Informal Institutions, Latent Variables and Political Methodology

Studying the Emergence of the Shadow Cabinet in Westminster Systems[*]

Andrew C. Eggers[†]   Arthur Spirling[‡]

Abstract

We consider the modeling of an 'informal institution' vital to the development and functioning of Westminster polities: the notion that the Shadow Cabinet is a 'government-in-waiting'. We clarify why such unwritten rules are hard to explore with usual approaches from political methodology by linking standard conceptions of such institutions to issues of latent variables in directed acyclic graphs. We provide a method for identifying Shadow Cabinet members using the surges in term use that begin with their speeches. Substantively, we show 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, and that the link between being part of this agenda-setting opposition leadership and holding cabinet office after the next general election became increasingly strong.

Word Count: 8447, excluding Abstract and SI A, B and C.

[†]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

Informal institutions are "socially shared rules, usually unwritten, that are created, communicated, and enforced outside of officially sanctioned channels" (Helmke and Levitsky, 2004, 727; see also Lauth, 2000). Our interest here is in the evolution of one particular informal institution, the study of which yields general methodological lessons for scholars of comparative politics: the rule that in Westminster systems, the 'Shadow Cabinet'—the group of frontbench spokespersons from the Official Opposition—forms the executive when the party currently in opposition next enters government. This relationship is at the core of Westminster democracy for reasons that are obvious from any textbook account of those systems (e.g. Lijphart, 1999). Yet it has never been part of statute law, and was not always the case in practice: in the 18th and 19th century, leaders of governments (Prime Ministers) were implicitly or explicitly selected directly by the Crown (Marriott, 1925), and later via some decision-making process within the majority party (see, e.g, Bagehot, 1873/2011; Jenkins, 1996). Prior to modern times, the presence of competing formal and informal institutions meant that conflicts over exactly which set of rules and actors had precedence was common (see, e.g., Erskine May, 1864/1986, on the 'bedchamber crisis').

In modern Westminster systems, where partisan voting is the norm (e.g. Butler and Stokes, 1969; Heath et al., 1991; Clarke et al., 2004)[1] and majoritarian electoral systems deliver disproportionate government numerical superiority in parliament (Bogdanor and Butler, 1983), along with disciplined backbenchers (Cowley, 2002), 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). Thus, the identity of the 'government-in-waiting', and the fact that it will become the executive once in office, has profound implica-

---

[1](though see Cain, Ferejohn and Fiorina, 1987)

tions for almost all actors in the system, including voters and legislators. 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.[2] Yet in stark contrast to the Cabinet (e.g. Alt, 1975; Cox, 1987; King, 1994; Jenkins, 1996; Jenks, 1903; Kam and Indridason, 2005; Berlinski, Dewan and Dowding, 2007; Dewan and Myatt, 2010), and with exceptions (e.g. Lowell, 1908; Turner, 1969; Punnett, 1973; Johnson, 1997), there has been little work on the opposition *per se*. This is especially true in terms of literature on the origins and development of the Shadow Cabinet and the informal institution to which it is vital.

If we can chart the changing nature of the informal institution linking opposition to government, we can simultaneously answer other questions of interest pertaining to the particular mechanics by which this institution emerged. In particular, we seek to understand how *electoral* forces—such as the massively increasing suffrage that characterizes the Victorian period—affect political development. A helpful spill-over of charting the rise of the 'government-in-waiting' rule from non-existent to 'almost formal', in *legislative* politics is 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)—becomes more evenly balanced, with appropriate focus on both government and its alternative.

Although of great potential substantive interest, executing large-n studies of informal in-

---

[2]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.

stitutions is extremely difficult (although see e.g. Desposato, 2006; Stokes, 2006), not least because, almost by definition, they leave less of a 'paper trail' of official documentation. In the specific case of the Shadow Cabinet, only in very recent times has its membership or activities been recorded for outside observers.[3] The result is that researchers must make more uncertain inferences about who, exactly, constitutes the body itself and what it is doing. This problem is compounded in Westminster systems by the fact that the opposition *per se* is procedurally weak and hard to observe 'in action': 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 either very consistently zero or simply non-existent. 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.[4] Consequently studying the opposition and its role in informal institutions is extremely challenging.

Despite the importance of "identifying and measuring" (Helmke and Levitsky, 2004, 733) informal institutions, political methodologists have been somewhat reticent about applying their expertise to the problem. Perhaps part of the disconnection is that a unified approach— using terminology that both 'sides' can grasp in a productive way—has been lacking. We attempt to improve matters below by describing informal institutions in a way familiar to methodologists: in particular, via directed acyclic graphs (e.g. Blackwell, 2013) in which an observed variable for a given unit (speech contents) may be used to make inferences about an unobserved one (Shadow Cabinet management) and its relationship with an outcome of

---

[3]Indeed, even the leader of the 'Official Opposition' was not recognized formally by Erskine May—the parliamentary procedure guide used in Britain—until 1937.

[4]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.

interest (Cabinet membership). As suggested by this strategy, a second contribution below is to provide a text-as-data measurement strategy (see, e.g. Quinn et al., 2010; Grimmer and Stewart, 2013) using the one million utterances between the approximate dates of the First and Fourth Reform Acts (1832–1918) in which the relevant informal institution first emerged and then swiftly evolved. We model these speeches using a measure that considers the 'burstiness' (Kleinberg, 2002) of different (government and opposition) actors over time: specifically, 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, in order to measure members' latent agenda-setting abilities.

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. More importantly though, our approach demonstrates how political methodologists and scholars of informal institutions may work together, using techniques from the former to test theories from the latter.

2   Shadow Cabinet: Substantive Background

Not least because it plays a larger role in policy making, and has done for a longer period, the the Cabinet has attracted much more scholarly attention 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 the informal institution of '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 focussed.

As with much of Westminster constitution-making, formal *de jure* recognition of entities

with political power and importance has traditionally come much later (if at all) than their *de facto* existence as a force. Thus, the informal practice by which a parliamentary opposition critiques 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 of the Shadow Cabinet 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.[5]

As alluded to above, oppositions—and Shadow Cabinets—are weak in procedural terms: 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, such as 'roll rates', successful legislation sponsoring (Volden, Wiseman and Wittmer, 2013) or negative agenda control (e.g. Cox and McCubbins, 2005; Wawro and Schickler, 2006). Since these measures take a value of (near) zero at Westminster, they cannot tell us much about who is organizing opposition to the government. Yet this is a key element of exploring the particular informal institution of interest here. What we do have is speeches, and we return to their use below once we have

---

[5]See Crisp (1983) and Bateman (2009) for similar discussion the Australian case, and Power (1966) for Canada.

given a more formal account of the general problem that informal institutions present for political methodologists.

