From Rivalry to Special Relationship:
How to Understand and Forecast Conflict and Cooperation with Legislative Speeches on Anglo-American Relations, 1803-1920

Abstract

Understanding how rivalries end, alliances form, and international cooperation emerges between states are central themes of international relations research. While a hostile relationship between states tends to perpetuate itself, cycles of violence have been reversed into cooperation. An important example of former enemies evolving into alliance partners was the Anglo-American relationship. Relations were generally hostile in the 19th century but began improving at the beginning of the 20th. This case offers a unique opportunity to track the motivations for these foreign policy changes because both the US and the UK maintained nearly verbatim records of legislative debates for this period. We propose to merge new developments in topic and position coding from natural language processing with information on legislature constituencies in both states. This will give us a fine-grained view of how individual policy makers with relatively well-defined constituencies viewed the other state, revealing who was most hostile during the rivalry, who was most friendly during the rapprochement, and how and when the distribution of these positions changed across time.

Michael P. Colaresi
Michigan State University

Benjamin O. Fordham
Binghamton University (SUNY)

Understanding how international rivalries end, alliances form, and cooperation emerges are among the most important explanatory tasks in the field of international relations. Previous research on these topics has been hampered by the need to rely on limited and highly aggregated indicators of the state of relations between states. Measuring the state of relations among states under "normal" conditions that do not draw much media attention has proven difficult, as has linking these everyday shifts to particular domestic political entrepreneurs within the states.

In this paper, we will outline a research design for overcoming these limitations using data drawn from legislative speeches. We propose to focus on relations between Britain and the United States during the 19th and early 20th centuries. During this period, Anglo-American relations moved from rivalry toward the extraordinary level of cooperation they have sustained since World War II. Both the British Parliament and the United States Congress have long maintained nearly verbatim records of their proceedings. Recent developments in topic and position coding from natural language processing present an opportunity to use these records to track the transformation in relations between the two states in a fine-grained way.

Collecting these data is a substantial undertaking. Completing it will certainly require more time and probably additional funding for research assistance and technical advice. At present, our efforts remain preliminary and incomplete. We hope this paper will elicit comments and suggestions that can improve our research design as well as informing our theoretical understanding of the processes that shaped the changing relationship. In the end, we hope to provide a better theoretical and empirical understanding of the way rivalries end, an improved historical understanding of Anglo-American relations, and a research design that can be applied to other international rivalries.

Previous Work Using Legislative Speech as Data

We propose to leverage new developments in Natural Language Processing (NLP) and Bayesian Time Series that help can help us extract political meaning from legislative speech in the US and UK. It is our hope that in the future this workflow will prove useful in additional contexts as well. Our research design builds on previous text-as-data approaches using legislatives speech found in the political science literature (Quinn, Monroe, Colaresi, Crespin, and Radev 2010 and Monroe, Colaresi, Quinn and Crespin 2009) as well as in the computer science community (Gerrish and Blei 2011, 2012).

This previous work on systematically extracting meaning and signals from political speech has shown significant promise. Quinn, Monroe, Colaresi, Crespin, and Radev (2010) develop and apply an unsupervised dynamic topic model of US Congressional speeches covering the time period 1997-2004. The project utilized a database of 118,000 speeches that included 70,000 words from the Congressional Record. The model produced a set of topics and that organizes speeches into categories. Key word lists, estimated with the underlying parameters of the model, identify the political content of the topics. With these, attention to issues and topics can be measured across time. In a related work, these authors found that pre-processing legislative speech with a topic model allowed them to uncover partisan frames within specific issues that would be missed if corpora as a whole where analyzed (Monroe, Colaresi, and Quinn 2009). Fader, Radev, Crespin, Monroe, Quinn and Colaresi (2007) utilized the hierarchical information

in legislative speech within an algorithm similar to Google's PageRank to identify the influential members of the Senate by estimating lexical centrality across legislators. Grimmer (2010) and Grimmer, Messing and Westwood (2012) use a distinct corpora of political press releases to topic code and test hypotheses related to the personal vote in late 20th century America.

The rich information in legislative speech has not gone unnoticed in computer science. For example, Gerrish and Blei (2011) and Wang, et al (2010) both use Senate speech in the US to predict future roll-call votes. In related work (Gerrish and Blei (2011), these researchers use Latent Dirichelet Allocation (LDA) to topic model Legislative speeches and bills together. Gerrish and Blei (2012) find that many legislators have issue-specific positions that do not conform to one simple aggregated dimension, such as a projection of the left-right spectrum. Instead, by tracking how legislators speak on specific issues, where both the legislators' positions and the issues are identified through modeling the text, out-of-sample prediction of roll-call votes improve.

Our Approach

Several of these previous approaches attempt similar inferences to our plans. Tracking individual and partisan frames and positions across time to test hypotheses on the domestic politics of the Anglo-American rivalry is similar to the work of Fader, et al (2007) on influential members of Congress, Monroe, Colaresi, and Quinn (2009) on partisan frames and Grimmer, Messing and Westwood (2012) on the personal vote. Similarly, our attempt to predict conflictual and cooperative events, out-of-sample, using legislative speech has much in common with Gerrish and Blei's (2011 and 2012) goals in using a dimensional reduction of legislative speech to forecast dichotomous votes.

Three critical differences in our approach, as compared to previous work with legislative speech, is that we will analyze (1) legislative speech in more than one legislature, (2) over a much longer time frame and (3) focus on foreign policy and national security content. By analyzing both UK and US legislative speeches, we open up the previously unexplored possibility of identifying the interactions between frames and positions across the Atlantic. We can track whether issues and positions begin in the US Senate and then transfer to the House of Commons, or vice versa. Similarly, we can test hypotheses related to the mechanisms by which cooperation spreads. While it would be possible using only information from the US legislature to analyze whether the British Second Reform Act in 1867 altered political positions on the rivalry in Washington, we could not rule out the possibility that any relationship was spurious. Instead, it could be that changes in UK positions that predated the Second Reform Act began to improve Anglo-American relations. Without looking at US and UK positions and frames together, this alternative mechanism could not be empirically investigated.

