

Democratic Reputations in Crises and War

Jonathan Renshon, University of Wisconsin–Madison

Keren Yarhi-Milo, Columbia University

Joshua D. Kertzer, Harvard University

Many international relations theories argue that leaders and publics use regime type to draw inferences about behavior in conflict, with implications for how democracies act as well as how they are treated by other states. We show that these beliefs can be studied as reputations, and we build a framework around reputations that adhere to regime types and whose content implicates not just resolve but a host of other important attributes and expected behaviors. We put democratic reputations under the microscope, fielding survey experiments on members of the Israeli Knesset as well as six national samples in four democracies. We find strong evidence of democratic reputations' existence and pervasiveness as well as insight into their content. Specifically, we find that the reputations are asymmetric: democracy is seen as considerably and consistently more favorable in war than in crises, suggesting that these regimes may have more difficulty signaling resolve than our theories suggest.

Political leaders and mass publics alike often look at world politics through the lens of regime type, frequently using states' domestic political institutions as indicative of their foreign policy behavior, underlying preferences, or resolve. In discussing his country's options in dealing with the rising German threat in the 1930s, the French defense minister Édouard Daladier told then British foreign affairs minister Anthony Eden that "no democratic country could indulge in" preventive war (Eden 1962, 44). Adolf Hitler made similar arguments, famously noting in *Mein Kampf* as well as during private deliberations with his advisors that because democracies were corrupt and weak, they would not take a stand against Germany, writing that they "will be unable to muster the courage for any determined act" (quoted in Press 2005, 76). More recently, the Clinton administration's democracy promotion agenda was directly tied to the president's belief—itself influenced by academic research on the topic—that "democracies rarely wage war on one another" (Clinton 1993).

International relations (IR) scholars have made similar arguments, offering a variety of theoretical models in which

actors use regime type as a heuristic to draw inferences about others' intentions, capabilities, or resolve. Both rationalist (Bueno de Mesquita and Lalman 1992, 156–57) and constructivist (Risse-Kappen 1995) variants of democratic peace theory argue that the reason why democracies are less likely to fight each other is because states use regime type as a heuristic for hawkish or belligerent preferences. Other versions of this same theory pinpoint the public's beliefs about regime type as critical in constraining the use of force (Mintz and Geva 1993) or argue that decision makers believe that democratic states are likely to try harder and expend more resources than comparable autocratic states (Bueno de Mesquita et al. 1999, 794). And a related literature on democratic credibility posits that targets believe threats issued by democracies to be more credible than those from other regimes (Fearon 1994; Schultz 1999). Even work on far-afield subjects such as terrorism emphasizes the importance of beliefs about regime type: Pape (2003), for example, argues that terrorists target democracies because they believe them to be especially sensitive to casualties.

These beliefs about whether democracy is a blessing or a curse in foreign policy are both theoretically and politically

Jonathan Renshon (renshon@wisc.edu) is a professor of political science, University of Wisconsin–Madison, Madison, WI 53706. Keren Yarhi-Milo (ky218@columbia.edu) is the Arnold A. Saltzman Professor of War and Peace Studies, Columbia University, New York, NY 10027. Joshua D. Kertzer (jkertzer@gov.harvard.edu) is a professor of government, Harvard University, Cambridge, MA 02138.

Studies discussed in this article were conducted in compliance with relevant laws and were approved or deemed exempt by the appropriate institutional or national research ethics committee. Replication files are available in the JOP Dataverse (<https://dataverse.harvard.edu/dataverse/jop>). The empirical analysis has been successfully replicated by the JOP replication analyst. An appendix with supplementary material is available at <https://doi.org/10.1086/722043>.

Published online December 1, 2022.

