Measuring Emotion in Parliamentary Debates Using Methods of Natural Language Processing

Ludovic Rheault[1], Kaspar Beelen[1,2], Christopher Cochrane[1] and Graeme Hirst[2]

[1]Department of Political Science, University of Toronto

[2]Department of Computer Science, University of Toronto

[**Please do not quote or circulate this paper without the authors' consent.**]

Abstract

An impressive breadth of interdisciplinary research suggests that emotions have an influence on political behaviour. Nonetheless, we still know very little about the emotional states of those actors whose decisions have a great impact on our societies: politicians in parliament. We address this problem by making use of methods of natural language processing and a digitized corpus of text data spanning a century of parliamentary debates in the United Kingdom. Algorithms for detecting emotion in speech and text have advanced considerably over the past few decades, with applications to social media analytics, stock market predictions and a wide range of other areas. However, existing methods have little exportability to the study of political discourse and debates, which differ markedly in terms of linguistic register and genre. Therefore, we develop a new approach that can be adapted to specific domains, allowing us to measure the emotional polarity of political debates over time. We use this approach to examine the change in aggregate levels of emotional polarity in the UK parliament, and to test a hypothesis about the emotional response of politicians to economic recessions. Our findings suggest that, contrary to popular belief, the mood of politicians has become more positive during the past decades, and that variations in emotional polarity can be predicted by economic business cycles.

Prepared for the Canadian Political Science Association 2015 Annual Conference at the University of Ottawa, Ottawa ON, Canada, June 2–4.

Measuring Emotion in Parliamentary Debates Using Methods of Natural Language Processing

Developments in natural language processing tools have opened the door to new possibilities in the humanities and social sciences, where the main subjects of analysis, human beings, communicate many of their thoughts and feelings with words. Recent examples of such headways include empirical analyses of gargantuan quantities of digitized books from the past centuries (Acerbi et al. 2013; Hughes et al. 2012; Michel et al. 2011) and stock market or box office predictions based on affective computing methods applied to millions of Twitter posts (Asur and Huberman 2010; Bollen, Mao, and Zeng 2011). The present study is an attempt to further break boundaries between social sciences and computer science, by focusing on the language of parliamentary proceedings. Our main goal is to adapt affective computing methods to the study of political discourse. We develop a methodology to produce domain-specific polarity lexicons and implement this approach using the entire corpus of proceedings of the British House of Commons during the past hundred years. Next, our paper illustrates the potential of this methodology by tackling a specific question about the emotional states of policy-makers. We argue that politicians not only represent the preferences of their constituents over issues debated in parliament, they also react emotionally to national and world events in a manner that is predictable. In essence, politicians mirror the feelings and apprehensions of civilians in the face of adversity. We test this claim by tracking down the dynamics of politicians' emotional responses during economic hard times. Our empirical results contribute to mounting evidence accumulated in social sciences about the linkages between emotion and human behaviour (see Loewenstein 2000; Neuman et al. 2007; McDermott 2004; Mercer 2005).

Our decision to focus on emotions in political discourse was propelled by the rise in importance of two influential streams of literature. The first one concerns the psychology of human behaviour. Since at least the 1960s, developments in the field of behavioural economics have brought the psychological aspects of decision-making to the forefront. Seminal works on bounded rationality (Simon 1967), prospect theory (Kahneman and Tversky 1979; 1984) and regret theory (Loomes and Sugden 1982) have all attempted to tackle the observed discrepancies between theories grounded in rationality and actual human behaviour. Lending additional credence to this field of research, a recent body of work stressed the conclusion that the emotionality of human decision-making has an intrinsic, neural basis (Bechara, Damasio, and Damasio 2000; Dawes et al. 2012; Krueger et al. 2009; Hsu et al. 2005; Takahashi 2013). The progress of neuroscience has even led some scholars to coin the term "neuroeconomics" to speak of its applications in the discipline of economics (Camerer, Loewenstein, and Prelec 2005). Implications for the study of political behaviour have been surveyed by McDermott (2004). A main concern behind this study is that making sense of the decisions made by elected politicians entails being able to tap, in one way or another, their emotions. Yet, most of the research on emotion and political behaviour has focused on individuals outside of real-life democratic institutions, such as survey respondents or experimental subjects.[1] We still know little about the emotional aspects of parliamentary debates. Until now, addressing this question has been hindered by the apparent difficulty of monitoring politicians from afar, let alone measuring their states of mind. The theory and methods that we introduce in this paper are an attempt to fill that gap, by focusing on large amounts of officially recorded political writings and by importing the practices developed in another field: natural language processing.

The second stream of literature stems from computer science and concerns the detection of sentiment and emotion in textual data. More specifically, *affective computing* refers to a wide range of computational tools for the measurement of emotional states and affective responses communicated by humans either through facial and corporal expressions, oral or written speech (see Cowie and Cornelius 2003; Ishizuka, Neviarouskaya, and Shaikh 2012; Jaywant and Pell 2012; Picard 1997; Schuller and Batliner 2014).[2] Recent works in this field have examined a variety of topics, from the use of emotions in musical lyrics (DeWall et al. 2011) to the spread of happiness in social media (Dodds et al. 2011; Bollen, Mao, and Zeng 2011) and the study of medical conditions such as depression in online communities (Nguyen et al. 2014). Of particular relevance for this study are the prior attempts to link affective computing to political and socio-economic topics. Bollen, Mao, and Pepe (2011), for instance, were able to measure meaningful emotional reactions among Twitter users in the face of real-world events such as the 2008 U.S. presidential campaign, shifts in market indicators of the recession and changes in oil prices. Variations in public mood as a response to economic and political events have also been studied in a growing number of papers (e.g. Lansdall-Welfare, Lampos, and Cristianini 2012; Mohammad et al. 2014; Johnson, Shukla, and Shukla 2014). Those works provide evidence that individuals react emotionally to macroeconomic conditions, justifying further our interest in finding out whether politicians exhibit similar attitudes inside parliamentary institutions.

[1] See, for example, Arceneaux and Wielen (2013); Bickford (2011); Black et al. (2011); Brader (2005; 2006; 2011); Camerer, Loewenstein, and Prelec (2005); Caplin and Leahy (2001); Civettini and Redlawsk (2009); Crawford (2000); Glimcher, Dorris, and Bayer (2005); Kaufman (1999); Ladd and Lenz (2008); Lodge and Taber (2005); Loewenstein (2000); Marcus, Neuman, and MacKuen (2000); Marcus (2002); Miller (2011); Pagano and Huo (2007); Petersen et al. (2012); Rick and Loewenstein (2008); Small and Lerner (2008).

[2] Those methods are sometimes referred to as emotional prosody or textual affect sensing.

Affective Computing and Political Debates

Our corpus consists of all available Hansards of the British House of Commons between 1909 and 2013 inclusive. It contains all the debates, oral questions and oral answers to written questions. The format of the Hansards—the official text archives of debates and speeches—was modified in 1909, when new standards were implemented for the verbatim record of the debates (United Kingdom 2010). This is why our corpus begins at that date. Those text documents have been digitized and stored in a markup language as part of the international project *Digging into Linked Parliamentary Data* (Dilipad).[3] The corpus comprises a total of approximately 925.4 million tokens, with an average of 8.8 million tokens each year. Considering the lemmatized version of the corpus—that is, the roots of words, which avoids duplicate counting of plural and conjugated verb forms—and counting only tokens appearing 10 times or more, UK parliamentarians have used a vocabulary of 102,404 tokens. The length of parliamentary sessions has increased over time, and so has the corpus size per year: the decade 1910-1919 had an average size of 5.7 million tokens, compared to 9.8 million during the decade 2000-2009. The Appendix contains additional details concerning data collection.

Many approaches have been developed for capturing sentiment or emotional polarity in textual data, using either machine learning classifiers trained on human-annotated corpora or lexicon-based affect sensing. One of the main challenges with those approaches, however, is that most of them are domain-specific. Sentiment analysis classifiers trained with corpora from a specific domain were shown to have limited exportability to texts using a different register or genre (Aue and Gamon 2005; Qiu et al. 2009). Polarity lexicons provide lists of words annotated with scores, which can be used either as a feature for classifiers or to compute frequency measures of polarity in a target corpus. Popular examples include Mohammad, Kiritchenko, and Zhu (2013)'s lexicon, in which words were annotated for six specific emotions and positive/negative polarity using crowdsourcing, SentiWordNet (Esuli and Sebastiani 2006), created using recursive algorithms based on the WordNet database, as well as the polarity lexicon of OpinionFinder (Wiebe, Wilson, and Cardie 2005). Like pre-trained sentiment analysis classifiers, those lexicons have limitations when used across domains. Parliaments are associated with a decorum, and political expressions convey specific meanings and interpretations that we need to take into account. For instance, those three lexicons give a negative score to a word such as `war`, and positive ones to `education` and `health`. Yet, a word like `war` will inevitably be used more frequently in times of war, since the topic is frequently discussed in parliament, but assuming that debates become more negative simply because of the increased presence of this word would be misleading. Similarly, nouns like `education` and `health` have different meanings in politics as they re-

[3]See `http://dilipad.history.ac.uk/`.

late to policy domains, identify the name of specific departments and ministerial functions. An increased use of the word `education` would provide little information about the tone of the debates taking place in the House of Commons. In short, we would like to avoid words that have a descriptive, domain-specific usage.