3   Formalizing Informal Institutions

The Helmke and Levitsky (2004) definition of informal institutions with which the paper opened is helpful for purposes of intuition, but thinking about the methodological implications of this concept requires a more technical treatment. We note first that a very simple way to think about 'rules', formal or informal, is as a relationship between one set of variables and another. A standard way to do this in statistics and related disciplines, is to use *directed acyclic graphs* (DAGs), which link variables (called 'vertices') with arrows ('edges') denoting probabilistic relationships between them. If we write $X \rightarrow Y$ using this framework, we are asserting that a probabilistic relationship exists between X and Y: they are, by our graph, not statistically independent.[6]

It is not difficult to come up with examples of such graphs that fit our understanding of rules: thus, if X is a variable coding for paying a bribe and Y was a variable coding for receiving prompt service from a bureaucrat, $X \rightarrow Y$ might describe a key element of a 'graft economy'. Of course, bribery for service is typically an informal institution, but this framework can also incorporate more formal rules. For example, abstracting away from various restrictions imposed by states in practice, the standard interpretation of the 14th Amendment of the US Constitution is that citizens may vote. Denoting one's citizenship status as X, and one's right to vote as Y, we might write the relationship again as $X \rightarrow Y$. In the special case of *de jure* laws the relationship is, in fact, deterministic (i.e. the relevant proba-

---

[6]In the Helmke and Levitsky (2004) presentation, there is much discussion of more nuanced elements of informal institutions—such as enforcement mechanisms—that we do not get into here: instead, we concentrate on the core concepts, and provide a basic framework for modeling.

bility distribution is degenerate) insofar as an 'if-then' relation exists between the variables: *if* a person is a citizen, *then* they may vote. Notice that even if we incorporate notions of *de facto* suffrage restriction via informal institutions—e.g. Jim Crow codes of behavior that restricted black voting well into the 20th Century—we still preserve the DAG framework described above.[7]

In the case of voting rights, we can imagine that X and Y are directly observable: the former might be measured from a passport or a birth certificate, while the latter can be observed in legislation. In the case of bribery, X, or X and Y, may not be directly observable (or at least not easily so). That is, as analysts, we may not be able to record 'values' for our X and Y variables from observational studies of human behavior. Seen in this light, the concern that Helmke and Levitsky (2004, 733) have for "identifying and measuring informal institutions" is thrown into stark relief. Even if we can observe Y, if we cannot observe X (the paying of a bribe), we will struggle both to 'identify and measure' the institution at hand, where these operations may be taken to literally mean uncovering the parameters of the probabilistic relationship that connects X and Y. In the language of DAGs, we say that X is *latent*, and we can denote the relationship between X and Y as a broken line: $X \dashrightarrow Y$.

Typically, when faced with a latent variable that is a important part of some data generating process as X is here, political methodologists attempt to infer its values from other, observed variables. To keep matters simple, suppose that there is one such observed variable, denoted Z, and that for the units in the study, their (latent) value of X determines their (observed) value of Z, but that Z does not have any direct relationship with Y itself. In

---

[7]Pushing further, there may be utility in simply redefining formal institutions as ones in which the relationship between X and Y is deterministic, while informal institutions are ones in which the relationship is probabilistic (which includes deterministic ones as a sub-category), but this is not required for our current presentation.

DAG terms, such a set of relations is compatible with multiple graphs, but a natural one that captures our intuitions might be $Z \dashleftarrow X \dashrightarrow Y$. If we can construct a measure for X, via our observed values of Z, we may be able to identify and measure the informal institution that links X and Y via regression or some other technique. Furthermore, with time series data on Z and Y, we may be able to comment on when that institution emerged, and how it changed over time, a task that Helmke and Levitsky (2004) assert as a key area of future research in comparative politics.

In what follows, we will take precisely the route implied by our comments about X, Y and Z. For us, Y is the status of an MP as part of the Cabinet in a particular time period. This is observable. Meanwhile, X is his status (or not) as a member of the Shadow Cabinet, something that occurs prior to Y. For our period, this is latent, and cannot be directly observed. Below, Z will be an MP's agenda-setting ability, derived from observational data on his speeches via a particular metric that we will define in some detail. Thus, an MP's speeches will help us infer whether he was a member of the Shadow Cabinet or not (along with other details on the Shadow Cabinet's evolution), and we will then study the relationship between this Shadow status and promotion to the Cabinet once his party is in power. In this way, we can assess the changing nature of an informal institution vital to the functioning of Westminster democracy.

Before moving ahead, notice that nothing in the above has alluded to causation *per se*: everything is couched in terms of probabilistic relationships. With some extra assumptions—in particular, 'Causal Markov-ness', 'Faithfulness' and 'Causal Sufficiency' (see Eberhardt, 2013, for discussion)—one may make causal statements. Here, we are primarily interested in understanding how predictive Shadow Cabinet membership is for Cabinet membership, rather than the precise *causal effect* of the former on the latter. Furthermore, we are well

aware that the particular set of dependence relationships assumed between Z, X and Y that we assume matters a great deal for any causal effect to be estimated.

Our general point from this section is that the problem of informal institutions can be formalized in a way that both comparative politics scholars and methodologists will understand. Furthermore that formulation in many cases will require careful and creative thought about latent variables and their measurement. Here our data are specific to legislative relations in Britain, but there are numerous examples of informal institutions that are similar in spirit: indeed, almost any unwritten rule that has manifest, observational implications fits into this framework. This includes, for example, the study of vote switching to chart changes in the Supreme Court's 'norm of consensus' (see Epstein, Segal and Spaeth, 2001). It also includes the practice of *amakudari* wherein retiring civil servants in Japan receive positions in corporations, possibly as a return for their preferential treatment of the firm while in office (Colignon and Usai, 2003). While we may not be able to directly observe favoritism by bureaucrats (X), it is possible that communications between the parties prior to appointment can be treated as manifestations of latent variables that might be measured (Z). Similarly, at a county or state level, the extent to which Jim Crow rules were in place (X) is presumably latent, though black voter turnout is observable (Y). A measurement strategy might proceed by considering the number of literacy tests utilized, or their complexity, to produce a variable Z in line with our approach.

4   Data

The data we use are described in Eggers and Spirling (Forthcoming) but 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.

For our purposes, the speeches are organized by 'parliamentary session', a period with a mean length of around 200 days. Each session begins 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.[8]

5   Methods

While our X, membership of the Shadow Cabinet, is latent, we can observe members making speeches which can inform us about X via our observed variable Z. Each MP has also has an observable set of covariates pertaining to their current role in the government (i.e. Y), 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.

---

[8]Including other opposition parties makes very little difference to the thrust of the substantive findings below.

5.1 Concept and Measurement

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. For a given term, we begin with a 'base rate' calculated as $\frac{n}{T}$, where n is the number of speeches using a particular word and T is the total number of speeches in the session. Thus, if there were a hundred mentions of the term 'boundary', and 10000 speeches, the base rate is $\alpha_0 = \frac{100}{10000} = 0.01$, corresponding to a mean wait time of $\frac{1}{0.01} = 100$ speeches. With the base rate in mind, we ask how the gaps between occurrences of the relevant term are changing as the session unfolds. In particular, 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 α_i, the rate, is proportional to a quantity s^i which we will define momentarily. Larger values of α imply smaller expected values on the wait time ($\frac{1}{\alpha}$) until the next event occurs.

For our purposes, s will be fixed at some value. We will estimate i, which will be different at different times for the same word (depending on the state of the system), and will be an integer greater than or equal to one. Put very crudely, the idea is to observe the series of gaps between uses of a term, and then to find values of i that when plugged into the rela-

tionship $\alpha_i = \frac{n}{T} s^i$ will fit the data, with respect to the (exponentially distributed) wait times that were seen in practice. Suppose we saw wait times of $20, 20, 10, 5, 10 \ldots$, and suppose further that s is fixed at some value. We can see that the third wait time (10) is half what the second one was (20), implying an increased value of i. Similarly, the fourth wait time (5) is half the previous one, implying that i has increased even further. The fifth wait time (10) suggests that i has declined, since the wait time has doubled. We will ultimately have a series of 'states' that describe our data, which is simply the vector of i values that we estimated.

