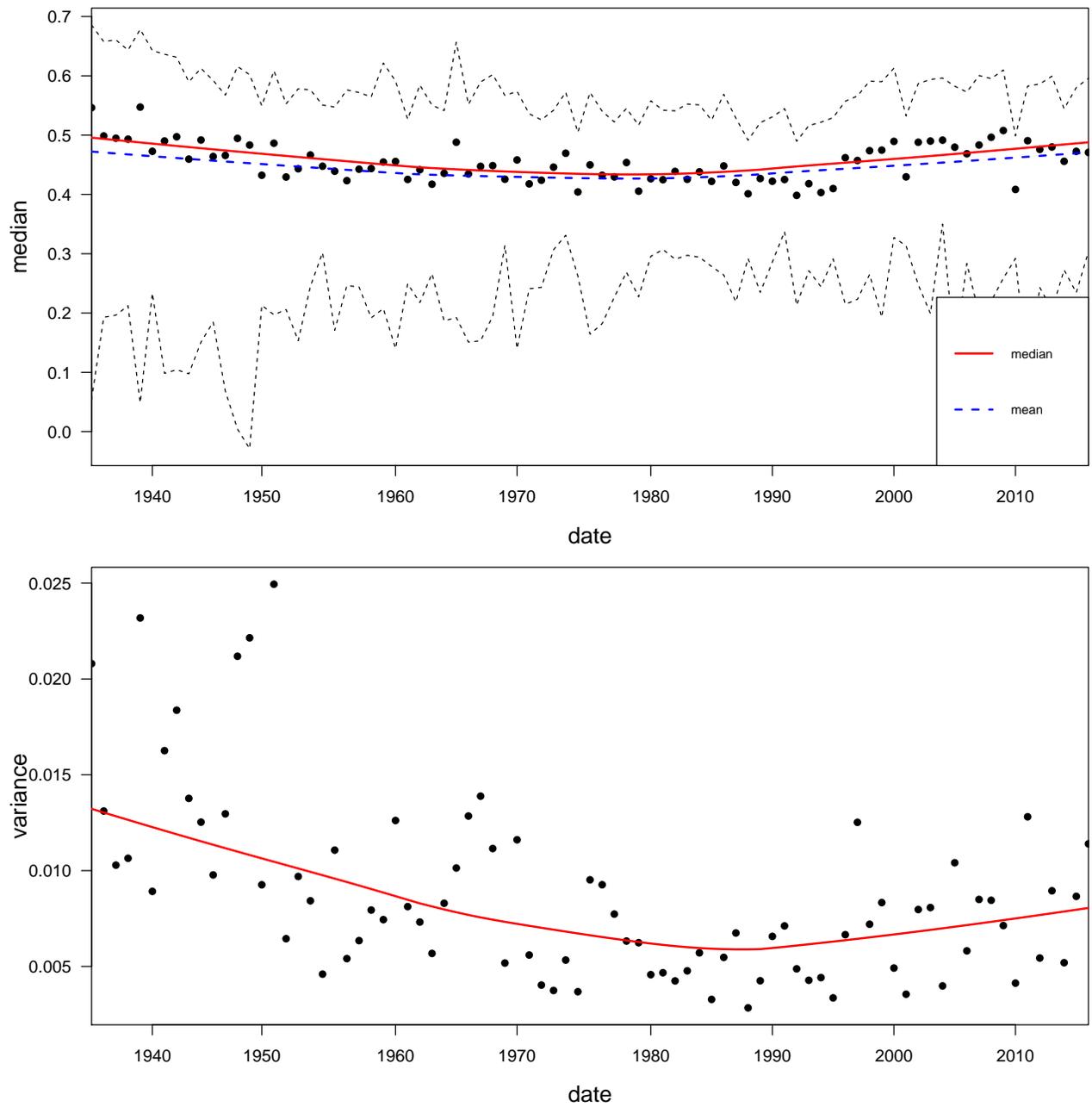


Figure 1: Time series of distinctiveness. In the top panel, the points represent the median, while the solid [red] line represents the lowess of the same. The [blue] broken line is the mean. The [black] broken lines at the top at bottom of the plot are the (empirical) 2.5th and 97.5th percentiles of the data. The bottom panel is the variance over time, plotted with a lowess.

In Table 1 we report the results from three specifications. In Column (1) the relevant regression is *pooled*: that is, we treat the entire sample of MPs as a cross-section and provide OLS estimates for the coefficients on the same. In Column (2) we add MP-level (i.e. unit-level) fixed effects. In Column (3), we use MP-level and session-level (that is, time) fixed effects. The results of an *F*-test and a Baltagi and Li (1990) Lagrange-Multiplier test suggest that time fixed effects are indeed warranted in this case.

Table 1: Effect of experience and demotion on distinctiveness of MPs. The first model pools all observations. The second adopts a panel structure with MP-fixed effects. The third uses MP-fixed and session-fixed effects.

<i>Dependent variable: distinctiveness</i>			
	(1)	(2)	(3)
experience	−0.0003*** (0.0001)	−0.001*** (0.0001)	−0.002*** (0.0002)
demoted	0.004 (0.003)	0.006 (0.004)	0.003 (0.004)
constant	0.443*** (0.001)		
MP-fixed effects	✗	✓	✓
Session-fixed effects	✗	✗	✓
Observations	16,022	16,022	16,022
R ²	0.001	0.006	0.069
F Statistic	4.995*** (df = 2; 16019)	39.602*** (df = 2; 13396)	11.889*** (df = 83; 13315)

Note: *** $p < 0.01$

Regardless of the specification, there is a negative effect of experience on distinctiveness ($p < 0.01$).² That is, as MPs serve longer terms, they become less and less interesting (relative to others) in terms of their speech. Being demoted at some point does not seem to change this dynamic. Substantively (for the best fitting model), a one standard deviation increase in experience (around

²As a robustness test, we also fit the panel regression using heteroscedastic ‘robust’ standard errors (in the sense of White, 1980). This makes no difference to our conclusions. Nor does, in addition, correcting for potential autocorrelation (in the sense of Arellano, 1987).