To overcome those problems, we create a domain-specific polarity lexicon from the vector representation of the entire corpus. The general approach that we follow here has been introduced in Turney and Littman (2003) and a related methodology is discussed in Maas et al. (2011). We start by creating the vector-space model of our corpus using the *GloVe* algorithm (Pennington, Socher, and Manning 2014) (details of our specification appear in the Appendix). This method converts the vocabulary of our corpus into numerical vectors based on the matrix of word-word co-occurrences. We compute vectors of 300 dimensions for each combination of lemma and part-of-speech. For simplicity, we use the expression lemma in what follows to speak of a lemma/part-of-speech pair. The second step of our methodology consists of creating a list of 200 unambiguous seed lemmas capturing positive and negative emotional polarity (100 lemmas for each pole). Using vector distances to detect word similarities, we attribute to all other lemmas in the vocabulary a score indicating how close they are to each of the two groups of seeds. The formula corresponds to:

$$s_i = \sum_{p=1}^{P} \frac{\boldsymbol{v}_i \cdot \boldsymbol{v}_p}{\|\boldsymbol{v}_i\| \|\boldsymbol{v}_p\|} - \sum_{q=1}^{Q} \frac{\boldsymbol{v}_i \cdot \boldsymbol{v}_q}{\|\boldsymbol{v}_i\| \|\boldsymbol{v}_q\|} \tag{1}$$

where $\|\boldsymbol{v}_i\|$ is the norm of vector $\boldsymbol{v}_i$ associated with lemma i, and where the seed lemmas for positive and negative emotions are indexed by $p = \{1, ..., P\}$ and $q = \{1, ..., Q\}$, respectively. The scores s_i are scaled into a $[-1, 1]$ index reflecting their emotional polarity. We retain the 2000 lemma/part-of-speech pairs with the highest and lowest scores, expanding our lexicon to 4200 words. To illustrate the output of this method, we report the first 20 lemmas with the highest and lowest scores in Table 1. By redistributing those scores to the lemmas across the original corpus, we are able to quantify the mood of parliamentary debates over time, which can be aggregated by sentence, speech, session, month, quarter, and so forth.

For the purpose of this study, we create both a measure of total emotionality and a measure of polarity. Let $\mathbf{1}\{w_i \in L\}$ designate an indicator function equaling one if the lemma w_i features in our lexicon, denoted by L, and zero otherwise. *Emotionality* is measured by counting the frequency of all lemmas belonging to the polarity lexicon, weighted by their absolute scores:

$$z_t = \frac{\sum_{i=1}^{n_t} \mathbf{1}\{w_{it} \in L\} |s_i| \theta_{it}}{\sum_{i=1}^{n_t} \mathbf{1}\{w_{it} \in L\}}, \tag{2}$$

where z_t captures the total level of emotionality in period t and n_t is the total number of lemmas in the parliamentary debates of period t. To account for negative clauses, we

Table 1: Highest Scores in Domain-Specific Polarity Lexicon

Positive			Negative		
Lemma	P-o-S	s_i	Lemma	P-o-S	s_i
congratulate	verb	1.00	compound	verb	–1.00
commend	verb	0.95	dreadful	adjective	–0.98
high-quality	adjective	0.94	grievous	adjective	–0.98
robust	adjective	0.91	appalling	adjective	–0.96
tribute	noun	0.90	mismanagement	noun	–0.89
balanced	adjective	0.90	ruinous	adjective	–0.88
worthwhile	adjective	0.90	intolerable	adjective	–0.88
welcome	verb	0.90	vexation	noun	–0.87
impressive	adjective	0.90	frightful	adjective	–0.86
superb	adjective	0.87	exacerbate	verb	–0.85
thank	verb	0.87	unfairness	noun	–0.85
informative	adjective	0.87	cruel	adjective	–0.85
constructive	adjective	0.87	cause	verb	–0.84
warm	adjective	0.86	misery	noun	–0.83
marvellous	adjective	0.85	aggravate	verb	–0.81
imaginative	adjective	0.85	grotesque	adjective	–0.80
warmly	adverb	0.84	scandalous	adjective	–0.80
co-operation	noun	0.84	stupidity	noun	–0.80
lively	adjective	0.83	unacceptable	adjective	–0.80
enable	verb	0.83	humiliation	noun	–0.80

The table reports the first twenty positive and negative lemmas/part-of-speech pairs and their scores.

constructed a parameter θ_{it} measuring the valence of lemma i in the corpus of period t. This parameter is set to 1 unless w_{it} is located between a word indicative of a negative clause and a punctuation mark, in which case it equals 0 (words indicating negative clauses include "not", "no", "never", "neither" and "nor"). On the other hand, *Polarity* is defined by taking the observed scores s_i instead of their absolute values:

$$y_t = \frac{\sum_{i=1}^{n_t} \mathbf{1}\{w_{it} \in L\} s_i \theta_{it}}{\sum_{i=1}^{n_t} \mathbf{1}\{w_{it} \in L\}}. \tag{3}$$

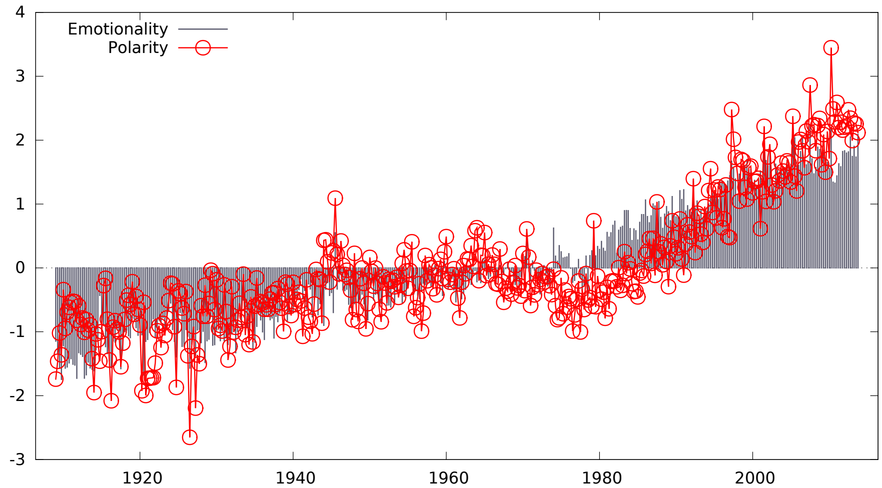
Higher values of y_t indicate that debates become more positive. The measure of polarity, just like the score variable s_i, can be negative or positive. However, positive words are used more frequently in English language and as a result, aggregate measures will tend to remain in the positive range as soon as the corpus length increases. Our interest lies in the temporal change in y_t; thus, the scaling of that measure is irrelevant. Importantly, notice that (2) and (3) use the count of lemmas as the denominator, to account for the fact that parliamentary sessions may differ in length from one period to the next. Additional details on the construction of our measures are provided in the Appendix.

Figure 1 reports our aggregate measures y_t and z_t, by quarter and by year. All measures have been normalized to facilitate comparisons. Arguably the most striking feature of those graphs is the clear rising trend in both indicators over time, suggesting that political debates have become more emotional, but also more positive in recent years. This last observation may sound counter-intuitive for contemporary observers of political affairs. Conventional wisdom suggests that modern political parties are prone to cynicism during debates in parliament, and the Prime Minister's Question Time draws attention for its lively tone. However, when considering the bigger picture of the 20th century, such a verdict must be qualified. In fact, the trends appear consistent with the major turns of events of the past century. The first two decades of our sample mark an era of negative polarity in the British House of Commons, in line with the social divide of those turbulent years. The 1910s and 1920s were punctuated by major labour disputes, including a nation-wide strike in 1926, the threat of socialism materialized with the arrival into power of a new Labour Party in the early 1920s, not to mention the First World War and the Irish War of Independence of 1920-1922. The current tone of parliamentary proceedings seems to be far more optimistic compared to the early 20th Century, which is consistent with a recent study claiming that elites have moderated their positions in Britain since the Thatcher era (Adams, Green, and Milazzo 2012). This finding—that debates have become more positive over time—appears robust, and resurfaces even when considering alternative indicators of emotional polarity based on the three general lexicons mentioned above, as Figure 7 in the Appendix illustrates.

