Democratization and Linguistic Complexity
The Effect of Franchise Extension on Parliamentary Discourse, 1832–1915

Arthur Spirling†

Abstract
We consider the impact of the Second Reform Act, and the doubling of the electorate it delivered, on the linguistic complexity of speeches made by members of parliament in Britain. Noting that the new voters were generally poorer and less educated than those who already enjoyed the suffrage, we hypothesize that cabinet ministers had strong incentives—relative to other members—to appeal to these new electors with simpler statements during parliamentary debates. We assess this claim with a data set of over half a million speeches for the period between the Great Reform Act and Great War, along with methods for measuring the comprehensibility of texts—which we validate in some detail. The theorized relationship holds: ministers become statistically significantly easier to understand (on average) relative to backbenchers, and this effect occurs almost immediately after the 1868 election. We show that this result is not an artifact of new personnel in the House of Commons.

†Department of Government, Harvard University. aspirling@gov.harvard.edu
1 Introduction

Few topics have featured as prominently in applied political science research as the causes and consequences of democratization (e.g. Lipset, 1959; Huntington, 1968; Przeworski et al., 2000; Boix and Stokes, 2003; Acemoglu and Robinson, 2005). Of particular interest is the (optimistic) notion that with franchise extension and competition comes increasing political responsiveness and accountability for citizens (e.g. Bartolini, 2000; Przeworski, 2009). And within this large literature the changes to Britain in the middle of the nineteenth century have captured a great deal of scholarly attention (e.g. Bagehot, 1873/2011; Seymour, 1915; Trevelyan, 1922; Gash, 1952; Woodward, 1962), with primary focus on the passing of the relevant legislation in 1867 (e.g. Himmelfarb, 1966; Smith, 1967; Walton, 1996; Moser and Reeves, 2014) and its effects on politicians and voters (e.g. Cox, 1987; McLean, 2001; Berlinski and Dewan, 2011; Camp, Dixit and Stokes, 2014). In part, this is because the ‘Westminster system’ that resulted has been widely emulated for its stability and decisiveness (e.g. Lijphart, 1999; Rhodes and Weller, 2005), and there is thus a natural interest in uncovering its development and possibly charting its future course. This is especially true of its characteristic institutions of ministerial responsibility and fierce frontbench competition in parliament and in the electorate.

In keeping with this interest, in the current paper we seek to understand how suffrage extension affected the behavior of backbench members of parliament (MPs) relative to ministers during the Victorian period. Our central idea is that members of the governing executive—the cabinet—had new incentives after the expansion of the electorate: they were required, as leaders of their parties, to appeal to a poorer, less educated median voter. We contend that they did so, in part, via simpler linguistic expressions in their parliamentary speeches. Meanwhile, because backbenchers increasingly understood that citizens were “voting for the
party, rather than for the man” (Cox, 1987, 136), we argue that these MPs were under considerably less pressure to adjust their speaking style.

We are hardly the first to investigate the new incentives for legislators and their principals introduced by voting reform. In the long term, the eventual rise of the parliamentary Labour party as an electoral force representative of working-class interests (see, e.g., Thompson, 1963; Cox, 1997), along with the commensurate decline of the Liberals (Searle, 2001), is probably the best known consequence of a larger franchise (when considered alongside other reforms such as the introduction of the secret ballot). By contrast, and with some exceptions (see, e.g., Hurst, 1965, on the effects of the secret ballot in Ireland) analysts find relatively little evidence of immediate change to other markers related to MP activity: these include roll call cohesion (Eggers and Spirling, 2014b), Liberal vote share (Berlinski and Dewan, 2011), the socioeconomic backgrounds of cabinet personnel (Berlinski, Dewan and Van Coppenolle, 2014) and party orientation in the electorate (Cox, 1987). On the one hand, these null findings are surprising: the almost doubling of the electorate via the Second Reform Act to include poorer, less educated voters was certainly predicted (or feared) by contemporaneous actors to have consequences for the nature of both substantive and descriptive representation (see, e.g., McLean, 2001, for a discussion of the ‘Adullamites’). On the other hand, scholars of the period do not typically have access to the kinds of fine-grained data that makes investigating possibly subtle shifts in behavior straightforward. Compare this situation, for example, with the study of American politics—in particular regarding the ‘Homestyle’ of Members of Congress, where researchers can either follow contemporary members and record their interactions (e.g. Fenno, 1978), experiment on them (see Grose, 2014, for an overview) and their constituents (e.g. Larson, 1990), or work with the large amounts of text produced by such elected officials in communication with their constituents and others (e.g. Grimmer, 2013). Similarly, researchers interested in Comparative politics
for the contemporary period can utilize manifestos (e.g. Budge et al., 2001; Benoit, Laver and Mikhaylov, 2009), parliamentary speeches (e.g. Slapin and Proksch, 2008), and detailed election studies with roughly similar questions across nations to estimate the degree to which different systems and different times respond to voter needs (e.g. Powell, 2000). Thus, while we have very strong priors for this vital period in the Westminster system’s history, testing our hypotheses is prohibitively difficult, and our empirical findings look ambiguous at best—and confusing at worst.

