Dimensions of Diplomacy

What the Wikileaks Cables can tell us about Information and Privacy in International Relations*

Michael Gill† Arthur Spirling‡

Abstract

We consider the role of information in international relations—a key concept in 'rationalist' models of interstate bargaining—from a unique empirical perspective. Noting that little systematic observational data exists regarding the contemporary private beliefs and private actions of state actors, we analyze one hundred and sixty thousand United States diplomatic cables from the Wikileaks organization for the period 2005 to 2010 to speak to several aspects of recent theoretical work in the area. In this preliminary analysis, we show that diplomatic secrecy consists of two distinct 'dimensions': substantive and procedural. The former deals with secrets *per se*, the publication of which would actively damage US interests, especially in terms of revealing the resolve or capabilities of the state. Procedural secrecy, meanwhile, deals with the diplomatic norm of confidentiality in meetings—regardless of their substantive content. We relate these two dimensions of diplomacy to different concepts of secrecy in the theoretical IR literature, and demonstrate that both play an important role in dictating the classification decisions of the US State department. In uncovering these substantive points, our paper presents new methodological tools of general interest to scholars in the field.

*First version: February 2014. This version: September 26, 2014.

†Department of Government, Harvard University. mzgill@fas.harvard.edu

‡Department of Government, Harvard University. aspirling@gov.harvard.edu

1 Introduction

Starting at least with the efforts of Fearon (1995) (see also Powell, 1999), the 'rationalist' approach to war has become the "dominant" framework for studying conflicts (Lake, 2010), inspiring a large theoretical literature that both extends the original model, and explores the precise nature of the mechanisms that underlie it (see, e.g., Reiter, 2003, for an overview). At the core of the approach, and thus at the core of subsequent efforts, are the assumptions of asymmetric information and commitment credibility (Powell, 2002; Freiden and Lake, 2005). Put simply, leaders have incentives to misrepresent their information (especially their resolve or capabilities), and they have problems committing to not fighting given specific strategic situations they may face. In light of the first of these issues, a body of related work has grown up around the "hand tying" (Schelling, 1966) effect of "audience costs" (Fearon, 1994*a*) which purportedly provide leaders with a way to effectively signal their resolve (see e.g. Smith, 1998; Schultz, 1999; Slantchev, 2006).

Precisely because actors within these approaches cannot send credible *private* signals, there is little role for 'diplomacy' as that term is typically understood: put crudely, any information that a leader—or that a diplomat on behalf of a leader—tries to communicate about state preferences or abilities to a foreign agent is deemed 'cheap talk' since there are no consequences from not revealing the truth in such discussions (Fearon, 1994*a*; Ramsay, 2004). Given the long history of diplomacy (see Black, 2010, for an overview), and the fact that leaders continue to regard it as a vital part of international relations today (e.g. Kissinger, 1994), it is perhaps unsurprising that scholars have found this feature of the rationalist framework unsatisfying. Thus in recent times, researchers have turned their attention to modeling diplomacy explicitly, with contributions that allow for information transmission based on secret negotiations as an alternative to public declarations (Kurizaki, 2007), as a

precursor to bargaining (Ramsay, 2011), as a precursor to a breakdown in other aspects of an international relationship (Trager, 2010), or as a mechanism by which a reputation for 'honesty' may be established and maintained (Sartori, 2002; Guisinger and Smith, 2002).

Despite the theoretical advances in understanding diplomacy, empirical evidence of the particular mechanisms discussed—or indeed of diplomacy in general—remains scant (although see Sartori, 2005). This is in stark contrast to the relative wealth of data-driven scholarship on the canonical bargaining model in international relations (Fearon, 1994*b*; Werner, 1999; Goemens, 2000; Reed, 2003; Smith and Stam, 2004; Ramsay, 2008; Reiter, 2009) and on audience costs in particular (see, e.g., Schultz, 2001; Tomz, 2007; Downes and Sechser, 2012). While such work has not meet with universal approval (e.g. Baum, 2004; Snyder and Borghard, 2011; Trachtenberg, 2012), the imbalance between theory and empirics in the study of diplomacy is nonetheless a concern. For example, we know little about the relationship between subject matter and secrecy (literally we cannot answer the question "which topics are discussed privately, and which are public?"), nor do we know how information is gathered, aggregated and communicated to domestic and foreign leaders. Of course, at a practical level, it is not hard to see why theories involving secret communication are harder to test than they are to propose. By definition, the private information of leaders, obtained or transmitted by their diplomats, is not typically available to observers. This is *a fortiori* the case for information that the state regards as sensitive: it may never be released on the grounds that to do so would damage national interests, endanger its citizens (though see Shapiro and Siegel, 2010, for discussion) or more cynically, because it might allow popular challenge to elite control of foreign policy (Gibbs, 1995). Thus researchers must either work with a few (perhaps non-representative) sampled cases or with systematically censored data; both have baleful consequences for any subsequent process of statistical inference.

Given the centrality of the concept to IR, and given the obvious data problem noted above, there is clear value in characterizing exactly how diplomacy 'works' in an empirical sense. This paper seeks to provide a first step in this process, by utilizing a new data set yet to be exploited by political scientists: the release of over a quarter of a million U.S. diplomatic cables that covering the period 1966 through to 2010 by the Wikileaks foundation. Precisely because the leak was unauthorized, the data is uncensored: the information flows between the United States Department of State and U.S. embassies include documents meant only for officials with high level security clearance. While the cables do not necessarily deal with historical examples of 'crisis bargaining'—the main scenario to which studies about private information speak—they do pertain to serious matters of state security and foreign policy within the modern era and are extremely useful in that regard.

As we demonstrate below from examining this data, private diplomatic communication may be characterized as having two dimensions: *substantive* secrecy and *procedural* secrecy. The first of these dimensions captures the idea that different audiences are allowed access to information about different subjects. For example, diplomatic communication pertaining to "Internal Government Affairs" or "Arms Controls" is much more likely to be kept from the public than cables dealing with "Trade Expansion" or "Travel". That such variation exists is not *per se* surprising, but our study allows researchers a properly nuanced understanding of the structure of diplomacy, beyond the broad abstract brush strokes that IR theories must necessarily use in this area. In so doing, we learn about leader preferences: what they wish citizens to know and not to know about the information they have and the actions they take. Furthermore, we show that the particular censorship decisions made are compatible with 'mainstream' rationalist approaches to international relations. The second dimension of diplomacy, that of *procedural* secrecy, concerns the way in which diplomats go about obtaining and disseminating information, regardless of subject area. Thus, procedural secrecy

describes how confidential and public documents differ *within* a given topic. We show that the primary component of procedural secrecy is information pertaining to the meetings that diplomats (specifically ambassadors and other agency staff) have with foreign officials, and the data that is obtained from said actors and/or passed to them in private. Put crudely, censored cables are not about "proper nouns" like particular leaders, countries, weapons or threats; what matters is that discretion is assured for all actors. This finding provides evidence consistent with the recent theoretical work on diplomacy noted above, while adding depth to those accounts: i.e., for the first time on a large scale and with recent uncensored data, we show that private signals are indeed a crucial and seemingly effective part of international relations, and that it is precisely the operating principle of confidentiality without public confirmation that facilitates the day-to-day business of statecraft. Taken broadly our results suggest, in line with recent literature critical of the 'audience cost' paradigm (Snyder and Borghard, 2011), that while public disclosure is actively pursued in some areas, private negotiation plays an important—and possibly equal—role in diplomatic interactions.

In undertaking the study, our paper innovates methodologically and suggests new ways of working with 'texts-as-data' (Grimmer and Stewart, 2013). In particular, we use machine learning techniques—such as random forests (Breiman, 2001) and the 'lasso' (Tibshirani, 1994)—to identify how 'important' terms discriminate between restricted and unrestricted documents. We provide novel ways of comparing texts, based on matching on the tokens within them, such that political scientists can think sensibly and systematically about the (marginal) 'effect' of secrecy on a document.

Our paper proceeds as follows: in Section 2 we briefly review relevant literature, and describe what it might predict for the research questions at hand. Section 3 describes our data, while Section 4 sets out our methodological approach. Results are found in Section 5, while

Section 6 concludes.

2 Literature and Orientation

In this section, we derive predictions from the literature on private information in International Relations that we might 'test' on our data set. Before starting that process, we underline the obvious caveat that the cables are not part of any single 'crisis bargaining' scenario. This is a problem insofar as most theories and studies we draw from are designed to tackle precisely those types of situations (in line with the original contribution of Fearon, 1994*a*). Nonetheless, we would claim that analysis of the documents can tell us about the continuous process of information generation and exchange that forms a backdrop for the more discrete episodes of escalation or diplomacy around (potential) confrontations.

2.1 Audience Costs and their Skeptics

The essence of an 'audience cost' is the price a leader pays when he makes a threat during a crises, and then does not follow it through (Fearon, 1994*a*; Schultz, 2001; Tomz, 2007). Because leaders will suffer if they do not act in the ways they threatened to, issuing public statements becomes a way to "lock in" and credibly commit to courses of action, even if to do so would not necessarily be in the best interests of their state or its citizens (see Slantchev, 2006, for discussion of this point). Skeptics of audience cost theory (e.g. Snyder and Borghard, 2011; Trachtenberg, 2012) emphasize that, when dealing with issues of foreign policy, politicians would generally prefer to preserve 'room for manoeuver' if possible. To be clear, this prediction may also arise from positions more sympathetic to traditional audience cost theory: 'rationalist' scholars have certainly presented models for cases in which public threats can cause problems for those issuing them (Kurizaki, 2007; Slantchev, 2010). Nonetheless the difference in prediction from the 'main line' literatures is clear: audience

cost theorists generally predict that international states will devote more time and effort to active public signalling of positions than the skeptics. Of course, neither set of theories—pro- or anti-audience cost—is precise in terms of empirical prediction: put crudely, they do not tell us when or where, in terms of timing or subject matter, governments will release public statements, as opposed to retaining information for internal consumption only. To obtain some empirical traction, we assume that audience costs theorists would generally predict non-uniform levels of secrecy across topics of diplomatic discussion: that is, if leaders are able (in the sense of Slantchev, 2012) to keep certain matters secret they will do, but they will gladly put others in the public domain in order to generate 'lock in' from their publics (and the publics of other nations). By contrast, taken at its most literal, the anti-audience cost position is presumably that governments ought to keep everything secret if they are able: all the better to allow for alternative courses of action later on.

Notice that our maintained assumption here is that diplomatic communication, either implicitly or explicitly, connotes notions of resolve, or 'preferences' more broadly; that is, we assume that the very act of sending a cable about a given subject allows external actors who observe it (or who could have observed it, were it public) to update their beliefs about the strategic and international concerns and priorities of the sending nation.

2.2 Theories of Diplomacy

In contrast to rationalist models which have little room for diplomatic relations, in recent times some scholars of formal theory have turned their attention to the notion that private meetings may indeed allow for the communication of information. Sartori (2002) constructs a model in which leaders, or their diplomats, have a long-term, iterated relationship. In that world, building a reputation for 'honesty' matters, and thus actors will reveal the truth about their position and resolve. In related work, Guisinger and Smith (2002) give a mechanism

by which dishonest leaders might be removed by their citizens. Ramsay (2011) considers a model that is similar to that of Sartori (2002), insofar as it relies on the threat of a reversion to a 'bad' equilibrium if negotiation breaks down, though the overall mechanism is different and does not rely on long term interactions. In particular, Ramsay (2011) proposes a 'simple diplomacy' model in which, prior to a potential conflict, a state can indicate that it is open to negotiation—even if that negotiation may result in war anyway. Being willing to discuss the issue at hand signals to the opposing party that war may be avoidable, and also allows coordination prior to fighting. Trager (2010) considers a model in which private communication over some dispute affects not simply the issue at hand, but also the states' perception of each other. In contrast to reputation-based account of Sartori (2002), the issue is not that leaders are concerned with being caught 'lying'. Rather it is that are aware that if their threat is taken seriously, it may result in the threatened state taking actions (such as forming military alliances) that are costly to the state sending the private signal. Separate to the logic of private meetings as allowing for information exchange (concerning reputation or preferences or something else), Kurizaki (2007) presents a model in which secrecy is 'efficient' because it insulates leaders from their publics. Thus, collectively, they may be able to access a much wider set of peaceful outcomes for their negotiations than would be available if they were compelled to present their plans to citizens for approval.