The Journal of Politics, volume 85, number 1, January 2023. © 2022 Southern Political Science Association. All rights reserved. Published by The University of Chicago Press for the Southern Political Science Association. <https://doi.org/10.1086/722043>

consequential, with implications for not just how other actors treat democracies but how democracies themselves behave. However, although beliefs about the consequences of regime type are critical, they are rarely studied directly. We study democracy from a new angle, focusing not on the question of whether democracies behave distinctively in foreign affairs (Lake 1992; Reiter and Stam 2002), but rather, putting democratic leaders and publics' beliefs about democracy under the microscope. Conceptually, we show that those beliefs can be studied as *reputations*: socially shared beliefs about an actor's characteristics and behavioral tendencies. Our framework therefore broadens the discussion of reputation in IR, following Dafoe, Renshon, and Huth (2014) in allowing reputations to belong to any kind of agent (rather than just states or leaders) and to be about a diverse set of attributes and expected behaviors (rather than just resolve). Studying beliefs about democracies as democratic reputations focuses our attention on a set of research questions relevant to a broad range of IR theories: whether democracies have reputations in the first place, what the content of those beliefs are, and how they vary across relevant populations of leaders and mass publics. We describe these as the *existence*, *content*, and *prevalence* of democratic reputations.

Empirically, studying democratic reputations—and, by implication, whether actors use regime type as a heuristic when assessing resolve or predicting military outcomes—requires disentangling democracy from its correlates. However, although democracy is many things, randomly assigned is not one of them: democratic states are wealthier and tend to be found in democratic “neighborhoods,” were united by common interests throughout the Cold War, and happen to include the current global hegemon. Thus, like Mintz and Geva (1993), Tomz and Weeks (2013), and others, we study beliefs about democracy using experimental methods.

Unlike much experimental work in IR, however, we include evidence from a sample uniquely positioned to give us direct insight into the beliefs of leaders: 89 current and former members of the Israeli Knesset. This builds on other recent elite experiments (Renshon 2015; Sheffer et al. 2018) as well as work using the same sample of Israeli leaders (Kertzer, Renshon, and Yarhi-Milo 2021; Tomz, Weeks, and Yarhi-Milo 2020; Yarhi-Milo, Kertzer, and Renshon 2018) to examine related issues such as elite perceptions of public opinion, costly signals and resolve. Our participants are not only elite in every sense of the term (ranking all the way up to prime minister) but also have a history of making decisions about war and peace, with over two-thirds of the sample having served on the Foreign Affairs and Defense Committee. Following the guidance in Kertzer and Renshon (2022), we take advantage of our participants' experience and domain-specific

knowledge by employing an experiment directly related to international conflict, letting us test what elite decision makers believe about the role of regime type in crises and war—and in a country outside the United States (particularly important given concerns about the American-centric nature of many of our conclusions about democracy in IR; see Levin and Trager 2019). We also field our study on six different national samples across four democracies facing very different security environments: two representative samples of the Israeli public, two in the United Kingdom, one in South Korea, and a nationally diverse sample in the United States.

Our results add to the growing literature on the link between regime type, crisis behavior, and war, as well as addressing methodological questions related to elite experiments. Our experimental designs give us leverage on the question of whether democratic reputations exist and what the content of those reputations is. Across our seven studies, we find that democratic reputations exist but are asymmetric: democracies have consistently more favorable reputations in war than in crises. Democracies are seen as more likely to win on the battlefield; better prepared for conflict in terms of training, morale, and allies ready to come to their defense; and more selective about the wars they fight. In crises, however, democracies are seen as being at less of an advantage: their threats are seen as less credible, their preferences more dovish, and depending on the sample, they are either seen as lacking a reputation for resolve or having a reputation for a lack of resolve.

DEMOCRATIC REPUTATIONS

Despite reputation being a foundational topic in the study of IR over many decades, scholars have traditionally conceptualized it quite narrowly, oftentimes reducing it to one type of reputation (for resolve; Schelling 1966), which can only be acquired in one way (through past actions; Weisiger and Yarhi-Milo 2015) and which belongs to only one type of actor (states; Mercer 1996). Much of the work on reputation in international security begins with a standard caveat that while one may have a reputation for any “persistent characteristic or behavioral tendency” (Dafoe et al. 2014, 375), the authors will focus on a reputation for resolve specifically and use “reputation” as a shorthand. While recent work has broadened the study of reputation by allowing actors to have reputations for multiple considerations simultaneously (Brutger and Kertzer 2018; Jervis, Yarhi-Milo, and Casler 2021), or extended theories of reputation to leaders rather than just countries (Guisinger and Smith 2002; Renshon, Dafoe, and Huth 2018; Yarhi-Milo 2018), they otherwise have not strayed very far and rarely explicitly conceptualize reputations as adhering to regime types.