Diagnosing the problems with extant studies of the effects of democratization is not difficult; solving these maladies is far from trivial. Put crudely, scholars are typically restricted by limited data on elite—i.e. MP—responses to suffrage expansion: studies are either intensive with coverage of short time periods (e.g. Schonhardt-Bailey, 2008; Berlinski and Dewan, 2011), or more extensive in terms of numbers of observations but necessarily less fine-grained in terms of both the information and inferences that are possible (e.g. Cox, 1987). In the former case, researchers face the obvious difficulty that suffrage expansion may not have immediate consequences for politician behavior—perhaps because some degree of ‘learning’ must take place. In the case of the broad studies, though impressive in scope they are likely to miss subtle, small changes to the way that individual agents perceive the situations they face and adjust their actions accordingly. In both cases then, there is a danger that effects that do exist are ‘missed’. Putting aside these specific issues of study scope, it is far from obvious where we should look for evidence of new incentives and behaviors: ideally, we would have a large number of observations from which we can plausibly measure ‘responsiveness’ directly and in a way that allows us to compare both across individuals and across time.

In this paper, we make progress where other attempts have faltered and show evidence consistent with our contention of a differential behavioral affect on ministers versus back-
benchers. We do this in a way made possible with an analysis of a data set of hundreds of thousands of speeches—along with MP covariates—from the House of Commons between 1832 and 1915 (see Eggers and Spirling, 2014a). We focus on speeches as outputs precisely because they allow politicians to respond instantly to changing circumstances, without the various implementation lags one must allow when studying policy or party-system shifts. Further, in contrast to the 24-hour news cycle politics of the present day, parliamentary speeches (and reports on them) were the primary way that voters of the time monitored the actions of their representatives. Speeches have another advantage: they are an equitable resource insofar as, subject to recognition rules in the Commons, anyone can (and did) undertake them, allowing us observations for essentially the entire population of MPs. In summary then, speeches are very much a key place, if not the only place, where we might see democratization having an immediate and noticeable effect.

Our innovation methodologically is to measure the ‘comprehensibility’ (or complexity) of the utterances using well known metrics from education research, that take into account the number of syllables relative to the number of words found in documents (see Flesch, 1948). These are straightforward to calculate, and have been used elsewhere in the study of speech (e.g. Lim, 2008), albeit not on so many texts. These scores are combined with panel data techniques that allow multiple individuals, making multiple speeches, to be compared over time as their roles in the chamber change. In particular, we show that almost immediately after the Second Reform Act, cabinet ministers altered their speech in a way that made those speeches simpler to understand for the median member of the electorate—that is, someone poorer and less educated than had previously voted. This finding provides crucial support for earlier hypotheses regarding the leadership role that cabinet members increasingly played (relative to backbenchers) in appealing directly to popular opinion, such that their parties could win national elections (Cox, 1987; Jenkins, 1996; Rush, 2001). In terms of point pre-
dictions, our estimates imply that, controlling for length of speech and other member-level variables, being a minister after 1868 was equivalent to moving from around the 48th to the 60th percentile of comprehensibility in the chamber as a whole: with predicted values approximately ten percent larger than backbenchers. This finding is robust to the usual standard error corrections, and to alternate specifications and measurement strategies. Importantly, we are able to rule out the possibility that the change in language is due to new personnel arriving in the House over time: rather, it is the result of new incentives for those already there.

Although the techniques we use here are not new, we are applying them to a large data set and in an innovative way. With that in mind, we spend some time below exploring their details and validating their use, before moving to our results and conclusion. Prior to that, however, we set the substantive scene for our study: the advent of the Second Reform Act in 1867.

2 Appealing to the Newly Enfranchised

The *Representation of the People Act* of 1867—colloquially known as the Second Reform Act—has attracted much scholarly attention on its origins and passing (e.g. Himmelfarb, 1966; Smith, 1966; Cowling, 1967; McLean, 2001; Moser and Reeves, 2014), its details and its effects (e.g Laski, 1928; Smith, 1967; Canandine, 1999; Aidt, Daunton and Dutta, 2010; Berlinski and Dewan, 2011).\(^1\) The features of interest for our purposes are two-fold: first, the massive expansion of the franchise from around one to two million men.\(^2\) Second, the reduction in the property requirement needed for voting. Thus, in urban ‘borough’ constituencies, the electorate now included householders, renters and lodgers paying at least

---

\(^1\)The Act itself dealt with English and Welsh matters; Scotland and Ireland saw reforms via the Representation of the People (Scotland) Act and Representation of the People (Ireland) Act, both of 1868.

\(^2\)For reference, the 1871 census recorded a total population of around 26 million for England, Scotland and Wales combined. Around 12.6 million were males (of any age).
£10 a year for their accommodation. Meanwhile, in the ‘county’ constituencies, the suffrage was extended to lease-holders, tenants and others with much smaller annual payments for their property than had been the case previously. In practice, and importantly for our work here, the Act “brought substantial working-class majorities to the electoral registers of almost all the boroughs” (Walton, 1996, 35), and in this way suffrage expansion was disproportionately greater in urban areas than elsewhere. Indeed, using figures from Bowley (1937) and Mackenzie (1921), Berlinski and Dewan (2011, 7) note that “it is clear that the extension of the franchise gave the vote to urban unskilled workers.”

A salient feature of these new voters was that they were, on average, less educated (and less literate) than pre-existing holders of the suffrage. To see this, consider calculations from Mitch (1992), who obtains a large sample of marriage certificates for the period 1869–1873, and measures male literacy from the ability of grooms to sign their own names. The class status of the men is inferred from the occupation listed for them on the same certificate. In the Mitch (1992, 24–25) approach, there is a hierarchy of five socioeconomic classes, the latter three of which are characterized as “petty shopkeepers, skilled manual trades, mining, most transport occupations”, “semi-skilled manual labor”, and “unskilled labor” respectively.\(^3\) In Table 1 we provide figures for all five of these groups,\(^4\) in terms of the number of individuals recorded as literate and illiterate in each. As can be seen from the table, the top occupation class (I and II) have very few individuals (around 1%) who lack literacy skills. By contrast, grouping the lower three classes together, we see illiteracy rates at around 20%. These proportions are statistically significantly different \((p < 0.01)\).\(^5\)

\(^3\)For completeness, class (I) are those occupations types which are “titled, high public office, military officers”, class (II) are “professions, commerce, clerical, farmers”.