Our indicators also appear to exhibit the characteristics of long-memory processes, as suggested by the slow decay of their autocorrelation functions (ACF) (Figure 6 in the Appendix). In the two quarterly series under consideration (y_t and z_t), the ACF spikes remain greater than 0.5 for at least 35 lags. The long-memory feature means that exogenous shocks affecting the mood of politicians can have a lasting effect on the nature of debates over time. The power spectra of the series are reported in Figure 2. Using linear regressions on a log scale, we estimated parameters from power spectral densities of the form $\frac{1}{f^\alpha}$, where f is the frequency and α a parameter which should approximate 1 in the presence of pink noise. The fitted values of α are close to one when using quarterly series (1.07 and 1.47, respectively for Polarity and Emotionality). This parallels a number of other societal processes that have previously been found to be pink noise (see Mathiesen et al. 2013). Since our measures are the product of a large number of micro-interactions between deputies with shared histories, the finding that mood spreads over time following a $1/f$ process is not surprising. We note, however, that the parameters α in the yearly series are closer to 2 (1.60 and 1.74, respectively for Polarity and Emotionality), a characteristic of Brownian noise. This can have implications for the choice of estimators in the rest of our analysis, as Brownian noise can be associated with unit root processes (Granger and Joyeux 1980). Accordingly, we tested each time-series for unit root processes. Table 5 in the appendix reports the results. The main measure of annual Polarity used in our empirical analysis below appears to follow a unit root process based on several specifications of both the Dickey-Fuller and KPSS tests.

Figure 1: Emotionality and Polarity in The British House of Commons, 1909-2013

(a) Quarterly



(b) Yearly

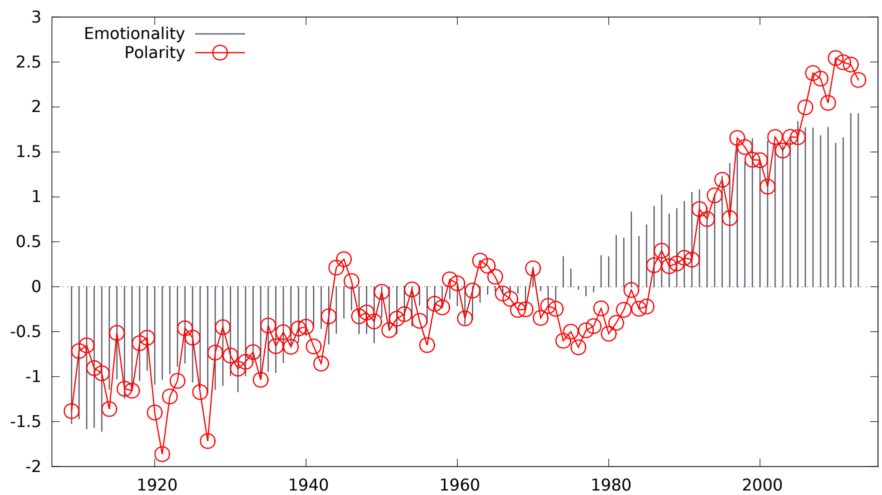
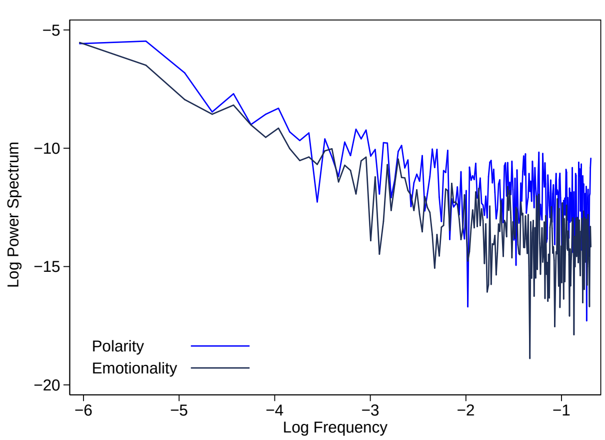
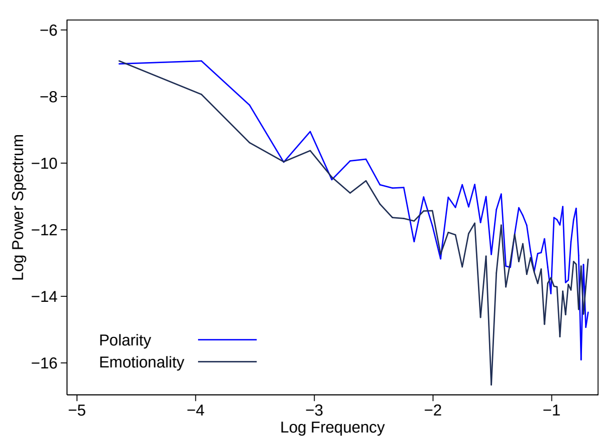


Figure 2: Power Spectral Densities (Log Scale)

(a) Quarterly



(b) Yearly



Emotional Response to Economic Hard Times

Our next objective is to explain changes in emotional polarity over time. In other words, we are considering an equation of motion in discrete time of the type

$$\Delta y_t = f(\Delta y_{t-1}, x_t), \tag{4}$$

where Δy_t is the first-difference operator of our Polarity measure, and x_t is a measure of national events affecting the mood of politicians. Specifically, we expect that the mood in parliament will respond to economic business cycles, that is, the core periodic transitions between economic recessions and expansions. Our theoretical argument is that those business cycles are a fundamental force affecting most spheres of activity in a polity; hence, recessions are likely to trigger a large amount of stress on workers and businesses that should realistically have repercussions in the House of Commons. The mandate of elected politicians is to represent their constituents. We usually think of this representation in terms of positions to be expressed on issues. In our view however, the emotional states of civilians are as likely to be echoed inside political institutions. This can take the form of harsher questions directed at the government in power and heightened discordance during debates over bills and motions. In short, we expect the mood to become negative during recessions, as opposed to periods of economic expansion.

We begin with the computation of autoregressive models including two exogenous components. The first is a simple binary indicator r_t equaling 1 if the British real gross domestic product (GDP) has decreased in a given year, indicating a recession, and 0 otherwise. Although the GDP may not be the most relevant measure to assess the evolution of living conditions (as opposed to per capita variables), public institutions such as the House of Commons or the Bank of England (which was an organ of the government until 1997) routinely employ the growth rate of this measure to assess the state of the economy. In addition to economic cycles, we also account for the occurrence of political cycles caused by the periodicity of democratic institutions. We include a second binary indicator e_t equaling 1 if a general election has been held during a given year, and 0 otherwise. Our empirical model corresponds to:

$$\Delta y_t = \alpha_0 + \sum_{k=1}^{l} \alpha_k \Delta y_{t-k} + \beta_1 r_t + \beta_2 e_t + \varepsilon_t, \tag{5}$$

where l is the lag length. Since the variables r_t and e_t are discrete events, typically lasting one year, we do not model their dynamics. We use a logarithmic and first-difference transformation of the y_t series, which amounts to a growth rate. The transformation produces a stationary series, and all of our models satisfy the usual stability conditions (with autoregressive roots inside the unit circle). We report OLS estimates computed with one and two lags in Table 2, which are equivalent to Arima(1,1,0) and (2,1,0) specifications. The estimated autoregressive parameters are negative, indicating that random shocks to the growth rate of Polarity will lead to oscillatory decays. The estimated coefficient for the Recession variable is negative, which is consistent with our expectation that the mood in parliament responds negatively to economic downturns. The value of the coefficient –0.015 in the first model indicates that a recession is associated with a 0.015 point decrease in the growth rate of polarity in the ongoing year. Put another way, during recessions the original Polarity indicator is 1.5% lower on average. Conversely, election years appear to increase the positivity of the mood by a similar order of magnitude (1.4%). Both estimates are statistically significant at the 95% level. We computed dynamic multipliers (impulse responses) for those two variables to illustrate the longevity of those effects over time (see Figure 3). Unlike the original series, the differenced Polarity has a short memory, and the effect of a recession vanishes after one year, as can be observed in our figures.

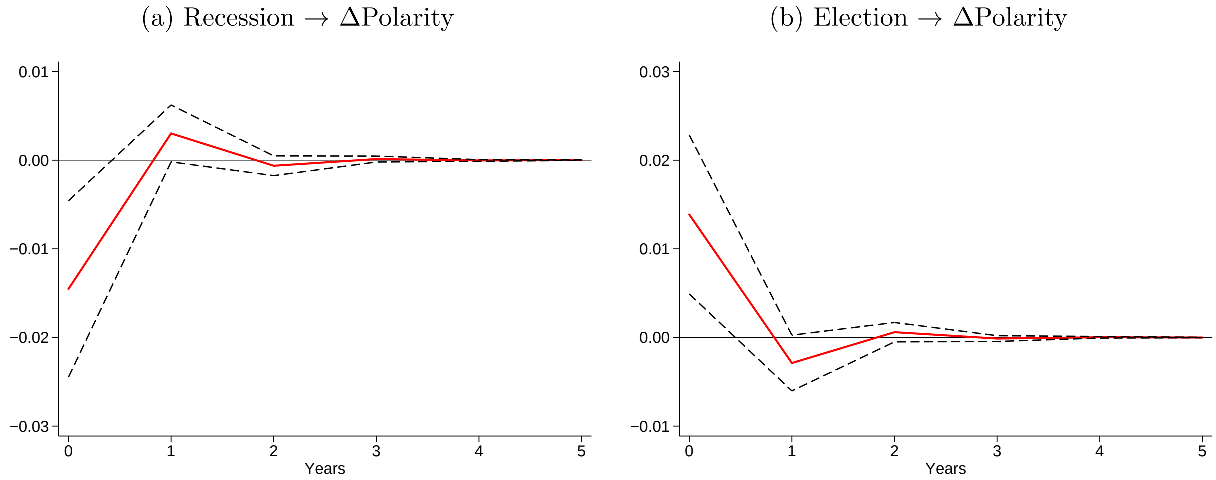
Table 2: Autoregressive Models of Polarity in UK Parliament

Δy_t	Model 1	Model 2
Δy_{t-1}	–0.208	–0.266
	(0.088)	(0.089)
Δy_{t-2}		–0.231
		(0.090)
Recession	–0.015	–0.011
	(0.005)	(0.005)
Election	0.014	0.014
	(0.005)	(0.004)
Intercept	0.002	0.001
	(0.002)	(0.002)
N	103	102
Log-Likelihood	259.337	259.703
AIC	–510.675	–509.406

Time-series autoregressions of Polarity (y_t) in the UK Parliament, with Recession (r_t) and Election (e_t) included as binary exogenous regressors. Standard errors are reported in parentheses.