However, there are reasons to question this triad of restrictive assumptions, in particular the two that are most

dogmatically followed in IR: that reputations “belong” only to states or leaders and that they are acquired exclusively through one’s past actions. On a conceptual level, if reputations are socially shared beliefs about an actor’s behavioral tendencies (Dafoe et al. 2014), they can adhere to any kind of actor even if most work focuses exclusively on state- or leader-owned reputations.¹ And if reputations can in fact adhere to any kind of agent or factor that shapes behavior (Jervis et al. 2021; Renshon et al. 2018, 326), regime type might be a particularly good candidate for several reasons. First, it is a highly salient and accessible indicator and one of the major ordering principles in contemporary global politics (Huntington 1991). There are a number of valuable distinctions to be made between different types of democratic (Martin 2000; Narang and Staniland 2018) and autocratic (Mattes and Rodríguez 2014; Weeks 2014) systems in foreign policy, but the contrast between democracies, on the one hand, and dictatorships, on the other, is especially stark. Second, IR scholars have long been interested in the ways democracies conduct their foreign policies systematically differently from nondemocracies (Doyle 1986), in areas ranging from conflict (Gelpi and Grieco 2001; Lake 1992) to cooperation (Martin 2000; for a review, see Hyde and Saunders [2020]). If this literature is correct, and democracies do behave distinctively compared to their non-democratic counterparts, this suggests that observers should catch on: regime types *should* have reputations.

Another common assumption in IR has been how reputations are acquired, with most work sharing the common assumption that actors “invest” in their reputation by taking specific actions that are observed by other parties (e.g., Sechser 2018). In this telling, reputations are developed through actions taken (or not) that are witnessed by observers and incorporated into a running tally of “past actions.” There are reasons to question the restrictive model of reputation building that most IR scholars use, however. For one, even if actors did calculate reputations in this manner, selective attention to events (Yarhi-Milo 2014) and bias in how they interpret complicated and mixed records would make it clear that the best way to study reputations is not by studying the past events themselves (e.g., democracies’ record of winning and losing in wartime) and assuming they track perfectly with reputations but rather by studying the reputational beliefs directly. The historical record is sufficiently complex that even IR scholars devoted to sifting through observational data do not neces-

sarily agree on the causal effect of democracy in conflict (Renshon and Spirling 2015).

Further, it is unlikely that how states act is the sole input into their reputations. Scholars in American and comparative politics often study the reputations of political parties (Lupu 2016; Snyder and Ting 2002), and these party brands are closer to what we might think of as stereotypes, cultivated and contested through discourse, rather than as a simple running tally of past actions. In an IR context, Democrats in the United States have been no less likely to initiate fatal militarized interstate disputes (MIDs) or international crises than Republicans, yet the Democratic Party continues to wrestle with its dovish reputation (Kertzer, Brooks, and Brooks 2021).² A growing literature on the role of gender stereotypes in crisis bargaining (Post and Sen 2020) provides another example: female leaders face incentives to escalate in international crises not because they have necessarily backed down in the past but because audiences have socially shared beliefs about gendered behavioral tendencies and may use gender as a heuristic when assessing resolve. In sum, if our interest is in the content of reputations, we are better served by directly examining the beliefs of leaders and publics about how democracies behave and what outcomes they are associated with rather than trying to backward induct such quantities through an examination of a very messy historical record that observers interpret and process selectively. Directly examining the beliefs themselves also supplements the standard experimental approach of manipulating past actions and studying the resultant inferences, which is unlikely to present a complete picture of reputations if other inputs exist.