The time-scale of our project is the second key distinction. Instead of 7 years as in Monroe, Colaresi and Quinn (2009), we are currently processing over a century of speeches in the UK, beginning in 1809 and continuing to 1920. Historical data on the US Congress is being processed beginning in 1833 with the Congressional Globe, accessed through the HathiTrust Digital Library and continuing until 1920. The goal for our project is to have nearly 100 years of daily legislative speech coverage in the US and the UK. In the future, there is no reason why this collection cannot continue from 1920 to the present. Burt Monroe, Ellen Lust, Tarek Masoud,

Kevin Quinn and Michael Colaresi have been collecting legislative speech from across the Middle East and North Africa region. This experience with this project helps to illuminate how multiple-legislature data can be stored and analyzed efficiently.

A third unique component of our project is that we intend to focus first on the Anglo-American relationship as opposed to summarizing the dimensionality of a political space as a whole (Gerrish and Blei 2012). There are several reasons why this focus is useful. First, while the infrastructure we will build-out in the course of the project (see below), including a publicly available and machine-readable corpora with date, debate title, political party and constituency tagging may be useful across a large array of problems, the aims of our project are to increase our understanding of practical political problems in international relations. Box (1976, 1980) argues that three steps need to be iterated repeatedly, building models, computing the parameters of interest and forming inferences, and then critiquing the results, in order to solve data-analysis problems. Blei (2014) calls this Box's loop and describes it as Build, Compute, Critique, Repeat. Building and critiquing require both substantive knowledge of the underlying domestic politics of international competition and cooperation as well as ultimately being able to understand where the models fall short, relative to the substantive process under investigation, to improve predictions that can be validated out-of-sample (Schrodt 2014). Without a substantive focus, Box's loop is simply spinning algorithmic wheels. Moreover, in an influential work on utilizing large-scale data to solve real-world problems, such as rivalry escalation, Schrodt (2014) suggests that blending substantive knowledge with careful data analysis within this workflow is of increasing importance to improved predictive performance.

How Data on Legislative Speeches Could Help Us Learn about Rivalry

We are interested in developing and testing both general arguments about the reasons cooperative relations can emerge out of rivalry and in assessing historically specific claims about the Anglo-American rapprochement. There is a substantial literature in international relations addressing the broader theoretical issues. Similarly, diplomatic historians have produced a large body of research on Anglo-American relations. Although these groups of scholars have typically employed different methods and understand their work differently, their explanatory tasks are actually complementary. The data we plan to collect can contribute to both of them.

Quantitative research on interstate rivalry in political science often relies on dichotomous indicators of events in relations between states. The militarized interstate disputes (MID) data are perhaps the most common source of information. Figure 1 provides an example for the Anglo-American rivalry using the dyadic MID data (Maoz 2005). The points represent the presence of absence of a disputes and the line is a lowess smooth with approximate 95 percent confidence interval. While there is important information in this militarized interstate dispute data, in addition to wars, alliances, and other major developments, it unlikely that these zeroes and ones, or summaries of the zeroes and ones like the smoothed expectation in the Figure contain enough information to track rivalry dynamics.

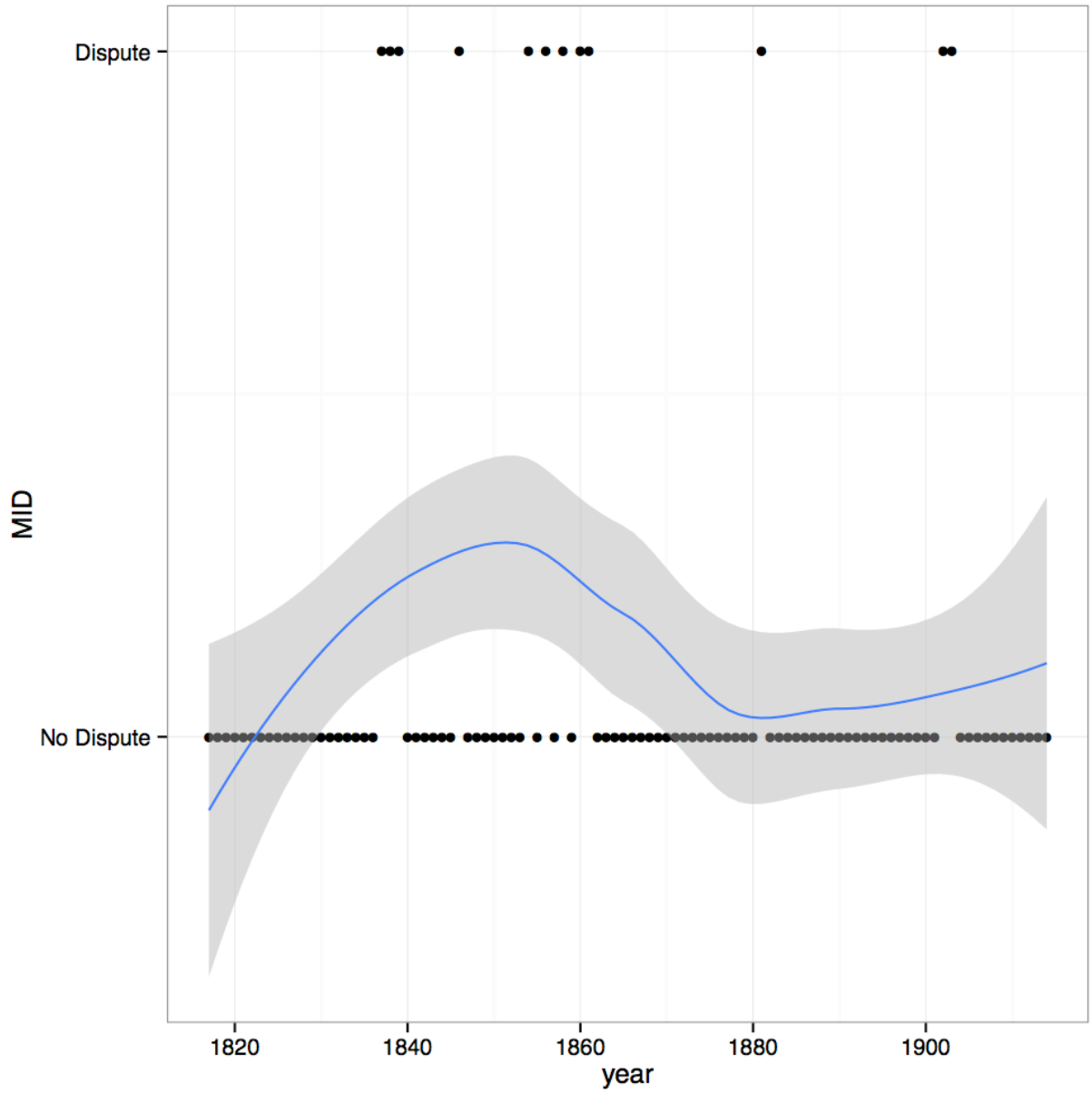

Figure 1: Plot of dyadic militarized interstate disputes between US and UK, 1817-1914, with lowess smooth and approximate 95 percent confidence interval