These theories, and the mechanisms they proffer, are not easy to 'test'. Nonetheless, they can be broadly split into two camps of empirical predictions. On the one hand, the models of Trager (2010) and Ramsay (2011) suggest that diplomatic exchanges are about information exchange albeit in different ways and for different reasons. That is, information may be transferred from one party to another *despite* the fact that the meetings take place in private. Taken to data on diplomacy, one implication of such work is that private diplomatic cables should pertain to the passing of information, and that this should presumably be

domain specific. Depending on the specification of the model one believes, this might be more or less information than in the public cables. Either way, one should expect to see substantive terms and 'proper nouns' dominating the set of words that discriminate between restricted and unrestricted cables. On the contrary, the Sartori (2002) and Kurizaki (2007) models imply that what matters is not substance, but procedure and reputation: i.e., that private cables are 'different' primarily because they deal with secret protocol of meetings *per se*. Furthermore, this should be relatively robust to the specific subject area of relations being discussed: diplomats will expect insulation against public buffering whatever the issue at stake.

2.3 Dimensions of Diplomacy

Our investigation below attempts to uncover to what extent these theories of interstate bargaining find support in the data. Perhaps unsurprisingly, our contention is that neither set of accounts has the monopoly on the truth. In allowing a place for both, we argue that diplomatic secrecy comes in at least two varieties, or 'dimensions', which are compatible with each other. The first pertains to the type of 'substantive' secrecy that bargaining theorists are familiar with: material from which state capabilities may be learned, and thus should not released for fear of weakening an international actor's position. In some cases, in line with an audience cost argument, the U.S. may make such positions public. Below we explore in more detail when precisely they do so, using tools from regression analysis. The other dimension of diplomacy is 'procedural', and refers to the notion that actors seek to discuss matters confidentially, regardless of area. We suspect that this practice is connected to the logic outlined by Kurizaki (2007), insofar as it allows actors to speak more freely, and agree to more 'damaging' positions, than they might otherwise be able to 'in public'. In stark contrast to the substantive dimension of secrecy, we will be required to show that, separate to any topical differences, what distinguishes confidential from public texts are the

norms of diplomatic meetings. Thus, these two dimensions of diplomacy are distinct yet related: procedural secrecy exists whatever the substantive matter at hand, and whatever its underlying importance in terms of international events.

3 Data

The core of our data are the WikiLeaks cables: 251,237 diplomatic cables sent by the U.S. State Department to U.S. embassies and missions around the world. The date range for the original data is from 1966 to 2010, and in Figure 1 we plot the total number of cables per month from that time period. In our work here, we focus on all cables released on or after January 1, 2005. We do this for two reasons: first, because coverage prior to the year 2000 is fairly sparse, and second, because we wished to guard against any change in protocols—concerning the content, security or nature of the cables—that the terrorist attacks of September 11, 2001 may have ushered in. All told, we are left with around 163,000 documents from which to draw inferences.

Technically speaking, cables may be classified into one of three categories, depending on the degree of damage to national security that "the unauthorized disclosure of which reasonably could be expected to cause."[1] Furthermore, any classified document must pertain to at least one of a series of topics which *inter alia* include military plans, intelligence, foreign relations of the United States, nuclear programs, weapons of mass destruction and vulnerabilities in national security. In descending order of the purported balefulness of unauthorized release, these categories are 'Top Secret', 'Secret' and 'Confidential'. If a cable does not meet the criteria for such restricted access, it is deemed 'Unclassified'. In our particular data, we have the following distribution: no Top Secret, 10,195 Secret cables, 87,270 Confidential cables,

[1] As described in Executive Order 13526, 2009.

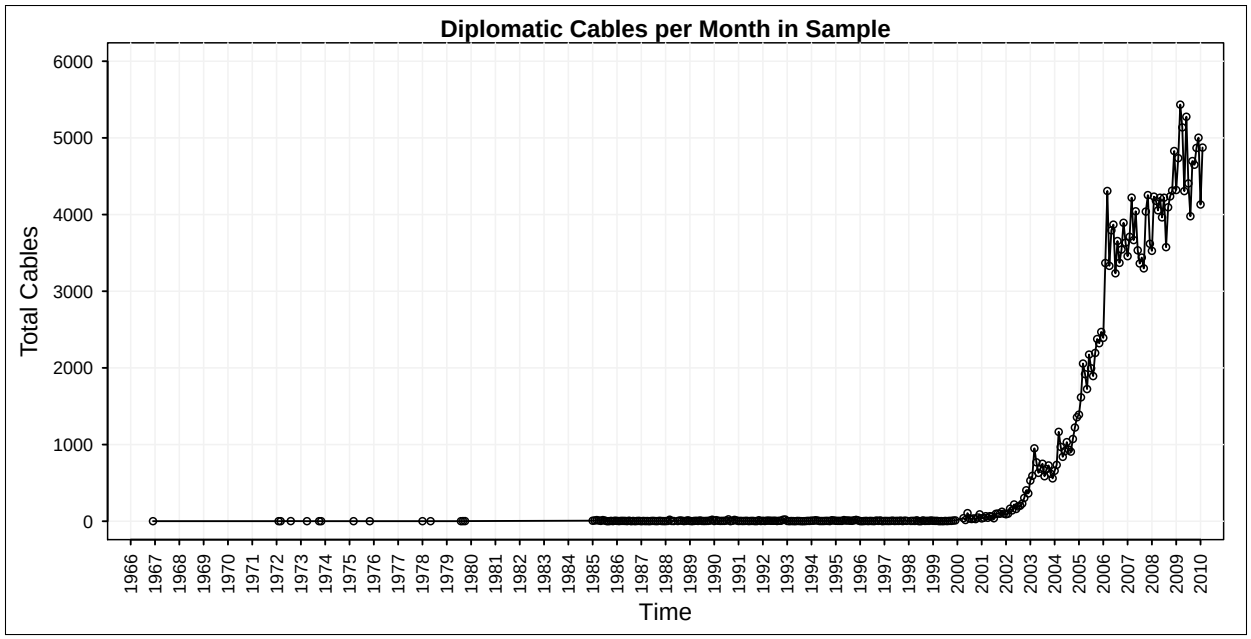

Figure 1: Number of cables per month, 1966–2010. Note that post-2001 period has much more dense coverage.

and 66,493 Unclassified. There are, in addition, some extra classifications that appear less frequently in the data, such as 'Confidential and Not For Foreign Distribution' , 'Unclassified for official use only' , and 'Secret and Not for Foreign Distribution'; we ignore these categories for our current efforts.

For our purposes below, we divide the categories into 'restricted' (R), which includes Secret and Confidential communications, and 'unrestricted' (U), which includes the unclassified documents only. The central idea here is to code documents into 'private' and 'public' information, respectively. This measure is somewhat crude, but given that theories in International Relations use similarly binary demarcations we think this is reasonable. To be clear, the fact that a cable is unrestricted does not mean that it is automatically made public: it is still a government document rather than a press release. But unclassified documents—so long as they are not 'For Official Use Only'—do make their way into the public domain, and are eligible for release under Freedom of Information requests. Put otherwise, our unre-

stricted case covers documents that the public (anyone without specific security clearances) could access; our restricted cables are those that are not released or releasable to the public.

Any given document has a series of subject matter 'tags' assigned to it by its authors, with guidelines for this process provided by the State Department.[2] From our perspective, these tags contribute meta-data that communicates the topic of the content therein. Examples include '`ADCO`' which refers to 'Diplomatic Courier Operations', '`PTER`' which refers to 'Terrorists and Terrorism', '`SMIG`' which pertains to 'Migration' and so on. There a total of 97 tags in our data, though their use varies widely in relative frequency terms (as we report in more detail below). The full list can be seen in Appendix A. The variety in tag number per document can be seen in Figure 2; inspection suggests that the modal number of tags is two or three. In Figure 3 we report the structure of the data in terms of the way that tags co-occur across cables. Areas of darkness in that plot are places where tags coincide. Our main observation is that tags in section 'P' (which denote 'Political' issues) and, to a lesser extent, tags in section 'E' (denoting 'Economic' matters) tend to coexist heavily with other subject indicators, suggesting that these issues play an important organizing role in the U.S. diplomatic service. Machine readable versions of the documents themselves are available at various websites for download, though some pre-processing is then required prior to any analysis. In particular, the tag information must be captured and removed, and some other cleaning performed. Much of what follows involves operations on the 'document-term matrix' (DTM) of the texts, which was 'stemmed' (meaning that words were pruned back to their 'roots' where possible, using the Porter stemmer), 'stopped' (meaning that function words

[2]These are literally 'TAGS', an acronym for 'Traffic Analysis by Geography and Subject', implemented for diplomatic communication in its modern form by an executive order (number 11652) in June 1974. Their justification was to "[p]ermit more rapid and discriminating distribution of messages", and to "[p]rovide statistics to both offices and posts on what is being communicated in the Department-field system"; they were to "[s]erve as headings for clustering the terms used by professional indexers to identify the content of substantive messages."

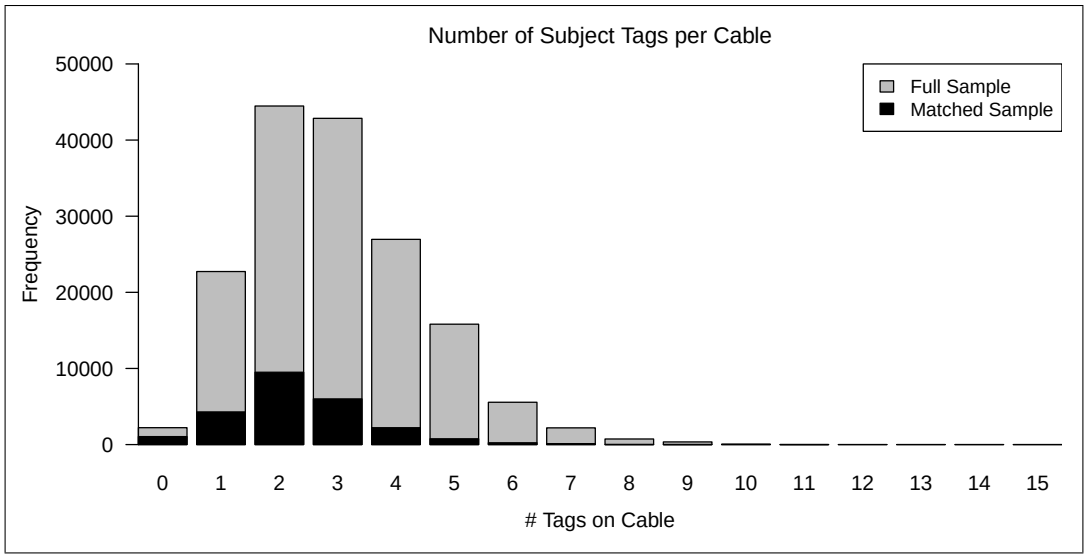

Figure 2: Distribution of tag numbers by document, matched and unmatched sample (defined below)

which are thought to contain little discriminating information were removed), and subject to a 'sparsity' condition of 99-percent (i.e., only words that occur in more than one-percent of all documents but in no more than 99-percent of all documents are included). The resulting DTM for analysis is matrix with dimensions $163,958 \times 3,755$.

4 Methods

Our claim above is that the secrecy endemic to diplomacy comes in at least two separable varieties: substantive secrecy—the notion that certain information is to be kept confidential because it would be *per se* damaging to security were it released—and procedural secrecy, which is concerned with the notion that secrecy protects foreign or domestic agents from outside consequences of their actions. To assess the evidence for these separate ideas, some care is required in terms of methods. Here we explain our approaches.

