A Boom with Review: How Retrospective Oversight Increases the Foreign Policy Ability of Democracies

Michael Colaresi Michigan State University

In the ongoing debate concerning whether democracies can carry out effective national security policy, the role of transparency costs has received little attention. I argue for a more nuanced understanding of how some democracies that possess specific investigative institutions, such as national security–relevant freedom of information laws, legislative oversight powers, and press freedoms, are able to avoid the problems of which democracy skeptics warn. Using a new dataset on national security accountability institutions in democracies within a Bradley-Terry framework, I find that national security oversight mechanisms raise the probability that a democracy wins international disputes as well as increasing the expected number of enemy casualties, as compared to democracies that lack effective oversight. Contra previous theories of foreign policy efficacy, I find that the chances for democratic foreign policy success are maximized when competitive elections are linked to institutions that increase the retrospective revelation of previously classified information.

In 1967, Israel launched a spectacularly successful strike on the Egyptian air force. After extensive training in secret, the Israeli air force destroyed approximately 300 Egyptian planes at a cost of only 19 Israeli planes in only a few hours. This initial move played a major role in the Israeli advances in the Six Days War and was only possible due to accurate Israeli intelligence about the location of Egyptian air resources as well as the fact that Egypt was ignorant of the strategic intelligence that Israel possessed.[1]

The actors and outcome of this one-sided engagement at first glance seem to fit the conventional story of democratic foreign policy success. Here was a democratic state effectively fighting a nondemocratic state as would be predicted by recent research that democracies win more of their wars (Reiter and Stam 2002) and have significant foreign policy advantages as compared to nondemocratic states (Huth and Allee 2003; Schultz and Weingast 2003). However, upon closer inspection, it is clear that most previous theories assume that domestic transparency is the engine of democratic foreign policy success.[2] Yet, in the Six Days War, as well as other democratic triumphs such as in the Cuban Missile Crisis (Garthoff 1989, 140–42, 145), the D-day invasion, the Battle of Midway (Edmonds 1986; Kahn 1991), and the "left hook" maneuver in the first Gulf War, it was secrecy and not transparency that played the pivotal role.[3] These examples suggest that the assumption of constant public access to national security information, embedded within the conventional definition of

Michael Colaresi is Associate Professor, Department of Political Science, Michigan State University, 303 South Kedzie Hall, East Lansing, MI 48824 (colaresi@msu.edu).

The author would like to thank the anonymous reviewers, the *AJPS* editorial staff, as well as Hein Goemans, Brett Ashley Leeds, Benjamin Fordham, David Clarke, Erik Gartzke, Steve Kautz, David Leblang, Burt Monroe, Patrick Regan, Cristina Bodea, Eric Chang, Nathan Monroe, Kevin Quinn, Scott Gates, Sabine Carey, and Charm Lastlie for helpful comments and criticisms that significantly improved the article. Thanks are also due to the seminar participants at SUNY-Binghamton for their time and advice, as well as the Vice President's Office for Research and the Department of Political Science at Michigan State University for financial support of the data coding efforts. The remaining errors are solely attributable to the inabilities of the author. The replication data for this article are available at http://dvn.iq.harvard.edu/dvn/dv/Colaresi.

[1] See, for example, the descriptions in Rabin (1996, 102–5) and Shlaim (2001, 241–42).

[2] Schultz (2001, 58) notes that his theory focuses on "public contestation," whereby "[t]he public nature of democratic politics ensures that the process of decision making is observable by foreign states and can thus influence international behavior." Similarly, Reiter and Stam posit that "vigorous discussion of alternatives and open dissemination of information in democratic political systems produce better decisions" (2002, 23).

[3] For example, had the Israeli public possessed accurate information a priori about the probable costs and benefit of the operation, it is doubtful the Egyptians or Syrians would not have learned of it and moved their planes, significantly altering the course of the conflict.

American Journal of Political Science, Vol. 56, No. 3, July 2012, Pp. 671–689

 DOI: 10.1111/j.1540-5907.2011.00567.x

democracy, neglects any role for secrecy in foreign policy success.

This article contributes to an enhanced understanding of the relationship between democracy and foreign policy efficacy by resolving the paradox of national security transparency inherent in conventional explanations of democratic foreign policy success. If secrets are valuable, and democracies are assumed to be transparent, in any given situation democracies either suffer transparency costs as information diffuses to an enemy, or keep necessary secrets but lose public accountability on national security issues (Desch 2002; Rosato 2003). If democratic leaders are institutionally *unable* to keep national security secrets, and democratic accountability is upheld on national security issues, enemies of democratic states are at a strategic advantage. While this accountability might provide bargaining leverage (Martin 2000; Putnam 1988), military preparations would be transparent as would be threat assessments. Democracies would be giving up the ability to surprise, feign weakness or strength, or more generally hide valuable foreign policy information at potentially substantial cost. Alternatively, if executives in democracies *are able* to keep national security secrets and avoid the associated transparency cost of revealing strategic information to the public and thus an enemy, they should also be able to dodge the public revelation of corruption or bluffing and thus domestic accountability. Consistent with this perspective, U.S. President Nixon ordered a secret nuclear alert in 1969 and then backed down without electoral punishment (Sagan and Suri 2003), and similarly, President Kennedy agreed to trade the removal of U.S. Jupiter missiles from Turkey, among other things, for the removal of USSR missiles in Cuba while avoiding near-term public scrutiny (Garthoff 1989).

To step out of the democratic secrecy dilemma, I argue that it is not solely immediate transparency that serves as the primary domestic distinction between states in the national security realm, but rather institutions that allow for retrospective oversight of national security policy and the punishment of foreign policy corruption. Previous research has treated democracies as homogenous across the domestic-foreign policy divide, assigning a democracy score to a country in a given year. However, great inter-issue variation exists even within a country, in a given year, between the handling of national security information and information on economic, social, and domestic policy. Most democracies include exceptions to freedom of information laws for national security information (Banisar 2004) and exempt the intelligence services from publicly reporting their budgets (Born and Leigh 2007), in stark contrast to the more modest executive information asymmetries on domestic issues. Democracies vary considerably across the globe in the dynamics and duration of the domestic-national security public-information gap.

These accountability institutions include limited freedom of information law exemptions for national security, foreign policy legislative oversight powers, and press freedoms. Due to the fact that secrets lose their value over time, democracies with retrospective oversight are able to mobilize a potentially skeptical public while paying minimal transparency costs without recourse to repression. Holding private strategic information while propelled by public mobilization should form a potent foreign policy combination for democracies with retrospective oversight. Conversely, democracies that fail to maintain oversight institutions are likely to be bogged down with either excessive transparency costs or public undermobilization.

I use the outcomes of international disputes to test whether national security oversight institutions raise the estimated foreign policy ability of democracies from 1970 to 2000. This is accomplished with a structured Bayesian Ordered Bradley-Terry model utilizing a new measure of retrospective security policy information institutions (r-SPII) as well as other correlates of foreign policy success. In each set of analyses, democracies with oversight are estimated, as expected, to have significantly greater foreign policy ability than their democratic counterparts that lack effective oversight in national security affairs. These results imply both that oversight institutions have significant empirical purchase when explaining foreign policy efficacy and that the exigencies of national security secrecy necessitate unique institutional mechanisms to induce accountability and effectiveness.

Democracy and Foreign Policy Ability: Deficient, Efficient, or Irrelevant?

For at least the last half-century a triangular debate has simmered concerning the ability of democratic states to formulate and execute effective foreign policy. On the one hand, scholars and politicians as varied as John F. Kennedy, Abraham Lincoln, Alexis De Toqueville, Hans Morgenthau, and Walter Lippmann have argued that democratic openness and transparency create ineffectual foreign policy.[4] These democratic deficiencies are

[4] For a summary of the disagreement on this issue, see Schultz and Weingast (2003).

supposedly the result of both national security leaks and quixotic public opinion spilling into foreign policy. Spying on other countries, hiding weaknesses, gaining the advantage of surprise, and bluffing all necessitate some level of secrecy and deception and are useful in international affairs (Gibbs 1995; Gill 1994; Mearsheimer 2004). Democratic openness, according to critics, forfeits the significant policy benefits of foreign policy secrecy.[5] Additionally, the democracy skeptics argue that since leaders in open democratic systems need to worry about being reelected, they cannot focus on creating the optimal foreign policy for the state (Ripsman 2009; Schweller 2004). When the uneducated public's preferences run counter to necessary international policy prescriptions, an office-seeking leader will be forced to sacrifice efficacy for job security.[6]

In contrast, Schultz and Weingast (2003), Bueno de Mesquita et al. (2003), Reiter and Stam (2002), Snyder (1991), and Lake (1992) have argued that democratic institutional components strengthen rather than weaken the foreign policy efficacy of democracies. Open and competitive democratic elections create incentives for leaders to formulate policies that are in the public interest. This process results in the promotion of general economic well-being rather than side-payments to political cronies and the launching of international interventions when there are public benefits rather than when intervention might only be beneficial to the leader or the leader's party.[7] Reiter and Stam succinctly summarize this argument when they argue, "democracies win wars because of the offshoots of public consent and leaders' accountability to the voters" (2002, 3).

A final group of theorists suggest that the form of government a state maintains is irrelevant to foreign policy ability because all states utilize equivalent national security decision-making processes. A series of powerful works by Kenneth Waltz purported to show that despite institutional differences, the United States and the United Kingdom had similar foreign policy capabilities (Waltz 1967) and that international anarchy forces all states into evolving functionally equivalent decision-making processes in order to survive (Waltz 1979). More recently, Desch (2002) has made the case for the irrelevance of domestic democratic institutional distinctions. He argues that on national security policy, material capabilities trump any domestic institutional differences. Further, since the public is aware that it has little knowledge of the intricacies of international polities, citizens defer to elites.

