

INFORMATION AND INTERNATIONAL CONFLICT

Christopher Liu

A DISSERTATION

in

Political Science

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2019

Supervisor of Dissertation

---

Avery Goldstein

David M. Knott Professor of Global Politics and International Relations

Graduate Group Chairperson

---

Alex Weisiger, Associate Professor of Political Science

Dissertation Committee

Avery Goldstein, David M. Knott Professor of Global Politics and International Relations

Alex Weisiger, Associate Professor of Political Science

Matthew Levendusky, Professor of Political Science

Michael C. Horowitz, Professor of Political Science

*For Ryon*

## ACKNOWLEDGMENT

I am grateful to my dissertation committee for supporting my efforts throughout the dissertation process and through my time as a graduate student. Avery's feedback consistently helped ground my research in the real world. Mike's comments pushed me to view my research from different perspectives and to defend my approach, strengthening the dissertation as a result. Alex gave tremendous insight, feedback, and suggestions at all stages of the project—the dissertation would be in much poorer shape in his absence. Matt provided insightful feedback on each paper at various points throughout the process, in addition to invaluable advice and support that helped me navigate the twists and turns of the past few years. My sincere thanks to you all.

A fundamental principle in game design is that players must be presented with obstacles and barriers—fun is the product of overcoming those obstacles. And so it would be remiss of me to not also thank those individuals who have presented me with such challenges—it's cliché, but I am where I am today because of you. Facing down those obstacles was not always fun, but standing here, at the end, sure is.

Lastly, I will forever be grateful to my partner Ryon. I am turning the page on this dissertation, but we are just starting the next chapter of our co-authored project together.

# **ABSTRACT**

## INFORMATION AND INTERNATIONAL CONFLICT

Christopher Liu

Avery Goldstein

Information plays a critical role in theories of interstate behavior. This dissertation examines the question: how does domestic politics structure and condition the role of information in international conflict?

In the first essay, I test an observable implication of audience cost theory by exploiting elections as a way to identify variation in the costliness of interstate signaling, under the assumption that election-year threats are more costly than threats made in non-election years. I find evidence consistent with audience cost theory in the behavior of democratic targets, but fail to find similar results for targeted autocracies. The findings overall suggest variation in domestic political institutions can explain whether and when information about threat credibility is received and interpreted by target states.

In the second essay, I examine a previously untested assumption of diversionary war theory by investigating whether and how initiation status affects the public's propensity to rally 'round the flag. Using a survey experiment with researcher manipulation of initiation status, I find heterogeneous rallies across partisan identification, with rallies limited to in-partisans under the initiation condition, but widespread rallies across the full sample when a leader is seen as responding to foreign provocation.

In the third essay, I develop conventional diversionary theory in two ways. First, I argue that diversion is fundamentally the manipulation of foreign policy salience, and that this objective is generally accomplished through actions short of war. Second, I argue that domestic constraints—specifically, a country's media system and the flow of information between elites and citizens—

condition the ability of leaders to reap the rewards of diversion. The results provide tentative support for the theory, and are consistent with media systems playing an important role in conditioning and structuring the relationship between domestic unrest and conflict.

## TABLE OF CONTENTS

<b>ACKNOWLEDGMENT .....</b>	<b>III</b>
<b>ABSTRACT .....</b>	<b>IV</b>
<b>TABLE OF CONTENTS .....</b>	<b>VI</b>
<b>LIST OF TABLES .....</b>	<b>VIII</b>
<b>LIST OF ILLUSTRATIONS .....</b>	<b>IX</b>
<b>CHAPTER 1: INTRODUCTION.....</b>	<b>1</b>
References .....	7
<b>CHAPTER 2: AUDIENCE COSTS AND THE ELECTORAL CYCLE .....</b>	<b>10</b>
Audience costs and its critics .....	11
Audience costs and the electoral cycle.....	14
Research Design .....	19
Results .....	23
Discussion .....	35
Conclusion.....	38
References .....	41
<b>CHAPTER 3: DO PUBLICS REWARD BELLIGERENCE? RALLY ‘ROUND THE FLAG AND DIVERSIONARY WAR .....</b>	<b>46</b>
The rally ‘round the flag effect.....	47
Conflict initiation vs. conflict reciprocation .....	50
Research design .....	52
Results .....	55

Discussion and conclusion .....	62
References .....	68
<b>CHAPTER 4: DIVERSIONARY THEORY UNDER MEDIATED POLITICS .....</b>	<b>72</b>
Literature review .....	75
Theory .....	78
Data.....	83
Results .....	89
Discussion and conclusion .....	98
References .....	101
<b>CHAPTER 5: CONCLUSION .....</b>	<b>107</b>
Common linkages .....	110
Implications .....	114
Areas for future research .....	116
References .....	117
<b>APPENDIX .....</b>	<b>118</b>
Appendix 2.1: Percentage and number of militarized compellent threats met with compliance .....	118
Appendix 2.2: Table of robustness checks .....	119
Appendix 3.1: Sample characteristics in comparison with US population.....	120
Appendix 3.2: News stories used in survey experimental conditions .....	121
Appendix 3.3: Logistic regression results .....	124
Appendix 4.1: Not free country-years .....	125
Appendix 4.2: Robustness checks.....	131

## LIST OF TABLES

Table 1: Percentage and number of reciprocated threats .....	22
Table 2: Logistic regressions, reciprocation on initiator election year .....	25
Table 3: Linear probability model, conflict on unrest and media freedom, with fixed-effects and robust standard errors.....	90
Table 4: Unrest, media freedom, and conflict intensity.....	95



## LIST OF ILLUSTRATIONS

Figure 1: Marginal effect of election-year and control variables on reciprocation by target regime type with 95 per cent confidence intervals .....	28
Figure 2: Marginal effect of election-year on reciprocation in interactive model with 95 per cent confidence intervals .....	30
Figure 3: Trump favourability by condition .....	56
Figure 4: Likelihood of voting for Trump by condition .....	57
Figure 5: Trump favourability by party ID and condition .....	59
Figure 6: Likelihood of voting for Trump by party ID and condition .....	60
Figure 7: Number of not free media states in the world, 1948-2010 .....	87
Figure 8: Change in probability of initiation by unrest and media freedom.....	92

## CHAPTER 1: Introduction

Geoffrey Blainey, paraphrasing the argument of Georg Simmel, once wrote that “the most effective way of preventing a war was to possess exact knowledge of the comparative strength of the two rival nations” (Blainey 1973, 118). In a world where all parties possess perfect and complete information of relative capabilities and resolve, conflict can always be avoided—insofar as fighting entails an absolute destruction of property, there is always a theoretical settlement that is preferable to war for all states. In bargaining models of war, the fundamental barrier to peace is the private nature of such information (Fearon 1995). Resolve may be tested, and it may be hinted at by examining situational and dispositional factors (Kertzer 2014), but at the end of the day estimating the resolve of an adversary is guesswork—one that carries enormous risk for those who make erroneous judgments. However, states can attempt to communicate private information about resolve to others by taking costly actions, such as actions that tie one's hands, or those that incur sunk costs (Fearon 1997). In such situations, information about the costs associated with such signals is key.

Clearly information plays a crucial role in the bargaining model of war. Information is also key in other theories that attempt to explain the causes of conflict. Proponents of diversionary war suggest that leaders may initiate foreign conflicts in order to distract a domestic population from woes at home to foes abroad. Implicit in diversionary theory, two distinct pieces of information are at play. First, leaders have information about the anticipated reactions of their domestic publics if foreign conflict occurs—that is to say, leaders anticipate that the public will abandon their previous concerns over domestic conditions in order to rally around the foreign enemy. Second, the public has information about the nature and existence of the foreign conflict itself.

Thus, although the bargaining model of war and diversionary war theory differ in significant ways, both make assumptions about the flow of information between important actors. In addition, both theories contain assumptions about the domestic public. In fact, I suggest that domestic audiences in the canonical versions of both theories are treated much the same way

that states are conceptualized in neorealist theory (Waltz 1979): unitary; internally undifferentiated; and largely lacking in agency. Fearon (1997) describes how public statements carry information about cost and about credibility, because a domestic public will punish a leader for backing down; there is little discussion of whether subsets of the public are more or less likely to do so, or if different publics in different states operate under varying logics. Diversionary theory turns on a domestic public that will reliably turn its attention, en masse, away from domestic troubles towards foreign tensions should the latter arise; less theoretical focus is placed on whether all publics and whether all members of any given public are willing to rally around their leader.

To be sure, developments in both literatures have made significant efforts to differentiate and disaggregate domestic publics. One body of work on audience costs attempt to identify whether democracies, with a more politically empowered domestic public, are able to more credibly signal resolve than autocracies (Eyerman and Hart 1996; Partell and Palmer 1999; Baum and Potter 2015; Schultz 2001; Weeks 2008; Weiss 2014). Similarly, scholars have probed the scope conditions of diversionary war by assessing whether democracies, given their salient publics, are also more prone to diversion (Russett 1990; Gelpi 1997; Oneal and Tir 2006; Lai and Slater 2006). More recently, a second vein of scholarship has focused on the microfoundations of audience cost by examining who among the domestic public is willing to punish a reneging leader, why, and under what conditions (Tomz 2007; Kertzer and Brutger 2016; Levy et al. 2015; Quek Forthcoming; Levendusky and Horowitz 2012; Trager and Vavreck 2011; Chaudoin 2014). To the extent that the microfoundations of diversionary theory have been tested, it has been done by scholars writing on rally 'round the flag effects who have examined rally heterogeneity across various subgroups (Baum 2002; Groeling and Baum 2008; Hetherington and Nelson 2003).

Building on prior work, this dissertation makes the case that scholarship on conflict and in international relations writ large can benefit from greater attention on the interaction of information and domestic politics. Specifically, *domestic politics structure and condition the role that information plays in international politics*—information never exists in a vacuum.

To demonstrate and support this claim, the three articles in this dissertation each examine a type of what I call *information failure*, by which I mean instances in which the role that information plays empirically differs from how it ought to operate according to theory. In particular, I identify three examples of information failure that can be produced by domestic politics.

*Domestic politics can explain when information fails to affect state behavior in predicted ways*

Audience cost theory suggest that public signals are credible and costly signals, and are one way truly resolved states can differentiate themselves from merely-bluffing low-resolve states. Correspondingly, an observable implication of the theory is that costly signals should be more effective, where effectiveness is defined as the target of the signal backing down.

The first article, or chapter two, tests this observable implication by leveraging election years in democracies as a way to identify variation in the costliness of threats, under the assumption that election-year threats are more costly than those made in non-election years. I also test whether differences in the domestic politics of target states play a role in determining the effectiveness of costly threats. By analyzing existing data sets on militarized disputes and elections, as well as original data of democratic election years, I find that threats made by democracies, during an election year and against democratic targets, are more effective than non-election year threats. At the same time, I fail to find evidence that costly threats are more effective against autocracies. Thus, the domestic politics of target states—in this case, the regime type of the target—can explain why costly threats can be effective against some states, but less effective in others.

*Domestic politics can explain why the same information can produce different effects for different audiences*

Diversionary war theory implicitly assumes that leaders who initiate conflicts abroad can elicit and benefit from rally 'round the flag effects, just like those observed in instances where a leader is responding to a foreign provocation. The extant literature on rally effects has been inconclusive in isolating the role of initiation status due to the limits of observational data.

Chapter three addresses this gap in the literature by manipulating initiation status in the context of a survey experiment. The findings suggest heterogeneous effects of initiation status across partisan identification. Specifically, conflicts initiated by the leader only produce rally effects among in-partisans, while conflicts in which the leader is responding to a foreign provocation produce observable rallies across the sample. The results thus suggest that domestic politics and partisan identities can shape the consequences of foreign conflicts—in other words, that information about foreign conflict is filtered through domestic politics to prismatic effect.

*Domestic politics can explain how and when the flow of information is constrained*

Studies of diversionary war have rarely problematized how reliable a diverting leader can reap the rewards of diversion. Diversionary theory assumes that leaders expect foreign aggression to ameliorate their domestic woes, but how the pattern of rally effects operate within diverse institutional arrangements remains underexplored.

Chapter four examines the consequences of varying domestic political institutions on this relationship between public opinion and foreign policy. In particular, it argues that media systems—the degree of autonomy with which journalists and media entities can operate in a given state—can place constraints on how reliable leaders in different states may reap the rewards of diversion. The empirical findings, while providing mixed support for the theoretical predictions, do suggest that media systems can condition and moderate the relationship between unrest and conflict that lies at the heart of diversionary theory. This paper thus provides evidence of how domestic politics structure the flow of information, and the consequences that such constraints may have for international conflict.

Individually, the three papers each make contributions to the literatures on audience cost, rally 'round the flag, and diversionary war. Each paper also has important implications for the others. First, chapter two suggest that autocrats may be less likely to recognize audience costs—at least costs induced by proximity to elections—while chapter four argues that autocrats are most likely to engage in diversionary foreign policy. It may be the case, then, that diversionary motivations explains the finding in chapter two. That is to say, it may be the case that autocrats fail to recognize the credibility of threats made by democracies in election years because they interpret such threats as being animated by diversionary motives and the desire to win elections, and thus treat them no differently than they view threats made in non-election years.

Second, chapter three suggests that reciprocated conflicts generate more widespread rally effects. Insofar as autocrats, all else being equal, care more about political survival than democratic leaders (due to the greater severity of punishment for the loss of political power in autocracies), autocratic leaders may be more willing to reciprocate costly and credible threats than democratic targets are—thus providing another explanation for the empirical patterns observed in chapter two.

Third, chapter four makes the case that IR scholarship should move to unpack the black box of regime type and to more clearly theorize how specific institutional arrangements within particular regimes can drive predicted outcomes. Adopting such a perspective suggests an alternative specification to the analyses conducted in chapter two. Instead of examining threat effectiveness among democratic versus autocratic targets, consider how credibility is perceived by states with varying degrees of media freedom. In particular, the importance of regime type for audience cost theory stems from the “plausible working hypothesis” that democracies enjoy stronger domestic audiences than authoritarian regimes (Fearon 1994). Regime type, however, is only one way to operationalize the concept of domestic audience strength; the theory presented in chapter four suggest that when leaders exercise political control over the mass media, audiences will necessarily be weak because of leader control over the flow of information about

foreign affairs that can be accessed by the mass public. Thus, the theoretical argument advanced in chapter four suggest ways to build on the research conducted in the first paper.

Lastly, chapters three and four together make complementary arguments that both support a reconceptualization of diversionary theory. The paper in chapter four argues that leaders can achieve the benefits of diversion through actions short of war that nevertheless manipulate foreign policy salience. In chapter three, salience is in effect purposefully manipulated by the researcher in the context of a survey experiment, and the treatment groups receive an artificially high dosage of foreign policy coverage. Insofar as the results of that chapter suggest rallies result from voter perceptions about a foreign policy event, chapter three provides some of the empirical microfoundations for the diversionary theory articulated in the last paper, which provides complementary evidence by examining behaviour at the country level.

Following the three papers, chapter five of the dissertation concludes by summarizing the main findings and limitations of the three studies; discussing linkages and common threads across the manuscripts; and articulating an agenda for future research.

## References

- Baum, Matthew A. 2002. "The Constituent Foundations of the Rally-Round-the-Flag Phenomenon." *International Studies Quarterly* 46(2):263–298.
- Baum, Matthew A. and Philip B. K. Potter. 2015. *War and Democratic Constraint: How the Public Influences Foreign Policy*. Princeton University Press.
- Blainey, Geoffrey. 1973. *The Causes of War*. London: MacMillan Press.
- Chaudoin, Stephen. 2014. "Policies or Promises? An Experimental Analysis of International Agreements and Audience Reactions." *International Organization* 68(1):235–256.
- Eyerman, Joe and Robert A. Hart. 1996. "An Empirical Test of the Audience Cost Proposition." *Journal of Conflict Resolution* 40(4):597–616.
- Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.
- Fearon, James D. 1997. "Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs." *Journal of Conflict Resolution* 41(1):68–90.
- Gelpi, Christopher. 1997. "Democratic Diversions: Governmental Structure and the Externalization of Domestic Conflict." *Journal of Conflict Resolution* 41(2):255–282.
- Groeling, Tim and Matthew Baum. 2008. "Crossing the Water's Edge: Elite Rhetoric, Media Coverage, and the Rally-Round-the-Flag Phenomenon." *Journal of Politics* 70(4):1065–1085.
- Hetherington, Marc J. and Michael Nelson. 2003. "Anatomy of a Rally Effect: George W. Bush and the War on Terrorism." *PS: Political Science and Politics* 36(1):37–42.
- Kertzer, Joshua D. 2014. *Resolve in International Politics*. Princeton: Princeton University Press.
- Kertzer, Joshua and Ryan Brutger. 2016. "Decomposing Audience Costs: Bringing the Audience Back Into Audience Cost Theory." *American Journal of Political Science* 60(1):234–249.
- Lai, Brian and Dan Slater. 2006. "Institutions of the Offensive: Domestic Sources of Dispute



- Initiation in Authoritarian Regimes, 1950-1992." *American Journal of Political Science* 50(1):113–126.
- Levendusky, Matthew S. and Michael C. Horowitz. 2012. "When Backing Down is the Right Decision: Partisanship, New Information, and Audience Costs." *Journal of Politics* 74(2):323–338.
- Levy, Jack S., Michael K. McKoy, Paul Poast and Geoffrey P. R. Wallace. 2015. "Backing Out or Backing In? Commitment and Consistency in Audience Costs Theory." *American Journal of Political Science* 59(4):988–1001.
- Oneal, John R. and Jaroslav Tir. 2006. "Does the Diversionary Use of Force Threaten The Democratic Peace? Assessing the Effect of Economic Growth on Interstate Conflict." *International Studies Quarterly* .
- Partell, Peter J. and Glenn Palmer. 1999. "Audience Costs and Interstate Crises: An Empirical Assessment of Fearon's Model of Dispute Outcomes." *International Studies Quarterly* 43(2):389–405.
- Quek, Kai. Forthcoming. "Type II Audience Costs." *Journal of Politics*.
- Russett, Bruce. 1990. *Controlling the Sword: The Democratic Governance of National Security*. Cambridge: Harvard University Press.
- Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy*. New York: Cambridge University Press.
- Tomz, Michael. 2007. "Domestic Audience Costs in International Relations: An Experimental Approach." *International Organization* 61(4):821–840.
- Trager, Robert F. and Lynn Vavreck. 2011. "The Political Costs of Crisis Bargaining: Presidential Rhetoric and the Role of Party." *American Journal of Political Science* 55(3):526– 545.
- Waltz, Kenneth. 1979. *Theory of International Politics*. McGraw-Hill.
- Weeks, Jessica P. 2008. "Autocratic Audience Costs: Regime Type and Signaling Resolve." *International Organization* 62(1):35–64.
- Weiss, Jessica C. 2014. *Powerful Patriots: Nationalist Protest in China's Foreign Relations*.

New York: Oxford University Press.

## CHAPTER 2: Audience Costs and the Electoral Cycle

Audience costs, it has been said, are the “dark matter” of international relations (Schultz 2012, 369). Under the theory’s original formulation, audience costs explain how leaders, under conditions of incomplete information and incentives to misrepresent, can credibly convey their resolve to adversaries (Fearon 1994). Since then, a growing literature has evolved to test and understand the implications and limits of the theory. At the same time, audience costs themselves are notoriously difficult to observe. This has led to a number of critiques that suggest the absence of evidence should, in fact, be seen as evidence of absence (Snyder and Borghard 2011; Trachtenberg 2012). In particular, critics contend that the elegance of a formal theoretic theory about interstate crisis bargaining fails to describe how states interpret and understand public threats when targeted. Such leaders, scholars suggest, do not believe public threats are made more credible by carrying potential audience costs.

Given the difficulties with directly observing leader beliefs, this paper takes an indirect approach. It asks the question: Do targeted leaders respond differently to high-cost vs. lowcost threats? If leaders do behave as audience cost theory expects, the answer should be in the affirmative. The paper then makes two novel contributions to the literature on audience costs. First, I exploit the presence of elections among democratic initiators in order to elicit variation in the costliness of threats, under the assumption that threats made in election years are more costly than those made in non-election years. Second, I theorize and test whether recognition of audience costs is contingent on target regime type.

Using data from the Militarized Interstate Disputes dataset and data on elections, I find mixed but theoretically intriguing support for audience cost theory. Democratic targets are significantly less likely to reciprocate threats made in an initiator’s election year, a result consistent with the explanation that these targets accurately understood such threats carry greater audience costs and thus have enhanced credibility. I fail to find evidence that autocratic targets differentiate between election-year and non-election year threats.

The paper proceeds as follows. First, I provide a brief overview of the development of and debates within the audience cost literature. Second, I defend the use of electoral cycles as a proxy for threat costliness and outline the reasoning for exploring target regime type differences. Third, I present the data used for evaluating my hypotheses. The fourth section details the results of the empirical analyses as well as robustness checks. The final section discusses these results and their implications both for audience cost theory and the democratic peace.

## **Audience costs and its critics**

In developing the concept of audience costs, Fearon (1994) introduced to the field of IR a theory that, to its defenders, “provided answers to ... fundamental puzzles about the nature of crises and war” (Schultz 2012, 369). Fearon begins by noting that “few international disagreements become wars” (1994, 578) because states have strong incentives to agree to nonmilitary settlements prior to conflict onset due to the costs and risks of waging war (see also Fearon 1995). At the same time, states have incentives to misrepresent their willingness to fight in order to obtain a more favorable settlement. Given these conflicting incentives, how can states credibly convey their resolve? Fearon points out that international crises are *public* events carried out before domestic audiences, and that as a result public actions by leaders are more credible. This is because leaders who then subsequently back down from a public threat risk punishment by their audiences at home. Because public statements by leaders carry these potential ‘audience costs,’ truly resolved states should be more willing to make such statements than weakly resolved states that merely wish to bluff.

At several places, and always briefly, Fearon develops the “plausible working hypothesis” (1994, 582) that domestic political structures condition bargaining effectiveness. If democracies in general have stronger domestic audiences than autocracies, the former’s inherently greater ability to credibly convey resolve in crises may help ameliorate the security dilemma between democracies, thus providing an explanation for the democratic peace (Babst 1964).

A number of studies since have attempted to quantitatively test Fearon's theory using observational data by exploiting this hypothesized distinction between regime type. Eyerman and Hart (1996) find that crises between two democracies involve fewer intra-crisis processes than disputes between a democracy and an autocracy. Partell and Palmer (1999) find that democracies are less likely to back down in disputes. More recently, scholars have sought to unpack the 'black box' of regime type by identifying pertinent institutional variation that can affect the generation of audience costs. Within democracies, a free and robust media environment (Baum and Potter 2015) and potent opposition parties (Schultz 2001) can increase the effectiveness of issued threats. Furthermore, due to variation in the number of stakeholders to whom different autocrats are beholden, at least some types of autocracies can seemingly approximate democracies in their ability to convey credible signals (Weeks 2008). While autocrats may not be punished at the ballot box like their democratic counterparts, they may credibly point to other manifestations of public opinion, such as mass nationalist protests, to credibly signal resolve (Weiss 2014). Some authoritarian regimes may also face more certain audience costs, since a smaller electorate can more reliably solve the collective action problem of imposing costs than a larger and more diffuse electorate under democracies.

In addition, a burgeoning experimental literature has sought to test the microfoundations of audience cost theory, including when and why voters punish leaders who back down. Tomz (2007) demonstrates that voters are willing to punish reneging leaders, while finding suggestive evidence that this arises because of citizens' concern about their country's reputation. In contrast, Kertzer and Brutger (2016) find evidence that audience costs arise from voters' aversion to both belligerence and leader inconsistency. The inconsistency explanation is further supported by Levy et al. (2015), who find that voters punish leaders not only when they back down from prior threats, but also when they fail to keep prior promises to stay out of conflicts (see also Quek Forthcoming). Other experimental studies have identified limitations to audience cost theory. Voters' punishments can be mitigated by presidential justifications (Levendusky and Horowitz

2012), support from oppositional partisan elites (Trager and Vavreck 2011), or if the electorate has strong prior policy preferences (Chaudoin 2014).<sup>1</sup>

At the same time, a number of scholars have raised questions about the extent to which audience cost theory explains actual crises. Examining 12 case studies of great power crisis bargaining, Trachtenberg (2012) concludes that leaders appear to avoid making definitive public statements, preferring ambiguity and flexibility to being locked in. Furthermore, the target of threats appears to privilege a variety of information, not just the possibility of punishment from domestic audiences. Snyder and Borghard (2011) similarly examine case studies of post-WWII crises, concluding that there is little evidence for the influence of audience costs in the decision-making of leaders. In addition to these qualitative critiques, Downes and Sechser (2012) argue that when analyzing true instances of ‘compellent threats’ (see Schelling 1960), support for audience cost theory is considerably weaker.

Defenders of audience cost theory respond by pointing to the inherent difficulties in observing the theory at work. Gartzke and Lupu (2012) argue that the supposed advantages of qualitative process tracing techniques used by Trachtenberg (2012) and Snyder and Borghard (2011) may be overstated, insofar as both quantitative and qualitative approaches make use of observable data and evidence. However, audience costs are particularly unlikely to be observed directly (Schultz 2001; Weeks 2008), given that “a public that learned of a leader’s intentions [to generate audience costs] would be less likely to play the role required of it by the theory” (Gartzke and Lupu 2012, 394). Consequently, concluding that audience cost thinking was absent based on a lack of evidence of leader intention to create such costs “requires us to assume that a leader would generally record his intentions, a problematic assumption that may bias results” (ibid, 395). The difficulty of capturing audience costs empirically has not been entirely intractable; Kurizaki and Whang (2015) for example take a structural approach to estimate the size of audience costs in observational data on crisis behavior. Yet, these attempts sidestep rather than specifically address the theory’s critics—specifically the skepticism over whether leaders understand and

---

<sup>1</sup> Some observational studies have contributed to understandings of the limits of audience cost as well, such as Croco (2011), which examines the importance of leader culpability.

recognize audience costs associated with interstate signaling, which is fundamentally a concern over leader beliefs. Given the challenge of directly observing such beliefs, one may instead look for indirect evidence by examining 'second-order' observables: what else should be true of the world, if we assume that leaders do recognize audience costs? Rather than attempt to find a written instance in which a targeted leader explicitly references the domestic audience of the threatening state, it may instead be more prudent to see whether leaders respond differently to threats with and without associated audience costs. If they do, such behavior would be consistent with leaders recognizing a distinction between threats made costly by audience costs and those that are not. While a research design utilizing indirect evidence may only be a second-best approach, attempting to cut through the Gordian Knot of audience cost debates by taking a degree of loss in causal inference may be an acceptable trade-off.

Existing quantitative studies that examine audience cost theory in the context of threat effectiveness have focused on the role of institutions such as opposition parties (Schultz 2001), media availability (Baum and Potter 2015), and selectorate size (Weeks 2008). Such studies provide evidence for between-country variation in threat effectiveness with patterns consistent with audience cost theory. Yet, there are two primary ways in which the existing literature lacks a direct test of the previously identified observable implication. First, there are a variety of reasons why targets may respond differently to a threat made by a personalist regime than one made by a military junta, and perceived differences in associated audience costs is only one of many possible explanations. More importantly, however, such a design implicitly assumes an overly deterministic relationship between institutions and costly signaling. What is needed is a way to measure and capture temporal variation in costliness of threats without reference to largely time-invariant institutional features of regime type.

