The Shadow Cabinet in Westminster Systems
Modeling Opposition Agenda-Setting using House of Commons Speeches, 1832–1915*

Andrew C. Eggers† Arthur Spirling‡

Abstract

We consider the emergence of the Shadow Cabinet, the alternative ‘government-in-waiting’ that characterizes Westminster polities. We use a new data set of over a million speeches uttered between 1832 and 1915 to chart the evolution of this ‘unwritten’ institution. In particular, we introduce a novel and well-validated measure of influence for individual members and provide evidence that the opposition responded to the Second Reform Act and the advent of a ‘party orientated electorate’ by strategically re-organizing in a way that mimicked the Cabinet’s structure. We then show that, in the decades immediately after this suffrage extension, the opposition were increasingly dominated by a smaller and more concentrated group of powerful front-benchers. Further, we demonstrate that the link between being one of these agenda-setting ‘outliers’ and holding cabinet office after the next general election became increasingly strong. We thus provide a metric for the study of agenda-setting by oppositions in parliamentary systems where their procedural rights are weak, while contributing to a substantive debate about the effects of voting reform on political development.

*First version: January 2014. This version: April 8, 2014.
†Assistant Professor, Government Department, London School of Economics. a.c.eggers@lse.ac.uk
‡Associate Professor, Department of Government, Harvard University. aspirling@gov.harvard.edu
1 Introduction

The electoral politics of Westminster systems are characterized as a competition between parties for unfettered access to concentrated executive power via control of the Cabinet (Lijphart, 1999). In such polities, voters are moved less by personal appeals of particular local candidates (though see Cain, Ferejohn and Fiorina, 1987), and more by their partisan affiliation (e.g. Butler and Stokes, 1969; Heath et al., 1991; Clarke et al., 2004). This behavior is typically combined with majoritarian electoral systems that deliver disproportionate government numerical superiority in parliament (Bogdanor and Butler, 1983), along with disciplined backbenchers (Cowley, 2002); thus the leadership of the winning party can expect comparatively long durations in government, and the ability to propose and enact legislation close to its ideal point (Powell, 2000). Given its centrality as motivation for party and individual career efforts in such systems (Kam, 2009), it is unsurprising that the Cabinet has attracted much scholarly attention—in terms of its constitutional role (Bagehot, 1873/2011), origins (Cox, 1987), development (e.g. Jenks, 1903; Alt, 1975; Jenkins, 1996; Adelman, 1997) and modern day operation (e.g. King, 1994; Kam and Indridason, 2005; Berlinski, Dewan and Dowding, 2007; Dewan and Myatt, 2010).

Of course, for the competitive dynamic above to function in practice, citizens must be provided with a prospective ‘government-in-waiting’ as an alternative choice to the incumbent Prime Minister and their Cabinet. However voters assess relative performance (see Lewis-Beck, Nadeau and Blanger, 2004; Lebo and Norpoth, 2006, for discussion), the ‘Shadow Cabinet’—the group of frontbench spokespersons from the Official Opposition—is thus of core importance to Westminster democracies. This is quite apart from other significant roles that the Shadow Cabinet plays: *inter alia*, organizing opposition to the government’s legislative plans in the division lobbies (see Potter, 1965; Brazier, 1999; Dewan and Spirling, 2011);
holding ministers to account in debates (Chester and Bowring, 1962; Franklin and Norton, 1993); and providing a formal link between the parliamentary party and its grassroots.\(^1\) Yet in stark contrast to the Cabinet, and with exceptions (e.g. Lowell, 1908; Turner, 1969; Punnnett, 1973; Johnson, 1997), there has been little work on the opposition as an institution. This is especially true in terms of literature on the origins and development of the Shadow Cabinet.

The purpose of this paper is to speak directly to this issue. Our question is a simple one: when and how did the Shadow Cabinet emerge as a \textit{de facto} force in Commons life? The motivation is three-fold: first, we wish to chart the rise of the opposition as an organized actor with unitary purpose in \textit{legislative} politics such that the literature on British political development—including the work on the export of its governance arrangements (e.g. Rhodes and Weller, 2005; Rhodes, Wanna and Weller, 2009)—is more evenly balanced, with appropriate focus on both government \textit{and} its alternative. Second, we seek to understand how \textit{electoral} forces—such as the massively increasing suffrage that characterizes the Victorian period—affect institutional development. In so doing, we wish to underline the crucial difference between \textit{de jure} and \textit{de facto} power (see Acemoglu and Robinson, 2006) of ‘informal’ institutions (Helmke and Levitsky, 2004) in this area of political science. Third, this is an opportunity to provide methodological tools, tailored to Westminster systems where oppositions are procedurally weak and debate is combative, that will allow other scholars to ask related questions in a satisfying manner.