The empirical record thus far has supported the democratic efficiency hypothesis, although with caveats. Recent research has found that democracies win more of their wars (Reiter and Stam 2002), convince other states to back down in militarized disputes (Schultz 2001), and receive greater concessions during negotiations over territorial issues (Huth and Allee 2003), as compared to nondemocratic states. However, even this accumulation of evidence has been challenged on empirical grounds. Desch (2002) and Downes (2009) argue that previous findings have suffered from methodological problems.[8]

Accountability versus Secrecy

In the face of this empirical evidence, one key contradiction in the democratic efficiency hypothesis remains. Specifically, it has been assumed that in democracies, transparency allows the public to view the actions of government agents and thus hold corrupt politicians to account. The process of accountability, built upon the foundation of democratic transparency, is hypothesized to enhance democratic credibility and military threats (Fearon 1994; Huth and Allee 2003; Ikenberry 2001; Martin 2000; Schultz 2001). For example, Ikenberry argues that democracies are able to generate greater cooperation from other states due to the transparency of the governance structure, where "secrecy is seen to be the exception, not the norm" (2001, 76). However, the exigencies of national security secrecy in democracies create important information asymmetries between the executive and the public. Not only do the United Kingdom, Ireland, and India have official secrets acts that punish the revelation of national security information,[9] but also all democracies use either executive decrees, criminal statutes, or constitutional or common law provisions to contain national security information within the government.[10] Further, the scale of national security secrecy is nontrivial. In FY2009 and

[5]To be specific, I define foreign policy effectiveness as the ability to gain positive security outcomes for a state (see Lake 1992). Foreign policy ability could manifest itself through the successful use of military threats or violent force, deception and guile, or diplomacy.

[6]Conversely, see Page and Shapiro (1995) and Holsti (1996).

[7]See, for example, Bueno de Mesquita et al. (2003).

[8]However, see arguments by Reiter and Stam (2009) and Schultz and Weingast (2003).

[9]Note that the UK Official Secrets Act is explicitly protected in the more recent Freedom of Information Act in Part II, section 44, and is thus still in force.

[10]A fuller list of 53 democracies, collected by the author, and a sampling of their statutes for keeping national security secrecy are included in the online appendix.

FY2010, the United States classified over 130 million documents to shield that information from public view.[11] One observer has complained that democracies now have what amounts to a "department of government secrets".[12] As Milner notes, "In foreign affairs particularly the executive branch . . . may well possess much private information about foreign countries, international negotiations, and foreign policy choices" (1997, 21).

Transparency Costs

If national security secrecy was unnecessary, then the tension between accountability and secrecy could be resolved purely in the direction of transparency and accountability. Yet, there is no corner solution in this case. Keeping national security secrets aids foreign policy effectiveness. The United States benefited from private information in resolving the Cuban Missile Crisis, the success of the D-day invasion, and the left-hook strategy in the First Gulf War. Israel has benefited from keeping national security secrets not only in the 1967 war but also dramatically in Operation "Cast Lead" in 2008.[13] Similarly, there are extant examples of public transparency having national security costs. The Silberman-Robb commission report is purported to have a classified appendix that lists over 100 incidents of media leaks that are said to have harmed U.S. interests and cost the United States possibly hundreds of millions of dollars.[14] For example, in 1974, the CIA's deep-sea reconnaissance of a Soviet nuclear submarine that had capsized and been abandoned in the Pacific was stopped when details of the operation were published in the *Los Angeles Times* (Johnson 1997, 200–201). The revelation of the Tawny Pipit Affair by Joseph McCarthy led the CIA to cancel plans to pass false intelligence to the USSR and China (Barrett 2005, 129). Perhaps one of the more colorful examples of transparency costs for the United States was the Kersten Amendment, which publicly directed Congress to covertly undercut the legitimacy of Communist countries. This turned into a diplomatic disaster for the United States at the UN, although a USSR-sponsored resolution condemning the amendment was rejected (Barrett 2005, 105–10). More dramatic was the alleged May Incident in 1943, where U.S. Congressman Andrew May revealed publicly that U.S. submarines were able to survive Japanese depth charges because the explosives were set to detonate at too shallow of a depth. Vice Admiral Charles A. Lockwood is said to have argued that this led Japan to change tactics and cost the United States 10 submarines and 800 crewmen (Blair 2001, 424). Additionally, between 1975 and 1981, there were several high-profile cases of CIA officers' identities and addresses being publicized in Greece, Jamaica, Nicaragua, and Mozambique with the immediate reaction being expulsion (in Mozambique), attempted assassination (in Jamaica and Nicaragua), or murder (in Greece) of these agents.[15]

Secrecy Costs

While naive public transparency has national security costs as private information is revealed to a potential enemy, secrecy itself is far from costless. First, a national security secrecy apparatus and classification system costs money to maintain. The United States spent over $8.6 billion on secrecy in 2008, and that figure does not include appropriations to classify information within the Central Intelligence Agency, the National Security Agency, and the Defense Intelligence Agency, among others.[16] Additionally, official secrecy, and the lack of information sharing that results, has been shown within a bureaucracy to increase transaction costs. As information transmission is constrained, decision makers may be without key pieces of information.[17] However, the maintenance and organizational costs of secrecy do not depend on regime type. Autocratic and democratic states alike must pay for the classification bureaucracy and operate in the information environment it creates.

Democracies have an additional debt that accrues with secrecy. While autocracies can mobilize support for foreign policies with repression, democracies must rely on public consent. Public mobilization in the form of taxes, political support, and military recruitment is often a necessary condition for successful foreign policy. As Lincoln argued in a different context, "Public sentiment is everything, with public sentiment, nothing can fail, without it, nothing can succeed" (Basler 1953, 224). However, national

[11] As reported by the Information Security Oversight Office, 2010 Annual Report, available at www.archives.gov/isoo.

[12] Stephen Aftergood interview, in the movie *Secrecy*.

[13] See "Disinformation, Secrecy and Lies: How the Gaza Offensive Came About," Barak Ravid. Haaretz Israel News Source. December 27, 2008.

[14] "Secrets and Leaks: The Costs and Consequences for National Security." Remarks by Congressman Pete Hoekstra, Chairman of the House Permanent Select Committee on Intelligence, to the Heritage Foundation. July 25, 2005.

[15] See Risen and Lichtblau (2009) and Szilagyi (2010).

[16] U.S. Budget Summary (2010).

[17] This was discussed in the 9/11 Commission Report in the United States and has served as one key motivation for the reorganization of intelligence services.

security secrecy, without additional caveats, robs an executive of the democratic tools to credibly signal to the public, as well as the international community, that threatening or cooperative gestures are in the national rather than a leader's private interest. For example, the executive could launch a war to gain political support, distribute the spoils of war to cronies, or incite an international rivalry to drive government centralization, all while the evidence of those motives remained classified. The government has an information surplus that could be used to drive foreign policy towards private rather than public ends. If the public were to support these policies, they would foot the bill, while the executive would disproportionately benefit. The public, particularly opposition partisans and independents, knowing that it is at an informational disadvantage vis-à-vis the executive, and that an executive could benefit from that advantage, will be less likely to support the executive's policy due to these incentives.[18]

National security information asymmetries raise the familiar principal-agent dilemma of adverse selection. The government, acting as an agent, has more information than the public, as principal. Further, this information could be turned into profit for the agent through overcharging the public for security or a lack thereof. Knowing this, the public may remain skeptical of supporting a policy or military budget put forward by the leader. Public approval for U.S. President Clinton dropped, rather than rallying, in August 1998 when cruise missiles attacked Afghanistan and Sudan during the Lewinsky scandal. While the policy might have been in the public interest, it might also have been a useful political diversion to avert public attention away from the scandal. Colaresi (2007) finds that this is a generalizable trend, where U.S. presidents get significantly smaller and in many cases negative public opinion rallies when launching foreign policy initiatives during scandals or just before elections. Similarly, Jackson (2000) blames secrecy, in part, for France's failure to mobilize adequately during the interwar years, particularly from 1934 to 1939. As the German threat increased, the French government received useful intelligence on Nazi intentions and capabilities particularly from the Army's Deuxieme Bureau. However, the public remained skeptical of government estimates of this threat, particularly after the Stavisky affair in 1934. This lack of preparation, at least in part, accelerated their defeat in the subsequent conflict.

[18]Therefore, similar to Akerlof's (1970) market for used cars example, the executive is attempting to sell a policy to the public, but has an informational advantage over whether that policy is a lemon or not. This constellation of private information and motives can lead to market collapse, which would be analogous to lower public support.

Summarizing the Puzzle

These contrasting costs would seem to put democracies at a disadvantage vis-à-vis nondemocracies, as critics argue. Where a nondemocracy may be able to use a repressive apparatus to mobilize the public while keeping national security secrets, democracies seem to be faced with either releasing national security information to mobilize the public or keeping national security secrets, but suffering reduced public mobilization. For example, the trade-off between private information and public mobilization was a central finding of the Doolittle Report presented to U.S. President Eisenhower in 1954[19] and later in the Cold War by Admiral Powell Carter, who noted, when discussing whether the U.S. government needed to discuss U.S. vulnerabilities to Soviet attack in 1981, "I think it is dangerous, but . . . [i]f we are going to mobilize the American people to support the large expenditures that are necessary to correct the force deficiencies, we have to convince them that force deficiencies exist."[20] Going beyond previous democracy skeptics, however, I argue in the next section that if we take a more nuanced view of national security–relevant democratic institutions, a solution to this dilemma in democracies is possible.

Telling Time: Transparency Cost Deflation and Retrospective Oversight

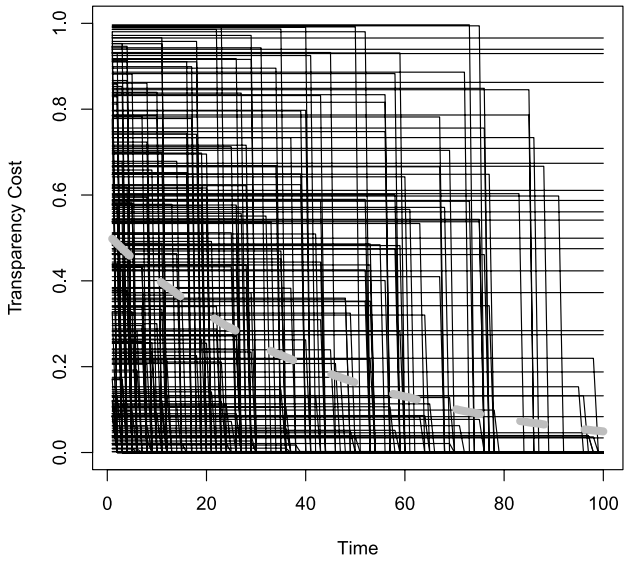
The two clues to unraveling the riddle of democratic foreign policy success lies in understanding the dynamics of transparency costs and the oversight institutions that can restore accountability at a reasonable foreign policy price. While national security secrecy is valuable to a state, the cost of revealing information declines over time—sources die or are removed from harm's way, the strategic situation changes, or an enemy learns the information by other means. The location of the D-day invasion, the breaking of the Enigma and Purple codes, and the frequency of the radio signal that U.S. Predator drones in Iraq and Afghanistan broadcast[21] on were all valuable secrets, but whose revelation costs declined over time. Once

[19]"The Report on the Covert Activities of the Central Intelligence Agency," September 30, 1954, available at http://www.foia.cia.gov/helms/pdf/doolittle_report.pdf, accessed on May 3, 2010.