\(^4\)Calculated from Table 2.3 in Mitch (1992).

\(^5\)This result is robust to including class (III) as one of the ‘prior voter’ groups; separately, it is robust to dropping class (V) from the analysis altogether.
Table 1: Number of literate and illiterate men in Mitch (1992) sample, by occupational class. The ‘prior voters’ are those classes likely already enfranchised prior to the Second Reform Act; the ‘new voters’ are those classes more likely to be part of the newly extended franchise.

If we interpret the lower labor classes as being comprised of those joining the suffrage as a result of the Second Reform Act (which accords roughly with the distinctions made by Berlinski and Dewan (2011)), we have clear evidence that these ‘new voters’ were less educated and less literate than those already part of the franchise. Although we cannot observe this directly, it seems safe to further assume that those that were literate in classes of low overall literacy had lesser proficiency in reading and writing that the literate in classes of widespread literacy. That is, we suppose that the binary indicator of literacy hides continuous variation whereby literate voters prior to 1868 were on average better able to read than the literate who joined the franchise after the Second Reform Act. This matters for our causal story below, in which speeches are reported to electors mostly in written (i.e. newspaper report) form: thus we require that the literate among the new voters struggle with complex linguistic expressions more than the literate who were already voting.

2.1 Ministers as the Focus of Electoral Competition

A related consequence of the franchise expansion was the development of new political behavior by electors; in particular, “voting for the party, rather than for the man” at the ballot box,
with citizens increasingly “using their votes to determine what did matter: party control of the executive” (Cox, 1987, 136). Whether or not the rise of the ‘party orientated electorate’ was caused simply by the expanded suffrage *per se* is debatable (see Cox, 1987, 94–95), but there is little doubt that it focussed attention on the cabinet and its members as the key actors in politics, and the ones responsible for winning (or losing) elections. Commensurate with this new role as the locus of voter choice was an ongoing increase in partisan cohesion (beginning in the 1850s) in roll call voting, with leaders in the House of Commons able to discipline their troops at levels approximately equal to those in modern British politics (see Eggers and Spirling, 2014b). Crucially for our account, backbenchers had much weaker incentives than ministers to adjust their language. This is because, at a time when national party appeals began to matter more than local connections or family name, those without cabinet rank were no longer as important as they previously had been for winning their own seats: it was their leaders on the frontbenches who would be the deciders of election success or failure for *everyone* in their party.

Our central idea is that these forces—new, less educated voters, and the “triumph of partisan politics” (Jenkins, 1996, ch 6) in the electorate and in parliament—meant that the Westminster executive was faced with fresh challenges and opportunities. In particular, the cabinet was required to adopt strategies such that it could appeal to electors and compete successfully for power at the ballot box. While others have investigated these strategies as they pertained to election spending (Hanham, 1978; Camp, Dixit and Stokes, 2014, e.g.), and some have specifically investigated the emergence of early manifesto-style addresses such as in the Midlothian Campaign (see Kelley, 1960; Matthew, 1997), we turn our attention to the changing nature of speeches in the House of Commons itself.
2.2 Observational Implications

Our hypothesis is that, relative to other MPs after the franchise extension, cabinet members reduced the linguistic complexity of their speeches. Ministers did this to ensure that the newly increased electorate—with its lower average educational level—could understand and be convinced by executive speeches. Put more crudely, democratization resulted in the ‘dumbing down’ of rhetoric and argument by ministers in a way designed to win votes at the ballot box. Whether this proposed mechanism is convincing depends on the plausibility of several links in the causal chain. First, readers may question the extent to which parliamentary speeches were in fact disseminated to the public at large. For the period under study, this is not a concern: indeed, Victorian Britain was notable for “universal press coverage” of Commons activity and “the explosive expansion of the press in the middle of the century” (Cox, 1987, 54–55). This press penetration extended to poorer voters, especially after the repeal of taxes that had kept prices artificially high until the middle of the century. Thus, by 1861 the cheapest of the weekly sheets, including those aimed specifically at working-class voters such as *Lloyd’s Newspaper* and *Reynold’s News* had circulations of 412,000 and 150,000 respectively (Hewitt, 2013, 105). Furthermore, there is little doubt that parliamentary speeches did indeed make the news. The importance of this point was not lost on politicians of the day: for example, commenting on the implications of a discussion regarding women’s suffrage in 1873, radical MP John Bright opined from the backbenches that “The substance of this debate will be carefully reported in the newspapers, the report will go to every town and village in the United Kingdom, and to every English-speaking country under British rule…” (cited in Jenkins, 1996, 18). Second, it is clear that members themselves were acutely aware that the expansion of the suffrage would bring less educated (if not necessarily illiterate) voters into the electorate: Robert Lowe, leader of the ‘Adullamite’ Liberal MPs skeptical of the Second Reform Act noted that those who would be newly enfranchised exhibited “venality…ignorance …drunkeness” and were in general “impulsive, unreflecting
and violent people” (cited in Saunders, 2011, 206). Among more sympathetic MPs, the debate was not over whether the pool of voters to whom politicians must appeal would change, but rather the extent of those changes (see e.g. McLean, 2001, 66–67 on the ‘rating’ required for enfranchisement). Third, there is evidence that party leaders were aware of the need to appeal to these new voters, albeit in somewhat limited ways that did not alienate other members of the electorate. Thus Disraeli—or at least his Home Secretary Richard Cross—embarked on a series of union and labor reforms in the 1870s (St John, 2010, 151–153) while Salisbury actively pursued the ‘respectable’ suburban but working class “villa vote” (see Shannon, 1996). On the Liberal side, the party made early, if perhaps ineffectual, attempts in some cities to “recruit candidates with working-class appeal” prior to the 1868 election (Moore, 2006, 25) and the rise of “New Liberalism” thereafter is a specific example of an ideology that sought to embrace new voters and their concerns (Sykes, 1997).6

All told then, senior politicians on the government side of the House of Commons had strong reasons to adapt their policies and language in the aftermath of the Second Reform Act in a way that was less true of their backbench colleagues in their own party and among the opposition. Assessing whether cabinet members (relative to others) did or not respond to these new forces requires an assessment of their speeches over time. It is to our data on this that we turn before explicating our measurement strategy.