9 sessions) decreases distinctiveness by around 0.02 (around one fifth of that variable's standard deviation). So these effects are not huge.

5.3 Model Inference: The Effect of Seniority is Changing

Figure 2 reports our final findings. Now, we divide the data into periods of uninterrupted party rule by one party. For each period, we record the coefficient on the experience of government party backbenchers in terms of its effects on distinctiveness. This point estimate is plotted via a letter symbolizing (L)abour or (C)onservative control at the time, and each estimate is plotted at the beginning of the period in question. Thus, for example, the Conservative control of the Premiership that began in 1979 and ended in 1997 is marked by a coefficient in 1979. For each regression, we report the 95% confidence interval on the figure. We see that for the vast majority of our data, the effect of seniority is negative: the coefficient is less than zero, implying that more senior members are less distinctive. This switches during the Blair government, with a positive coefficient for that period; it continues into the coalition government and beyond, though the coefficient is not distinguishable from zero.

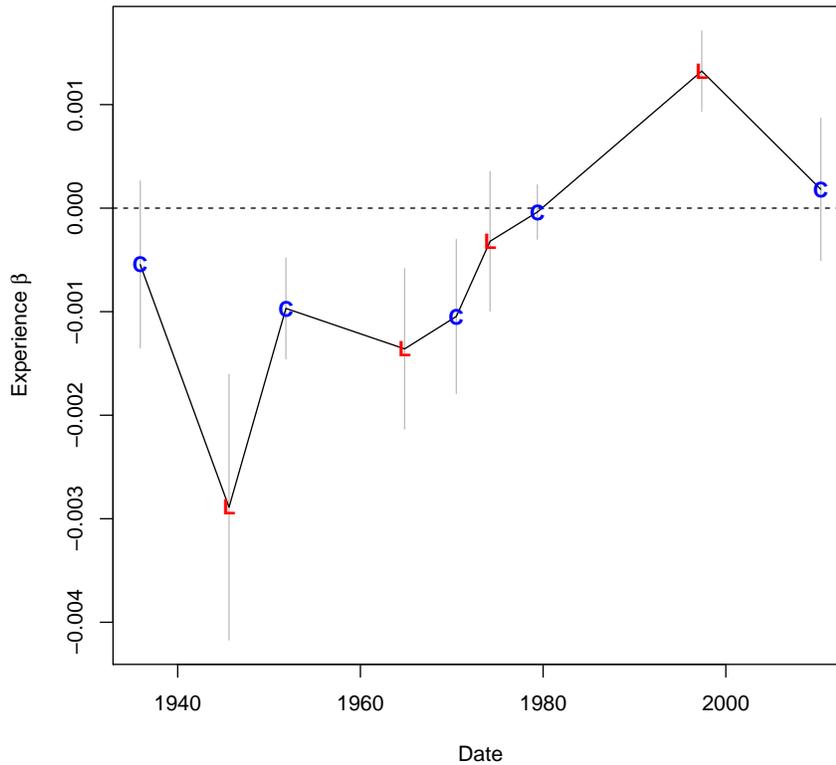


Figure 2: Effect of longer service on distinctiveness, over time. Data is broken up into periods of uninterrupted one party rule, with one coefficient (and 95% confidence interval) per period.

6 Discussion

Exactly what behavior is desirable in MPs has been debated at least since the time of Edmund Burke’s ‘Speech to the Electors of Bristol’ in 1774. Modern voters are not consistently clear about their preferences either: on the one hand, they say they like independent-minded MPs willing to express their positions (Vivyan and Wagner, 2015); on the other, they generally make their voting decisions based on other factors, such as party leader reputations for competence (Green and Jennings, 2012). Given this ambiguity, it is unclear whether MP distinctiveness (from others) is likely to be electorally helpful or not. Regardless, a first step in the process of investigation

is measuring ‘style’, validating that measure, and using it to draw aggregate and individual-level conclusions about how MPs have or have not changed over time.

Here we found there is very little evidence that MPs are becoming less interesting over time. Perhaps there are fewer ‘big beasts’, but most MPs are as different or as similar to their colleagues as they ever were: since 1935, the diversity of ways that MPs express their views is essentially constant. More interestingly perhaps, the effect of seniority is changing. We showed that for most of the 20th Century, longer-serving backbenchers tended to be less distinctive. But, at least since the Blair victory in 1997, the pattern is different: perhaps due to the intake of young MPs more solely focussed on career promotion, it is older MPs who emerge as more interesting speakers.

This latter finding is in line with work on the determinants of rebellion in recent times, from the likes of Cowley and Childs (2003) and Benedetto and Hix (2007). Therein weathered MPs beyond the risk-set for promotion, tend to be less compliant. Perhaps we are seeing the side-lining of socialization as the traditionally dominant force in MP careers (Eggers and Spirling, 2016). In that case, it would be interesting to know more about *what* distinctive MPs choose to talk about: is it simply a matter of linguistic choices, or something more related to topics of debate? And, presumably, we would like to know more about *who* is speaking differently, in terms of their (re)election prospects, their career histories, their ideological positions and so on. We leave such efforts for future work.

References

- Airoldi, Edoardo M, Annelise G Anderson, Stephen E Fienberg and Kiron K Skinner. 2006. “Who wrote Ronald Reagan’s radio addresses?” *Bayesian Analysis* 1(2):289–319.
- Airoldi, Edoardo M., Stephen E. Fienberg and Kiron K. Skinner. 2007. “Whose Ideas? Whose Words? Authorship of Ronald Reagan’s Radio Addresses.” *PS: Political Science and Politics* 40(3):501–506.
- Ansolahehere, Stephen, James Snyder and Charles Stewart. 2001. “Candidate positioning in US House elections.” *American Journal of Political Science* pp. 136–159.
- Arellano, Manuel. 1987. “PRACTITIONERS CORNER: Computing Robust Standard Errors for Within-groups Estimators.” *Oxford bulletin of Economics and Statistics* 49(4):431–434.