To reiterate, i is the exponent of s: for a fixed s, an increasing i means that a *geometric* decrease must have been seen in wait times: that is, to go from our base rate model to s^1 to s^2 to s^3 requires at least a halving of the gap if $s = 2$ (and more than a halving if s is chosen to be larger than 2). Clearly, there will be a great many terms that never exhibit any bursts (they are 'not bursty') because their arrival rate is simply too uniform. Thus if the term 'bill' occurs (exactly) every 100 speeches, then obviously the gaps between observations are not changing. As a result, the term exhibits no bursts in use. This logic potentially extends to any word, no matter how common: e.g. the word 'the' might be used uniformly in *every* speech, and thus will demonstrate no bursts. In this case, the base rate, which is very high, will be a perfectly adequate model for the data.

The second component of the process, γ, 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. A larger gamma is associated with relatively few upwards transitions. Meanwhile, the exponential component (determined by s) 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

14

within periods of lower intensity activity. In principle, both s and γ 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 γ implies that the burst needs to be sustained for a longer period to 'count'. Binder (2012) implements the relevant model in R (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 "boundary", and in Figure 1 we picture the latter of these terms for the 1884 session in which the Redistribution of Seats Act—dealing specifically with the redrawing of districts—was discussed: note the levels (literally, the states of the Markov process given by the equation involving s^i and the cost term, where 0 is the base rate) of the bursts that the word went through, and the varying lengths of the those bursty episodes. In principle, we can do this for every term and every session.

Our innovation is now to use these burstiness estimates to compare MPs with each other. To do that, we need a metric that allows us to compute a score for each member taking into account the relative burstiness of their contributions. For us, this is a weighted sum. For each MP, a burst that *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
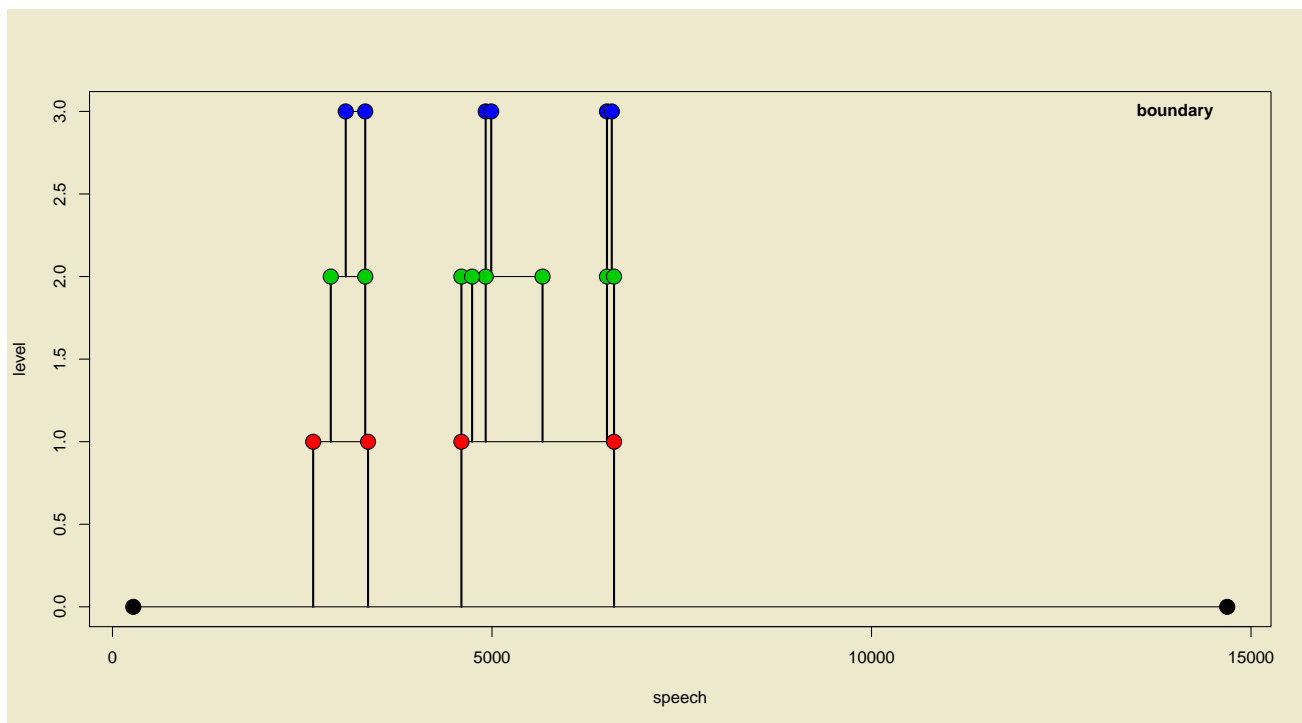
Figure 1: Burst levels (literally, the state of the Markov process at that point) and burst durations for the word 'boundary' in the final session of the 1880 parliament (1884).

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 uniform over time and will not be bursty. 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 Supp Info A, we give some pseudo-code to clarify the algorithm we used.

5.2   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. We report three particular

17

session	1846 (1841, 6)	1866 (1865, 1)	1885 (1885, 1)
terms (rank)	agriculturists (1) wheat (3) grain (5) farmer (6) prices (7)	suffrage (4) franchise (5) 1832 (7) redistribution (10) seats (11)	irishmen (2) 1782 (3) kingharmon (6) parnell (15) tenant (18)

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.

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 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 the earlier 1832 Great Reform Act. In 1885, the time of the controversial Government of Ireland Bill that would have delivered Home Rule to Ireland, we see surges of terms like 'irishmen' and their leader '[P]arnell' along with other terms specific to such discussion.

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. [9]

Moving on with our validation, we want to see that certain terms are bursty when we expect them to be. Consider Figure 2. There we report four terms with distinct burstiness 'signatures' over time. In each case, the y-axis is the burstiness of the word, calculated as its levels multiplied by the durations of those levels.[10] This is 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.

---

[9]There are methods that use speaker 'centrality' to detect influence Fader et al. (2007). Again a difference is that we focus on, and up-weight, legislators that are associated with surges in certain term use.

[10]Thus, a word that a burstiness of 1 for 2 periods, and then a burstiness of 2 for 6 periods would have a burstiness of $(1 \times 2) + (2 \times 6) = 14$. Note that this is literally the 'area under the curve' for the term specific equivalent of Figure 1.

As the third part of our validation exercise, we considered the burstiness profile of members of parliament, reporting our results in 3. William Gladstone and Benjamin Disraeli are both bursty during periods that they dominated the Commons (including as Prime Minister). The Irish parliamentary leader and strategist, Charles Parnell appears especially bursty during the 1880s, as expected. Finally, Samuel Plimsoll (MP for Derby) has a small but marked impact after the 1868 election, when he was responsible for pressuring the then government to introduce legislation to mandate 'waterlines' on merchat ships.