A limitation of our binary indicator of recession is that it overlooks the broader dynamics affecting the state of a national economy over time. We seek to extend our insights by contrasting our mood indicators with other historical measures. We selected four time series available for the entire period and capturing important aspects of economic conditions: a measure of labour disputes (in number of days lost due to strikes every year), the rate of unemployment, the crude death rate, and a chained measure of the gross domestic product. The series are reported using a heat map in Figure 4. The figure reflects the major transformation that took place during the last century in Britain. The early 20th century was

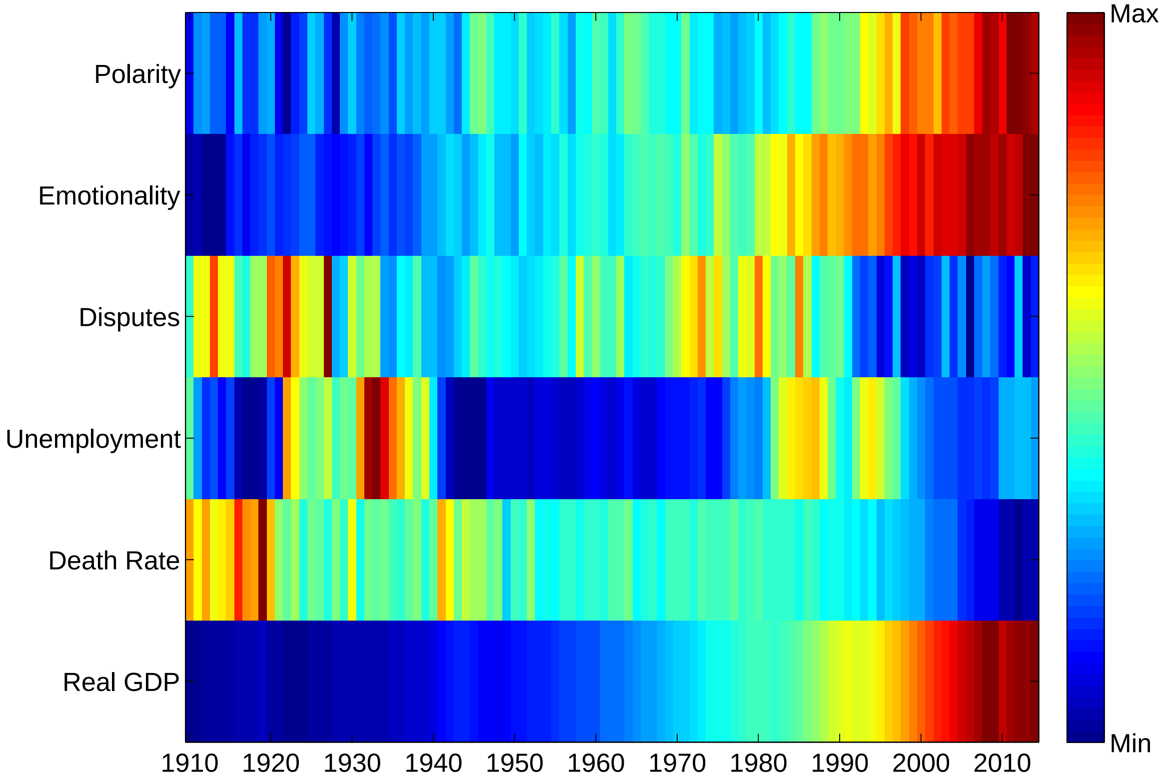
Figure 3: Dynamic Multipliers: Economic and Political Cycles

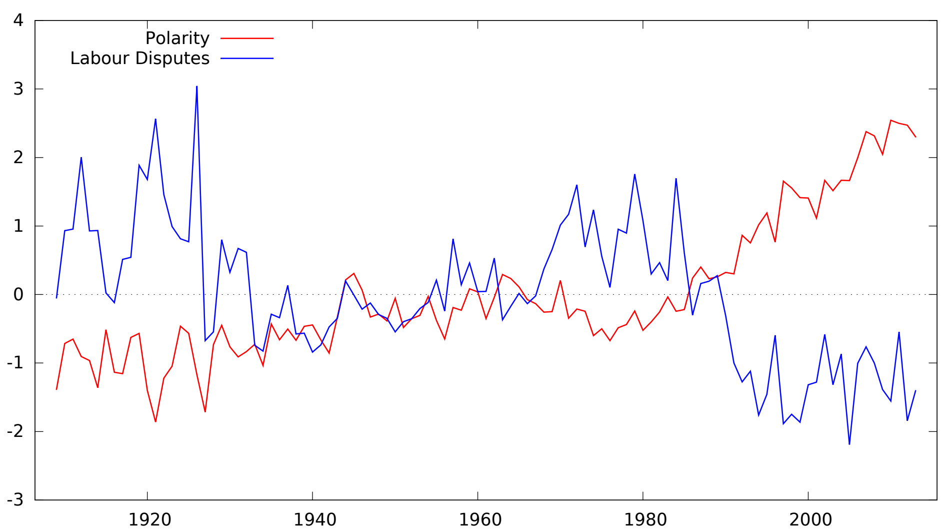


characterized by intense labour disputes and skyrocketing rates of unemployment during the Great Depression of the 1930s. On the other hand, the high crude death rate reflects not only the prevalence of lethal infectious diseases during this period, but also the casualties of the major wars punctuating the first half of that century. The picture changes drastically after the second world war, when most of those measures stabilize. At the same time, the mood of parliamentary debates becomes more positive. A second period of turbulence arises in the 1970s and 1980s, decades encompassing two important recessions. Accordingly, the mood becomes increasingly negative, before experiencing an upward trend. Meanwhile, the increasing trend in polarity over time matches the overall change in the size of the economy, as measured by the real GDP.

However, comparing figures visually does not give a definitive idea of the statistical association between those variables. For this reason we performed pairwise Granger causality tests between Polarity and those four historical measures. Since the series are possibly cointegrated, we used the method proposed by Toda and Yamamoto (1995). As can be seen in Table 3, only one relationship is revealed to be Granger-causal: the one going from labour disputes to emotional polarity. The idea that labour disputes are a relevant indicator to model the mood of politicians makes sense. Strikes and related labour conflicts are disruptive social activities that are fundamentally political, as well as emotionally laden for the actors involved. Moreover, labour conflicts can be readily interpreted on the traditional left-right dimension of political disagreement, to which the major political parties in Britain are historically attached. As a result, the mood associated with those conflicts is likely to transfer into the political arena, and to exacerbate existing ideological divisions in the House of Commons. We emphasize the relationship between the natural logarithm of the number of days lost due to labour disputes and our measure of emotional polarity by superposing

Figure 4: A Century of Change in the United Kingdom





the two time series in the bottom section of Figure 4. The scales have been normalized to facilitate comparisons. At first glance, there seems to be a substantively important fit between those two measures, both appearing negatively related. Moreover, episodes of intense work disputes coincide with the economic hard times of the early period of our sample and the recessions of the 1970 and 1980 decades, in particular.

Table 3: Granger Causality Tests

Cause	Effect	χ^2	d.f.	Pr > χ^2
Labour Disputes	Polarity	9.164	1	0.002
Polarity	Labour Disputes	0.047	1	0.828
Unemployment	Polarity	1.465	1	0.226
Polarity	Unemployment	1.842	1	0.175
Death Rate	Polarity	0.051	1	0.822
Polarity	Death Rate	0.257	1	0.612
GDP	Polarity	0.078	2	0.962
Polarity	GDP	3.114	2	0.211

The table shows tests of the null of Granger non-causality for integrated processes based on augmented pairwise VAR models (Toda and Yamamoto 1995). VAR lengths are selected based on the Schwarz Bayesian Information Criterion.