Once we shed the restrictive assumptions listed above and see reputations as belonging to actors other than states and leaders, as being about more than just resolve, and as being built through a combination of actions taken but also discourse and motivated processing of information, we can begin to see the outlines of democratic reputations—“socially shared beliefs about the characteristics, attributes, and behaviors” of democracies (Hilton and von Hippel 1996, 240)—implicit in a wide range of literatures in IR, even if the authors might not always characterize them as such. For example, both rationalist (Bueno de Mesquita and Lalman 1992, 156–57) and constructivist (Risse-Kappen 1995) variants of democratic peace theory argue that the reason why democracies are less likely to

1. A recent quantitative literature search estimated that 69% of extant IR work on reputations focused on state reputations, 21% addressed leader reputations, and the remainder fell into a miscellaneous “other” category. Renshon et al. (2018, 328).

2. One key distinction between reputation and stereotypes is that stereotyping requires social categorization, whereas reputation does not. As Taylor (1981, 83) notes, “we do not stereotype a person, we stereotype a person-as-a-member-of-a-group,” but since democracy itself constitutes the relevant social category here (Hayes 2012), we can think of the two as being equivalent to one another for the discussion that follows.

fight each other is because states use regime type as a heuristic for aggressive or belligerent preferences, which is to say that a reputation (for either hawkish or dovish behavior) has attached not to a state or to a leader but to a set of institutions under the umbrella of “regime type.” Democratic reputations similarly manifest themselves in Gelpi and Griesdorf’s (2001, 642) findings about democracies being more likely to be targeted in crises, which they attribute to the stereotype “that democracies prefer not to fight.” These reputations also arise outside of the confines of IR theory: democratic reputations are invoked everywhere from George Kennan’s justifications for covert interventions (Gaddis 2011, 294) to neoconservative justifications for democracy promotion (Caverley 2010). And they can apply to domestic or international audiences alike. One of the major concerns scholars have about the current “crisis of democracy” is that democratic institutions have acquired reputations for poor performance compared to their autocratic counterparts, causing their citizens to lose faith in the importance of democratic norms and principles (Hall 2013). In a recent article, Bush and Zetterberg (2021) study “reputations for democracy”: If you know an actor’s behavioral tendencies, what inferences do you draw about how democratic it is? We study the inverse question: If you know an actor is democratic, what inferences do you draw about its behavioral tendencies in crises and war?

A FRAMEWORK FOR STUDYING DEMOCRATIC REPUTATIONS

The brief discussion above points to several sets of hypotheses about democratic reputations. The first simply concerns the existence of these reputations. The democratic reputations hypothesis posits that observers—both elites and the general public—will use regime type as a heuristic to draw inferences about other states’ intentions, capabilities, and resolve. Knowing that an actor is democratic will affect observers’ predictions about its behavioral tendencies in crises and war. There are many reasons to expect this to be the case. On a general level, there is an “extraordinary readiness to characterize vast human groups” in terms of a few fairly broad “traits” (Tajfel 1963). These beliefs about the qualities and attributes of other groups are easily formed, difficult to change and serve the critical psychological purpose of helping us make sense of the behavior, motivations, and possible future actions of others. On the more specific issues of beliefs and stereotypes about democracies, our earlier discussion helps make clear that actors in IR seem ready to categorize others at least partially on the basis of their regime types. Of course, the democratic reputations hypothesis is by no means definitionally true. Just as plenty of work in the realist tradition was skeptical that reputations mattered or could form in the first place (Copeland

1997; Tang 2005), it might well be the case that observers see regime type as uninformative in predicting what a given actor is likely to do in conflict situations, perhaps because of the widespread heterogeneity among democracies themselves (Hyde and Saunders 2020).

H1a. Knowing a state is democratic will affect observers’ predictions about its foreign policy behavior.

H1b. Knowing a state is democratic will not affect observers’ predictions about its foreign policy behavior.