## **Audience costs and the electoral cycle**

This paper exploits elections in democratic initiators to identify variation in the level of audience costs associated with threats. It relies on one key assumption: *threats made by a democracy in an election year are more costly than threats made in a non-election year.*

Why might this assumption be true? Should a leader back down right before an election, they are much more likely to be punished for it at the ballot box than if they had backed down a year or two prior to the election. A voter's decision at the ballot box is determined by a number of factors, not the least of which being partisan identification and economic performance. For backing down to have appreciable effects on a voter's decision in the booth, foreign policy needs to be a particularly salient and accessible consideration in the mind of voters (Zaller 1992). Because accessibility is driven primarily by elite messaging and media coverage, the probability that a leader's backing down will be an accessible consideration by which voters evaluate candidates is greater when that backing down occurs proximate to the vote decision. In other words, the probability that voters punish a leader electorally for backing down decays as the period of time between backing down and voting increases.

While less intuitive, backing down immediately *after* an election may also be particularly costly for leaders. For the outgoing leader, a foreign policy failure at the end of her term may be viewed as a notable 'taint' on her political legacy. For the incoming leader, backing down means starting her tenure with a notable, visible setback during precisely the period of her term when the highest proportion of the electorate has no prior opinion towards her leadership (Mueller 1973). Given cognitive biases in how voters evaluate political leaders (Kunda 1990; Lodge and Taber 2000, 2005), any initial negative perceptions that result from backing down will affect how the public views and evaluates her actions for the remainder of her term. Such costs may not only manifest in leader approval. As Gelpi and Grieco (2015) note, backing down can result in 'competency costs' that significantly increase the difficulty leaders may encounter in pursuing their legislative agenda. Such costs are particularly punitive at the start of a leader's term in office, since the earlier in the term that competency costs are imposed, the greater the number of



potential legislative efforts that are affected.<sup>2</sup>In addition, it is worthwhile to note that in one of the most recent and precise studies of how audience costs vary across the electoral calendar, Chiozza (2017) finds that U.S. presidents are most likely to engage in costly signaling not only immediately prior to a presidential election, but also up to a year afterwards.

According to audience cost theory, if threats made in election years carry greater potential audience costs than those made in non-election years, election-year threats should also be more credible. If the targets of democratic initiators accept this logic, we should also observe systematic differences in how they respond to these two categories of threats. If no evidence is found that targets differentiate between election-year and non-election-year threats, that suggests the possibility that target leaders do not understand or do not value the signal of audience costs, bolstering the theory's critics.

**Hypothesis 1.** *Targets are less likely to reciprocate threats made in a democratic initiator's election year, compared to non-election years.*

There is some existing evidence for the importance of electoral cycles in crisis bargaining. Gaubatz (1991) finds that democracies are less likely to initiate wars when elections are less than one year away. Elections also heighten tension between the president and Congress (Rhode and Simon 1985), such that the signaling role of opposition parties (Schultz 2001) may be particularly important in election years. More recently, Chiozza (2017) examines U.S. presidential behavior across 66 international crises in the 20th and early 21st century. Like the present study, Chiozza argues that public statements of resolve are costlier when elections are proximate. However, this article goes beyond Chiozza in two crucial ways. First, Chiozza examines only US conflict initiation, while I examine conflicts initiated by all democracies starting

---

<sup>2</sup> To be sure, costs in the form of legislative defeats or decreases in leader approval are conceptually distinct from removal from office. However, a leader's approval directly affects her ability to pass favored legislation, while both legislative accomplishments and approval ratings affect probability of surviving elections. Furthermore, while the observational IR literature typically operationalizes costs to leader as removal from office, the experimental literature on audience costs operationalizes costs as decreases in approval.

in 1816. As a result, the sample size in the present study is approximately nine times larger. Second, Chiozza follows much of the remaining literature by focusing on whether the behavior of the initiator is consistent with audience cost theory. In contrast, I explore whether *targets* of costly signals also respond to threats in a manner predicted by audience cost theory. Insofar as the behavior of target states is an important element of game theoretic models of audience costs, yet have been underexamined by the empirical literature thus far, the present study represents an important contribution to the literature.<sup>3</sup>

Much of the audience cost literature has been concerned with understanding how institutional variation among initiators affects the ability to generate audience costs. If states vary in their ability to make costly signals, then it is at least plausible that states also differ when targeted by such threats, particularly in whether and how much they value any associated audience costs. As Snyder and Borghard (2011) note in their critique of audience cost theory, autocrats are less likely to be concerned with audience costs in interstate crises. In 1945, as the USSR, with forces still in parts of Iran, contemplated the establishment of a friendly pro-Soviet regime in that country, American diplomat Harry Hopkins met with Stalin and attempted to argue that American public opinion would react negatively to Soviet machinations in Iran. In response, Stalin angrily dismissed the possibility that public opinion could constrain U.S. policy. The Soviets saw public opinion not as “a major independent factor the government had to contend with, but rather was something that could be manipulated for instrumental purposes” (Trachtenberg 2012, 25).

Autocratic leaders, who retain office not through the goodwill of the electorate but through the support of a smaller winning coalition, can maintain power through transfers of private goods to members of the winning coalition, which holds true irrespective of whether a leader backs down or follows through on her public threats (Bueno de Mesquita et al. 2003). Furthermore, Snyder and Borghard point out that leaders tend to overestimate the degree of similarity between

---

<sup>3</sup> In effect, the present study entails an empirical trade-off compared with Chiozza, opting for a far larger sample size and temporal/spatial coverage at the expense of a more precise measurement of the independent variable. No single research design is perfect; it is the complementarity between individual studies that pushes forward a research agenda.

themselves and others. Thus, it's possible that while all states have incentives to try to adjudicate the credibility of an adversary's threat, democratic leaders are much more likely to take potential audience costs into account while doing so. Because autocrats' own foreign policy behavior is largely unconcerned with public opinion, they are less likely to see the same logic behind the threats made by adversaries. Insofar as election timing is used in the present study to generate variation in costs, we may expect democratic targets to systematically vary their responses to credible (election-year) and not credible (non-election year) threats, while autocratic targets would be less likely to make such a differentiation.

**Hypothesis 2.** *Democratic targets are less likely to reciprocate threats made in a democratic initiator's election year, compared to non-election years and compared to autocratic targets.*

There are two reasons to further investigate this question. The first is that prior literature has largely overlooked the possibility that institutional differences among *targets* of costly threats may help shed light on the debates surrounding audience cost theory. Snyder and Borghard (2011, 444) specifically restrict their sample case studies to democratic–autocratic dyads. Of the 12 case studies examined by Trachtenberg (2012), only one (the Fashoda crisis of 1898 between Britain and France) is between two democracies. Quantitative studies, by focusing on regime-type and institutional variation among initiators, have similarly overlooked variation in the other half of these dyadic interactions (Baum and Potter 2015; Schultz 2001; Weeks 2008). Second, differences in how audience cost signaling is interpreted by targets of different regime types may help develop Fearon's 'plausible working hypothesis' into a stronger explanation for the dyadic democratic peace (Maoz and Abdolali 1989). As Gartzke (2007) points out, if peace results from the superior signaling abilities of democracies, why are they unable to effectively signal to autocrats? In his original model, Fearon finds that as the difference in ability to generate audience costs between two adversaries increases, the higher-audience cost state is more likely

to escalate while the lower is more likely to back down (1994, 585). However, the latter proposition is true if and only if the lower-audience cost state recognizes that escalatory signals are credible due to associated audience costs. If the lower-audience cost state discounts or ignores the costs associated with such signals, the result is a monotonic increase in the probability of conflict onset as the the audience cost differential between the two states increases. Consequently, one possible explanation for the empirical finding of dyadic peace between democracies is the greater ability of democracies to evaluate the costliness (and hence credibility) of threats made by other democracies, relative to authoritarian targets of similar threats.

## Research Design

To test my hypotheses and to facilitate comparison with prior work, I follow previous observational studies on audience costs as closely as possible (Baum and Potter 2015; Schultz 2001; Weeks 2008). In these studies, the unit of analysis is dyadic dispute, defined as instances in which the initiator uses, explicitly threatens, or displays force against a target state's government, forces, or territory (Jones, Bremer and Singer 1996). As noted above, issues with selection bias complicate observational studies of audience cost. Leaders often prefer to avoid making explicitly hand-tying signals (Snyder and Borghard 2011) and are strategic in choosing when to elevate a potential conflict by making an explicit threat. Furthermore, elections can affect the ex ante costs and benefits of issuing a threat. The solution adopted by these scholars is to examine threat reciprocation by the target in order to capture the effectiveness of the initiator's threat. The intuition is that, *conditional on a threat having already been issued*, the more effective that threat is perceived to be by the target state, the more likely it is for the target to back down. In contrast, reciprocation by the target suggest that the threat failed and was ineffective.<sup>4</sup>

---

<sup>4</sup> Downes and Sechser (2012) suggest that the lack of reciprocation may not necessarily be the target backing down, but may actually be evidence of foot-dragging on the part of the target. However, because the target state faces its own domestic pressures and depending on the nature of the crisis, foot-dragging

My primary dataset is the Correlates of War Project's data on militarized interstate disputes (Ghosn, Palmer and Bremer 2004) from the period of 1816 to 2010. I drop multilateral disputes, focusing only on those MID's with one state as the initiator and one state as the target. Because I exploit democratic elections as a way of operationalizing differences in the costliness of various threats, it is appropriate to also restrict the sample to only those disputes initiated by democracies. This, in turn, requires an operationalization of democracy. I follow Schultz (2001) by defining democratic regimes as those that meet two criteria: First, having at least one chief executive chosen through competitive elections matching two or more major parties or candidates; and second, having relatively stable and enduring political groups that compete for political influence. I obtain this data via the Polity IV dataset (Marshall, Gurr and Jaggers 2014). Those who do not meet the criteria are defined as autocracies. Restricting my disputes sample to those initiated by democracies leaves me with 565 disputes during the period. Of these, 181 were made against targets that were also democracies, representing just under one-third of the sample.<sup>5</sup>

In keeping with prior literature, I operationalize my dependent variable of threat effectiveness by examining whether or not the target of the threat responds militarily to the threat. To that end, I analyze the variable RECIP in the MID's dataset which takes on a value of 1 if the target state takes a "codeable action" against the initiating state and 0 otherwise (Jones, Bremer and Singer 1996, 186). The expectation is that, *ceteris paribus* and conditional on a threat already made, a more credible and costly signal will be less likely to be resisted. Thirty-seven per cent of the disputes in the sample were reciprocated (N = 209).

---

may not always been a realistic strategy. Furthermore, the decision to foot-drag or provide no response is itself a strategic choice by the target to avoid explicit confrontation or escalation with the initiator, one that may be motivated by the perceived credibility of the initiator's threat.

<sup>5</sup> Target regime type was initially missing for 69 disputes. I manually inspect each of the 69 disputes to confirm that in the majority of cases missingness resulted from civil war/anarchy or foreign occupations; in these cases I code the state as non-democratic. Belize in 2000 and Liechtenstein in 1992 were missing from the Polity IV dataset; I code these states as democratic. I also code as democracies (again, under coding rules of Schultz (2001)): Cyprus during the 1960–65 period of intercommunal violence; Sweden during WWI; and Rhodesia. Excluding these 69 cases does not change the substantive results of the following analyses (see Robustness Checks).

My independent variable is a dichotomous indicator of whether or not the initiator has an election for the chief executive in the year of the dispute. From 1975 to 2010, I obtain the variable from the Database of Political Institutions (Beck et al. 2001). Due to the temporal limitations of the DPI, for the majority of disputes in the present dataset I collect original data on whether or not the dispute initiator had an election in the year of the dispute, using publicly available sources and following the same criteria as the DPI. Of the 565 disputes initiated by democracies during the period covered, 139 (roughly 25 per cent) were initiated in election years.<sup>6</sup>

As a first cut at the data and the plausibility of the hypotheses, Table 1 presents the percentage and number (in brackets) of disputes in which the initiator's threat was reciprocated across a number of categories. The first row displays the percentage and number of disputes in which the initial threat is reciprocated for the whole sample and by the whether the democratic initiator experiences an election in the same year. Overall, the initiator's electoral cycle appears to have little effect on whether the target reciprocates, in contrast to the prediction made under Hypothesis 1. This apparent uniformity, however, belies significant differentiation across democratic and authoritarian targets. As the second row indicates, consistent with Hypothesis 2 democratic targets appear less likely to respond to election-year threats, compared with non-election year threats (26 vs. 35 per cent). Authoritarian targets, in contrast, are somewhat *more* likely to respond to election-year threats (42 vs. 38 per cent).

---

<sup>6</sup> According to the DPI, from 1975 to 2010 21.9 per cent of democratic country-years had an election; that 25 per cent of disputes occurred in election years suggests that democratic leaders do not systematically avoid initiating disputes when elections are proximate.

Table 1: Percentage and number of reciprocated threats

	All	No Election	Election
All	37.0 (565)	36.9 (426)	37.4 (139)
Democracy	33.1 (181)	35.3 (139)	26.2 (42)
Autocracy	38.8 (384)	37.6 (287)	42.2 (97)

Cells display percentage of threats in each category that are reciprocated, with raw totals below in brackets.

To further test the two hypotheses, I estimate a series of logistic regression models with cluster-robust standard errors to account for error correlation between disputes initiated by the same state. The dependent variable is whether or not the target state responded to the initiator's threat and the independent variable is whether or not the democratic initiator had an election for the chief executive in the same year as the dispute. Although logistic regression models are appropriate for estimating the effect of independent variables on the probability of a binary response, I choose the specification mainly for consistency with past studies. Estimation via probit and linear probability models yield similar results (not shown).

Again in keeping with past literature, I include a standard set of control variables used in dyadic studies of interstate conflict. I control for the distribution of power within the crisis dyad, including whether each side was a major or minor power, as well as the initiator's share of capabilities. I code major powers following Schultz (2001)<sup>7</sup>, and obtain measures of state capabilities from version 4.0 of COW's National Material Capabilities Dataset (Singer 1987; Singer, Bremer and Stuckey 1972)<sup>8</sup>. I control for border contiguity using COW's Direct Contiguity Dataset (Stinnett et al. 2002). Lastly, I control for the issues at stake in the dispute using measures in the MID dataset.<sup>9,10</sup>

## Results

---

<sup>7</sup> Major powers consist of: US, UK, France (except 1940–45), Germany (1816–1918; 1925–1945; 1990–present), Austria (prior to 1918), Italy (1860–1943), Russia (except 1917–1922), China (after 1950), Japan (1895–1945; 1990–present).

<sup>8</sup> The initiator's 'share of capabilities' in the dyad is defined as the Composite Index of National Capabilities (CINC) score of the initiator divided by the sum of the CINC scores of the initiator and the target.

<sup>9</sup> The MID dataset includes a categorical variable capturing whether the initiator is seen as a revisionist state, defined as one that "demonstrate through its behavior a desire to change the status quo in a significant way" (Ghosn, Palmer and Bremer 2004, 138). Schultz (2001) disaggregates this measure, resulting in four dummy variables that each take a positive value if the initiator expresses a desire to claim territory, attempts to overthrow a regime, changes or does not abide by the target state's policies, or exhibits miscellaneous revisionist claims.

<sup>10</sup> In my main specifications I exclude measures of alliance portfolio similarity, which is yet another standard control variable in directed-dyadic analyses. Because alliance data is only available until 2000, its inclusion results in the listwise deletion of 109 observations. I subsequently include these measures as robustness checks.



Table 2 presents the results of a series of logistic regression models of reciprocation on elections in the threat initiator and control variables. Column one contains estimates for the regression among the full sample. Reiterating the finding from Table 1, the election-year coefficient is small and insignificant. Among all disputes initiated by democracies captured in the MID dataset, there is little indication that threats made in the initiator's election year are more or less likely to be met with a militarized response than threats made in non-election years. In other words, there is no evidence in support of Hypothesis 1.

Columns two and three contain regression estimates among autocracies and democracies respectively<sup>11</sup>. I find no evidence that authoritarian targets are more or less likely to reciprocate election-year threats compared with non-election year threats. In contrast, and in support of Hypothesis 2, election years do have an impact on the behavior of democratic targets. This effect is significant at  $p < 0.05$  significance level while the negative sign on the coefficient suggests that the effect is in the hypothesized direction. The results in columns two and three together show that democratic targets systematically differentiate between threats made in election years and non-election years while authoritarian targets do not. This empirical pattern is consistent with the possibility that democratic targets, compared to authoritarian targets, are more likely to view election-year threats as particularly credible, which in turn suggests that democratic leaders, more so than their authoritarian counterparts, place value on the audience costs associated with public interstate threats.

---

<sup>11</sup> The advantage of subsetting the sample by regime type is the ability to allow for variation in the effects of the control variables on reciprocation between democratic and authoritarian targets. The interactive model, in contrast, assumes these control variables do not differ between subsets.

**Table 2: Logistic regressions, reciprocation on initiator election year**

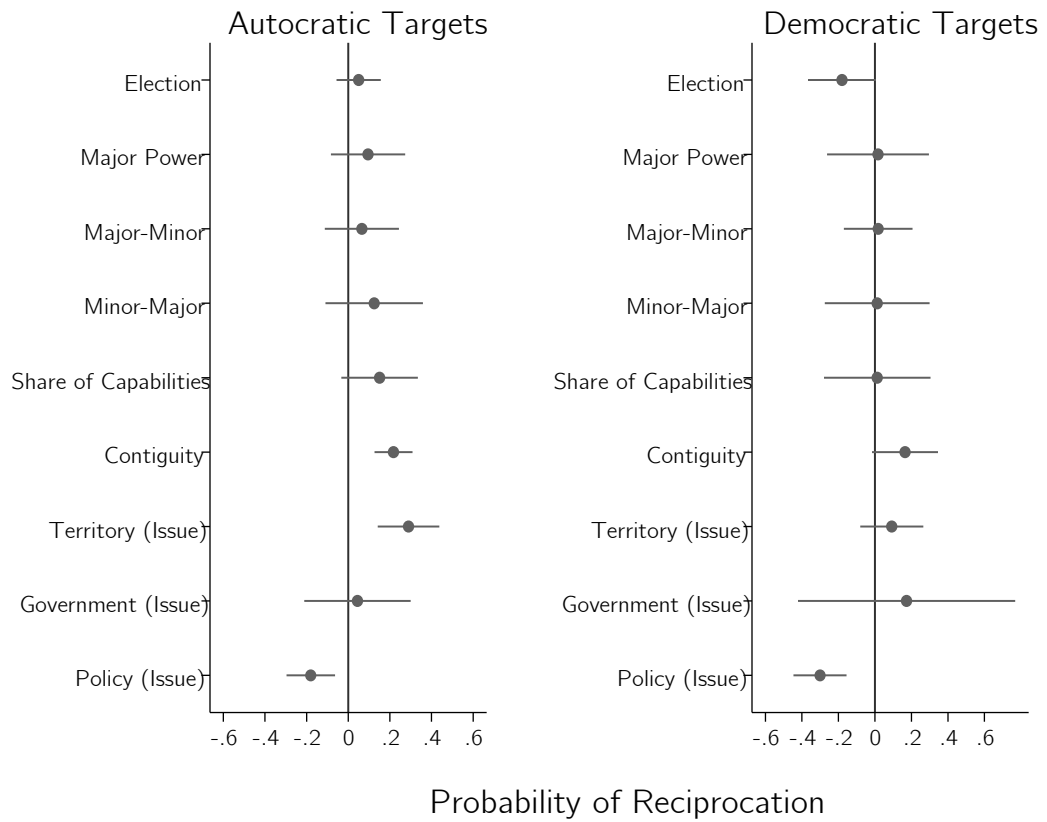
	(1)	(2)	(3)	(4)
	Full	Aut Target	Dem Target	Interaction
Election	-0.04 (0.21)	0.21 (0.24)	-0.88* (0.44)	0.20 (0.23)
Dem Target				-0.25 (0.22)
Election x Dem Target				-0.90† (0.49)
Major-Major	0.24 (0.37)	0.40 (0.39)	0.08 (0.70)	0.28 (0.41)
Major-Minor	0.13 (0.33)	0.28 (0.39)	0.09 (0.47)	0.18 (0.35)
Minor-Major	0.33 (0.37)	0.53 (0.51)	0.06 (0.72)	0.37 (0.38)
Share of Capabilities	0.58 (0.37)	0.64 (0.40)	0.06 (0.73)	0.50 (0.37)
Contiguity	0.72*** (0.18)	0.93*** (0.21)	0.81† (0.45)	0.83*** (0.19)
Territory (Issue)	0.95*** (0.25)	1.23*** (0.32)	0.45 (0.44)	0.97*** (0.24)
Government (Issue)	0.21 (0.49)	0.19 (0.53)	0.85 (1.49)	0.25 (0.55)
Policy (Issue)	-0.97*** (0.20)	-0.77** (0.27)	-1.47*** (0.39)	-0.96*** (0.20)

Other (Issue)	-0.88 (0.67)	-0.07 (0.87)		-0.75 (0.77)
Constant	-1.23*** (0.31)	-1.52*** (0.41)	-0.80 (0.67)	-1.21*** (0.31)
<i>N</i>	539	369	170	539

Standard errors in parentheses. 'Other' dropped from dem target model due to lack of positive cases.

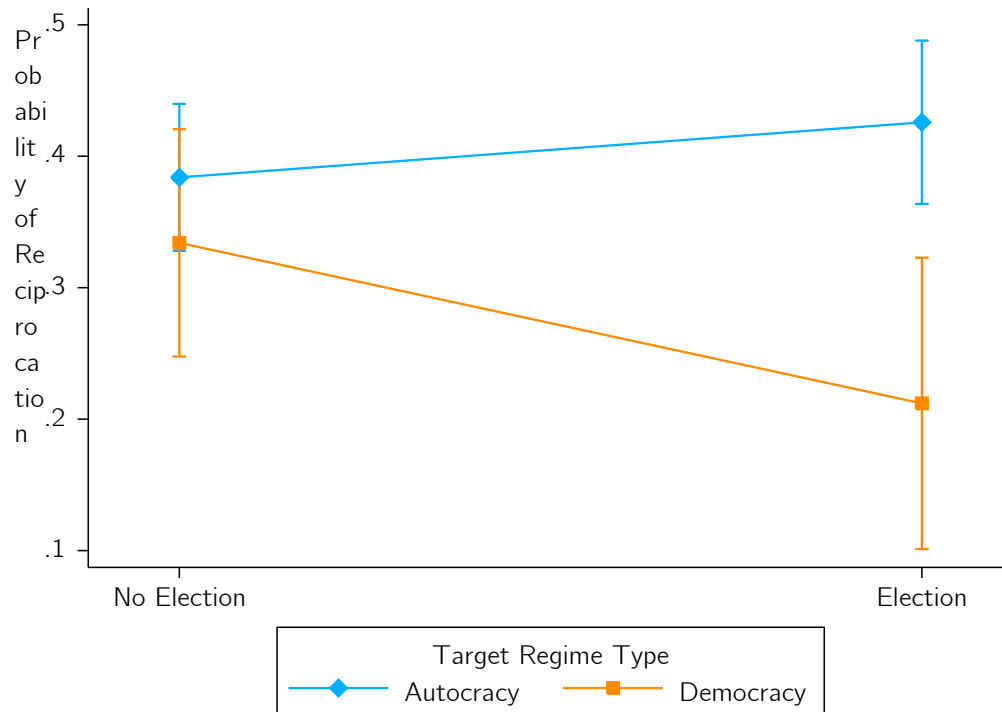
†  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Because logistic regression coefficients are difficult to interpret, I also calculate the marginal effect of a one unit change in the regressors on the probability of reciprocation, holding covariates at their means. Figure 1 graphically displays the results. I estimate that an election occurring in the same year as the dispute reduces the probability that a democratic target will reciprocate by 18 per cent. Among authoritarian targets, the point estimate approximates zero. The results also indicate that the substantive effect of election timing for democratic targets compare favorably with the effect sizes for the control variables. Among democratic targets, contiguity with the initiator increases the probability of a militarized response by 16 per cent, while threats made with an intent to influence the policies of the target state are 30 per cent less likely to be reciprocated. For autocratic targets, contiguity increases the probability of reciprocation by 22 per cent while policy influencing threats are 18 per cent less likely to be met with a militarized response. Interestingly, threats made over territory are 29 per cent more likely to be reciprocated by autocrats, even though this variable was insignificant among the democratic subset.



**Figure 1: Marginal effect of election-year and control variables on reciprocation by target regime type with 95 per cent confidence intervals**

As the effect of the control variables did not dramatically differ across democratic and autocratic targets, I also estimate a model with an interaction between the election-year dummy and target regime type. Column four of Table 2 presents the regression coefficients, while Figure 2 displays the results graphically. For autocratic targets, there is little difference between election-year and non-election year threats. Democratic targets, however, are less likely to respond in the former (21 vs. 33 per cent,  $p < .077$ ). Furthermore, there is little difference in the probability that democracies and autocracies respond to non-election year threats, but a significant difference between the two for election years (21 vs. 43 per cent,  $p < .001$ ).



**Figure 2: Marginal effect of election-year on reciprocity in interactive model with 95 per cent confidence intervals**

## Addressing Selection Effects

Leaders have an incentive to avoid situations in which audience costs can directly be observed; thus, scholars such as Schultz (2001) and Weeks (2008) attempted to circumvent such selection bias by examining dispute reciprocation. Across a large number of disputes, states more capable of generating audience costs should be more successful. While the present design leverages this insight, it raises two new concerns with selection effects. First, disputes made in election years may be systematically different than those made in non-election years; this may be the case if audience cost-conscious leaders select disputes more likely to succeed in election years. Second, disputes made against democracies may be different than those made against autocracies; this may arise due to fundamental differences in the foreign relations that democracies may have with these two sets of states, or because democratic publics are less tolerant of hostile disputes with other democracies (Tomz and Weeks 2013).

To assess the severity of potential selection bias, I compare election-year and non-election year disputes in their ratio of democratic targets and in their average hostility levels. If election-year threats systematically differ from non-election year threats on these two attributes, it suggests the possibility that apparent differences in reciprocation may be driven more by unobserved differences in the types of threats issued when elections are proximate versus when they are distant. I find that 32.6 per cent of threats in non-election years are made against democracies, while 30.2 per cent of election-year threats are against democratic targets. I also find that the average hostility level of an election-year dispute is 3.53 on a scale ranging from 1 to 5, while the average hostility level of a dispute made in non-election years is 3.55. Finally, I find that the average hostility level of the target's response to an election-year threat is 2.02, while the average hostility of the target's response to threats made in non-election years is 1.99. None of these differences are statistically significant.

To address the second possible selection bias—are disputes made against democracies different than those made against autocracies?—I compare these two sets of disputes on the



same attributes as above. I find that 23.2 per cent of threats made against democracies are in election-years, while 25.3 per cent of threats made against autocracies are made proximate to elections. I also find that the average hostility of a dispute between democracies is 3.52, while the average hostility of a democratic-autocratic dispute is 3.55. Finally, the average hostility of a democratic target's response to all threats is 1.87, while the average hostility of an autocrat's response to all threats is 2.06. Once again, none of these differences are statistically significant. Overall, these series of placebo tests fail to find evidence suggesting systematic differences in the types of threats that democratic initiators make in election years, or against democratic targets, provide a plausible alternative explanation for the observed results.