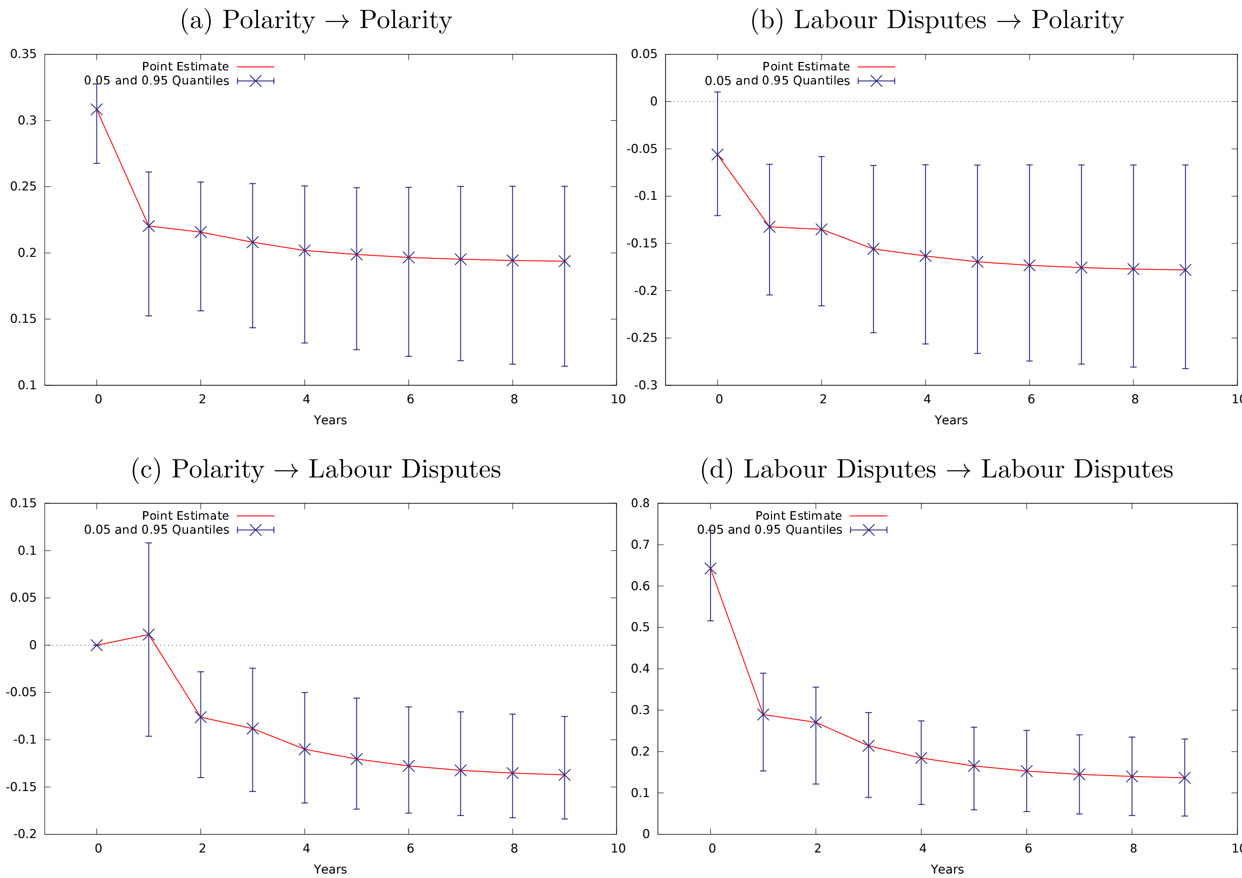
To examine this relationship further, we compute a vector error correction model (VECM) of the system composed of Polarity (y_t) and labour disputes, which we denote x_t. After examination of information criteria, we choose a parsimonious specification with one lag in the first-difference equations of the VECM, which amounts to two lags in the VAR representation of the VECM. Like the Polarity series, our indicator of labour disputes appears to follow a unit root process according to both the augmented Dicker-Fuller and KPSS tests. We report the results as well as the output of the Johansen cointegration test in the Appendix (see Table 5). The system of equations appears cointegrated, which is evidence supporting the idea that the relationship between those series is non-spurious. The long-run relationship between labour disputes and polarity can be summarized by the estimated cointegrating equation

$$y_t = -0.071 - 1.370x_t, \tag{6}$$

meaning that both series are negatively related in the long run. To illustrate the dynamics of the system, we compute the impulse responses of the VECM with bootstrapped error bands and report these in Figure 5. As can be seen, the impulse responses do not converge to zero, which is expected when a system is non-stationary (Lütkepohl 2005). Focusing on the upper-right figure, a one standard deviation shock in the intensity of labour disputes generates an immediate response of –0.05 in our indicator of polarity. This value eventually converges to –0.18 (our variables are normalized, so that this point estimate means a –0.18

standard deviation change in Polarity). Since the polarity measure is a unit root process, the effect of a shock in labour disputes is permanent, and does not decay over time. As can be seen from the four impulses, this is the case for any shocks to the system. We note that the opposite impulse response, from polarity to labour disputes, is also negative. However, as indicated earlier, we are confident only about the existence of a Granger-causal relationship *from* the labour disputes indicator *toward* emotional polarity.

Figure 5: VECM Impulse Responses: Labour Disputes and Emotional Polarity



(a) Polarity → Polarity

(b) Labour Disputes → Polarity

(c) Polarity → Labour Disputes

(d) Labour Disputes → Labour Disputes

Discussion

The method discussed in this paper to measure emotion in political discourse has several benefits. It is relatively simple to use, it can be applied to different domains as long as a sufficiently large corpus exists, and it allows scoring the level of polarity of lemmas on a continuous scale. In fact, it is possible to use this approach to generate lexicons measuring different features of language, not just emotional polarity, as long as researchers can provide polar seed words to serve as a basis for computing word similarities from the vector-space models. Applying this method to the corpus of the British Hansards from 1909 to 2013 leads to insightful findings. In particular, we found that politicians react emotionally to economic cycles, a result supported by models using both a binary indicator of recession and a more politicized indicator of labour conflicts. There are numerous implications to consider in future research. For instance, if politicians react emotionally to economic downturns, it matters to reassess whether those emotions have in turn an impact on crucial decisions made during those periods. The indicators that we proposed in this paper could be used to pursue fine-grained analyses of this type. Moreover, when focusing on the quarterly series, we found some evidence that emotional polarity follows a pink noise process, which is consistent with earlier findings about many social phenomena. However, this conclusion seems affected by the periodicity of the data. We resolved this problem by considering empirical methods that can accommodate integrated processes, but additional research could provide more insights on this particular question. Overall, given the importance of economic legislation and the ripple of impacts it begets on societies, we believe that improving our comprehension of the factors that alter the mood of policy-makers is an important research agenda for the social sciences.

References

Acerbi, Alberto, Vasileios Lampos, Philip Garnett, and R. Alexander Bentley. 2013. "The Expression of Emotions in 20th Century Books." *PLoS ONE* 8(3): 1–6.

Adams, James, Jane Green, and Caitlin Milazzo. 2012. "Has the British Public Depolarized Along With Political Elites? An American Perspective on British Public Opinion." *Comparative Political Studies* 45(4): 507–530.

Arceneaux, Kevin, and Ryan J. Vander Wielen. 2013. "The Effects of Need for Cognition and Need for Affect on Partisan Evaluations." *Political Psychology* 34(1): 23–42.

Asur, Sitaram, and Bernardo A. Huberman. 2010. Predicting the Future With Social Media. In *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. pp. 492–499.

Aue, Anthony, and Michael Gamon. 2005. Customizing Sentiment Classifiers to New Domains: A Case Study. In *Proceedings of Recent Advances in Natural Language Processing (RANLP-2005)*.

Bechara, Antoine, Hanna Damasio, and Antonio R. Damasio. 2000. "Emotion, Decision Making and the Orbitofrontal Cortex." *Cerebral Cortex* 10(3): 295–307.

Bickford, Susan. 2011. "Emotion Talk and Political Judgment." *The Journal of Politics* 73(4): 1025–1037.

Black, Ryan C., Sarah A. Treul, Timothy R. Johnson, and Jerry Goldman. 2011. "Emotions, Oral Arguments, and Supreme Court Decision Making." *The Journal of Politics* 73(2): 572–581.

Bollen, Johan, Huina Mao, and Alberto Pepe. 2011. Modeling Public Mood and Emotion: Twitter Sentiment and Socio-Economic Phenomena. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*. pp. 450–453.

Bollen, Johan, Huina Mao, and Xiaojun Zeng. 2011. "Twitter Mood Predicts the Stock Market." *Journal of Computational Science* 2(1): 1–8.

Brader, Ted. 2005. "Striking a Responsive Chord: How Political Ads Motivate and Persuade Voters by Appealing to Emotions." *American Journal of Political Science* 49(2): 388–405.

Brader, Ted. 2006. *Campaigning for Hearts and Minds: How Emotional Appeals in Political Ads Work*. Chicago: The University of Chicago Press.

Brader, Ted. 2011. "The Political Relevance of Emotions: 'Reassessing' Revisited." *Political Psychology* 32(2): 337–346.

Camerer, Colin, George Loewenstein, and Drazen Prelec. 2005. "Neuroeconomics: How Neuroscience Can Inform Economics." *Journal of Economic Literature* 43: 9–64.

Caplin, Andrew, and John Leahy. 2001. "Psychological Expected Utility Theory and Anticipatory Feelings." *The Quarterly Journal of Economics* 116(1): 55–79.

Civettini, Andrew J. W., and David P. Redlawsk. 2009. "Voters, Emotions, and Memory." *Political Psychology* 30(1): 125–151.