[20]"The Defense of the USA, Part I." CBS Report. 1981. Aired June 14, 2011.

[21]See for example, "Insurgents Hack US Drones." Siobhan Gorman, Yochi Dreazen, and August Cole. *Wall Street Journal*, December 17.

FIGURE 1 Threads of Quickly Decaying Information (Black-solid) and the Average Transparency Cost Over Time (Gray-dashed)



an invasion occurs, the war ends, or the enemy learns the information from other sources, public transparency is less expensive.

To illustrate this point, we can simulate transparency costs over time. Each bit of information receives a randomly drawn value from a uniform distribution on the (0, 1) interval. Then at each time point, the information may quickly lose value with probability p, due to a source moving out of danger or a battle ending. The mean transparency cost for this information is a smooth decreasing function over time. Figure 1 plots 1,000 of these bits of information (black), as well as the aggregate mean transparency cost function (red). Therefore, even though specific conflicts or intelligence operations may produce an array of information, the expected value of the transparency cost declines smoothly over time.

The Institutions of National Security Oversight

Depreciating transparency costs allows the public to access national security information, ex post, and hold leaders or parties accountable for private interest policies. However, information revelation is neither deterministic, automatic, nor immediate in any democracy around the world. First, the formal classification of national security information erects a barrier to publicity by design. Prying that information free from the executive takes time and tools. I suggest there are three useful levers that can be wielded by extraexecutive actors for this purpose in democratic states: (1) specific national security–relevant legislative oversight powers, (2) effective freedom of information laws that do not fully exempt national security policy, and (3) laws that protect press freedom. Together, these oversight mechanisms increase the probability of the retrospective revelation of national security information and particularly a government's misuse of national security secrecy. In the following sections, I summarize the relevant institutions and then turn to the repercussions of retrospective accountability on national security issues for understanding foreign policy ability.[22]

While the triad of legislative oversight, freedom of information, and press freedom are usually rolled into discussions of democratic institutions and accountability (Banisar 2004; Ikenberry 2001; Reiter and Stam 2002), there is remarkable and meaningful variation in the national security relevance of these institutions across democracies. For example, the legislature in Greece has a number of useful powers for overseeing domestic political actions, including permanent committees that may investigate and provide both minority and majority reports. However, Article 133 of the Standing Orders for the Parliament states that the legislature cannot access information that the executive deems "relating to diplomatic, military or pertinent to national security."[23] Similarly, stable democracies such as Denmark, Romania, and the United Kingdom join Chile, Ghana, Peru, Kenya, and Thailand, among others, in having rules that explicitly limit legislative oversight on national security issues as compared to domestic issues. In contrast, the United States and the Netherlands have robust legislative oversight mechanisms specifically targeted to the post hoc revelation of national security information. In the United States, there are standing committees on defense and intelligence that have the power to hold both public and private hearings and compel testimony.[24] A legislature with enervated national security oversight powers has little chance of

[22]This section is not meant to imply that all national security information revelation occurs domestically. Crescenzi et al. (2011) make a similar argument with regard to conflict mediators and international organizations, such as the African Union and the Organization of American States, and their role in providing information. However, a 2010 Pew survey found that over 95% of U.S. citizens got their news from domestic sources (see "Americans Spending More Time Following the News," Pew Research Center 2010, 138). Thus, the domestic provision of information is likely to be an important part of the story. I return to this point in the conclusion.

[23]Official English translation downloaded from www.parliament.gr/english/organwsh/parliam.asp on March 22, 2007.

[24]See Johnson (1997) and Barrett (2005).

prying open the safe of national security secrets either after a war or when international tensions have migrated to different areas. For example, France, which at the time included weak national security oversight powers, failed to complete its investigation into the causes of their defeat in the Battle of France (Jackson 2000). Conversely, as the Vietnam War came to a close, the U.S. Congress unleashed the Pike and Church Committees, which released public reports on U.S. intelligence activities and spurred greater legislative oversight in the future (Barrett 2005). Therefore, legislative oversight is strongest where a permanent committee is specifically tasked with national security oversight and holds investigatory powers that could be used to create public reports on government actions.

Perhaps most explicitly, freedom of information laws around the world include exemptions for national security information. Broadly, there are three types of exemptions. The most stringent exempts all national security information from being revealed under the law and allows the executive to decide if the exemption holds. Examples of this all-encompassing exemption have been used in Canada, Greece, Switzerland, and the United States before 1972. A less restrictive rule, known as the harm test, only allows information to be withheld from public view if releasing the information will cause appreciable harm to government policy. Austria, France, India, Italy, Japan, and the United States (between 1972 and 2002) used this harm test. Finally, the balance test, used currently in Australia, Belgium, and New Zealand, is the most open exemption and only allows information to be kept secret if the harm is greater than the public benefit of releasing the information (Banisar 2004). The National Security Archive, a nonprofit group in the United States, has published more than 700,000 pages of previously secret documents from Freedom of Information requests.[25] Additionally, partial implementation of freedom of information laws, prohibitive costs, and restrictions on use can reduce their efficacy.

A free press, unencumbered by legal, political, or economic obstacles, can also release previously secret information to the public. On national security issues, government censorship, often in the form of prior restraint or penalties for divulging classified information, limits the effectiveness of the press. For example, while in the United States the Pentagon Papers case (Prados and Portier 2004) as well as Seymour Hersh's articles on previously covert U.S. actions in Haiti as well as allegations of domestic spying by the FBI (Johnson 1997) publicized previously classified information, an agreement between the media and the Israeli defense force in 1966 set up a system of prior restraint whereby a military censor can redact newspaper articles. In the United Kingdom, MI6 official Richard Tomlinson and Foreign Office clerk Sarah Tisdall served time in jail for releasing secrets to the public through the press.[26]

In overcoming national security secrecy, these three analytically distinct processes of retrospective information diffusion are highly interrelated in practice and rely on competitive elections to be political meaningful. Legislative oversight that cannot be broadcast to the public through the press is less effective. Facts released by either the press or a legislative body that cannot be independently verified through a freedom of information request may be less persuasive. With that knowledge, in a competitive electoral democracy, the public has the opportunity to turn out a specific leader or party from government, reversing previous policies that a leader might care about if foreign policy corruption is uncovered.

Private Information, Public Mobilization, and Foreign Policy Ability

The ebb of national security secrecy and the rising tide of publicly available information over time have important repercussions for domestic accountability and foreign policy ability. First, in states with effective legislative oversight, freedom of information laws, and press freedoms, it is more probable that a leader launching a war for private gain at public expense will be discovered ex post. While immediate secrecy makes executive foreign policy malfeasance possible,[27] retrospective accountability increases the costs of this behavior to the executive. If a president launched a war for private political or economic gain, effective legislative oversight would allow for soldiers, experts, and government officials to give public evidence, and press freedom would allow the press to publish both the legislative investigation and any information gleaned from whistleblowers. Finally, the public could directly query the government for policy planning documents, budgets, and strategy memos through freedom of information laws that do not exempt national security issues in full. Brody and Shapiro (1989) and Page and Shapiro (1995) provide evidence of the press, interest groups, and the legislature working together to

[25] A digital archive can be found at http://www2.gwu.edu/ nsarchiv/.

[26] See "Troubled History of Official Secrets Act." BBC. November 18, 1998. Accessed May 1, 2010, http://news.bbc.co.uk/2/hi/uk_news/216868.stm.

[27] In fact, other work formally modeling this process illustrates that either foreign policy corruption or underspending occurs in all equilibria. However, corruption is minimized as oversight institutions increase in effectiveness.

investigate previous policy decision and motives involving aid to the Contras in Nicaragua as well as during the latter years and aftermath of the Vietnam War.

All democratic foreign policy decisions face some level of dissent and skepticism (Gaubatz 1999; Holsti 1996, 183). The informative variation comes in the shape of a state's ability to convince the public, especially skeptical opposition partisans, to support a policy. Without accountability on national security issues, leaders in democracies are less able to credibly signal that their policies are in the public interest. This is likely to lead to domestic foreign policy schisms such as existed in France in the late nineteenth century where rival domestic coalitions disagreed whether Germany or Great Britain was the primary enemy (Colaresi, Rasler, and Thompson 2007, 104).

Knowing that networks of oversight institutions raise the costs of foreign policy corruption, the public can increase its certainty that proposed foreign policy programs are targeted at providing the public good of security rather than private leadership benefits. For example, in the United States, when the president announces that a new threat has arisen across the globe, the public will be unaware of the particulars of this threat assessment, and some might even doubt whether the threat has been manufactured for political gain; however, retrospective accountability institutions make it more likely that the public will support military preparations to meet and supersede the threat.[28] Johnson (2006, 30–31, 34) notes that the growth of the national security state and the Pentagon budget at the dawn of the Cold War became integrally intertwined with the Legislative Reorganization Act and the empowering of committees, particularly the Senate Armed Services Committee. Fisher (2007, 284–85) finds that U.S. foreign policy setbacks were more likely when the president sealed policy from democratic checks and failed to engage with the public, the press, and the legislative branch. Wright (1994) argues that increasing legislative accountability and executive openness in the United Kingdom would boost public engagement and mobilization. National security oversight can produce a virtuous cycle whereby the public can restrain corruption through ex post accountability, and the executive gains increased public support and mobilization even from a skeptical public.[29]

This argument qualifies the previous assumption that all democratic leaders are accountable to the public on national security issues (Reiter and Stam 2002; Schultz 2001) while opening theoretical space for private information to play a crucial role in foreign policy ability. Thus, it is only democracies with significant oversight institutions that can avoid the Scylla of excessive transparency costs and the Charybdis of public undermobilization to gain a foreign policy advantage over other types of states. The dual ability to count on private information (e.g., breaking codes, strategic surprise, threat assessment) and public mobilization (e.g., appropriations, political support, military recruitment) in foreign policy should be a potent mix. As democracies with oversight are prepared, supported, informed, and one step ahead, democracies absent national security accountability are likely to be underprepared, domestically divided, ignorant, and one step behind.