To understand the latter point of our contribution, note that part of the reason for the dearth of scholarly attention to the origins and evolution of this key institution is a lack of

\(^{1}\)In the case of the British Labour party (see Quinn, 2012) or Canadian Liberal party, leaders seek the endorsement of ‘ordinary’ members in a formal vote.
data. Whatever the informal institutional arrangements in place, the “Shadow Cabinet”—almost by definition—produces less of a ‘paper trail’ of records concerning activities and personnel than its government counterpart.\(^2\) The result is that researchers must make more uncertain inferences about who, exactly, constitutes the body itself and what it is doing. Furthermore, in Westminster systems, the sheer concentration of procedural power in the Cabinet means that ‘usual’ metrics for examining the strength of opposition organization—like ‘roll rates’ (Cox and McCubbins, 2005) or strategic use of committee control (e.g. Krehbiel, 1992) in the US Congress—are not available. Put more succinctly, since oppositions almost *always* lose against governments—in terms of what gets on to the legislative agenda and what becomes law—there is seemingly little variation in legislative output to explain or explore over time. Of course, oppositions are doing other things that are important but do not manifest themselves so obviously, and it is this more latent data that we put to use below.

Given these constraints, our methodological approach is necessarily innovative, and is part of the text-as-data movement in the discipline (see, e.g. Grimmer and Stewart, 2013; Quinn et al., 2010). In particular, we take an explicitly agenda-setting approach to parliamentary speech, of which we have the entirety of the one million utterances between the approximate dates of the First and Fourth Reform Acts (1832–1918). We model these speeches using a measure that considers the ‘burstiness’ (Kleinberg, 2002) of different (government and opposition) actors over time: in particular, we introduce a validated method for scoring individuals via their spoken contributions to debate in the House of Commons. This metric relies on the relative ‘spike’ in activity around particular terms that members of parliament (MPs) use, to measure their latent agenda-setting abilities.

\(^2\)Indeed, even the leader of the ‘Official Opposition’ was not recognized formally by Erskine May until 1937.
Ultimately, we provide theory and evidence to suggest that the 1868 Second Reform Act, and its associated introduction of a “party orientated electorate” (in the sense of Cox, 1987) was crucial for the establishment of a hierarchial opposition leadership, with small numbers of senior individuals increasingly dominating exchanges from the 1870s onwards. More specifically, we show that after 1868, (a) the opposition as a whole was able to wrestle back some noticeable control of the agenda from the cabinet; (b) a small group of opposition individuals emerged who, relative to their co-partisan colleagues, increasingly dominated debates; (c) the relationship between being one of these individuals and taking a role in the ‘next’ cabinet controlled by their party was increasingly strong. These statistical findings are new, and help clear up substantive debates in the field in a way that is data-driven and less speculative than previous literature.

Our paper proceeds as follows: in the next section we briefly review the literature on the emerged of organized opposition in Westminster systems, before setting up the central research debate. In Section 3 we describe our data of over one million speeches, linked to unique MP records. In Section 4 we report on our methodological approach to the estimation problem and show that the measurement strategy is valid, before reporting our results in Section 5. We conclude in Section 6.

2 Literature and Orientation

Not least because it plays a larger role in policy making, and has done for a longer period, the literature surrounding the development of the Cabinet is much more voluminous than its opposition counterpart. In political science, the most widely cited account is that of Cox (1987) (though, see also Bagehot, 1873/2011; Redlich, 1908; Fraser, 1960; Rush, 2001), who argues that the Cabinet as agenda-setter emerged in the 1830s as a attempt to solve a
common resource problem—of too many MPs taking up too much time with self-promoting minutiae—in the aftermath of the Great Reform Act. A puzzle that arises from this accepted assertion is the timing and precise form of the Shadow Cabinet’s emergence as a *de facto* organization. On the one hand, we might expect it to arise fairly immediately, motivated by the sudden threat of institutional dominance by a powerful executive. Certainly, scholars of other Westminster institutional developments—like the advent of (aggressive) parliamentary questions—have made the case that they arose relatively quickly from the need of non-Cabinet members to keep the executive in check (see Chester and Bowring, 1962). Similarly, certain institutional behaviors, such as cohesive division voting against the government’s legislation (Berrington, 1968) and the commensurate use of government whipping to make Cabinet bills into Cabinet acts (Cox, 1992) started not long after the rationalization noted by Cox (1987). On the other hand, historians argue that the notion of a ‘government-in-waiting’ did not emerge until much later: at least until after the Second Reform Act (1868) and the “triumph of partisan politics” Jenkins (1996). We might thus expect a delayed evolution of opposition leadership. Either way, the period between the First and Fourth Reform Acts is crucial, and it is on this period that—like many other scholars of British Political Development—our study is focused.