Cowie, Roddy, and Randolph R. Cornelius. 2003. "Describing the Emotional States That Are Expressed in Speech." *Speech Communication* 40(1-2): 5–32.

Crawford, Neta C. 2000. "The Passion of World Politics: Propositions on Emotion and Emotional Relationships." *International Security* 24(4): 116–156.

Dawes, Christopher T., Peter John Loewen, Darren Schreiber, Alan N. Simmons, Taru Flagan, Richard McElreath, Scott E. Bokemper, James H. Fowler, and Martin P. Paulus. 2012. "Neural Basis of Egalitarian Behavior." *Proceedings of the National Academy of Sciences* 109(17): 6479–6483.

DeWall, C. Nathan, Richard S. Jr. Pond, W. Keith Campbell, and Jean M. Twenge. 2011. "Tuning in to Psychological Change: Linguistic Markers of Psychological Traits and Emotions Over Time in Popular U.S. Song Lyrics." *Psychology of Aesthetics, Creativity, and the Arts* 5(3): 200207.

Dodds, Peter Sheridan, Kameron Decker Harris, Isabel M. Kloumann, Catherine A. Bliss, and Christopher M. Danforth. 2011. "Temporal Patterns of Happiness and Information in a Global Social Network: Hedonometrics and Twitter." *PLoS ONE* 6(12): 1–26.

Esuli, Andrea, and Fabrizio Sebastiani. 2006. SentiWordNet: A Publicly Available Lexical Resource for Opinion Mining. In *In Proceedings of the 5th Conference on Language Resources and Evaluation (LREC06).* pp. 417–422.

Glimcher, Paul W., Michael C. Dorris, and Hannah M. Bayer. 2005. "Physiological Utility Theory and the Neuroeconomics of Choice." *Games and Economic Behavior* 52(2): 213–256.

Granger, Clive W. J., and Roselyne Joyeux. 1980. "An Introduction to Long-Memory Time Series Models and Fractional Differencing." *Journal of Time Series Analysis* 1(1): 15–29.

Hills, Sally, and Ryland Thomas. 2014. "The UK Recession in Context—What Do Three Centuries of Data Tell Us?" In *Research and analysis: The UK recession in context*. The Bank of England Quarterly Bulletin 2010 Q4 pp. 277–291.

Hsu, Ming, Meghana Bhatt, Ralph Adolphs, Daniel Tranel, and Colin F. Camerer. 2005. "Neural Systems Responding to Degrees of Uncertainty in Human Decision-Making." *Science* 310(5754): 1680–1683.

Hughes, James M., Nicholas J. Foti, David C. Krakauer, and Daniel N. Rockmore. 2012. "Quantitative Patterns of Stylistic Influence in the Evolution of Literature." *Proceedings of the National Academy of Sciences* 109(20): 7682–7686.

Ishizuka, Mitsuru, Alena Neviarouskaya, and Mostafa Al Masum Shaikh. 2012. "Textual Affect Sensing and Affective Communication." *International Journal of Cognitive Informatics and Natural Intelligence* 6(4): 81–102.

Jaywant, Abhishek, and Marc D. Pell. 2012. "Categorical Processing of Negative Emotions from Speech Prosody." *Speech Communication* 54(1): 1–10.

Johnson, Christopher, Parul Shukla, and Shilpa Shukla. 2014. "On Classifying the Political Sentiment of Tweets.".

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47(2): 263–292.

Kahneman, Daniel, and Amos Tversky. 1984. "Choices, Values, and Frames." *American Psychologist* 39(4): 341–350.

Kaufman, Bruce E. 1999. "Emotional Arousal as a Source of Bounded Rationality." *Journal of Economic Behavior & Organization* 38(2): 135–144.

Krueger, Frank, Aron K. Barbey, Kevin McCabe, Maren Strenziok, Giovanna Zamboni, Jeffrey Solomon, Vanessa Raymont, and Jordan Grafman. 2009. "The Neural Bases of Key Competencies of Emotional Intelligence." *Proceedings of the National Academy of Sciences* 106(52): 22486–22491.

Ladd, Jonathan McDonald, and Gabriel S. Lenz. 2008. "Reassessing the Role of Anxiety in Vote Choice." *Political Psychology* 29(2): 275–296.

Lansdall-Welfare, Thomas, Vasileios Lampos, and Nello Cristianini. 2012. Effects of the Recession on Public Mood in the UK. In *Proceedings of the 21st International Conference Companion on World Wide Web*. WWW '12 Companion New York: ACM pp. 1221–1226.

Lodge, Milton, and Charles S. Taber. 2005. "The Automaticity of Affect for Political Leaders, Groups, and Issues: An Experimental Test of the Hot Cognition Hypothesis." *Political Psychology* 26(3): 455–482.

Loewenstein, George. 2000. "Emotions in Economic Theory and Economic Behavior." *The American Economic Review* 90(2): 426–432.

Loomes, Graham, and Robert Sugden. 1982. "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty." *Economic Journal* 92(368): 805–824.

Lütkepohl, Helmut. 2005. *New Introduction to Multiple Time Series Analysis.* Berlin: Springer.

Maas, Andrew L., Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, and Christopher Potts. 2011. Learning Word Vectors for Sentiment Analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1 (HLT11).* Stroudsburg, PA, USA: Association for Computational Linguistics pp. 142–150.

Manning, Christopher D., Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. The Stanford CoreNLP Natural Language Processing Toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations.* pp. 55–60.

Marcus, George E. 2002. *The Sentimental Citizen: Emotion in Democratic Politics.* University Park: Penn State University Press.

Marcus, George E., W. Russell Neuman, and Michael MacKuen. 2000. *Affective Intelligence and Political Judgment.* Chicago: University of Chicago Press.

Mathiesen, Joachim, Luiza Angheluta, Peter T. H. Ahlgren, and Mogens H. Jensen. 2013. "Excitable Human Dynamics Driven by Extrinsic Events in Massive Communities." *Proceedings of the National Academy of Sciences* 110(43): 17259–17262.

McDermott, Rose. 2004. "The Feeling of Rationality: The Meaning of Neuroscientific Advances for Political Science." *Perspectives on Politics* 2(4): 691–706.

Mercer, Jonathan. 2005. "Prospect Theory and Political Science." *Annual Review of Political Science* 8: 1–21.

Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, The Google Books Team, Joseph P. Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, Steven Pinker, Martin A. Nowak, and Erez Lieberman Aiden. 2011.

"Quantitative Analysis of Culture Using Millions of Digitized Books." *Science* 331(6014): 176–182.

Miller, Patrick R. 2011. "The Emotional Citizen: Emotion as a Function of Political Sophistication." *Political Psychology* 32(4): 575–600.

Mohammad, Saif M., Svetlana Kiritchenko, and Xiaodan Zhu. 2013. NRC-Canada: Building the State-of-the-Art in Sentiment Analysis of Tweets. In *Proceedings of the Seventh International Workshop on Semantic Evaluation Exercises (SemEval-2013)*. Atlanta, Georgia, USA: .

Mohammad, Saif M., Xiaodan Zhu, Svetlana Kiritchenko, and Joel Martin. 2014. "Sentiment, Emotion, Purpose, and Style in Electoral Tweets." *Information Processing & Management* In Press.

Neuman, W. Russell, George E. Marcus, Ann N. Crigler, and Michael MacKuen, eds. 2007. *The Affect Effect: Dynamics of Emotion in Political Thinking and Behavior*. Chicago: The University of Chicago Press.

Nguyen, Thin, Dinh Phung, Bo Dao, Svetha Venkatesh, and Michael Berk. 2014. "Affective and Content Analysis of Online Depression Communities." *IEEE Transactions on Affective Computing* 5(3): 217–226.

Pagano, Sabrina J., and Yuen J. Huo. 2007. "The Role of Moral Emotions in Predicting Support for Political Actions in Post-War Iraq." *Political Psychology* 28(2): 227–255.

Pennington, Jeffrey, Richard Socher, and Christopher D. Manning. 2014. Glove: Global Vectors for Word Representation,. In *Conference on Empirical Methods in Natural Language Processing (EMNLP 2014)*.

Petersen, Michael Bang, Daniel Sznycer, Leda Cosmides, and John Tooby. 2012. "Who Deserves Help? Evolutionary Psychology, Social Emotions, and Public Opinion about Welfare." *Political Psychology* 33(3): 395–418.

Picard, Rosalind. 1997. *Affective Computing*. Cambridge: The MIT Press.

Qiu, Guang, Bing Liu, Jiajun Bu, and Chun Chen. 2009. Expanding Domain Sentiment Lexicon through Double Propagation. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI -2009)*.

Rick, Scott, and George Loewenstein. 2008. "The Role of Emotion in Economic Behavior." In *Handbook of Emotions*, ed. Michael Lewis, Jeannette M. Haviland-Jones, and Lisa Feldman Barrett. New York: Guilford Press pp. 138–156.

Schuller, Björn, and Anton Batliner. 2014. *Computational Paralinguistics: Emotion, Affect and Personality in Speech and Language Processing.* Hoboken, NJ: John Wiley & Sons.