3 Data

Our data consists of speeches made in the House of Commons between 1832 and 1915. Thus, the Second Reform Act of 1868 occurs approximately half way into our time series, allowing a large window before and after in which to assess any effects on speech style. The speech data

6See also Jennings (426 1962) on Chamberlain’s ‘Radical Program’ for the 1885 General Election.
is matched to individual MPs, which is then matched to various covariates including party of election, cabinet status, and competitiveness of constituency elections. These measures and the data are described by Eggers and Spirling (2014a) (which draws on Craig (1989, 1974); Cook and Keith (1975); Butler and Butler (1994)). For the data set in its original form, there are 860,192 speeches for 4233 MPs, with an average of 203 speeches per member. For our analysis below, we restrict ourselves to members running for either the Conservative or Liberal party in elections, excluding various idiosyncratic versions of those labels, along with nationalist parties and the (early) Labour party. This leaves 675,997 speeches, from 3613 members, for an average of 187 speeches per MP. We are confident that restricting our data is appropriate for at least two reasons: first, because only the Liberals or Conservatives could plausibly form the cabinet during the period under study and thus these parties constitute the key actors for our work; second, our findings below are robust to including those other parties as part of the opposition. We make very few further ‘adjustments’ to our data set. In particular, we impose no minimum length on speeches (empirically, the minimum number of words is 1, the maximum is 11,000, with a mean around 248) and remove only one session from our analysis: the very short (just 129 speeches) first session after the indecisive 1892 general election, at which time Salisbury awaited a no confidence vote before resigning as Prime Minister.

In time-series terms, our data is divided into parliamentary ‘sessions’ each of which last approximately one year and which collectively comprise ‘parliaments’ (which begin after general elections). In the period under study, different sessions within the same parliament have different parties in cabinet because Victorian politicians did not always go to the country for a new popular mandate after their Prime Minister resigned or lost the confidence of the House. An example of this would be the ascension of Disraeli (who followed Lord Derby to the premiership) after Lord Russell’s Liberal government fell in 1866—without any
intervening election. Because ministers (then as now) could leave office at any time with no more general consequence for the session itself, our measurement of who is a minister in any given session is relatively inclusive. That is, the metric includes anyone who served at least one day in the cabinet during that session.

Having described our data, our next task is to provide a metric for measuring, and comparing, the comprehensibility of speeches made in the House of Commons.

4 Methods and Measurement

Starting at least with Sherman’s (1893) “Objective Study of English Prose and Poetry”, scholars of literature and education have been interested in the notion that texts could be statistically analyzed and their “readability” measured. Although this key quantity of interest has been variously defined—depending in part on the relevant researcher’s motivation (e.g. Dale and Chall, 1949)—at its core, readability refers to the comprehensibility of a text; literally, the ease with which it may be understood by a reader with varying levels of education. A number of metrics have been proposed for assessing comprehensibility (e.g. Lively and Pressey, 1923; Dale and Chall, 1949; Gunning, 1952; McLaughlin, 1969) with that of Flesch (1948) being the most famous and widely used (Klare, 1963). Flesch’s (1948) formula, given in Equation (1) yields a score for any given body of text that is known as the Flesch Reading Ease (FRE) statistic. In the original application from which it was derived, the value of the statistic was found to have lower bound of 0, and an upper bound of 100, though this need not be the case in other data sets.7 Though we will use the score directly

7The formula results from a study undertaken by Flesch in which he regressed the average grade level of school children who could answer at least 75% of multiple choice questions regarding comprehension of texts they read on a constant and the two bracketed variables in the equation. In that context, a score of 100 means that the document could be understood by a student with a fourth grade education and thus could be described as “barely functionally literate” (Flesch, 1948, 225).
in what follows, we note that educational researchers typically convert this output to a (minimum) number of years of US schooling—known as a Flesch-Kincaid Grade Level—that a student would require to find a given document comprehensible (see Kincaid et al., 1975).

\[
\text{Flesch score} = 206.835 - 1.015 \left( \frac{\text{total number of words}}{\text{total number of sentences}} \right) - 84.6 \left( \frac{\text{total number of syllables}}{\text{total number of words}} \right)
\]  

(1)

Inspection of Equation (1) suggests that the Flesch score is not difficult to calculate using modern processing programs, assuming some machine-readable version of the text exists.\(^8\) As can readily be seen, for a fixed number of words in a document, increasing the number of syllables of those words, and grouping the words into fewer sentences both increase the complexity of the text in question.