One concern readers may have is that burstiness is simply a stand-in (i.e. a proxy) for 'speechiness': that is, our metric measures nothing more than the ability to speak a great deal. This is not the case: the correlation between the burstiness of MPs and the number of speeches they make varies between 0.13 (1841) and 0.93 (1832) over the sessions as a whole. These variables are not measuring the same thing: while it is true that an MP may be non-bursty because he makes no speeches, making lots of speeches is no guarantee of being bursty. In particular, an MP who makes (perhaps thousands of) speeches that are simple responses, or contain terms that are not picked up by others, will not be bursty.[11]

6   Results

We have established a metric for measuring the agenda-setting ability of individual MPs. Ultimately, we want to use it to explore the ways in which the informal institution of interest—i.e. that Shadow Cabinet members become Cabinet members—evolved over time. This requires three interrelated steps: first, we need to show how and when that the opposition as a whole organized, and collectively paid more attention to agenda control. Second, within
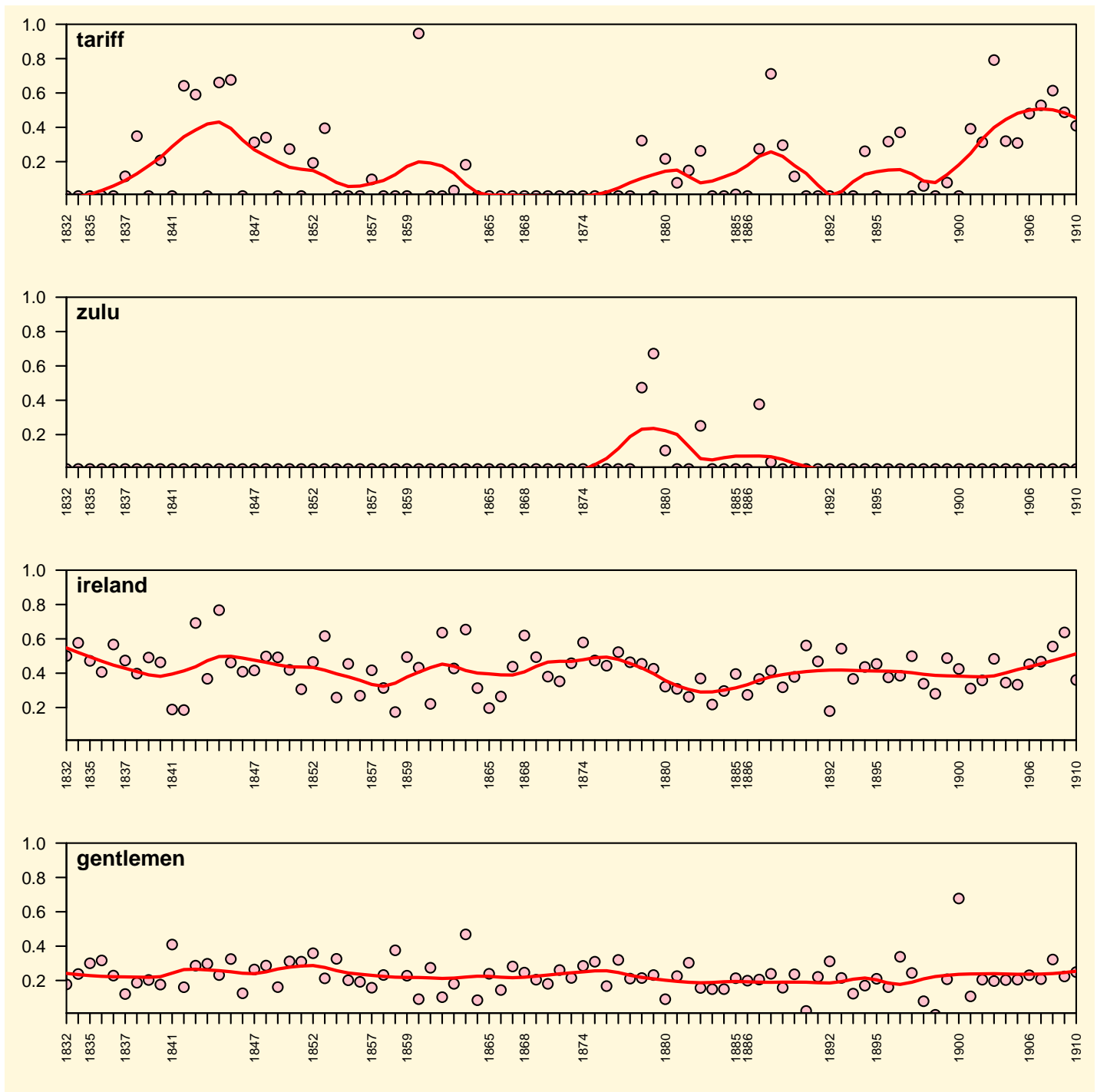
---

[11]See Supp Info B for more detailed information.

Figure 2: 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.
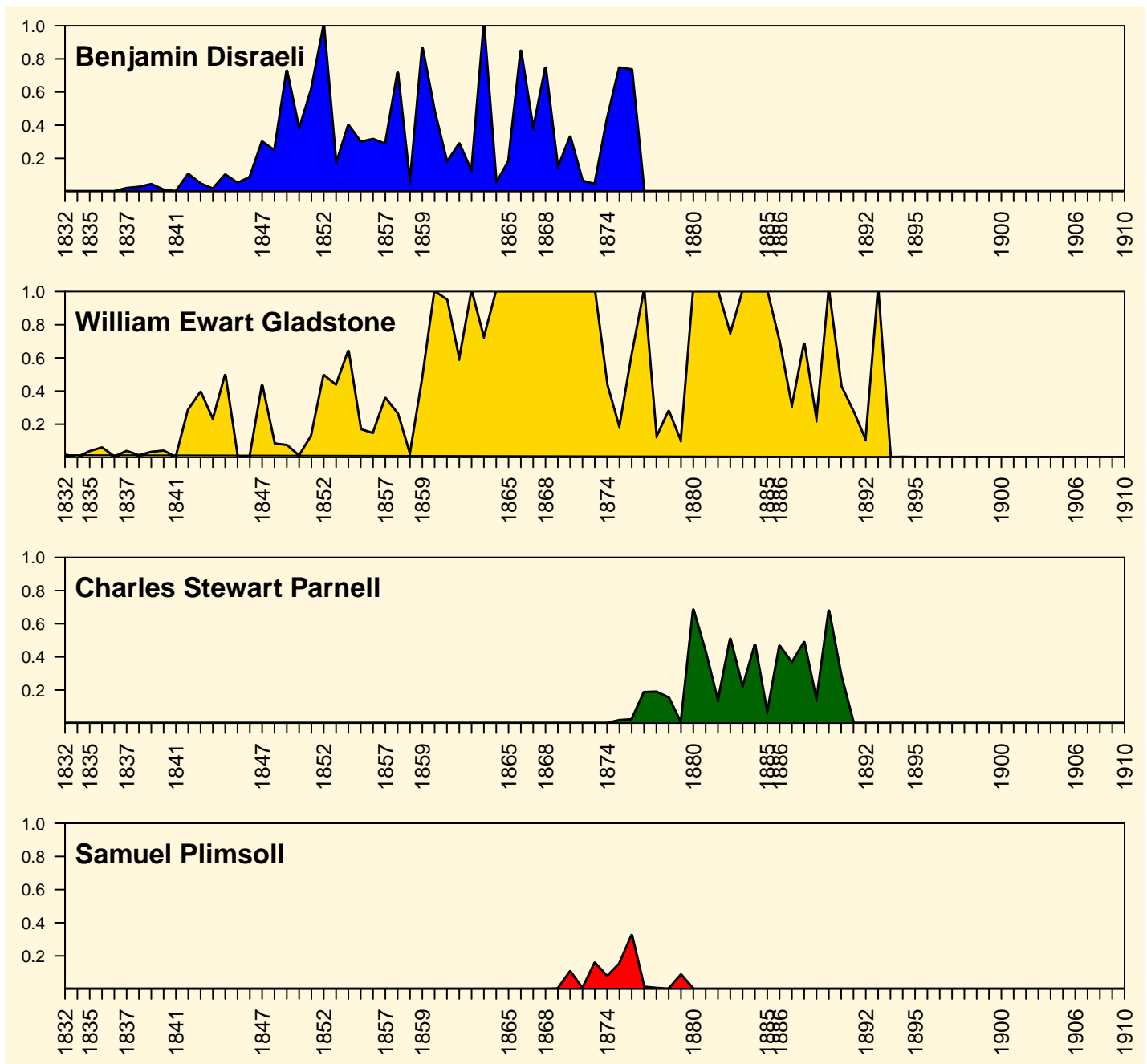
Figure 3: The burstiness profile of different MPs over time. The y-axis of the plots is the 'standardized' burstiness of the term, a rescaled metric where a value of 1 corresponds to the most bursty MP that session, while a value of 0 refers to the least bursty MP. The x-axis labels correspond to the beginnings of the various parliaments over the period.