### 2.1 Two Westminster Data Problems

Resolving this problem of date and mechanism is difficult in the Westminster context for two reasons of data. First, as with much of Westminster constitution-making, formal *de jure* recognition of entities with political power and importance has traditional come much later (if at all) than their *de facto* existence as a force. Thus, an informal notion of a parliamentary opposition that critiqued the government has a long history: it was well underway by the 1720s, with the present day term of the ‘His Majesty’s loyal Opposition’ first appearing in debate in 1826 (Johnson, 1997, 488–490). In contrast, the Leader of the Opposition was not
mentioned in statute until the 1937 *Ministers of the Crown Act*, which *inter alia* guaranteed him a salary. The term ‘Shadow Cabinet’ was used as early as the 1880s, though not with any legal basis, and it initially referred to a set of ex-ministers, now out of office as their party was no longer in government (see Brazier, 1997, Ch 3). Initial meetings were more informal than modern practice (and records of them are scant), but in the post-Second World War period in Britain, opposition parties gave chosen senior MPs specific policy responsibilities and titles with the expectation that they would fulfill a similar ministerial role should their party win a subsequent general election.³

The second data problem stems from the sheer relative power of the Cabinet. Westminster governments are typically single-party (Lijphart, 1999), and face few serious institutional impediments to imposing their will (Powell, 2000). A consequence is that the opposition rarely achieves legislative ‘victories’, and thus one cannot usefully measure outcomes that would be seen in other parliaments. In the case of the US Congress, these would include ‘roll rates’ that involve the minority party defeating the majority, the ability of the minority party to ‘kill’ or substantially alter a bill in committee via negative agenda control (e.g. Cox and McCubbins, 2005), the ability to sponsor legislation successfully (Volden, Wiseman and Wittmer, 2013), or the use of the filibuster to end debate (Wawro and Schickler, 2006). But for the opposition in Westminster, the counts for these quantities would be zero, or near zero. Thus, using roll call votes, committee reports or the success of (private) member bills will not reveal much about our question.

What we do have in the case of the House of Commons is speech data, and if we are to draw conclusions about the relative power and organization of the government and opposi-

tion it is to such information that we must turn. The next section considers the source of our data for this purpose, and the section after considers how might use it systematically.

3 Data

The data we use were obtained from Eggers and Spirling (Forthcoming). The essence is this: we have access to over one million House of Commons speeches uttered between 1832 and 1915. They have been disambiguated in terms of speaker, which in turn has been matched to a unique MP identity. Other information pertaining to these MPs includes their party affiliation in any given parliament, along with their ministerial service record. The speech records are machine-readable, and can be processed using software tools discussed below.

Note that for our purposes, the speeches are organized by ‘parliamentary session’, a period with a mean length of around 200 days. Each session beginning after a general election and though not corresponding to a calendar year, we have approximately one session per annum to consider. We obtained dates for the sessions from the usual sources for the period: Cook and Keith (1975) and Butler and Butler (1994). Thus for any given day, we know the identity of the government and opposition parties, and thus any contemporary MPs. In what follows, we will limit our analysis to MPs running in general elections under either a Conservative or Liberal label (as originally demarcated by Craig (1989), Craig (1974) and Walker (1978)), the two parties who actually held ministerial positions during this time and thus for whom the concept of ‘Shadow Cabinet’ makes most sense.4

4Including other opposition parties makes very little difference to the thrust of the substantive findings below.
4 Methods: Concept and Measurement

As background, consider Members of Parliament (MPs) giving speeches over the course of a session in parliament. Each MP has a set of covariates pertaining to their current role in the government, if they are part of the governing party. Our central concern is understanding which MPs ‘lead’ debate in parliament. Our strategy trades on the idea that influential individuals will raise concerns, terms, topics and issues which MPs will subsequently ‘talk about’ in that debate and ones that follow.

One way to approach this measurement problem is to see speeches in the House of Commons as analogous to a stream of arriving data the contents of which requires modeling. In computer science, a popular way to examine such streams is to consider their ‘burstiness’, in the sense of Kleinberg (2002). The idea is to model the arrival times at which certain words—considered as a type of event—appear. Words that surge in use suddenly are said to “burst” or to be “bursty”, which in practice means that the ‘gaps’ between seeing the word are becoming shorter and shorter. Depending on the nature of the stream process, there are different statistical models that may be fit to the data to determine burstiness.

When data arrives as a continuous process—rather than as, say, batches every year—Kleinberg (2002) suggests an ‘infinite-state model’ in which bursts are state transitions in a hidden Markov process. The Markov process assumes that, when in state \( i \), gap times, \( x \), are exponentially distributed with pdf \( f(x) = \alpha_i e^{-\alpha_i x} \) where \( \alpha_i \), the rate, is proportional to \( s^i \). Larger values of \( \alpha \) imply smaller expected values on the wait time until the next event occurs. The second component of the process, \( \gamma \), is a cost term associated with moving ‘up’ in intensity in terms of the underlying rate—no cost is imposed for the system to move down in intensity. Being concerned with cost implies that the system will have relatively few
transitions. Meanwhile, the exponential component encourages the fitting of a model to the data that reflects the actual sequence of gaps observed. The resulting minimization problem takes both parts into account, and thus attempts to fit the data with as few transitions as possible. Note that the bursts in this model are nested: that is, bursts of higher intensity occur within periods of lower intensity activity.

In principle, both \( s \) and \( \gamma \) could be estimated. In practice, analysts set values for these parameters (in the original presentation, \( s = 2, \gamma = 1 \)). Notice that a larger value of \( s \) implies that changes in gap times will have to be larger in magnitude in order for a ‘burst’ to be said to have occurred. A larger value of \( \gamma \) implies that the burst needs to be sustained for a longer period to ‘count’. \textit{Binder (2012)} implements the relevant model in R (\textit{R Core Team, 2013}) and we use his package in some of what follows.