Finally, if my theory is correct, we may expect that the average hostility of a democratic target's response during an initiator's election year is lower than the hostility of an autocratic target's response during elections. This is because if democracies are generally disinclined to reciprocate election-year threats because they recognize those threats as costly and credible, then when they do choose to respond they should also be more cautious in order to avoid escalation into conflict. Comparing these two values, I find that the average hostility level of a democratic target's response during an initiator's election year is 1.76, and the average hostility of an autocratic target's response during an initiator's election year is 2.14 ( $p = 0.06$ ). The fact that this auxiliary prediction of the theory also appears to be consistent with the data should strengthen confidence in the above interpretation of the findings.

## **Robustness Checks**

I assess the robustness of the above findings in a number of ways. First, I attempt to re-run the analyses using Todd Sechser's Militarized Compellent Threats (MCT) dataset (Sechser 2011). Downes and Sechser (2012) criticize the MID dataset for including disputes which do not meet Schelling (1960)'s definition of a 'compellent threat' and which, the authors conclude, are inappropriate for testing theories of coercive diplomacy. One drawback to the MCT dataset,

however, is the comparably smaller number of cases included as a result of the stricter coding rules. After dropping those compellent threats made by nondemocracies, I am left with 56 observations. Appendix 2.1 replicates Table 1 in this paper using MCT data.<sup>12</sup> Unfortunately, there is only one case of a dispute made against a democracy in the initiator's election year,<sup>13</sup> rendering regression analyses intractable due to lack of variation. It is encouraging, however, that election-year threats are met with greater compliance than those made in non-election years (50 vs. 41 per cent among the full sample)—results that are generally consistent with above findings.

Additionally, I conduct a series of tests to assess the robustness of my results from the MID dataset. First, I re-estimate the models presented in Columns one through three of Table 2 using unadjusted standard errors. Second, I use an alternative dichotomous measure of reciprocation following Schultz (2001, 263), which takes on a positive value if and only if the target state responds with a militarized action, including the threat, display, and actual use of force.<sup>14</sup> Third, I drop from analyses those disputes for which I manually input the target regime type<sup>15</sup>. Fourth, again following Schultz (2001), I drop those disputes that were part of the World Wars. Fifth, I add measures of S-weighted<sup>15</sup> alliance portfolio similarity as well as each state's status quo evaluations (measured as their alliance portfolio similarity to the leading state in the system at the time) as additional control variables.

The results of these robustness checks are presented in Appendix 2.2. Across all specifications, the election year variable remains significant for democratic targets but insignificant among autocratic targets and for the full sample, although when using unadjusted standard errors or the alternative reciprocation measure election year is only significant at  $p < 0.10$  confidence level. Similarly, the effects of the control variables generally remain across the

---

<sup>12</sup> Downes and Sechser (2012) suggest that the reciprocation measure from the MID dataset does not indicate with certainty that the initial demand was rejected, instead preferring a measure of whether the target responds with full, partial, or non-compliance. Of the 56 disputes, five were met with partial compliance and 19 were met with full compliance; I collapse these two categories to create a dichotomous measure of compliance.

<sup>13</sup> South Africa against Lesotho in 1994

<sup>14</sup> The standard reciprocation measure takes on a positive value if the target state takes any "codeable action" (Jones, Bremer and Singer 1996, 186). <sup>15</sup> See footnote 4.

<sup>15</sup> See Signorino and Ritter (1999).

models. Contiguity and disputes over territorial claims are generally more likely to be met with reciprocation, while responding is less likely when the threat is made over policy preferences in the target state.

There are several other potential concerns to the validity of the results. First, it is likely the case that democratic leaders are strategic in selecting from the universe of potential disputes those that they want to elevate into actual confrontations (Gelpi and Griesdorf 2001). One may be concerned that this selection mechanism is correlated with the key independent variable of election years, which would be the case if leaders seek to initiate disputes as a diversion from domestic discontent when elections are proximate (Levy 1989; Smith 1996). Note, however, that election-year disputes do not appear to be drastically rarer or more frequent than election years. Furthermore, the diversionary logic alone poorly explains the observed patterns. Insofar as it suggests the escalation of disputes that, absent diversionary incentives, may well be left alone, such threats should be even less credible and more likely to be met with a response. Threats made for diversionary purposes in election years, in other words, should bias against observing the negative relationship between elections and reciprocation. A slightly different taken on this diversionary alternative explanation is that leaders may be more willing to make such threats against rivals (which, for democratic initiators, are likely to be autocracies), while threats against democracies will be made only when the initiator is sincere. While this is a plausible explanation, the empirical findings observed in chapter four suggest that leaders of states with free media (most of which are democracies) are particularly unlikely to engage in foreign conflict during times of domestic turmoil. It seems similarly likely that democratic leaders in particular will wish to avoid issuing diversionary threats during election years.

Second, one may be concerned about reverse causality. It may be the case that leaders of parliamentary democracies who succeed in disputes decide to call early elections in an attempt to benefit from approval boosts following the foreign policy win. Alternatively, parliamentary leaders may face no-confidence motions after a foreign policy defeat that then precipitates an election. However, both of these explanations cannot explain why early elections are called only

against democracies but not autocracies. In particular, because democratic publics view autocratic adversaries as more threatening than similar democracies (Tomz and Weeks 2013), the potential rally effects should be greater after foreign policy victories against autocracies. In other words, leaders who wish to call early elections after winning a dispute should be *more likely* to do so after winning against autocrats, not less.

Third, one may wonder whether the observed results are affected by leader tenure. In particular, it may be the case that in fixed term systems, leaders in their last term are less beholden to the public. If true, such leaders should be less effective at generating audience costs, with the implication again being a bias against observing greater effectiveness for election-year threats.

Fourth, the present study only considers elections among initiators; it does not consider the electoral cycle of target states. In cases where the democratic initiator issues a threat against a democratic target during the latter's election year, the target has additional incentives to not back down and appear weak vis-à-vis the domestic public. Thus, we should expect the target during their election year to be more likely to reciprocate—again, with the implication of a bias against observing greater effectiveness (i.e. no target reciprocation) for threats made in an initiator's election year.

## Discussion

In sum, I fail to find evidence that threats made in a democratic initiator's election years are more effective than those made in non-election years, both for the full sample and against autocratic targets. However, I do find evidence that such threats are significantly more effective against democratic targets. This empirical pattern is consistent with democratic targets (1) ascertaining that election-year threats carry greater audience costs, and thus are more credible; and (2) being better able to recognize these costs (or accord greater weight and value to such costs) than their autocratic counterparts.

To be sure, there are limits to the present findings. The study uses elections as an observable measure of the level of audience costs associated with particular threats. Because election-year is a dichotomous measure—there either is or is not an election in the same year as the threat—the results should also be interpreted in light of a dichotomous, coarse measure of costliness. That is to say, one should interpret the empirical findings as evidence that democratic targets are able to adjudicate between high- vs. low-cost threats. Additionally, the use of elections as a proxy for costliness implies that reciprocation here should be interpreted as the response to *threats made costly by elections specifically*—in other words, a *local* average treatment effect. Thus, it sheds little light on costliness induced by other means.

Elections are often seen, both by specialists and among the public, as a *sine qua non* of democracy, while those that exist under dictatorship are generally viewed as instruments to advance authoritarian control (Schedler 2002; Gandhi and Lust-Okar 2009), not mechanisms by which domestic audiences can punish unpopular leaders. Autocrats may be skeptical of or unable to recognize democratic elections as bolstering the costliness of interstate signaling, but that may not mean that they abstain from the logic of audience cost entirely—instead, instances of mass public protests and demonstrations may serve for dictators a similar hand-tying function as electoral institutions do for democrats (Weiss 2014). Other authoritarian regimes may exploit audience cost dynamics through features such as factions within the ruling elite, or scheduled and institutionalized succession processes.

At the same time, to the extent that election timing is simply exploited to generate variation in the costs associated with threats, one should be cautious in generalizing or drawing broad conclusions about elections or electoral politics specifically, at least from the present results. To the extent that certain other autocracies may face election-like pressures at varying points and to varying degrees (i.e. leader transitions), future research may attempt to replicate the present findings by using a different institutional mechanism for generating variation in costs. For example, a replication or extension may look at a subset of autocratic initiators, and examine

whether costliness induced by domestic unrest and unpopularity are also less likely to be met with reciprocation by democratic targets.

Additionally, the study design follows previous observational studies of audience cost by examining reciprocation rates of threats made—in other words, it examines the perceptions of costs and threat credibility from the perspective of the target. Because we cannot observe threats that are not made, the results speak to how threat credibility is received and perceived by the target, but it does little to say whether the state who issued the threat was doing so with awareness of audience cost dynamics. Thus, the present findings seem unlikely to fully appease skeptics of audience cost theory. On the other hand, a second potential limitation may be more adequately addressed. Because we cannot observe threats not made, we do not know the criteria used by leaders to decide whether or not to issue a threat, or to what extent that criteria is correlated with either election timing or regime type of target. When attempting to address selection effects, I use the observed data to show that (1) the ratio of democratic and autocratic targets are not different between election-year and non-election year threats; (2) election-year threats are not more hostile; (3) ratio of election and non-election threats do not differ between democratic and autocratic targets; (4) average hostility between democratic dyad disputes is not different than hostility between democratic-autocratic dyads; and (5) average hostility of responses by democracies are not different than average hostility of responses by autocracies. If the criteria used by leaders to issue or not issue a threat is correlated with either the electoral cycle or with target regime type, we should be able to observe differences among issued threats along those two dimensions.

Limitations aside, the present study suggests mixed but theoretically interesting implications for audience cost theory. On the one hand, there is little evidence that autocrats, who are the targets for most threats issued by democracies, recognize and operate according to the logic of audience costs. To that extent, the study presents quantitative evidence to support the conclusions that critics of the theory such as Trachtenberg (2012) and Snyder and Borghard (2011) reach using qualitative case studies of democratic–authoritarian crises. On the other hand,

that democracies appear to respond to threats in a manner consistent with audience cost theory suggest caution in throwing the baby out with the bathwater, while pointing to a continuing need for scholars to be cognizant of how decisions over research scope—intentional or otherwise—impacts the conclusions obtained.

Perhaps the most interesting implication of the present findings concerns the relationship between audience costs and the democratic peace. Since the bivariate relationship between democracy and peace was first uncovered by Babst (1964), one of the most important qualifications has been the dyadic as opposed to monadic nature of the effect—although democracies do not fight each other, they are just as belligerent as autocracies when facing non-democratic adversaries (Maoz and Abdolali 1989). Gartzke (2007) notes that, if peace arises purely because of the stronger domestic audiences of democracies and consequently their superior signaling ability in interstate bargaining processes, the prediction should be one of monadic as opposed to dyadic peace. Insofar as the present study suggests that the effectiveness of costly signaling is conditional on the regime type of the signal's intended target, it strengthens the audience cost explanation. If autocrats systematically fail to recognize costly signals as carrying potential audience costs, then they may be more likely to underestimate the level of resolve in democracies. Democratic dyads, in contrast, will be better able to reach a bargaining settlement prior to outbreak of war. Both *sending* and *receiving* of credible signals have been central to theories of crisis bargaining—misperception of such signals that correlates with regime type may be one explanation for why democracies do not fight other democracies.

## Conclusion

A key debate surrounding audience cost theory focuses on whether targets of costly signaling interpret such signals as carrying hypothesized audience costs (Snyder and Borghard 2011; Trachtenberg 2012). I argue that if we assume leaders do indeed recognize audience costs, then one observable implication is that leaders will respond differently to high-cost threats than to those that are comparatively low-cost. The paper makes two contributions to the audience

cost literature. First, I leverage the presence of elections in democracies to elicit observable variation in the costliness of threats, under the assumption that threats made in election years are more costly than those made in non-election years. Second, while previous quantitative work examine institutional variation among initiators, I explore regime type differences among the *targets* of threats. I find no evidence that autocrats differentiate between election-year and non-election year threats, but consistent and robust evidence of this adjudicate by democracies. The results are consistent with the possibility that democracies are better able than autocracies to recognize audience costs associated with interstate signaling.

Thus far, I have suggested that the failure to find evidence of autocrats placing weight on election timing implies that such leaders either do not understand or do not value audience costs. Future work may wish to adjudicate between these two possibilities. Given that leaders of all states have strong incentives to accurately diagnose the capability and resolve of potential adversaries, the possibility that autocrats are somehow unable to follow the logic of punitive domestic audiences seems limited. More likely is the possibility that autocrats discount public costly signals under the belief that such claims are mere bluffs, and that political leaders even in democracies should be able to shape public opinion rather than be constrained by them. Counterintuitively, then, what may make public signals by democracies to autocracies more credible is private assurances regarding the legitimacy of the audience costs associated with such threats (see also Yarhi-Milo 2013). While the present findings address the question of whether targets understand audience costs, a number of critiques that have been levied against the theory remain standing. For example, Trachtenberg (2012) suggest that leaders rarely choose to make public threats as part of an intentional hand-tying strategy. Given the inherent difficulties in observing leader intentions directly, future studies may attempt a similar approach as the present article by looking for second-order observable implications. For example, if initiators do intentionally leverage audience costs for credibility in interstate signaling, they may be more likely to make such signals during periods of high potential audience costs, such as when elections are temporally proximate. Though one would need to disentangle the audience cost explanation from



that of diversionary war—leaders may make more threats when elections are close in order to benefit from rally 'round the flag effects, rather than to take advantage of enhanced credibility of interstate signaling—addressing and testing the various critiques that have been levied against audience cost theory will surely continue to occupy the attention of scholars.

## References

- Babst, Dean. 1964. "Elective Governments—A Force for Peace." *Wisconsin Sociologist* 3(1):9–14.
- Baum, Matthew A. and Philip B. K. Potter. 2015. *War and Democratic Constraint: How the Public Influences Foreign Policy*. Princeton University Press.
- Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer and Patrick Walsh. 2001. "New Tools in Comparative Political Economy: The Database of Political Institutions." *World Bank Economic Review* 15(1):165–176.
- Bueno de Mesquita, Bruce, James D. Morrow, Randolph M. Siverson and Alastair Smith. 2003. *The Logic of Political Survival*. Cambridge: MIT Press.
- Chaudoin, Stephen. 2014. "Policies or Promises? An Experimental Analysis of International Agreements and Audience Reactions." *International Organization* 68(1):235–256.
- Chiozza, Giacomo. 2017. "Presidents on the Cycle: Elections, Audience Costs, and Coercive Diplomacy." *Conflict Management and Peace Science* 34(1):3–26.
- Croco, Sarah E. 2011. "The Decider's Dilemma: Leader Culpability, War Outcomes, and Domestic Punishment." *American Political Science Review* 105(3):457–477.
- Downes, Alexander B. and Todd S. Sechser. 2012. "The Illusion of Democratic Credibility." *International Organization* 66(3):457–489.
- Eyerman, Joe and Robert A. Hart. 1996. "An Empirical Test of the Audience Cost Proposition." *Journal of Conflict Resolution* 40(4):597–616.
- Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88(3):577–592.
- Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.
- Gandhi, Jennifer and Ellen Lust-Okar. 2009. "Elections Under Authoritarianism." *Annual Review of Political Science* 12:403–422.

- Gartzke, Erik. 2007. "The Capitalist Peace." *American Journal of Political Science* 51(1):166–191.
- Gartzke, Erik and Yonatan Lupu. 2012. "Still Looking for Audience Costs." *Security Studies* 21(3):391–397.
- Gaubatz, Kurt Taylor. 1991. "Election Cycles and War." *Journal of Conflict Resolution* 35(2):212–244.
- Gelpi, Christopher and Joseph M. Grieco. 2015. "Competency Costs in Foreign Affairs: Presidential Performance in International Conflicts and Domestic Legislative Success, 1953–2001." *American Journal of Political Science* 59(2):440–456.
- Gelpi, Christopher and Michael Griesdorf. 2001. "Winners or Losers? Democracies in International Crises, 1918–94." *American Political Science Review* 95(3):633–647.
- Ghosn, Faten, Glenn Palmer and Stuart A. Bremer. 2004. "The MID3 Data Set, 1993–2001: Procedures, Coding Rules, and Description." *Conflict Management and Peace Science* 21(2):133–154.
- Jones, Daniel M., Stuart A. Bremer and J. David Singer. 1996. "Militarized Interstate Disputes, 1816–1992: Rationale, Coding Rules, and Empirical Patterns." *Conflict Management and Peace Science* 15(2):163–215.
- Kertzer, Joshua and Ryan Brutger. 2016. "Decomposing Audience Costs: Bringing the Audience Back Into Audience Cost Theory." *American Journal of Political Science* 60(1):234–249.
- Kunda, Ziva. 1990. "The Case for Motivated Reasoning." *Psychological Bulletin* 108(3):480–493.
- Kurizaki, Shuhei and Taehee Whang. 2015. "Detecting Audience Costs in International Disputes." *International Organization* 69(4):949–980.
- Levendusky, Matthew S. and Michael C. Horowitz. 2012. "When Backing Down is the Right Decision: Partisanship, New Information, and Audience Costs." *Journal of Politics* 74(2):323–338.
- Levy, Jack S. 1989. Diversionary Theory of War: A Critique. In *Handbook of War Studies*, ed. Manus I. Midlarsky. Routledge pp. 259–288.

- Levy, Jack S., Michael K. McKoy, Paul Poast and Geoffrey P. R. Wallace. 2015. "Backing Out or Backing In? Commitment and Consistency in Audience Costs Theory." *American Journal of Political Science* 59(4):988–1001.
- Lodge, Milton and Charles Taber. 2000. Three Steps Toward a Theory of Motivated Political Reasoning. In *Elements of Reason*, ed. Arthur Lupia, Matthew D. McCubbins and Samuel L. Popkin. Cambridge University Press.
- Lodge, Milton and Charles Taber. 2005. "The Automaticity of Affect for Political Leaders, Groups, and Issues: An Experimental Test of the Hot Cognition Hypothesis." *Political Psychology* 26(3):455–482.
- Maoz, Zeev and Nasrin Abdolali. 1989. "Regime Types and International Conflict, 1816-1976." *Journal of Conflict Resolution* 33(1):3–35.
- Marshall, Monty G., Ted Robert Gurr and Keith Jagers. 2014. "POLITY IV Project: Political Regime Characteristics and Transitions, 1800-2013: Dataset Users' Manual." <http://www.systemicpeace.org/inscr/p4manualv2013.pdf>.
- Mueller, John. 1973. *War, Presidents, and Public Opinion*. New York: Wiley & Sons.
- Partell, Peter J. and Glenn Palmer. 1999. "Audience Costs and Interstate Crises: An Empirical Assessment of Fearon's Model of Dispute Outcomes." *International Studies Quarterly* 43(2):389–405.
- Quek, Kai. Forthcoming. "Type II Audience Costs." *Journal of Politics*.
- Rhode, David and Dennis Simon. 1985. "Presidential Vetoes and Congressional Response: A Study of Institutional Conflict." *American Journal of Political Science* 29:397–427.
- Schedler, Andreas. 2002. "Elections Without Democracy: The Menu of Manipulation." *Journal of Democracy* 13(2):36–50.
- Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge: Harvard University Press.
- Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy*. New York: Cambridge University Press.
- Schultz, Kenneth A. 2012. "Why We Needed Audience Costs and What We Need Now."

- Security Studies* 21(3):369–375.
- Sechser, Todd S. 2011. "Militarized Compellent Threats, 1918-2001." *Conflict Management and Peace Science* 28(4):377–401.
- Signorino, Curtis S. and Jeffrey M. Ritter. 1999. "Tau-b or Not Tau-b: Measuring the Similarity of Foreign Policy Positions." *International Studies Quarterly* 43(1):115–144.
- Singer, David J. 1987. "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816-1985." *International Interactions* 14:115–132.
- Singer, David J., Stuart Bremer and John Stuckey. 1972. Capability Distribution, Uncertainty, and Major Power War, 1820-1965. In *Peace, War, and Numbers*, ed. Bruce Russett. Sage.
- Smith, Alastair. 1996. "Diversionary Foreign Policy in Democratic Systems." *International Studies Quarterly* 40(1):133–153.
- Snyder, Jack and Erica D. Borghard. 2011. "The Cost of Empty Threats: A Penny, Not a Pound." *American Political Science Review* 105(3):437–456.
- Stinnett, Douglas M., Jaroslav Tir, Philip Schafer, Paul F. Diehl and Charles Gochman. 2002. "The Correlates of War Project Direct Contiguity Data, Version 3." *Conflict Management and Peace Science* 19(2):58–66.
- Tomz, Michael. 2007. "Domestic Audience Costs in International Relations: An Experimental Approach." *International Organization* 61(4):821–840.
- Tomz, Michael and Jessica L. P. Weeks. 2013. "Public Opinion and the Democratic Peace." *American Political Science Review* 107(4):849–865.
- Trachtenberg, Marc. 2012. "Audience Costs: An Historical Analysis." *Security Studies* 21(1):3–42.
- Trager, Robert F. and Lynn Vavreck. 2011. "The Political Costs of Crisis Bargaining: Presidential Rhetoric and the Role of Party." *American Journal of Political Science* 55(3):526– 545.
- Weeks, Jessica P. 2008. "Autocratic Audience Costs: Regime Type and Signaling Resolve." *International Organization* 62(1):35–64.

- Weiss, Jessica C. 2014. *Powerful Patriots: Nationalist Protest in China's Foreign Relations*. New York: Oxford University Press.
- Yarhi-Milo, Keren. 2013. "Tying Hands Behind Closed Doors: The Logic and Practice of Secret Reassurance." *Security Studies* 22(3):405–435.
- Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.

### **CHAPTER 3: Do Publics Reward Belligerence? Rally 'Round the Flag and Diversionary War**

On the eve of 9/11, George W. Bush was not a particularly popular president. A bruising battle to pass sweeping tax reductions through Congress left him at around 50 per cent approval—squarely in the middle of the pack compared with his predecessors in office. Just one week later, a staggering 90 per cent of the American people placed their support behind the president.

Bush after 9/11 is the most dramatic example of the 'rally 'round the flag' effect: the phenomenon of publics raising their support of a country's leader following a dramatic foreign confrontation. First documented systematically in Mueller (1973)'s study of the Korean and Vietnam wars, the rally literature since has focused on understanding the conditions under which large, pervasive, and enduring rallies are found.

At the same time, whether and how initiation status affects the rally phenomenon remains an open question. Do publics reward belligerent leaders initiating conflict as much as they do for leaders responding to a foreign adversary? The question is significant due to the rally literature's connection to IR scholarship on diversionary war, which suggests that leaders have an incentive to engage in aggressive foreign policies when facing domestic discontent. Crucially, this theory revolves around an assumption that such belligerent leaders can nevertheless benefit from a rally effect (Levy 1989). Thus, showing whether publics do in fact differentiate between leaders who instigate and those who respond to foreign conflicts has significant implication for diversionary theory specifically as well as the relationships between public opinion and foreign policy broadly. I address this question empirically through a survey experiment. One key challenge facing previous studies is the difficulty in disentangling initiation status from all other salient attributes of rally events. An experimental approach is appropriate in this case as researcher manipulation of initiation status in a survey scenario allows one to credibly identify the causal effect of initiation

versus reciprocation on the subsequent rallies. To the best of my knowledge, this is the first experimental assessment of rally effects in a non-terrorism context.

The results suggest heterogeneous rallies across partisanship. For the sample as a whole, I find no evidence that individuals rally when President Trump is initiating conflict, though a substantial rally effect is observed when Trump is responding to foreign aggression. However, this is driven largely by the Democratic skew in the sample. The aggregate finding largely holds true among self-identified Democrats, and there is suggestive evidence that they may even punish Trump when he is seen as the initiator. Self-identified Republicans, on the other hand, rally behind Trump regardless of whether or not Trump initiates or reciprocates conflict. Additionally, I find substantially larger rally effects among Republicans than Democrat—a noteworthy contrast to prior work finding the largest rallies to be produced by out-partisans (Baum 2002; Hetherington and Nelson 2003).

The paper proceeds as follows. The following section briefly reviews existing findings concerning the rally phenomenon, including prior work that addresses the role of initiation status. Next, I introduce the research design. The section after presents the findings of the survey experiment. I conclude by discussing the implications of the findings for both rally effects and diversionary war.

## **The rally ‘round the flag effect**

The claim that statesmen may reap benefits at home from conflicts abroad permeates historical accounts of war and diplomacy. One of the most well-known examples is the Russian interior minister V. K. Plehve, who on the eve of his country’s 1904 conflict against Japan saw value in “a little victorious war to stem the tide of revolution” (quoted in Blainey 1973). Though Russia’s imperialist ambitions in China may have been the primary motivation for the conflict, and while popular support in Russia quickly turned to popular discontent upon suffering a string of unexpected military defeats at the hands of an Asian power, Plehve’s statement comes



close to smoking gun evidence that political leaders may sometimes anticipate domestic dividends from foreign conflicts.

One of the earliest systematic studies of what would become known as the ‘rally ‘round the flag effect’ is by John Mueller. Examining an expansive number of US public opinion surveys, Mueller finds consistent evidence that the mass public’s approval of the US president increases following what he terms ‘rally points’ (1970; 1973). These points—events that trigger rally effects—are operationalized as consisting of three necessary criteria: (1) it is international in nature; (2) it involves the US and the president; and (3) is specific, dramatic, and sharply focused. Mueller also attributes the cause of rally effects to a “desire to support a country’s leadership in time of trouble” (Mueller 1973).

Since then, a number of scholars have extended the study of rally effects to countries other than the United States. Lai and Reiter (2005) analyze 50 years of public opinion data in the UK, with results that are largely consistent with prior findings in the American context. In contrast, Sprecher and DeRouen (2002) find a ‘reverse’ rally effect in Israel, with uses of force against Arab military actions producing more domestic protests. Recent work has also looked at rally dynamics in authoritarian contexts. Hou and Quek (2015) find that Chinese subjects randomly assigned to read reports on a terrorist attack become more nationalistic and supportive of policies that infringe on civil liberties. However, Yue (2016) finds little evidence that the Chinese public rallies when exploiting an overlap between the fielding of a nationally representative survey and the exogenous timing of a foreign policy crisis, though he was only able to obtain measures for support of local government.