- Baltagi, Badi H and Qi Li. 1990. "A Lagrange multiplier test for the error components model with incomplete panels." *Econometric Reviews* 9(1):103–107.
- 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.
- Butler, David and Donald Stokes. 1969. *Political Change in Britain: Forces Shaping Electoral Choice*. London: Macmillan.
- Cain, Bruce, John Ferejohn and Morris Fiorina. 1987. *The personal vote: Constituency service and electoral independence*. Harvard University Press.
- Canes-Wrone, Brandice, David W Brady and John F Cogan. 2002. "Out of step, out of office: Electoral accountability and House members' voting." *American Political Science Review* 96(1):127–140.
- Clarke, Harold D. and Marianne C. Stewart. 1998. "The Decline of Parties in the Minds of Citizens." *Annual Review of Political Science* 1(1):357–378.
- Cowley, Philip. 2002. "Revolts and rebellions." *London, Politics* .
- Cowley, Philip and Mark Stuart. 2004. "When sheep bark: the parliamentary Labour Party since 2001." *British Elections & Parties Review* 14(1):211–229.
- Cowley, Philip and Sarah Childs. 2003. "Too Spineless to Rebel? New Labour's Women MPs." *British Journal of Political Science* 33(3):345–365.
- Cox, D. R. and A. Stuart. 1955. "Some Quick Sign Tests for Trend in Location and Dispersion." *Biometrika* 42(1-2):80–95.
- Cox, Gary W. 1987. *The efficient secret: The cabinet and the development of political parties in Victorian England*. Cambridge University Press.
- Crowe, Edward. 1986. "The web of authority: Party loyalty and social control in the British House of Commons." *Legislative Studies Quarterly* pp. 161–185.
- Eggers, Andrew C and Arthur Spirling. 2016. "Party cohesion in Westminster systems: inducements, replacement and discipline in the house of commons, 1836–1910." *British Journal of Political Science* 46(3):567–589.
- Eggers, Andrew C and Arthur Spirling. 2017. "Incumbency Effects and the Strength of Party Preferences: Evidence from Multiparty Elections in the United Kingdom." *The Journal of Politics* 79(3):903–920.
- Gaines, Brian J. 1998. "The Impersonal Vote? Constituency Service and Incumbency Advantage in British Elections, 1950-92." *Legislative Studies Quarterly* 23(2):167–195.
- Green, Jane and Sara B Hobolt. 2008. "Owning the issue agenda: Party strategies and vote choices in British elections." *Electoral Studies* 27(3):460–476.

- Green, Jane and Will Jennings. 2012. "Valence as Macro-Competence: An Analysis of Mood in Party Competence Evaluations in Great Britain." *British Journal of Political Science* 42(2):311-343.
- Hanretty, Chris, Benjamin E Lauderdale and Nick Vivyan. 2017. "Dyadic representation in a Westminster system." *Legislative Studies Quarterly* 42(2):235-267.
- Heath, Anthony, John Curtice, Roger Jowell, Geoffrey Evans, John Field and Sharon Witherspoon. 1991. *Understanding Political Change: the British Voter, 1964-1987*. Oxford: Pergamon.
- Heath, Oliver. 2015. "Policy Representation, Social Representation and Class Voting in Britain." *British Journal of Political Science* 45(1):173-193.
- Jackson, Nigel and Darren Lilleker. 2011. "Microblogging, Constituency Service and Impression Management: UK MPs and the Use of Twitter." *The Journal of Legislative Studies* 17(1):86-105.
- Kam, Christopher J. 2009. *Party discipline and parliamentary politics*. Cambridge University Press.
- Kam, Christopher, William T Bianco, Itai Sened and Regina Smyth. 2010. "Ministerial selection and intraparty organization in the contemporary British parliament." *American Political Science Review* 104(2):289-306.
- Kellermann, Michael. 2012. "Estimating ideal points in the British House of Commons using early day motions." *American Journal of Political Science* 56(3):757-771.
- Kendall, Maurice. 1975. *Rank Correlation Methods*. 4th ed. London: Charles Griffin.
- King, Anthony. 1981. "The rise of the career politician in Britain and its consequences." *British Journal of Political Science* 11(3):249-285.
- Lijphart, Arend. 2012. *Patterns of democracy: Government forms and performance in thirty-six countries*. Yale University Press.
- Mann, H.B. 1945. "Non-parametric tests against trend." *Econometrica* 13(3):163-171.
- Mosteller, Fred and David Wallace. 1964. *Inference and Disputed Authorship: The Federalist*. Reading, Mass.: Addison-Wesley Publishing Company.
- Mosteller, Frederick and David L. Wallace. 1963. "Inference in an Authorship Problem." *Journal of the American Statistical Association* 58(302):275-309.
- Norris, Pippa and Joni Lovenduski. 1995. *Political recruitment: Gender, race and class in the British Parliament*. Cambridge University Press.
- Norton, Philip. 1975. *Dissension in the House of Commons: intra-party dissent in the House of Commons' division lobbies, 1945-1974*. London: Macmillan.