that opposition, we need to explore the ways that agenda-power became concentrated in a 'leadership' group. That is, we need to assess whether and when a Shadow Cabinet could have been said to emerge. Third, given that we have established that the opposition organized, and that they did so under a Shadow Cabinet, we need to show that the latter became ministers at the exchange of power and that this relationship was non-constant over time.

6.1 Opposition Burstiness over Time

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.

In the upper panel of Figure 4 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 4 is that the cabinet was *always* more bursty than the opposition, on average: notice that the line is never below one. Given the
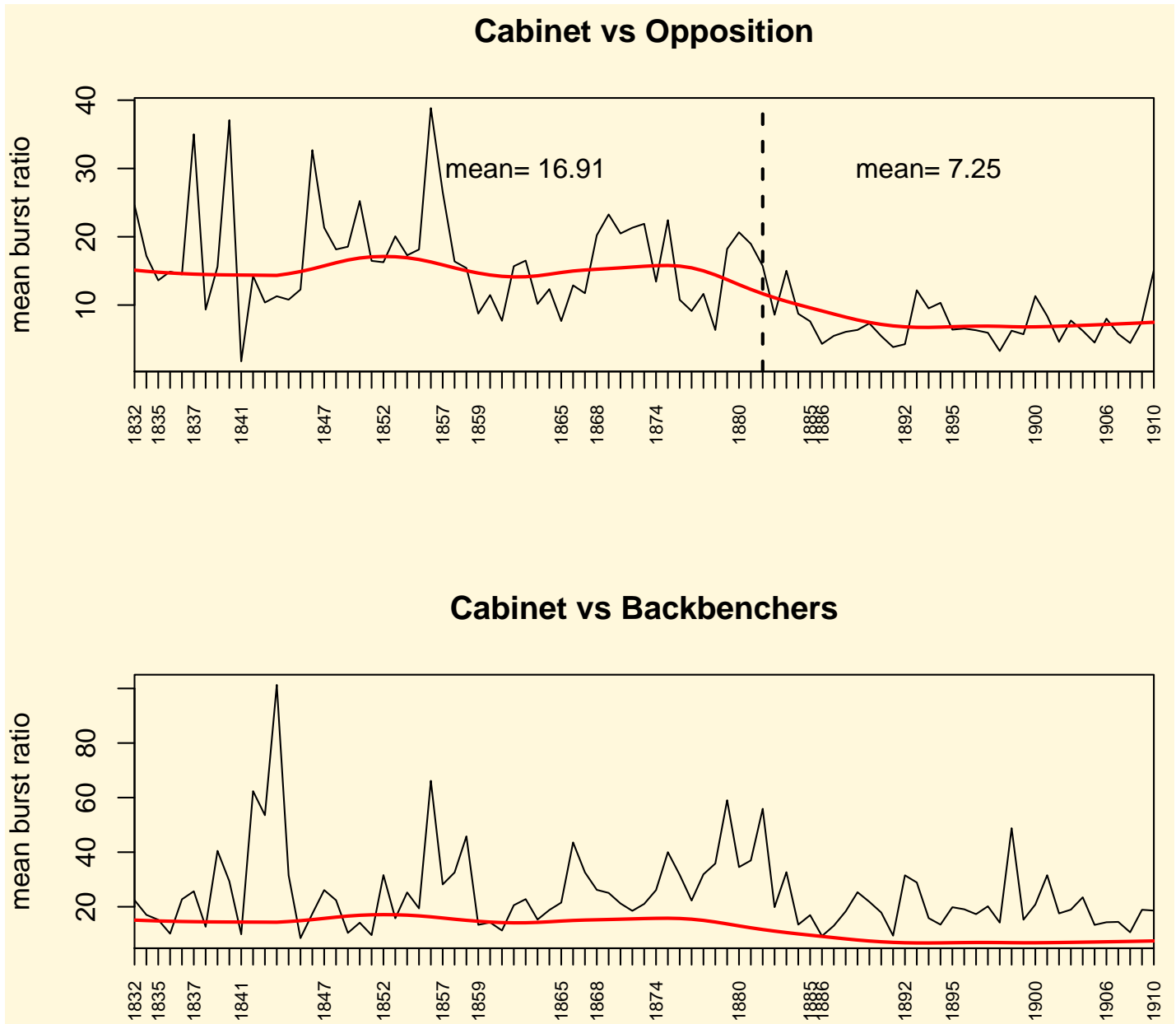
Figure 4: Ratio of (mean) burstiness: cabinet to opposition, cabinet to (Government party) backbenchers. One change point found in the opposition ratio time series, marked on the plot with the broken line and mean of ratio given on either side.

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).[12] 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 4, where we consider the (mean) ratio of the cabinet to government backbenchers. 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.

6.2   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

---

[12]See Zeileis et al. (2002) for implementation.

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 5 we examine the evidence for such a claim.