In addition to questions over how well rallies can be observed in non-American contexts, this literature has identified at least four limitations to rally effects, some of which continue to be controversial. First, a number of studies find that most rallies following a foreign crisis are extremely small. Using monthly data on US presidential approval and crisis involvement from 1953 to 1994, James and Rioux (1998) find that crises produce approval bumps that average only 3 to 4 per cent. Furthermore, Lian and Oneal (1993) and Oneal and Bryan (1995), using two

different datasets on foreign conflicts, estimate the average rally effect to be 0 and 1.4 per cent respectively.<sup>16</sup>

Second and closely related to concerns over average rally sizes are questions regarding the duration of rally effects. Examining rallies by the British public, Lai and Reiter (2005) estimate that these effects are concentrated in the first two months. On average, the authors find that increases in support decline by more than half after the first month, and all but disappears by the third. On the other hand, Hetherington and Nelson (2003) show that rallies for Kennedy following the Cuban Missile Crisis and for George H. W. Bush after the Gulf War lasted for six to eight months, while George W. Bush's rally following the 9/11 attacks can still be seen over a year later. Furthermore, they problematize Lai and Reiter's study—which measures support as vote intention for the ruling party—by suggesting that rally effects are generally personal to leaders and have diminished coattail effects on in-party Congressional candidates.<sup>17</sup>

A third qualification to rallies identified by the literature is the effect's heterogeneity across different demographic groups. Mueller (1973) finds that rallies decay more quickly for out-partisans than for those who identify with the party of the president (see also Kam and Ramos 2008). However, quicker decay of rally effects among out-partisans can also be consistent with a larger initial rally among out-partisans (Baum 2002; Hetherington and Nelson 2003). Thus, the initial surge in leader approval following a foreign crisis is more pronounced among out-partisans due to ceiling effects among the leader's base, but, perhaps unsurprisingly, this bump among out-partisans also disappears sooner.<sup>3</sup> In addition to partisan differences, Baum (2002) also find that rally effects are strongest among those individuals who are ambivalent in terms of approval of the president and those who are moderately attentive to politics. Finally, rally sizes are also

---

<sup>16</sup> One implication is that studies of rally effects can overestimate effect sizes if authors select only prominent examples of foreign policy crises. A second implication is that, if average rallies approximate zero but some crises experience large positive rallies, then at least in some instances conflicts can also produce 'negative rallies'—a finding consistent with both Sprecher and DeRouen (2002) as well as the experiences of Interior Minister Plehve (see also Brody and Shapiro 1991)

<sup>17</sup> While true that voters in parliamentary systems are only able to indirectly elect executives by casting a vote for members of particular parties, it still may be the case that support for a political party is a weaker measure of rally effects than a survey item assessing leader approval directly. <sup>3</sup>However, see James and Rioux (1998) for contrary finding.

influenced by the interaction of journalistic incentives and elite rhetoric in the aftermath of the foreign crisis (Groeling and Baum 2008).

One final limitation to rally effects concerns how these vary in accordance with conflict intensity and level of escalation. Given that high intensity conflicts are more likely to dominate the news cycle, the finding that crises that enjoy high and prominent media coverage (covered on the front pages of the New York Times, for example) experience larger rally effects than less prominent disputes suggests that rally size is positively associated with conflict intensity (Lian and Oneal 1993; Oneal and Bryan 1995). Baker and Oneal (2001) makes this claim the most forcefully, arguing that rallies hold for war but not actions short of war such as threats and displays of force. Similarly, Lai and Reiter observe that “lesser international episodes which do not receive substantial media coverage and/or are not seen to threaten British interests directly do not cause rallies” (2005, 264). The one exception is James and Rioux (1998) who posits a non-monotonic relationship between rally size and conflict escalation; the authors suggest that rallies increase as leader actions escalate, but the outright use of force decreases rally size.

### **Conflict initiation vs. conflict reciprocation**

Mueller (1973)’s tripartite definition of a ‘rally event’ does not differentiate events triggered by the president and those in which the president merely responds to a crisis initiated by the other side. Much of the literature since then proceeds in a similar vein, with units of analyses often being all militarized disputes or crises involving a country of interest, regardless of how each dispute is initiated.

Examining British public opinion, Lai and Reiter (2005) observe that rally effects were larger during the Falkland War than in the Gulf War. This leads the authors to speculate that rallies are larger when a country is the target of a crisis rather than the initiator. Such an explanation has intuitive appeal insofar as some dovish subset of a population may punish leaders whose actions are perceived to be belligerent (Kertzer and Brutger 2016).

Furthermore, Lai and Reiter suggest that when a state is the target in a dispute, it is easier for a leader to justify militarized action as necessary on national security grounds.

On the other hand, a number of studies demonstrate that George H. W. Bush benefitted from rallies as a result of the Gulf War, despite US initiation of that conflict (Parker 1995; Mueller 1994; Zaller 1993, see also Iyengar and Simon 1993; Krosnick and Brannon 1993). Considering the most prominent examples of rally events in US history sheds no further light. After the Pearl Harbor attacks, Roosevelt experienced a personal bump in approval and popularity in addition to the well-known surge in public support for US participation in World War II. Similarly, George W. Bush's rally following the 9/11 attacks was the largest rally ever received by a US president (Hetherington and Nelson 2003). On the other hand, George H. W. Bush's rally for the Gulf War and Lyndon Johnson's rally when the US initially deepened its involvement in Vietnam are both cases where publics rallied to leaders who ostensibly are the initiators of conflict abroad. Furthermore, Lai and Reiter (2005)'s argument that Presidents can better justify militarized action when responding to conflicts seems to overlook the information advantage that leaders have vis-à-vis the public and mass media at the start of a conflict—and their ability to use this information advantage to justify military action regardless of the conflict's context (Baum and Groeling 2008; Canes-Wrone 2005; Levendusky and Horowitz 2012). Perhaps the greatest challenge for answering this question, however, lies in the differences between conflicts. After all, events as disparate as 9/11, Vietnam, Pearl Harbor, and the Gulf War involve a multiplicity of factors in addition to the US's status as initiator or target that may affect subsequent rallies; given the finite number of conflicts in history it is difficult to disentangle the effect of initiation status from these other factors in observational data.

Why should we care whether conflict initiation affects a public's propensity to rally around the flag? The rally phenomenon is closely related to the theory of diversionary war. According to this theory, leaders have an incentive to engage in foreign aggression when facing domestic discontent at home. Rally effects provide the reason and justification for foreign aggression—by rallying a domestic population around a foreign scapegoat, a leader may wish to

divert attention away from economic and social woes at home (Chiozza and Goemans 2003; Levy 1989; Smith 1996). Diversionary war has most famously been used to explain Argentine initiation of the Falklands (Oakes 2006, but see Fravel 2010), various European conflicts of the late 19th and early 20th century (Blainey 1973; Fischer 1975; Mayer 1969; Levy 1989; Taylor 1954), and even the timing of military operations in the Gulf War (Tyler 1992). What has been underappreciated thus far—and by IR scholars in particular—is that whether rallies hold not only when leaders respond to international conflicts but also when leaders initiate conflict themselves is a question that remains unsettled.

It is worth noting that if publics are indeed less willing to rally to leaders who are initiating rather than responding to conflict, that is not necessarily a death knell to diversionary theory. The assumption is that diverting leaders who initiate conflict operate under the *expectation* of domestic dividends afterwards; it does not predict that leaders will *necessarily reap* such benefits.<sup>18</sup> That individuals may operate in accordance to felt beliefs as opposed to demonstrable facts should hardly be a surprise. However, ascertaining whether or not the mass public differentiates between initiation and reciprocation when rallying sheds light not only on an open question in the public opinion and foreign policy literature, but may also suggest new avenues of research in IR. If publics are less likely to rally when initiating, perhaps experienced leaders may be less prone to diverting behavior as a result of 'on the job' learning. Such a finding may also explain differences in diversionary propensities across regime type and regime capacity (Gelpi 1997; Lai and Slater 2006; Levy 1989; Oakes 2012) insofar as leaders and their state structures systematically vary in how able they are to collect public opinion data and how willing they are to take such data to account.

## Research design

---

<sup>18</sup> Critics of diversionary war as diverse as Blainey (1973) and Chiozza and Goemans (2003) claim to falsify the theory by attempting to demonstrate the latter.

The extant rally literature primarily analyzes observational data<sup>19</sup>—most quantitative though a few qualitative. For the question posed by the present project—are rallies affected by initiation status?—such designs face crucial limitations. As noted above, credibly identifying the causal effect of initiation status on rallies in observational data is difficult due to the uniqueness of each rally event in history. Experimental methods are a solution to such challenges. Researcher manipulation of the causal factor as well as random assignment of subjects to conditions allows for the identification of the causal effect of initiation status on outcomes of interest.

I recruit subjects for the study using Amazon’s Mechanical Turk (MTurk) service.<sup>20</sup> MTurk has quickly become a popular method across the social and behavioural sciences for recruiting large samples at relatively low cost. ‘Requesters’ may submit simple tasks to an online labour market and ‘Workers’ then agree to complete the tasks in exchange for a very small amount of remuneration (for an overview of MTurk see Mason and Suri 2012). It is necessary to note that MTurk users are not representative of the US population (Berinsky, Huber and Lenz 2012).<sup>21</sup> However, recruitment through MTurk is useful for assessing treatment effects among a non-random sample. Furthermore, MTurk samples are generally more representative of a national population than either in-person convenience samples or student samples Berinsky, Huber and Lenz (2012). Lastly, three factors—unique Worker IDs, ability of Requesters to reject submission of individual Workers, and reputation requirements to complete certain tasks—combine to place strong incentives on Workers to complete tasks with honesty and accuracy (Suri, Goldstein and Mason 2011).

The survey was fielded in the first and second week of March 2017. After a standard battery of demographic items, subjects are randomly assigned to one of three experimental conditions. In all three groups, subjects are exposed to a hypothetical news article. To maximize plausibility and realism, I generate the stories by using JavaScript to edit in browser an article

---

<sup>19</sup> See Hou and Quek (2015) and Merolla and Zechmeister (2009) for exceptions.

<sup>20</sup> The sample is restricted to MTurk users from the US and those with at least 90 per cent approval from prior tasks.

<sup>21</sup> Appendix 3.1 displays sample characteristics and a comparison with representative US data. The MTurk sample for this study appears more similar to the US population than those examined by Berinsky, Huber and Lenz (2012).

from the website of the *New York Times*.<sup>22</sup> The resulting stories are thus fictitious with content written and manipulated by the researcher, but with a professional and plausibly realistic typeface.<sup>23</sup> In the two treatment conditions, subjects are told that President Trump is moving US aircraft carriers to disputed waters in the South China Sea—specifically to the waters surrounding the Scarborough Shoal, which is contested by China, the Philippines, and Taiwan. In the ‘reciprocation’ treatment, subjects are additionally told that the actions are in response to recent missile tests that China has conducted in the disputed region.<sup>24</sup> Thus, initiation status is the only manipulated factor in this study.<sup>25</sup> Subjects in the control condition read an unrelated pop culture story.

I assess the dependent variable of support for the leader in two ways. First, subjects are asked whether or not they have a favourable opinion of Trump. Second, subjects are asked whether or not they would vote for Trump, should a hypothetical election be held today.<sup>26</sup> Between the administering of the treatments and the measurement of the dependent variables, subjects are also administered items that attempt to test potential mediators of rally effects, including patriotism, issue salience, evaluation of Trump’s competence, and whether subjects blame other countries or the US for the country’s economic and security concerns. Following the conclusion of the survey, subjects are fully debriefed about the hypothetical nature of the news

---

<sup>22</sup> Could the choice of *The Times* as the typeface template ‘tip off’ subjects? I suggest that such a possibility is unlikely. First, we should not overestimate the fraction of the US population that are regular readers of the newspaper. Second, those that are most likely to recognize the template as *The Times* are presumably readers of the paper, while those that are most likely to react negatively to the paper are also those who are least likely to read the paper in real life—and thus, the least likely to recognize and identify the origins of the typeface.

<sup>23</sup> See Appendix 3.2 for the news articles.

<sup>24</sup> I choose these particular actions—the sending of aircraft carriers and Chinese missile tests—based on American and Chinese actions during the 1996 Taiwan Straits Crisis.

<sup>25</sup> One concern may be that the results are subject to compound treatment effects, if the treatment manipulates both initiation status and leader justification of militarized action. To mitigate such concerns, both treatment stories report that Trump justifies the action by citing the need to “stand up to China,” and both treatments report the Philippines as welcoming the announcement.

<sup>26</sup> These variables are dichotomized as a result of a natural bimodal distribution in subject responses. An additional benefit is comparability with past estimates of rally size from observational studies that utilize aggregate public opinion data.

articles and are directed to an actual New York Times article explaining the South China Sea dispute should they desire to learn more about the issue.<sup>27</sup>

A total of 1428 subjects were recruited for participation in the study. The reported results are from analyses of the 1245 subjects who passed all three attention checks (see Berinsky, Margolis and Sances 2014). Results are generally somewhat stronger for the full sample.

## Results

Do publics rally differently to leaders initiating conflict as opposed to those that respond to conflicts initiated by the other side? As a first cut at this question, I estimate logistic regressions<sup>28</sup> of Trump favourability and likelihood of voting for Trump on two dummy indicators for the treatment groups. In both cases, I find a significant rally effect for the reciprocation treatment, but not the initiation treatment. Relative to the control group, the reciprocation treatment increases the probability of having a favourable opinion of Trump by six per cent ( $p=0.047$ ) and of voting for Trump in a hypothetical election held today by eight per cent ( $p=0.015$ ).<sup>29</sup> These effect sizes compare favourably with the largest estimates of rally effects in prior observational studies short of George W. Bush's rally after 9/11 (James and Rioux 1998; Lian and Oneal 1993; Oneal and Bryan 1995).<sup>30</sup>

---

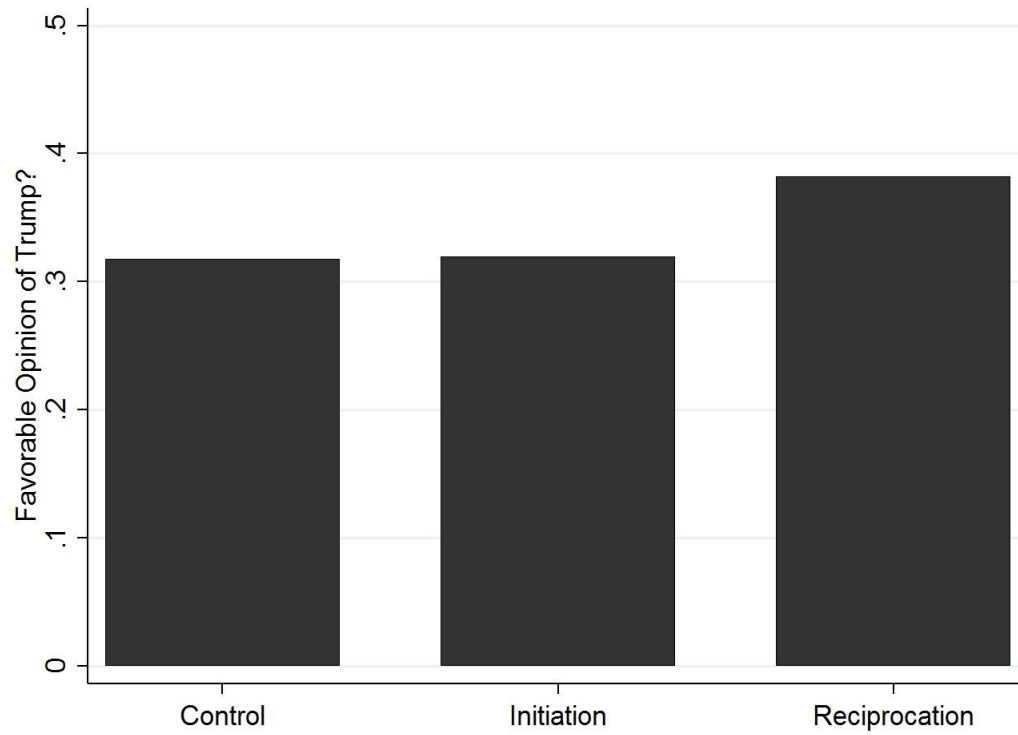
<sup>27</sup> [www.nytimes.com/2016/07/15/world/asia/south-china-sea-dispute-arbitration-explained.html](http://www.nytimes.com/2016/07/15/world/asia/south-china-sea-dispute-arbitration-explained.html)

<sup>28</sup> Regression results are available in Appendix 3.3.

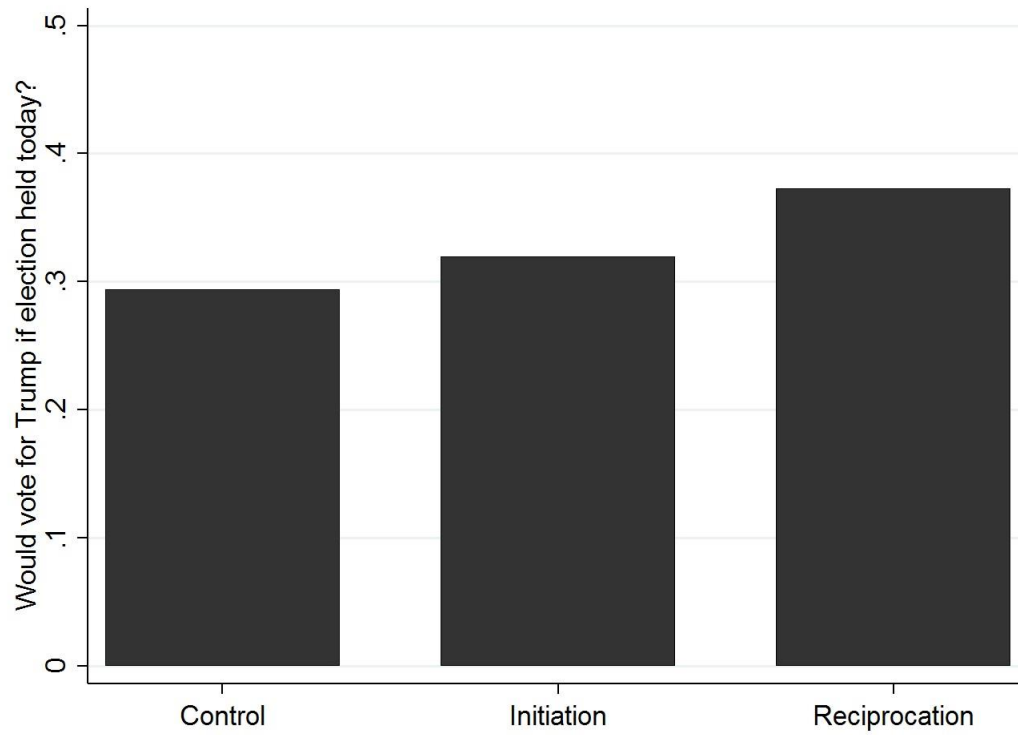
<sup>29</sup> Subjects in the reciprocation treatment are also 6 per cent more likely to have a favourable impression of Trump ( $p=0.061$ ) and 5 per cent more likely to vote for Trump ( $p=0.110$ ) than those in the initiation treatment.

<sup>30</sup> Oneal and Bryan (1995) find that rally events that are prominent enough to receive front page coverage in *The New York Times* produce an average rally effect of 8 per cent. One implication of the present finding, then, is that it is media salience and coverage, rather than attributes of the rally events, that drives the larger rallies from among this subset.





**Figure 3: Trump favourability by condition**



**Figure 4: Likelihood of voting for Trump by condition**

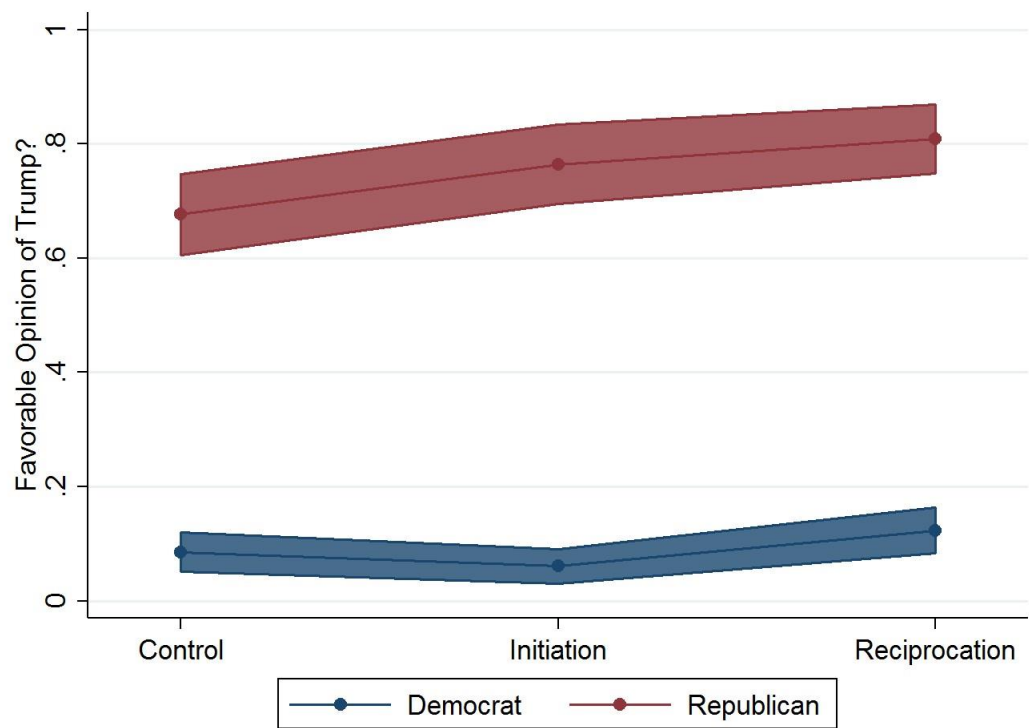
Figures 3 and 4 display respectively the proportion of subjects in each condition with favourable opinions of Trump and that would vote for Trump in an hypothetical election today. Perhaps unsurprisingly for a MTurk sample that skews Democrat,<sup>31</sup> approval and support for Trump was on average low. Only 32 per cent of subjects indicate favourable opinions of Trump in the control and initiation conditions, but this increases to 38 per cent in the reciprocation condition. Similarly, 29 per cent of subjects indicate that they would vote for Trump in an election held today, but this increases to 32 per cent in the initiation condition and 37 per cent in the reciprocation condition.

Given these results, one may be tempted to conclude that publics only rally for crises instigated by foreign adversaries, and not for those initiated by their own leaders. However, it is possible that different dynamics underlie the rally phenomenon for different subgroups. In particular, past studies point to partisan identification as a key moderator influencing how foreign crises affect support for leaders (Baum 2002; Hetherington and Nelson 2003; Kam and Ramos 2008; Mueller 1973). To investigate this possibility, I estimate logistic regressions of Trump favourability and likelihood of voting for Trump on an interaction of party ID and experimental condition. Figure 5 displays the proportion of subjects with a favourable opinion of Trump by condition and partisanship, while Figure 6 does the same for the 'would vote for Trump' variable.<sup>32</sup>

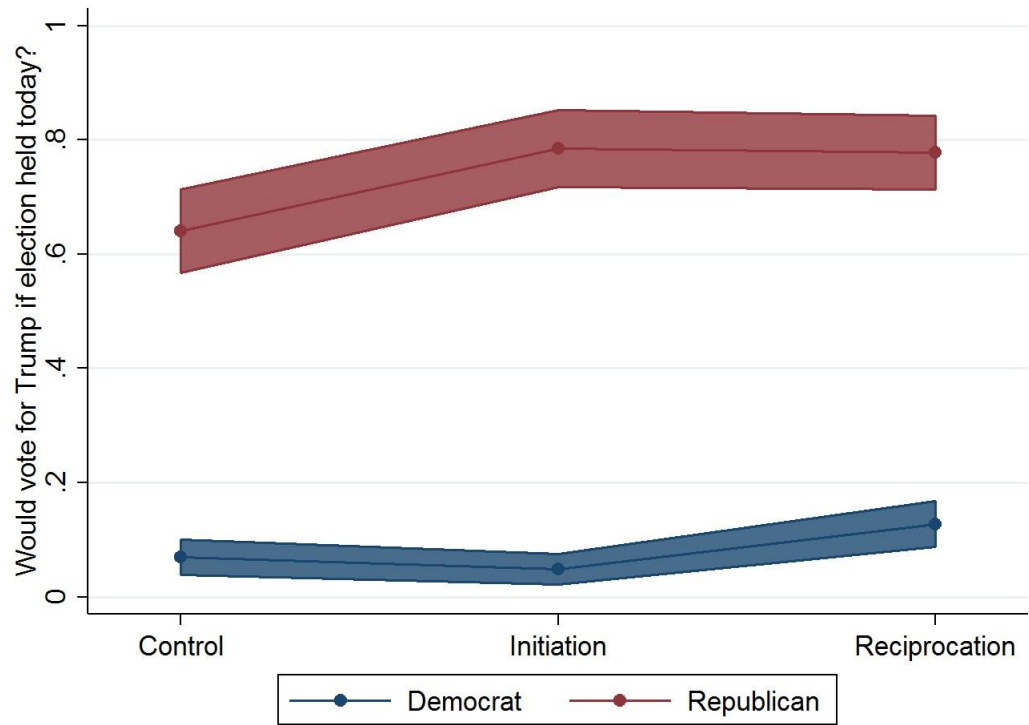
---

<sup>31</sup> Thirty-eight per cent of the sample identified as Republican or Republican-leaning while the remainder identified as Democrat or Democratic-leaning.

<sup>32</sup> Shading indicates 95 per cent CI



**Figure 5: Trump favourability by party ID and condition**



**Figure 6: Likelihood of voting for Trump by party ID and condition**

The results provide evidence for heterogeneous effects across the partisan divide. Only nine per cent of Democrats in the control condition have a favourable opinion of Trump. This drops to six per cent in the initiation condition, and rises to 12 per cent in the reciprocation condition. Significance tests of the contrasts between experimental conditions suggest the difference between reciprocation and initiation conditions are significant ( $p=0.017$ ). Similarly, only seven per cent of Democrats in the control condition indicate that they would vote for Trump, dropping to five per cent in the initiation condition and rising to 13 per cent in the reciprocation condition, with reciprocation significantly different from both the initiation treatment ( $p=0.003$ ) and from control ( $p=0.030$ ). Thus, Democrats appear to exhibit rallies for Trump only under the reciprocation treatment—largely mirroring results seen when analyzing the whole sample.

For Republicans, the story looks much different. Sixty-eight per cent of Republicans in the control condition hold a favourable opinion of Trump, which rises to 76 per cent in the initiation condition and 81 per cent in the reciprocation condition. Both the initiation versus control and the reciprocation versus control contrasts are significant ( $p=0.090$  and  $p=0.007$  respectively). Similarly, 64 per cent of Republicans in the control condition would vote for Trump in an election held today, rising to 78 per cent in both the initiation and control conditions. Again, both initiation and reciprocation differ significantly from the control condition ( $p=0.006$  and  $p=0.007$  respectively). The results suggest that while Democrats only rally to leaders responding to foreign aggression, Republicans appear willing to rally behind Trump irrespective of initiation status.

I test the robustness of these findings in two ways. First, following the recommendation of Berinsky, Margolis and Sances (2014), I repeat the above analyses at different thresholds of attention check completion (i.e. at least two passes; at least one pass). In most cases effect sizes increase, and in all cases significant contrasts between groups remain so. Second, I test the robustness of the partisan differences by estimating logistic regressions among subsets of Democrats and Republicans instead of interacting partisan identification with experimental conditions. The results do not substantively differ from those presented above.

To summarize: Among the full sample, I find no evidence that initiating conflict results in a rally effect, but strong evidence that publics rally to leaders who are responding to foreign aggression. This result however obscures heterogeneous rally dynamics across partisanship. Democrats appear to only rally to leaders under the reciprocation scenario, with suggestive evidence that they may punish a Republican executive who initiates conflict. In contrast, Republicans rally to their in-partisan leader under cases of foreign conflict, regardless of initiation status.