Conceived in the usual way, burstiness is a property of streams of events—with one example being words in speeches (or indeed, MPs in debates as we’ll explain below). We can, for example, examine the burst pertaining to the word “Ireland” or “reform”. Indeed, for validation purposes that we get to in the following sections, we record the burstiness of every term, for every session. To compare MPs with each other, we need a metric that allows us to compute a score for each member taking into account the relative burstiness of their contributions. We do this as a weighted sum. In particular, for each MP, a burst that \textit{begins} with a speech made by him is scored as the length of that burst multiplied by its intensity. All such bursts are then summed and a total score produced. As an example, consider an MP making 100 speeches. Suppose that a word from one of his speeches launches a burst of intensity level 2 for a time period of 30. A different word from the same speech launches a burst of intensity level 3 for a time period of 4. Meanwhile a word from another speech launches a burst of intensity level 3 for a time period of 5. His overall burstiness is thus
calculated as \((2 \times 30) + (3 \times 4) + (3 \times 5) = 87\). Note that bursts are hierarchical: a burst of level \(n\) can only occur within a burst of level \(m\), where \(m < n\). A consequence is that MPs cannot be given ‘credit’ for decreasing the intensity with which a particular word is used relative to the current period in which they are speaking. Note further, that if one MP boosts a term’s use to, say, level 2, while a second MP then boosts it further to level 3, the first MP receives ‘credit’ only for the level 2 burst, while the second receives credit only for the level 3. This is simply a measurement strategy that accords with our notions of MPs building on the points of others, which we believe requires certain oratory skill and which our metric rewards.

In terms of preprocessing, we do nothing to our texts except remove punctuation and convert everything to lower case. In particular, we do not remove stop words since their use, if they are indeed stop words in the usual sense, should remain relatively constant over time. Nor do we stem the terms, the idea being that we wish to observe particular uses of terms rather than generic concepts that can be spoken in several ways. In Appendix A, we give some pseudo-code to clarify the algorithm we used.

4.1 Validation

The claim is that our burstiness metric captures some notion of ‘agenda setting’ by MPs, and ‘agenda content’ in terms of the words that come up in debate. We now validate our approach by demonstrating that (a) during given periods, the ‘right’ words are bursty; that (b) for given words, the ‘right’ sessions show them to be bursty at that time; and that (c) the ‘right’ individual MPs are bursty at the ‘right’ times. By “right” in the foregoing sentence, we mean ‘in ways that are congruent with our expectations and knowledge of the period’. Beginning with our first validation exercise, consider Table 1. Here, we report three particular sessions—in 1846, 1866 and 1885—and terms that appeared near the top of the burstiness rank order for those periods. We see immediately from the first column that
<table>
<thead>
<tr>
<th>session</th>
<th>terms (rank)</th>
<th></th>
<th>terms (rank)</th>
<th></th>
<th>terms (rank)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1846 (1841, 6)</td>
<td>1866 (1865, 1)</td>
<td>1885 (1885, 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>agriculturists (1)</td>
<td>suffrage (4)</td>
<td>irishmen (2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>wheat (3)</td>
<td>franchise (5)</td>
<td>1782 (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>grain (5)</td>
<td>1832 (7)</td>
<td>kingharmon (6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>farmer (6)</td>
<td>redistribution (10)</td>
<td>parnell (15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>prices (7)</td>
<td>seats (11)</td>
<td>tenant (18)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Very bursty (highly ranked) terms from various sessions in the 19th Century. Note that the columns refer to the periods pertaining to the Corn Laws, the Second Reform Act and the Government of Ireland Bill, respectively.

MPs were discussing (in a bursty way) ‘wheat’ (ranked 3) and ‘grain’ (ranked 5), during a period when the Corn Laws were under serious discussion. Similarly, just prior to the Second Reform Act of 1867, they raised issues pertaining to the ‘franchise’ and ‘seats’, and also seemingly alluded to the (Great) Reform Act of ‘1832’. Finally, in 1885, the time of the controversial Government of Ireland Bill that would have delivered Home Rule to Ireland, we see discussion of ‘irishmen’ and their leader ‘[P]arnell’, while mentions of ‘1782’ presumably pertain to the previous Anglo-Irish constitutional settlement by which the Irish parliament was at the time allowed a version of self-governance. The term ‘kingharmon’ refers to Edward King-Harmon MP, an Irish Nationalist turned Unionist, then active in British politics.

It is worth contrasting the exercise that produced Table 1 relative to the use of ‘topic models’ in political science (Quinn et al., 2010; Grimmer and Stewart, 2013). In our approach, terms are rewarded if they ‘suddenly’ appear with relative intensity; in this way, a specific term used consistently in every session such as ‘budget’ or ‘trade’ or ‘education’, would not necessarily be bursty. By contrast, a topic model would almost certainly have a topic allocated to, or defined by, such concepts. That is, topic models do a good job of summarizing ‘what’ was discussed in some general way, while burstiness captures dynamics in which terms were
intensely discussed and that dominated the agenda for spurts of time.