Guidelines for interpreting the statistic may be found in several sources (including Flesch, 1949, 149–150); Cann, Goelzhauser and Johnson (2014, 663) give the following: “Texts with FRE scores ranging from 0 to 30 are considered very difficult to read, 31 to 50 are difficult, 51 to 60 are fairly difficult, 61 to 70 are standard, 71 to 80 are fairly easy, 81 to 90 are easy, and 91 to 100 are very easy.” To anchor these categories conceptually, note that Cann, Goelzhauser and Johnson (2014) place the average academic political science article at around 33, on a par with judicial opinions, while the New York Times has a mean FRE of about 48 and childrens’ books such as Peter Pan and The Wind in the Willows have FRE scores approaching 80. Giving context for these scores outside of the school setting, Dalecki, Lasorsa and Lewis (2009, 6) calculate that “85 percent of Americans today can read at the 50–60 reading ease level, 72 percent at the 30–50 level, and 28 percent at the lowest (0–30) level”.

\(^8\)Indeed there are several online calculators for this task, and it is included as standard in some word processing software. Here we use the implementation given by Rinker (2013) for the R statistical environment.
As suggested by the citations above, this paper joins a literature in social science that makes use of Flesch scores. It is also not the first piece to make the assumption that tools designed originally for measuring ease of reading can be meaningfully applied to texts that were spoken. For example, Jansen (2011) considers the clarity of central bankers answers to questions at legislative hearings. Closer to the subject matter of the current paper, Lim (2008) considers the evolution of rhetoric in Presidential speech-making since the founding of the republic. As a practical matter of course, the most common method by which parliamentary speeches would come to the attention of voters at the time would be via written reports in newspapers.

4.1 FRE Scores for Parliamentary Speeches

Applying the formula implied by Equation (1) to the parliamentary speeches for our study requires some preliminary preprocessing decisions on how to deal with the texts. While we do not stem the documents, or remove stop words, we do convert some parliamentary terms of art that contain period punctuation: thus, ‘Hon.’ becomes ‘Honorable’, ‘Rt.’ becomes ‘Right’, ‘Mr.’ becomes ‘Mister’ and so on. This allows more accurate calculation of the number of sentences in a speech, since it avoids miscounting periods. We then split each speech into sentences using the usual punctuation marks plus semi-colons and vertical bars, which are used in the early periods of our data to break up long utterances. Finally, we strip whitespace (other than single spaces), and drop speeches that contain no alphabetic characters (these typically arise when members give answers as numbers).

Looking over the entire period, the distribution of FRE statistics for our speeches is given in Table 2. We note that mean and median are both around 52, with a standard deviation around 20. We note that the minimum (-301.80) and maximum (205.80) imply a range larger

---

9We also drop some terms with periods, like ‘St.’
Table 2: Summary of FRE statistics for speeches in our data.

<table>
<thead>
<tr>
<th>Minimum</th>
<th>First Quartile</th>
<th>Median</th>
<th>Mean</th>
<th>Third Quartile</th>
<th>Maximum</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>-301.80</td>
<td>42.58</td>
<td>52.24</td>
<td>52.62</td>
<td>62.12</td>
<td>205.80</td>
<td>19.92</td>
</tr>
</tbody>
</table>

Figure 1: Distribution of FRE statistics for parliamentary data: boxplot and histogram. The x-axis for the later is the FRE statistic of the speech. Note that the bulk of the FRE statistics are between 0 and 100.

than in the original Flesch (1948) study, although the boxplot (left) and histogram (right) in Figure 1 suggest that such values are outliers: note that the bulk of the distribution is between 0 and 100.

In Figure 2 we report the (by session) mean speech comprehensibility for cabinet (square points) and non-cabinet (circular points) MPs over the period under study. We also include smoothed loess lines to capture general trends. The main observation is that, somewhere around the 1860s, the average cabinet speech becomes more comprehensible than the average non-cabinet one, whereas prior to that time the means had been very similar. Immediately then, we have some (albeit) crude evidence in favor of our hypothesis above.
Given the length of the period under study, it would be surprising if speeches had not changed in ways other than their comprehensibility. In fact, they became on average shorter: cabinet and non-cabinet speeches had a mean length of around 500 words in the immediate aftermath of the First Reform Act, and were reduced in length in a fairly smooth and consistent fashion over time. By the turn of the 20th century, both cabinet and non-cabinet speeches reach a low of around 150 words (on average). Figure 3 displays these trends clearly. A natural concern might be that any changes to comprehensibility of speeches (measured by the FRE statistic) are an artifact of this shortening. On inspection, we doubt this is the case: the correlation between speech length and reading ease is very weak, at around 0.05. This is true regardless of whether we include ‘outliers’ (as identified in the boxplot of Figure 1) in the calculation. In any case though, we include speech length as a variable in some of our regressions below.
Figure 3: Average length of speeches for cabinet and non-cabinet MPs over time. The $y$-axis is the (mean) number of words per speech.

4.2 Validation

An immediate concern about the use of Flesch scores is that, though they are well validated outside of this application, they do not measure linguistic complexity for our period. In Table 3 we report scores for a set of speeches by William Gladstone, a parliamentary giant of the age. In particular we chose speeches from his period as Prime Minister, in various sessions after the 1868 election, which capture a distribution of complexity. In the table, we can see that speeches of similar length, but varying numbers of syllables, have very different scores. Of course, to modern eyes, all the speeches seem somewhat convoluted and indirect in their manner; while we cannot experiment on 19th century voters to verify this rank ordering, it seems plausible at least that such electors would struggle more (as they read, word-to-word)
<table>
<thead>
<tr>
<th>word count</th>
<th>score</th>
<th>session</th>
<th>speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>29.46</td>
<td>1868 (3)</td>
<td>I deny that the Government ever attempted to throw the responsibility on the officers.</td>
</tr>
<tr>
<td>13</td>
<td>50.47</td>
<td>1868(5)</td>
<td>I have heard nothing of the rumour alluded to by the honourable Baronet.</td>
</tr>
<tr>
<td>15</td>
<td>67.53</td>
<td>1868 (1)</td>
<td>I am glad to think I may have misunderstood my right honourable and gallant Friend.</td>
</tr>
<tr>
<td>14</td>
<td>89.89</td>
<td>1868 (3)</td>
<td>I think it right to state that I shall not consent to the withdrawal.</td>
</tr>
</tbody>
</table>

Table 3: Samples of Gladstone’s speeches, of approximately similar length, with a variety of comprehensibility scores: from the least to the most easy to understand.

with terms like ‘attempted’ and ‘responsibility’ rather than ‘think’ and ‘it’ and ‘right’ and so on.