- Norton, Philip. 1980. *Dissension in the House of Commons: 1974-1979*. Oxford: Oxford University Press.
- Norton, Philip. 1999. "The individual member in the British house of commons: Facing both ways and marching forward." *The Journal of Legislative Studies* 5(3-4):53–74.
- Norton, Philip and David M Wood. 1993. "Back From Westminster: constituency service by British members of parliament." *Lexington: The University Press of Kentucky* .
- Pattie, Charles, Edward Fieldhouse and R. J. Johnston. 1994. "The Price of Conscience: The Electoral Correlates and Consequences of Free Votes and Rebellions in the British House of Commons, 1987-92." *British Journal of Political Science* 24(3):359–380.
- Proksch, Sven-Oliver and Jonathan B Slapin. 2012. "Institutional foundations of legislative speech." *American Journal of Political Science* 56(3):520–537.
- Rheault, L, Beelen K, Cochrane C and Hirst G. 2016. "Measuring Emotion in Parliamentary Debates with Automated Textual Analysis." *PLOS ONE* 11(12).
- Rush, Michael. 2001. *The Role of the Member of Parliament Since 1868: From Gentlemen to Players*. Oxford: Oxford University Press.
- Rush, Michael and Philip Giddings. 2011. *Parliamentary Socialisation*. London: Palgrave Macmillan UK pp. 174–194.
- Rush, Michael and Sarah Childs. 2004. *A Changing Culture*. London: Palgrave Macmillan UK pp. 50–66.
- Searing, Donald. 1994. *Westminster's world: understanding political roles*. Harvard University Press.
- Shaw, Eric. 2001. "New Labour: New Pathways to Parliament." *Parliamentary Affairs* 54(1):35–53.
- Slapin, Jonathan, Justin Kirkland, Joseph Lazzaro, Patrick Leslie and Tom O'Grady. 2017. "Ideology, Grandstanding, and Strategic Party Disloyalty in the British Parliament." *American Political Science Review* p. 116.
- Spirling, Arthur and Kevin Quinn. 2010. "Identifying intraparty voting blocs in the UK House of Commons." *Journal of the American Statistical Association* 105(490):447–457.
- Vivyan, Nick and Markus Wagner. 2012. "Do voters reward rebellion? The electoral accountability of MPs in Britain." *European Journal of Political Research* 51(2):235–264.
- Vivyan, Nick and Markus Wagner. 2015. "What Do Voters Want From Their Local MP?" *The Political Quarterly* 86(1):33–40.
- Weissberg, Robert. 1978. "Collective vs. dyadic representation in Congress." *American Political Science Review* 72(2):535–547.
- White, Halbert. 1980. "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity." *Econometrica: Journal of the Econometric Society* pp. 817–838.

Supporting Information

A Mathematical Derivation of the Style Estimator

Consider the context c of a particular parliamentary session. Let S_c denote the reference set of speakers for this context: all speaking back-benchers from the governing party with at least 10 speeches. Let t denote the target speaker, any member of S_c . Our formal goal is to define the distinctiveness of speaker t relative to speaker set S_c in context c . For the purposes of defining the distinctiveness measure, we suppose that each speaker s belongs to a unique context $c(s)$; a member appearing in two parliamentary sessions is treated as two speakers.³

To start, we will need a probabilistic model relating a speaker to the text of a speech. For purposes that will become clear momentarily, recall that a word ‘type’ is a distinct entity or concept in a text, while a word ‘token’ is an instance of that entity. Thus, the phrase ‘Dog eat dog world’ contains four tokens (the words), but only three types (the second dog is conceptually similar to the first).

We begin, by reducing the text of each speech to a sequence of word tokens drawn from fixed vocabulary V_c specific to the context c . In our application, V_c is the total set of word types for session c . That is, the union of the types used by all the MPs who speak during that period. We ignore all spoken words outside this vocabulary. Let I_s denote the set of all speeches from speaker s . For speech $i \in I_s$ and word type $v \in V_c$, let x_{iv} denote the number of word tokens in speech i equal to v ; let $n_i = \sum_{v \in V_c} x_{iv}$ denote the length of speech i .

Our first simplifying assumption is that each speaker $s \in S_c$ has a set of word type probabilities that determine how speeches from s are generated. It will be convenient to parameterize these probabilities in terms of their natural logarithms. Specifically, take a particular speech i given by speaker $s \in S_c$. Let w denote a randomly-chosen word token from this speech, and suppose that the probability that this word is $v \in V_c$ in terms of its logarithm by

$$\log \Pr(w = v | s) = \eta_{sv},$$

the same for all speeches by s in context c . Denote the speaker-specific vector of such log-probabilities by η_s .

Suppose that we know η_s for each speaker (or that we have estimated these vectors using speakers’ empirical word frequencies). Our next task will be to use these quantities together with the speeches from context c to define the distinctiveness of each speaker.

To define the distinctiveness of target speaker t with respect to reference set S_c in context c , we start by taking the simple case where the reference set contains only two speakers, $S_c = \{s, t\}$, and the context contains only a single speech, i . Suppose that we randomly pick a word token from speech i . If, on the basis of this token, it is easy to identify whether t is the speaker, then we will say that t is distinctive. If, on the other hand, it is difficult to identify whether t is the speaker, then we will say that t is typical (not distinctive). In particular, suppose that we randomly pick word type v from the speech. Using Bayes’ rule and equal prior probabilities for whether s or t is the speaker, the log posterior odds ratio that t is the speaker are given by $(\eta_{tv} - \eta_{sv})$. The expected

³This is simply for measurement purposes in the sense that we have to define the unit of observation: when making *inferences* from the data below, we will use fixed effects to look at within MP variation.

value of this quantity for a random word type drawn from speech i is

$$\frac{1}{n_i} \sum_{v \in V_c} x_{iv} (\eta_{tv} - \eta_{sv}).$$

We define this quantity as the distinctiveness of target speaker t relative to reference set $\{s, t\}$ in the context of speech i .

With the simple case covered, it is straightforward to generalize our distinctiveness measure to larger reference speaker sets and larger contexts. In the first direction, for an arbitrary reference set S_c containing t , we take the average pairwise distinctiveness for a randomly chosen alternative $s \in S_c$:

$$\frac{1}{|S_c|} \sum_{s \in S_c} \left\{ \frac{1}{n_i} \sum_{v \in V_c} x_{iv} (\eta_{tv} - \eta_{sv}) \right\},$$

where $|S_c|$ denotes the size of set S_c . The second generalization, to larger reference sets, can be obtained by taking the expectation over a randomly-chosen speech $i \in I_t$. This gives our final measure of distinctiveness, which, after re-arranging the sums, can be expressed as

$$\mathbb{D}_t = \mathbb{D}_t(S_c; I_t) = \frac{1}{|I_t|} \frac{1}{|S_c|} \sum_{i \in I_t} \sum_{s \in S_c} \sum_{v \in V_c} f_{iv} (\eta_{tv} - \eta_{sv}),$$

where $f_{iv} = x_{iv}/n_i$. We can further simplify this expression by defining

$$\bar{f}_{tv} = (1/|I_t|) \sum_{i \in I_t} f_{iv} \quad \bar{\eta}_{cv} = (1/|S_c|) \sum_{s \in S_c} \eta_{sv}.$$

In this case,

$$\mathbb{D}_t = \sum_{v \in V_c} \bar{f}_{tv} (\eta_{tv} - \bar{\eta}_{cv}).$$

With this final expression we can compute \mathbb{D}_t by taking the difference between η_t and the average η_s over all speakers $s \in S_c$, then taking the dot product with the empirical frequencies \bar{f}_{tv} computed from target speaker's speeches I_t . It is \mathbb{D}_t that becomes our 'distinctiveness' dependent variable in what follows.