Using Disputes to Illuminate Latent Foreign Policy Ability

The above discussion leads to the hypothesis that the foreign policy ability of democracies with national security oversight institutions should be, ceteris paribus, higher than democracies that lack these types of institutions. While this seems to be a simple comparison, the task of measurement is significantly complicated by the fact that foreign policy ability is a latent trait. States have strong incentives to hide and misreport their foreign policy ability for strategic reasons, so as to shield vulnerabilities, unleash surprises, or bluff. Further, while historical measures of military and economic capabilities are available, they quantify only one piece of foreign policy effectiveness. In 1990, the USSR remained the second strongest state in the world according the the Correlates of War material capabilities data, but was internally divided and in the throes of dissolution.

However, there are times when we receive signals as to the relative foreign policy ability of states. As eloquently put by Waltz when discussing the distribution of international power, "Like a flash of lightning, crises reveal the landscape's real features" (1979, 152). As armies mobilize and threats are made, bluffs are called, and military preparations peak. In 1967, while Egypt had four times

[28] This occurs because there is now a disincentive for a leader to propose foreign policies that accrue private benefits at public expense.

[29] For elections to be a useful deterrent to national security corruption, both information on that corruption needs to be publicized with high enough probability, and conditional on receiving that information, the public must be willing to punish a leader, her party, policies, or legacy to a sufficient degree. Research on public responses to foreign policy information suggest that, indeed, the public responds reasonably to new information (Baum 2004; Baum and Groeling 2004; Brody and Page 1975; Page and Shapiro 1995). The public need not be perfect, only responsive enough to deter private interest policy action on the part of the executive.

the aggregate military capabilities of Israel,[30] Israel had a mobilized population, valuable strategic intelligence, and a successful plan. The resulting conflict illuminated the effectiveness gap at the time, in Israel's favor (Shlaim 2001, 241–42). Dispute outcomes provide a snapshot of the relative ability of states.

A Bayesian Ordered Bradley-Terry Model of Foreign Policy Ability

This hypothesis raises the question of how to empirically estimate latent foreign policy abilities from the outcome of international disputes. To answer this, let each state i at some time t have foreign policy ability λ_{it}. In a dyadic foreign policy crisis, where y_{ijt} indexes the outcomes of a particular dispute for state i against state j at time t, the propensity for a winning outcome for state i will be a function of the difference between state i's ability and the opponent state j's abilities at time t, where $j \neq i$, and some stochastic component ϵ_{ijt}.[31]

The mapping from the latent abilities to a measured outcome, y_{ijt}, would be

$$y_{ijt} = \begin{cases} Lose & \text{if} \quad \lambda_i - \lambda_j \leq -\epsilon_{ijt} \\ Win & \text{if} \quad \lambda_i - \lambda_j > -\epsilon_{ijt} \end{cases}$$

This implies that

$$\begin{aligned} P(y_{ijt} = Win) &= Pr(\lambda_i - \lambda_j > -\epsilon_{ijt}) \\ &= F(\lambda_i - \lambda_j) \end{aligned} \tag{1}$$

and

$$\begin{aligned} P(y_{ijt} = Lose) &= Pr(\lambda_i - \lambda_j \leq -\epsilon_{ijt}) \\ &= 1 - F(\lambda_i - \lambda_j) \end{aligned}$$

where $F(\cdot)$ is a cdf that maps the difference in abilities to the associated probabilities.

The discussion in the previous section suggests that the abilities of each democratic state will vary, among other things, with the oversight institutions present in that state. If we drop the time subscript for clarity and let the observed institutional traits of state i be R_i and those other observed traits that affect ability, such as material military capabilities, be C_i, then we can express the ability of democracy i as

$$\lambda_i = \mu_1 + \beta_R R_i + \beta_C C_i + \epsilon_i. \tag{2}$$

[30] As measured again by the Correlates of War material capabilities data.

[31] Here we are assuming that the greater the difference in the latent foreign policy ability of two states, the greater the probability that the stronger state gains a positive outcome of the dispute, subject to ceiling and floor effects.

The β terms represent potentially nonscalar linear weights that link the observed traits to the outcome of interest,[32] and ϵ_i represents the unmeasured portion of state foreign policy ability. μ_1 represents the intercept.[33]

Substituting (2) into (1) yields

$$\begin{aligned} P(y_{ij} = Win) = F(&\mu_1 - \mu_2 + \beta_R(R_i - R_j) \\ &+ \beta_C(C_i - C_j) + \epsilon_i - \epsilon_j). \end{aligned}$$

In this formulation, β weights that are positive increase a state's ability and thus the probability of a win, all else equal.

This formulation is similar to the well-known structured Bradley-Terry model for paired comparisons that has been used to analyze the abilities of sports teams (Agresti 2002), journal citation patterns (Stigler 1994), and animal behavior (Stuart-Fox et al. 2005).[34] For present purposes there are two additional complications. The first can be seen in the additive random effect terms (ϵ_i and ϵ_j) for each country in the dispute. Ignoring these disturbances, as has been done in the past,[35] assumes the observed covariates measure state ability without error. This would seem unlikely, a priori. The inclusion of the random effect allows us to look at the posterior distribution of the variance of the ability parameters for democracies and nondemocracies. The second complication is that all international conflicts do not end in a victory for one side or the other. In fact, international draws are the modal category for militarized interstate disputes. Therefore, our unobserved latent abilities will result in wins, draws, or losses. As previously suggested by Tutz (1986), we can handle ties in Bradley-Terry models by adding a threshold,[36] $\tau > 0$, to model the difficulty of generating a win for either side. Thus, we can reexpress y_{ijt} as the observed index of state i losing, drawing, or winning a

[32] For example, in one analysis below there are six control variables and thus C_i would be a $6 \times N$ matrix (N being the number of observations) and β_C would be a 1×6 matrix of weights. Each element of β_C is a coefficient in a generalized linear mixed-model framework.

[33] I allow the intercepts, μ_1 and μ_2, to vary across position in the dyad; thus, there may be distinct intercepts in the abilities equation for the member of the dyad which has the higher oversight score, in the first position, versus the lower score, in the second position. Only the difference $\mu_1 - \mu_2$ is identified, and thus it is this difference that is presented in the supporting information. Zeroing out this term does not influence the inferences drawn from the data.

[34] See Agresti (2002) and Turner and Firth (2010) for additional details.

[35] Turner and Firth (2010) include a useful discussion of Springall (1973), where the random effects are assumed to be zero.

[36] More thresholds can be added if additional categories are relevant. This point is returned to in the discussion of net fatalities.

dispute as

$$y_{ijt} = \begin{cases} Lose & \text{if} \quad \lambda_i - \lambda_j + \epsilon_{ijt} < -\tau \\ Draw & \text{if} \quad \tau > \lambda_i - \lambda_j + \epsilon_{ijt} > -\tau \\ Win & \text{if} \quad \lambda_i - \lambda_j + \epsilon_{ijt} > \tau \end{cases}$$

leading to

$$\begin{aligned} P(y_{ijt} = Lose) &= Pr(\lambda_i - \lambda_j + \epsilon_{ijt} \leq -\tau) \\ &= F(-\tau - (\lambda_i - \lambda_j)) \end{aligned}$$

$$\begin{aligned} P(y_{ijt} = Draw) &= Pr(\tau > \lambda_i - \lambda_j + \epsilon_{ijt} > -\tau) \\ &= F(\tau - (\lambda_i - \lambda_j)) \\ &\quad - F(-\tau - (\lambda_i - \lambda_j)) \end{aligned}$$

and

$$\begin{aligned} P(y_{ijt} = Win) &= Pr(\lambda_i - \lambda_j + \epsilon_{ijt} > \tau) \\ &= 1 - F(\tau - (\lambda_i - \lambda_j)) \end{aligned}$$

Substituting (2) into the above equations produces

$$\begin{aligned} P(y_{ijt} = Loss) &= F(-\tau - \mu_1 - \mu_2 - \beta_R(R_i - R_j) \\ &\quad - \beta_C(C_i - C_j) - \epsilon_i + \epsilon_j) \\ P(y_{ijt} = Draw) &= F(\tau - \mu_1 - \mu_2 - \beta_R(R_i - R_j) \\ &\quad - \beta_C(C_i - C_j) - \epsilon_i + \epsilon_j) \\ &\quad - F(-\tau - \mu_1 - \mu_2 - \beta_R(R_i - R_j) \\ &\quad - \beta_C(C_i - C_j) - \epsilon_i + \epsilon_j) \\ P(y_{ijt} = Win) &= 1 - F(\tau - \mu_1 - \mu_2 - \beta_R(R_i - R_j) \\ &\quad - \beta_C(C_i - C_j) - \epsilon_i + \epsilon_j) \end{aligned}$$

This model can be estimated within a Bayesian framework by selecting an inverse link function for $F(\cdot)$ and placing priors on the unobserved parameters. Specifically, I replace $F(\cdot)$ with the inverse logit[37] and use the following prior specification, where β is a vector of coefficients and intercepts, and ϵ is a stacking of the random effects: $\beta \sim \mathcal{N}(0, 1 \times 10^{-8})$, $\epsilon \sim \mathcal{N}(0, \nu^2)$, $\tau \sim |\mathcal{N}(0, .01)|$, $\nu \sim \mathcal{U}(0, 10)$.

The priors on β are diffuse and contain little information,[38] and the prior on τ is a Half Normal distribution. I will refer to this model as a Bayesian Ordered Bradley-Terry Model (BOB-T) for brevity. This model allows us to compute the probability of a democratic state winning a dispute against another state conditioning on its national security–relevant oversight institutions and other observable traits.

[37]Other binomial-family link functions such as the probit or cauchy are also possible. A probit specification leads to an ordered version of the Thurstone-Mosteller model.

[38]The priors for the normal distribution are given with precisions, rather than variances. Non-Bayesian models are presented in the supporting information.

Further, the framework is flexible enough to accommodate more nuanced indicators of relative abilities between states, such as the difference between state i and state j's fatalities in a dispute. For the measurement of net fatalities, we can define $F(\cdot)$ now as the normal cdf and include the observed thresholds for the fatality intervals.[39] This type of analysis is analogous to inferring abilities from point differentials in sports and is similar to interval regression with two additive random effects.