Another way to assess the validity of our approach is to consider the scores given to contemporaneous texts that are not parliamentary speeches but are aimed at voters in general. To the extent that the relative scores given to those documents are in line with our priors regarding their intended audiences, we have evidence that the metric is a reasonable one. Here, we look at two members serving over approximately the same period: Keir Hardie, a Labour MP (with intermittent service between 1892 and 1915) and Arthur Balfour, a Conservative member (between 1874 and 1922). For Hardie, we use the text of eight books that he wrote between 1905 and 1911 published primarily by the (then) Independent Labour party, and presumably aimed at working class voters. For Balfour, we use (seven) essays for mid-brow magazines and lectures given to various university audiences between 1882 and 1891, presumably aimed at a more middle class audience.\(^{10}\) For each of the works, we calculate the Flesch scores and report their distributions in Figure 4. As can readily be seen, Balfour’s writings with a median score of around 46 are considerably more complex than Hardie’s with a median of approximately 61.\(^{11}\) Parametric and non-parametric tests of means ($p < 0.01$)

\(^{10}\)Bibliographical details of the texts can be found in Appendix A.

\(^{11}\)In passing, we note that Charles Dickens’ fiction of this period has a mean score of around 77, suggesting that the politicians in question were writing and speaking in a considerably more complex way.
Figure 4: Boxplot showing difference between complexity of works by Keir Hardie (Labour party) and Arthur Balfour (Conservative party).

confirm this observed difference.

5 Results

We begin with our estimation of the session-by-session regressions. That is, for each time period in our data, we regress the comprehension scores for the speeches on the cabinet status of the MP in question (a binary variable) with a series of controls—party, competitiveness of constituency and the word count of the utterance itself. Because members make multiple speeches per session, we cluster the standard errors at the MP level. Our estimated coefficients on cabinet are displayed (with 95% confidence intervals) in Figure 5. The solid horizontal line marks zero. Our first observation from the figure is that the point estimates begin below zero, and around the 1860s rise into positive territory and stay there for the remaining time periods in the data. In words, being a cabinet minister is initially associated with making speeches that are (on average) more difficult to comprehend than those of other
members; subsequently, cabinet speeches are easier to understand. Obviously, in many cases, the confidence intervals cross the zero line, but a general pattern is apparent.

Figure 5: Estimated $\hat{\beta}$ on cabinet status in session-by-session linear regressions [with 95% confidence intervals]. Solid horizontal line marks zero. Note the general rise in coefficients from below, to above zero, around 1868.

To clarify the timing of the change, we conducted a set of structural break tests on the data (in the sense of Bai and Perron, 2003).\textsuperscript{12} We began with the estimated $\hat{\beta}$s from Figure 5 themselves each of which was regressed on a constant term. Using standard defaults, the first session of the 1868 parliament emerged as the optimal break. Second, we re-formed the data into a session-by-session time series, with one data point per session. Specifically, each of the observations was composed of the mean comprehensibility of speech in that session, the median cabinet member status (i.e. non-membership) and all controls at their means. Using this explicitly time series representation of the data, the estimated break point is the fourth session of the 1874 parliament.

\textsuperscript{12}See Zeileis et al. (2002) for implementation.
This *prima facie* evidence is helpful, and is in line with our main hypothesis. Nonetheless readers may reasonably object that it is inefficient and possibly misleading to break up the data on a session-by-session basis, especially if subsequent structural break tests ignore the estimation uncertainty in the coefficients—as they do here. A more philosophically appropriate test then is to combine all the sessions and assess the possibility of time-specific effects directly. With that in mind, we now re-estimate the regression with the inclusion of an interaction term involving the product of a member’s cabinet status and a dummy for the first session of 1868—the point at which we hypothesize the change occurred. We do this with and without the controls. The results of those regressions, with standard errors again clustered by MP, are presented in Table 4.

<table>
<thead>
<tr>
<th></th>
<th>Reform Act Interaction</th>
<th>With controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>51.3976***</td>
<td>51.5920***</td>
</tr>
<tr>
<td></td>
<td>(0.2176)</td>
<td>(0.5636)</td>
</tr>
<tr>
<td>Cabinet member</td>
<td>−0.8189*</td>
<td>−0.7802</td>
</tr>
<tr>
<td></td>
<td>(0.4176)</td>
<td>(0.4271)</td>
</tr>
<tr>
<td>Reform Act dummy</td>
<td>0.7183*</td>
<td>0.5183</td>
</tr>
<tr>
<td></td>
<td>(0.3374)</td>
<td>(0.3592)</td>
</tr>
<tr>
<td>Cabinet× Reform Act</td>
<td>5.3039***</td>
<td>5.2229***</td>
</tr>
<tr>
<td></td>
<td>(0.7195)</td>
<td>(0.7172)</td>
</tr>
<tr>
<td>Liberal MP</td>
<td>0.4494</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3788)</td>
<td></td>
</tr>
<tr>
<td>word count</td>
<td>−0.0013***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0001)</td>
<td></td>
</tr>
<tr>
<td>competitiveness</td>
<td>0.0222</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3708)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>670216</td>
<td>670216</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.0084</td>
<td>0.0103</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>0.0084</td>
<td>0.0103</td>
</tr>
</tbody>
</table>

Dependent variable is comprehensibility of speech. MP clustered standard errors in parentheses.