Richer events data include many less consequential incidents that receive press coverage. Researchers often weight and aggregate these events in order to measure the state of relations over time. These data contain more information than the MID data do, but they still have similar limitations. Like militarized disputes, events reported in the news media are necessarily episodic. Characterizations of the state of relations between states that rely on these events will have difficulty characterizing periods when few newsworthy events are taking place. The difficulty is that hostile or friendly relations persist in the actors' plans and expectations even in "normal" times when the media find little to report. Data on legislative speeches can help overcome this

problem. Legislators certainly comment on major developments, but they also make statements about the underlying state of relations with major rivals even when nothing worth reporting is going on. Indeed, institutional rules about appropriations and other matters may require regular discussions of the state of relations even when there is no precipitating event. Legislative statements can thus provide a more comprehensive view of relations between rivals, including normal times as well as periods of crisis.

Another limitation of most events data used in research on rivalry arises from their use of the state as the principal actor. The events they record are interactions between states. While these events are clearly important, data on them does not reveal much about the processes within states that produced the decisions and events they chronicle. We can observe that a war began, or a treaty was signed, but we do not know which domestic actors favored or opposed the policies and decisions that produced these outcomes. By contrast, legislative statements are linked to a specific speaker, allowing us to track domestic political differences over how to approach the rival.

Historical research on Anglo-American relations relies primarily upon the archival record of state decision-making, supplemented by the public statements and private reflections of other observers. Like the data on dichotomous events in the political science literature, these sources can tell us much about how the rivalry ended, but they also have limits. Not every relevant statement or conversation is recorded for future reference. Even among those that are recorded, not all come to the attention of historians. When scholars think they discern a pattern in the mass of detail that these sources offer, it is difficult to know whether it holds broadly, or is merely an artifact of the process that determined what sources survived. The verbatim record of legislative deliberation does not entirely overcome this problem, because the legislature is not the only arena for policy-relevant deliberation, but it does offer a comprehensive treatment of this aspect of the policymaking process. It thus offers a useful way to test whether patterns that appear in other sources are as general as they seem.

The potential value of data on legislative speeches is easy to see when one considers some of the theoretical arguments advanced to explain how rivalries end in general, and how the Anglo-American rivalry ended in particular. Its usefulness is most obvious in assessing arguments about the underlying mechanisms that drive domestic political actors to support or oppose improving relations with a rival. For example, Colaresi (2004) has argued that political actors known to prefer reconciliation with a rival are in a difficult position because they may be electorally punished if the rival fails to reciprocate offers to cooperate. Schultz (2005) argues that moderate hawks, who have no reputation for wanting to end the rivalry, are actually in a better position to make peace. These arguments apply to the Anglo-American rivalry because Southern Democrats in the United States and Liberals in Britain were known to prefer improved relations. Hypotheses about the behavior of hawks and doves under different circumstances could be tested using data on legislative speeches.

Another widely discussed pathway toward ending an interstate rivalry concerns the impact of domestic institutional change. Political science research suggests that democratization should inhibit the development of an interstate rivalry or hasten its end (Prins and Daxecker 2008; Hensel, Goertz, and Diehl 2000). Historians of the Anglo-American rivalry have argued that the

Reform Bills democratizing British institutions also contributed to friendlier relations with the United States, mainly by diminishing the influence of the aristocracy on British policy and removing one basis for American suspicion of Britain (e.g., Perkins 1968, 10; Allen 1955, 114). These claims contrast with arguments about the durability of enemy images formed during a rivalry (e.g., Dreyer 2010; Kuenne 1989; Thies 2001). These rest on individual psychological processes rather than institutional incentives. Data on legislative speeches could be used to test whether changes in British institutions indeed led sitting politicians in Britain and the United States to change their rhetoric, or whether the effect of these institutions had to wait for the arrival of new legislators.

Because all of these arguments rest on domestic political processes within the rival states, the data we seek to gather have immediate and obvious uses. Indeed, these arguments cannot be tested using conventional data that takes the state or dyad as its unit of analysis. Of course, not all arguments about the end of rivalry and the emergence of cooperation focus on domestic political processes. Many others focus instead on international events and conditions. Systemic shocks, such as major wars and periods of systemic change are associated with the end of interstate rivalries (e.g., Colaresi 2001; Goertz and Diehl 1995; Thies 2001), as are the failure of coercive conflict management strategies (Goertz, Jones, and Diehl 2005), instances of major conflict between the rivals (Morey 2011), the emergence of other rivals who are more threatening (Akcinaroglu, Radziszewski, and Diehl 2013; Bennett 1996), economic downturns (Bennett and Nordstrom 2000) and even events like natural disasters (Akcinaroglu, DiCicco, and Radziszewski 2011). Shocks of this sort have also played a role in historical accounts of the Anglo-American rivalry. The increasing relative power of the United States (e.g., Perkins 1968, 9) and instances of nearly militarized conflict such as the Venezuelan Boundary Crisis of 1895-6 are commonly held to have contributed to improved relations over time (e.g., Campbell 1974, 175-88; Temperley 2002, 76-8; Dunning 1914, 311-7).