In the upper portion of Figure 5, 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. Formal time-series tests on the means of each session show that there is one break point, demarcated by a broken line during the third session of the parliament meeting in 1880. 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 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. 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
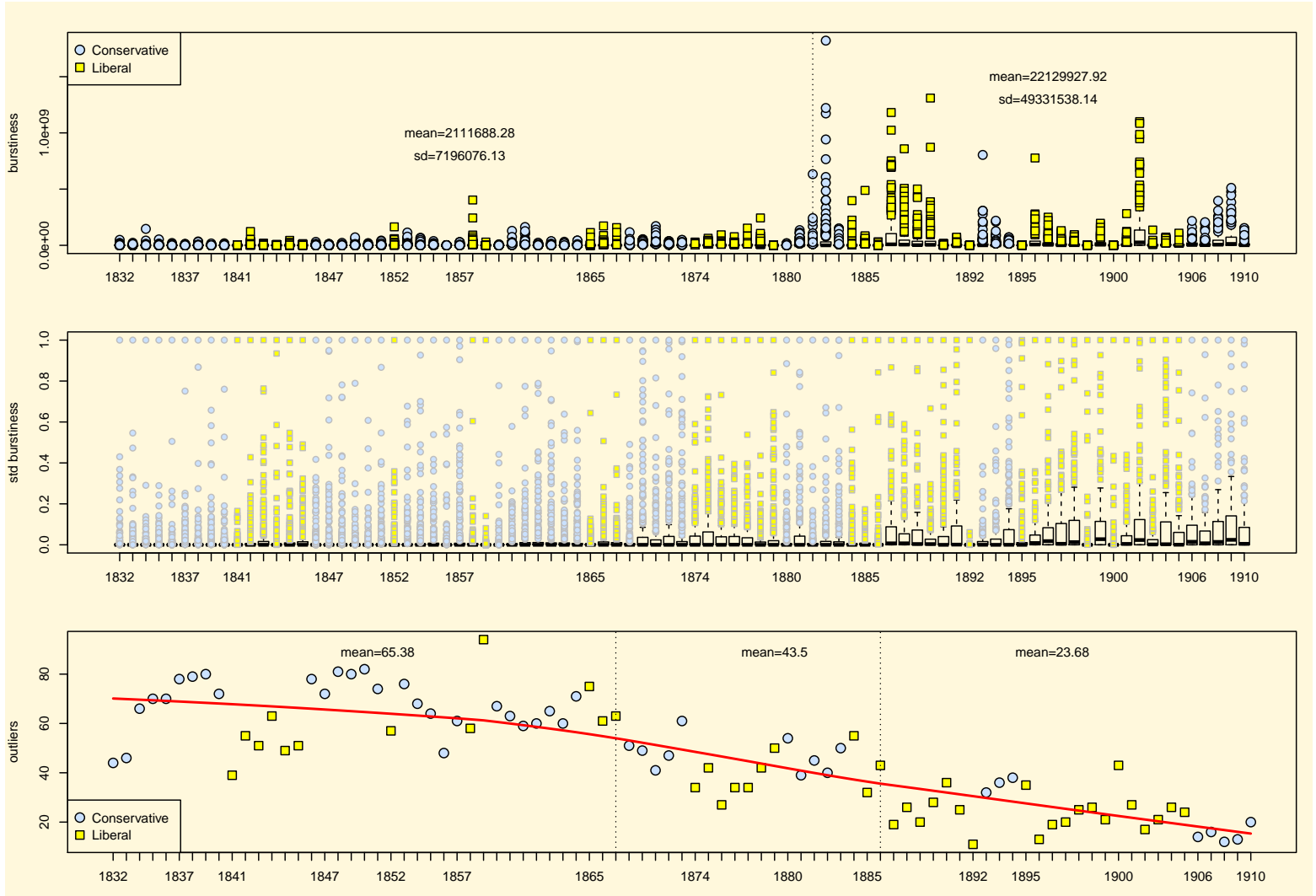
Figure 5: Concentration of agenda-setting power in the opposition over time. Top panel shows changing distribution of burstiness for the opposition; middle panel shows the same relationship but using standardized data, by session; lower panel shows (declining) number of outliers over time—consistent with the emergence of a 'shadow cabinet'.

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 5 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.

6.3 Burstiness and Future Cabinet status

One way to verify our presumption—that the outliers from Figure 5 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 'sesssion 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 an MP be promoted next time his party is in office, and, in fact, as time passes he is unconditionally more likely to be promoted. In Model 2, we try an alternate measure of 'leadership-ness', the number of speeches given by the MP. As we see from the higher AIC, this model does not fit the data as well as the previous one, implying that using the burstiness outlier metric provides useful extra leverage over more traditional alternatives.

In Model 3 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.[13] Thus the *net* effect of being an outlier is that one was more likely 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. Finally, in Model 4, we add a variable measuring whether or not (1 or 0) the member had previously served in the Cabinet. As expected, the coefficient is positive, but note that, crucially, being an outlier

---

[13]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 0.11, while for the last switch (in 1910) the predicted probability is around 0.19.

still matters (our coefficient is statistically significant). All told, our burstiness outlier metric is sound and helpful: it genuinely measures some notion of being in the Shadow Cabinet that is not simply captured by the number of speeches given, or having been previously selected as a Cabinet minister.[14]

	Model 1	Model 2	Model 3	Model 4
(Intercept)	−5.0724***	−3.7663***	−6.0391***	−6.5841***
	(0.2837)	(0.2154)	(0.5090)	(0.5713)
outlier	2.6113***		3.9262***	3.4753***
	(0.2129)		(0.5661)	(0.6597)
years	0.0179***	0.0038	0.0370***	0.0390***
	(0.0048)	(0.0047)	(0.0087)	(0.0098)
speech count		0.0253***		
		(0.0024)		
outlier×years			−0.0281**	−0.0330*
			(0.0105)	(0.0129)
prior service				4.4446***
				(0.2842)
N	3076	3076	3076	3076
AIC	871.6884	922.2606	866.3145	575.8072
BIC	944.0651	994.6373	962.8167	696.4349
$\log L$	−423.8442	−449.1303	−417.1573	−267.9036

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.

6.4 Summary

We have three interrelated results:

1. though the cabinet was always more bursty than the opposition, the latter became *relatively* more assertive in agenda-control terms around the time of the Second Reform Act (1868).

---

[14]In Supp Info C we verify that our outlier findings are robust to including raw burstiness as a an additional control and as alternate measure of Shadow Cabinet membership.

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 key informal institution of interest—the purported relationship between being in the Shadow Cabinet, and being in the Cabinet when the party in question took power— was present for the entire period, and became increasingly strong over time (in that an outlier in later sessions was more likely to find himself promoted to cabinet office than an outlier in earlier sessions)

7   Discussion

Informal institutions form the core of practical politics in Westminster systems where statute law is often silent: this includes the role of parties (there is none, constitutionally speaking), the role of whipping (never officially acknowledged) and the role of the Prime Minister itself (which has never been formally defined). This means that scholars of these polities, and comparative politics more generally, have a particularly pressing interest in understanding how inferences may be made about these norms and rules, if they are to plot their emergence and evolution over time. In this paper, we considered the role of the Shadow Cabinet as a 'government-in-waiting', a vital organization that ensures citizens 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).

First, we showed that informal institutions are helpfully modeled as graphs between variables, and that the special problems they present are partly due to the fact that at least one

of these variables is latent and thus cannot be directly observed. Our solution was to use new text-as-data methods, and applied to over a million speeches by members of parliament between the First and Fourth Reform Acts. Using a 'burstiness' metric, we showed that, after the 1870s, an increasingly small group of opposition 'leaders' closed the gap in terms of agenda-setting with their partisan competition in the Cabinet. Intriguingly, 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; Adelman, 1997; Rush, 2001; McLean, 2001; Aidt, Daunton 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, 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 'institutionalism' (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. Second, we took an explicitly 'agenda-setting' approach—a topic of very general interest to political scientists (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, we have not looked at the screening and selection mechanisms by which MPs joined the Shadow Cabinet: a specific 'career path' focus for the Victorian period, in line with more modern work (e.g. Benedetto and Hix, 2007; Kam, 2009) is called for. 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, to see how different systems converge or diverge in term use over time. We leave such efforts for future work.