Moving on with our validation, we want to see that certain terms are bursty when we expect them to be. Consider Figure 1. There we report four terms with distinct burstiness ‘signatures’ over time. In each case, the y-axis is the burstiness metric outlined above, then rescaled, or standardized, between 0 and 1 within a given session. Thus, as terms approach a burstiness value of ‘one’ they are the most bursty term that session, the second most bursty term would typically have a score of just shy of one (e.g. 0.98), the third most term just below that and so on. The x-axis labels correspond to the beginnings of the various parliaments (generally following general elections) over the period. The [pink] dots are the transformed scores per session, and the solid [red] lines are lowess curves. In the first panel, we consider the term ‘tariff’, which was used repeatedly and intensively in two different periods: first, during the ‘corn laws’ debates of the 1840s, and then at the start of the twentieth century, when Joseph Chamberlain in particular argued for a system of ‘imperial preference’ for Empire goods (see, e.g, Howe, 1998, for discussion). The term ‘zulu’ appears high on the parliamentary agenda in the early and mid-1880s—during the exact period that the British were at war with this group—and then disappears. The word ‘ireland’ is bursty throughout the entire Victorian era, and this seems entirely reasonable given that the ‘Irish Question’, and Irish MPs, were a constant concern during this time. Finally, in the last panel, we note that the word ‘gentlemen is similarly constantly on the agenda though its burstiness is very low—implying that its use is not especially intense. This makes sense for a generally procedural word that is used fairly consistently over time.

As the third part of our validation exercise, we considered the burstiness profile of members of parliament themselves. In Figure 2 reports some results for validation. The first two panels are for Benjamin Disraeli and William Gladstone, who both served as Prime Minister
Figure 1: The burstiness profile of different terms over time. The $y$-axis of the plots is the ‘standardized’ burstiness of the term, a rescaled metric where a value of 1 corresponds to the most bursty term that session, while a value of 0 refers to the least bursty term. The $x$-axis labels correspond to the beginnings of the various parliaments over the period. The [pink] dots are the actual standardized scores, and the solid [red] lines are lowess curves.
at various times. It’s worth noting that Gladstone was (in relative) terms generally more bursty than Disraeli, and spent time in the late 1860s and 1880s as the most bursty member of the Commons. Moving to the third panel, Charles Parnell was an Irish MP serving as the leader of the Nationalist contingent in the 1880s, actively involved with obstructionism of the time: the graphic reflects his contributions. Finally, in the fourth panel we chart the career of Samuel Plimsoll, MP for Derby. Otherwise unremarkable, this backbencher relentlessly pressured the Disraeli government of the 1870s to pass legislation that would introduce legal ‘waterlines’ on merchant ships. He left parliament soon after success was achieved on this front. Again, our graphic reflects our priors: showing Plimsoll’s to have marked though small burstiness after the 1868 election.

5 Results

We begin by considering the agenda-setting ability of the opposition, and the way that this changes over time. Of course, our metric above is ‘absolute’: it calculates a raw number pertaining to individuals, or groups of individuals, and their ability to raise issues which draw attention in parliament. In practice, this means that burstiness may be generally higher under two conditions: first, when (exogenously) there are more things to be bursty about—e.g. a war occurs, or a famine, or some other event of note; second, when there are more opportunities to talk, since this lengthens the period (in speech terms) when bursts may come to exist. Given these facts, we consider the burstiness of the opposition relative to the cabinet. In particular, we begin this section by taking the ratio of mean opposition burstiness to mean cabinet burstiness for every session in our data.
Figure 2: The burstiness profile of different MPs over time. The $y$-axis of the plots is the ‘standardized’ burstiness of the term, a rescaled metric where a value of 1 corresponds to the most bursty MP that session, while a value of 0 refers to the least bursty MP. The $x$-axis labels correspond to the beginnings of the various parliaments over the period.
5.1 Opposition Burstiness over Time

In the upper panel of Figure 3 we plot that quantity: it appears as the [black] undulating line, that peaks and troughs, moving left to right, reaching its zenith around 1857 (when the cabinet was about 40 times more bursty), and its nadir around 1885 (when the cabinet was about 5 times more bursty). Note that for clarity, we demarcate the x-axis using general election dates for the period.

The first observation from the upper panel of Figure 3 is that the cabinet was always more bursty than the opposition, on average: notice that the line is never below one. Given the dominance of the cabinet over procedure from the 1830s onwards, this is not per se surprising: ministers have more opportunities to be bursty, and presumably by the very nature of their jobs have more ready access to information that can become bursty (e.g. reports of officials figures or policies). However, moving left to right, we see a generally decreasing ratio: the smooth [red] lowess line makes the point very clear. Put otherwise, throughout this entire period, the opposition is increasingly bursty, relative to the cabinet. To place this result on a more sound statistical footing, we conducted structural break tests (in the sense of Bai and Perron, 2003).\footnote{5See Zeileis et al. (2002) for implementation.} We found one break in the ratio data, dating to the first session of the parliament beginning after the 1874 General Election: in the figure, we present this point as a broken line and note that the mean ratio dropped by over 50%, from 16.91 to 7.25 after the change point.