Our second set of hypotheses concerns the content of the reputations, or as Dafoe et al. (2014) and Jervis et al. (2021) asked, a reputation for what? Our chief focus concerns beliefs about how democracies are likely to fare in crisis and war relative to their autocratic counterparts. Here, two contradictory sets of predictions are suggested by different literatures.

On one side are the literatures on audience costs, domestic constraints, and democratic credibility (Fearon 1994), which make the argument that democratic leaders will be less likely to back down once they have engaged another state in a crisis as a result of the penalty that they would expect to pay at the polls, a punishment based on an assumption that democratic publics will punish leaders who make threats and do not follow through. Following this logic through, our respondents should see democracies as less likely to back down once already engaged in a public crisis (in accordance with Schultz [2001]). This is particularly true for leaders who—in order for audience cost theory to work—must see democracies as less likely to back down in crises because they know that those democratic leaders would not be in the crisis in the first place unless they intended to follow through.

Yet, audience cost theory is not without its critics: if democratic publics are seen as having preferences about policy rather than just inconsistency (Kertzer and Brutger 2016), democracies might have reputations for weakness in crises rather than strength. Gelpi and Griesdorf (2001, 641), for example, suggest that leaders will hold the “belief that democracies prefer not to fight” (and thus would be more likely to back down in a crisis). A number of theories imply that democracies will have a reputation for being especially sensitive to casualties compared to other regimes (Pape 2003) and more generally to the human and financial costs of war (Mueller 1971; Valentino, Huth, and Croco 2010). Other literature on variation within autocracies suggests that authoritarian leaders might be more susceptible to domestic constraints than theories of crisis bargaining typically assume, which would have the effect of reducing or even erasing any advantage democracies are perceived to have (Weeks 2008). Taken together, these works

imply that democracies should be seen as being at a disadvantage in crises as a result of dovish preferences.

The IR literature has similarly mixed expectations for democratic reputations in war. Work on military effectiveness implies we should find evidence of democracies as having reputations for greater effectiveness in battle as a result of democratic culture and institutions (Reiter and Stam 2002). The implications of selectorate theory (Bueno de Mesquita et al. 1999) are similarly optimistic, suggesting that our respondents should believe that democratic states are likely to try harder and expend more resources than nondemocratic regimes and should be more selective about the wars they fight. Yet other scholarship is skeptical of these findings (Desch 2008), and the historical record is mixed (Renshon and Spirling 2015).

Popular conceptions of democracy in war often emphasize the extent to which democracies are forced to fight wars with one hand behind their back, constrained by a cost-intolerant public and an open press eager to publicize any missteps. Democratic leaders—and their publics—thus might know or suspect that “wars can be hazardous to [their] political career” (Chan and Safran 2006, 139), particularly when casualties mount (Gartner and Segura 1998). Yet, a competing perspective notes that, in general, democratic leaders seem to face comparatively low odds of “irregular removal” from office (exile, prison, or death) compared to autocratic leaders (Goemans 2008), and their tenure is less obviously affected by war outcomes (Debs and Goemans 2010). Just as with crises, however, a third set of literature suggests more of a middle ground, arguing that there may be fewer differences than we assume between democracies and autocracies, with Weeks (2012, 327) arguing that “autocratic audiences consisting primarily of civilians are scarcely more likely to forgive unnecessary or failed uses of force than democratic domestic audiences.” Just as with crises, then, existing IR scholarship leads to contradictory expectations about the content of democratic reputations in war.

H2a. Democracies will be seen as more likely to stand firm in crises.

H2b. Democracies will not be seen as more likely to stand firm in crises.

H3a. Democracies will be seen as more likely to win in war.

H3b. Democracies will not be seen as more likely to win in war.

Three points are worth emphasizing here. First, these hypotheses are all bidirectional, reflecting the divergent expect-

ations offered by different strands of literature in IR. Our aim here is therefore an exploratory one, seeking to empirically adjudicate between the discipline’s competing expectations about the role that regime type plays in how countries are perceived in foreign policy. Second, these expectations about how democracies will fare in crises and war are each driven by specific mechanisms implicated by the existing literature—perceptions of domestic constraints, the credibility of democratic threats, dovish preferences, democratic selectivity, and military effectiveness. To appropriately reflect the complexity of these arguments, we incorporate these mechanisms into our experimental design and analysis. This has the benefit of helping us better understand the more general patterns we observe, giving us a more fine-grained picture of the content of democratic reputations.