*p < .05; **p < .01; ***p < .001

Table 4: Table of estimates for regression of comprehensibility of speech on cabinet status and time dummy (for Second Reform Act) and interaction between the two—with and without controls.
For completeness, we begin with the version with the controls (second column). Note first that there is no statistical significance attached to the ‘Liberal MP’ variable: that is, the party identification of the member does not seem to be systematically associated with complexity. The same goes for ‘competitiveness’ of constituency, which is a measure of the average number of candidates running in a seat at the general election (and has a mean of around 1.5 for our period). We observe that the coefficient on ‘word count’ (literally, the number of words in the speech) is statistically significant, but negative: that is, longer speeches are (on average) easier to comprehend than shorter ones. Before getting to the main variables of interest, we note that the two models have essentially identical fit statistics: the adjusted-$R^2$ of the restricted model is 0.0084 while adding the extra variables on the right hand side pushes this only to 0.0103. We thus focus on the simpler version—without controls—for interpretation purposes since those extra variables add little to the analysis (implying that our main finding below is reasonably robust).

Putting aside the uncertainty estimates for the moment, we see that the point predictions are as we would expect given our hypothesis. A minister prior to the Second Reform Act has a lower average comprehensibility ($\hat{y} = 50.6$) than one serving after that date ($\hat{y} = 56.6$). Furthermore, a cabinet member after the Second Reform Act is more comprehensible on average than a backbencher either before ($\hat{y} = 51.4$) or after ($\hat{y} = 52.11$) suffrage expansion. Taken as a causal effect, this implies that promotion to cabinet ceteris paribus results in an increase in ease of understanding of around eight percent. This is not a trivial substantive effect and corresponds with a change from around the 48th to the 60th percentile in the score distribution (for speeches made after the 1868 election).

Returning to uncertainty estimates, in Figure 6 we provide an estimated marginal effect plot (taking into account the standard error clustering) and we see that our priors find sup-
port: while ministers—if anything—are slightly less comprehensible relative to backbenchers prior to reform, they are clearly more understandable after.

![Diagram showing estimated marginal coefficient on Cabinet before and after the Second Reform Act.](image)

Figure 6: Estimated marginal effect of cabinet membership on speech comprehensibility, before and after the Second Reform Act.

5.1 Ruling out ‘new types’

Thus far, an implicit assumption for our work has been that the change to cabinet ministers’ utterances was (primarily) a product of individuals responding to new incentives in the electorale. An alternative hypothesis is that, in fact, the Second Reform Act introduced new ‘types’ of individuals to the House of Commons with different latent features and that it the changing make-up of the chamber that yields the results we saw above. There are at least two ways to investigate this possibility, to which we now turn.

First, we consider all individuals who served in a cabinet position at least once after the 1868 election: i.e. after the Second Reform Act took effect. Using a paired $t$-test with ministerial office as the treatment, we compare their mean speech comprehensibility when in the cabinet with their average when serving as a backbencher. The mean difference uncovered is (an increase of) 2.61 on average, which is statistically significant ($p < 0.01$). Since
this test keeps the individuals themselves constant, and combined with the fact that cabinet ministers serving in this period are not disproportionately more likely to have been elected to parliament after 1868 relative to backbenchers they serve alongside,\textsuperscript{13} it provides strong circumstantial evidence that cabinet office (after 1868) had some effect regardless of the fixed characteristics of the MPs involved.

To put this finding on even surer footing, we now turn to a more systematic study of fixed effects. In particular, we regress the session mean comprehensibility score on a session dummy, and then on a session dummy plus fixed effects for the MPs. The idea here is that if the regression with the MP fixed effects has different coefficients for the time dummies, we may conclude that the latent features of individuals are important for explaining the data we saw. In Figure 7 we present a plot of the coefficient on the session dummies for both regressions, with their 95\% confidence intervals. The broken lines represent the intervals for the fixed effects case, and the solid lines are those without member effects. In every case, these intervals overlap: that is, we have no evidence that adding member fixed effects matters relative to the more general time dynamic portrayed above.

\textsuperscript{13}To clarify, we potentially have confounding here if ministers in the post-1868 period are more likely than contemporaneous backbenchers to have been elected in or after the 1868 election. This is not the case: if anything, the reverse is true—67\% of the cabinet ministers had their first session in parliament after the 1868 election, compared with 92\% of those not serving in the cabinet.
5.2 Summary of Findings

This section has various specifications and model tests. A summary of results is as follows:

1. the period of expanding franchise after 1867 is associated with a decline in the complexity of language used by ministers, relative to their speeches prior to that time
2. the period is also associated with a statistically significant decline in ministerial linguistic complexity relative to members of parliament who are not part of the cabinet
3. this finding is not a artifact of ‘new types’ of individuals entering the House of Commons and speaking in ‘new’ ways.

6 Discussion

Observers of modern democracies speak anxiously of the ‘dumbing down’ of political discourse (Lim, 2008). They fear that important yet subtle debates and distinctions are increasingly lost and that elected officials no longer lead opinion in a thoughtful way. A more
optimistic take on recent trends is to regard the simplification of political language as helpful (or perhaps vital) for the engagement of citizens with increasingly constrained time budgets and interests outside of governance (e.g. Temple, 2006). Whatever the truth, as political scientists we have strong reasons to hope and to believe that politicians respond to voters as much as voters respond to their representatives. This is true in both theory (e.g. Meltzer and Richard, 1981) and in empirical work (e.g. Canes-Wrone, Brady and Cogan, 2002) that stresses the importance of congruence between the preference held by constituents and the actions taken by politicians. A natural consequence of this logic is that when new types of citizens join the electorate—in the modern period, typically via immigration (e.g. Tam Cho, 1999)—officeholders will compete for their support and alter their platforms in a way that reflects this underlying change.