Research Design

Measuring Foreign Policy Disputes and Success

To measure foreign policy success after 1970, I rely on the Dyadic Militarized Interstate Dispute data (DYADMID) version 2.0 (Maoz 2005).[40] These data include measures of both the participants and outcomes of events where force was threatened, shown, or utilized. The dyadic conflicts are coded only for those states that came into conflict and the coding procedure avoids pairing all states on each side of a militarized dispute against each other.[41] I code policy success when a state is either victorious in a dispute or forces the other side to yield. Similar dyadic militarized dispute outcome data have been used in research by Senese and Quackenbush (2003) and Hensel (1994).[42] This gives us a sample of 449 dyadic disputes, of which 44 end in victory for one side or the other.

As an alternative to using dispute outcomes, I also measure foreign policy success by using net enemy fatalities.[43] Drawing inferences from enemy fatalities less own fatalities within militarized interstate disputes is complicated by the available data. The MID 3.02 data include information on six different ranges of casualties. Further, we do not observe the upper bound of the highest fatality

[39]This is discussed in the supporting information in more detail.

[40]The DYADMID data were cross-checked with version 3.02 of the original MID data—Ghosn, Palmer, and Bremer (2004)—with respect to outcomes. If there was a discrepancy, the DYADMID coding was used.

[41]For example, this approach would make Poland a victor over Germany in World War II.

[42]The DYADMID dataset is preferred to looking at wars only, due to the limited number of Correlates of War (COW) interstate war events involving at least one democracy after 1970. For example, data from Bennett and Stam (1998) only include eight cases of democratic war participants within the study window.

[43]I use the MID Ghosn, Palmer, and Bremer (2004) definition of fatalities as battle-related deaths rather than the more amorphous casualties that often include injuries.

interval. These ranges are supplemented with exact fatalities in a subset of cases. Therefore, instead of having a scalar dependent variable for each case, we now have two values that define the net fatality interval. I use both Maoz (2005) and Ghosn, Palmer, and Bremer (2004), along with historical research, to code the net fatalities into nine categories. These observed boundaries of net fatality categories serve as the τ_k's in the interval-level BOB-T model discussed in the supporting information. I keep all dyadic disputes from Maoz (2005) that included one democracy, regardless of their oversight institutions, as differentiated by a Polity score (Jaggers and Gurr 1995) equal to or greater than 6.[44]

Measuring Security Policy Information Institutions

In order to test the foreign policy effectiveness of national security oversight institutions, measures need to be developed that are sensitive to the institutional differences between domestic and foreign political processes in democracies. As international relations research has begun to focus on domestic-level predictors, scholars have borrowed measures of political institutions from comparative scholars. For example, the Polity measure of democracy has become a focal point of international political research on conflict (Jaggers and Gurr 1995). This measure and others have proved entirely appropriate and eminently useful for numerous research topics. However, Polity or other similar measures obscure the specific distinctions between domestic and international political processes in general institutional tendencies. As discussed above, some states with very high democracy scores include both formal and informal mechanisms to shield international policy decisions from public view.

A newly created database of retrospective security policy information institutions (r-SPII) uniquely measures the national security issue-specific distinctions between democracies from 1972 to 2000.[45] The database was constructed through historical research and secondary analysis of preexisting datasets on the three retrospective information revelation tools detailed above. These are (1) freedom of information laws and their national security exemptions, (2) legislative oversight powers specifically on security issues, and (3) press freedom/free speech, as discussed in the previous sections. A full discussion of the coding of these variables can be found in the supporting materials and in the online codebook (Colaresi 2011). Both the freedom of information and legislative oversight components were coded through historical research into the relevant laws and parliamentary procedures.[46] The press freedom component relies on the Freedom House civil liberties measure.[47] This measures the right of citizens to voice opposition, the key component in any retrospective diffusion of information on government foreign policy malfeasance. The underlying data include measures ranging from 0 to 1 on all three subcomponents.

All three components are scaled from 0 to 1 and averaged to create an index of security policy information institutions that ranges from 0 to 1. Information has been collected for 70 democracies.[48] Only democracies[49] are coded. This allows us to explore the variation within a category that has been previously treated as homogenous. Democracies may or may not have national security–relevant institutional mechanisms, as I highlight here.[50] The annual correlation between the SPII index and the Polity democracy score (ranging between 5 and 10 for the sample of democracies) varies around a mean of 0.66. For robustness, I also estimate a model that controls for the Polity democracy scores in the ability equations. This would guard against the inferential possibility that democracies with oversight are merely more democratic, as measured by Polity, and more effective on those grounds alone, compared to democracies absent

[44] In the few disputes that had democracies facing off against each other, state i was the state with the greater oversight score; see below. This is done merely for ordering and to avoid the double counting of observations, since this is merely one dispute.

[45] On an annual basis.

[46] For freedom of information laws, Banisar (2004) and the citation therein were useful starting points.

[47] In many ways, the Freedom House political press freedom measure would have been preferable, but the detail, scale, and range of coding have changed so considerably over time that temporal consistency cannot be ensured. The civil liberties measure was rescaled to run from 0 to 1, with 1 representing the greatest liberties.

[48] More details on the construction of the index as well as the list of democracies are available in the supporting information, as well as from the author, and will be published online. The major secondary sources for the index include Banisar (2004), Kurian (1998), the annual Freedom of the World surveys, and the collection of Library of Congress country studies.

[49] With a Polity democracy score greater than 5.

[50] Note that nominally information-producing institutions in nondemocracies are very rare empirically and of a significantly different quality, as compared to the oversight institutions discussed herein. For example, Zimbabwe passed the Access to Information and Privacy Protection Act in 2002, which might sound like a freedom of information law, but it is designed to punish both the diffusion of government information and dissent rather than protect the public's right to know (Banisar 2004). In fact, the UK freedom of information group Article 19 described the Zimbabwe law as "the leading weapon of the government and the ruling ZANU PF party in their ongoing campaign to stifle independent media reporting in Zimbabwe" ("The Access to Information and Protection of Privacy Act: Two Years On," accessed July 30, 2010).

oversight. The SPII index for state i is R_i in the BOB-T models.[51]

Control Variables

I also measure[52] several other variables whose absence might confound or inflate the relationship between the national security oversight institutions index and foreign policy ability.[53] I weigh a state's prestige and status by including whether a state is a major power or not.[54] To control for the material strength of a state, I measure the natural log of each state's capabilities in the dispute (Desch 2002).[55] Additionally, dispute outcomes may be linked to opponents that are not already involved in the dispute. For example, a state facing a strong external threat, such as Israel, from one or many rivals may find foreign policy success more difficult to achieve. Therefore, I control for the strength of a country's external rivals.[56]

Previous research by Colaresi, Rasler, and Thompson (2007) and Hensel et al. (2008), among others, has found evidence that disputes can be related to each other over time. There are three relevant dimensions to consider: (1) simultaneous, cross-dispute effects; (2) deterministic trends in outcome probabilities the longer the elapsed time since the previous dispute; and (3) within-dyad conflict accumulation over time. Therefore, I record both the number of disputes in which a state becomes initially involved in each year and the time since the last dispute between the pair of disputants to measure components of the first two effects. Even a major power like the United States cannot simultaneously start and easily win multiple disputes. Further, recent crises between the same disputants might signal a conflict that is difficult to resolve and unlikely to lead to victory on either side. I additionally code the number of previous disputes between the two states as well as the number of previous losses by one state in the dyad. These help to code for the third dimension of dependence.

Within a dispute, I measure the potential for first-mover advantage by including whether a state both moves first[57] and is an originator to the dispute rather than a joiner, per the MID data. Similarly, the potential for asymmetric effort in a dispute should be evident if one state is using force in a dispute, but the other state does not. Therefore, I code a variable marking whether a state was using force.[58] As a robustness check on the results, I also subset the analyses by disputes where force was used on one side as well as by disputes where the United States was not involved. The role of the United States might sway outcomes towards more democratic victories, or it might be argued that states with high oversight only win small-scale disputes that fail to escalate towards uses of force. Summary statistics for the relevant variables are included in the supporting materials.

Results

A full table of results for four specifications of the BOB-T model of winning an interstate dispute is presented in the supporting materials.[59] In each specification, the representing national security oversight institutions is positive, suggesting that democracies with greater oversight have an increased probability of winning their international disputes, as compared to democracies with few institutional mechanisms for national security oversight. Figure 2 plots the posterior distribution of the oversight institutions coefficient for the four models. The first

[51] Additionally, the state with the higher oversight score is associated with intercept μ_1 while the other state gets μ_2.

[52] Note that these variables enter the ability equations for a dispute participant, i, or j, measuring λ_i or λ_j, respectively.

[53] It is important to note that several of these variables, including material capabilities, may be posttreatment in the sense that oversight institutions aid in mobilizing foreign policy resources. If, as expected, higher values of the r-SPII index lead to greater mobilization and material capabilities, and both then lead to higher foreign policy ability, the partial effect of the oversight institutions index when controlling for material capabilities will be attenuated. We should be able to see this when comparing the substantive size of the r-SPII coefficients across specifications. I include these variables in select specifications because they may give a conservative picture of the empirical benefits of oversight institutions.

[54] Major power is coded dichotomously. The information comes from the COW project (Correlates of War Project, n.d.).