A.1 Standard Errors

The distinctiveness measure $\mathbb{D}_t = \mathbb{D}_t(S_c; I_t)$ is an empirical average over all observed speeches I_t by the target speaker t and all other speakers in the context S_c of the quantity $\sum_{v \in V_c} f_{iv} (\eta_{tv} - \eta_{sv})$. There are at least three sources of variability in that affect \mathbb{D}_t :

1. The observed speeches I_t can be considered as a sample of all speeches that could have been delivered by the target t in the same context.
2. The reference speakers S_c (in our context, the other back-benchers in the governing party) can be considered a sample of all potential reference speakers that could have been present.
3. The speaker-specific log word type frequencies η_{sv} are estimates based on empirical frequencies; these estimates depend on the actual observed speeches by s , which, again, can be

considered as a sample of all potential speakers by s .

To assess the variability in our computed value in \mathbb{D}_t we make the simplifying approximation that the largest sources of variability come from the random process that determined I_t and S_c ; we ignore variability in determining the estimates of η_{sv} .

Set $D_{is} = \sum_{v \in V_c} f_{iv}(\eta_{tv} - \eta_{sv})$. The quantity \mathbb{D}_t is an average of D_{is} over sets I_t and S_c . We will condition on the sizes of the sets I_t and S_c but otherwise we will treat these sets as random. Specifically, set $n = |I_t|$ and $m = |S_c|$; take I_t to be a set of independent identical draws from some population \mathbb{I}_t and take S_c to be a set of m independent draws from some population \mathbb{S}_c . Treat D_{is} as a deterministic function of the speech $i \in \mathbb{I}_t$ and the speaker $s \in \mathbb{S}_c$.

To assess the uncertainty in \mathbb{D}_t due to the variability in I_t and S_c first define for each $i \in \mathbb{I}_t$.

$$D_i(S) = \frac{1}{m} \sum_{s \in S} D_{is}$$

If S is a set of size m drawn independently and identically from population \mathbb{S}_t then define the expectation and variance over random S as

$$\mathbb{E}\{D_i(S)\} = \mu_i, \quad \text{var}\{D_i(S)\} = \frac{\sigma_i^2}{m}.$$

where μ_i and σ_i^2 are the mean and variance of D_{is} as s ranges over population \mathbb{S}_c .

Express the distinctiveness over a random set speaker set S of size m drawn as before and a random speech set I of size n drawn independently and identically from population \mathbb{I}_t as a random variable

$$D = D(I, S) = \frac{1}{n} \sum_{i \in I} D_i(S).$$

Note that $\mathbb{D}_t = D(I_t, S_c)$. Now, $\text{var}(D) = \mathbb{E}\{\text{var}(D | I)\} + \text{var}\{\mathbb{E}(D | I)\}$, where the outer expectation and variance on the right hand side are over the random set I . Using the independence of the speeches yields

$$\mathbb{E}(D | I) = \frac{1}{n} \sum_{i \in I} \mu_i, \quad \text{var}(D | I) = \frac{1}{n^2} \sum_{i \in I} \frac{\sigma_i^2}{m}.$$

Hence,

$$\text{var}(D) = \frac{1}{n} \text{var}(\mu_i) + \frac{1}{nm} \mathbb{E}(\sigma_i^2),$$

the variance and expectation being computed over a random i drawn from population \mathbb{I}_t . Define estimate $\hat{\mu}_i = D_i(S_c)$ and set $\hat{\sigma}_i^2$ to be the empirical variance of D_{is} as s ranges over S_c . We estimate $\text{var}(\mu_i)$ by the empirical variance of $\hat{\mu}_i$ and we estimate $\mathbb{E}(\sigma_i^2)$ by the empirical mean of $\hat{\sigma}_i^2$. This gives us an estimate of the variance of \mathbb{D}_t ; we use the square root of this quantity as a standard error for \mathbb{D}_t .

A.2 Estimation Notes: Vocabulary Standards

We use n -fold cross-validation to select the vocabulary V_c for each session c according to the following procedure:

From the raw set of all words spoken during each session, we compute the frequency of each type. We select types above a given percentile p as a candidate vocabulary v_p for $p \in \{50, 55, \dots, 90, 95, 99\}$, e.g. the vocabulary above the 50th percentile is comprised of terms that occur more frequently than the median.

For the cross-validation procedure, we partition the data from each session into n folds (in our case, 5 folds each comprised of 20 percent of speeches). For a given v_p , we hold out one of the folds and fit a model to the other 4 folds of data, repeating this procedure 5 times per v_p . We use the fitted model to predict the MPs who spoke speeches in the held-out fold of data, and compute the mean prediction rate obtained. This procedure is repeated for each $v_{p \in P}$ for each session. We select as V_c the vocabulary with the maximum mean prediction rate for a given session.