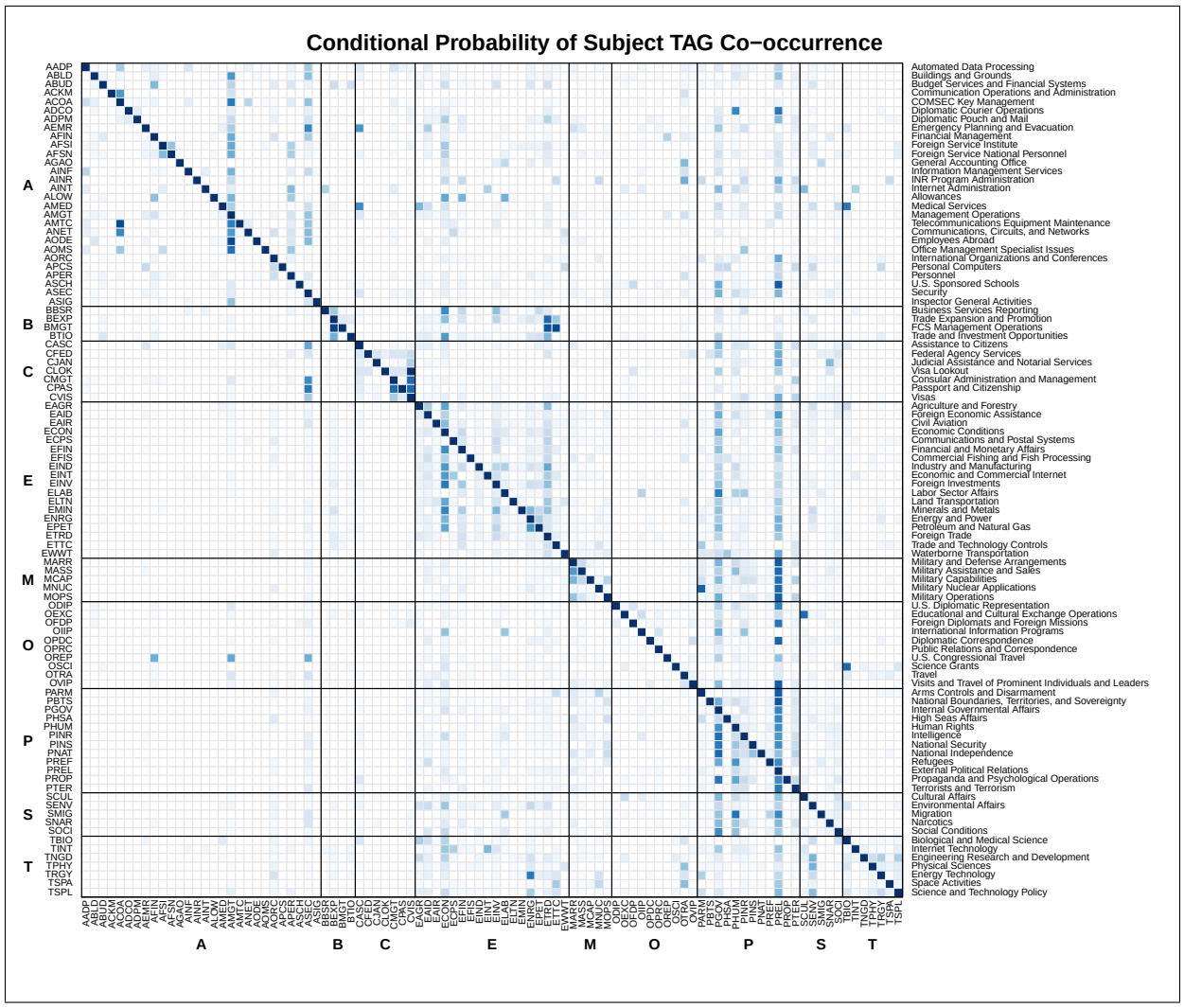

Figure 3: This figure shows the conditional probability of U.S. State Department subject TAG co-occurrences in the post-2005 sample ($n = 163,958$). Subject tags are are presented in alphabetical order with their official U.S. State Department meanings listed in the righthand column. Each cell in the figure represents the conditional probability that a *column subject* will be tagged given that a *row subject* has already been tagged. Darker shaded cells indicate a higher conditional probability.

4.1 Substantive Secrecy

We first examine the question of *substantive secrecy*—i.e., how the topic or substance of a diplomatic communication, all else equal, influences its probability of restriction. The objective is to quantify both the magnitude and direction of how official U.S. State Department communication subject tags affect cable secrecy. In suit, we regress each cable's observed restriction status on its subject tags and location of origin. This fixed-effects least squares equation can be written as follows:

$$R_i = \alpha_0 + \sum \beta_t Tag_{it} + \gamma_j + \varepsilon_{ij} \tag{1}$$

where R_i is a dummy variable for cable i that takes the value of 1 if the cable is restricted and 0 if unrestricted, Tag_{it} is a subject tag dummy variable for cable i and tag t, and γ_j is the fixed effect for embassy j. Standard errors are clustered at the embassy level. Given that each covariate in this regression is binary, each regression coefficient $\hat{\beta}_t$ is a sample estimate of the difference between two conditional expectations: the conditional probability a document will be restricted given the presence of a subject tag minus the conditional probability of restriction without that subject tag present.[3]

4.2 Procedural Secrecy

Recall that procedural secrecy concerns the diplomatic norms of confidentiality in meetings. If it exists as a quantity that can be identified in our data, then it should emerge as a key discriminator between restricted and unrestricted cables. However, if there is, indeed, a

[3]That is, $\hat{\beta}_t = Pr(R_i = \widehat{1|Tag_t} = 1, \mathbf{X}) - Pr(R_i = \widehat{1|Tag_t} = 0, \mathbf{X})$. If standard assumptions hold—e.g., unconfoundedness, overlap, or "selection on observables" (Rosenbaum and Rubin, 1983; Heckman and Robb, 1985)—$\hat{\beta}_t$ tells us on average how much document restriction varies by each subject tag in our sample. Although the outcome of interest is binary, OLS is appropriate when the conditional expectation function (CEF) of each regressor with respect to the outcome exhibits is linear (see, e.g., Angrist and Pischke, 2009, Chapter 3). OLS suitably estimates whereby sample average effect of restriction on each subject tag in the context of our data, as each regression coefficient represents a conditional mean.

subject tag imbalance between restricted and unrestricted cables (as suggested above), this implies that a simple comparison of word frequencies between restricted and untristricted documents is unlikely to isolate how text varies marginally as a function of secrecy status, since observed differences are likely to arise directly from initial differences in subject matter.

Thus, the question we ask in this section is: having adjusted for cable subject matter (given an observed sequence of subject tags on a document) and locations of origin, all else equal, can restricted diplomatic communications be distinguished from unrestricted communications? This question may be thought of as estimating the *marginal effect* of secrecy on the content of a restricted communication. In other words, given two documents indexed with identical subject tags and originating from the same source, are there specific textual features that systematically distinguish restricted cables from unrestricted cables? If such textual features exist, is there anything substantively unifying about these features? In particular, does whatever differentiates these communications properly be considered 'procedural' in nature?

4.2.1 Exact Matching on Subject Tags and Origin

To assess whether secrecy, on the margin, is associated with differences in document composition, we restrict our sample to *exactly matched* subsets of cables within each embassy in our sample. More precisely, for each embassy (i.e., each cable's location of origin), we implement the algorithm outlined in Figure 11 in Appendix D to construct datasets of cable pairs that are exactly matched on official U.S. State Department subject tags and their embassies of origin, but differ on their restriction level. The objective of this matching procedure is to restrict the full sample such that there is perfect subject overlap on cables in our study. As a result of the matching procedure, within each embassy, for each restricted cable there will exist an unrestricted cable that has an identical subject tag pattern. We rely only on the State Department's official subject tags for this procedure. If two or more unrestricted

matches are found for a single restricted cable, we select the match that is written most closely in time to the restricted cable's date of authorship. For the results presented in this study, matching is performed without replacement, and datasets are stored and analyzed at the embassy level (although pooled analyses are also appropriate with the resulting data).

Since we wish to make inferences about textual differences between restricted and unrestricted cables on the margin—i.e., once cable subjects have been accounted for—the within-embassy matched sampling design has intuitive appeal. The sampling design allows for a meaningful examination of procedural secrecy. Adjusting the sample directly for differences in subject matter and controlling for embassy-level effects, the design allows us to isolate differences in textual composition that are likely to arise from a document's handling status alone. Intuitively, the aim of our exactly-matched sampling design is to "control" for substantive differences in cables that may be present in the unmatched sample—differences that may arise from hypothetical variation in reporting rules, document disclosure standards, authorship style, or political priorities at the embassy level. If systematic textual differences remain between restricted and unrestricted cables after subject and location have been accounted for, these differences are likely to arise from procedural rules that are separate from subject-specific handling rules.

The formal appeal of exact matching is that it is nonparametric and approximates the act of "blocking" in randomized experiments (Cox, 1958; Imai, King and Stuart, 2008).[4] Ex-

[4]In the context of parametric adjustment, Rosenbaum and Rubin (1983, 1984) show that matching on a correctly specified propensity score (i.e., a unit's conditional probability of being assuaged to treatment, given its covariates) is sufficient to allow for the unbiased of the average effect of a treatment for a given population of interest, i.e., eliminate confounding. But in settings with observational data, a researcher rarely knows whether an appropriate functional relation has been specified in model-based matching procedures (Rosenbaum, 2002). The appeal of an exact covariate matching procedure is that if the appropriate set of conditioning measures has been identified, the unobserved functional relation between between covariates and the assignment to treatment is ignorable due to *perfect balance* on conditioning variables. Under general conditions, exact matching procedures are both *equal percent bias reducing* (Rubin, 1976) and *monotone*

act matching is often untenable in applied research, however, since the sampling procedure can dramatically reduce a researcher's final sample size, and the procedure tends to rely on initially large sample sizes. Unsurprisingly, this was a concern for our modeling attempts, along with the possible danger that many documents dealing with sensitive substantive areas would be jettisoned from the final analysis because no match could be found for them. Further, we were concerned that certain 'important' embassies would be, relative to the original dataset, heavily under-represented.

Neither of these concerns appear to be true of the matched sample. In Figure 4 we report the reduction in subject tag imbalance of the exactly matched sample, in addition to information on which subject tags remain present. In the upper-left subplot, the thin-transparent lines in red (restricted) and blue (unrestricted) correspond to embassy-level averages of individual tag frequencies in unrestricted and restricted cables. The thicker vertical colored lines in the foreground denote sample averages. The lower-left plot provides much of the same information but in slightly different form: background lines correspond to embassy-level imbalances (subject tag differences in means between restricted and unrestricted cables within embassy) whereas the thicker bar plot in the foreground is the sample level difference in means. These two subplots demonstrate there is subject imbalance between unrestricted and restricted cables both on aggregate levels and, generally speaking, at individual embassy levels. In the post-2005 sample, "A – Administrative Affairs" tend to be more public, "B – Business Services" tend to be more public, "C – Consular Affairs" tend to be more public, "E – Economic Affairs" tend to be more public, "M – Military and Defense Affairs" tend to be more private, "O – Outreach" tends to be more public , "P – Political Affairs" tend to be

imbalance bounding (Iacus, King and Porro, 2011). These traits are not generally true for most distance-based or model-based (parametric adjustment) matching methods, which has led several scholars to conclude that exact matching is close to an "ideal" matching procedure in observational settings (e.g., Stuart, 2010; Imai, King and Stuart, 2008).

more private, "S - Social Affairs" tend to be more public, and "T - Technology and Science" issues are slightly more public on average.