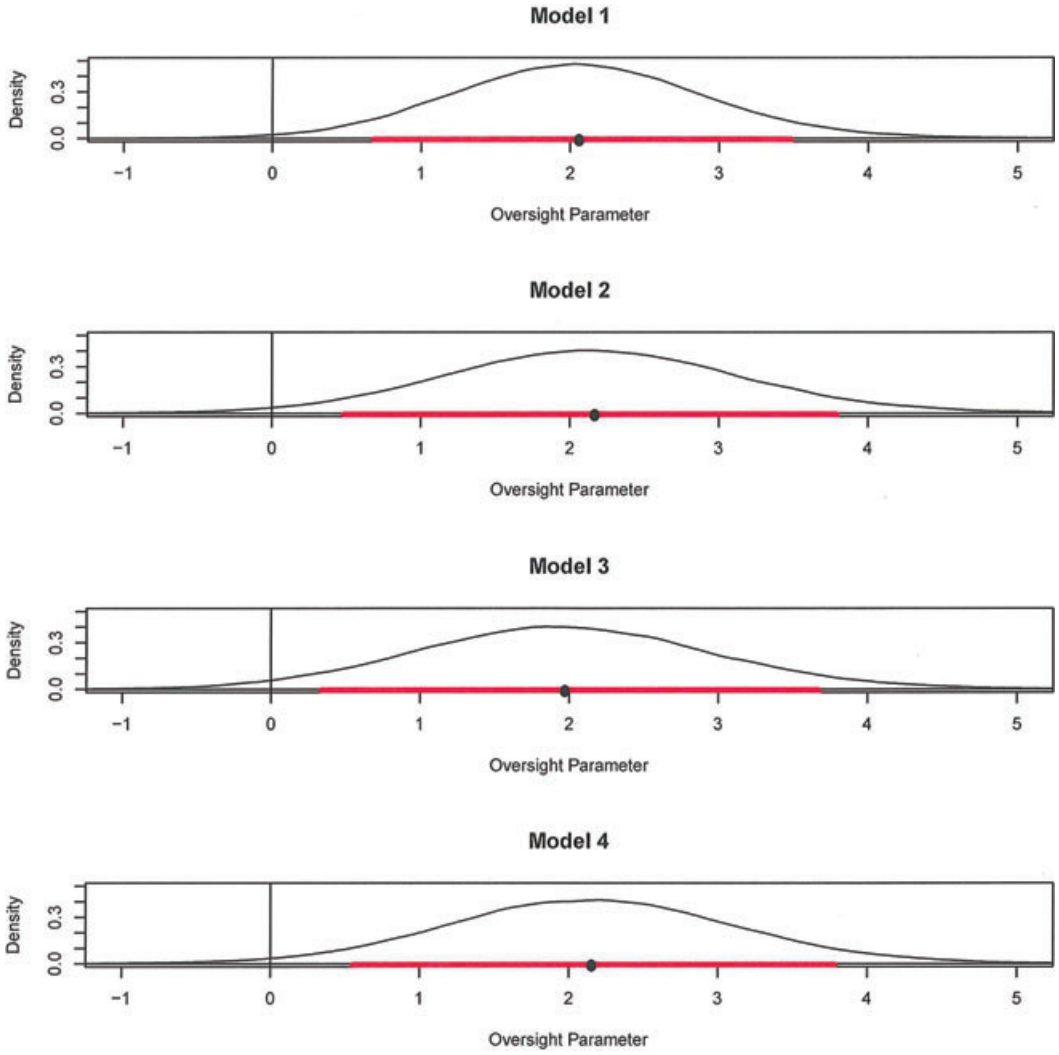
[55] Power capabilities come from the well-known COW capabilities index, version 3.02 (Singer 1987).

[56] Coded from Thompson's list of strategic rivalries (Thompson 2001). Specifically, external threat/rival strength is measured by taking the natural log of the sum of all rival capabilities. A small start-up value is added to all threat observations (.01). This represents the theoretical expectation that all states feel some small quantity of external pressure, even if they are not involved in a rivalry.

[57] This implies that the state is on "side a" in the MID data.

[58] This was measured as a 3 or above on the Hostility Level variable in the DYADMIDs data (Maoz 2005).

[59] These results were computed using JAGS 2.0 (Plummer 2003) running the Gibbs sampler for 505,000 iterations, discarding the first 5,000 runs and then keeping every 10th run for posterior inference on the 50,000 remaining samples. This chain passes the run-length criterion of Heidelberger and Welch (1983). A smaller pilot run that was discarded was analyzed using the Raftery and Lewis (1992) diagnostic and suggested that this length of run would be more than adequate for posterior inference.



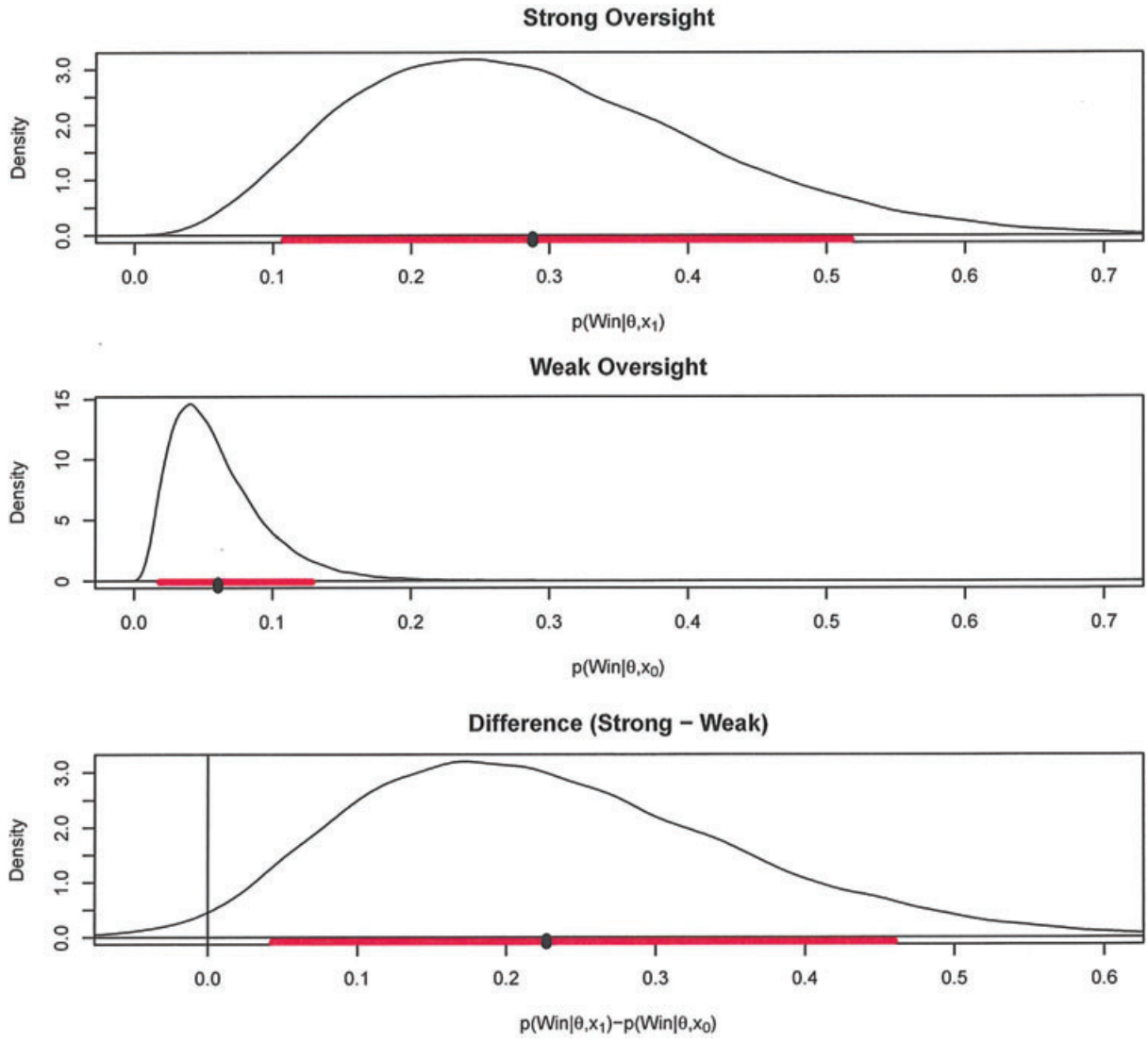
Note: The thick black lines on the x-axis indicate 90% highest posterior density (HPD) intervals, and the black dots represent the means. A tabular summary of the posterior distributions of the parameters can be found in the supporting materials.

model includes only the national security oversight institutions index. The second adds the structural controls such as material capabilities, major power status, external threat burden, and other ongoing militarized disputes as well as controls for the previous number of disputes and losses within the dyad. The third model adds the intracrisis variables, who moved first and whether asymmetric force was present, as described above. Model 4 presents a more parsimonious parameterization that probes whether controlling for major power status was masking the effect of capabilities on foreign policy ability. Consistently, the changes in the specification do not substantively alter the mean or spread of the posterior estimate for oversight institutions. In Models 1, 2, and 4, there is less than a 0.01 probability, respectively, that the effect of oversight institutions on conflict outcomes is zero or less. In Model 3, this probability rises only slightly to 0.02. Thus, given the data and priors, there are posterior probabilities greater than 0.98 that the effect of oversight institutions on foreign policy ability is greater than zero.

Substantively, we are interested in how the underlying probability of victory changes as we increase national security oversight institutions. This is explored in Figure 3, which plots the posterior probability of a democratic state with strong oversight (top) and a democratic state lacking oversight (middle) winning a dispute against a nondemocracy while enjoying a capability advantage.[60] The bottom panel of the figure represents the estimated difference between these two scenarios. A democracy with

[60]Specifically, I calculate the posterior probability of a democracy winning an interstate dispute with the other variables set at their means in a dispute where the democracy uses force but does not move first. The posterior distributions of the parameters come from Model 4.

FIGURE 3 The Posterior Probability of Winning a Dispute as Oversight Increases



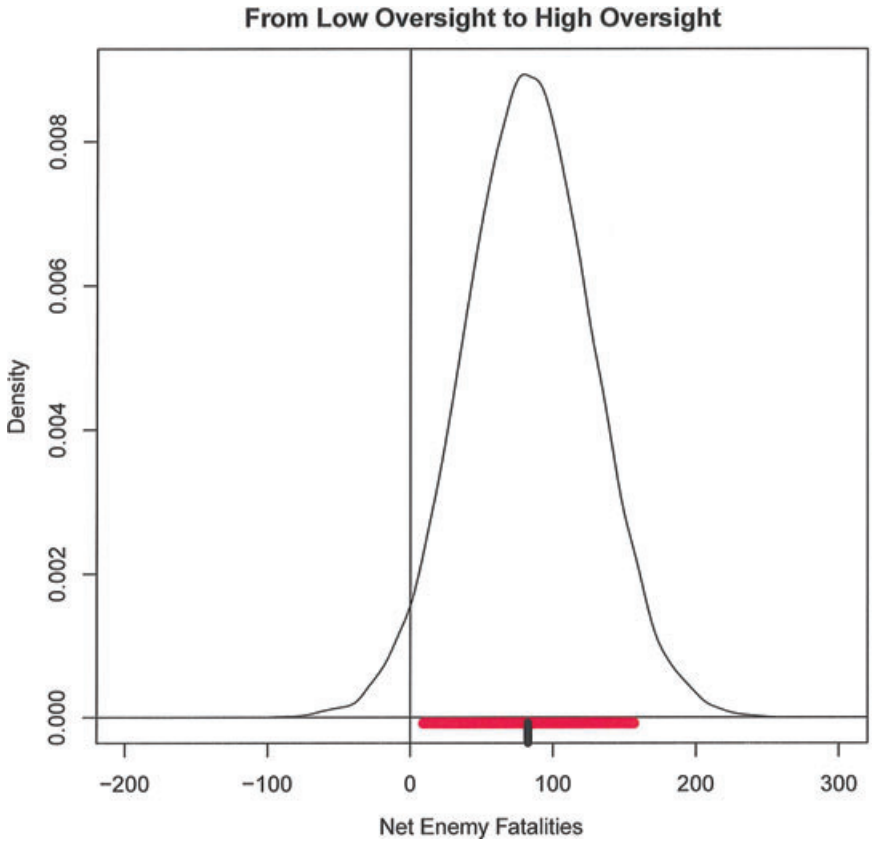
Note: The estimated posterior probability of winning a dispute for a democracy with strong oversight (top panel) and a democracy that lacks oversight (middle panel). The bottom panel illustrates the difference between the posterior distributions of these scenarios. The thick black lines on the x-axis represent 90% HPD intervals, and the black dots represent the means.