A third and final point is that the relationship between democratic reputations in the domains of crises and wars (hypotheses 2 and 3 taken together) can help shed light on a significant theoretical question in the study of reputation: To what extent do reputations “transfer” (Wiegand 2011) from the contexts in which they were generated? Typically, this is addressed through research on whether reputations for resolve are transferable to different geographic regions or issue areas. Results have been mixed, with extant studies indicating that states do attempt to transfer reputations for resolve (Wiegand 2011; see also Huth 1988) but other more skeptical voices arguing that reputations can only exist within one geographic region or within a single issue area (Snyder and Diesing 1977) or that the type of reputations most likely to be transferable are those for honoring commitments (Gibler 2008; or what Jervis [2002, 305] would call a “signaling reputation”). To the extent that reputations in crises do not affect reputations in war, it would suggest the difficulty that regime types might have in generating reputational credit in one domain and spending it in another.

Our final question concerns the pervasiveness of the reputations, or who exactly holds these beliefs about democracies and how much overlap exists across different audiences? As figure 1 shows, in some cases our theories implicate democratic reputations in the eyes of nondemocratic leaders: for example, whether nondemocratic leaders perceive democratic threats to be more credible than nondemocratic ones (Schultz 2001). Yet, as the top-left panel of figure 1 shows, we also care about democratic reputations in the eyes of democratic leaders themselves. Just as the existence of party brands affects political parties’ strategic incentives—parties with reputations for hawkishness, for example, may be better able to ratify arms control agreements (Kreps, Saunders, and Schultz 2018)—democracies’ reputations affect not just how they are treated by other states but also how democracies themselves conduct foreign policy. For example, if democratic leaders buy into the

	Leaders	Mass public
... in democracies	Decision-makers using regime type as a heuristic for hawkish preferences (e.g. Bueno de Mesquita & Lalman 1992) Neoconservative rationales for democracy promotion (Caverley 2010) Almond-Lippman consensus: stereotype of democratic publics as weak/inconsistent	Mass public variants of democratic peace theory (e.g. Mintz & Geva 1993, Hayes 2012) “Crisis of democracy” literature: have democracies acquired reputations for poor performance among their publics? (e.g. Hall 2013) Democratization literature: public beliefs about democratic superiority necessary for democratic institutions to function (e.g. Rose, Misherl & Haerpfer 1998)
... in non-democracies	Decision-makers using regime type as a heuristic for hawkish preferences (e.g. Bueno de Mesquita & Lalman 1992) Democratic threats seen as more credible than those from non-democracies (e.g. Fearon 1994, Schultz 1999) Democracies more likely to be targeted in crises because they prefer not to fight (e.g. Gelpi & Griesdorf 2001)	Mass public variants of democratic peace theory in non-democratic states (e.g. Bell and Quek 2018) “Crisis of democracy” literature: have democracies acquired reputations for poor performance among non-democratic publics? (e.g. Diamond 2008) Democratization literature: public beliefs about democratic superiority necessary for democratic institutions to function (e.g. Rose, Misherl & Haerpfer 1998)

Figure 1. Examples of democratic reputations

Almond-Lippman consensus’s concerns about democratic weakness in foreign policy crises, they should strive to insulate the public from foreign policy and engage wherever possible in private rather than public diplomacy.