In the exactly-matched sample, we see both embassy-level and aggregate level subject imbalances have been eliminated. The upper-right plot shows the within-embassy subject proportions are identical between unrestricted and restricted cables, which is why the colored lines appear purple (due to perfect overlap between the red and blue lines). The exact subject balance is also true for the exactly matched sample average. The lower-right plot demonstrates this point in an extreme form: there exist no embassy-level or aggregate level subject imbalances in the exactly-matched sample. The analysis gives us confidence that inferences following from the matched sample will be appropriate to a broad class of diplomatic communications. The distribution of subject tags in the matched sample map to substantively meaningful political issue areas. The majority of cables in the matched samples have to do with Economic Affairs, Military and Defense Affairs, and Political Affairs—each topic within foreign policy that are closely related to formal theories of rational diplomacy. On the other hand, the exact-sampling design is less capable of making credible counterfactual statements about Administrative Affairs and Outreach.[5]

In Figure 5 we report the embassies, and their relative prevalence, in our matched data. Importantly, we see some 'big' embassies—including the U.S. State Department itself—are

[5]The reason subject tag overlap is important is because exact matching will allow us to inspect textual differences akin to *treatment effects on the treated*. Treatment effects on the treated are not the same as the average effect of treatment unconditionally, nor are they the average effect of treatment in the sample. SATT (sample average treatment effect on the treated) is an estimate of how much potential outcomes would differ for the set of treated units in the sample if they were instead to become untreated: $\tau_{ATT} = \mathbb{E}[Y(T = 1|T = 1, X)] - \mathbb{E}[Y(T = 0|T = 1, X)]$, using super-population notation . In the present study, therefore, with our exactly-matched sample, the design allows us to inspect questions like the following: If a set of treated (i.e., restricted) documents like those in our sample were instead to become untreated (i.e., unrestricted), on what textual dimensions would we expect to the documents to vary? This is not the same as the unconditional marginal effect of secrecy.

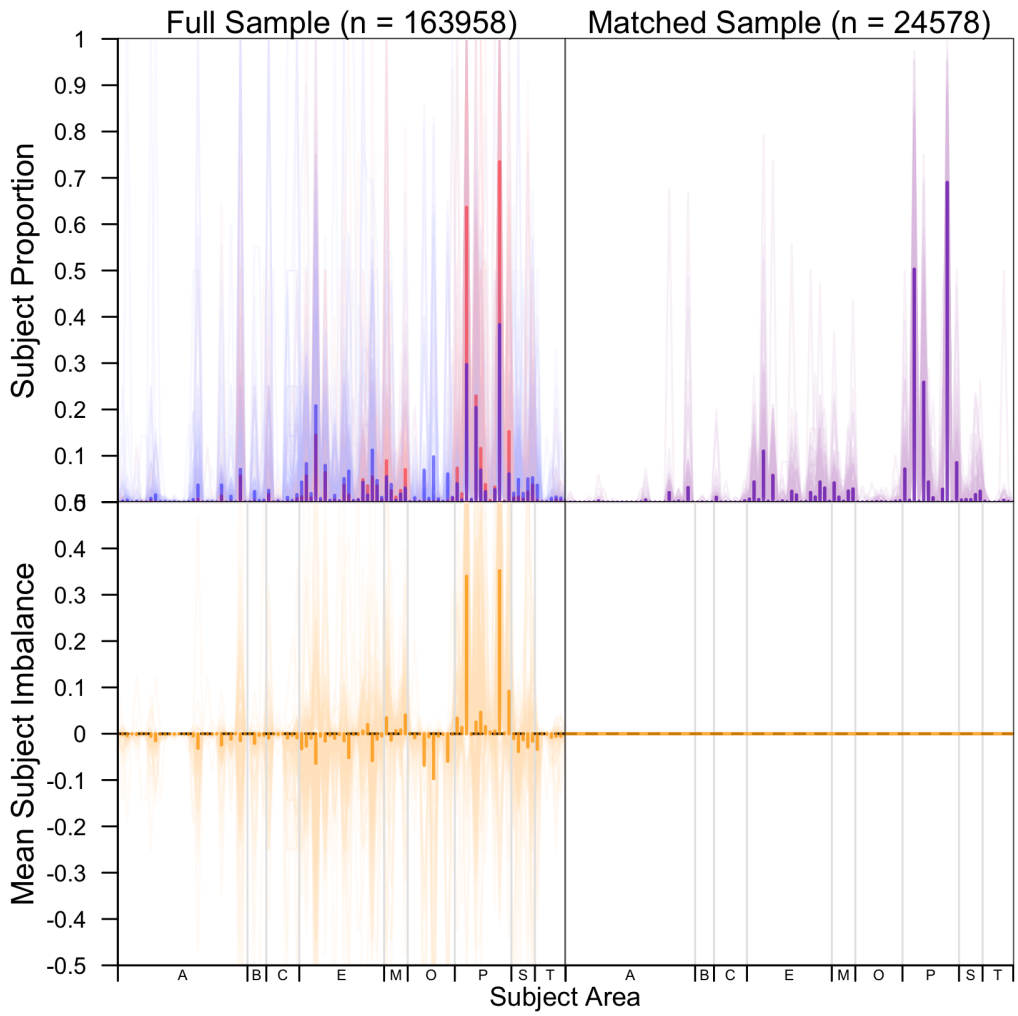

Figure 4:

represented; in particular, Ankara, Baghdad, Paris, Cairo and Moscow (all centers of activity in the original data) appear. Taken alongside the results of Figure 4, this presents strong evidence that the matched sampling procedure does not leave the general patterns of the whole sample too far behind, and is due to the fact that there are strong within-embassy subject correlations. The diplomatic locations contained in the study sample are represented in a manner proportionate to their overall representativeness in the full sample.

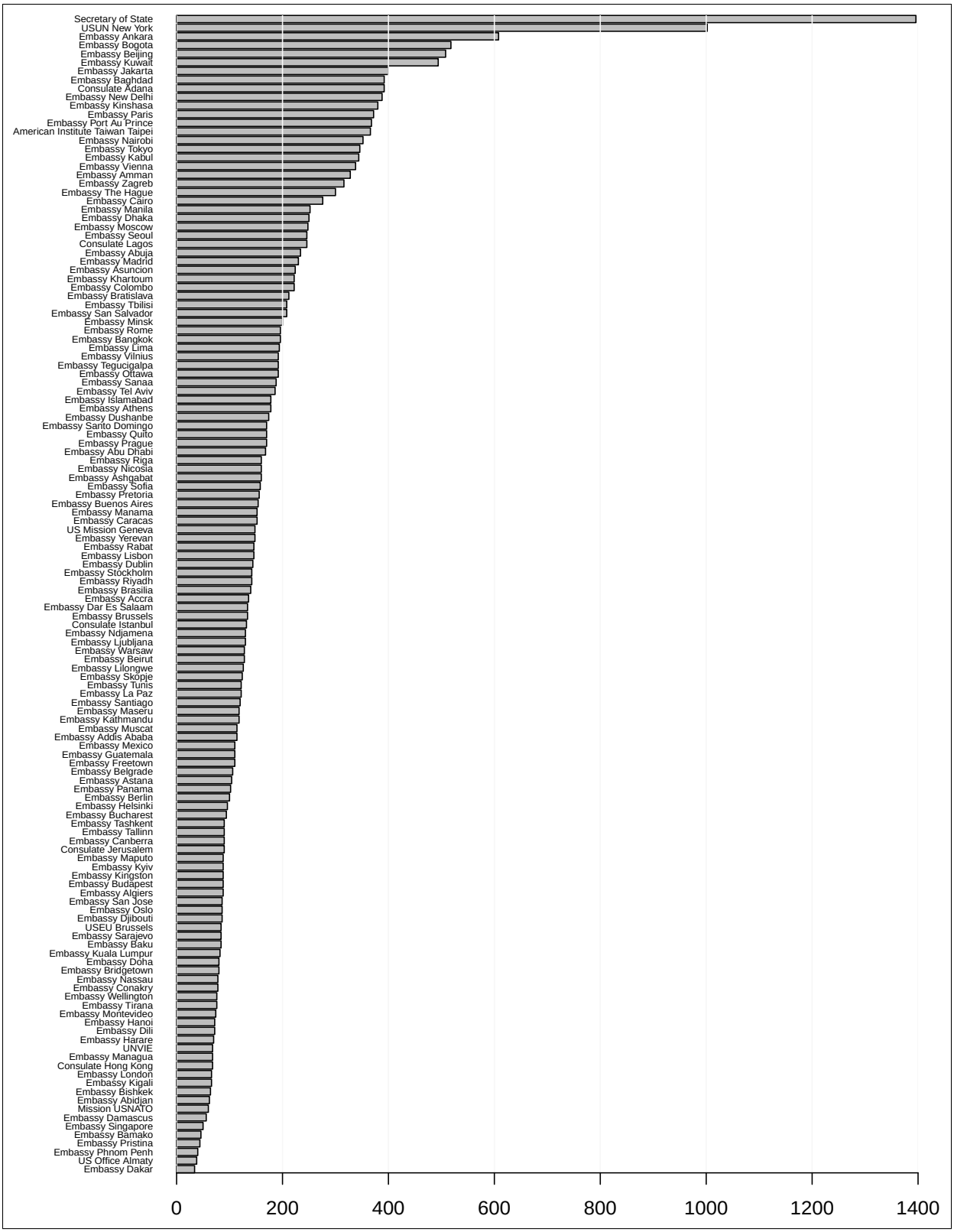

Figure 5: Counts of cables by embassy in the matched sample.

4.2.2 Supervised Learning and Penalized Regression

For each of the matched samples described in Section 4.2.1, we implement a set of supervised learning models to identify which words are most *important* to (i.e., predictive of) cable secrecy. The matrix of words used in this classification setting is taken from the full post-2005 document-term matrix described before, but now only includes rows that satisfy the within-embassy, exactly-matched sampling design. On the 'left hand side' we have the (binary) restriction status of a given document which we intend to predict with the words within that document. Quantitatively, we observe how within-sample classification error rates vary as a function of which words are included in the model; qualitatively, we wish to make statements about how a document's restriction status would likely change if particular words within these documents were to vary.

Two supervised learning methods are applied to these data: the "random forest" (hereafter RF) algorithm (Breiman, 2001), and the "lasso" (Tibshirani, 1994). Results from both procedures are used alongside the topic model estimates described below to make statements both at the world-level and topic-level about how secrecy, on the margin, influences the content of diplomatic communications. The RF and lasso procedures require brief explanation as they are not widely used in political science research (though inevitably many technical details will be left for readers consult in the works cited). Both are widely used in "small n, large p" settings—cases in which there may be there may be a greater number of possible parameters than observations in the sample.

The RF algorithm is a decision tree and resampling-based classification procedure which relies on repeatedly dividing the observed sample of data into random bootstrapped training datasets and fitting decision trees to each random training set, then aggregating the classification results over all independent training sets. In the statistical learning literature,

this procedure is commonly referred to as bootstrapped aggregation (i.e., "bagging"), and can be widely applied to improve the classification precision of various models, regression included. A RF algorithm procedure deviates from bagging alone by also randomly sampling the parameter space included in each iteration of this bagging procedure (e.g., Ho, 1998). One result of procedures like RF is it allows researchers to think about the relative variable importance of predictors in a classification setting. Due to the fact that at each bagged iteration of the procedure there are random subsets of the feature space included in the decision-trees, not all predicting variables (i.e., "words" in our context) are likely to be included as predictors at each stage of the algorithm. Overall, a predictor's *variable importance* can be thought of as a result of this process: an estimate of the marginal reduction in classification error that results from a single word's inclusion to the classification procedure overall, given the random inclusion of other predictor variables.

The lasso is a form of penalized regression, similar to ridge regression, whereby regression coefficients are weighted by "shrinkage factors" such that regression coefficients are weighted towards zero (Tibshirani, 1994; Hastie, Tibshirani and Friedman, 2009). The lasso is commonly used for feature selection in high-dimensional learning problems to decrease the variance of a particular classifier. In our context, the procedure is similar to an ordinary least squares regression procedure in which the best-model is determined by that which minimizes the in-sample sum of squared residuals, except regression coefficients are penalized according to prior rules (i.e., the shrinkage factor and tuning factor) on the minimum coefficient size a variable is allowed to have to be included in the final classification model.