## **Discussion and conclusion**

The present findings have interesting implications for a number of theories concerning the relationship between public opinion and foreign policy. First, the finding of heterogeneous rallies across partisans makes an intriguing interjection in scholarship on rally effects. Extant work has shown that rallies are larger and decay more quickly among individuals who identify with the opposite party as the one that holds the executive office (Baum 2002; Hetherington and Nelson 2003; Mueller 1973; Kam and Ramos 2008). However, in the present study rally effects are substantially larger among Republicans than among Democrats (cf. James and Rioux 1998). It is unclear what exactly explains this apparent countervailing finding. One simple explanation is that out-partisans start from a lower baseline, and so we observe “larger” rallies from them under the reciprocation scenario. From that perspective, the present finding that Democrats double their approval of Trump in the reciprocation condition is consistent with past findings. It may also be the case that Trump is unique insofar as even Republican subjects in the sample appear to have relatively lukewarm baseline levels of support, suggesting that ceiling effects are less of a constraint on the size of in-partisan rallies that he can elicit.<sup>33</sup> Another possibility is that given current levels of political polarization, attitudes toward an out-party president are more resistant to

---

<sup>33</sup> It is worth noting that both Democrats and Republicans appear willing to rally to Trump at least in some circumstances, despite the affect-laden image of Trump in many voters' minds. In that sense, Trump can be taken as a 'hard test' of rally effects among out-partisans. This also highlights a strength of the present study's design relative to many other experimental studies in IR that use vignettes with hypothetical leaders (cf. Tomz 2007).

change than ever before. Lastly, prior work has looked at heterogeneity of rally effects across subgroups without systematically varying features of rally events themselves. The present findings suggest that heterogeneity across partisanship may interact or be driven by such features, such as initiation status. Thus, unpacking how other salient aspects of conflicts affect rally dynamics—such as the identity of the adversary, or whether or not the conflict is multilateral or bilateral—remain a potential avenue for future research to explore.

Furthermore, the treatments in the experiment represent relatively mild stimuli, when considering the real world cases that the treatments are trying to capture. While this presents a hard test of the theory, and partially offsets other study characteristics that bias in favor of finding strong treatment effects (such as a captive and receptive audience), it raises the possibility that characteristics other than initiation are being manipulated, and that those are driving the observed results. For example, it may be the case that subjects are responding to variation in perceived aggressiveness demonstrated by Trump, and that Republican respondents are more likely to reward trump for aggression than Democrats.

The results of the study have ambiguous implications for diversionary war theory. Taking the findings at face value, it appears to suggest that leaders who initiate foreign conflict in order to reap rally benefits at home may only gain among in-partisans, and may even suffer decreases in support from out-partisans punishing apparent belligerence<sup>34</sup>. How this would impact diversionary theory is unclear. On the one hand, in-partisans may be the most politically relevant group. Leaders may calculate that maintaining the loyalty of in-partisan supporters may be more important and feasible than trying to rally the other side. On the other hand, it does suggest that diversionary action may be inefficacious in situations where domestic dissatisfactions stem largely from partisans of the out-party. Trump, therefore, may wish to think twice before contemplating diversionary war as a strategy to boost sinking approval ratings.,

---

<sup>34</sup> Compared to control, Democrats in the initiation condition had 3 per cent lower approval, and this difference was not statistically significant. However, it also represents a loss of one third of Trump's support among Democrats in the control condition.



A corresponding implication is that a more precise conceptualization of diversionary theory may wish to take into account the source of domestic discontent—specifically, whether that discontent stems specifically from a leader’s base or is more widespread in nature. To the extent that diversion represents an initiated conflict, an important scope condition to diversionary theory may be that the source of domestic discontent should largely stem from in-partisans.

One of the most striking findings is that Democrats rally to Trump under the reciprocation condition. Although the measure of approval in that condition is not large, given Trump’s extremely low baseline approval among Democrats reciprocating a conflict still represents a doubling of Trump’s approval among Democrats in this study. Given the high degree of affective polarization in American politics today, especially under President Trump, the present findings suggest that there remains limits to the gravitational force of partisan identification: voters’ views of an out-party President are not unmovable; it just may take a foreign conflict to move them. It also suggest that the effect sizes in the present study may underestimate similar rally effects for a more typical leader.

It is important to note that the key limitation to the current study involve external validity—not only whether an online convenience sample can approximate the US population, but more importantly, whether a survey administering a news article concerning carrier movements can approximate rally events that we observe in the world. Do findings from such a survey tell us anything about public opinion dynamics following an event such as 9/11? Would such a survey experiment evoke the same reaction and emotion as the vividness of real world events? There are three critical ways in which rally phenomenon in the real world operates in ways that differ significantly from the scenario in the present study. First, leaders have the ability to frame rally events, engage in finger-pointing, and provide justifications to their publics for their actions (Baum and Groeling 2008; Levendusky and Horowitz 2012). The news articles in the survey attempt to maximize the distinction between initiation and reciprocation for the purposes of eliciting strong treatment effects; in real life, a complicated interaction of leader, elite, and journalistic incentives may make maintaining such a clean distinction difficult if not untenable. Indeed, in the real world

leaders will face strong incentives to manipulate perceptions of initiation and reciprocation, and any observed rally effects will be in part determined by the success of the leader in framing the initiation status of a given conflict.

Second, publics are not ‘captive audiences’ to foreign policy news stories as subjects in the study are (a condition compounded by MTurk samples being particularly attentive). A combination of variables that include availability and access of media technologies, journalistic incentives, and individual taste in terms of media consumption can together limit the extent to which any given individual may learn or even hear about a particular rally event. Thus, it seems unlikely that a real-world action similar in intensity to that portrayed in the treatments would generate the same rally effect size, and the extent to which such an action would generate rallies at all remains an open question.

Even setting aside the information advantage of leaders and the selective reception of news by the mass public, it is possible that different voters may be presented the same set of facts about a conflict event and come to different conclusions about the identity of the initiator. For example, a strict definition of initiation would hold that the US initiated the Gulf War against Iraq, though most observers in the US would view the conflict as a US *response* to Iraqi aggression—just against Kuwait, rather than the US specifically. To the extent that perceptions of who initiated among individual voters is of theoretical interest here, rather than any “objective” standard of initiation that can be applied to events, the divergence in views and perceptions among real world voters complicates the dynamics captured in this study. Indeed, divergent perceptions of initiation may (at least partially) explain the heterogeneous effects across partisanship; perhaps Republicans in the initiation condition still see Trump’s actions as a response to a general Chinese geopolitical threat.

In addition, the use of a US sample raises questions about what the results reveal about publics across the world. Clearly, the US is somewhat unique in having clear in-group out-group dynamics under a two-party system, at least relative to many European democracies. This may mean that the results found here—out-partisans rally for reciprocated threats, in-partisans rally

always—may not translate as cleanly to systems with multiple viable political parties and perhaps weaker partisan identities. However, many countries with multiple political parties find themselves in *de facto* two-party competitions where a right-leaning coalition competes against a left-leaning coalition for voters, and this situation is made more analogous with *de jure* two party systems due to the often short lifespans of parties. Ultimately, the extent to which present results travel to non-US populations is an empirical question, one that can only be answered through replication with other samples.

Finally, the present findings raises at least two additional questions to explore in future work. First, how does the identity of the target country affect rally results? China was selected for the present study given that it is one of the few plausible states with against which the US may be involved in a militarized dispute. However, past studies show that publics in democracies are more belligerent towards non-democracies (Tomz and Weeks 2013). Furthermore, even if militarized action against a democracy may be somewhat implausible for US publics, it still may be the case that different rally dynamics would be observed for US action against Russia or Iran, given differences in the history of bilateral relations and demographics of the target states (particularly race and religion). Thus, future work may wish to unpack how traits of the adversary affect rally effects.

Second, would the partisan differences remain similar under a Democratic president? One interpretation of the finding—Democrats rally when leaders respond to foreign aggression; Republicans rally irrespective of initiation status—is that voters are simply quicker and more willing to trust and endorse the foreign policy actions of an executive from their own party. If true, this implies that partisan dynamics should reverse when a Democratic executive is involved in a rally event. Thus, the expectation is that Democrats would have rallied to Obama regardless of whether or not he initiates or reciprocates a foreign conflict, while Republicans would restrict rallying to only the latter situations. It's not entirely clear whether public opinion during the Obama presidency confirms such an expectation. A second interpretation is that the partisan differences arise out of individual-level characteristics that covary with partisan identification. The past

several decades has been characterized by ideological sorting into parties, as liberals increasingly became Democrats and conservatives to Republicans (Levendusky 2009); whether hawkishness, belligerence, or militarism as predispositions have similarly been involved in a process of partisan sorting, either due to covariation with political ideology or for other reasons, remains to the best of my knowledge an open question<sup>35</sup>. Future research may wish to better disentangle these two possible interpretations of the present findings.

---

<sup>35</sup> The survey instrument also included a measure of pacifism; topline results and the interaction of pacifism with the treatments were not reported, given lack of significant interactions. However, the failure to find a moderating effect of pacifist orientation can be taken as one point of evidence against this second interpretation of the present results (that it is driven by partisan sorting along hawkishness and dovishness).

## References

- Baker, William D. and John R. ONeal. 2001. "Patriotism or Opinion Leadership? The Nature and Origins of the 'Rally 'Round the Flag' Effect." *Journal of Conflict Resolution* 45(5):661–687.
- Baum, Matthew A. 2002. "The Constituent Foundations of the Rally-Round-the-Flag Phenomenon." *International Studies Quarterly* 46(2):263–298.
- Baum, Matthew A. and Tim Groeling. 2008. *War Stories: The Causes and Consequences of Public Views of War*. Princeton: Princeton University Press.
- Berinsky, Adam J., Gregory A. Huber and Gabriel S. Lenz. 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20:351–368.
- Berinsky, Adam J., Michele F. Margolis and Michael W. Sances. 2014. "Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self-Administered Surveys." *American Journal of Political Science* 48(3):739–753.
- Blainey, Geoffrey. 1973. *The Causes of War*. London: MacMillan Press.
- Brody, Richard A. and Catherine R. Shapiro. 1991. *Assessing the President: The Media, Elite Opinion, and Public Support*. Stanford: Stanford University Press chapter The Rally Phenomenon in Public Opinion.
- Canes-Wrone, Brandice. 2005. *Who Leads Whom? Presidents, Policy, and the Public*. Chicago: University of Chicago Press.
- Chiozza, Giacomo and Hein E. Goemans. 2003. "Peace Through Insecurity: Tenure and International Conflict." *Journal of Conflict Resolution* 47(4):443–467.
- Fischer, Fritz. 1975. *War of Illusions: German Policies from 1911 to 1914*. Chatto & Windus Press.
- Fravel, Taylor. 2010. "The Limits of Diversion: Rethinking Internal and External Conflict." *Security Studies* 19(2).

- Gelpi, Christopher. 1997. "Democratic Diversions: Governmental Structure and the Externalization of Domestic Conflict." *Journal of Conflict Resolution* 41(2):255–282.
- Groeling, Tim and Matthew Baum. 2008. "Crossing the Water's Edge: Elite Rhetoric, Media Coverage, and the Rally-Round-the-Flag Phenomenon." *Journal of Politics* 70(4):1065–1085.
- Hetherington, Marc J. and Michael Nelson. 2003. "Anatomy of a Rally Effect: George W. Bush and the War on Terrorism." *PS: Political Science and Politics* 36(1):37–42.
- Hou, Yue and Kai Quek. 2015. "The Effects of Terror Attacks on Attitudes in China."
- Iyengar, Shanto and Adam Simon. 1993. "News Coverage of the Gulf Crisis and Public Opinion: A Study of Agenda-Setting, Priming, and Framing." *Communication Research* 20(3):365–383.
- James, Patrick and Jean Sebastien Rioux. 1998. "International Crises and Linkage Politics: The Experiences of the United States, 1953-1994." *Political Research Quarterly* 51(3):781–812.
- Kam, Cindy D. and Jennifer M. Ramos. 2008. "Joining and Leaving the Rally: Understanding the Surge and Decline in Presidential Approval Following 9/11." *Public Opinion Quarterly* 72(4):619–650.
- Kertzer, Joshua and Ryan Brutger. 2016. "Decomposing Audience Costs: Bringing the Audience Back Into Audience Cost Theory." *American Journal of Political Science* 60(1):234–249.
- Krosnick, Jon A. and Laura A. Brannon. 1993. "Impact of the Gulf War on the Ingredients of Presidential Evaluations: Multidimensional Effects of Political Involvement." *American Political Science Review* 87(4):963–975.
- Lai, Brian and Dan Reiter. 2005. "Rally 'Round the Union Jack? Public Opinion and the Use of Force in the United Kingdom, 1948-2001." *International Studies Quarterly* 49:255–272.
- Lai, Brian and Dan Slater. 2006. "Institutions of the Offensive: Domestic Sources of Dispute Initiation in Authoritarian Regimes, 1950-1992." *American Journal of Political Science* 50(1):113–126.

- Levendusky, Matthew S. 2009. *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans*. Chicago: University of Chicago Press.
- Levendusky, Matthew S. and Michael C. Horowitz. 2012. "When Backing Down is the Right Decision: Partisanship, New Information, and Audience Costs." *Journal of Politics* 74(2):323–338.
- Levy, Jack. 1989. *Handbook of War Studies*. Routledge chapter The Diversionary Theory of War: A Critique, pp. 259–288.
- Lian, Bradley and John R. Oneal. 1993. "Presidents, the Use of Military Force, and Public Opinion." *Journal of Conflict Resolution* 37(2):277–300.
- Mason, Winter and Siddharth Suri. 2012. "Conducting Behavioral Research on Amazon's Mechanical Turk." *Behavior Research Methods* 44(1):1–23.
- Mayer, Arno J. 1969. "Internal Causes and Purposes of War in Europe, 1870-1956: A Research Assignment." *Journal of Modern History* 41(3):291–303.
- Merolla, Jennifer L. and Elizabeth J. Zechmeister. 2009. *Democracy at Risk: How Terrorist Threats Affect the Public*. Chicago: University of Chicago Press.
- Mueller, John. 1970. "Presidential Popularity from Truman to Johnson." *American Political Science Review* 61(1):18–34.
- Mueller, John. 1973. *War, Presidents, and Public Opinion*. New York: Wiley & Sons.
- Mueller, John. 1994. *Policy and Opinion in the Gulf War*. Chicago: University of Chicago Press.
- Oakes, Amy. 2006. "Diversionary War and Argentina's Invasion of the Falkland Islands." *Security Studies* 15(3):431–463.
- Oakes, Amy. 2012. *Diversionary War: Domestic Unrest and International Conflict*. Stanford: Stanford University Press.
- Oneal, John R. and Anna Lilian Bryan. 1995. "The Rally 'Round the Flag Effect in US Foreign Policy Crises, 1950-1985." *Political Behavior* 17:379–402.

- Parker, Suzanne L. 1995. "Toward an Understanding of 'Rally' Effects: Public Opinion in the Persian Gulf War." *Public Opinion Quarterly* 59(4):526–546.
- Smith, Alastair. 1996. "Diversionary Foreign Policy in Democratic Systems." *International Studies Quarterly* 40(1):133–153.
- Sprecher, Christopher and Karl DeRouen. 2002. "Israeli Military Actions and Internationalization-Externalization Processes." *Journal of Conflict Resolution* 46(2):244–259.
- Suri, Siddharth, Daniel G. Goldstein and Winter A. Mason. 2011. "Honesty in an Online Labor Market." Papers from the Workshops at the AAAI Conference on Artificial Intelligence.
- Taylor, A. J. P. 1954. *The Struggle for Mastery in Europe, 1848-1918*. London: Clarendon Press.
- Tomz, Michael. 2007. "Domestic Audience Costs in International Relations: An Experimental Approach." *International Organization* 61(4):821–840.
- Tomz, Michael and Jessica L. P. Weeks. 2013. "Public Opinion and the Democratic Peace." *American Political Science Review* 107(4):849–865.
- Tyler, Patrick E. 1992. "U.S. Said to Plan Raids on Baghdad Over Inspections." *New York Times* p. 1.
- Yue, Jiahua. 2016. "Rally 'Round the Red Flag? Revisiting Public Opinion in China During International Crises.".
- Zaller, John. 1993. "The Converse-McGuire Model of Attitude Change and the Gulf War Opinion Rally." *Political Communication* 10(4):369–388.



## CHAPTER 4: Diversionary Theory Under Mediated Politics

On March 1, 2017, the Washington Post reported that Jeff Sessions, President Trump's newly-confirmed Attorney General, met twice with the Russian Ambassador in July and September 2016 (Entous, Nakashima and Miller 2017). The latter meeting occurred during "the height of what U.S. intelligence officials say was a Russian cyber campaign to upend the U.S. presidential race." The revelations were particularly damaging for Sessions given his claim, during his Senate confirmation hearings a month prior, that he "did not have communications with the Russians" (Abramson 2017).

On March 4, two days after Sessions recused himself from any investigation into Russia's role in the 2016 election, President Trump alleged—in a 6:30 a.m. tweet on a Saturday—that his phones had been wire tapped by then-President Obama in the lead-up to the Election Day.<sup>3637</sup> Reactions from Congress ranged from confusion to anger. Republicans such as Marco Rubio and Tom Cotton admitted that there was no evidence supporting Trump's claim, while Democrat Nancy Pelosi argued that Trump's tweet was a diversion tactic: "You make up something, then you have the press write about it, then you say everybody is writing about this charge. It's the tool of an authoritarian, to just have you always be talking about what you want them to be talking about it" (Cheney 2017). Indeed, the tweets immediately before and after the one containing the wire tapping claim both draw parallels between Sessions' and Obama's actions vis-a-vis Russia.<sup>38</sup>

A number of observers saw this episode as just another instance in which Trump, facing mounting critical coverage on one issue, makes an outrageous claim to divert and distract the media and the public. Richard Wolffe, writing for The Guardian, describes it thus: "When the

---

<sup>36</sup> His exact claim: "Terrible! Just found out that Obama had my 'wires tapped' in Trump Tower just before the victory. Nothing found. This is McCarthyism!" <https://twitter.com/realDonaldTrump/status/>

<sup>37</sup>

<sup>38</sup> In the tweet prior, Trump claims that Sessions' meeting with the Russian ambassador was set up by the Obama administration. In the latter tweet, Trump compares Sessions' two meetings with the number of meetings that the same ambassador had met with Obama.

media coverage, or the congressional pressure, gets too tough, a simple tweet is enough to send the press stampeding in the other direction” (Wolffe 2017). A CNN article argues that Trump’s playbook is to “go on offense with a new conspiracy theory” (Wolf 2017). Bloomberg sees this tactic as one of generating controversy over superficial matters while distracting the public from substantive issues and policy (Bershidsky 2017).

Bershidsky’s article draws a connection between Trump’s diversionary rhetoric with the theory of diversionary war. Leaders, according to the theory, may wish to divert the public’s attention away from woes at home to foes abroad. Diversionary motives have been suggested as a cause to a number of interstate conflicts, from the Russo-Japanese War to the Falklands to Russia’s militarized actions in Ukraine and Syria in recent years. Political science and related disciplines have now accumulated a large body of work, both qualitative and quantitative, that seeks to test and assess this theory. Indeed, Bershidsky is not the first to apply the theory to the Trump administration. Writing in the *Monkey Cage*, Dennis Foster argues that Trump’s psychological traits—specifically distrustfulness and conceptual simplicity—makes him particularly likely to engage in diversionary war (Foster 2016; Foster and Keller 2014).

What is striking about the episode of Trump’s wire tapping tweet is not only the ways in which it exemplifies the logic of diversionary theory, but also how it simultaneously reveals the theory’s limitations, at least in its common formulation. First, as Bershidsky notes, “Trump doesn’t need to start wars.” Diversionary theory has traditionally been thought of as an explanation for war, rather than of foreign policy behavior broadly construed. If diversion is about distracting populations by manipulating the salience of different issues, why should starting a war be the only way to accomplish this goal? Second, Trump’s efforts at diversion have arguably mixed results. According to Buzzsumo, the most widely shared New York Times news story in the first week of March 2017 was indeed about Trump’s wire tapping claim; the story about Sessions’ recusing himself had less than half the shares. On the other hand, the Washington Post article which broke the story about Sessions’ meetings with the Russian ambassador has twice the shares as the Post story about the wire tapping claim, while the most popular articles on both websites in the

period starting from Trump's inauguration were ones about Trump's immigration executive order. Thus, despite the media skill that some observers ascribe to Trump, there are clear limitations in the extent to which an executive under a free media regime can manipulate the issues that journalists choose to cover, the tone of coverage, and the level of viewership and reach that various stories enjoy. Diversionary theory has traditionally overlooked both the information environment in which leaders find themselves as well as the degree to which issue salience among the public is a key component for the theory to operate. The payoffs for refining diversionary theory include improving our understanding of the conditions under which leaders engage in interstate conflict, of the mass public's role in foreign policy, and of the relationship between domestic and international politics.

I build on the existing literature in two ways. First, I argue that diversion generally does not take the form of war, which are costly endeavours that may result from multiple logics and motivations. Diversion at its core is the purposeful manipulation of foreign policy salience. Thus, I suggest that leaders who wish to divert the public's attention away from domestic ills, rather than initiating costly conflicts, may instead engage in a number of acts ranging from sabre-rattling, 'Trumpian' rhetoric to low-intensity kinetic actions like drone strikes in order to raise the salience of foreign policy while lowering the salience of domestic politics. By refocusing theory away from diversionary war to diversionary foreign policy broadly understood, I expand the range of leader and state behaviour that can be explained by domestic politics.

Second, I argue that leaders face domestic constraints that condition their propensity to engage in diversionary action. Specifically, a country's media system and the flow of information between elites and citizens condition the ability of leaders to reap the rewards of diversion. Much of the existing literature sees democratic regimes as being more likely to divert due to greater susceptibility to public opinion, but when Pelosi suggests that diversion is a "tool of the authoritarian," she is closer to the mark than political scientists realize. Leaders in regimes with free media systems face low incentives to undertake diversionary action. Journalistic independence in terms of what stories are covered and how stories are framed means that

leaders under free media regimes are unlikely to reap the rewards of diversion. Leaders who exercise control over the media, however, can mandate high volume and positive valence of media coverage in order to maximize the benefits of a diversionary policy.

Using a cross-national time series data set of international conflict, domestic unrest, and media freedom, I attempt to answer three questions in order to test the theory. First, does variation in media freedom condition the relationship between unrest and conflict? Second, are states with political control over the media more likely to engage in conflict as unrest increases? Third, are there differences in the relationship between media freedom, conflict, and unrest for conflicts of different intensity?

The results provide suggestive evidence of how domestic institutions such as media systems can structure the relationship between public opinion and foreign policy. My findings suggest that media freedom does condition the relationship between unrest and conflict—particularly for low-intensity conflicts. However, I do not find evidence that this relationship is driven by greater belligerence during high unrest periods in states without free media.

## **Literature review**

Diversion from domestic ills is one of the oldest, most popular, and most controversial explanations for the causes of war. The example that is most often held as smoking gun evidence of a diversionary war is the statement by Russian interior minister V. K. Plehve, on the eve of his country's 1904 conflict against Japan, that his country needs "a little victorious war to stem the tide of revolution" (quoted in Blainey 1973). Indeed, the idea that leaders may initiate foreign conflicts in anticipation of domestic dividends has been leveraged to explain conflicts as diverse as the Argentine initiation of the Falklands (Oakes 2006), various European conflicts of the late 19th and early 20th century (Blainey 1973; Fischer 1975; Mayer 1969; Levy 1989; Taylor 1954), and even the timing of military operations in the Gulf War (Tyler 1992).

Reviewing the political science literature on diversionary war, Levy (1989) suggest that there is mixed evidence in support for the theory, at least among quantitative studies. Nearly 30 years later, this conclusion still largely holds true. A number of studies cast doubt on diversionary war entirely (Chiozza and Goemans 2003; Fravel 2010; Gaubatz 1991) while those that find supporting evidence usually focus on delineating scope conditions. Some argue that domestic turmoil needs to be very large for leaders to resort to diversion (Dassel and Reinhardt 1999; Davies 2002). Others suggest that ethnic fragmentation may play a role in providing opportunities for diversionary action, either due to transborder ethnicking ties (Haynes 2017) or from the existence of vulnerable ethnic minorities at home (Tir and Jasinski 2008). Scholars have also examined the role of regime type, with most arguing that democracies are particularly liable to engage in diversion (Russett 1990; Gelpi 1997; Oneal and Tir 2006), although Lai and Slater (2006) argue that variation exists among authoritarian regimes as well. Relatedly, Oakes (2012) argues that regimes with low state capacity are more likely to have to resort to diversionary wars, while high capacity states have greater options that they can substitute for war when facing domestic unrest.

A related but distinct literature has also developed that attempts to answer whether leaders do, in fact, benefit from foreign conflicts due to the propensity of publics to 'rally 'round the flag.' Mueller (1973) finds evidence that dramatic international events involving the US consistently increased support for American presidents. Since then, scholars have examined rallies outside of the US (Hou and Quek 2015; Lai and Reiter 2005; Sprecher and DeRouen 2002), the duration of rally effects (Hetherington and Nelson 2003; Kam and Ramos 2008; Lai and Reiter 2005), and heterogeneity of rallies across different subgroups (Baum 2002; Groeling and Baum 2008; Hetherington and Nelson 2003). One of the most striking findings to emerge is the moderating role of media salience. A number of studies by John Oneal and coauthors find that while the average rally size approximates zero, they need to be consistently large when the foreign policy crisis enjoys high media coverage (Baker and Oneal 2001; Lian and Oneal 1993; Oneal and Bryan 1995).

A number of explanations have been put forth by both IR and public opinion scholars for why publics rally to leaders during times of crises. One explanation centres on psychology. In IR, Coser (1956, see also Simmel (1898)) was one of the first to suggest that conflict may serve as a form of socialization, maintaining and strengthening group identities; this idea is supported by later psychological theories of social identity (Tajfel and Turner 1979). More recently, Kam and Ramos (2008) show that national identities become a more important determinant of presidential approval following rally events. A second explanation for rally effects focuses on scapegoating. Taylor (1961) argues that Hitler and the Nazi Party was able to rally the German people around the perception that Germany's domestic woes were the result of punishments imposed on the country by Allied powers through the Treaty of Versailles, while more recent examples include the phenomenon of 'China bashing' during the American electoral cycle. A third explanation conceptualizes diversionary war as a principal-agent problem. The public desires competent executives, but competence is private information. A diversionary foreign policy can be used to publicly demonstrate leader competence; the resulting rallies obtain because of public approval of leader competence (Richards et al. 1993; Smith 1996). Related, Merolla and Zechmeister (2009) show that rally events spur the public to desire strong, authoritarian leaders. A final explanation for rally effects focus on the role of the media. Specifically, media coverage of foreign policy can raise leader approval by changing the importance that voters attach to different issues (agenda-setting), and corresponding, changing the issues that voters use to evaluate political candidates (priming).

These two literatures—one on diversionary warfare, one on rally 'round the flag effects—have developed largely independently of each other. By focusing on the effect of conflicts on public support, the rally literature largely overlooks the possible reciprocal nature of these two variables—a relationship posited by diversionary war theory. Furthermore, the majority of the rally literature looks at the relationship between American public opinion and foreign policy, yet there is no strong a priori reason for assuming that findings concerning the mass public and interstate behavior of the US accurately account for elite-mass dynamics in other polities. In fact, as I argue

below, there is strong reason to believe that variation across regimes in the flow of information between elites and the mass public can generate different incentives for leaders contemplating diversionary foreign policy. At the same time, it is precisely the role of information that is largely overlooked in the traditional diversionary literature. In particular, studies that test for the effect of domestic turmoil on foreign conflict always implicitly assumes—but rarely if ever explicitly states—that leaders hold beliefs and expectations about the benefits of conflict. That is to say, hypothesizing that leaders will engage in aggressive foreign policies when facing domestic unpopularity implicitly assumes that leaders expect the foreign aggression to ameliorate their domestic woes in some way. Whether or not leaders operating within diverse institutional arrangements can all uniformly expect to benefit from conflict is a question that remains unexplored.