strong oversight institutions is estimated to have a .28 probability of winning such an encounter, with an 90% central credible interval (CCI) of (0.11, 0.52), as compared to a 0.06 estimated posterior probability of victory for a democracy with no oversight advantage.[61] Moving from an institutional setting that lacks oversight to one with strong oversight is estimated to increase the probability of a foreign policy victory by .22 with a 90% CCI of (0.04, 0.46) in absolute terms and by 350% in relative risk terms. These results are consistent with the hypothesis that oversight institutions, through retrospective rather than simultaneous public accountability in national security policy, uniquely allow democracies to accrue foreign policy benefits from both private information and public mobilization. Democracies that fail to maintain these oversight institutions are less likely to win their foreign policy disputes.

To probe the plausibility of this finding further, I conducted a battery of additional tests. First, I used net fatalities in disputes as an alternative measure of dispute

[61] Here the 90% central credible interval is (0.02, 0.13).

FIGURE 4 The Posterior Expected Value of Additional Net Enemy Fatalities (Enemy Fatalities Minus Own Fatalities) for a Democracy with Oversight versus a Democracy without Oversight Over National Security Policy from the Interval Regression Version of the Bradley-Terry Model



Note: The thick black line on the x-axis represents the 90% HPD interval, and the black dot represents the mean.

outcomes.[62] In this interval regression version of the BOB-T model, the results remained consistent with the foreign policy efficacy of oversight institutions with the posterior probability that the national security oversight coefficient is greater than zero being more than .96%. Figure 4 plots the posterior distribution of the estimated effect of increasing the national security oversight index from its minimum to its maximum. This suggests that increasing the strength of oversight institutions from 0 to .8 (approximately equal to the Netherlands in 2000) leads to an expected increase in net enemy fatalities of 93 with a 90% CCI of (10, 157).[63]

[62]Replacing wins, draws, or losses with an interval that matched the scale of net own versus enemy fatalities. This is described in the data section above and further in the supporting information.

[63]The mean of the posterior distribution for the national security oversight coefficient was 103, with a standard deviation of 56 and a 90% CCI of (12, 196).

Three other robustness checks were conducted using the original BOB-T model. First, I dropped all cases where force was not used by at least one side. There are two related reasons for this. First, it could be the case that democracies with oversight are only winning inconsequential crises that are not worth fighting about. Second, despite the reduction in sample size, the effectiveness of oversight institutions should remain stable in militarized crises since it is in these cases that the benefits of secrecy and mobilization are manifest. When a BOB-T model is estimated using only disputes where force was used, the mean of the posterior distribution for the national security oversight coefficient was 2.73, with a standard deviation of 1.45 and a 90% CCI of (.45, 5.18). The posterior probability of this parameter being less than or equal to zero is less than 0.023. Therefore, the posterior inferences we draw from all disputes and only militarized disputes is substantively similar. Additionally, I removed all cases

where the United States was an actor in the dispute. Using all non-U.S. disputes, the parameter for oversight institutions was estimated to be 3.13 with a standard deviation of 2.12. This yields a posterior probability of oversight failing to increase foreign policy ability of less than .06%. In the force and non-U.S. subset analyses, the posterior probability that the national security oversight coefficient is greater than zero was greater than .97 and .94, respectively. Again, this provides evidence that is consistent with the theory.

In the final robustness check, I controlled for the Polity score for the democratic state in the dyad. This variable ranges in the sample from 6 to 10. In this analysis, not only did the effect of oversight institutions remain stable, where the posterior parameter mean was 2.13 with a 90% HPD of (0.21, 4.25) and a prior probability that the oversight parameter was greater than zero equal to 96%, but also the partial effect of the democracy score on the probability of a state winning a dispute was inconclusive. The parameter for democracy had a posterior mean and standard deviation of −0.01 with an 80% HPD of (−0.12, 0.11). Thus, the posterior probability mass for the Polity democracy score was centered around zero, while the mass for the national security oversight index was positive. This is consistent with the inference that oversight institutions are important components of democratic success, mirroring the results illustrated in Figure 3.[64]

The results for the other variables in the BOB-T models are mixed. Plots of the posterior distribution of the other variables that are of interest along with their 80% HPD are included in the supporting materials. The most consistent finding is that the use of force and moving first had meaningful, although opposite effects, on the probability of victory. Not surprisingly, using force increased the probability of victory. On the other hand, moving first was associated with a lower win probability. Additionally, becoming involved in multiple disputes simultaneously is estimated to decrease the probability of a foreign policy victory. The parameters measuring the number of previous disputes and losses had inconsistent effects across Models 2 and 3. However, the posterior probability that both of these parameters, along with the time since the previous dispute, were simultaneously equal to zero was less than .85.

Taken together, the results robustly support the importance of national security oversight institutions in understanding democratic success. Using dispute outcomes, fatalities, and distinct subsets of observations over several different specifications, oversight institutions are estimated to increase the expected foreign policy ability of democracies. Further, these oversight institutions are not merely a proxy for more democratization, since even when controlling for the democracy level of the states, oversight continues to predict greater foreign policy ability and success.

[64] A discussion of the random effects can be found in the supporting materials. I ran similar BOB-T models with and without the random effect terms and found that the uncertainty estimates around the coefficients shrank by approximately 20%.

Conclusion

The allied World War II slogan "loose lips might sink ships" has been echoed by some international relations scholars in their pessimistic outlook on democratic foreign policy efficacy (Kissinger 1995; Lippmann 1922; Morgenthau 1967). It is hard to argue against the notion that all states need some level of immediate secrecy to carry out effective national security policy. Complete operational transparency would compromise troop movement, weapon effectiveness, planning, and any element of surprise. On the other hand, as scholars of domestic political accountability stress, secrecy comes with potential costs, including the increased opportunities for leaders to act in their private interest and at a decreased propensity to mobilize skeptical citizens. Recently, the ambivalent role of secrecy in foreign policy has reached the headlines with the Wikileaks scandals. This research helps to broaden the simplistic argument that all national security leaks are equally costly. A focus on transparency cost deflation suggests that old information that comes to light will be, on average, less detrimental to U.S. interests than newer information. Therefore, the Afghanistan War Diary leaks, with the unredacted names of informants, may be more detrimental to security than the older diplomatic cable leaks. Additionally, a focus on accountability suggests that intermittent leaks of information to the public, while potentially incurring costs, also have benefits. Specifically, increased public confidence that secret information will eventually be released reassures potentially skeptical citizens that the executive is less likely to abuse secrecy for private ends.

Instead of arguing that secrecy should be avoided unconditionally or redacted in perpetuity, this article makes the case that transparency costs can be reduced through the delayed publication of national security secrets, and further that the mobilization-inhibiting costs of secrecy can be reduced through strong national security oversight institutions in democratic contexts. Further, I provide evidence that these national security oversight institutions increase the probability of foreign policy victory for

democracies. Democracies that lack mechanisms to uncover classified information, ex post, are less able on the international security stage.

This project has implications for research on a broad range of issues in international relations, constitutional studies, and comparative politics. For example, scholars in the democratic peace literature have attempted to test which types of institutions promote peace between states. One theory holds that accountability mechanisms are the key restraining influence on international war in democracies (Huth and Allee 2003; Russett and Oneal 2000). A focus on the national security oversight institutions suggests that a complementary constellation of mechanisms, rather than one magic electoral bullet, may be necessary to avoid executive corruption or undermobilization in democracies. Without retrospective information assurances, a leader is unlikely to be held accountable for private interest, rather than public interest, foreign policy. Knowing this, the public may doubt the motives of an executive's foreign policy choices. The expected outcome of democratic policy, under a regime of effective oversight, may not be unconditionally peaceful, but instead may reduce the private justifications for war, as compared to nondemocratic circumstances.

References

Agresti, A. 2002. *Categorical Data Analysis*. 2nd ed. New York: Wiley.

Akerlof, George. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84(3): 488–500.

Banisar, David. 2004. "Global Survey: Freedom of Information and Access to Government Record Laws Around the World." Available at www.freedominfo.org.

Barrett, David M. 2005. *The CIA and Congress: The Untold Story*. Lawrence: University Press of Kansas.

Basler, Roy P., ed. 1953. "The Lincoln–Douglas Debate at Ottawa, August 21, 1858." In *The Collected Works of Abraham Lincoln*. Vol. 3. New Brunswick, NJ: Rutgers University Press.

Baum, Matthew. 2004. "Public Opinion, Presidential Rhetoric, and the Domestic Politics of Audience Costs in US Foreign Policy Crises." *Journal of Conflict Resolution* 48(5): 603–31.

Baum, Matthew, and Tim Groeling. 2004. "Crossing the Water's Edge: Elite Rhetoric, Media Coverage and the Rally-Round-the-Flag Phenomenon, 1979-2003." 100th annual meeting of the American Political Science Association, Chicago.

Bennett, D. Scott, and Allan Stam. 1998. "The Declining Advantages of Democracy: A Combined Model of War Outcomes and Duration." *Journal of Conflict Resolution* 42(3): 344–66.

Blair, Clair, Jr. 2001. *Silent Victory: The US Submarine War Against Japan*. Annapolis, MD: Naval Institute Press.