B Caveats Regarding our Style Estimates

With reference to the main text, and Supporting Information A, we stress the following about our estimates of distinctiveness:

1. **Relative Distinctiveness** It is impossible to talk about how “distinctive” a speaker is in an *absolute* sense. We can only measure her distinctiveness relative to a particular reference set of speakers. In the case of Mosteller and Wallace (1963), Madison was distinctive relative to the reference set that contains Madison and Hamilton. But Madison need not be distinctive relative to other contemporaries.
2. **Distinctiveness heterogenous *within* members** Even for a fixed reference set, a speaker can exhibit varying degrees of distinctiveness depending on the speech. Over a session of parliament, any given member will give some largely procedural speeches and some essentially free-form speeches. For the procedural speeches, the members tend to speak in the same manner; for the others, the members exhibit more heterogeneity in the speaking style. Any given member, in the context of a procedural speech, will not exhibit particularly distinctive style. In the context of open debate, she may exhibit a markedly different style and level of distinctiveness relative to her peers. Even for the same speaker and reference set, distinctiveness can vary depending on the speech.
3. **Convergence Assumptions and Validation** A reliable measure of distinctiveness should have the property that with enough data, it converges to a fixed value. In this paper, we define the distinctiveness of target speaker t relative to reference speaker set S_c in context c . The context c will be a particular parliamentary session, and the reference set S_c will be the set of members from the same session with the same party as t . We have at our disposal a set of speeches drawn from this context. We define distinctiveness such that, provided certain independence assumptions are in force, the measure converges in probability to a fixed value as the number of speeches from context c increases. To the extent that our assumptions are reasonable approximations of reality, our measure will accurately quantify distinctiveness. Even when our assumptions are unreasonable, it may still be the case that our distinctiveness measure approximates that which we hope to capture. We rely on external validation to argue this case.

C Variable Summaries

In Table 2, for each variable, we take the *mean* over all (Government party, backbench) MPs in a session, and then report various summaries of those means.

Description	Minimum	Median	Mean	Max
Experience	2.40	9.92	9.88	15.68
Proportion Demoted	0.00	0.13	0.13	0.31
Distinctiveness	0.39	0.44	0.44	0.51
Average number of speeches per MP (per career)	10.00	38.74	52.08	482.94
Average number of speeches per MP (per session)	26.14	57.87	56.68	135.87
Total speeches per session	3424	10979	11124	29212
Speaking MPs per session	102.0	187.0	193.0	288.0
Mean Speech Length	115.8	188.9	183.6	229.6

Table 2: Summaries for variables in our data

D Validation: MPs

For each MP t , in each session, we have an estimate of their distinctiveness in log-odds terms: our \mathbb{D} , above. For current purposes, however, we focus on something related but more concrete and directly interpretable: the proportion of their speeches which are correctly predicted as being from them relative to all other MPs (‘PCP’ in the tables below). We use 5-fold cross-validation to fit a model to texts from a given session, predict the speakers of held-out texts using this model, and calculate each speaker’s rate of correct predictions; we report each speaker’s mean PCP. To validate these estimates, we consider their extrema—their minimums and maximums. In the subsection tables below, we list the twenty names of the MPs who were most distinct (most interesting) and least distinct (most boring) by this measure (subject to having made a minimum of 20 speeches). We do this for two sessions: one in 1995–1996 and one in 1998–1999—thus, either side of the Blair landslide. We also list the number of mentions of each MP in the *Times* newspaper archives (via Gale Group Digital Archive) for the same period, specifically the ‘Politics and Parliament’ subsection of the ‘News’. Note that we searched for the person’s (professional) first name and last name together (as a bigram). Our maintained assumption here is that more interesting MPs will tend to be discussed more: whether in the news, editorials or in parliamentary sketches. Of course, there are various reasons why that assumption might not hold and indeed, there are technical issues with this measure. For example, a politician might be mentioned for something they did prior to the relevant search. And, on inspection, it is apparent that we sometimes miss mentions since the searching of an OCR’d newspaper has less than perfect recall. Nonetheless, we hope that with our qualitative comments this helps validate our measure.

D.1 Tory Backbenchers, 1995–1996

Comparing the counts from Tables 3 and 4, a Wilcoxon Signed-rank test returns a statistically significant result for the former having a higher mean ($p < 0.05$).

More substantively, we note the presence of several well-known Eurosceptics in Table 3. These include some of the so-called ‘Maastricht Rebels’ like John Wilkinson and Andrew Hunter who abstained or voted against their Conservative government on the relevant treaty (note that Barry Legg, Teresa Gorman, Teddy Taylor, Bill Cash and Rupert Allason were all just outside our top 20). Also present are other Eurosceptics like Patrick Nicholls, Eric Pickles and Edward Leigh (who was fired by Major from an under-secretary role for his views). Michael Colvin, unusually for the time, was opposed to restricting gun ownership rules after the Dunblane Massacre in 1996. David Mellor voted similarly to Colvin. Barry Field had initially decided he was sufficiently interesting to challenge John Major for the leadership of the Tory party itself, but ultimately withdrew following the emergence of John Redwood. Peter Bottomley was described (by the *Independent*, “The maverick with ‘five ideas: four good, one mad’”, 11 July, 1993) as being notable for “his delight in surprising colleagues with a range of apparently perverse causes.” Our most ‘interesting’ MPs are Edward Heath and David Wilshire. Heath is a former Prime Minister, and someone who actively criticized the Conservative governments of Thatcher and Major for being too economically liberal at the expense of social cohesion. Meanwhile, David Wilshire was a right wing MP responsible for the initial introduction of Section 28 (which meant that local authorities could not “promote homosexuality or . . . promote the teaching in any maintained school of the acceptability of homosexuality”). He also criticized initial plans (in 1995) to allow Hong Kong Chinese to settle in the UK after the hand-back of the territory to China.

Among the least interesting MPs, Edwina Currie is perhaps the only one worthy of further comment. Currie had been controversial cabinet minister (forced to resign in 1988), and by the mid-1990s was a novelist with two popular tomes written. This perhaps inflates her mentions in the *Times*. We note candidly that Nicholas Budgen, a known Euroskeptic, appears in this list too—in contrast to his rebellious colleagues who generally appear to be more interesting overall.

Table 3: Most interesting MPs November 1, 1995–October, 31, 1996 in parliament, by proportion of speeches correctly predicted.