With both RF and the lasso, we obtain embassy-level estimates of word-level dependencies to document restriction. In the context of RF, each exactly-matched dataset for embassy j has a corresponding vector of *word importances*, where importance is defined as an estimate of

each variable's in-sample average marginal error reduction. In the context of the lasso, each embassy has a corresponding vector of penalized partial regression coefficients. For both the RF and the lasso procedures, we refer to this collection of embassy importance vectors as the embassy *importance matrix*. Each row in this matrix represents a given embassy, and each column is a measure of a word's relative importance to prediction accuracy in the embassy's matched sample. Each cell entry is then the Random Forest importance measure for that term in that embassy. To obtain sample-average estimates of word-level importances to prediction, we weight weight the results of each embassy-level importance vector by its relative share of all cables in the exactly matched sample. The prevalence of any given embassy in the matched sample, therefore, proportionately weights the importance terms associated with that embassy (thus, for example, we will up-weight the importance terms associated with the State Department itself and other embassies near the top of Figure 5). Using the sample-weighted results of the RF within-embassy, exactly-matched classification procedure, we then took the top 30 of these terms (recall that they are all positively signed, regardless of their actual signed effect on classification), and looked up their corresponding coefficients from the lasso regressions. The lasso regression coefficients are similarly weighted as sample averages.

4.2.3 Supplementary Analysis: Topics

Some supplementary analyses are performed to address of the marginal effect of secrecy. In particular, we *topic model* our sample of $n = 163,958$ cables, using the most common probabilistic topic model in contemporary text analysis research, Latent Dirichlet Allocation, henceforth referred to as LDA (Blei et al, 2003). Information on our topic modeling procedure is outlined in Appendix C. Results of the topic modeling procedure are used as a descriptive aid to categorize the words we find to be predictive of document restriction.

5 Results

5.1 Substantive Secrecy

Recall that testing for substantive secrecy boils down to testing whether or not the probability a diplomatic cable is withheld from the public is measurably predicted by the *subject* of the cable communication, adjusting for the cable's location of origin and other factors. Figure 6 presents this analysis, where each point corresponds to an estimate of the sample average effect of a subject TAG on the probability of the cable's restriction. Around each estimate is the 95-percent confidence interval. In terms of coefficient direction, note that the broken line in the center of the plot denotes a point estimate of zero 'effect': tags to the right of this line are generally associated with restricted documents (on average); the presence of tags to the left, generally predict an unrestricted status for the cables. Tags highlighted in red indicate coefficients that are statistically differentiable from zero. Our first observation is that there are a large number of statistically significant predictors: almost every subject matter tag is associated with increasing or decreasing the probability that a particular cable is restricted. Second, we note that the direction of the effects are somewhat in line with our priors. Thus we see that cables concerning "Terrorists and Terrorism", "Military Capabilities", "Intelligence," and "National Independence," for example, are more likely to be kept private than cables concerning "Migration" "Narcotics," "Personnel," or "Environmental Affairs." In particular, we see that dispatches dealing with 'core' state secrets, especially pertaining to information, capabilities and threats are restricted. On the other hand, cables that discuss more 'public good' orientated matters—wherein we can imagine that sharing information may not be damaging, and may in fact be optimal—tend to be unrestricted. In this latter category are tags that seem to require or be synonymous with publicity and the dissemination of information: "International Information Programs", "Public Relations and Correspondence", "International Organizations and Conferences", "Educational and Cul-

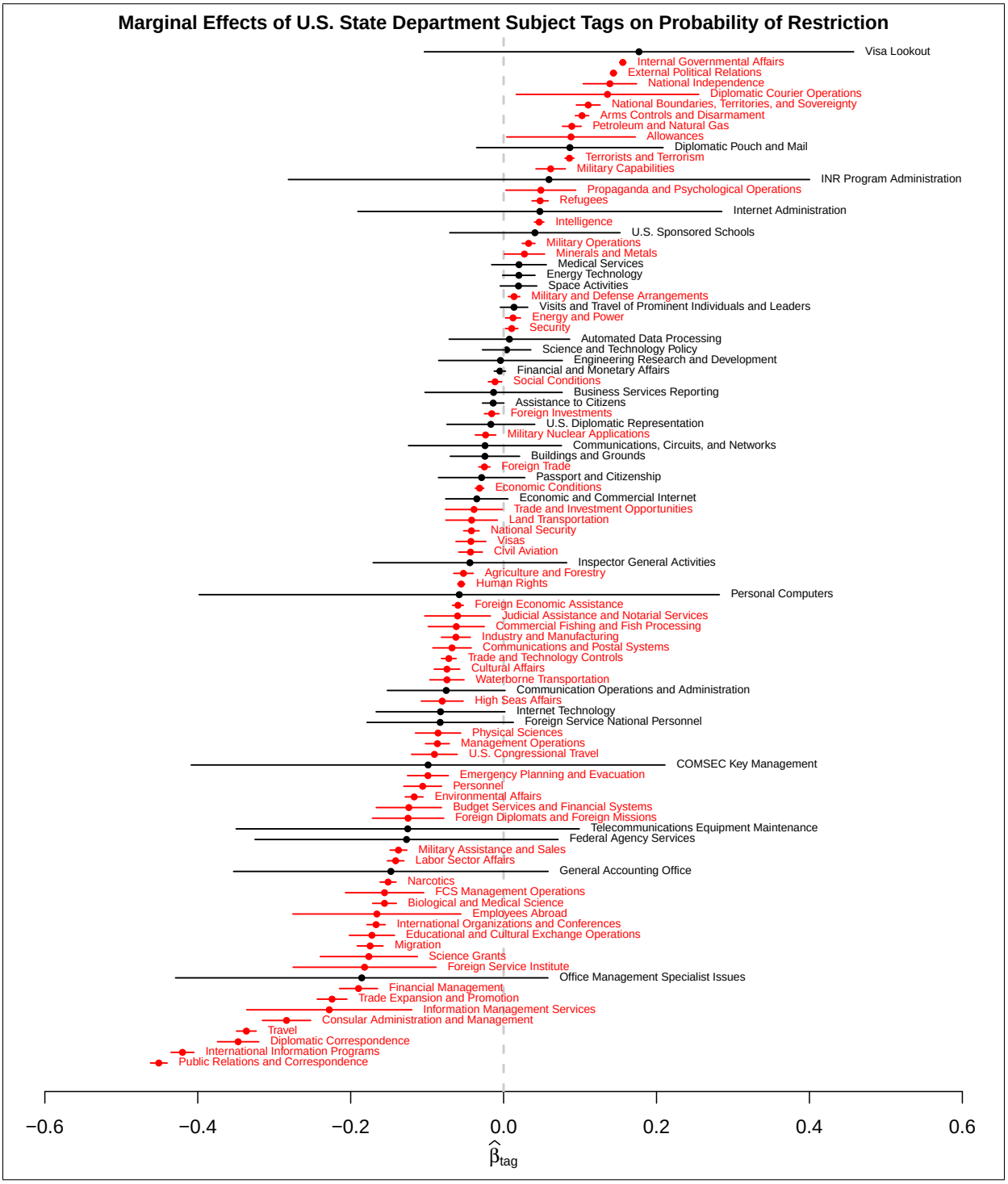

Figure 6: Substantive content as a predictor of secrecy status: red point estimates are statistically distinguishable from zero. 95% CIs around each estimate. All cables written in and after 2005 used in estimates, along with cable-origin fixed-effects. The central broken line corresponded to a $\hat{\beta}$ of zero.

tural Exchange Operations" and so on.

The fact that cable substance drives at least some part of diplomatic secrecy should not come as a surprise to theorists of rational diplomacy. As noted above, most contemporary theoretical treatments of crisis diplomacy concern agents' incentives to misrepresent their resolve, capabilities, or information in bargaining settings: our results here suggest the U.S. acts in a way compatible with that logic.

5.2 Matched Sample Results: Procedural Secrecy

In terms of procedural secrecy, an overview of our main results may be found in Figure 7. Recall that we used the RF algorithm to identify the thirty 'most important' tokens in terms of their ability to discriminate between the unrestricted and restricted cables status of a document. In the second column of the plot, these are clearly seen and include words such as 'said', 'told', 'ambassador', 'want', 'note', 'meet', 'want', 'ask', 'discuss', 'concern', 'state', 'agre[e]' 'support', 'however', 'thank', 'request', 'possibl[e]', 'like' and so on. Our immediate observation is that in stark contrast to our tag regressions, these words do not connote substantive state secrets *per se*; rather, they refer to the holding of meetings and the general protocols of diplomatic exchange with foreign nationals. Related to this idea, note the presence of terms such as 'poloff' (the Embassy's Political Officer), 'usg' (United States Government) and 'minist' (minister): actors who we expect to be involved in daily embassy interactions. On the left of the figure, we report the lasso (point) estimate associated with the terms. When these points are to the right of the vertical line, the use of that word (on average) increases the probability that a document is restricted; when the points are to the left, this suggests that the word is associated (on average) with a decrease in probability that a document is restricted. Examining that part of our results, we note that terms such as 'said' and 'told', 'request', 'like' are used disproportionately in restricted cables. To us,

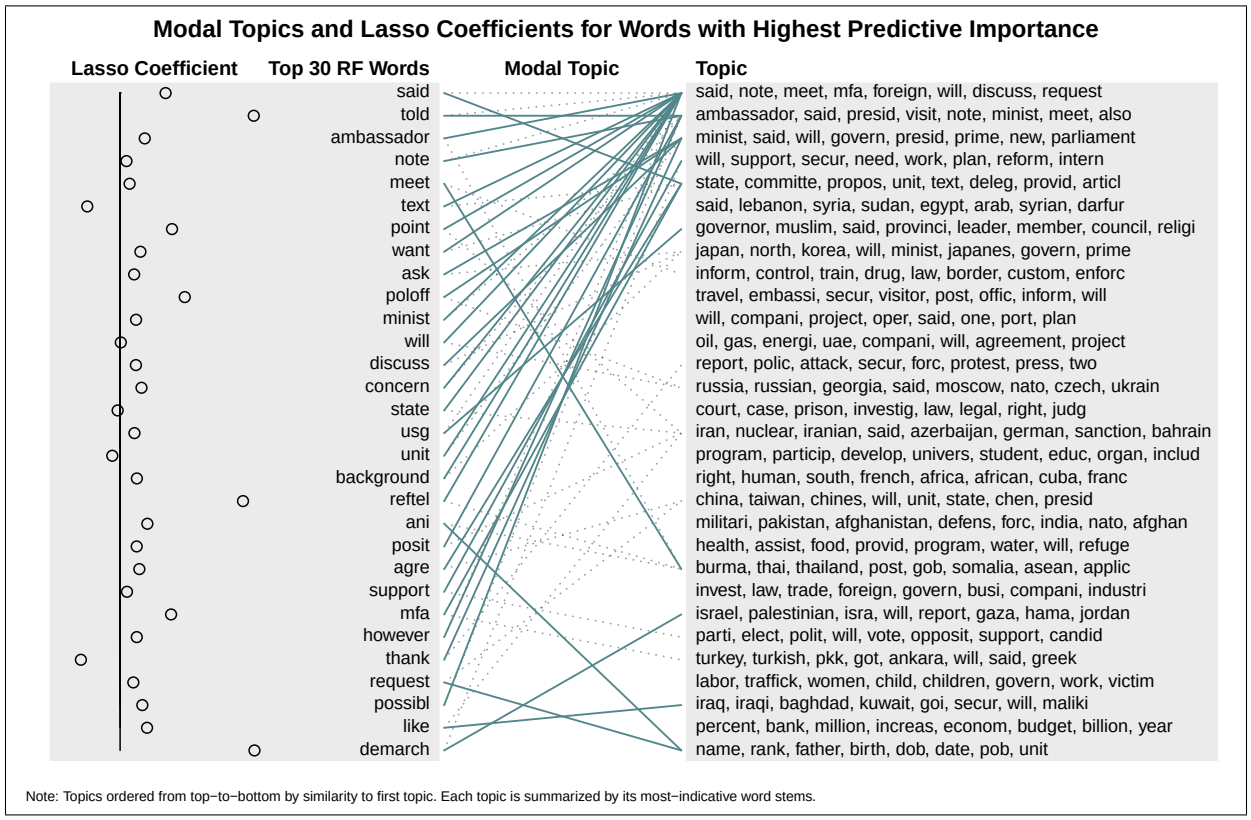

Figure 7: Procedural secrecy: matched sample results. For each of the words listed on the lefthand side of the plot, a solid line maps that word to its most likely topic (given estimates from the LDA model described in Appendix C). A dotted line maps each word to its second most likely topic. The topics listed on the righthand side of the graph are ordered in a specific manner: the uppermost topic is the mode of the modal topic assignments (i.e., the topic that is most frequently the modal topic assignment for the top RF terms), while subsequent topics are presented in descending order according to their semantic similarity to the first topic. Semantic similarity determined by the cosine similarity between topic vectors. The plot reveals remarkable concordance on the following: words that are most predictive of secrecy tend to be used in semantically similar topics, and those topics tend to concern the official business of foreign leaders, their meetings, and words relating to information exchange.

this is evidence that once one controls for substantive area, secrecy is mostly about keeping meetings private and confidential, regardless of whether anything intrinsically 'secret' is being discussed.