Supporting Information (SI)

Supp Info 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:

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

1. draw the ith row of `tdm`, which is a binary vector of occurrences. Thus, supposing there were 10 speeches in the corpus, we might have $\{$`0,0,1,1,1,1,0,0,0,0`$\}$, 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 ith 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).

Supp Info B   Correlation between burstiness and 'speech-iness' over time

Figure 6 reports the correlation between all MPs' burstiness and the number of speeches they gave, for the various sessions in the data. Notice that the mean is around 0.7, implying that though the variables are correlated, they do not apparently measure identical concepts.

Supp Info C   Robustness: using 'raw' burstiness in regression

Recall that, in our account, being a 'leader' (and thus in the Shadow Cabinet) is associated with being an *outlier* in burstiness terms. In Table 3 we consider the robustness of this definition by fitting models using raw burstiness as an alternative measure of leadership (both additionally and as a substitute for outlier status).[15] As we see in MOdel I, the coefficient on outlier status is essentially unchanged, though a little extra explanatory power

---

[15]Outlier status is obviously post-treatment with respect to raw burstiness, and so one should not be overly confident in interpreting the coefficients.

Figure 6: Speechiness is not burstiness: correlation between burstiness of MPs and number of speeches they gave, over time.

	Model I	Model II
(Intercept)	−6.0205***	−3.6912***
	(0.5076)	(0.2269)
outlier	3.9864***	
	(0.5649)	
years	0.0364***	0.0079
	(0.0087)	(0.0050)
burstiness	0.0000***	0.0000***
	(0.0000)	(0.0000)
outlier×years	−0.0358***	
	(0.0108)	
burstiness×years		−0.0000†
		(0.0000)
N	3076	3076
AIC	856.9049	987.1880
BIC	977.5326	1083.6902
log L	−408.4525	−477.5940

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

Table 3: Raw Burstiness as a control and alternate measure of 'leadership' on opposition benches.

is added (AIC is lower). All in all, this suggests that it is being an outlier that 'matters', rather than simply being bursty from the opposition backbenches. Finally, we consider Model II 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.

References

Adelman, Paul. 1997. *Gladstone, Disraeli and Later Victorian Politics*. New York: Longman.

Aidt, Toke, Martin Daunton and Jayasri Dutta. 2010. "The Retrenchment Hypothesis and the Extension of the Franchise in England and Wales." *The Economic Journal* 120(547):990–1020.

Alt, James. 1975. Cabinet Studies: A Reader. London: Macmillan chapter Continuity, Turnover and Experience in the British Cabinet, 1868–1970, p. 3354.

Bagehot, Walter. 1873/2011. *The English Constitution (2nd Edn)*. Accessed June 6, 2011.: http://socserv.mcmaster.ca/ econ/ugcm/3ll3/bagehot/constitution.pdf.

Bai, Jushan and Pierre Perron. 2003. "Computation and Analysis of Multiple Structural Change Models." *Journal of Applied Econometrics* 18:1–22.

Bateman, Joel. 2009. "In The Shadows: The Shadow Cabinet in Australia." *Parliament of Australia: Department of Parliamentary Services* .

Benedetto, Giacomo and Simon Hix. 2007. "The Rejected, the Ejected and the Dejected: Explaining Government Rebels in the 2001-5 British House of Commons." *Comparative Political Studies* 40(7):755–781.

Berlinski, Samuel and Torun Dewan. 2011. "The Political Consequences of Franchise Extension: Evidence from the Second Reform Act." *Quarterly Journal of Political Science* 6(34):329–376.

Berlinski, Samuel, Torun Dewan and Keith Dowding. 2007. "The length of ministerial tenure in the United Kingdom, 1945-97." *British Journal of Political Science* 37(4):245–262.

Berrington, Hugh. 1968. "Partisanship and Dissidence in the Nineteenth Century House of Commons." *Parliamentary Affairs* XXI:338–74.

Binder, Jeff. 2012. *bursts: Markov model for bursty behavior in streams*. R package version 1.0.

Blackwell, Matthew. 2013. "A Framework for Dynamic Causal Inference in Political Science." *American Journal of Political Science* 57(2):504–520.

Bogdanor, Vernon and David Butler. 1983. *Democracy and elections: Electoral systems and their political consequences*. New York: Cambridge University Press.

Brazier, Julian. 1997. *Ministers of the Crown*. Oxford: Clarendon Press.

Brazier, Rodney. 1999. *Constitutional Practice: The Foundations of British Government*. New York: Oxford University Press.

Butler, David and Donald Stokes. 1969. *Political Change in Britain.* London: Macmillan.

Butler, David and Gareth Butler. 1994. *Twentieth Century British Historical Facts.* 8th ed. London: MacMillan Press.

Cain, Bruce, John Ferejohn and Morris Fiorina. 1987. *The Personal Vote: Constituency Service and Electoral Independence.* Cambridge: Harvard University Press.

Chester, Norman and Nona Bowring. 1962. *Questions in Parliament.* Oxford: Clarendon Press.

Clarke, Harold, David Sanders, Marianne Stewart and Paul Whiteley. 2004. *Political Choice in Britain.* Oxford: Oxford University Press.

Cobb, Roger, Jennie-Keith Ross and Marc Howard Ross. 1976. "Agenda Building as a Comparative Political Process." *American Political Science Review* 70(1):126–138.

Colignon, Richard and Chikako Usai. 2003. *Amakudari: The Hidden Fabric of Japan's Economy.* Ithaca: Cornell University Press.

Cook, Christopher and Brendan Keith. 1975. *British Historical Facts 1830-1900.* London: Palgrave-MacMillan.

Cowley, Philip. 2002. *The Rebels.* London: Politicos.

Cox, Gary. 1987. *The Efficient Secret: The Cabinet and the development of Political Parties in Victorian England.* Cambridge: Cambridge University Press.

Cox, Gary. 1992. "The Origin of Whip Votes in the House of Commons." *Parliamentary History* 11(2):278–285.

Cox, Gary and Mathew McCubbins. 2005. *Setting the agenda: Responsible party government in the US House of Representatives.* Cambridge: Cambridge University Press.

Craig, Frederick. 1974. *British Parliamentary Election Results, 1885–1918.* London: MacMillan Press.

Craig, Frederick. 1989. *British Parliamentary Election Results, 1832–1885.* London: Macmillan Press.

Crisp, Leslie. 1983. *Australian National Government.* Melbourne: Longman.

Desposato, Scott. 2006. *Informal Institutions and Democracy: Lessons from Latin America.* Baltimore, Maryland: Johns Hopkins University Press chapter How Informal Electoral Institutions Shape the Brazilian Legislative Agenda, pp. 125–142.

Dewan, Torun and Arthur Spirling. 2011. "Strategic Opposition and Government Cohesion in Westminster Democracies." *American Political Science Review* 105(2):337–358.

Dewan, Torun and David Myatt. 2010. "The Declining Talent Pool of Government." *American Journal of Political Science* 54(2):267–286.

Doring, Herbert. 1995. Institutions and Policies: Why We Need Cross National Analysis. In *Parliaments and Majority Rule in Western Europe.* New York: St Martin's Press.

Eberhardt, Frederick. 2013. "Experimental Indistinguishability of Causal Structures." *Philosophy of Science* 80(5):684–696.

Eggers, Andrew and Arthur Spirling. Forthcoming. "Electoral Security as a Determinant of Legislator Activity, 1832–1918: New Data and Methods for Analyzing British Political Development." *Legislative Studies Quarterly* Special Issue on British Political Development.