PCP	Name	Times Mentions
0.41	Piers Merchant	1
0.41	Phil Gallie	3
0.43	Norman Fowler	3
0.43	Anthony Steen	2
0.43	Terence Higgins	6
0.43	Eric Pickles	0
0.44	James Pawsey	3
0.45	Andrew Hunter	4
0.46	Peter Brooke	5
0.47	Peter Bottomley	2
0.48	Edward Leigh	6
0.48	Robert Hughes	2
0.48	Patrick Nicholls	5
0.48	John Wilkinson	4
0.51	David Howell	4
0.52	David Mellor	3
0.64	Michael Colvin	0
0.66	Barry Field	4
0.72	Edward Heath	19
0.75	David Wilshire	4

Table 4: Most boring MPs November 1, 1995–October, 31, 1996 in parliament, by proportion of speeches correctly predicted.

PCP	Name	Times Mentions
0.00	Alan Haslehurst	0
0.00	Michael Lord	2
0.00	Peter Tapsell	1
0.00	Tim Yeo	2
0.00	Charles Hendry	1
0.00	Nicholas Budgen	7
0.00	Matthew Carrington	0
0.00	Michael Jopling	2
0.00	Gary Waller	2
0.00	Anthony Durant	0
0.00	Hartley Booth	0
0.00	John Butcher	0
0.00	Donald Thompson	0
0.00	Edwina Currie	14
0.00	Jerry Hayes	2
0.00	Peter Butler	3
0.01	John Whittingdale	1
0.02	Michael Neubert	3
0.03	Alan Duncan	3
0.03	Nick Hawkins	0

D.2 Labour Backbenchers, 1998–1999

Comparing the counts from Tables 5 and 6, a Wilcoxon Signed-rank test returns a statistically significant result for the former having a higher mean ($p < 0.05$). More substantively, we note the presence of several Labour ‘rebels’ among the most distinct. These include Tony Benn, Diane Abbott, John McDonnell, Roger Berry and Tam Dalyell, all of whom consistently voted against the Labour government’s plan to reform the welfare state.⁴ Peter Temple-Morris was a party switcher, and ‘interesting’ for that reason—he was elected as a Tory MP in 1997, but then crossed the floor to Labour the same year. The most interesting MPs here include Stuart Bell, who was the Church Estates Commissioner, meaning that he was one of the managers of the Church of England’s property. David Hinchliffe, chairman of the select committee on Health, was subsequently extremely critical of the Blair government’s proposed reforms to the NHS. Finally, Barry Jones was the chair of the Intelligence and Security Committee of Parliament.

The set of least interesting MPs contains fewer obvious ‘stars’, and consists mostly of loyalists like Doug Naysmith, Ivan Lewis and Charlotte Atkins. Exceptions include Ronnie Campbell who would subsequently become rebellious regarding the Iraq War, and Dennis Canavan who was dismissed (in 1999) from the Labour Party after a dispute over whether he could be an official party

⁴Jeremy Corbyn had a PCP of 0.216 for this session, placing him around the 60th percentile for interestingness by this metric.

candidate for the Scottish Parliament (although he was not particularly notable as a government critic prior to this development).

Table 5: Most interesting MPs November 1, 1998–October, 31, 1999 in parliament, by proportion of speeches correctly predicted.

PCP	Name	Times Mentions
0.48	Hazel Blears	1
0.48	Denzil Davies	0
0.49	Geraint Davies	0
0.50	Gareth Thomas	1
0.51	Diane Abbott	3
0.51	Peter Temple-Morris	0
0.53	Tony Benn	19
0.53	Martin Linton	0
0.53	David Winnick	5
0.54	Harry Barnes	1
0.57	Norman Godman	1
0.58	Tam Dalyell	3
0.59	Bill Rammell	0
0.60	Joan Walley	0
0.62	John McDonnell	1
0.66	Roger Berry	1
0.70	David Hinchliffe	1
0.72	Donald Anderson	3
0.73	Stuart Bell	3
0.75	Barry Jones	1

Table 6: Most boring MPs November 1, 1998–October, 31, 1999 in parliament, by our measure.

PCP	Name	Times Mentions
0.00	Charlotte Atkins	0
0.00	Nigel Beard	1
0.00	Ben Bradshaw	1
0.00	Ronnie Campbell	3
0.00	Dennis Canavan	0
0.00	Ivor Caplin	1
0.00	Helen Brinton	0
0.00	Barbara Follett	0
0.00	Linda Gilroy	0
0.00	Win Griffiths	0
0.00	John Heppell	0
0.00	Helen Jackson	0
0.00	Jenny Jones	1
0.00	Jackie Lawrence	1
0.00	Ivan Lewis	1
0.00	Stephen McCabe	0
0.00	Shona McIsaac	0
0.00	John Maxton	2
0.00	Doug Naysmith	0
0.00	James Plaskitt	0

E Validation: Tokens

We can also inspect the ‘most influential’ terms, both within certain sessions and over the data as a whole. Term influence is estimated separately for each speaker–term pair in each session, and is calculated for term v and speaker s by multiplying the mean (over speaker s ’s speeches) term frequency count $\frac{1}{|I_s|} \sum_{i \in I_s} x_{iv}$ with the mean-centered term frequency rate η_{sv} (the latter term is $\eta_{sv} - \frac{1}{|S|} \sum_{s \in S} \eta_{sv}$). Thus, the difference between the speaker’s rate of use of word v and the mean rate of use of v is weighted by how many times the speaker actually used the word v . One consequence of this arrangement is that stop words tend to be ‘influential’ because they are so commonly used, boosting the mean frequency rate across all speakers.

In Table 7 and Table 8 we list the ‘top 50’ most influential words from the pre- and post-Blair election landslide, that we discussed in Supporting Information D.