To check this intuition, we looked up the modal topic—from the topic model described earlier—in which our 'top' words appeared. If we are correct that secrecy is partly about a norm of discretion rather than content, we would expect to see most of the terms mapping to a single (or perhaps a few) 'administrative' topic(s), rather than topics pertaining to matters of substantive import. On the right-hand side of the plot, we see this is almost entirely the case. There, the solid black lines lead from each word to the topic it most likely belongs; the dashed lines are from each word to second most likely topic. We see first that with a few exceptions, all of the words 'belong' to the first, second, or third topics. Inspecting those more closely, we note that those topics generally consist of administrative nouns and verbs, rather than subjects of interest: thus, we find "said" in the first, second, and third topic as a leading word, while 'will' appears in the fourth topic. Importantly, the words that we have identified as discriminating between unrestricted and restricted cables do *not* appear alongside obviously substantive subject matter such as pertains to the Middle East (topic six or topic seven), the Pacific rim (topic 8), nuclear proliferation (topic 16) or Russian aggression (topic 14). Of course, we do see that some terms are likely to appear within certain substantive topics (such as 'meet', which appears in a 'Burma' topic and 'demarch[e]' which appears in an Israel topic towards the bottom of the plot). Such occurrences are not the norm, however.

In terms of the theories presented earlier, our finding here seems most closely compatible with the work of Sartori (2002) and Kurizaki (2007) insofar as privacy seems to be intrinsically valued by diplomats, rather than because it allows *per se* information exchange.

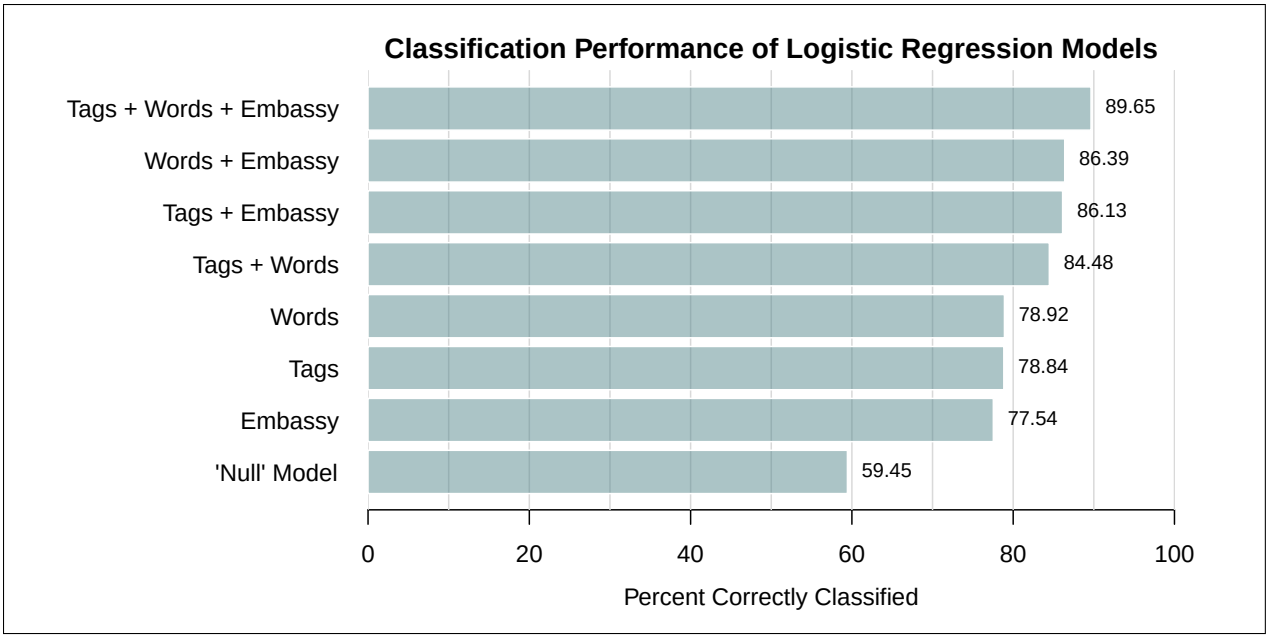

Figure 8: Comparison of Logistic Regression Models (matched sample). Null model represents the modal category of the data set (i.e. restricted).

5.3 Share of Secrecy: Substance vs Procedure

Above, we made the claim that while some of observed diplomatic censorship is a consequence of the need to protect state secrets, at least part of it results from the need to keep meetings confidential as a procedural requirement, regardless of what is to be discussed. In our final set of results, we attempt to estimate the relative contribution that these two separate elements make to the practice of restricting information from public view. In Figure 8 we report a comparison of models with this in mind. Here, 'Tags' refers to the tag covariates we noted earlier, 'Embassy' are simply embassy fixed effects, and 'Words' are the top 50 words selected by the Random Forest procedure above. In all cases, the numbers to the right of the bars refer to the percent correctly predicted (unrestricted and restricted) by a given (logit) model in the entire sample of 160000 documents.

Unsurprisingly, we see that a model with tags, the word information and the embassy

fixed effects does best in terms of the proportion of documents it can classify correctly, at around 90%. The null model, the sample proportion of restricted cables is 59%, and clearly the statistical model improves substantially upon this. More interesting from our perspective is a comparison of the second and third bar ('Words + Embassy' and 'Tags + Embassy'), and the fifth and sixth ('Words' and 'Tags') since the performance of the models using the RF words and tags are so similar. That is, it seems that whether we use the substantive topics alone, or the words that we identified as connoting secret meetings rather than substance, our model performs similarly. This suggests, at the very least, that both substantive and procedural secrecy matter for diplomatic communication, and that both the audience cost theories and more recent work on communication have some support in the data.

6 Discussion

Conflict and bargaining is at the core of international relations, and the discipline has amassed an impressive array of theoretical models that make use of, or provide findings for, 'information' and its exchange between actors. This paper opened by noting that, despite this voluminous literature, there is little statistical work on the subject, and that this is hardly surprising given that secrets—by definition—are difficult to research. In this paper, we made use of the WikiLeaks 'Cablegate' disclosure of diplomatic communications, a new and modern dataset that has a good deal of 'uncensored' content (that was not systematically edited), to examine the empirical support for various conceptions of secrecy and communication. We argued that diplomatic confidentiality, i.e. information actively kept from the public, is used in at least two situations or 'dimensions': first, in a way pertaining to *substance* and second, pertaining to *procedure*. In the former case, documents that deal with issues that could damage U.S. capabilities were they available to others, are disproportionately kept secret. Meanwhile, in cases where publicity is helpful to the U.S. government,

perhaps because it creates an 'audience cost' and encourages commitment to a costly path of action, cables are made available. In the second case, that of the procedural dimension, diplomats ensure that the circumstances and process of meetings in general—regardless of their actual subject content—are not disclosed. To be clear, we found evidence of both dimensions in our data, and were able to characterize their content and nature. In this way, both the recent literature that emphasizes the importance of diplomacy, and the earlier 'rationalist' literature that has it playing little role, finds some support here.

Apart from the preliminary analysis our paper provided, it also contributed methodologically to a growing area of political science: that of text analysis. In particular, we were faced with a situation in which documents had to be compared within particular subject areas, such that their discriminatory terms could be uncovered. We used an exact matching algorithm to get at the estimand of interest. In our case, the subject matter was determined by the U.S. State department (via the TAGS system), but the problem is obviously more general than this. For example, one might be interested in the success (or otherwise) of different bills in Congress or the public opinion reception of speeches from primary candidates. Clearly, the subject matter between documents differs and needs to be 'controlled' for in some sense. We provided one way of proceeding in such a situation.

Of course, analytically, we have only scratched the surface here. Though we have documented the nature and structure of secrecy and the cables themselves, there is much more to do. First, while we argue that the 'more secret' topics in the TAGS system seem to deal more fully with capabilities than the 'least secret', we are necessarily vague on the details. We would like to know more about why exactly some subjects are kept from public view, and whether such decisions accord with IR theory in the area: for example, is topic secrecy actually dictated by a desire to avoid revealing capabilities on a particular subject, or is

it more connected to notions of resolve, or even just the information gathering machinery itself? Second, we argued that the content of the cables is consistent with both the audience cost literature, and its recent critics, insofar as we find some evidence that the U.S. attempts to broadcast its views (and thus possibly create such 'audience costs') where helpful, but not always. That is, it seems to preserve 'room for manoeuver' on some topics. Subsequent analysis might weigh in more helpfully on this debate by considering the constraints that the U.S. actually faces in the various areas of international relations with which it deals: for example, we might be interested to know whether, in fact, issues that the U.S. is seemingly 'open' about with the public are simply those where it cannot be otherwise given commonly held knowledge about the U.S. position (or it weaknesses) in the wider world. This is ultimately a call to incorporate more topic specific covariates and circumstances in the analysis. Finally, while we have emphasized the importance of private diplomatic meetings as part of the arsenal of U.S. international relations practice, we have done little to explain how or why they are used. That is, we are not much the wiser as to which of the various theories of diplomatic exchange is correct, if any. We leave such questions for future work.

A Cable Tags

	Meaning		Meaning
1) AADP	Automated Data Processing	51) ELTN	Land Transportation
2) ABLD	Buildings and Grounds	52) EMIN	Minerals and Metals
3) ABUD	Budget Services and Financial Systems	53) ENRG	Energy and Power
4) ACOA	Communication Operations and Administration	54) EPET	Petroleum and Natural Gas
5) ACKM	COMSEC Key Management	55) ETRD	Foreign Trade
6) ADCO	Diplomatic Courier Operations	56) ETTC	Trade and Technology Controls
7) ADPM	Diplomatic Pouch and Mail	57) EWWT	Waterborne Transportation
8) AEMR	Emergency Planning and Evacuation	58) MARR	Military and Defense Arrangements
9) AFIN	Financial Management	59) MASS	Military Assistance and Sales
10) AFSI	Foreign Service Institute	60) MCAP	Military Capabilities
11) AFSN	Foreign Service National Personnel	61) MNUC	Military Nuclear Applications
12) AGAO	General Accounting Office	62) MOPS	Military Operations
13) AINF	Information Management Services	63) ODIP	U.S. Diplomatic Representation
14) AINR	INR Program Administration	64) OEXC	Educational and Cultural Exchange Operations
15) AINT	Internet Administration	65) OFDP	Foreign Diplomats and Foreign Missions
16) ALOW	Allowances	66) OIIP	International Information Programs
17) AMED	Medical Services	67) OPDC	Diplomatic Correspondence
18) AMGT	Management Operations	68) OPRC	Public Relations and Correspondence
19) AMTC	Telecommunications Equipment Maintenance	69) OREP	U.S. Congressional Travel
20) ANET	Communications, Circuits, and Networks	70) OSCI	Science Grants
21) AODE	Employees Abroad	71) OTRA	Travel
22) AOMS	Office Management Specialist Issues	72) OVIP	Visits and Travel of Prominent Individuals and Leaders
23) AORC	International Organizations and Conferences	73) PARM	Arms Controls and Disarmament
24) APCS	Personal Computers	74) PBTS	National Boundaries, Territories, and Sovereignty
25) APER	Personnel	75) PGOV	Internal Governmental Affairs
26) ASCH	U.S. Sponsored Schools	76) PHSA	High Seas Affairs
27) ASEC	Security	77) PHUM	Human Rights
28) ASIG	Inspector General Activities	78) PINR	Intelligence
29) BBSR	Business Services Reporting	79) PINS	National Security
30) BEXP	Trade Expansion and Promotion	80) PNAT	National Independence
31) BMGT	FCS Management Operations	81) PREF	Refugees
32) BTIO	Trade and Investment Opportunities	82) PREL	External Political Relations
33) CASC	Assistance to Citizens	83) PROP	Propaganda and Psychological Operations
34) CFED	Federal Agency Services	84) PTER	Terrorists and Terrorism
35) CJAN	Judicial Assistance and Notarial Services	85) SCUL	Cultural Affairs
36) CLOK	Visa Lookout	86) SENV	Environmental Affairs
37) CMGT	Consular Administration and Management	87) SMIG	Migration
38) CPAS	Passport and Citizenship	88) SNAR	Narcotics
39) CVIS	Visas	89) SOCI	Social Conditions
40) EAGR	Agriculture and Forestry	90) TBIO	Biological and Medical Science
41) EAID	Foreign Economic Assistance	91) TINT	Internet Technology
42) EAIR	Civil Aviation	92) TNGD	Engineering Research and Development
43) ECON	Economic Conditions	93) TPHY	Physical Sciences
44) ECPS	Communications and Postal Systems	94) TRGY	Energy Technology
45) EFIN	Financial and Monetary Affairs	95) TSPA	Space Activities
46) EFIS	Commercial Fishing and Fish Processing	96) TSPL	Science and Technology Policy
47) EIND	Industry and Manufacturing		
48) EINT	Economic and Commercial Internet		
49) EINV	Foreign Investments		
50) ELAB	Labor Sector Affairs		