Epstein, Lee, Jeffrey A. Segal and Harold J. Spaeth. 2001. "The Norm of Consensus on the U.S. Supreme Court." *American Journal of Political Science* 45(2):362–377.

Erskine May, Thomas. 1864/1986. *The constitutional history of England since the accession of George Third, 1760-1860.* Littleton, Colorado: F.B. Rothman.

Fader, Anthony, Dragomir Radev, Burt L. Monroe and Kevin M. Quinn. 2007. MavenRank: Identifying influential members of the US senate using lexical centrality. In *In Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning.* pp. 658–666.

Franklin, Mark and Philip Norton. 1993. Parliamentary Questions. Oxford: Clarendon Press chapter Questions and Members.

Fraser, Peter. 1960. "The Growth of Ministerial Control in the Nineteenth-Century House of Commons." *English Historical Review* 75(296):444–463.

Gash, Norman. 1952. *Politics in the Age of Peel: A Study in the Technique of Parliamentary Representation, 1830-1850.* London: Longmans, Green, and Co.

Godbout, Jean-Francois and Bjorn Hoyland. 2013. "Parties and Voting in Parliament." *Working Paper* Princeton University.

Grimmer, Justin and Brandon Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21(3):267–27.

Hall, Peter and Rosemary Taylor. 1996. "Political Science and the Three New Institutionalisms." *Political Studies* 44(5):936–57.

Heath, Anthony, John Curtice, Roger Jowell, Geoffrey Evans, John Field and Sharon Witherspoon. 1991. *Understanding Political Change: the British Voter, 1964–1987.* Oxford: Pergamon.

Helmke, Gretchen and Steven Levitsky. 2004. "Informal Institutions and Comparative Politics: A Research Agenda." *Perspectives on Politics* 2(4):725–740.

Holzhacker, Ronald. 2005. "The power of opposition parliamentary party groups in European scrutiny." *The Journal of Legislative Studies* 11(3-4):428–445.

Howe, Anthony. 1998. *Free Trade and Liberal England, 1846-1946.* Oxford: Oxford University Press.

Jenkins, Terence. 1996. *Parliament, party and politics and Victorian Britain.* New York: Manchester University Press.

Jenks, Edward. 1903. *Parliamentary England: the Evolution of the Cabinet System.* New York: G Putnam's Sons.

Jensen, Jacob, Ethan Kaplan, Suresh Naidu and Laurence Wilse-Samson. 2012. "Political Polarization and the Dynamics of Political Language: Evidence from 130 years of Partisan Speech." *Brookings Papers on Economic Activity* Fall.

Johnson, Nevil. 1997. "Opposition in the British Political System." *Government and Opposition* 32(4):487–510.

Kam, Christopher and Indridi Indridason. 2005. *Legislative Studies Quarterly* XXX(3):327–363.

Kam, Christopher J. 2009. *Party Discipline and Parliamentary Politics.* Cambridge: Cambridge University Press.

King, Anthony. 1994. Cabinet Ministers and Parliamentary Government. Cambridge: Cambridge University Press chapter Ministerial Autonomy in Britain.

Kleinberg, Jon. 2002. Bursty and hierarchical structure in streams. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining.* New York, NY, USA: ACM pp. 91–101.

Krehbiel, Keith. 1992. *Information and Legislative Organization.* Ann Arbor: University of Michigan Press.

Krehbiel, Keith. 1998. *Pivotal Politics: A Theory of US Lawmaking.* Chicago: University of Chicago Press.

Lauth, Hans-Joachim. 2000. "Informal Institutions and Democracy." *Democratization* 7(4):21–50.

Lijphart, Arend. 1999. *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries.* New Haven, CT: Yale University Press.

Lowell, Abbott. 1908. *The Government of England.* New York: Macmillan.

Mahoney, James and Kathleen Thelen. 2009. A Theory of Gradual Institutional Change. In *Explaining Institutional Change: Ambiguity, Agency and Power*, ed. James Mahoney and Kathleen Thelen. Cambridge: Cambridge University Press.

Marriott, John. 1925. *English political institutions: An Introductory Study*. Oxford: Oxford University Press.

McLean, Iain. 2001. *Rational Choice and British Politics*. Oxford: Oxford University Press.

Pollack, Mark. 1997. "Delegation, agency, and agenda setting in the European Community." *International Organization* 51(1):99–134.

Potter, Allan. 1965. Political Oppositionsin Western Democracies. New Haven: Yale University Press chapter Great Britain: Opposition with a Capital "O".

Powell, Bingham. 2000. *Elections as Instruments of Democracy: Majoritarian and Proportional Visions*. New Haven: Yale University Press.

Power, Charles. 1966. *A party politician: The memoirs of Chubby Power*. Toronto: Macmillan of Canada.

Punnett, Robert. 1973. *Front-bench Opposition: The Role of the Leader of the Opposition, the Shadow Cabinet, and Shadow Government in British Politics*. London: Heinemann.

Quinn, Kevin, Burt Monroe, Michael Colaresi, Michael H. Crespin and Dragomir Radev. 2010. "How to Analyze Political Attention with Minimal Assumptions and Costs." *American Journal of Political Science* 54:209–228.

Quinn, Thomas. 2012. *Electing and Ejecting Party Leaders in Britain*. New York: Palgrave Macmillan.

R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Redlich, Josef. 1908. *The Procedure of the House of Commons: A Study of its History and Present Form*. London: Constable & Co.

Rhodes, Rod, John Wanna and Patrick Weller. 2009. *Comparing Westminster*. Oxford: Oxford University Press.

Rhodes, Rod and Patrick Weller. 2005. Westminster Legacies: Democracy and Responsible Government in Asia and the Pacific. University of New South Wales: University of New South Wales Press chapter Westminster Transplanted and Westminster Implanted: Exploring Political Change.

Rush, Michael. 2001. *The Role of the Member of Parliament Since 1868: From Gentlemen to Players*. Oxford: Oxford University Press.

Schonhardt-Bailey, Cheryl. 2003. "Ideology, Party and Interests in the British Parliament of 1841-1847." *British Journal of Political Science* 33:581–605.

Schumpeter, Joseph. 1942. *Capitalism, Socialism and Democracy*. New York: Harper.

Stokes, Susan. 2006. *Informal Institutions and Democracy: Lessons from Latin America.* Baltimore, Maryland: Johns Hopkins University Press chapter Do Informal Rules Make Democracy Work? Accounting for Accountability in Argentina, pp. 125–142.

Turner, Duncan. 1969. *The Shadow Cabinet in British Politics*. New York: Routledge.

Volden, Craig, Alan Wiseman and Dana Wittmer. 2013. "When Are Women More Effective Lawmakers Than Men?" *American Journal of Political Science* 57(2):326–341.

Walker, Brian. 1978. *Parliamentary Elections in Ireland, 1801-1922*. Dublin: Royal Irish Academy.

Wawro, Gregory and Eric Schickler. 2006. *Filibuster: Obstruction and Lawmaking in the US Senate*. Princeton: Princeton University Press.

Zeileis, Achim, Friedrich Leisch, Kurt Hornik and Christian Kleiber. 2002. "strucchange: An R Package for Testing for Structural Change in Linear Regression Models." *Journal of Statistical Software* 7(2):1–38.