An obvious concern on seeing such a result is that there is nothing ‘special’ about the opposition: perhaps the cabinet’s agenda-setting ability was in secular decline from the 1870s onwards? We can go some way to refuting this suggestion by studying the lower panel of Figure 3, where we consider the (mean) ratio of the cabinet to government backbenchers.
Figure 3: Ratio of (mean) burstiness: cabinet to opposition, cabinet to (Government party) backbenchers. One change point found in the opposition ratio time series, marked on the plot with the broken line and mean of ratio given on either side.
Notice that both the underlying ratio, and the smoothed lowess, are essentially constant. We find no breakpoints here using the usual formal tests. Ultimately then, we can conclude that the change in the ratio for the opposition is something specific to that side of the House of Commons, and not a general artefact of changing cabinet roles or priorities at the time.

5.2 Opposition Outliers as a ‘Front Bench’

Having established that the opposition was increasingly aggressive in its agenda-setting just after the Second Reform Act, we next seek the precise mechanics of that change. That is, we wish to understand exactly how the opposition asserted its control. Recall that one possibility is that it increasingly mimicked the government party’s authority structure by establishing an ‘executive’ core of frontbenchers to set policy and rebuff the cabinet, while a pliant majority of opposition backbenchers formed up behind them. In Figure 4 we examine the evidence for such a claim.

In the upper portion of Figure 4, we report boxplots of the burstiness of opposition parties (specifically, the Conservatives and Liberals) over time. The points (circles and squares) denote outliers, defined in the usual way as points above (and below) 1.5 times the interquartile range of the given session. Note immediately that, in practice, all outliers are in the right tails of their distributions: that is, the median opposition member has a very low burstiness for the entire period (and, indeed, it is close to zero on this measure). Second, we see a surge in the magnitude of the outliers around 1880: indeed, some of the largest burstiness scores are recorded between 1880 and 1892. Switching to formal time-series tests on the means of each session, we note that there is one break point, demarcated by a broken line during the third session of the parliament meeting in 1880. Looking at the standard deviation of the burstiness yields an almost identical finding, albeit the change point corresponds to the second session of 1880. Finally, we report the changing means and standard deviations
Figure 4: Concentration of agenda-setting power in the opposition over time. Top panel shows changing distribution of burstiness for the opposition; middle panel shows the same relationship but using standardized data, by session; lower panel shows (declining) number of outliers over time—consistent with the emergence of a ‘shadow cabinet’.
themselves: prior to the break, we have a mean of 211688.28, while the standard deviation is around seven million. By contrast, the latter part of the time series has a mean and standard deviation an order of magnitude higher. Thus we conclude that the ‘average’ burstiness of the opposition was increasing, while simultaneously showing more variance. Given that the floor value of the metric is zero, the implication of the top panel is that some individuals are increasingly ‘pulling away’ from average members.

To make this point clearer, consider the middle panel where we have ‘standardized’ the measure by session, meaning all MPs fall between zero and a burstiness of one. The pattern we note above holds more starkly: there are always outliers, but they are obviously fewer by the turn of the 2th century than in the 1830s. Moreover, the outliers are becoming generally more similar to each other, and further from the respective medians.

To test our intuitions more precisely, the bottom panel of Figure 4 reports the number of opposition outliers over time. Clearly, there is a downward trend: beginning around 70 outliers, the opposition has around 50 outliers by the 1870s, and less than 20 by the end of the period. Again, we use a formal structural break test which in this case revealed two breaks: one in the last session of the 1865 parliament, and the second in the first session of the parliament meeting after the 1886 election. In both cases, the mean is reduced. Importantly for our purposes, the average number of outliers is reduced to the approximate size—below 20—that we would expect for a ‘shadow cabinet’ of spokesmen on various issues of governance. To reiterate: here we find that the date of the Second Reform Act (1868) was a crucial transition point for the emergence of a small(er) set of bursty individuals on the opposition benches, congruent with the existence of a cadre of senior MPs in leadership roles.
5.3 Burstiness and Future Cabinet status

One way to verify our presumption—that the outliers from Figure 4 are a ‘cabinet-in-waiting’—is to show that, in fact, they went on to fill cabinet roles when their party found itself next in government. To examine this possibility, we considered the 14 times that power switched, in the sense that a new party previously in opposition now formed the government, during the period. For the opposition members in each ‘switching’ session, we pooled the data and regressed their (binary) status as a cabinet member in the next session on their (binary) status as an outlier in the previous period, along with a time indicator, and burstiness as a robustness check. We make no claims that our efforts here are causally identified: there are surely many reasons why MPs do or do not get promoted to ministerial office when their leader becomes Prime Minister. However, such an analysis can establish whether or not the evidence is broadly consistent with our claims.