In other cases, we care about democracies’ reputations in the eyes of the general public (the right half of fig. 1). This is obviously the case in the democratization literature, which argues that citizens’ socially shared beliefs about democracy (e.g., its superiority to alternative forms of government) are the necessary “software” without which democratic institutions cannot operate (Rose, Mishler, and Haerpfer 1998), which is why debates about whether support for democratic institutions has declined over time have drawn so much attention (Foa and Mounk 2016). However, these beliefs of average citizens also matter in IR because of the importance of public opinion for theories of democratic constraint (Baum and Potter 2015): if publics and leaders share the same beliefs about democracies in crises and war, it suggests public opinion is unlikely to act as a constraint on leaders’ abilities to act on those beliefs. It is also valuable to use mass public data to learn how widespread the reputations of democracies are (do they extend across countries and types of democracy?). To address these theories properly and provide a first cut on democratic reputations, our research design (described in more detail below) focuses on the top half of figure 1, centering on a unique pairing of democratic elite and mass public samples. Our final

set of hypotheses thus concerns the overlap in the stereotypes across the two different audiences.

A widespread debate in political science concerns whether the beliefs of elites and the mass public overlap, with some recent work showing both striking similarities in beliefs across surveys and experiments in multiple samples (Kertzer 2022; Kertzer et al. 2021) and others suggesting large differences (Page and Bouton 2007). In the more general work linking elite and public opinion, a host of research on elite cues suggests that we should see substantial overlap among the two groups as a result of the public following the lead of elites, particularly in the domain of IR where the public may be less well informed (Guisinger and Saunders 2017; Zaller 1992). However, the issue is far from settled, and other notable work has demonstrated that there are limits to the “top-down” model of foreign policy beliefs (Kertzer and Zeitzoff 2017) and ways in which political elites and the general public differ along dimensions that affect the formation of their beliefs and preferences. Elites are, for example, typically much more powerful, which shapes their views on a number of issues (Renshon 2015), and may be less subject to common decision biases (Carnevale, Inbar, and Lerner 2011; Sheffer et al. 2018).

H4a. Democratic reputations in crises and war will not significantly differ between democratic leaders and democratic publics.

Here is the situation:

- Two countries are currently involved in a public dispute over a contested territory. The dispute has received considerable attention in both countries, because of the risk that disputes like these can escalate to the use of force.
 - Country **A** is a [democracy/dictatorship]. Country **B** is a dictatorship.
 - Both countries have moderately powerful militaries, with large armies, moderate-sized navies, and well-trained air forces.
 - Neither country is a close ally of the United States.
 - Country **A** is slightly larger than Country **B**, though their economies are approximately the same size.
 - Country **A** has moderate levels of trade with the international community. Country **B** has high levels of trade with the international community.
 - The last time the two countries were involved in an international dispute, different leaders were in power.
1. Given the information available, what is your best estimate about whether Country A will stand firm in this dispute, ranging from 0% to 100%?
 2. If the dispute were to escalate and war were to break out, what is your best estimate about whether Country A will win, ranging from 0% to 100%?

Figure 2. Experimental vignette and primary outcomes

H4b. Democratic reputations in crises and war will significantly differ between democratic leaders and democratic publics.

RESEARCH DESIGN

We study democratic reputations in crises and war by fielding an original experiment in seven samples, from four different democracies, across a three year period. Our experiment studies democratic reputations in the eyes of leaders and the mass public (the top row of fig. 1), by manipulating the regime type of an actor in a military dispute and testing how it affects the inferences respondents draw. Fielding the study on both the elite and mass samples allows leverage on the additional question of whether these reputations are consistent across different groups in democracies.³ Our elite sample consists of 89 current and former members of the Knesset in Israel. As we note below, this is an unusual sample even by the standards of elite experiments in IR: two-thirds of our respondents had experience serving as members of the Knesset's Foreign Affairs and Defense Committee. Our least "elite" participant is a member of parliament; our most, a prime minister. The other six surveys were fielded on mass public samples—two nationally representative samples in Israel, two in the United Kingdom, one in South Korea, and one nationally diverse (although not representative) sample in the United States. As we discuss below, by fielding the same experiment in a range of democracies in very different foreign policy contexts, we not only probe the generalizability of our results but also draw inferences regarding beliefs about democracies in crises and

war outside the confines of the United States (Narang and Staniland 2018).