While arguments like these can indeed be tested using state- or dyadic-level data that lacks information about domestic politics, they nevertheless have domestic political implications. Some domestic political actors must recognize and respond to the events and conditions that ultimately lead to the end of the rivalry. These external pressures rarely affect all political actors in the state identically. Those most directly affected are most likely to respond. Finding that they do not would raise real questions about whether the proposed theoretical mechanism really functions as suggested.

Arguments about domestic divisions over the appropriate response to international pressure are especially well developed in the case of economic interdependence. The most common argument about the impact of economic interdependence on conflict is pitched at the dyadic level: economic interdependence should increase the opportunity cost of conflict and contribute to more friendly relations (e.g., Oneal and Russett 1997). Some historians of the Anglo-American rapprochement have identified the two countries' economic ties as one reason the rivalry did not escalate to war after 1815 and ultimately ended peacefully (e.g., Campbell 1974, 201-4), though others have questioned its importance (e.g., Allen 1955, 69). However, economic interdependence can only contribute to the end of a rivalry if the actors who have the greatest stake in maintaining the economic relationship respond to it in the way liberal theorists have suggested. For instance, in the case of the United States and China, there is evidence that

members of Congress whose constituents benefit from trade with China are less likely to treat the country as a security threat (Kleinberg and Fordham 2013). Similarly, those whose constituents confront Chinese competition are more likely to criticize Chinese human rights violations (Cutrone and Fordham 2010). Some historians have applied a similar line of argument to the Anglo-American rivalry, noting for example that Southerners often argued for better relations with Britain because of their enormous stake in the cotton trade (e.g., Dunning 1914, 203-4). Data on legislative speeches coupled with information about the constituency of the speakers offers an inviting way to test whether the domestic politics of trade played a role in the Anglo-American rapprochement.

Similar arguments could be developed about the transmission of other international shocks into the policymaking process through the agency of those most affected by them. For instance, those most concerned about the emergence of a new rival, as historians have argued Britain was about Germany, should be most likely to favor conciliation with the existing rival. Similarly, those whose constituents are most in need of assistance during an economic downturn should be most likely to favor conciliatory policies toward a rival that will free up resources for domestic programs. Data on legislative speeches are well suited for testing these and other hypotheses concerning the domestic politics of responding to international conditions.

How Data on Legislative Speech Can Help Us Forecast International Events

In addition to testing hypotheses on rivalry process across multiple scales, from the individual legislator level to party differences to directed dyadic conflict and cooperation in the Anglo-American case, legislative speech offers the broader promise of improving on forecasts of international political and social events. If there is information, for example on rivalry escalation and de-escalation, in legislative speech, then this should be helpful in tracking the underlying hostility in relations that currently is unobserved and largely unmeasured. Recently, several studies have pointed out poor out-of-sample performance of many annual models of political violence (Ward, Greenhill, Bakke 2010, Gleditsch and Ward 2011).

Our approach builds upon and complements the event data tradition of forecasting social and political events. Recent approaches using event data disaggregated into sub-annual observations have proved relatively effective at forecasting events (O'Brien 2010, Schrodt 2011). The workflow to create event data is summarized in Schrodt (2011) and has been used in track conflict dynamics in a number of rivalries, as well as forecast future hostility (see Goldstein and Freeman 1991, Schrodt and Gerner 2004, and Brandt, Colaresi and Freeman, 2008). Recent large-scale data collections at the global level include the classified W-ICEWS data (see O'Brien 2010 and Schrodt 2011) and the publicly available SPEED project (Nardulli and Slana 2013).

Modern event data models leverage the structure of information found in news sources. Since event data coding is the closest analogue to our proposed workflow for legislative speech data, it will be useful to detail event data procedures, although Schrodt (2011) includes a more complete description than offered here. For event data coding, a news story is acquired either through webscraping or arrangement from a data service such as Lexis-Nexis. The story is then split into sentences or distilled to its lead sentence, which are filtered to try and screen out irrelevant information, such as sports scores and retrospectives as well as duplicates.

The sentences that survive the filter are parsed by a program, such as the open-source TABARI (Schrodt), which include sets of rules and elaborate dictionaries that identify parts of speech and named entities (such as countries or leaders), to code the source of an action, the target of that action, and the categorize the action/event. The inputs to these programs are non-trivial as the actors and actions need to be delimited. One useful categorization scheme for events is the Conflict and Mediation Event Observations (CAMEO) ontology, (Gerner, Schrodt and Yilmaz 2009), which includes 20 major categories subdivided in to many smaller more precise events. A recent actor dictionary for ICEWs was reported to have 20,000 entries (Schrodt 2011, 7). Event data coders than provide information on the source, event and target of the event, sometimes with post-processing that again attempts to reduce duplicate records (Schrodt 2011, 9). A model utilizing the counts of aggregated event types or scales, often augmented by additional information, can then be trained on a set of observations on a process to be forecasted. This model can then be used to forecast out-of-sample, or combined with other models to create an ensemble prediction (see Montgomery, Hollenback, and Ward 2012).

One of the strengths of this workflow is that journalists are trained to write lead sentences and structure their writing to identify the who, what, when and where of a story. Programs like TABARI can then leverage that structure. However, one significant downside of solely utilizing event data to forecast political and social events is that, as Schrodt (2011, 23) notes, "[n]ews reports are only a tiny, tiny fraction of all the events that occur daily, and are non-randomly selected by reporters and editors."

We suggest three characteristics of legislative speech corpora make them likely candidates to contain information that can supplement and extend forecasts of rivalry behavior and political and social events more generally. First, event data is limited to recent time periods due to the need for electronic news feeds. Almost all even data projects begin in the mid-twentieth century. While historical archives that are comparable to Lexis-Nexis sources could theoretically be made available, this has not yet happened. In addition, the structured journalistic standards and dictionaries that event data coding currently relies upon would need to be optimized for these changes. In contrast, legislative speech corpora can stretch back centuries. For example, in the United States we are currently working on processing files that begin in 1840, and the UK there are xml files for speeches in the House of Commons and the House of Lords beginning in 1809. This allows us to train models on an additional set of historical circumstances and still learn about out-of-sample forecasting performance. This enhanced time frame also has the potential to illuminate whether models that account for the evolution of rivalry and changes in domestic politics provide superior forecasting fits (Hensel 1999 and Colaresi 2006).