Simon, Herbert A. 1967. "Motivational and Emotional Controls of Cognition." *Psychological Review* 74(1): 29–39.

Small, Deborah A., and Jennifer S. Lerner. 2008. "Emotional Policy: Personal Sadness and Anger Shape Judgments about a Welfare Case." *Political Psychology* 29(2): 149–168.

Takahashi, Hidehiko. 2013. "Molecular Neuroimaging of Emotional Decision-Making." *Neuroscience Research* 75(4): 269–274.

Toda, Hiro Y, and Taku Yamamoto. 1995. "Statistical Inferences in Vector Autoregressions with Possibly Integrated Processes." *Journal of Econometrics* 66(1-2): 225–250.

Turney, Peter D., and Michael L. Littman. 2003. "Measuring Praise and Criticism: Inference of Semantic Orientation from Association." *ACM Trans. Inf. Syst.* 21(4): 315–346.

United Kingdom. 2010. The Official Report. Technical report House of Commons Information Office, Factsheet G17 General Series.

Wiebe, Janyce, Theresa Wilson, and Claire Cardie. 2005. "Annotating Expressions of Opinions and Emotions in Language." *Language Resources and Evaluation* 39(2-3): 165–210.

Appendix

Preprocessing of the Text Data

Our corpus comes from two different sources. The Hansards for the period 1936-2013 were formatted using a markup language by the team of the Dilipad project (`http://dilipad.history.ac.uk/`), using the files previously processed by the independent project *They Work For You* (`http://www.theyworkforyou.com/`). We processed the remaining Hansards from 1909 to 1935 using the digitized archives available on the website of the UK parliament (`http://www.hansard-archive.parliament.uk/`). We cleaned the early files for structural mistakes such as broken sentences or irregular spacing, using a custom Python script. We also duplicated the years from 1936 to 1938 to confirm that the final corpora collected from both sources are virtually identical. A few volumes are missing from the online archives, and the final values of our quarterly measures were linearly interpolated. The digitization is of good quality, although the corpus is not entirely free of typographic errors, likely caused by the optimal character recognition routines that were used to create the archive. Foreign words with accentuation are the most problematic and were excluded from our analysis.

We split the text into sentences and words using the Apache OpenNLP sentence detector and the tokenizer from the same library. Lemmatization and part-of-speech (PoS) tagging were performed using the Stanford CoreNLP software (Manning et al. 2014). Both pieces of software are written in Java and available freely to researchers (respectively at `https://opennlp.apache.org/` and `http://nlp.stanford.edu/software/corenlp.shtml`). We computed the vector space model using the *GloVe* algorithm, the source code of which, written in C, is also available to researchers (`http://nlp.stanford.edu/projects/glove/`). As mentioned in the text, we created a model with 300 dimensions, considering word-word co-occurrences up to 15 words to the left and to the right. Our vocabulary is restricted to lemma–PoS pairs occurring 10 times or more, and that are present in both the first half of the sample (1909-1961) and the second half (1962-2013). As a result, period-specific expressions are in effect disregarded, which prevents them from driving our indicators.

A sensitive part of our approach consists of selecting the initial seed words that will enable the creation of a polarity lexicon. Table 4 lists all the seeds we used for this study. To select the seeds, we started with the most common words with an unambiguous positive or negative polarity. Those words (`good`, `love`, and `happy` for the positive pole; `bad`, `hate`, and `sad` for the negative pole) were searched recursively for synonyms using two open-source dictionaries and thesauruses, the The Collaborative International Dictionary of English and WordNet 3.0. We considered words as seeds only if they respected the following rules:

1. Seed words need to have an unambiguous polarity, which means that multiple meanings

of the same word used as the same PoS must not have opposite polarities.

2. Seed words cannot be the name of an institution, parliamentary procedure or political topic (excluded are words such as 'war', 'dispute', 'unemployment', and so forth).

3. Seed words have to be general, basic, and common words of everyday language.

Notice that since we distinguish between parts of speech, we may still include a word that has opposite polarities when used as a verb as opposed to a noun, for instance, by including the orientation-relevant word/PoS pair. Once a list of potential seeds was established, we reduced its size to 100 for both positive and negative items by selecting the most frequent in the English language. To have an estimate of their frequency, we queried the Google Ngram database for the period 1909-2008. We report the relative frequencies along with each seed in Table 4.

We took a number of additional steps to prevent the contamination of our measures by the idiosyncrasies of parliamentary life. First, we removed all non-informative expressions used as formal epithets to address members of parliament, which are used frequently in the Hansards. Those include expressions such as "My Honourable Friend", "The Right Honourable Member", and so forth. Members of parliament are required to used them by protocol, hence they cannot be associated with emotions. Virtually all instances of the word `honourable` left in the final corpus are used in an actual sentence, rather than being a form of speech required by the decorum of the House of Commons. We also removed all indicators of nationalities (e.g. `Americans`, `Czecho-slovakian`, and so forth) as they should be theoretically neutral. Finally, we removed all digits and proper nouns from our procedure when computing polarity lexicons. All the scripts required for those steps and the computation of our measures were written in R, Python or C, with the language chosen to optimize speed.

Data Sources for Economic Variables

1. The Labour Disputes indicator measures the number of days lost due to labour disputes in the United Kingdom per year. The series is taken from the Labour Market Statistics Dataset published by the Office for National Statistics (ONS) of the United Kingdom, released on April 17, 2015 (`http://www.ons.gov.uk/ons/rel/lms/labour-market-statistics/april-2015/dataset--labour-market-statistics.html`). It corresponds to the series labeled BBFW.

2. Real Gross Domestic Product (GDP) is a chained volume measure in millions of pounds with the reference year 2006, as compiled in the dataset "Three Centuries of Data" published by the Bank of England (`http://www.bankofengland.co.uk/research/`

Table 4: Polarity Seed Lemmas by Google Ngram Relative Frequencies

Positive						Negative					
Lemma	P-o-S	Frequency	Lemma	P-o-S	Frequency	Lemma	P-o-S	Frequency	Lemma	P-o-S	Frequency
well	adv.	0.715	wonderful	adj.	0.031	problem	noun	0.223	hate	verb	0.017
good	adj.	0.555	friendly	adj.	0.030	death	noun	0.216	complaint	noun	0.017
important	adj.	0.337	pleasant	adj.	0.029	difficult	adj.	0.142	painful	adj.	0.017
best	adj.	0.261	creative	adj.	0.028	loss	noun	0.116	worry	verb	0.017
better	adj.	0.243	worthy	adj.	0.027	bad	adj.	0.089	unfortunate	adj.	0.017
true	adj.	0.234	friendship	noun	0.026	fear	noun	0.085	neglect	verb	0.016
love	verb	0.205	sympathy	noun	0.026	failure	noun	0.079	prejudice	noun	0.015
able	adj.	0.192	nice	adj.	0.025	enemy	noun	0.071	disaster	noun	0.015
help	verb	0.188	honour	noun	0.025	wrong	adj.	0.068	distress	noun	0.015
strong	adj.	0.147	comfort	noun	0.025	difficulty	noun	0.068	hatred	noun	0.014
solution	noun	0.138	honest	adj.	0.024	pain	noun	0.065	tragic	adj.	0.014
importance	noun	0.129	genuine	adj.	0.024	ill	adj.	0.063	shame	noun	0.014
respect	noun	0.123	healthy	adj.	0.024	risk	noun	0.062	breach	noun	0.013
truth	noun	0.115	intelligent	adj.	0.023	danger	noun	0.060	contempt	noun	0.013
strength	noun	0.101	welcome	adj.	0.023	error	noun	0.057	unhappy	adj.	0.013
effective	adj.	0.099	helpful	adj.	0.023	evil	adj.	0.054	frightened	adj.	0.013
success	noun	0.099	encourage	verb	0.022	criticism	noun	0.047	regret	noun	0.013
freedom	noun	0.092	praise	noun	0.022	false	adj.	0.046	corruption	noun	0.013
significant	adj.	0.091	dignity	noun	0.021	weak	adj.	0.041	restriction	noun	0.012
interesting	adj.	0.084	prosperity	noun	0.021	dangerous	adj.	0.041	poorly	adv.	0.011
useful	adj.	0.078	comfortable	adj.	0.020	excess	noun	0.040	fraud	noun	0.010
successful	adj.	0.075	reliable	adj.	0.019	damage	noun	0.040	miserable	adj.	0.010
beautiful	adj.	0.073	succeed	verb	0.019	lose	verb	0.038	stupid	adj.	0.010
appropriate	adj.	0.068	delight	noun	0.019	worse	adj.	0.037	injustice	noun	0.010
fair	adj.	0.067	merit	noun	0.018	afraid	adj.	0.036	ugly	adj.	0.010
happy	adj.	0.059	lovely	adj.	0.018	fail	verb	0.034	wicked	adj.	0.010
perfect	adj.	0.058	splendid	adj.	0.018	sick	adj.	0.033	disadvantage	noun	0.009
gain	verb	0.055	sympathetic	adj.	0.017	unfortunately	adv.	0.030	disappointment	noun	0.009
excellent	adj.	0.053	generous	adj.	0.017	confusion	noun	0.029	unfair	adj.	0.009
superior	adj.	0.051	vigorous	adj.	0.017	burden	noun	0.029	nonsense	noun	0.009
fairly	adv.	0.050	perfection	noun	0.017	anxiety	noun	0.028	ridiculous	adj.	0.009
reasonable	adj.	0.050	appreciate	verb	0.016	terrible	adj.	0.027	undesirable	adj.	0.009
secure	verb	0.049	loving	adj.	0.016	suffer	verb	0.027	imperfect	adj.	0.009
efficiency	noun	0.049	magnificent	adj.	0.016	fault	noun	0.026	harmful	adj.	0.009
valuable	adj.	0.049	integrity	noun	0.015	anxious	adj.	0.026	horrible	adj.	0.009
properly	adv.	0.047	talent	noun	0.015	destroy	verb	0.025	disastrous	adj.	0.008
improvement	noun	0.046	kindly	adv.	0.015	worst	adj.	0.025	unsatisfactory	adj.	0.008
safe	adj.	0.043	fortunately	adv.	0.014	excessive	adj.	0.025	hopeless	adj.	0.008
desirable	adj.	0.039	grateful	adj.	0.014	threat	noun	0.025	complain	verb	0.008
satisfactory	adj.	0.039	glorious	adj.	0.013	mistake	noun	0.025	fearful	adj.	0.008
wise	adj.	0.039	fortunate	adj.	0.013	inferior	adj.	0.023	unjust	adj.	0.008
protect	verb	0.038	clever	adj.	0.012	weakness	noun	0.023	irrelevant	adj.	0.008
truly	adv.	0.036	sincere	adj.	0.012	anger	noun	0.022	corrupt	adj.	0.008
satisfaction	noun	0.036	confident	adj.	0.012	hurt	verb	0.022	unreasonable	adj.	0.008
efficient	adj.	0.035	delightful	adj.	0.012	angry	adj.	0.021	restrict	verb	0.007
joy	noun	0.035	strengthen	verb	0.011	tragedy	noun	0.020	careless	adj.	0.007
improve	verb	0.033	respected	adj.	0.011	abuse	noun	0.020	grim	adj.	0.007
enjoy	verb	0.032	admirable	adj.	0.010	inadequate	adj.	0.020	wretched	adj.	0.007
happiness	noun	0.031	smart	adj.	0.009	sad	adj.	0.020	discomfort	noun	0.007
glad	adj.	0.031	satisfying	adj.	0.009	harm	verb	0.020	brutal	adj.	0.006