The relevant part of our results can be seen in Table 2. In Model 1, we use outlier status and ‘session number’ since the Great Reform Act—literally, the number of sessions of parliament that have occurred since 1832 (thus the first session is our data is given the value ‘1’, the second is ‘2’ and so on). We see a positive effect of both variables: that is, being an outlier helps get you promoted next time your party is in office, and, in fact, as time passes one is unconditionally more likely to be promoted. In Model 2 we add our key interaction term between time and outlier status. As expected and consistent with our claims, the coefficient on being an outlier remains positive and significant. The coefficient on session number is similarly positive and significant, and is larger in this specification. Importantly the interaction effect is significant, and smaller than the combined effect of being an outlier and the session number.\(^6\) Thus the net effect of being an outlier is that one was more likely

\(^6\)That is, when we consider $\Pr(y = 1) = \frac{1}{1 + \exp(-\beta X)}$ we see that it is increasing as years pass for an outlier: the predicted probability for the first switching session is around 0.1, while for the last switch (in 1910) the predicted probability is around 0.2.
to be promoted to office as time passed. Notice that this model has a smaller AIC than the previous effort, suggesting it is a better fit to the data. Moreover, a likelihood ratio test favors the model with the interaction.

Model 3 considers the robustness of this result by including a control for raw burstiness.\(^7\) As we see, the coefficient on outlier status is essentially unchanged, though a little extra explanatory power is added (AIC is lower). Finally, we consider Model 4 that does not use our outlier status variable at all, and relies solely upon an MPs burstiness, in addition to the time variable and the interaction. Note that this model does a relatively poorer job (in terms of fit) than the variant using the outlier metric we explained above.

## 5.4 Summary

Our results show that three interrelated facts hold for the period under study:

1. though the cabinet was always more bursty than the opposition, their *relative* ability to set the legislative agenda changed markedly around the time of the Second Reform Act (1868). After this, the opposition was generally stronger as an agenda-setting force.

2. within the opposition, ‘outliers’—that is, extremely bursty individuals—became fewer in number over time, with marked shifts downwards at the time of the Second Reform Act, and in the mid-1880s. By the turn of the twentieth century, a group of individuals approximately the size of a ‘shadow cabinet’ (in terms of shadowing the major offices of state) had emerged.

3. the relationship between setting the legislative agenda with opposition speech, and

\(^7\)Outlier status is obviously post-treatment with respect to raw burstiness, and so one should not be overly confident in interpreting the coefficients.
<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-5.1984***</td>
<td>-6.1150***</td>
<td>-6.0979***</td>
<td>-3.7618***</td>
</tr>
<tr>
<td></td>
<td>(0.2853)</td>
<td>(0.5105)</td>
<td>(0.5091)</td>
<td>(0.2262)</td>
</tr>
<tr>
<td>outlier</td>
<td>2.6767***</td>
<td>3.9185***</td>
<td>3.9831***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.2131)</td>
<td>(0.5671)</td>
<td>(0.5660)</td>
<td></td>
</tr>
<tr>
<td>session</td>
<td>0.0184***</td>
<td>0.0364***</td>
<td>0.0359***</td>
<td>0.0074</td>
</tr>
<tr>
<td></td>
<td>(0.0048)</td>
<td>(0.0088)</td>
<td>(0.0087)</td>
<td>(0.0049)</td>
</tr>
<tr>
<td>outlier × session</td>
<td>-0.0265*</td>
<td>-0.0344**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0105)</td>
<td>(0.0108)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>burstiness</td>
<td></td>
<td></td>
<td>0.0000***</td>
<td>0.0000***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0000)</td>
<td>(0.0000)</td>
</tr>
<tr>
<td>burstiness × session</td>
<td></td>
<td></td>
<td>-0.0000†</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0000)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>AIC</th>
<th>log L</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3344</td>
<td>884.8504</td>
<td>-430.4252</td>
</tr>
<tr>
<td></td>
<td>3344</td>
<td>880.3381</td>
<td>-424.1690</td>
</tr>
<tr>
<td></td>
<td>3344</td>
<td>870.2763</td>
<td>-415.1382</td>
</tr>
<tr>
<td></td>
<td>3344</td>
<td>1005.1339</td>
<td>-486.5669</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
† significant at $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$

Table 2: Coefficients [standard errors] for logistic regression of cabinet membership on outlier status, burstiness, years past since Great Reform Act and interaction terms.

being in the cabinet when the party in question took power was positive, and became increasingly strong over time (in the sense that an outlier in later sessions was more likely to find himself promoted to cabinet office than an outlier in earlier sessions)

6 Discussion

It is axiomatic that representative democracy requires citizens have an alternative to the present government at election time—even if this does not ultimately mean that the people’s will is implemented as policy (Schumpeter, 1942). In the case of Westminster systems,
this takes a special form: the Shadow Cabinet. Here, we attempted to provide theory and
evidence on the development of this crucial institution that matters for the normative jus-
tification for, and day-to-day functioning of, many of the world’s democracies. Using new
text-as-data methods, and over a million speeches by members of parliament between the
First and Fourth Reform Acts, we showed that the 1870s were a vital period of transition
after which the opposition front-bench organized itself as a ‘government-in-waiting’. In par-
ticular, we showed that over time an increasingly small group of opposition ‘leaders’ closed
the gap in terms of agenda-setting with their partisan competition in the Cabinet. Intrigu-
ingly, though the Cabinet began its characteristic dominance of procedure in the 1830s (Cox,
1987), it was not until suffrage was broadened and the electorate became more sensitive to
party voting at the ballot box that the Shadow Cabinet was given sufficient impetus to
emerge as an institutional force. Our work joins a large literature on the effects of suffrage
expansion on political behavior and policy making at Westminster (e.g. Gash, 1952; Adel-
man, 1997; Rush, 2001; McLean, 2001; Aidt and Dutta, 2010; Berlinski and Dewan, 2011),
and by moving the focus to the opposition similarly contributes to the study of comparative
parliamentary politics (e.g. Doring, 1995; Holzhacker, 2005).