Second, where news reports include only newsworthy events, legislative speeches cover political and social topics that may not be newsworthy yet. For example, on May 25, 2011, and in other speeches from 2011 to 2012, Senators Ron Wyden and Mark Udall, both members of the Senate Select Intelligence Committee, discussed over-reach by the government in their use of Patriot Act provisions and surveillance. This was significantly before news broke after the Edward Snowden's leaks on several National Security Agency programs. In August and September, 2013, despite headlines such as "Senate committee approves resolution authorizing U.S. strike on Syria" (Washington Post, September 14, 2013), speeches in the Senate and the House reflected deep skepticism across about US strikes on Syria after the reported chemical weapons use in that

country's civil war. In the Senate this included skeptical reactions from Senators Tom Udall and Christopher Murphy, both from US President Obama's own party. Ultimately, the disagreement in the legislature presaged a backing down of US rhetoric and a negotiated comprise instead of air strikes. These anecdotes might suggest more broadly that signals in legislative speech might provide more lead time to warn of oncoming events, as compared to news stories.

Third, legislative speeches include a wider band of content on particular issues as compared to news stories about the event. There is a developed literature on how the media filters information (Tuchman 1975, Gamson and Modigliani 1987, McCombs, Shaw and Weaver 1997, and Smidt 2014), and in particular international information (Baum and Groeling 2012). However, legislative speeches are organized as debates, with different sides, competing frames and evidence and counter-evidence offered. In fact, the value placed on the heterogeneity of information provided in legislatures so great is that many, including the United States (Article 1, section 6 of the US Constitution) and the United Kingdom (Parliamentary Privilege), speech is protected even if it breaks slander laws or reveals national security secrets. The heterogeneity in legislative support for international threats has been theorized to be a useful signal of resolve (Schultz 2000). In addition, trends in the content of speeches can be informative. For example, Perkins (1968, 72) and Bourne (1970, 170-175) both suggest that increased positive sentiment both among elites and the public in the US helped to improve relations.

Therefore, we can also use the Anglo-American rivalry, and the availability of legislative speech corpora for over two centuries, as a test case for tracking and forecasting conflict and cooperation over time. Of course, where news stories collapse and summarize information, legislative debates unfold multiple frames for the same issue. The complexity of extracting signals from the content of speeches suggests will be a challenge. Despite the practical challenges, discussed below, the theoretical and practical benefits of utilizing this information to learn about the domestic politics of rivalry processes and explore the potential for utilizing legislative speech to forecast political events is potentially transformative.

Workflow and Research Design

In this section we outline our proposed workflow and research design. These are organized into data collection and storage, model formation and estimation, and critique and revision steps, as suggested by Blei (2014). In the critique and revisions section we outline some potential avenues to explore in the future as well as justify some of the flexibility that we have built into the workflow.

Data Sources

Since we are interested in tracking the changes in Anglo-American relations from rivalry to alliance, we would like to access complete corpora of legislative speech for the House of Commons and House of Lords in the UK and the Senate and House of Representatives in the US that begin as early in the 19th century as possible and continue until 1920 and the beginning of World War I. In the case of the UK, the Hansard is the publication of all legislative speech in Parliament. Libraries of the House of Commons and House of Lords in 2005 began digitizing and then made available the text of the Hansard beginning in 1803 and continuing through and past 1920. These texts are available in xml format, simplifying pre-processing. Uncompressed,

the speeches from 1803 to 1920 are approximately 5 terabytes and include over 1 million speeches by members in both houses on any topic.

Legislative speech corpora of the United States Senate and House of Representatives are less accessible. Legislative speech since 1873 to the present has been collected in the Congressional Record, and further back, from 1837-1873 a publication known as the Congressional Globe included records of debates in both houses. Until 1837, there was no contemporaneous systematic record kept of US debates, however information on these debates where collected later in the Debates and Proceedings in the Congress of the United States, also known as the Annals of Congress, which was published in two sections during the nineteenth century. However, only recent speeches in the late twentieth and early twenty-first centuries are made publicly available in machine-readable form (see Quinn, Monroe, Colaresi, Radev and Crespin 2010). There are several library holdings that have machine readable and searchable databases of both the Congressional Record and the Congressional Globe, including HeinOnline, but these database provides have not allowed academic access to their corpora and prevent scraping of the data.

Recently, the Google Books project and the HathiTrust as allowed us access to the corrected OCRed versions of the Congressional Globe and Congressional Record going back to 1837 and continuing until 1920. These records are simply flat text files, with very limited systematic formatting. However, these would be the most complete machine-readable public records of legislative speech available. We are currently working with the Digital services staff at the Michigan State University to procure the complete plain text corpora on our servers.

In addition to the Legislative speech data from each country, we are also interested in tracking the party and constituency of each member. For the UK, this information has been collected and organized into a relational database by Eggers and Spriling (Forthcoming). In the United States a publicly available database is available at, https://github.com/unitedstates/congress-legislators. In addition, this information is held in the Historical Atlas of Political Parties in the United Sates, 1789-1989 and are available through VoteView for each member.

Pre-processing

Extensive pre-processing is necessary to analyze the speech data. In addition, as we would like to make the data collection effort as useful as possible to researchers across different topics in the future, as well as anticipate future revisions we might want to make to our models. Therefore we follow several steps that keep as much information as possible at each step.

After the data is downloaded, we turn to parsing and identifying speeches that are relevant to Anglo-American relations. While it may seem obvious, the definition of a speech must be set out explicitly. In debates, particularly in the UK, there are often interruptions. Thus a speech could refer to contiguous words spoken by the same member without interruption, but that would split what we commonly think of as speeches into parts. In the algorithms we will be using, splitting speeches into smaller quantities will reduce the information we are able to extract. Therefore, we first define a speech as all ordered words spoken by the same person in a debate on a given day. Speeches can then be broken up into what we might call contiguous utterances in the future. For our purposes, we concatenate utterances by the same person in a day within a debate into a

speech, even if they were interrupted. For completeness, we define debate by the headings in the legislative speech records and the date provided in the source.