Born, Hans, and Ian Leigh. 2007. "Intelligence Accountability: A Comparative Perspective." In *Strategic Intelligence: Understanding the Hidden Side of Government*, ed. Loch K. Johnson. Westport, CT: Greenwood, 141–64.

Brody, Richard A., and Catherine R. Shapiro. 1989. "Policy Failure and Public Support: The Iran-Contra Affair and Public Assessments of President Reagan." *Political Behavior* 11(2): 353–69.

Brody, Richard, and Benjamin Page. 1975. "The Impact of Events on Presidential Popularity." In *Perspectives on the Presidency*, ed. Aaron Wildavsky. Boston: Little, Brown, 136–68.

Bueno de Mesquita, Bruce, Alastair Smith, Randolph Siverson, and James Morrow. 2003. *The Logic of Political Survival*. Cambridge, MA: MIT Press.

Colaresi, Michael. 2007. "The Benefit of the Doubt: Testing an Informational Theory of the Rally." *International Organization* 61(1): 99–144.

Colaresi, Michael P. 2011. "Codebook for the Security Policy Information Institutions (SPII) Data." Unpublished manuscript. Michigan State University.

Colaresi, Michael P., Karen Rasler, and William R. Thompson. 2007. *Strategic Rivalries in World Politics: Position, Space and Escalation*. Cambridge: Cambridge University Press.

Correlates of War Project. n.d. "State System Membership List, v2004.1." http://correlatesofwar.org. Forthcoming.

Crescenzi, Mark, Kelly Kadera, Sara Mitchell, and Clayton Tyne. 2011. "A Supply Side Theory of Mediation." *International Studies Quarterly* 55(4): 1069–94.

Desch, Michael. 2002. "Democracy and Victory: Why Regime Type Hardly Matters." *International Security* 27(2): 5–47.

Downes, Alexander B. 2009. "How Smart and Tough Are Democracies? Reassessing Theories of Democratic Victory in War." *International Security* 33(4): 9–51.

Edmonds, Robin. 1986. "Yalta and Potsdam: Forty Years Afterwards." *International Affairs* 62(2): 197–216.

Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88(3): 577–93.

Fisher, Louis. 2007. *Presidential War Powers*. Lawrence: University of Kansas Press.

Garthoff, Raymond. 1989. *Reflections on the Cuban Missile Crisis*. Washington, DC: Brookings Institution Press.

Gaubatz, Kurt. 1999. *Elections and War: The Electoral Incentive in the Democratic Politics of War and Peace*. Palo Alto, CA: Stanford University Press.

Ghosn, Faten, Glenn Palmer, and Stuart Bremer. 2004. "The MID3 Data Set, 1993-2001: Procedures, Coding Rules, and Description." *Conflict Management and Peace Science* 21(1): 133–54.

Gibbs, David N. 1995. "Secrecy and International Relations." *Journal of Peace Research* 32(2): 213–26.

Gill, Peter. 1994. *Policing Politics: Security Intelligence and the Liberal Democratic State*. New York: Frank Cass.

Heidelberger, P., and P. D. Welch. 1983. "Simulation Run Length Control in the Presence of an Initial Transient." *Operations Research* 31(3): 1109–44.

Hensel, Paul. 1994. "One Thing Leads to Another: Recurrent Militarized Disputes in Latin America, 1816-1986." *Journal of Peace Research* 31(3): 281–98.

Hensel, Paul R., Sara Mitchell, Thomas Sowers, and Clayton Tyne. 2008. "Bones of Contention: Comparing Territorial, Maritime and River Issues." *Journal of Conflict Resolution* 52(1): 117–43.

Holsti, Ole. 1996. *Public Opinion and American Foreign Policy*. Ann Arbor: University of Michigan Press.

Huth, Paul, and Todd Allee. 2003. *The Democratic Peace and Territorial Conflict in the Twentieth Century*. Cambridge: University of Cambridge Press.

Ikenberry, G. John. 2001. *After Victory: Institutions, Strategic Restraint and the Rebuilding of Order after Major Wars*. Princeton, NJ: Princeton University Press.

Jackson, Peter. 2000. *France and the Nazi Menace: Intelligence and Policy Making, 1933-1939*. Oxford: Oxford University Press.

Jaggers, Keith, and Tedd Robert Gurr. 1995. "Transitions to Democracy: Tracking Democracy's Third Wave with the POLITY III Data." *Journal of Peace Research* 32(6): 469–82.

Johnson, Loch K. 1997. *America's Secret Power*. Cambridge: Cambridge University Press.

Johnson, Robert David. 2006. *Congress and the Cold War*. Cambridge: Cambridge University Press.

Kahn, David. 1991. *Seizing the Enigma*. New York: Houghton Mifflin.

Kissinger, Henry. 1995. *Diplomacy*. New York: Simon and Schuster.

Kurian, George Thomas. 1998. *World Encyclopedia of Parliaments and Legislatures*. Vol. 2. London: Inter-Parliamentary Union.

Lake, David. 1992. "Powerful Pacifists: Democratic States and War." *American Political Science Review* 86(1): 24–37.

Lippmann, Walter. 1922. *Public Opinion*. New York: Free Press.

Maoz, Zeev. 2005. "Dyadic Militarized Interestate Disputes (version 2.0)." Dataset available at http://psfaculty.ucdavis.edu/zmaoz/dyadmid.html (accessed June 15, 2010).

Martin, Lisa. 2000. *Democratic Commitments*. Princeton, NJ: Princeton University Press.

Mearsheimer, John. 2004. "Lying in International Politics." 100th annual meeting of the American Political Science Meeting, Chicago.

Milner, Helen. 1997. *Interests, Institutions and Information: Domestic Politics and International Relations*. Princeton, NJ: Princeton University Press.

Morgenthau, Hans. 1967. *Politics among Nations*. 4th ed. New York: Knopf.

Page, Benjamin, and Robert Shapiro. 1995. *The Rational Public: Fifty Years of Trends in Americans' Policy Preferences*. Chicago: University of Chicago Press.

Plummer, Martyn. 2003. "JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling." Proceedings of the 3rd International Workshop on Distributed Statistical Computing, March 2022, Vienna, Austria.

Prados, John, and Margaret Pratt Portier. 2004. *Inside the Pentagon Papers*. Lawrence: University of Kansas Press.

Putnam, Robert D. 1988. "Diplomacy and Domestic Politics: The Logic of Two-Level Games." *International Organization* 42(3): 427–60.

Rabin, Yitzhak. 1996. *The Rabin Memoirs*. Berkeley: University of California Press.

Raftery, A. E., and S. M. Lewis. 1992. "One Long Run with Diagnostics: Implementation Strategies for Markov Chain Monte Carlo." *Statistical Science* 7(1): 493–97.

Reiter, Dan, and Allan C. Stam. 2002. *Democracies at War*. Princeton, NJ: Princeton University Press.

Reiter, Dan, and Allan C. Stam. 2009. "Correspondence: Another Skirmish in the Battle over Democracies and War." *International Security* 34(1): 194–200.

Ripsman, Norrin M. 2009. "Neoclassical Realism and Domestic Interest Groups." In *Neoclassical Realism, the State and Foreign Policy*, ed. Steven E. Lobell, Norrin M. Ripsman, and Jeffrey W. Taliaferro. Cambridge: Cambridge University Press, 170–93.

Risen, James, and Eric Lichtblau. 2009. "Media Incentives and National Security Secrets." *Harvard Law Review* 122: 2228–49.

Rosato, Sebastian. 2003. "The Flawed Logic of the Democratic Peace." *American Political Science Reivew* 97(4): 585–602.

Russett, Bruce, and John Oneal. 2000. *Triangulating Peace*. New York: W. W. Norton.

Sagan, Scott, and Jeremi Suri. 2003. "The Madman Nuclear Alert: Secrecy, Signaling, and Safety in October 1969." *International Security* 27(4): 150–83.

Schultz, Kenneth. 2001. *Democracy and Coercive Diplomacy*. Cambridge: Cambridge University Press.

Schultz, Kenneth, and Barry Weingast. 2003. "The Democratic Advantage." *International Organization* Winter(1): 3–42.

Schweller, Randall. 2004. "Unanswered Threats: A Neoclassical Realist Theory of Underbalancing." *International Security*, 29(2): 159–201.

Senese, Paul D., and Stephen L. Quackenbush. 2003. "Sowing the Seeds of Conflict: The Effect of Dispute Settlements on Durations of Peace." *Journal of Politics* 65(3): 696–717.

Shlaim, Avi. 2001. *The Iron Wall: Israel and the Arab World*. New York: Norton.

Singer, J. David. 1987. "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816-1985." *International Interactions* 14: 115–32.

Snyder, Jack L. 1991. *Myths of Empire: Domestic Politics and International Ambition*. Ithaca, NY: Cornell University Press.

Springall, A. 1973. "Response Surface Fitting Using a Generalization of the Bradley-Terry Paired Comparison Model." *Applied Statistics* 22(1): 59–68.

Stigler, Stephen. 1994. "Citation Patterns in the Journals of Statistics and Probability." *Statistical Science* 9(1): 94–108.

Stuart-Fox, D. M., D. Firth, A. Moussalli, and M. J. Whiting. 2005. "Multiple Signals in Chameleon Contests: Designing and Analysing Animal Contests as a Tournament." *Animal Behavior* 71(4): 1263–71.

Szilagyi, Andrew. 2010. "Blowing Its Cover: How the Intelligence Identities Protection Act Has Masqueraded as an Effective Law and Why It Must Be Amended." *William and Mary Law Review* 51(6): 2269–2312.

Thompson, William R. 2001. "Identifying Rivalries in World Politics." *International Studies Quarterly* 45(2): 557–86.

Turner, Heather, and David Firth. 2010. "Bradley-Terry Models in R: The BradleyTerry2 Package." Unpublished manuscript. University of Warwick.

Tutz, Gerhard. 1986. "Bradley-Terry-Luce Models with an Ordered Response." *Journal of Mathmatical Psychology* 30(3): 306–16.

U.S. Budget Summary. 2010. Government Print Office, available at http://www.gpoaccess.gov/usbudget/fy10/pdf/budget/.

Waltz, Kenneth. 1967. *Foreign Policy and Democratic Politics*. Boston, MA: Little, Brown.

Waltz, Kenneth. 1979. *The Theory of International Politics*. Reading, MA: Addison Wesley.

Wright, Tony. 1994. *Citizens and Subjects: An Essay on British Politics*. London: Routledge Press.