Our work has several broader implications. First, the Shadow Cabinet was initially an
“informal institution” in the sense of Helmke and Levitsky (2004): its particular structure
and personnel was in no way ‘officially’ enforced or encouraged. A consequence that we were
required to overcome above, is that the ‘usual’ sources of data were not available. To the
extent that we care about modeling such things for other examples of informal institutions,
our text-as-data work here will be helpful for others. Second, and related, we demonstrated
an important case in which an (informal) institution arose ‘organically’ as a counterpoint to
a pre-existing organization—the Cabinet—when an external stimulus was presented (in our
case, a party orientated electorate). Our work thus joins a literature that deals with ‘insti-
utionalism' (see Hall and Taylor, 1996), and the specific mechanisms by which institutions evolve (see Mahoney and Thelen, 2009). Again, we think our measurement strategy is a way to proceed when faced with the task of charting the development of such organizations over time. Third, we took an explicitly ‘agenda-setting’ approach—a topic of very general interest to political scientists than ‘agenda-setting’ (e.g. Cobb, Ross and Ross, 1976; Pollack, 1997; Krehbiel, 1998; Cox and McCubbins, 2005). Typically measuring the extent to which bodies or individuals have the power to do so is difficult—especially in parliamentary systems where, in day-to-day operations, oppositions lose and governments win. We have gone part way to resolving that issue.

This paper raises several interesting questions that we have left unanswered. First, though we have shown that being a part of the Shadow Cabinet was increasingly important for holding ministerial office after a general election, we have not looked at the criteria by which members were selected into the Shadow Cabinet in the first place. Scholars of Westminster systems have devoted a substantial amount of attention to ‘career paths’ of MPs in modern times (e.g. Benedetto and Hix, 2007; Kam, 2009) and doing the same for the foundational Victorian period would be interesting. Second, our technique allows for helpful (weighted) word-based summaries of debates. Our focus here was on the relative burstiness of sets of individuals, but it would presumably be beneficial to those interested in ideological changes in Westminster legislatures over time (e.g. Schonhardt-Bailey, 2003; Godbout and Hoyland, 2013) to use a metric like ours to get a sense of exactly how—i.e. on what issues—MPs became divided or unified as their parties evolved. Finally, with the speech records of other legislatures—such as the US Congress (Jensen et al., 2012)—increasingly available online, it would be intriguing to compare the burstiness of terms in a comparative context: this would allow us, for example, to understand how different nations prioritized different matters at different times and thus garner insights into the national (and particular) vs international
(and common) currents of events that inform policy making across the world. We leave such efforts for future work.
A  Pseudo-code for burstiness calculation

Let $tdm$ by the term-document matrix of the speeches, such that each row is a word, and each column a speech. A given $i^{th}, j^{th}$ cell-entry of $tdm$ is a binary indicator $\{0,1\}$ of whether or not word $i$ appeared in speech $j$ (multiple uses are treated similarly to single occurrences). The steps to calculate our statistics are as follows:

```plaintext
for (i in 1:number of rows in tdm){

1. draw the $i$th row of $tdm$, which is a binary vector of occurrences. Thus, supposing there were 10 speeches in the corpus, we might have $\{0,0,1,1,1,0,0,0,0,0\}$, with the use of some term appearing in documents 3,4,5,6 and then nowhere else.

2. calculate the burstiness of this term, as described above. That is, for each individual burst, multiply its level by its duration. Then sum these terms for all bursts that occur for the term. Denote this sum as $b$.

3. allocate $b$ to the appropriate location in a document vector (that is, a vector of length equal to the number of documents in the corpus). Notice that this will require simply adding it to whatever the ‘running total’ for that document currently is (since a given document may have multiple bursty terms).

4. record the time of the start of the maximum or ‘peak’ burst of the $i$th term, and the time of the end of that burst. Using a look-up table, record the MP making the speech that began the peak burst.

}
```

The result of this algorithm is (a) a table of bursty terms (i.e. all terms with non-zero burstiness), each with a starting and ending point of their peak burst, and the identity of the MP who began the peak burst; (b) a table listing every speech and the burstiness of each
(which will generally be zero for at least some speeches). Finally, a look-up table is used to aggregate the results of the speech table by MP: that is, each speech is mapped to a unique MP, and his score derived by summing the total burstiness of all the speeches he gave (some of which may be zero scored).
References