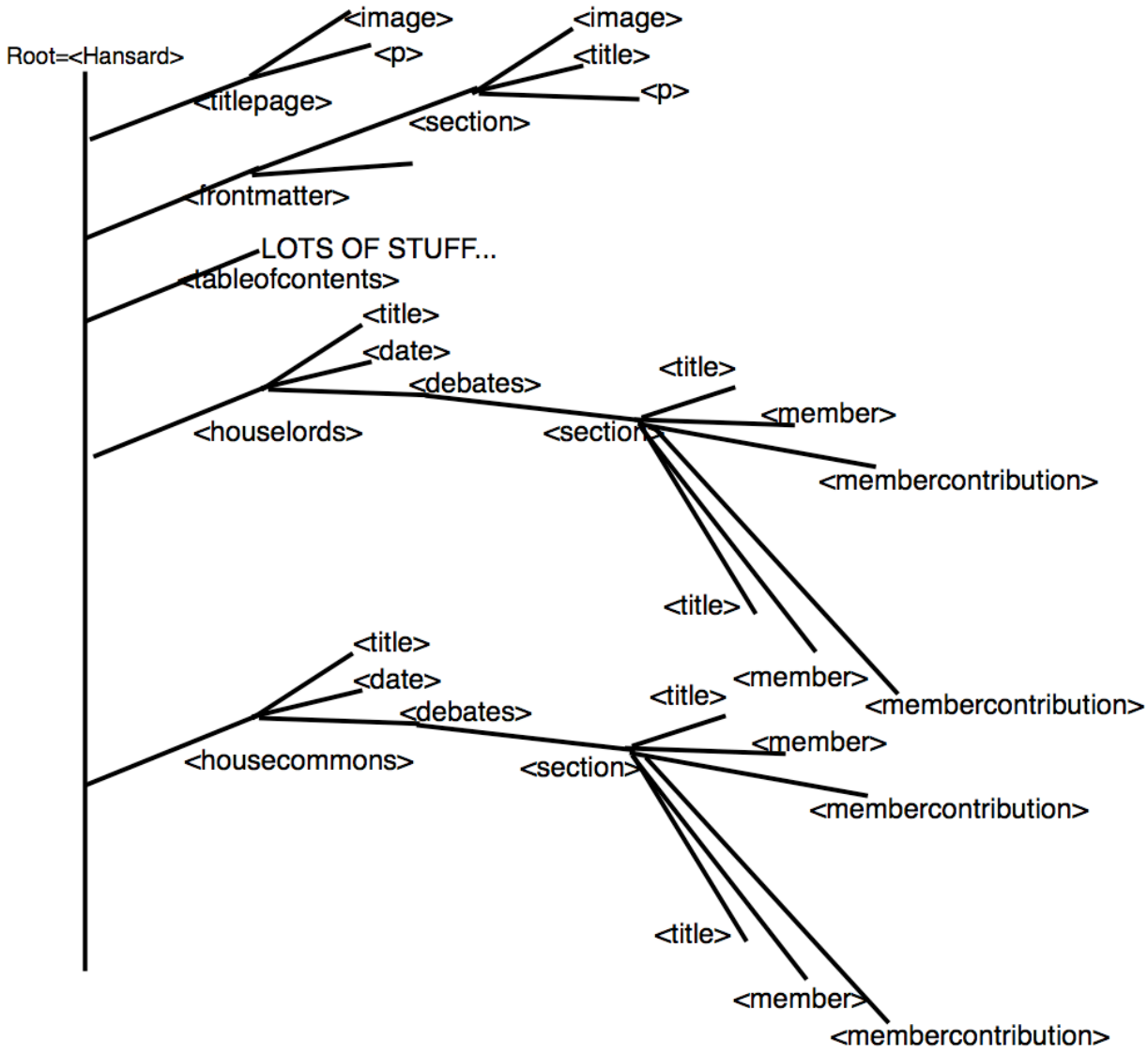

Figure 2: A partial representation of the xml tree for one day's Hansard file.

Using this definition of speech, for the UK source data, we first parse the xml tree with is available for each day of proceedings. The tags identify the date (<date>) and location (<houselords> or <housecommons>) , of the speeches, as well as the speaker (<member>), and then the content of the speeches (<membercontribution>). There are some small complications with parsing the tree as formatting tags are included in the speeches and the speeches are not nested within speakers. We handle these by stripping the specific formatting tags, such as (<bold>) before parsing the true using regular expression matching and after parsing the speeches checking that the <member> coding and <membercontribution> vectors have the same

length. We use Python with the packages lxml and elementtree to parse the xml files and the mechanize package for web scraping.

To identify relevant content we use a dictionary approach. We created a dictionary that include 150 key phrases and terms including different names for the United States of America (e.g., "America"), names of US Presidents, US Secretaries of State, and other keywords such as forts that were under contention (e.g., "Fort Bowyer") and trade issues (e.g., "cotton") that were commonly discussed. We also built in a filter for phrases that included single words that appeared to be refer to the US but do not (e.g., "Washington Square").

Once a speech is identified as relevant, we will have a simple python program format the content of that speech into an analyzable matrix. First, we filter out a list of stop words and punctuation. Stop words are common tokens with little or no content (such as can). We use the list of 127 stop words as well as the punctuation symbols defined in the NLTK package, with one change. We keep "not" in the texts, as negation can significantly change the meaning of the next word. This information is then used below.

After the stop words have been removed, we stem the remaining tokens. Stemming is the process of reducing words to their common root, so "attack" and "attacking" are both reduced to "attack". We utilize the Porter stemmer, again, as implemented in the NLTK package. Since we have not thus far altered the order of the words, we then concatenate "not", itself a stem, to the next adjacent stem. Thus the stemmed three word ordered pair "not a threat" would be stemmed and then concatenated to be "not_threat". We then count the number of stems in each speech. In the future it is very easy to aggregate negated and non-negated stems, however if we did not separate this information now, it would be very time-consuming to disaggregate the count.

Parsing in the US case is much more difficult as there are not tags for relevant information. We are currently exploring different pattern matching approaches. In addition, we are developing a similar dictionary for the US that will be ready when the US Legislative Speech data become available.