## **Theory**

I propose amending diversionary theory in two ways. First, I argue that diversionary wars in the strictest sense are few and far in between. Not only do wars have multiple explanations and causes, they entail significant costs for the country and leader. At its core, diversionary theory is about the manipulation of salience. Diversion occurs when actions by the leader successfully raises public attention towards foreign policy issues and lowers attention towards woes domestically—I suggest that this outcome can be achieved by actions other than outright wars. However, the ability of leaders to successfully engage in diversion is constrained by domestic political institutions. Should a leader choose to enact a diversionary foreign policy, the relationship between the leader, mass media, and the public affects how likely it is that the leader reaps the benefits of diversion.

## **Moving away from diversionary war**

Levy (1989) suggests that qualitative and historical scholarship have found stronger support for diversionary war than studies using other methodologies. However, nearly every historical instance of war that one scholar has attributed to diversionary motives, another has claimed otherwise. This is true even for the 'smoking gun' example of Russia in its 1904 conflict against Japan. Blainey (1973) outlines a number of reasons to doubt Interior Minister Plehve's 'war to stem revolution' quip as insight into Russia's motives for entering war. While Russian leaders were no doubt anxious over the possibility of popular discontent, Blainey points to Russia's imperialist ambitions in northeast China as the primary motivation for the conflict. Additionally, there are potential biases in the source of the quote, which originated in the memoirs of a former Russian Minister of Finance Count Witte. Not only did Witte record the statement eight years after the fact, he was also a noted opponent of Plehve and personally disapproved of the war, both of which gives him motive to portray the interior minister in an unfavourable light. Indeed, concerns about biases in historical accounts is one of the primary criticisms that Blainey makes against diversionary theory broadly. Alexander Kinglake, who made an influential argument about Louis Napoleon's diversionary motives for initiating the Crimean War, also happened to have a personal vendetta against the French ruler: "Kinglake's explanation of the Crimean War, it is said, was simply a way of maliciously taking revenge on [Napoleon]" (p. 75). Blainey similarly believes diversionary explanations for the Franco-Prussian War of 1870 to be motivated by Prussia's desire to absolve itself from blame for the conflict. The difficulty of studying and observing leader beliefs, the normative implications of a leader who wars out of a selfish desire to mislead the populace, and the incentives for various actors to misrepresent motivations combine to make historical evidence for diversionary war a fraught terrain to navigate (see Lustick 1996).

Not only are diversionary motives difficult to observe in the historical record, they are also difficult to disentangle from other causes of war. Political science has traditionally thought about the causes of war as being systemic, rooted in the nature and ordering of the international system. Theories of hegemony suggest that wars occur when differential growth and



development rates across states results in a disequilibrium between international structure and a redistribution of state power (Gilpin 1981; Organski and Kugler 1980). Waltzian neorealism sees wars resulting out of desire for security under international anarchy, though this explanation focuses more on war's recurrence rather than any particular instantiation (Waltz 1988). Fravel (2010) critiques diversionary explanations of the Falklands on similar grounds; the Argentine initiation of the conflict in 1982 seems better explained as a response to Britain appearing to take sovereignty of the islands off the negotiating table a year prior to the start of war. More recent efforts to integrate explanations of war at different levels of analysis have failed to rule decisively in favor of any one theory or set of theories (Bennett and Stam 2004). Instead, war seems to be a phenomenon characterized by multiple causes both international and domestic, as well as the interaction of these causes with each other. Weisiger (2013) outlines several distinct pathways or mechanisms by which war can occur. While incentives to divert is one of the pathways that he considers, the explanatory emphasis is on commitment problems, which are found to produce the most catastrophic conflicts.

Perhaps the biggest difficulty facing diversionary war theory is the deterrent effect of the high cost of waging war. War, it is often noted, entails the absolute destruction of property and as such is always less efficient than bargaining to a mutually preferable settlement (Fearon 1995). Waging war in order to improve one's domestic standing seems self-defeating, insofar as the decision may result in the loss of lives, the destruction of economic assets, and the disruption of the domestic economy. One estimate places the cost to date of American wars in the Middle East since 2001 at \$5.6 trillion (Crawford 2017), while the Iraq and Afghanistan conflicts have claimed 4494 and 2424 American lives respectively (Department of Defense 2019). Fighting wars also carry indirect costs for the leader. Executives who are known to initiate international disputes because of domestic politics may acquire a reputation for bellicosity or even irrationality, which can reduce the credibility of any future commitment to pacifism and restraint. These reputational costs may also be intrinsic; it seems plausible that most leaders' self-image would be harmed should they acquire a reputation for belligerence, and given that diversionary motivations have

often been used to characterize leaders in a negative light, leaders may also be concerned for their historical legacy and the image of their rule that is carried into posterity.

One may respond by suggesting that leaders who face the prospect of losing power may nevertheless choose to incur these costs and attempt to 'gamble for resurrection.' Scholars since at least Kant have suggested that the costs of war are not borne by leaders but rather by the people. Furthermore, the consequences of losing political power may be so dire in some cases that war and its associated costs may seem like a worthwhile risk (Goemans 2000). I argue, however, that 'resurrection' can be bought at significantly cheaper cost. At its core, diversion is about the manipulation of salience, or the attention that the public pays to various political issues. Specifically, leaders are thought to divert in order to redirect the attention of the mass public away from troubles in society or the domestic economy towards international affairs and the threat of a foreign aggressor. However, there are a wide range of foreign policy actions within the toolbox of leaders that fall short of war but can nevertheless affect issue salience. Baggott (2016), for example, shows how US leaders can engage in diversionary 'cheap talk' in order to affect approval. Leaders can also take one step beyond rhetoric by publicly showing and displaying force. These can range from troop mobilization to military or nuclear alerts. Chinese foreign policy in particular is rife with creative demonstrations of force, such as sending troops to repeatedly cross over a disputed border with India, or the recent confiscation of a US underwater drone by the PLA navy. Finally, leaders can choose to engage in limited uses of force that nevertheless fall short of a declaration of war. Drone strikes may be one example of such an action. All of these actions have the potential to affect the relative salience of foreign policy and domestic issues without incurring the costs associated with war.

Why would leaders choose distraction with foreign policy crises if domestic alternatives—such as playing up ethnic or racial tensions, or engaging in redistributive policies—are available? I suggest that foreign vs. domestic diversion are not mutually exclusive—leaders do not have to forswear foreign policy diversion just because domestic levers are available options. Leaders prefer, all else being equal, a larger menu of options and a wider degree of

manoeuvrability (Snyder and Borghard 2011). The question then becomes why leaders would ever pursue foreign policy diversion at all, to which I raise two possible explanations. First, domestic diversion can carry greater costs than foreign diversion. A leader risks antagonizing and writing off the future support of a segment of the electorate or selectorate should they choose to engage in domestic diversion; for example, enflaming in-group out-group tensions may mean losing the support of the out-group in the future. However, foreign publics are not voters, and they are not part of a leader's selectorate—they represent easy targets for scapegoating. Second, domestic diversion may suffer from diminishing returns. Inflaming ethnic tensions is not a strategy that a leader can pursue in perpetuity. In such cases, foreign diversion may be a likely alternative, even if they do not present the most preferred course of action for a given leader.

### **Domestic constraints to the benefits of diversion**

Leaders can manipulate foreign policy salience through actions other than war. However, they face certain domestic constraints that limit their ability to reliably reap the rewards from diversion. The flow of information and the structure of a country's media institutions and environment is crucial to the causal story implicit in diversionary theory—not only does the issue need to resonate with the public, but news of the issue needs to reach the public in the first place. Journalists under a free media system select from the universe of all possible stories those they deem to be most newsworthy (Baum and Groeling 2008; Graber and Dunaway 2015; Shoemaker and Vos 2009). This 'gatekeeping' literature in political communication suggests that many foreign policy issues may not be seen as interesting enough to warrant a significant allocation of precious print space or broadcast minutes. Additionally, coverage of political leaders has become increasingly negative across Western democracies in the latter part of the 20th century (Hallin and Mancini 2004), such that negativity and critical coverage has often been seen as part of the newsworthiness criteria (Groeling and Baum 2008; Soroka 2014). This negativity bias in media

suggests that, in addition to uncertain volume of news coverage, what coverage does occur will often be critical of the leader.

In other countries, the structure of political and media systems accord leaders with significant levels of control over the content, timing, and salience of media coverage. In such cases, political control over the media allows leaders to mandate high levels of media coverage when such coverage is politically desirable. Furthermore, political control ensures uniformly positive coverage of the foreign conflict, overcoming psychological and market incentives for negative, critical news. Thus, 'negative' censorship (the prohibition of critical coverage) and 'positive' censorship (mandating politically favorable coverage of a foreign conflict) work together to ensure that leaders who exercise political control over media institutions may more reliably expect to reap the benefits of diversionary actions.

Given this theoretical approach, the present study tests three hypotheses:

H1: The relationship between domestic unrest and conflict initiation is moderated by media freedom

H2: States without free media are more likely to initiate conflict as unrest increases, while states with free media see a weaker or no relationship between unrest and conflict

H3: The moderating effect of media freedom on the relationship between domestic unrest and conflict initiation is stronger for low-intensity conflicts

## **Data**

To test my hypotheses, I collect and merge a number of existing datasets to create one capturing monadic dispute initiation by all countries in the world from 1948 to 2010. This temporal range is appropriate given the scope of the theory tested here. The postwar era represents the

rise of mass media consumption in most parts of the world, which was previously limited by variation in economic development and in media technologies. My argument, which focuses specifically on how the media as an actor conditions the relationship between leaders and publics in the area of foreign policy, thus operates in eras of world history where mass media is, indeed, for the masses. This temporal range also sidesteps historiographical debates about the role of domestic political considerations in the World Wars, as well as the volatility of statehood in this period as states entered and exited the international system. A monadic dataset is similarly appropriate, given my theoretical interest and focus on how domestic political institutions conditions international behavior; the aim is to explain *when* and *why* states engage in conflictual behavior, rather than to explain *against whom*.

For my dependent variable of dispute initiation, I use the Correlates of War Project's data on militarized interstate disputes (Ghosn, Palmer and Bremer 2004). The data defines a dispute as "united historical cases in which the threat, display, or use of military force short of war by one member state is explicitly directed towards the government, official representatives, official forces, property, or territory of another state" (Jones, Bremer and Singer 1996, 168). My dependent variable is binary, taking on a value of one for any country-year which initiates at least one dispute. In total, just under 16 per cent of the country-year observations initiated at least one dispute (N = 1465). I generate a dichotomous specification because theoretically it is unclear why a diverting leader should need more than one dispute.

If diversion is the manipulation of foreign policy salience, not only are there diminishing returns (and greater risks) to disputes beyond the first, but it may also be more efficacious to manipulate salience by controlling the intensity of one dispute rather than to create a second. Furthermore, examining the data I find that only four per cent of the country-year observations in the dataset contains two or more disputes. Nevertheless, I include the count of the number of disputes initiated by a given country-year as an alternative specification of the dependent variable.

For the independent variable of unrest, I use Arthur Banks' Cross-National Time Series dataset, which has been most often used in past studies of diversionary war. Also in keeping with prior literature (see Haynes 2017), I create the variable by generating an additive index of the count of protests, strikes, and demonstrations observed in a given country-year; the three are measured separately in the CNTS dataset. The index is thus a general measure of the level of unrest experienced by a given country-year. There are, however, at least three potential issues with the CNTS unrest data. First, the majority of country-year observations (approximately 70 per cent) have not even a single instance of unrest observed, which is unlikely to be true *prima facie*. Second, the large number of observations with no observed unrest is particularly concerning if the factors that explain why unrest is uncaptured in certain country-years is correlated with the moderator of media freedom—a possibility that is extremely likely. If unrest in states with political control over their media institutions is systematically underreported in the CNTS dataset, that suggests a bias against the ability of regression analyses to discern a positive correlation between unrest and dispute initiation for these states. Furthermore, it suggests that this bias should be stronger for states with political control of the media compared to states with media freedom. Both possibilities thus work against my ability to find evidence consistent with my theoretical expectations. A third potential issue with the unrest index is the presence of a few observations with extremely large values; the maximum value on the unrest index is 85, yet 95 per cent of the country-year observations have five or less incidents<sup>39</sup>. As a result, in subsequent analyses I use a log-transformed specification of the index.

To measure media freedom, I use the Global Media Freedom Dataset (Whitten-Woodring and Van Belle 2017). The dataset begins in 1948 and runs until 2012, and codes the level of media freedom in a given country-year as “Free,” “Partly Free,” or “Not Free.” According to the coding criteria of the dataset, states are coded as “Free” if media criticism of government and officials are a common part of the political dialogue; “Partly Free” if criticism can and does occur, but is limited by social, legal, or economic costs that can be imposed upon media organizations;

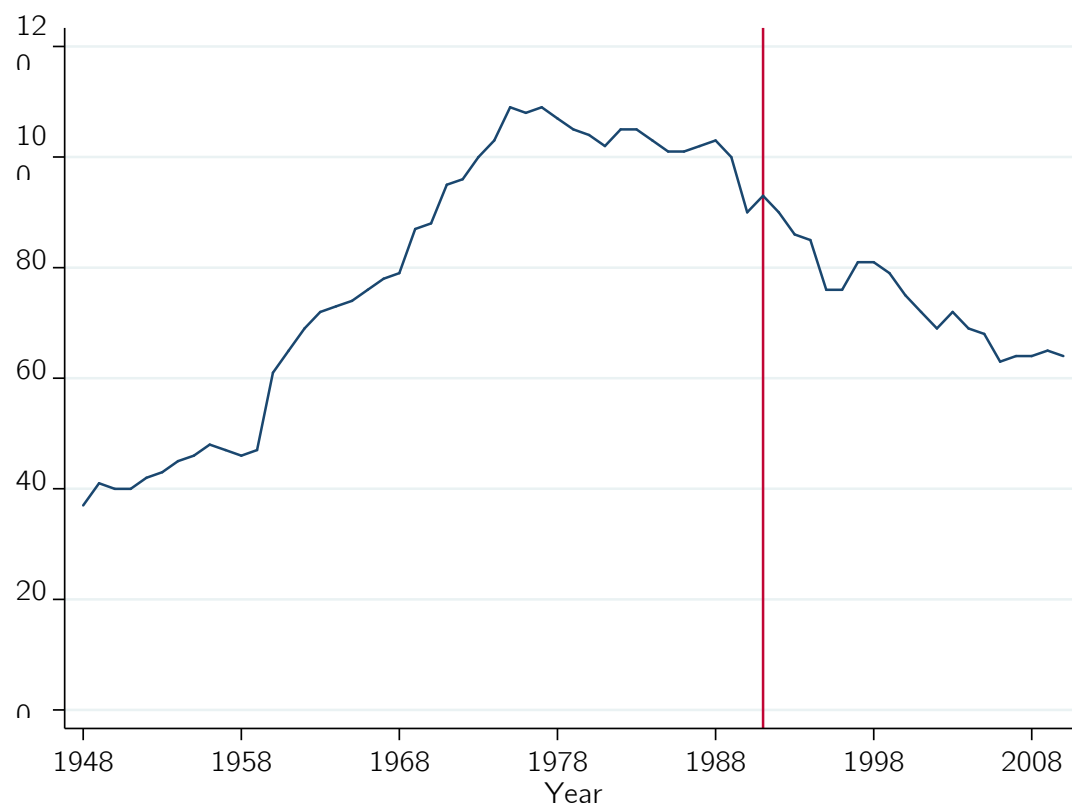
---

<sup>39</sup> The kurtosis of the distribution of the unrest index is 173; the kurtosis of the normal distribution is 3.

and “Not Free” if it is not possible to directly criticize the government, and if the media is directly or indirectly controlled. Approximately 26 per cent of the country-year observations in my dataset are categorized as Free, while 53 per cent are considered Not Free<sup>40</sup>. A visual inspection of the data suggest that media freedom is strongly correlated with regime type; most states that fall into the Not Free category are the Soviet satellites, Communist and authoritarian regimes in Asia, Africa, and the Middle East, and the Latin American states during their authoritarian periods in the '70s and '80s. However, a number of democracies appear as having Not Free media in the dataset, including Colombia (2000-2005), Portugal (1976-1994), and Burundi (2005-2010). On the other hand, a number of non-democracies are coded as having free media, including Panama (1948-1968), Senegal (1981-1999), and Thailand (2006-2009). Furthermore, most states are coded as Not Free for one period within the temporal range, though a few—such as Argentina and Pakistan—enter and exit Not Free status three or more times. Figure 7 portrays the number of Not Free states across time. As the graph indicates, there is an increase in the number of states with Not Free media from approximately 1960 to 1978, with a slow decrease following the end of the Cold War (indicated by the red line). Thus, the data shows that despite the intuitive and often-assumed overlap between regime type and media freedom, this overlap is fuzzy; the two concepts are both theoretically and empirically distinct; and there is additional empirical traction obtained by considering the role of media systems specifically rather than regime type broadly.

---

<sup>40</sup> A list of the countries coded as Not Free, as well as the temporal period during which they are measured as such, appear in Appendix 4.1.



**Figure 7: Number of not free media states in the world, 1948-2010**



In my analyses, I also include a number of control variables that may explain unrest or media freedom and conflict initiation. Because I am theoretically motivated to disentangle coarsely-measured regime type categories frequently used in international relations and to focus on the effects of specific domestic political institutions, it is important to control for a country's overall level of democracy (which I conceptualize in Schumpeterian fashion as the degree of openness in the formal rules governing political selection and competition). To that end, I control for the Polity score of each country-year. Because freedom of the press is often considered to be a *sine qua non* of liberal democracy, I verify that the coding criteria for the Polity index includes only measures capturing these formal regulations. I also control for whether a country is considered a major power in a given year under the coding criteria of Schultz (2001) as well as a country's share of the national capabilities of all states in the international system in a given year (Singer, Bremer and Stuckey 1972). Because protests and foreign conflict can be spurious to regime transitions, particularly for newly democratizing states (Mansfield and Snyder 2005), I control for recent regime transitions with a dichotomous variable that takes on a value of 1 if the country experienced a regime transition in the current or past three years<sup>41</sup>. Finally, I control for each country-year's population using data from the World Bank, given that states with larger population are likely to score higher on the unrest index and may engage in more conflictual behavior.

Given the data structure, in my main analyses I estimate linear probability models with country fixed-effects and robust standard errors to account for unobserved time-invariant confounders. The reported results thus represents a conservative estimate of the conditional effect of media freedom, insofar as the fixed-effects specification is likely to produce weaker findings.<sup>42</sup> A linear probability model is appropriate for the dichotomous DV, and avoids the

---

<sup>41</sup> I use Polity's definition of regime transition, which conceptually is a "substantive, normative change in political authority" and which empirical is defined as a three-point or greater change in the Polity score, with each continuous, sequential change in the same direction occurring within three years or less of the previous change.

<sup>42</sup> This is compounded by my use of a monadic conflict data set, which is more appropriate theoretically than dyadic data, but reduces the number of total observations in the sample.

dropping of countries that occurs under alternative fixed-effects specifications such as a logistic model due to the lack of dispute initiation by these countries.

## Results

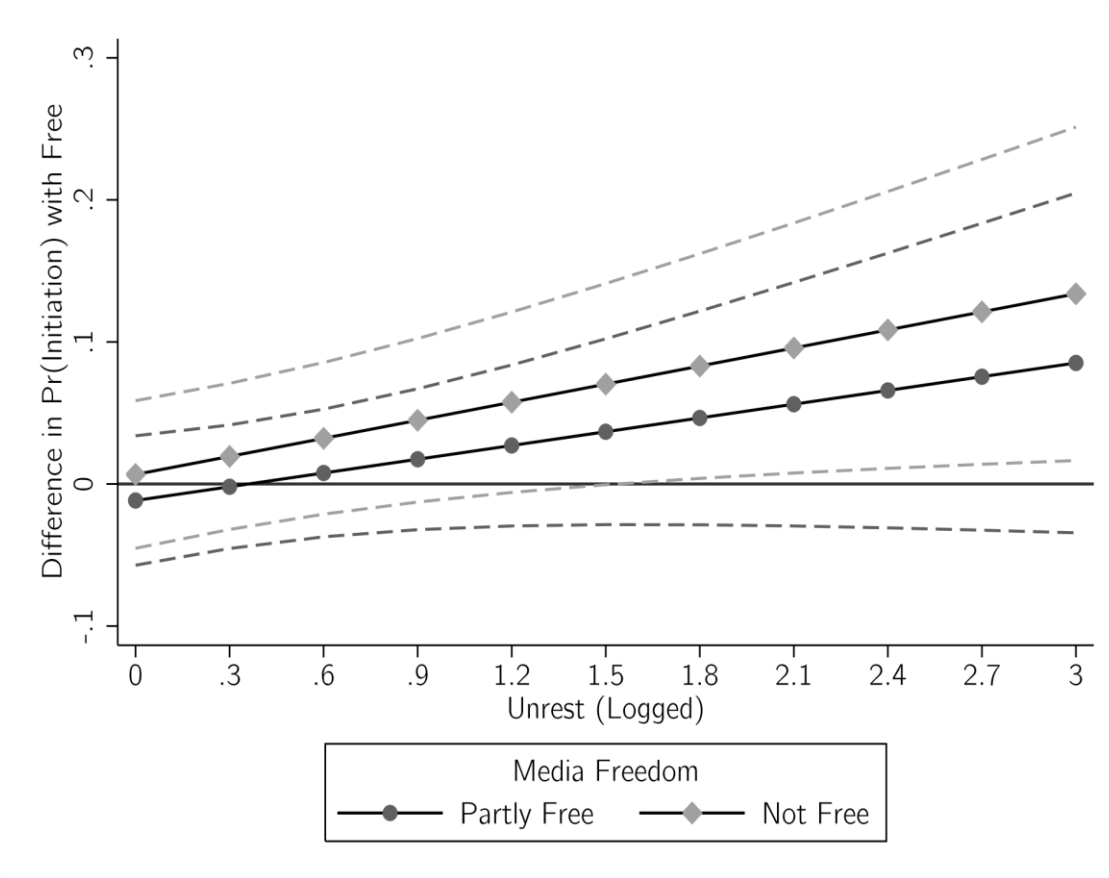
Table 3 presents the main results. In column one, I estimate a model that tests the traditional diversionary hypothesis, regressing conflict initiation on the unrest index. I fail to find evidence of any significant relationship between the two, with a point estimate that approximates zero. In column two, I include an interaction of the continuous unrest index and the three-category ordinal measure of media freedom, representing a test of H1. The first term displays the estimated impact of a one unit increase in the logged unrest variable among states with free media. Here, I find a small negative effect that is significant at  $p < .1$ , suggesting that states with free media appear *less likely* to initiate conflict as unrest increases. The second and third terms capture the conflict propensity of party free and not free states, relative to states with free media, when unrest is zero. The null findings here suggest that states with different media systems do not appear to have differing conflict propensities absent unrest. Finally, the fourth and fifth terms display the interaction effect between media freedom and domestic unrest. The results suggest that there is a significant difference in conflict initiation between free and not free states as unrest increases.

Table 3: Linear probability model, conflict on unrest and media freedom, with fixed-effects and robust standard errors

	(1)	(2)
	Diversionary	With Media Freedom
Unrest (Logged)	0.00 (0.01)	-0.03+ (0.02)
Media Freedom Partly Free		
Not Free		-0.01 (0.02) 0.01 (0.03)
Unrest x Media Freedom Partly Free		
Not Free		0.03 (0.02) 0.04* (0.02)
Polity Score	-0.00 (0.00)	-0.00 (0.00)
Major Power	0.31*** (0.01)	0.28*** (0.01)
National Capabilities	-0.16 (1.94)	0.56 (1.77)
Recent Regime Transition	0.01 (0.01)	0.01 (0.01)
Population	-0.00 (0.00)	-0.00 (0.00)
Constant	0.16***	0.16***

	(0.01)	(0.02)
<i>N</i>	6433	6418

Standard errors in parentheses  
<sup>+</sup>  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$



**Figure 8: Change in probability of initiation by unrest and media freedom**

Figure 8 displays the difference in the probability of initiating conflict between Free (the omitted category) and the other two types of states as unrest increases, with dotted lines denoting 95 per cent confidence intervals. The results show that Not Free states are significantly more likely to initiate conflict than Free states for values of the logged unrest index above approximately 1.7, which represents approximately 6 per cent of the country-year observations. While Partly Free states similarly see an increase vis-à-vis Free states as unrest increases, this difference remains statistically insignificant. However, I also estimate the marginal effect of a change in one unit of the logged unrest index on conflict by media freedom categories. Although the effect of unrest for Not Free states is estimated to be positive (.02), the effect does not reach statistical significance at conventional levels ( $p = .19$ ). Instead, I find some evidence that unrest has a *negative* effect on conflict initiation among Free states ( $-.03$ ,  $p = 0.08$ ).

As discussed above, because of the quality of the unrest data and the possibility that it systematically biases against finding a positive effect of unrest on conflict, particularly among Not Free states, the above result is in many ways unsurprising. Rather, it is tentative evidence that the logic of diversion operates differently for different states, and that a key reason for this difference lies in variation in domestic institutions governing the operation and autonomy of media entities.

### **Conflict intensity**

I conceptualize diversionary conflict as the manipulation of foreign policy salience for domestic dividends. If domestic salience manipulation is at least one motivation for conflict initiation, we should see greater probability of conflict initiation when domestic conditions are such as to threaten the leader's hold on power, such as periods of high domestic unrest. However, leaders should also wish to minimize the risks and costs associated with foreign conflicts. These include material costs associated with outright fighting and property destruction as well as audience costs associated with a leader appearing to back down from the conflict and conceding

to the adversary, and such costs increase monotonically as conflict intensity increases. At the same time, the benefits of diversion in the form of heightened foreign policy salience can be obtained, especially for leaders with political control over the media, at relatively low levels of conflict intensity, and further escalatory steps will have greater diminishing returns. Consequently, there is in theory a threshold of conflict intensity beyond which the costs of conflict outweigh the benefits. The result is an opportunity to conduct a ‘placebo test’ of the findings reported above: the significant conditional effect of media freedom on the relationship between unrest and conflict should be less likely to obtain in high-intensity conflicts, relative to low-intensity conflicts.

Table 4 presents the results of such a test. I generate four alternative specifications of the conflict initiation variable based on two different ways to operationalize conflict intensity. In the first two columns, I test whether media freedom conditions the relationship between unrest and conflict initiation among conflicts that resulted in fatalities (column one) and those that did not (column two). In the second two columns, I assess the conditional role of media freedom for conflicts that involved the use of force, including wars (column three), and those that involved only the threat or display of force (column four); I term these high- and low-hostility conflicts respectively, using the terminology from the MID dataset.