Table 7: Most influential terms, 1995–1996.

the	of	to	and	that	i	we	in	is	a
it	hon	for	not	my	be	have	bill	are	will
they	was	he	as	labour	people	our	friend	on	local
would	european	has	education	which	committee	council	right	service	government
by	do	defence	schools	or	who	there	about	been	member

Table 8: Most influential terms, 1998–1999.

the	of	to	that	and	i	in	a	we	is
it	not	for	london	have	be	people	was	health	are
they	hon	will	women	my	on	children	bill	local	government
he	as	house	industry	who	by	scottish	our	police	with
their	wales	friend	would	right	but	about	there	has	committee

As is clear and expected, we indeed see a large number of stop words in both tables: ‘the’, ‘of’, ‘to’ and so on. More compellingly though, we also words that make sense as important or distinctive during this time. For example, Tory backbenchers made ‘european’ influential, which accords to the protracted and fractious discussion of the various EU treaties at this time. Meanwhile, ‘defence’—a common right wing concern—turns up as an influential word in 1995, but not in 1998 (among Labour MPs). After Blair’s election, it is words like ‘london’, ‘scottish’ and ‘wales’ that are most influential. Given that the Labour government was discussing and passing legislation to create new governing institutions in the capital, Scotland and Wales, this makes some sense. We also typical left-wing concerns pertaining to ‘women’, ‘children’ and ‘health’ at this time.

A related validation exercise is to study the way that some of the (non-stop) terms we identified above behave with respect to the entire series. In Figure 3 we do exactly that, with a lowess imposed on the data for visualization purposes. In particular, we plot the ‘average influence’—defined as the mean of a term’s influence over all speakers within a session—of the various tokens for every session of parliament under study. Starting at the top left, we see that ‘defence’ surged in influence during the Cold War, and then fell away after it ended, around 1990. Moving right, we see ‘european’ coming to a peak during the Tory rebel years of the early 1990s, during the Maastricht negotiations. Moving to the bottom left, ‘Scottish’ peaks in the late 1970s and late 1990s—the periods in which some devolved rule was offered to Scotland (and referendums took place on the same). Finally, ‘health’ peaks in the late 1990s/early 2000s as the Blair government abolished the internal market, added large real term budget increases, but allowed independent foundation status for hospitals. This latter move was very unpopular with Labour rebels, and in 2003 they almost defeated the government’s bill in the Commons that dealt with this policy change.

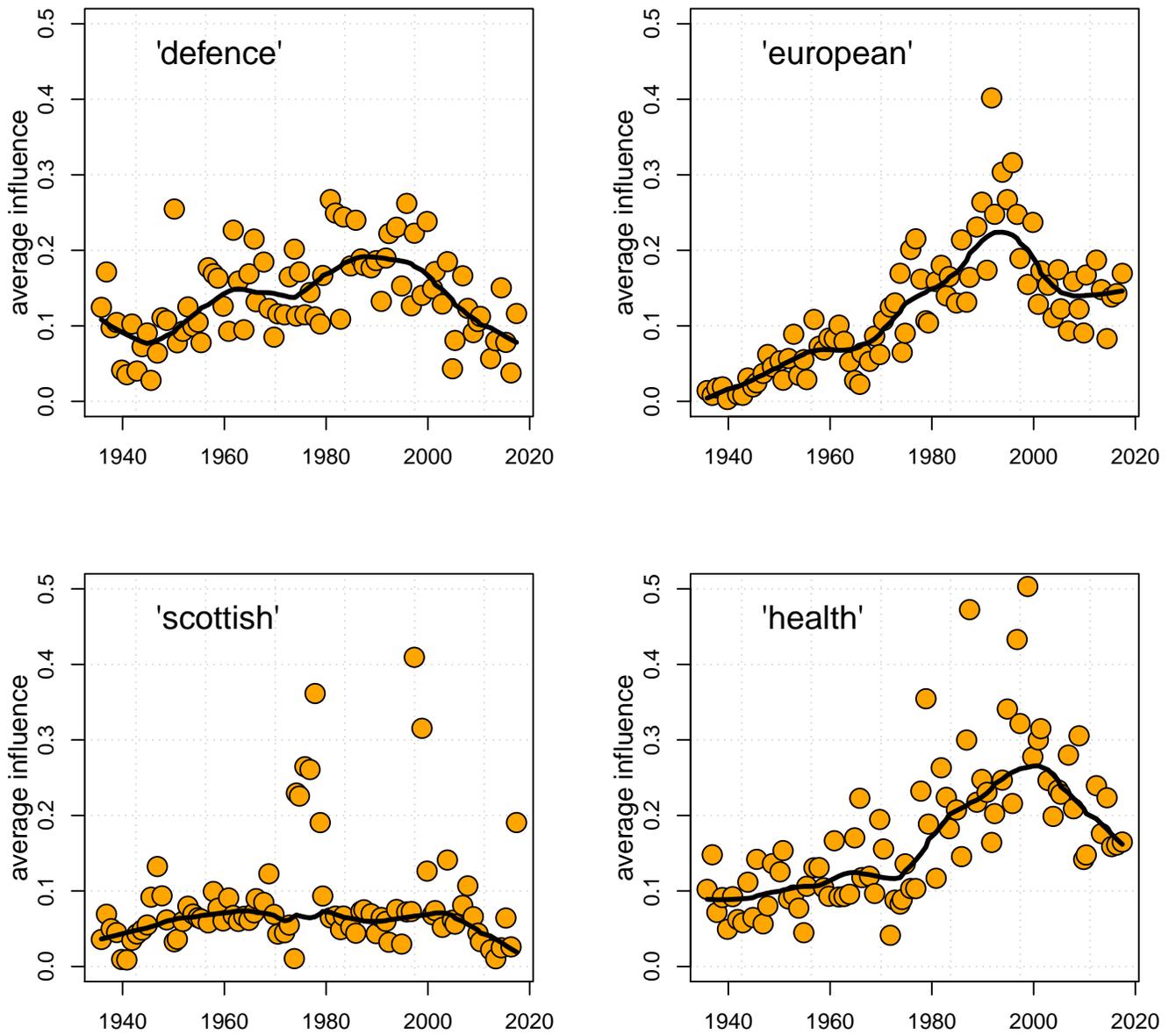


Figure 3: Average influence of some key terms over time. The non-stop word tokens from the sessions of 1995 and 1998 accord with our priors, in terms of their behavior over the entirety of the data. For example, ‘defence’ ceases to be (as) influential once the Cold War ends.