We T separate Nx(W+3) matrices which holds the count of each of the W stems, for each of the N speeches, in each of the T days when the legislature is in session. The matrices are stored in sparse format. The last three columns in the stem count dimension include the content of the <member> and <debate> tag from the original xml for future use and a document ID that uniquely identifies the record. We build the dictionary of all words spoken sequentially in time, and then fill-in zeros for words in matrices in that had not been spoken yet. The storage of the original speech text and the stem counts is nontrivial as is moving the data between external drives and workstations and local machines and the MSU computing cluster.

Models and Estimation

We are ultimately interested in a hierarchical model of legislatives speech that has three levels and can be useful for forecasting international events. The levels involve, first, a set of K themes that may change over time but that divide the issue space discussed in the legislative chambers, where each of the N speeches is potentially on a combination of these themes. Within these themes we are interested in a second level, the potentially time-varying systematic partisan

differences in word choice. We can use indicator g to delineate membership identity in a set of length G potential groups which we are interested in comparing across. At the third level, we want to identify the changes in language for individuals i who are nested within a group g across the K topics.

One relatively simply framework that would allow us to model legislative speech into these levels is Latent Dirichlet Allocation (LDA) (Blei, Ng, Jordan 2003, and Blei 2012). LDA has been used across a multitude of corpora (see Blei 2012) and extended to take into account corpora that arrive sequentially, allow the content of topics to evolve over time, and incorporate covariates. Recently Bagozzi and Schrodt (2012) have suggested that LDA provides a useful topical summary of newswire sentences on international events and can uncover patterns that map to underlying conflict-cooperation spectrums. These models allow us to extract the themes from the text.

For our purposes, a useful extension of LDA has recently been suggested by Roberts, Stewart, Tingley and Airoldi (2014). This is a Structured Topic Model (STM) that allows covariates to predict variations in both the prevalence of the topic as well as the content of the topic. For our partisan group comparisons, we are particularly interested in how parties approach a theme by using different terms or frames. The methods of Monroe, Colareis and Quinn (2009) will also be useful in tracking these partisan frames. We also plan to develop a Dynamic Structured Topic Model that merges the insights from the Dynamic LDA model with the covariate structure in the STM.

Critique and Revision

We can validate and critique our model using the historical record, expert accounts, and posterior predictive checks, in-sample. This will involve analyzing whether the LDA models uncover useful structure, whether the partisan divides identified are plausible and if we are able to locate the entrepreneurs that led changes in themes across the rivalry. We plan to utilize cross-validation strategies to compare models and in particular to validate more complicated modeling strategies that involve covariates and dynamics.

Most importantly, we are also interested in exploring the usefulness of these models with respect to forecasting international events. New approaches to forecasting events with big data and model specification and uncertainty have been proposed. These On-Line Bayesian Dynamic Logit Models and Ensemble Bayesian Model Averaging approaches can help us narrow down the vast model space of which words, themes and trends help to predict, for example, militarized disputes in the Anlgo-American Rivalry and other events of interest. We will investigate revising the model to avoid over-fitting of Sample data and explore different numbers of underlying themes. We can also explore whether parts of speech tagging and alternative tokens, such as n-grams, or further negation use will be helpful.

Plans for the Future: Beyond the Atlantic

The workflow, tools and evidence we will acquire in this project will have direct application to the study of streams of international relationship outside of Anglo-American relations. First, the same data we are collecting for these relationships can be used to analyze themes in US relations with other countries including the Soviet Union. Similarly, UK relations with France, Germany

and the Russia/USSR can be analyzed. Israel keeps a long-running corpus of legislative speech that may be of interest to many. Democratic rivalry relations, for example between Turkey and Greece, may be particularly rich locations, where mining legislative speech might pay particularly inferential dividends. Australia, Canada and New Zealand also maintain similar information that could be used to track both the breakdown of the British Empire, and along with other Scandinavian and European countries can illuminate the building, maintenance and evolution of NATO. Newer legislatures in the Middle East/North Africa region, as well as legislatures in hybrid democratic systems, such as Russia, would also be interesting contexts for the application of our workflow and tools. In the best case, we imagine being able to track legislative information around the world to help us forecast international events in near real-time.

Conclusion

Legislative speech data have much to tell us about rivalry and many other topics. This source of data covers not only newsworthy events, but also a wide range of additional topics that sophisticated political actors considered worthy of attention even though they were not in the news. These speeches can thus illuminate the state of relations even during periods when few of the events included in most datasets are taking place. Moreover, the identity of the speaker provides a way to link the policy perspectives in each speech to specific political entrepreneurs and their constituents, allowing us a test a range of hypotheses not possible with existing data.

Once it is completed, we hope that this project will advance our knowledge of interstate rivalry in at least four ways:

- It will allow us to test existing arguments from the political science literature about why rivalries end in fine-grained ways that bring us closer to the underlying causal process than has been possible with existing data.

- It will allow us to test the generality of patterns historians have suggested about the end of the Anglo-American rivalry and the emergence of the "special relationship."

- It will allow us to develop and test forecasting models using legislative speech data to predict conflictual and cooperative events over a long span of time.

- It will develop a research design that can be applied to other rivalries for which there are similar data. These comparable data include speeches in other legislatures, but might also other types of commentary by major political actors such as official statements by foreign ministries, executives, and other government officials, editorial commentary in major media outlets, and the like.

References

Akcinaroglu, Seden, Jonathan M. DiCicco, and Elizabeth Radziszewski. 2011. "Avalanches and Olive Branches: A Multimethod Analysis of Disasters and Peacemaking in Interstate Rivalries." *Political Research Quarterly* 64(2): 260-75.

Akcinaroglu, Seden, Elizabeth Radziszewski, and Paul Diehl. 2013. "The Effect of Rivarly on Rivalry: Accommodation and the Management of Threats." *Foreign Policy Analysis* 10(1): 81-100.

Allen, H.C. 1955. *Great Britain and the United States*. New York: St. Martin's Press.

Bennett, D. Scott. 1996. "Security, Bargaining, and the End of Interstate Rivalry." *International Studies Quarterly* 40(2): 157-83.

Bennett, D. Scott. 1997. "Measuring Rivalry Termination, 1816-1992." *Journal of Conflict Resolution* 41(2): 227-54.