**Table 4: Unrest, media freedom, and conflict intensity**

	(1)	(2)	(3)	(4)
	Fatal	Non-Fatal	High-Hostility	Low-Hostility
Unrest (Logged)	0.00 (0.01)	-0.02+ (0.01)	-0.00 (0.01)	-0.04* (0.02)
Media Freedom Partly Free	-0.01+ (0.01)	0.00 (0.02)	-0.04+ (0.02)	0.02 (0.02)
Not Free	-0.01 (0.01)	0.02 (0.02)	0.01 (0.02)	0.00 (0.02)
Unrest x Media Freedom Partly Free	0.01 (0.01)	0.02 (0.02)	0.03 (0.02)	0.02 (0.02)
Not Free	0.01 (0.01)	0.03+ (0.02)	0.02 (0.02)	0.04* (0.02)
Polity Score	-0.00 (0.00)	.00 (0.00)	-0.00 (0.00)	0.00+ (0.00)
Major Power	0.01 (0.01)	0.28*** (0.01)	0.17*** (0.01)	0.20*** (0.02)
National Capabilities	-0.52 (0.52)	1.26 (2.28)	-2.03 (1.27)	5.28** (1.71)
Recent Regime Transition	0.00 (0.01)	.01 (0.01)	0.01 (0.01)	0.00 (0.01)
Population	-0.00* (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)
Constant	0.05***	0.13***	0.12***	0.04*



	(0.01)	(0.02)	(0.02)	(0.02)
<i>N</i>	6418	6418	6418	6418

Standard errors in parentheses  
+  $p < 0.10$ , \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Overall, the pattern of results is consistent with my theoretical expectations. The moderating effect of media freedom on the relationship between unrest and conflict initiation is significantly different between Free and Not Free states, as indicated by the significance of the interaction terms, but only for the initiation of non-fatal or low-hostility crises. I fail to find a significant difference between states with different levels of media freedom in the initiation of fatal and high-hostility MIDs. Furthermore, while high-intensity conflicts are rarer than low-intensity conflicts, the relative frequency of these two sets of disputes cannot explain all of the observed differences, given that the directionality of estimates between the odd and even columns are frequently different. Overall, these findings provide support for H3, suggesting that media systems are particularly influential in conditioning the relationship between unrest and low-intensity conflicts.

### **Robustness checks**

I assess the robustness of the main finding—reported in column two of Table 3—in a number of ways; the results of these robustness checks are reported in Appendix 4.2. In column one, I re-run the analysis without fixed effects. In column two and three, I try alternative specifications of the unrest variable, with the former using a measure that is both logged and lagged by one year, and the latter that is the unmodified index. In column four and five I include measures of alliance portfolio similarity with the leader of the international system, which is the United States for the entire temporal period<sup>43</sup>. In column six I use an alternative conflict initiation variable from the International Crisis Behavior dataset. In column seven I run the model using a simple count of the number of MIDs initiated in each country-year.

These results demonstrate that the conditional effect of media freedom remains robust in some, but not all, of the models. As shown in columns one through five, this result is robust to

---

<sup>43</sup> I include this measure here rather than in the main models due to the alliance data only being available until 2000.

the dropping of fixed-effects specification, to alternative specifications of the unrest variable, and to the inclusion of alliance portfolio similarity as an additional control. Indeed, in some cases, such as when using the unmodified unrest index, the results appear stronger and more precisely estimated than in the main specification reported above.

However, the result does not appear robust to alternative DVs. No significant interaction is seen when the dependent variable is a count of the number of MIDs initiated in a country-year, nor when using the definition of conflict from the ICB dataset. The former case is unsurprising, given the extreme skewness in this variable<sup>44</sup>; less than 4 per cent of the country-year observations feature more than one conflict, yet one observation contains 23 instances of conflict initiation. Indeed, this skewness is what motivated the dichotomous specification in the main analyses. Instances of crises in the ICB data set are comparatively rarer than instances of disputes in the MID dataset,<sup>45</sup> and this reduction in observable instances of the dependent variable under ICB definitions may contribute to the failure to find the same relationships observed throughout the other robustness checks.

## Discussion and conclusion

Recall the three questions I set out to answer in this chapter: (1) Does media freedom condition the relationship between unrest and conflict? (2) If (1) is true, is it because states without free media are more likely to initiate conflict as unrest increases? (3) If (1) is true, is the supporting evidence stronger for low-intensity conflicts? The results of my empirical analyses suggest support for (1) and (3), but not (2). I find evidence that variation in press freedom, *independent of regime type* (as defined by a regime's method of executive selection and recruitment), is an important moderator of the relationship between unrest and conflict in the post-war period. I also find evidence that this moderating role of media freedom is true only for low-intensity conflicts—exactly what we should see if leaders purposefully engage in low-level conflict

---

<sup>44</sup> Kurtosis = 183

<sup>45</sup> Curiously, the ICB codebook does not appear to include a definition of “crisis.”

during domestic downturns but are wary of the costs of war. However, I fail to find evidence that the importance of media freedom is driven by states without free media being particularly belligerent during times of high domestic unrest; the upward slope of unfree states is too gentle to reach statistical significance. Instead, I find some evidence suggesting that states with free media are *less likely* to engage in conflict as unrest increases.

Does this final result invalidate diversionary theory? Other studies have found stronger disconfirming findings without decisively settling the debate. Furthermore, there are reasons why the current analysis may struggle to uncover a true positive relationship for the unfree states. Diversion is one of many potential explanations for conflict—conflict data thus present a noisy environment to statistically tease out the impact of one potential causal factor. Additionally, the strengths of the unrest index used is that it facilitates comparison with prior studies in the diversionary literature, and that it is one of the few datasets attempting to capture levels of protests and demonstrations across the world and for a lengthy time period. The disadvantage is that true instances of unrest are severely under-reported, and given that these measures are based on news reports, under-reporting is almost certainly to be greater for states without free media. The result is a bias against finding a positive relationship between conflict and unrest among unfree states.<sup>46</sup>

In addition, my theory may over-predict the foreign policy behavior of a number of unfree states. There are a number of states who maintain isolationist foreign policies and avoid interactions, conflictual or otherwise, with other actors internationally. If having a closed internal political system makes a regime more likely to be closed externally as well, it may be the case that the current “not free” category conceals significant heterogeneity in conflict propensity. Conversely, further disaggregating this category may improve the precision with which I can estimate the relationship between unrest and conflict for the not free states that are of theoretical interest. At the very least, the mixed success of the empirical tests using cross-national time

---

<sup>46</sup> Note, however, that the unrest data is also more likely to pick up the most severe and visible instances of domestic unrest, which are also the most theoretically appropriate unrest events, insofar as they are most likely to threaten a leader’s hold on power.

series data suggest the importance of more closely examining one polity in particular to see if evidence of diversionary dynamics exist.

Furthermore, my specification uses country fixed effects, which implies that the observed results are primarily driven by cases in which media freedom varies over time. This further adds to the difficulty in observing significant findings, insofar as it shrinks the number of states that can drive the predicted relationship. It also raises additional questions concerning the characteristics of the set of countries that experience change in their media systems. On the one hand, countries may experience such changes when they experience regime change overall; these scenarios may imply a correlation between unrest and conflict that is due purely to the belligerence of newly formed regimes and is spurious to any diversionary motivations. On the other hand, these types of conflicts are particularly likely to afflict states that have transitioned into democracy (Mansfield and Snyder 2005). Insofar as democracy and free media covary, this suggest a bias in favor of an apparent diversionary relationship that is stronger for states with free media, contrary to the pattern observed in the present study.

Finally, the empirical analysis is focused on a specific time period, much of it during the Cold War and an era of bipolarity. Given Cold War dynamics, in which conflicts outside of the two major super powers may escalate and spiral, we should expect particularly states to be particularly conflict averse, thus adding to the difficulty with establishing clear evidence on the factors that predict conflict initiation.

Nevertheless, the results provide evidence of the important moderating role of media systems. Though not in exactly the predicted fashion, the findings in this chapter do support the claim that domestic political institutions can structure the flow of information, with important ramifications for our understanding of international relations.

## References

- Abramson, Alana. 2017. "Here's Exactly what Jeff Sessions Said about Russia at His Confirmation Hearing." *Time* .
- Baggott, Erin. 2016. "Diversionary Cheap Talk: Domestic Discontent and US Foreign Policy." .
- Baker, William D. and John R. Oneal. 2001. "Patriotism or Opinion Leadership? The Nature and Origins of the 'Rally 'Round the Flag' Effect." *Journal of Conflict Resolution* 45(5):661–687.
- Baum, Matthew A. 2002. ""The Constituent Foundations of the Rally-Round-the-Flag Phenomenon." *International Studies Quarterly* 46(2):263–298.
- Baum, Matthew A. and Tim Groeling. 2008. *War Stories: The Causes and Consequences of Public Views of War*. Princeton: Princeton University Press.
- Bennett, D. Scott and Allan C. Stam. 2004. *The Behavioral Origins of War*. Ann Arbor: University of Michigan Press.
- Bershidsky, Leonid. 2017. "Trump's Forever War of Diversion:." *Bloomberg* .
- Blainey, Geoffrey. 1973. *The Causes of War*. London: MacMillan Press.
- Cheney, Kyle. 2017. "Lawmakers Stunned, Baffled by Trump's Wiretap Allegations." *Politico*.
- Chiozza, Giacomo and Hein E. Goemans. 2003. "Peace Through Insecurity: Tenure and International Conflict." *Journal of Conflict Resolution* 47(4):443–467.
- Coser, Lewis A. 1956. *The Functions of Social Conflict*. New York: Routledge Press.
- Crawford, Neta C. 2017. "United States Budgetary Costs of Post-9/11 Wars Through FY2018: A Summary of the \$5.6 Trillion in Costs for the US Wars in Iraq, Syria, Afghanistan and Pakistan, and Post-9/11 Veterans Care and Homeland Security." *Costs of War*.
- Dassel, Kurt and Eric Reinhardt. 1999. "Domestic Strife and the Initiation of Violence at Home and Abroad." *American Journal of Political Science* 43:56–85.
- Davies, Graeme A. M. 2002. "Domestic Strife and the Initiation of International Conflicts: A Directed Dyad Analysis, 1950-1982." *Journal of Conflict Resolution* 46:672–692.

Department of Defense. 2019. "Casualty Status as of 10 a.m. EDT July 3, 2019."

**URL:** <https://dod.defense.gov/News/Casualty-Status/>

Entous, Adam, Ellen Nakashima and Greg Miller. 2017. "Sessions Met with Russian Envoy Twice Last Year; Encounters He Later Did Not Disclose." *Washington Post* .

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

Fischer, Fritz. 1975. *War of Illusions: German Policies from 1911 to 1914*. Chatto & Windus Press.

Foster, Dennis. 2016. "Would President Trump Go to War to Divert Attention from Problems at Home?" *Washington Post* .

Foster, Dennis M. and Jonathan W. Keller. 2014. "Leaders' Cognitive Complexity, Distrust, and the Diversionary Use of Force." *Foreign Policy Analysis* .

Fravel, Taylor. 2010. "The Limits of Diversion: Rethinking Internal and External Conflict." *Security Studies* 19(2).

Gaubatz, Kurt Taylor. 1991. "Election Cycles and War." *Journal of Conflict Resolution* 35(2):212–244.

Gelpi, Christopher. 1997. "Democratic Diversions: Governmental Structure and the Externalization of Domestic Conflict." *Journal of Conflict Resolution* 41(2):255–282.

Ghosn, Faten, Glenn Palmer and Stuart A. Bremer. 2004. "The MID3 Data Set, 1993-2001: Procedures, Coding Rules, and Description." *Conflict Management and Peace Science* 21(2):133–154.

Gilpin, Robert. 1981. *War and Change in International Politics*. Cambridge: Cambridge University Press.

Goemans, H. E. 2000. *War and Punishment: The Causes of War Termination and the First World War*. Princeton: Princeton University Press.

Graber, Doris and Johanna L. Dunaway. 2015. *Mass Media and American Politics*. CQ Press.

- Groeling, Tim and Matthew Baum. 2008. "Crossing the Water's Edge: Elite Rhetoric, Media Coverage, and the Rally-Round-the-Flag Phenomenon." *Journal of Politics* 70(4):1065–1085.
- Hallin, D. C. and P. Mancini. 2004. *Comparing Media Systems: Three Models of Media and Politics*. Cambridge: Cambridge University Press.
- Haynes, Kyle. 2017. "Diversionary Conflict: Demonizing Enemies or Demonstrating Competence?" *Conflict Management and Peace Science* 34(4):337–358.
- Hetherington, Marc J. and Michael Nelson. 2003. "Anatomy of a Rally Effect: George W. Bush and the War on Terrorism." *PS: Political Science and Politics* 36(1):37–42.
- Hou, Yue and Kai Quek. 2015. "The Effects of Terror Attacks on Attitudes in China.".
- Jones, Daniel M., Stuart A. Bremer and J. David Singer. 1996. "Militarized Interstate Disputes, 1816-1992: Rationale, Coding Rules, and Empirical Patterns." *Conflict Management and Peace Science* 15(2):163–215.
- Kam, Cindy D. and Jennifer M. Ramos. 2008. "Joining and Leaving the Rally: Understanding the Surge and Decline in Presidential Approval Following 9/11." *Public Opinion Quarterly* 72(4):619–650.
- Lai, Brian and Dan Reiter. 2005. "Rally 'Round the Union Jack? Public Opinion and the Use of Force in the United Kingdom, 1948-2001." *International Studies Quarterly* 49:255–272.
- Lai, Brian and Dan Slater. 2006. "Institutions of the Offensive: Domestic Sources of Dispute Initiation in Authoritarian Regimes, 1950-1992." *American Journal of Political Science* 50(1):113–126.
- Levy, Jack S. 1989. Diversionary Theory of War: A Critique. In *Handbook of War Studies*, ed. Manus I. Midlarsky. Routledge pp. 259–288.
- Lian, Bradley and John R. Oneal. 1993. "Presidents, the Use of Military Force, and Public Opinion." *Journal of Conflict Resolution* 37(2):277–300.
- Lustick, Ian. 1996. "History, Historiography, and Political Science: Multiple Historical



- Records and the Problem of Selection Bias." *American Political Science Review* 90(3):605– 618.
- Mansfield, Edward D. and Jack Snyder. 2005. *Electing to Fight: Why Emerging Democracies Go to War*. Cambridge: MIT University Press.
- Mayer, Arno J. 1969. "Internal Causes and Purposes of War in Europe, 1870-1956: A Research Assignment." *Journal of Modern History* 41(3):291–303.
- Merolla, Jennifer L. and Elizabeth J. Zechmeister. 2009. *Democracy at Risk: How Terrorist Threats Affect the Public*. Chicago: University of Chicago Press.
- Mueller, John. 1973. *War, Presidents, and Public Opinion*. New York: Wiley & Sons.
- Oakes, Amy. 2006. "Diversionary War and Argentina's Invasion of the Falkland Islands." *Security Studies* 15(3):431–463.
- Oakes, Amy. 2012. *Diversionary War: Domestic Unrest and International Conflict*. Stanford: Stanford University Press.
- Oneal, John R. and Anna Lilian Bryan. 1995. "The Rally 'Round the Flag Effect in US Foreign Policy Crises, 1950-1985." *Political Behavior* 17:379–402.
- Oneal, John R. and Jaroslav Tir. 2006. "Does the Diversionary Use of Force Threaten The Democratic Peace? Assessing the Effect of Economic Growth on Interstate Conflict." *International Studies Quarterly* .
- Organski, A. F. K. and Jacek Kugler. 1980. *The War Ledger*. Chicago: University of Chicago Press.
- Richards, Diana T., Clifton Morgan, Rick K. Wilson, Valerie L. Schwebach and Garry D. Young. 1993. "Good Times, Bad Times, and the Diversionary Use of Force: A Tale of Some Not-So-Free Agents." *Journal of Conflict Resolution* 37(3):504–535.
- Russett, Bruce. 1990. *Controlling the Sword: The Democratic Governance of National Security*. Cambridge: Harvard University Press.
- Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy*. New York: Cambridge University Press.

- Shoemaker, Pamela J. and Timothy Vos. 2009. *Gatekeeping Theory*. Taylor Francis.
- Simmel, Georg. 1898. "The Persistence of Social Groups." *American Journal of Sociology* 3(5):662–698.
- Singer, David J., Stuart Bremer and John Stuckey. 1972. Capability Distribution, Uncertainty, and Major Power War, 1820-1965. In *Peace, War, and Numbers*, ed. Bruce Russett. Sage.
- Smith, Alastair. 1996. "Diversionary Foreign Policy in Democratic Systems." *International Studies Quarterly* 40(1):133–153.
- Soroka, Stuart N. 2014. *Negativity in Democratic Politics*. Cambridge: Cambridge University Press.
- Sprecher, Christopher and Karl DeRouen. 2002. "Israeli Military Actions and Internationalization-Externalization Processes." *Journal of Conflict Resolution* 46(2):244– 259.
- Tajfel, Henri and John Turner. 1979. An Integrative Theory of Intergroup Conflict. In *The Social Psychology of Intergroup Relations*, ed. William G. Austin and Stephen Worchel. pp. 33–47.
- Taylor, A. J. P. 1954. *The Struggle for Mastery in Europe, 1848-1918*. London: Clarendon Press.
- Taylor, A. J. P. 1961. *The Origins of the Second World War*. Simon Schuster.
- Tir, Jaroslav and Michael Jasinski. 2008. "Domestic-Level Diversionary Theory of War: Targeting Ethnic Minorities." *Journal of Conflict Resolution* 52(5):641–664.
- Tyler, Patric E. 1992. "U.S. Said to Plain Raids on Baghdad Over Inspections." *New York Times* p. 1.
- Waltz, Kenneth N. 1988. "The Origins of War in Neorealist Theory." *The Journal of Interdisciplinary History* 18(4):615–628.
- Weisiger, Alex. 2013. *Logics of War*. Ithaca: Cornell University Press.
- Whitten-Woodring, Jennifer and Douglas A. Van Belle. 2017. "The Correlates of Media Freedom: An Introduction of the Global Media Freedom Dataset." *Political Science Research and Methods* 5(1):179–188.

Wolf, Byron Z. 2017. "Trump's Playbook: Go On Offense with a New Conspiracy Theory."  
*CNN*.

Wolffe, Richard. 2017. "Out of Control? Or is Trump's Tweeting Designed to Distract?" *The Guardian*.

## CHAPTER 5: Conclusion

In chapter two, I intervene in debates surrounding the existence of audience costs—specifically, on the question of whether targets of costly signals recognize such signals as carrying costs. If results demonstrate that leader responses to costly threats differ systematically from threats that do not carry such costs, the findings would provide evidence consistent with and in support of audience cost theory. If results do not find such systematic differences in responses across threats that vary in costs, the results would be difficult for audience cost theory to explain, and bolster the theory's many critics.

In the paper, I exploit the electoral cycle among democratic threat initiators to generate within-country variation in threat costliness. I argue that threats made in election years carry greater costs for the leader than those made in non-election years. By analyzing how effective each of these two types of threats are, I find mixed evidence that speaks to both proponents and critics of audience cost. Costly threats by democracies are more effective against other democracies, consistent with democratic targets recognizing the costs associated with such threats—the same is not true for autocratic targets. In uncovering the surprising heterogeneity of threat effectiveness across target regime type, the paper suggests one important way that domestic politics and information in international politics interact: differences in domestic political institutions can explain how the same information—in this case, costs associated with election-year threats—can be effective in some cases and fail in others.

There are important limitations to the paper. Because the key concept at play—leader beliefs about the credibility of signals—is fundamentally unobservable, researchers must instead test second-order hypotheses that are consistent implications of the theory. While the research design used in the paper is based on canonical observational studies of audience costs (Schultz 2001; Weeks 2008), there are two basic concerns that a reader may have. First, the claim that election-year threats are costlier than those made in non-election years is one that is defended and supported by extant literature (Chiozza 2017; Gaubatz 1991), but fundamentally untested

and unmeasured in this study. Second, even if one assumes that costs vary with the electoral cycle, the study is unable to isolate that costs are the *only* factor that vary across the two categories of threats.

In my view, both fundamental limitations stem from the intractable difficulty that researchers face in directly measuring costs associated with interstate signals. Because such costs are unobservable, one has to make assumptions about conditions under which costs should vary. Yet this approach prevents researchers from methodologically manipulating costs in isolation from other factors. To the extent that such limitations are inherent to the empirical literature on audience costs, I suggest that the paper still makes important theoretical contributions. In particular, heterogeneous effectiveness across target regime type empirically demonstrates how domestic politics can condition the role and effectiveness of information in international politics.

Chapter three identifies and attempts to address a lacuna in the literatures on diversionary war and rally 'round the flag effects. Specifically, diversionary theory assumes that leaders can successfully elicit rally effects by initiating conflicts abroad. However, the rally literature has little to say on how initiation status affects subsequent rallies. I suggest that the key difficulty that the rally literature has faced on this question stems from its reliance on observational data. Because conflicts in the historical record differ along many dimensions, among which initiation status is only one; and because there are a finite number of historical conflicts to analyze, observational data is poorly suited towards isolating the causal impact of a specific conflict characteristic. Consequently, observational studies provide little traction on questions about whether conflicts that are initiated produce similar rallies to those that are started by the other side.

In the paper, I further argue that experimental methods are well-suited to addressing the aforementioned methodological challenges. I administer a survey experiment, where the treatments experimentally manipulate the initiation status of a hypothetical conflict between China and the US, thus isolating the impact of initiation on public opinion. The results suggest

heterogeneous rallies across partisan identification—a primer driver of public opinion on political issues, and a key lens through which political information is viewed by voters. Consistent with the rally literature, I find that a reciprocated conflict raises approval of the president among the full sample. However, conflicts that are initiated by the president produces rally effects only among voters who share a partisan affiliation with the president. This in-partisan rally effect is unique and surprising in the literature on rally effects (Baum 2002). More generally, the results demonstrate how domestic politics can explain why the same information produces different effects among different audiences.

The study contains one fundamental limitation: I do not experimentally manipulate the partisan identification of the president. Thus, the study is unable to adjudicate between two plausible explanations of the results. It may be the case that in-partisans are always willing to rally to their leader. It may also be the case that Republican voters are uniquely hawkish, and that is why one observes rally effects among them regardless of initiation status. A natural extension of this project, then, is to experimentally manipulate the partisanship of the president as well. If the preferred interpretation regarding in-partisan rallies is correct, we should see Democrats exhibiting the same willingness to rally, under a Democratic president, as this chapter observes for Republicans. If instead the alternative explanation focusing on the covariation of foreign policy preferences with partisan identification is correct, we should continue to see Republicans being more willing to rally, even under a Democratic president.

In chapter four, I advance the literature on diversionary war in two ways. First, I argue that diversion is fundamentally *the manipulation of foreign policy salience*, and that this objective can be achieved by actions short of war. Given the high costs of waging war, diversion is unlikely to be the primary motive for any high-intensity interstate conflict—this is particularly true if diversion can be bought for the cheap. Second, I argue that domestic political institutions constrain the ability of leaders to reap the rewards of diversion in critical ways. In particular, leaders under a free media system face structural barriers in their ability to generate salient, positive news coverage of low-intensity foreign policy actions due to the autonomy of journalists

in such systems. On the other hand, leaders who exercise political control over media institutions can mandate high levels of positive coverage for foreign policy when they desire, thus enhancing their ability to benefit from diversionary policy.

I test predictions from this theory using a data set of international conflict, domestic unrest, and media freedom. The results provide tentative support, insofar as media freedom conditions the relationship between unrest and conflict. However, while theory predicts that this conditionality should take the form of not free states exhibiting a stronger diversionary relationship, I find stronger evidence supporting the inverse: that free states are *less likely* to initiate conflict when unrest increases.

The study in chapter four has a number of limitations. Because diversion is just one of many potential explanations for conflict, and because it is likely to be present in conjunction with other explanations in the case of any single given event, conflict data presents a particularly noisy data environment in which to look for diversionary effects. Furthermore, the unrest data used in the study relies on media reports, and by construction it is almost certain to underreport domestic turmoil in states that lack free media. The implication of such systematic underreporting would be to weaken any real relationship between unrest and conflict for such states. However, the empirical results of chapter four does demonstrate that the observed interaction between media system, unrest, and conflict is stronger for low-intensity conflicts, consistent with theoretical expectations. More broadly, the evidence demonstrates that media systems condition the relationship between domestic politics and foreign policy in critical ways. To that end, the study provides one example of how domestic politics can constrain and structure the role of information in international politics.

## **Common linkages**

One common thread running throughout the three studies is the important role played by regime type heterogeneity. In doing so, the studies build on a large body of work across literatures on audience cost, diversionary war, and rally effects. Domestic political institutions

matter, not only by changing the size of selectorates, but in fundamentally structuring the flow of information between and within states. Domestic institutions, in other words, affect what information is transmitted; the form that that information takes; who receives what pieces of information; and how different audiences and stakeholders react once information is received.

A second commonality between the studies is the critical role played by the mass media. To borrow from Lippmann (1922), there is the world outside, and there is the image of the world in our heads—for the mass public, media is the critical channel through which they learn about the world of international politics. From that perspective, the relative sparse treatment of media institutions and media effects by scholars of international relations is somewhat surprising. The third chapter alludes to the important role that media entities have in shaping narratives and perceptions about conflict. In the context of survey experiments, researchers have the ability to manipulate perceptions of initiation—in the real world, individual journalists and editors fulfill the exact same function. Although leaders enjoy an information advantage vis-a-vis the media at conflict onset (Groeling and Baum 2008), actors in the media play a critical role in deciding what information to communicate to the mass public, and how, throughout the entire duration of a given conflict. At a more general level, chapter four demonstrates that media systems—that is, the relationship between media entities and actors of political power—play important roles in constraining the relationship between domestic politics and foreign policy. Indeed, if autonomous media entities can help reduce incentives to divert during times of unrest, then the Fourth Estate may not only be a safeguard of democracy, but also a safeguard of pacifism—a claim with sobering implications for the present day.

In addition, the three studies all demonstrate how domestic politics can pressure and affect leader incentives in myriad and complex ways. At a basic level, the manifestation of the public's political power during election years can affect the credibility attached to leader threats and statements. Furthermore, public opinion can incentivize leaders to pursue foreign policy bellicosity, though not in clear and deterministic ways. Rather, leaders who contemplate diversionary action must consider whether her institutional environment is conducive to rally



benefits, and if so, from which subset of the population is she most likely to see the greatest dividends. In conjunction, chapters three and four suggest that the benefits to diversion for a democratic leader is likely to be small, and that in general leaders of democracies are ill-served by contemplating foreign aggression as a way to address domestic unpopularity.

Chapter four argues that leaders with political control over the media can more reliably reap the rewards of diversion. This claim, if true, may represent a possible explanation for the heterogeneous effects found in chapter two. The results in that paper suggest that democracies are less likely to respond to threats made by an initiator in their election year (in other words, costly threats) while autocrats do not appear to differentiate between election-year and non-election year threats. This may be because autocrats interpret election-year threats as instantiations of a diversionary foreign policy by the initiator. This can then lead the autocrat to downplay the credibility of the threat, raising the probability of reciprocation. Democratic targets, on the other hand, are less prone to diversionary motivations given the constraints of a free media system—they may correspondingly be less likely to attribute such motivations to democratic initiators, and treat election-year threats as credible signals of resolve.

Chapter three raises a different but equally intriguing explanation for the findings in chapter two. In the former, I find that initiated conflicts generate in-party rallies, but reciprocated conflicts generate rallies among the general sample—in other words, reciprocated conflicts generate bigger and more widespread rally effects than initiated conflicts. Bringing this insight into chapter two suggests the possibility that the decision to reciprocate to a threat represents a tradeoff between a boost in domestic popularity and a risk of escalation into conflict. When viewed from such a perspective, a possible explanation emerges for the differences in reciprocation propensity between democratic and autocratic targets. For democratic leaders, domestic support is important—all leaders wish to stay in power—but the loss of power for democrats is less punishing than it is for autocrats; there are many instances, therefore, in which the risk of escalation outweighs the risk of losing political power. If true, we would expect democratic leaders to be particularly likely to be wary of responding to credible, election-year

threats. Autocrats, on the other hand, may face severe punishment or even death should they lose political power—Gaddafi is the most recent and prominent example of this dynamic. Autocrats, therefore, should be particularly attracted to the possible rally effects engendered by reciprocation to a foreign threat. In short, autocrats should be more willing to reciprocate more threats than democrats, because they care about domestic rally effects more—there is in theory a subset of costly and credible threats that would lead a rational democratic leader, calculating the risks and rewards, to back down, and a similarly rational autocratic leader to reciprocate. This subset of threats may be driving the heterogeneous reciprocation rates observed in chapter two between democratic and autocratic targets.