Figure 9: List of diplomatic cable tags and their meanings

B Cable restrictiveness by Embassy

Cable Totals and Restriction Frequency by Place of Origin: post−2005

Origin	# U	# R	% R	Origin	# U	# R	% R	Origin	# U	# R	% R
1) Embassy Baghdad	4970	823	0.14	90) Embassy Wellington	219	410	0.65	179) Mission USOSCE	108	16	0.13
2) Secretary of State	1772	3456	0.66	91) Embassy Maputo	183	442	0.71	180) REO Kirkuk	116	5	0.04
3) Embassy Tokyo	1845	3151	0.63	92) Embassy Brussels	386	224	0.37	181) Embassy Bujumbura	54	56	0.51
4) Embassy Ankara	2689	1787	0.40	93) Embassy Guatemala	186	392	0.68	182) Embassy Luxembourg	62	48	0.44
5) American Institute Taiwan, Taipei	1431	1502	0.51	94) Embassy Budapest	440	135	0.23	183) Consulate Frankfurt	17	88	0.84
6) Embassy Paris	1685	1248	0.43	95) Embassy Stockholm	300	274	0.48	184) Consulate Shenyang	97	7	0.07
7) Embassy Moscow	2172	449	0.17	96) Embassy Bucharest	434	137	0.24	185) Consulate Ho Chi Minh City	68	32	0.32
8) Embassy Tel Aviv	1245	1280	0.51	97) Embassy Lisbon	243	322	0.57	186) Embassy Lome	18	81	0.82
9) Embassy Beijing	1737	787	0.31	98) Embassy Panama	345	217	0.39	187) Embassy Grenada	0	97	1
10) Embassy Madrid	747	1705	0.70	99) Embassy Oslo	403	134	0.25	188) Consulate Kolkata	13	78	0.86
11) USUN New York	1068	1364	0.56	100) Embassy Conakry	394	136	0.26	189) Embassy Windhoek	49	39	0.44
12) Embassy Bangkok	1380	985	0.42	101) Embassy Djibouti	290	237	0.45	190) Consulate Rio De Janeiro	7	79	0.92
13) Embassy New Delhi	1546	773	0.33	102) Embassy San Jose	138	385	0.74	191) Consulate Vladivostok	0	84	1
14) Embassy Jakarta	1498	615	0.29	103) USEU Brussels	337	183	0.35	192) Consulate Monterrey	46	36	0.44
15) Embassy Kuwait	1037	1058	0.51	104) Consulate Sao Paulo	30	489	0.94	193) Consulate Karachi	75	4	0.05
16) Embassy Cairo	1574	413	0.21	105) Embassy Hanoi	186	332	0.64	194) Consulate Nogales	0	78	1
17) Embassy Beirut	1746	144	0.08	106) Embassy Riga	251	258	0.51	195) Embassy Mbabane	12	60	0.83
18) Embassy Kabul	1072	703	0.40	107) Embassy Doha	403	102	0.20	196) Consulate Halifax	0	71	1
19) Embassy Amman	1113	642	0.37	108) Embassy Maseru	107	389	0.78	197) Embassy Valletta	31	39	0.56
20) Consulate Jerusalem	1232	519	0.30	109) Embassy Dar Es Salaam	223	270	0.55	198) Consulate Lahore	60	7	0.10
21) Embassy Caracas	1563	162	0.09	110) Embassy London	309	172	0.36	199) Mission UNESCO	20	47	0.70
22) Embassy Seoul	1066	640	0.38	111) US Mission Geneva	224	256	0.53	200) Consulate Chennai	16	47	0.75
23) Embassy Dhaka	778	913	0.54	112) Mission USNATO	363	109	0.23	201) Embassy Port Moresby	39	23	0.37
24) Embassy Bogota	1032	636	0.38	113) Embassy Pristina	324	136	0.30	202) Consulate Mumbai	22	38	0.63
25) Embassy The Hague	702	942	0.57	114) Embassy Accra	211	247	0.54	203) Consulate Cape Town	14	45	0.76
26) Embassy Islamabad	1089	514	0.32	115) US Interests Section Havana	411	45	0.10	204) Consulate Munich	23	35	0.60
27) Embassy Mexico	308	1281	0.81	116) Embassy Asmara	406	49	0.11	205) Embassy Brazzaville	3	51	0.94
28) Embassy Colombo	1081	453	0.30	117) Embassy Ndjamena	241	212	0.47	206) Consulate Dusseldorf	0	51	1
29) Embassy Buenos Aires	636	881	0.58	118) Embassy Abidjan	318	133	0.29	207) REO Mosul	49	2	0.04
30) Embassy Khartoum	1093	411	0.27	119) Consulate Adana	0	450	1	208) Consulate Hamburg	8	41	0.84
31) Embassy Abuja	903	431	0.32	120) Embassy Nouakchott	391	55	0.12	209) Consulate St Petersburg	0	45	1
32) Embassy Ashgabat	1023	305	0.23	121) Embassy Singapore	267	156	0.37	210) Consulate Surabaya	0	42	1
33) Embassy Baku	1177	145	0.11	122) Consulate Istanbul	287	128	0.31	211) Embassy Praia	5	35	0.88
34) Embassy Kathmandu	1062	253	0.19	123) Embassy Tripoli	352	62	0.15	212) Consulate Toronto	9	29	0.76
35) Embassy Vienna	503	812	0.62	124) Embassy Helsinki	224	189	0.46	213) US Delegation, Secretary	30	6	0.17
36) Embassy Nairobi	673	627	0.48	125) Embassy Kigali	274	128	0.32	214) Consulate Johannesburg	2	33	0.94
37) Embassy Berlin	936	331	0.26	126) Embassy Freetown	209	192	0.48	215) Embassy Podgorica	0	35	1
38) Embassy Damascus	810	415	0.34	127) Embassy Belgrade	144	254	0.64	216) Consulate Montreal	8	23	0.74
39) Embassy Manila	776	432	0.36	128) Mission Geneva	354	43	0.11	217) Embassy Kolonia	20	11	0.35
40) Embassy Kinshasa	702	497	0.41	129) Embassy Dakar	255	138	0.35	218) Consulate Quebec	14	16	0.53
41) Embassy Rome	807	359	0.31	130) Embassy Tallinn	173	209	0.55	219) Consulate Guadalajara	2	27	0.93
42) Embassy Rangoon	1059	82	0.07	131) UNVIE	271	103	0.28	220) Consulate Naha	25	4	0.14
43) Embassy Manama	967	171	0.15	132) Embassy Suva	280	67	0.19	221) Embassy Bangui	2	24	0.92
44) Embassy Santiago	274	861	0.76	133) Consulate Jeddah	298	47	0.14	222) Consulate Milan	18	6	0.25
45) Embassy Muscat	585	526	0.47	134) Embassy Canberra	184	161	0.47	223) Consulate Thessaloniki	0	24	1
46) Embassy Abu Dhabi	700	394	0.36	135) Embassy Lilongwe	108	231	0.68	224) Consulate Guayaquil	18	4	0.18
47) Embassy Bridgetown	271	778	0.74	136) Embassy Niamey	82	252	0.75	225) Consulate Strasbourg	16	6	0.27
48) Embassy Pretoria	426	608	0.59	137) Embassy Skopje	191	143	0.43	226) Consulate Vancouver	5	17	0.77
49) Embassy Tashkent	853	169	0.17	138) Embassy Vatican	271	63	0.19	227) UN Rome	0	21	1
50) Consulate Lagos	707	313	0.31	139) Embassy Bamako	216	115	0.35	228) Consulate Yekaterinburg	0	20	1
51) Embassy Harare	750	268	0.26	140) Embassy Paramaribo	60	268	0.82	229) Consulate Durban	11	7	0.39
52) Embassy Tegucigalpa	616	399	0.39	141) Embassy Montevideo	169	152	0.47	230) Consulate Dhahran	12	5	0.29
53) Embassy Asuncion	391	623	0.61	142) Embassy Phnom Penh	127	194	0.60	231) US Mission CD Geneva	16	1	0.06
54) Embassy Yerevan	775	230	0.23	143) REO Basrah	288	23	0.07	232) Consulate Curacao	3	13	0.81
55) Embassy Brasilia	374	626	0.63	144) Embassy Nassau	154	145	0.48	233) Consulate Tijuana	4	10	0.71
56) Embassy Santo Domingo	229	756	0.77	145) Embassy Georgetown	92	196	0.68	234) Consulate Calgary	0	13	1
57) Embassy Tbilisi	722	255	0.26	146) Embassy Tirana	153	125	0.45	235) Consulate Naples	8	5	0.38
58) Embassy Athens	662	310	0.32	147) Consulate Shanghai	247	26	0.10	236) Consulate Barcelona	2	10	0.83
59) Embassy La Paz	770	181	0.19	148) Embassy Dili	127	128	0.50	237) Consulate Auckland	0	11	1
60) Embassy San Salvador	354	578	0.62	149) Embassy Antananarivo	155	96	0.38	238) Consulate Ciudad Juarez	0	9	1
61) Embassy Port Au Prince	475	433	0.48	150) Embassy Chisinau	212	33	0.13	239) Consulate Florence	2	7	0.78
62) Embassy Riyadh	714	192	0.21	151) REO Hillah	203	39	0.16	240) Consulate Recife	1	8	0.89
63) Embassy Managua	581	313	0.35	152) US Office Almaty	121	117	0.49	241) Embassy Majuro	2	7	0.78
64) Embassy Sanaa	593	298	0.33	153) Embassy Yaounde	141	87	0.38	242) US Delegation FEST TWO	5	4	0.44
65) Embassy Addis Ababa	607	282	0.32	154) Embassy Port Of Spain	80	140	0.64	243) Consulate Fukuoka	1	7	0.88
66) Embassy Kyiv	761	122	0.14	155) Embassy Cotonou	30	189	0.86	244) Consulate Sydney	4	4	0.50
67) Embassy Ottawa	473	404	0.46	156) Embassy Kampala	118	96	0.45	245) Consulate Leipzig	0	7	1
68) Embassy Rabat	590	282	0.32	157) Consulate Chiang Mai	114	98	0.46	246) Consulate Hermosillo	0	6	1
69) Embassy Kingston	257	599	0.70	158) Embassy Vientiane	127	72	0.36	247) Consulate Melbourne	5	1	0.17
70) Embassy Quito	598	256	0.30	159) Embassy Libreville	141	57	0.29	248) Consulate Perth	4	2	0.33
71) Embassy Zagreb	363	480	0.57	160) Embassy Lusaka	114	82	0.42	249) Embassy Koror	0	6	1
72) Embassy Prague	566	268	0.32	161) Embassy Banjul	163	26	0.14	250) Consulate Belfast	4	1	0.20
73) Embassy Minsk	509	312	0.38	162) Embassy Gaborone	74	112	0.60	251) Consulate Hamilton	0	5	1
74) Embassy Bishkek	637	154	0.19	163) Embassy Luanda	119	63	0.35	252) Consulate Marseille	1	4	0.80
75) Embassy Lima	376	381	0.50	164) Embassy Ulaanbaatar	91	91	0.50	253) Consulate Sapporo	1	4	0.80
76) Embassy Warsaw	586	170	0.22	165) Consulate Dubai	149	27	0.15	254) Embassy Apia	0	5	1
77) Embassy Bratislava	482	266	0.36	166) Consulate Guangzhou	92	84	0.48	255) Consulate Matamoros	0	4	1
78) Embassy Kuala Lumpur	517	209	0.29	167) Embassy Bern	144	29	0.17	256) Embassy Malabo	0	4	1
79) Embassy Vilnius	426	295	0.41	168) Embassy Copenhagen	132	39	0.23	257) Consulate Nuevo Laredo	0	3	1
80) Embassy Sofia	465	240	0.34	169) Iran RPO Dubai	163	2	0.01	258) Consulate Osaka Kobe	2	1	0.33
81) Embassy Tunis	520	182	0.26	170) Consulate Peshawar	158	3	0.02	259) Consulate Amsterdam	0	2	1
82) Embassy Sarajevo	542	138	0.20	171) Embassy Belmopan	41	119	0.74	260) Consulate Merida	0	2	1
83) Embassy Algiers	560	118	0.17	172) Consulate Chengdu	144	8	0.05	261) ** Dhahran	1	0	0
84) Consulate Hong Kong	308	361	0.54	173) Embassy Reykjavik	78	73	0.48	262) American Consulate Hyderabad	0	1	1
85) Embassy Nicosia	459	210	0.31	174) Embassy Port Louis	74	74	0.50	263) Consulate Krakow	0	1	1
86) Embassy Astana	291	374	0.56	175) Embassy Monrovia	73	74	0.50	264) Department of State	0	1	1
87) Embassy Dushanbe	350	307	0.47	176) Embassy Bandar Seri Begawan	91	48	0.35	265) DIR FSINFATC	0	1	1
88) Embassy Ljubljana	420	216	0.34	177) Consulate Casablanca	70	66	0.49	266) US Office FSC Charleston	0	1	1
89) Embassy Dublin	348	286	0.45	178) Embassy Ouagadougou	72	63	0.47				