Bennett, D. Scott, and Timothy Nordstrom. 2000. "Foreign Policy Substitutability and Internal Economic Problems in Enduring Rivalries." *Journal of Conflict Resolution* 44(1): 33-61.

Blei., D.2014. Build, compute, critique, repeat: Data analysis with latent variable models. *Annual Review of Statistics and Its Application* 1:203-232.

Blei, D., Ng, A, and Jordan, M. 2003. Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, January 2003.

Box, G.E.P. 1976. Science and Statistics. *Journal of the American Statistical Association*, Vol. 71, No. 356. pp. 791-799

Box, G.E.P. 1980. Sampling and Bayes' Inference in Scientific Modeling and Robustness. *Journal of the Royal Statistical Society*, Series A, 143(Pt. 4):383-430.

Brinton, Crane. 1948. *The United States and Great Britain*. Norwood, MA: Norwood Press.

Campbell, Charles S. 1974. *From Revolution to Rapprochement*. New York: John Wiley & Sons.

Colaresi, Michael P. 2001. "Shocks to the System: Great Power Rivalry and the Leadership Long Cycle." *Journal of Conflict Resolution* 45(5): 569-93.

Colaresi, Michael P. 2004. "When Doves Cry: International Rivalry, Unreciprocated Cooperation, and Leadership Turnover." *American Journal of Political Science* 48(3): 555-70.

Cutrone, Ellen A., and Benjamin O. Fordham. 2010. "Commerce and Imagination: The Sources of Concern about International Human Rights in the United States Congress." *International Studies Quarterly* 54(3): 633-55.

Dreyer, David R. 2010. "Issue Conflict Accumulation and the Dynamics of Strategic Rivalry." *International Studies Quarterly* 54(3): 779-95.

Dunning, William Archibald. 1914. *The British Empire and the United States.* New York: Charles Scribner's Sons.

Fader, Anthony; Dragomir R. Radev; Michael H. Crespin; Burt L. Monroe; Kevin M. Quinn; and Michael Colaresi. 2007. "MavenRank: Identifying Influential Members of the US Senate

Using Lexical Cen- trality." *Proceedings of the Conference of Empirical Methods in Natural Language Processing (EMNLP '07)*. Prague, Czech Republic.

Gerrish, S and D. Blei. 2012. How they vote: Issue-adjusted models of legislative behavior. *Neural Information Processing Systems*.

Gerrish, S. and D. Blei. 2011. Predicting legislative roll calls from text. *International Conference on Machine Learning*.

Goertz, Gary, and Paul F. Diehl. 1995. "The Initiation and Termination of Enduring Rivalries: The Impact of Political Shocks." *American Journal of Political Science* 39(1): 30-52.

Goertz, Gary, Bradford Jones, and Paul F. Diehl. 2005. "Maintenance Processes in International Rivalries." *Journal of Conflict Resolution* 49(5): 742-69.

Hassan, Ahmed; Anthony Fader, Michael H. Crespin, Kevin M. Quinn, Burt L. Monroe, Michael Colaresi, and Dragomir R. Radev. 2008. "Tracking the Dynamic Evolution of Participant Salience in a Discus- sion." *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)*. 313-320.

Hensel, Paul R., Gary Goertz, and Paul F. Diehl. 2000. "The Democratic Peace and Rivalries." *Journal of Politics* 624): 1173-88.

Hopkins, Daniel J., and Gary King. 2010. "A Method of Automated Nonparametric Content Analysis for Social Science." *American Journal of Political Science* 54 (1): 229–47.

Grimmer, Justin. 2010. "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases." *Political Analysis* 18 (1): 1–35.

Grimmer, Justin, Messing, Solomon, and West, Sean. 2012. "How Words and Money Cultivate a Personal Vote: The Effect of Legislator Credit Claiming on Constituent Credit Allocation" *American Political Science Review* Vol. 106. No. 4.

Kleinberg, Katja B., and Benjamin O. Fordham. 2013. "The Domestic Politics of Trade and Conflict." *International Studies Quarterly* 57(3): 605-19.

Kuenne, Robert E. 1989. "Conflict Management in Mature Rivalry." *Journal of Conflict Resolution* 33(3): 554-66.

Monroe, Burt; Michael Colaresi; and Kevin M. Quinn. 2009. "Fightin' Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict." *Political Analysis*. 16: 372-403.

Morey, Daniel S. 2011. "When War Brings Peace: A Dynamic Model of the Rivalry Process." *American Journal of Political Science* 55(2): 263-75.

Oneal, John R., and Bruce M. Russett. 1997. "The Classical Liberals Were Right: Democracy, Interdependence, and Conflict, 1950-1985." *International Studies Quarterly* 41(2): 267-93.

Perkins, Bradford. 1968. *The Great Rapprochement*. New York: Atheneum.

Prins, Brandon C., and Ursula E. Daxecker. 2010. "Committed to Peace: Liberal Institutions and the Termination of Rivalry." *British Journal of Political Science* 38(1): 17-43.

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

Schrodt, Phil. 2011. "Automated Production of High-Volume, Near-Real-Time Political Event Data". Working Paper.

Schrodt, Phil. 2012. "Precedents, Progress and Prospects in Political Event Data." *International Interactions* 38(4).

Schrodt, Phil. 2014"Seven Deadly Sins of Contemporary Quantitative Analysis." *Journal of Peace Research*.

Schrodt, Brandt, and Freeman. 2011. "Real time, time series forecasting of inter- and intra-state political conflict." *Conflict Management and Peace Science* 28: 41-64.

Schultz, Kenneth A. 2005. "The Politics of Risking Peace: Do Hawks or Doves Deliver the Olive Branch?" *International Organization* 59(1): 1-38.

Eggers, A, and Spirling, A. Forthcoming. "Ministerial Responsiveness in Westminster Systems: Institutional Choices and House of Commons Debate, 1832-1915", *American Journal of Political Science*.

Temperley, Howard. 2002. *Britain and American Since Independence*. New York: Palgrave.

Thies, Cameron G. 2001. "A Social Psychological Approach to Enduring Rivalries." *Political Psychology* 22(4): 693-725.