Here we studied this very broad phenomenon for an historically important period: the Victorian age of democratization in Britain. Unlike other studies that relied on shorter periods or coarser data at higher levels of aggregation, our findings were unambiguous: cabinet members, after a doubling of the electoral roll in 1867, began to make parliamentary speeches with different properties than before. In particular, entirely in fitting with predictions from the literature, their utterances became easier for the median member of the electorate to understand. This median voter was a man of the working class, with less access to education (and literacy) than had previously been the case in the electorate. As far as we know, our paper is the first to provide systematic evidence of a ‘reform effect’ on the language used in parliament. Crucially, we noted that this change was not due to new types of MPs—with different priorities or experiences—entering the Commons after electoral reform. Indeed, our auxiliary analysis suggested that it is the same members acting in new ways upon finding their way to the Front Bench that is responsible for the decrease in complexity in speeches.
As is inevitable with observational data, it is no easy task to be confident about the causal process that undergirds an empirical pattern. Certainly, our findings are not artifacts of superficial changes to speech records: for example, it is not simply that ministers make shorter speeches over time which is then picked up (artificially) in our complexity metric. Furthermore, we have reason to believe that incentives to simplify presentation are strongest for cabinet members: in Westminster systems they are held accountable for government policy, and their performance—especially on the economy—is the best predictor of future general election success. Unsurprisingly then, it is ministers who most sought to appeal to voters. On the other hand, showing evidence consistent with a theory is not the same as showing that the theory is correct. In particular, we do not know whether ministers consciously altered their linguistic style and what, precisely the impetus for this was: perhaps Prime Ministers such as Disraeli and Gladstone, who seemed acutely aware of the new electoral calculus (see, e.g., McLean, 2001, on Disraeli’s introduction of a new dimension to British politics), took the lead and advised their colleagues to speak more simply (or promoted those from the backbenches that could). Alternatively, the stimulus may be been less direct—perhaps a result of civil service professionalisation and the increasing role of the bureaucracy in serving and advising ministers in terms of their relationship with the House.\footnote{Richard Crossman, for example, argues that ministers dealing with their civil servants find themselves constrained in terms of the plausible policy choices they may chose (and presumably then present to the electorate) (see Crossman, 1975).} Understanding the precise mechanism requires more fine-grained data than we have here, though studying modern speech-writing and speech-giving by politicians may help us understand how they think about the audience to which they must appeal.

Of course, the tone or complexity of speeches is only one part of what it means for parliamentarians to be ‘responsive’ to voters. More important for material welfare is policy. Here, the extent of linguistic complexity is likely less helpful than a study of both topics.
of debate (as in Quinn et al., 2010), and of bills that became acts (relative to those that
didn’t). Again, textual methods can be helpful, and the speeches and related data we have
used provide the beginnings of a resource to get at such quantities of interest. We leave such
efforts for future work.
A Validation: Texts from Hardie and Balfour

The Balfour texts are from Balfour (1893) and are as described in Table 5. The Hardie texts are as described in Table 6.

<table>
<thead>
<tr>
<th>Date</th>
<th>Title</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1887</td>
<td>The Pleasures of Reading</td>
<td>Address at St Andrews University</td>
</tr>
<tr>
<td>1883</td>
<td>Berkeley’s Life and Letters</td>
<td>National Review</td>
</tr>
<tr>
<td>1887</td>
<td>Handel</td>
<td>Edinburgh Review</td>
</tr>
<tr>
<td>1882</td>
<td>Cobden and the Manchester School</td>
<td>Nineteenth Century</td>
</tr>
<tr>
<td>1885</td>
<td>Politics and Political Economy</td>
<td>National Review</td>
</tr>
<tr>
<td>1891</td>
<td>A Fragment on Progress</td>
<td>Address at Glasgow University</td>
</tr>
<tr>
<td>1888</td>
<td>The Religion of Humanity</td>
<td>Address at Church Congress, Manchester</td>
</tr>
</tbody>
</table>

Table 5: Texts by Arthur Balfour, used for validation of FRE statistics.

<table>
<thead>
<tr>
<th>Date</th>
<th>Title</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1909</td>
<td>My Confession of Faith in the Labour Alliance</td>
<td>Published by ILP</td>
</tr>
<tr>
<td>1905</td>
<td>Can a Man be a Christian on a Pound a week</td>
<td>Published by ILP</td>
</tr>
<tr>
<td>1910</td>
<td>Common Good</td>
<td>Published by National Labour Press</td>
</tr>
<tr>
<td>1908</td>
<td>ILP and All About It</td>
<td>Published by ILP</td>
</tr>
<tr>
<td>1909</td>
<td>India: Impressions and Suggestions</td>
<td>Published by Home Rule for India League</td>
</tr>
<tr>
<td>1905</td>
<td>John Bull and His Unemployed</td>
<td>Published by ILP</td>
</tr>
<tr>
<td>1911</td>
<td>Killing No Murder</td>
<td>Published by ILP</td>
</tr>
<tr>
<td>1910</td>
<td>Karl Marx: the man and his message</td>
<td>Published by ILP</td>
</tr>
</tbody>
</table>

Table 6: Texts by Keir Hardie, used for validation of FRE statistics. “ILP” refers to the Independent Labour Party.
References