Note: locations in red indicate cable 'senders' with at least 100 restricted and 100 unrestricted cables in sample.

Figure 10: Frequency of restricted versus unrestricted cables by place of origin, for all cables post-2005.

C Topic Model

The field of quantitative text analysis has grown substantially in recent years. In this literature, applied researches extensively use Latent Dirichlet Allocation (Blei, Ng and Jordan, 2003) as a generative model to extract "themes" or "topics" from a collection of documents.[6] The model assumes documents are composed of latent topics that are chosen with probabilities following a Dirichlet distribution, and multinomial choice probabilities for word choice conditional on a topic. More precisely, the framework from Blei et al (2003) has the number of words N in a document be Poisson(ξ), the latent topic probabilities θ be Dirichlet(α), the topics z_n be Multinomial(θ), and the words w_n be Multinomial(β), conditional on z_n. Then, with M documents, they have that $p(\mathcal{C}|\alpha,\beta) = \prod_{d=1}^{M} \int p(\theta_d|\alpha) \left(\prod_{n=1}^{N_d} \sum_{z_{dn}} p(z_{dn}|\theta_d) p(w_{dn}|z_{dn},\beta) \right) d\theta_d$. Computational difficulties arise in this setting, but there are ways to deal with them; we refer the reader to Blei et al (2013) for further details.

When a researcher estimates an LDA model, the topics returned are characterized by the multinomial probabilities for all words within each topic, as well as the posterior distribution of topics conditional on a certain word. In practice, the researcher determines the number of topics a priori, although recent efforts have been made to assess how an approximate number of topics may be present in a sample of data (see, e.g., Hoffman et al., 2013).

D Exact Matching Algorithm

[6] See, for example, Blei 2012 "Probabilistic Topic models" for an overview, and Quinn et al. (2010) for recent political science applications of topic models.

1. Let N_j be the set of cables from embassy j that occur during or after the year 2005 in the sample, where $|N_j|$ is the number of cables originating from location j.

2. For each of the $|N_j|$ documents in the sample, record the subject tags present on each diplomatic cable.

3. For all restricted cables in N_j, find all unrestricted cables in N_j that exactly match on subject tags and year of creation.

4. From the subset of restricted cables in N_j with at least one unrestricted exact match,

   (a) Randomly draw a restricted cable and find the unrestricted, exact-matched cable that is written most closely in time (i.e., the cable that minimizes the absolute value between the difference in release days). Each cable may be matched with or without replacement.

   (b) Continue this process until there are no-more restricted and unrestricted cables to pair together.

5. Record the list of exactly-matched pairs of cables, if applicable.

6. Repeat this process many times, storing the exactly-matched dataset that minimizes the average difference in cable date between the restricted and unrestricted samples.

Figure 11: Pseudocode for Exact Matching Algorithm

References

Angrist, Joshua D. and Jörn-Steffan Pischke. 2009. *Mostly Harmless Econometrics: An Empiricists Companion*. Princeton University Press.

Baum, Matthew. 2004. "Going Private: Public Opinion, Presidential Rhetoric and the Domestic Politics of Audience Costs in US Foreign Policy Crises." *Journal of Conflict Resolution* 48(5):603–631.

Black, Jeremy. 2010. *A History of Diplomacy*. Univeristy of Chicago Press.

Blei, David M., Andrew Y. Ng and Michael I Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 2(4-5):993—1022.

Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45(1):5–32.

Cox, David R. 1958. *Planning of Experiments*. Wiley Publications in Statistics.

Downes, Alexander and Todd Sechser. 2012. "The Illusion of Democratic Credibility." *International Organization* 66:457–489.

Fearon, James. 1994*a*. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88:577–92.

Fearon, James. 1994*b*. "Signaling Versus the Balance of Power and Interests: An Empirical Test of a Crisis Bargaining Model." *Journal of Conflict Resolution* 38(2):236–69.

Fearon, James. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.

Freiden, Jeffry and David Lake. 2005. "International Relations as a Social Science: Rigor and Relevance." *Annals of the American Academy of Political and Social Science* 600(1):136–156.

Gibbs, David. 1995. "Secrecy and International Relations." *Journal of Peace Research* 32(2):213–228.

Goemens, Hein. 2000. *War and Punishment: The Causes of War Termination and the First World War*. Princeton, NJ: Princeton University Press.

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

Guisinger, Alexandra and Alastair Smith. 2002. "Honest Threats: The Interaction of Reputation and Political Institutions in International Crises." *Journal of Conflict Resolution* 46(2):175–200.

Hastie, Trevor, Robert Tibshirani and Jerome Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. Springer.

Heckman, James J. and Richard Robb. 1985. "Alternative Methods for Evaluating the Impact of Interventions." *Journal of Econometrics* 30(1):239–267.

Ho, Tin. 1998. "The Random Subspace Method for Constructing Decision Forests." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20(8):832–844.

Hoffman, Matthew D., David M. Blei, Chong Wang and John Paisley. 2013. "Stochastic Variational Inference." *Journal of Machine Learning Research* 14:1303–1347.

Iacus, Stefano M., Gary King and Giuseppe Porro. 2011. "Multivariate Matching Methods That Are Monotonic Imbalance Bounding." *Journal of the American Statistical Association* 106(493):345–361.

Imai, Kosuke, Gary King and Elizabeth A. Stuart. 2008. "Misunderstandings between experimentalists and observationalists about causal inference." *Journal of the Royal Statistical Society, Series A* 171(2):481–502.

Kissinger, Henry. 1994. *Diplomacy*. New York: Simon & Schuster.

Kurizaki, Shuhei. 2007. "Efficient Secrecy: Public versus Private Threats in Crisis Diplomacy." *American Political Science Review* 101(3).

Lake, David. 2010. "Two Cheers for Bargaining Theory: Assessing Rationalist Explanations of the Iraq War." *International Security* 35(3):7–52.

Powell, Robert. 1999. *In the Shadow of Power,*. Princeton, NJ: Princeton University Press.

Powell, Robert. 2002. "Bargaining Theory and International Conflict." *Annual Review of Political Science* 5:1–30.

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

Ramsay, Kristopher. 2004. "Politics at the Water's Edge: Crisis Bargaining and Electoral Competition." *Journal of Conflict Resolution* 48(August):459–86.

Ramsay, Kristopher. 2008. "Settling It on the Field: Battlefield Events and War Termination." *Journal of Conflict Resolution* 52(6):850–879.

Ramsay, Kristopher. 2011. "Cheap Talk Diplomacy, Voluntary Negotiations, and Variable Bargaining Power." *International Studies Quarterly* 55:1003–1023.

Reed, William. 2003. "Information and Economic Interdependence." *Journal of Conflict Resolution* 47(1):54–71.

Reiter, Dan. 2003. "Exploring the Bargaining Model of War." *Perspective on Politics* pp. 27–43.

Reiter, Dan. 2009. *How Wars End*. Princeton: Princeton University Press.

Rosenbaum, Paul R. 2002. *Observational Studies*. 2nd ed. Springer.

Rosenbaum, Paul R. and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1):41–55.

Rosenbaum, Paul R. and Donald B. Rubin. 1984. "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score." *Journal of the American Statistical Association* 79(387):516–524.

Rubin, Donald B. 1976. "Multivariate Matching Methods That Are Equal Percent Bias Reducing, I: Some Examples." *Biometrics* 32(1):109–120.

Sartori, Anne. 2002. "The Might of the Pen: A Reputational Theory of Communication in International Disputes." *International Organization* 56(1):121–149.

Sartori, Anne. 2005. *Deterrence by Diplomacy*. Princeton: Princeton University Press.

Schelling, Thomas. 1966. *Arms and Influence*. New Haven, CT: Yale University Press.

Schultz, Kenneth. 1999. "Do Democratic Institutions Constrain or Inform? Contrasting Two Institutional Perspectives on Democracy and War." *International Organization* 53(2):233–266.

Schultz, Kenneth. 2001. "Looking for Audience Costs." *Journal of Conflict Resolution* 45:32–60.

Shapiro, Jacob and David Siegel. 2010. "Is this Paper Dangerous? Balancing Secrecy and Openness in Counterterrorism." *Security Studies* 19(1):66–98.

Slantchev, Branislav. 2006. "Politicians, the Media, and Domestic Audience Costs." *International Studies Quarterly* 50:445–477.

Slantchev, Branislav. 2010. "Feigning Weakness." *International Organization* 64:357–388.

Slantchev, Branislav. 2012. "Audience Cost Theory and Its Critics." *Security Studies* 21:376–382.

Smith, Alastair. 1998. "International Crises and Domestic Politics." *American Political Science Review* 92(3):623–38.

Smith, Alistair and Allan Stam. 2004. "Bargaining and the Nature of War." *Journal of Conflict Resolution* 48(6).

Snyder, Jack and Erica D. Borghard. 2011. "The Cost of Empty Threats: A Penny, Not a Pound." *American Political Science Review* 105(3).

Stuart, Elizabeth A. 2010. "Matching methods for causal inference: A review and a look forward." *Statistical Science* 25(1):1–21.

Tibshirani, Robert. 1994. "Regression Shrinkage and Selection Via the Lasso." *Journal of the Royal Statistical Society, Series B* 58:267–288.

Tomz, Michael. 2007. "Domestic Audience Costs in International Relations: An Experimental Approach." *International Organization* 61:821–40.

Trachtenberg, Marc. 2012. "Audience Costs: An Historical Analysis." *Security Studies* 21(1):3–42.

Trager, Robert. 2010. "Diplomatic Calculus in Anarchy: How Communication Matters." *American Political Science Review* 104(2).

Werner, Suzanne. 1999. "The Precarious Nature of Peace: Resolving the Issues, Enforcing the Settlement, and Renegotiating the Terms." *American Journal of Political Science* 43(3):912–934.