Chapter four also makes the case that IR scholarship should move beyond regime type towards understanding specific institutional arrangements, such as media systems. How might such a perspective change the approach taken in chapter two? In that paper, I adopt the democracy-autocracy distinction in order to more directly build on and advance the audience cost literature that explores how domestic political structures condition bargaining effectiveness (Fearon 1994). However, democracy/autocracy is really a dichotomous operationalization of the underlying concept—political strength of domestic audiences—under the assumption that democracies have stronger domestic audiences than autocracies. “Domestic audience strength” could be conceptualized continuously, however, and indeed previous work has shown the importance of disaggregating within autocracies (Weeks 2008). Alternatively, media freedom can be just a plausible proxy for the strength of the domestic audience as regime type is. When a leader exercises political control over the media, they can mask media coverage of their decision to back down, preventing punishment and the levying of audience costs by their domestic audiences. Leaders under a free media system, in contrast, are those that can reliably expect to be punished for backing down. Viewed from this perspective, it is not surprising that audience cost theory was developed and remains prominent in a country with free media; *that the public actually needs to learn about their leader backing down in order to punish them for backing down* is an implicit assumption of the theory that should have always been explicit. A future replication

of chapter two, adopting the approach advocated in chapter four, would examine reciprocation rates of election and non-election year threats between states with free and not free media systems, rather than regime type.

Finally, chapter three and chapter four together approach diversion from different perspectives but reach complementary conclusions. Chapter four argues that “diversion can be bought for the cheap”—that is to say, that leaders can successfully achieve the benefits traditionally articulated by scholars of diversionary war without necessarily engaging in war, simply through the initiation of actions short of war that nevertheless manipulate the foreign policy salience of the public. In chapter three, subjects in treatment groups of the experimental design are necessarily treated to a high dosage of foreign policy salience, and corresponding rally effects are observed. Thus, chapter four provides indirect macro-level evidence for the articulated, broader version of diversionary theory, while chapter three provides evidence that lends support to the empirical microfoundations of this approach.

## **Implications**

Cumulatively, the three papers present important implications for how to understand the ongoing US-China trade dispute. Differences in domestic political institutions and context between the two states may exacerbate issues with miscommunication and interstate signaling. Paper one demonstrates how autocrats do not appear to systematically differentiate between credible and not credible signals sent by democracies; it is possible that similar dynamics explain apparent miscalculations on the Chinese side over the resolve of the US and President Trump during the early stages of the dispute.

At the same time, differences in media systems between the two sides results in China enjoying three distinct advantages vis-à-vis the US. First, political control over the media means that Chinese leaders can minimize the pain of audience costs should they back down from a previous position by limiting news information about the event. Thus, China enjoys the flexibility

and manoeuvrability that is coveted by all leaders (Trachtenberg 2012). Second, political control also allows China to more reliably reap any rally effects that may result from the ongoing dispute, by carefully tailoring the timing and tone of coverage about the trade conflict. The existence of this benefit thus slightly reduces the incentive and need for China to reach a bargaining settlement. Third, the relative openness of the media environment in the US can be exploited by China. For example, China tailored its first round of tariffs on US exports from states that are predominantly Republican. While the direct effect of these tariffs may only operate on a small number of farmers, China benefitted from widespread media coverage in the US about the effects of the tariffs. The US media, in effect, became a force multiplier of the effect that China desired to produce on American public opinion through its enactment of the tariffs.

Most of the public discourse surrounding the duration of the trade war focuses on each side's endurance and tolerance for the economic pain of tariffs. As the three papers in this dissertation demonstrate, however, pain and costs only represent one side of the equation—leaders may also reap domestic benefits from foreign policy conflicts. When considering the Chinese side, resolution or settlement is most likely when the costs incurred by US punitive actions exceeds the benefits that China gains from maintaining the salience of the US threat among the domestic Chinese population. It is possible that this threshold was achieved by US targeted actions against ZTE and Huawei. The particular acute costs associated with US bans against these two companies in particular may be due to the political influence enjoyed by the leadership at the two firms. It may also be driven by the prominence and visibility of those two companies within China, which may make it difficult for the Chinese leadership to control coverage about events affecting those two companies. The particular visibility of US threats against these two firms may have made it particularly difficult for the Chinese side to accept costs in this case.

## **Areas for future research**

Based on these common linkages, the dissertation suggests three general avenues for future scholarship to explore. First, the importance of regime type differences is now generally appreciated by international relations scholars, but future work should unpack the black box of regime type. Chapter four attempts to do this by focusing on media systems, a particular domestic political institution that, while covarying with regime type, is more causally proximate and relevant for the theory in question. In many instances, particular institutional arrangements are doing the heavy lifting in international relations theory, but only implicitly. Moving beyond regime will allow future scholars to more clearly identify relationships between domestic politics and international relations; test and validate existing theories; and generate new ideas on how particular institutional arrangements affect the interplay of world politics.

Second, the importance of the media as an actor with agency in international affairs is ripe for further scholarship. Challenges with measurement and design are no more considerable for analysis of the media as with other concepts in international relations. With the recent trend towards understanding microfoundations in IR, and given the media's role as the information conduit connecting the mass public to world affairs, unpacking the effects and consequences of the media on international relations appears to be a logical direction for future scholars to pursue.

Third, fundamental questions remain surrounding how domestic politics structure the incentives of leaders. Why do we observe autocracies responding to costly threats differently than democracies? What are the implications of empirically disaggregating the mass public for theories of international relations that rely on a unitary and compliant domestic audience? At a high level, this dissertation makes the case that domestic politics interact with information, with important ramifications for our understanding of international politics and interstate behaviour. A number of research agendas and questions follow from adopting such a perspective, and each provide fertile ground for future scholars to explore.

## References

- Baum, Matthew A. 2002. "The Constituent Foundations of the Rally-Round-the-Flag Phenomenon." *International Studies Quarterly* 46(2):263–298.
- Chiozza, Giacomo. 2017. "Presidents on the Cycle: Elections, Audience Costs, and Coercive Diplomacy." *Conflict Management and Peace Science* 34(1):3–26.
- Gaubatz, Kurt Taylor. 1991. "Election Cycles and War." *Journal of Conflict Resolution* 35(2):212–244.
- Groeling, Tim and Matthew Baum. 2008. "Crossing the Water's Edge: Elite Rhetoric, Media Coverage, and the Rally-Round-the-Flag Phenomenon." *Journal of Politics* 70(4):1065–1085.
- Lippmann, Walter. 1922. *Public Opinion*.
- Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy*. New York: Cambridge University Press.
- Weeks, Jessica P. 2008. "Autocratic Audience Costs: Regime Type and Signaling Resolve." *International Organization* 62(1):35–64.

## APPENDIX

### Appendix 2.1: Percentage and number of militarized compellent threats met with compliance

	All	No Election	Election
All	42.9 (56)	40.9 (44)	50.0 (12)
Democracy	46.2 (13)	41.7 (12)	100.0 (1)
Autocracy	41.9 (43)	40.6 (32)	45.5 (11)

Cells display percentage of threats in each category that are met with compliance, with raw totals below in brackets.

## Appendix 2.2: Table of robustness checks

	Unadjusted SE			Alt Recip			No Manual Regime			No World Wars			With Alliances		
	(1) Full	(2) Aut	(3) Dem	(4) Full	(5) Aut	(6) Dem	(7) Full	(8) Aut	(9) Dem	(10) Full	(11) Aut	(12) Dem	(13) Full	(14) Aut	(15) Dem
Election	-0.04 (0.22)	0.21 (0.27)	-0.88 <sup>†</sup> (0.47)	0.03 (0.20)	0.28 (0.21)	-0.76 <sup>†</sup> (0.44)	-0.17 (0.22)	0.07 (0.26)	-0.97 <sup>*</sup> (0.47)	-0.12 (0.18)	0.11 (0.23)	-0.89 <sup>*</sup> (0.45)	-0.04 (0.21)	0.24 (0.25)	-0.88 <sup>*</sup> (0.45)
Major-Major	0.24 (0.36)	0.40 (0.45)	0.08 (0.65)	0.21 (0.37)	0.43 (0.40)	-0.04 (0.66)	0.60 <sup>†</sup> (0.34)	1.34 <sup>***</sup> (0.41)	0.14 (0.70)	-0.15 (0.51)	-0.25 (0.43)	0.21 (0.70)	-0.12 (0.49)	0.57 (0.56)	-0.58 (1.07)
Major-Minor	0.13 (0.29)	0.28 (0.37)	0.09 (0.52)	0.09 (0.33)	0.25 (0.41)	0.00 (0.44)	0.13 (0.29)	0.39 (0.37)	0.33 (0.45)	0.27 (0.30)	0.39 (0.39)	0.23 (0.45)	-0.01 (0.39)	0.69 (0.46)	-0.81 <sup>†</sup> (0.46)
Minor-Major	0.33 (0.39)	0.53 (0.50)	0.06 (0.72)	0.33 (0.38)	0.63 (0.52)	-0.13 (0.72)	0.29 (0.36)	0.59 (0.54)	-0.11 (0.70)	0.25 (0.42)	0.58 (0.60)	-0.21 (0.81)	0.05 (0.42)	0.15 (0.55)	-0.27 (0.87)
Share of Capabilities	0.58 (0.42)	0.64 (0.53)	0.06 (0.76)	0.63 <sup>†</sup> (0.36)	0.81 <sup>*</sup> (0.39)	-0.07 (0.74)	0.42 (0.37)	0.52 (0.51)	-0.44 (0.71)	0.62 (0.40)	0.67 (0.45)	0.11 (0.74)	0.42 (0.35)	0.44 (0.46)	-0.03 (0.78)
Contiguity	0.72 <sup>***</sup> (0.23)	0.93 <sup>***</sup> (0.29)	0.81 <sup>†</sup> (0.47)	0.73 <sup>***</sup> (0.18)	0.93 <sup>***</sup> (0.22)	0.83 <sup>†</sup> (0.44)	0.66 <sup>***</sup> (0.17)	1.06 <sup>***</sup> (0.20)	0.75 <sup>†</sup> (0.45)	0.63 <sup>***</sup> (0.23)	0.80 <sup>*</sup> (0.32)	0.80 <sup>*</sup> (0.39)	0.56 <sup>*</sup> (0.27)	1.07 <sup>***</sup> (0.25)	-0.31 (0.51)
Territory (Issue)	0.95 <sup>***</sup> (0.29)	1.23 <sup>***</sup> (0.37)	0.45 (0.49)	0.93 <sup>***</sup> (0.25)	1.24 <sup>***</sup> (0.33)	0.37 (0.45)	0.73 <sup>***</sup> (0.25)	0.93 <sup>***</sup> (0.31)	0.48 (0.48)	0.93 <sup>***</sup> (0.25)	1.15 <sup>***</sup> (0.35)	0.44 (0.44)	0.99 <sup>***</sup> (0.28)	1.35 <sup>***</sup> (0.37)	0.26 (0.52)
Government (Issue)	0.21 (0.56)	0.19 (0.62)	0.85 (1.49)	0.20 (0.49)	0.18 (0.54)	0.62 (1.53)	0.29 (0.59)	-0.00 (0.55)	0.00 (.)	-0.04 (0.46)	-0.12 (0.46)	0.84 (1.47)	0.13 (0.60)	-0.03 (0.81)	0.05 (1.54)
Policy (Issue)	-0.97 <sup>***</sup> (0.22)	-0.77 <sup>***</sup> (0.27)	-1.47 <sup>***</sup> (0.43)	-0.95 <sup>***</sup> (0.20)	-0.78 <sup>***</sup> (0.26)	-1.41 <sup>***</sup> (0.40)	-1.16 <sup>***</sup> (0.24)	-1.07 <sup>***</sup> (0.37)	-1.40 <sup>***</sup> (0.40)	-1.05 <sup>***</sup> (0.22)	-0.85 <sup>***</sup> (0.31)	-1.53 <sup>***</sup> (0.40)	-1.01 <sup>***</sup> (0.20)	-0.88 <sup>***</sup> (0.23)	-1.70 <sup>***</sup> (0.48)
Other (Issue)	-0.88 (1.13)	-0.07 (1.25)		-0.90 (0.67)	-0.10 (0.89)		-0.99 (0.72)	-0.20 (0.99)		-1.04 (0.75)	-0.30 (1.03)		-0.90 (0.63)	-0.16 (0.60)	
Alliance Similarity															
Status Quo (Initiator)															
Status Quo (Target)															
Constant	-1.23 <sup>***</sup> (0.38)	-1.52 <sup>***</sup> (0.48)	-0.80 (0.69)	-1.26 <sup>***</sup> (0.31)	-1.66 <sup>***</sup> (0.40)	-0.66 (0.70)	-0.92 <sup>***</sup> (0.30)	-1.30 <sup>***</sup> (0.39)	-0.57 (0.69)	-1.13 <sup>***</sup> (0.34)	-1.36 <sup>***</sup> (0.47)	-0.81 (0.65)	-0.41 (0.38)	-0.78 (0.50)	0.43 (1.06)
N	539	369	170	539	369	170	465	303	161	490	327	163	457	327	130

Standard errors in parentheses. Columns 1 through 3 present main models but with unadjusted standard errors. Columns 4 through 6 present models that use an alternative measure of reciprocity as the dependent variable, one which takes a value of 1 if the target state responded to the threat with its own threat, display, or use of force and zero otherwise. Columns 7 to 9 exclude disputes involving target states whose regime type was originally missing and was manually input. Columns 10 through 12 present models with disputes as part of World War I and II excluded, following Schultz (2001). Columns 13 through 15 present models with alliance portfolio control variables included.

<sup>†</sup> p < 0.10, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001



### Appendix 3.1: Sample characteristics in comparison with US population

	MTurk Sample	US Population <sup>21</sup>
Median Age	37.1	37.6
Percentage Women	48.5	50.8 <sup>22</sup>
Percentage Non-Hispanic/Latino White	72.8	62.3
Percentage Bachelor's Degree or Higher	60.7	29.8
Conservatives	29.4	41.7 <sup>23</sup>
Republicans	37.8	43.0

<sup>23</sup>Based on US Census Bureau 2015 estimates

<sup>23</sup>Results are percentage female, as census does not ask for gender

<sup>23</sup>Estimates of ideology and partisanship from ANES 2016

## Appendix 3.2: News stories used in survey experimental conditions

### Initiation treatment

# ***TRUMP ORDERS CARRIERS TO DISPUTED WATERS IN SALVO AGAINST CHINA***

---

By THOMAS MILLMORE

WASHINGTON — President Trump is ordering two U.S. aircraft carrier battle groups to disputed waters in the South China Sea, the White House announced in a statement earlier this evening.

In a step that is sure to provoke Beijing, the U.S. will move Carrier Group Seven into waters surrounding the Scarborough Shoal, a territory that is claimed by the People's Republic of China, Taiwan, and the Philippines. Carrier Group Eleven is expected to join it in one week's time.

"China has gotten away with being a bully in the region for too long," Mr. Trump said in the statement. "This shoal belongs to the Philippines, one of our strongest allies in the region, and China has no right to it."

In addition to disputes with the Philippines, China also has territorial claims in the South China Sea that overlaps with those of Vietnam, Malaysia, and Brunei. Many of these disputes have been simmering for decades.

Mr. Trump's move appears to be welcomed by the Philippines. "China's claims and actions concerning the Scarborough Shoal contravene international law," said a statement released by the Department of Foreign Affairs. "Today the U.S. demonstrates a commitment not only to the defense of its allies around the world, but also to international justice."

"Era of weak American leaders is OVER. Time to stand up to China!!" Mr. Trump later announced in a tweet.

## ***TRUMP ORDERS CARRIERS TO DISPUTED WATERS IN RESPONSE TO CHINA***

---

By THOMAS MILLMORE

WASHINGTON — In response to recent Chinese missile tests in the region, President Trump is ordering two U.S. aircraft carrier battle groups to disputed waters in the South China Sea, the White House announced in a statement earlier this evening.

In a step that is sure to provoke Beijing, the U.S. will move Carrier Group Seven into waters surrounding the Scarborough Shoal, a territory that is claimed by the People's Republic of China, Taiwan, and the Philippines. Carrier Group Eleven is expected to join it in one week's time. The move comes three days after China conducted missile tests in waters 45 miles away from the Shoal.

"China has gotten away with being a bully in the region for too long," Mr. Trump said in the statement. "This shoal belongs to the Philippines, one of our strongest allies in the region, and China has no right to it. We must respond to Chinese aggression with strength and resolve."

In addition to disputes with the Philippines, China also has territorial claims in the South China Sea that overlaps with those of Vietnam, Malaysia, and Brunei. Many of these disputes have been simmering for decades.

Mr. Trump's move appears to be welcomed by the Philippines. "China's claims and actions concerning the Scarborough Shoal contravene international law," said a statement released by the Department of Foreign Affairs. "Today the U.S. demonstrates a commitment not only to the defense of its allies around the world, but also to international justice."

"Era of weak American leaders is OVER. Time to stand up to China!!" Mr. Trump later announced in a tweet.

## *MOONLIGHT TAKES OSCAR GOLD IN STUNNING UPSET*

---

By THOMAS MILLMORE

LOS ANGELES — In a shocking twist ending to an otherwise predictable Oscar night, “Moonlight” took home the Academy Award for Best Picture — but only after the heavily-favored “La La Land” was erroneously proclaimed the winner.

Warren Beatty, tasked with presenting the award, appeared confused when he opened the sealed envelope. Beatty then passed the card inside to co-presenter Faye Dunaway, who proclaimed “La La Land” as the winner. The modern musical about star-crossed lovers chasing Hollywood dreams began the night as frontrunner for the Best Picture trophy.

While the producers of “La La Land” were making their acceptance speeches, visible scurrying of panicked stagehands was the first sign that something was wrong.

When one producer ended his speech with the words, “we lost, by the way,” gasps from the audience could be heard.

Beatty and Dunaway, apparently, were handed the envelope with the name of Emma Stone, who won the Best Actress award for her role in “La La Land.” Hence the ensuing confusion.

The actual Best Picture winner is “Moonlight,” a harrowing coming-of-age story about a gay black youth in Miami. The film was also awarded Best Supporting Actor for star Mahershala Ali.

“La La Land” did not end the night empty-handed, however. It won a total of six Oscars overall, including Best Director for Damien Chazelle, who at 34 is the youngest winner of the award in the Academy's history.

### Appendix 3.3: Logistic regression results

Favourability and Likelihood of Voting for Trump, Full Sample		
	Favourability	Vote Intention
Initiation	0.01 (0.15)	0.12 (0.15)
Reciprocation	0.28** (0.14)	0.36** (0.15)
Constant	-0.76** (0.10)	-0.88*** (0.11)
Observations	1245	1245

Standard errors in parentheses

\*\*\* $p < 0.01$ , \*\*  $p < 0.05$ , \* $p < 0.1$

Favourability and Likelihood of Voting for Trump by Partisanship		
	Favourability	Vote Intention
Republican	3.11*** (0.28)	3.17*** (0.29)
Initiation	(-0.37) (0.35)	(-0.38) (0.38)
Reciprocation	0.41 (0.29)	0.67** (0.31)
Republican x Initiation	0.80* (0.43)	1.10* (0.46)
Republican x Reciprocation	0.29 (0.39)	0.01 (0.39)
Constant	-2.37*** (0.22)	-2.59*** (0.24)
Observations	1245	1245

Standard errors in parentheses

\*\*\* $p < 0.01$ , \*\*  $p < 0.05$ , \* $p < 0.1$

#### Appendix 4.1: Not free country-years

Country	Years
Afghanistan	1948-2010
Albania	1948-1999
Algeria	1962-2010
Angola	1975-2010
Argentina	1948-1955; 1969; 1974-1983
Armenia	1991-2010
Azerbaijan	1991-2010
Bahrain	1971-2010
Bangladesh	1971-2008
Belarus	1991-2010
Benin	1960-1980
Bhutan	1971-2008
Bolivia	1948-1981
Bosnia	1995-1997
Brazil	1969-1977
Brunei	1984-2010
Bulgaria	1948-1994
Burkina Faso	1960-1977; 1982-1990
Burma	1962-2010
Burundi	1962-2010
Cambodia	1953-2010
Cameroon	1960-2010
Cape Verde	1975-1990



Country	Years
Central African Republic	1960-1992; 1996; 2001-2010
Chad	1960-1989; 1998-2010
Chile	1973-1989
China	1949-2010
Colombia	1948-1957; 2000-2005
Comoros	1975-1992
Congo	1969-1993; 1997-2000
Cote D'Ivoire	1960-2010
Croatia	1995-1999
Cuba	1952-1953; 1957-2010
Cyprus	1960-1975
Czechoslovakia	1948-1989
Democratic Republic of Congo	1960-2010
Djibouti	1977-1994; 1999-2010
Dominican Republic	1948-60
Ecuador	1948-1956
Egypt	1948-2006; 2010
El Salvador	1948-1949; 1978-1992
Equatorial Guinea	1969-2010
Eritrea	1993-2010
Estonia	1992-1994
Ethiopia	1948-2010
Fiji	1987-1999; 2000; 2009-2010
Gabon	1961-1990; 1998-2010
Gambia	1995-2010
Georgia	1991-1994

Country	Years
Germany (East)	1954-1989
Ghana	1960-1965; 1972-1978; 1982-1992
Greece	1948-1995
Grenada	1974-1984
Guatemala	1955-1998; 2003
Guinea	1958-2009
Guinea-Bissau	1974-1990; 1995-2001; 2003
Guyana	1967-1994
Haiti	1948-2005
Honduras	1948-1949; 1963-1966; 1973-1993; 2009-2010
Hungary	1948-1994
India	1975-1977
Indonesia	1950-1997
Iran	1952-2010
Iraq	1948-2010
Jordan	1948-2010
Kazakhstan	1991-2010
Kenya	1963-1993; 1997-2004
Kiribati	1999
Korea (North)	1948-2010
Korea (South)	1949-1992
Kuwait	1961-2010
Kyrgyzstan	1995-2010
Laos	1954-2010
Lebanon	1976-2003
Lesotho	1966-1992; 1998



Country	Years
Liberia	1948-1971; 1982-2009
Libya	1969-2010
Lithuania	1991-1994
Macedonia	1994-1996
Madagascar	1973-1989; 2009-2010
Malawi	1964-2002
Malaysia	1969-2010
Maldives	1965-2007
Mali	1960-1991
Mauritania	1960-2005
Mauritius	1970-1975
Mexico	2010
Moldova	1992-1994; 1998-2009
Mongolia	1948-1988
Morocco	1956-2010
Mozambique	1975-1991; 1996-1998
Nepal	1960-1979; 1993-2005
Nicaragua	1948-1979; 1982-1989; 1994-1995
Niger	1961-2000; 2007-2009
Nigeria	1979-2001
Oman	1971-2010
Pakistan	1948-1988; 1996-1999; 2004-2010
Panama	1970-1989
Paraguay	1948-1989
Peru	1948-1980; 1997-2000
Philippines	1971-1983

Country	Years
Poland	1948-1997
Portugal	1948-1994
Qatar	1971-1996; 2002-2010
Rawanda	1962-2010
Romania	1948-1995
Russia	1948-1994; 1997-2010
Samoa	1997-1999
Sao Tome and Principe	1975-1990
Saudi Arabia	1948-2010
Senegal	1960-1980
Seychelles	1976-1991
Sierra Leone	1967-2002
Singapore	1965-2010
Slovakia	1993-1999
Slovenia	1992-1994
Somalia	1970-2010
South Africa	1961-1993
Spain	1948-1978
Sri Lanka	1948-1990; 1995-2001; 2006-2010
Sudan	1959-1964; 1973-1984; 1989-2010
Suriname	1980-1988
Swaziland	1968-2010
Syria	1948-1958; 1961-2010
Taiwan	1949-1990
Tajikistan	1991-2010
Tanzania	1961-1991

Country	Years
Thailand	1948-1997; 2010
Togo	1960-2010
Tunisia	1956-2010
Turkey	1948-2001
Turkmenistan	1991-2010
Uganda	1966-1978
Ukraine	1991-1994; 1997-2003
United Arab Emirates	1971-2010
Uruguay	1973-1983
Uzbekistan	1991-2010
Vanuatu	1988-1991
Venezuela	1949-1957; 1992; 2002-2010
Vietnam	1954-2010
Vietnam (South)	1956-1975
Yemen	1990-2010
Yemen (North)	1948-1989
Yemen (South)	1967-1989
Yugoslavia	1948-1999
Zambia	1965-1991; 1997-2010
Zanzibar	1963-1964
Zimbabwe	1966-2010

## Appendix 4.2: Robustness checks

	(1) No FE	(2) Lagged Unrest	(3) Unmod Unrest	(4) Alliance (S-Weighted)	(5) Alliance (tau)	(6) ICB DV	(7) Count DV
Unrest	-0.03 (0.02)	-0.00 (0.00)	-0.01* (0.00)	-0.03 (0.02)	-0.03 (0.02)	-0.00 (0.01)	-0.03 (0.03)
Media Freedom							
Partly Free	0.00 (0.02)	-0.00 (0.02)	-0.01 (0.02)	-0.05* (0.03)	-0.05* (0.03)	-0.02 (0.01)	0.00 (0.03)
Not Free	0.03 (0.03)	0.01 (0.03)	0.01 (0.03)	-0.04 (0.03)	-0.05 (0.03)	0.01 (0.02)	0.04 (0.04)
Unrest x Media Freedom							
Partly Free	0.04 (0.03)	0.00 (0.03)	0.01* (0.01)	0.04 (0.03)	0.04 (0.03)	0.01 (0.01)	0.02 (0.03)
Not Free	0.04* (0.02)	0.03* (0.02)	0.01** (0.00)	0.04* (0.02)	0.04* (0.02)	0.01 (0.01)	0.03 (0.04)
Polity Score	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	0.00 (0.00)
Major Power	0.27*** (0.06)	0.30*** (0.02)	0.29*** (0.01)	0.15*** (0.02)	0.14*** (0.02)	-0.00 (0.01)	0.43*** (0.03)
National Capabilities	2.98*** (0.89)	0.00 (1.99)	1.56 (1.85)	-0.72 (1.25)	-0.76 (1.24)	-0.12 (1.59)	7.51 (6.02)
Recent Regime Transition	0.01 (0.01)	0.00 (0.01)	0.01 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.02 (0.02)
Population	-0.00 (0.00)	-0.00 (0.00)	-0.00 (0.00)	-0.00* (0.00)	-0.00* (0.00)	-0.00 (0.00)	-0.00 (0.00)
Alliance Similarity							
Constant	0.12*** (0.02)	0.15*** (0.02)	0.15*** (0.02)	0.20*** (0.03)	0.21*** (0.02)	0.06*** (0.02)	0.17*** (0.04)
N	6418	6399	6418	5323	5323	6418	6418

Standard errors in parentheses

+ p < 0.10, \* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001