The table shows seed lemmas/part-of-speech pairs used to create the domain-specific lexicon, along with the Google Ngram frequency of each lemma, per thousand words, averaged over each year between 1909 and 2008.

`Documents/onebank/threecenturies.xlsx`) retrieved on February 1, 2015. The series is described in detail in Hills and Thomas (2014). It covers the period 1909-2009. In order to preserve the same reference year, the values for 2010 to 2013 were extrapolated based on the annual rates of growth of the chained GDP in volumes (code ABMI) from the ONS's United Kingdom Economic Accounts Time Series Dataset (Q4 2014) published on March 31, 2015.

3. The Crude Death Rate is measured by counting the total number of deaths reported

for England and Wales in the ONS's 20th Century Mortality Data Files and the 21st Century Mortality Files, Deaths Dataset, 2001-2013 (released on October 30, 2014). The crude death rate is obtained by dividing the death count by mid-year population estimates (in thousands) for England and Wales contained in the same Mortality Files (Population Datasets).

4. The Unemployment rate is taken from the "Three Centuries of Data" dataset for the period 1909-2009. Data for 2010 to 2013 come from the above-mentioned ONS's Labour Market Statistics Dataset, series BCJE, for consistency with the source used in the main dataset.

5. The Recession variable is measured by coding values as 1 when the rate of growth of the real GDP variable is negative on a given year, 0 otherwise.

6. The Election variable equals 1 if at least one general election was held in a given year (some years had more than one general election), and 0 otherwise.

Figure 6: Autocorrelation Functions

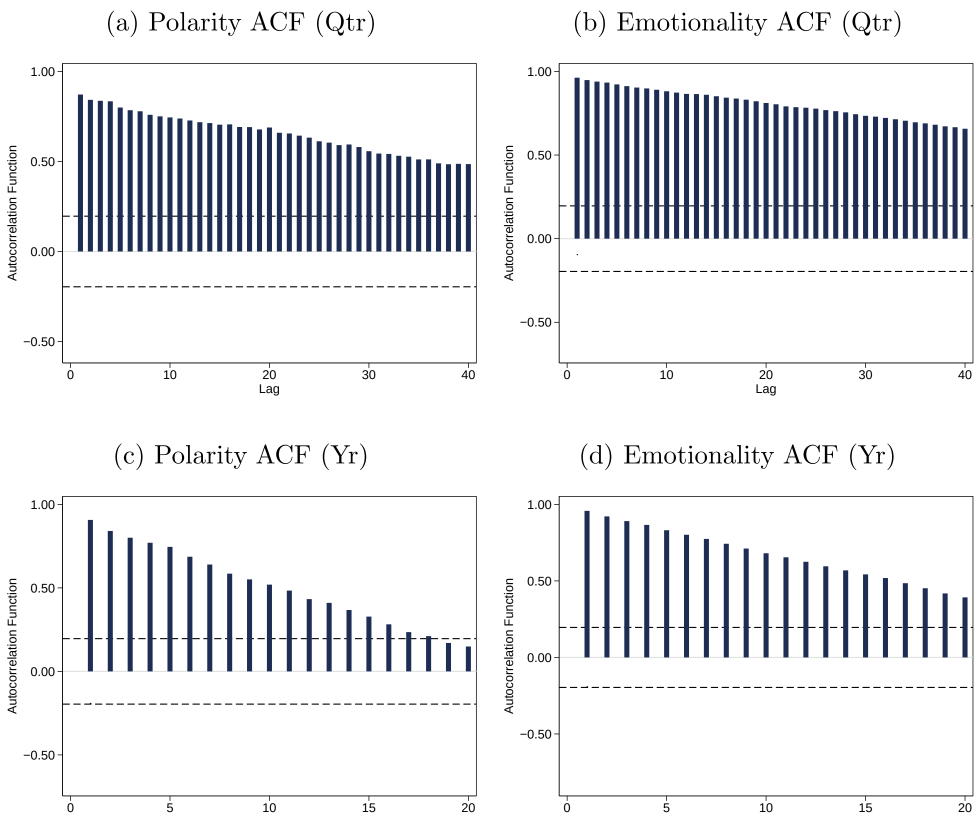
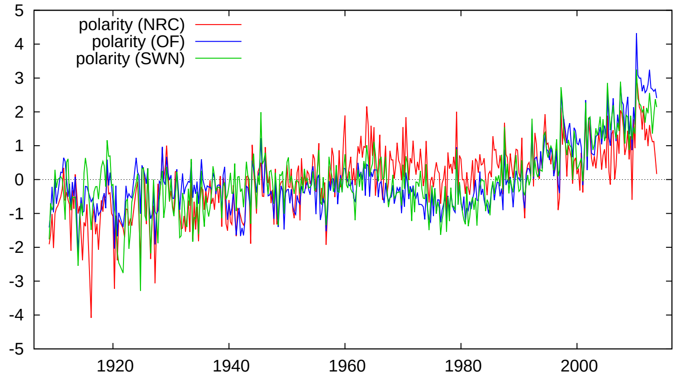


Figure 7: Emotional Polarity in the British House of Commons: Alternative Polarity Lexicons



Alternative measures of emotional polarity computed using three popular affect computing lexicons that are not specific to the domain of parliamentary debates: Mohammad, Kiritchenko, and Zhu (2013)'s NRC lexicon, OpinionFinder (OF) and SentiWordNet (SWN).

Table 5: Unit Root and Cointegration Tests

	Unit Root Tests					
	ADF (H_0: Non-Stationarity)			KPSS (H_0: Stationarity)		
Variable	Lag Order	No Trend	Trend	Lag Order	No Trend	Trend
Polarity	2	0.972	0.683	2	2.67	0.493
Emotionality	2	0.960	0.501	2	1.28	0.288
Labour Disputes	2	0.223	0.174	2	3.41	0.612
Unemployment	2	0.072	0.236	2	0.22	0.219
Crude Death Rate	2	0.594	0.197	2	2.41	0.212
Real GDP	2	0.999	0.943	2	3.29	0.819
5 % C.-V. (KPSS)					0.463	0.146
	Johansen Cointegration Test (Polarity, Labour Disputes)					
Cointegration Rank	Lag Order	Trace Statistic	p-Value	Lag Order	Trace Statistic	p-Value
H_0: Rank > 0	1	40.912	0.000	2	19.582	0.011
H_0: Rank > 1	1	1.449	0.235	2	0.306	0.585

The upper panel reports the MacKinnon approximate p-values for the Augmented Dickey-Fuller (ADF) unit root tests of the null of a unit root process, and the KPSS test statistic of the null of stationarity along with the 5% critical values. The bottom panel reports trace statistics of the Johansen cointegration rank tests. All variables have been normalized. The Labour Disputes series has been previously transformed on the natural log scale.