The experiment

All respondents across all seven samples were presented with the same experiment, albeit in different languages (Hebrew in Israel, Korean in South Korea, and English in the United Kingdom and the United States).⁴ All subjects read a vignette (reproduced in fig. 2), in which the regime type of country A was experimentally manipulated in a manner consistent with both other experimental designs (e.g., Johns and Davies 2019; Kertzer et al. 2021) and recent guidance suggesting that using unnamed countries does not risk any loss of experimental control (Brutger et al. 2022).⁵ In addition to our regime type manipulation, we described a number of other characteristics of the individual states, as well as their relationship. For example, we control for alliance ties by stating that neither country is an ally of the United States, in order to avoid two concerns in particular. First is information leakage (Dafoe, Zhang, and Caughey 2018; Kertzer and Brutger 2016; Tomz and Weeks 2013). A common critique of the literature on democracies in war is that our theories of democracy are just really theories of US hegemony in disguise and that the salutary effects of democracy are simply the benefits that accrue to the United

3. Crucially, the reputational effects we examine are identified off of experimental manipulations of the *target of the reputational inference* in the vignette rather than the actual regime type of the respondents.

4. Our experiments were fielded using the Qualtrics platform (aside from the few Knesset members who chose to fill out paper copies). In each study except for the Israel II sample, respondents took part in other unrelated experiments as part of their participation. The module orders were fixed because of the multimode nature of the Knesset study, which means that we cannot examine potential order effects.

5. See app. D for a broader discussion of issues related to the manipulation of regime type and our choice to fix alliance ties.

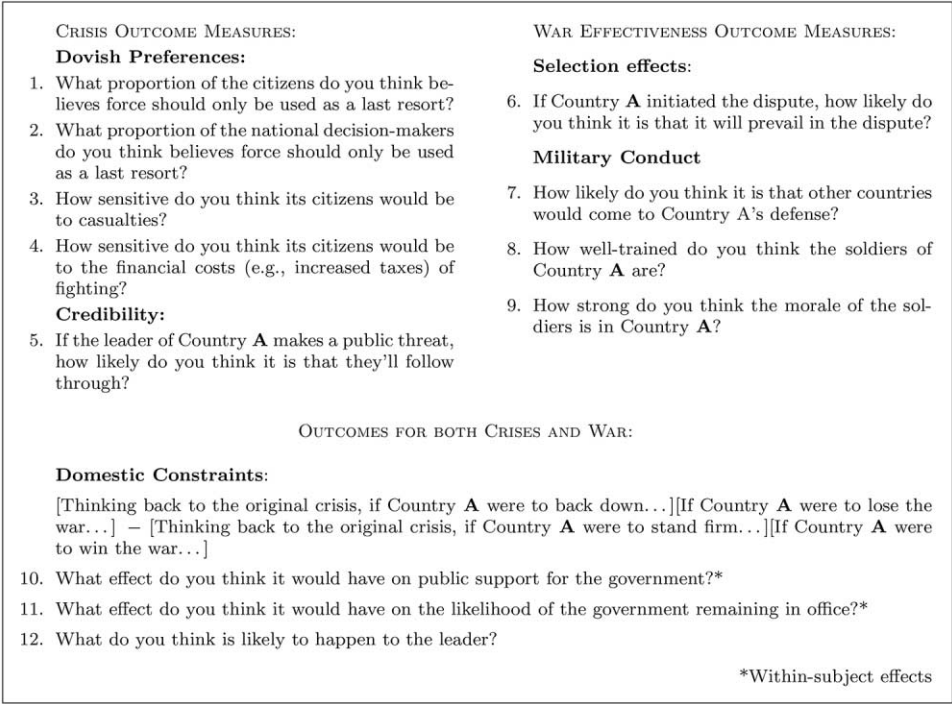


Figure 3. Secondary measures

States and its friends. Without providing anything about each country's alliance ties, one concern would be that respondents would assume the country was a US ally if it was a democracy, and assume it was not an ally if it was a dictatorship, leading to information leakage and bundled treatment problems; holding alliance ties fixed across both conditions avoids this concern.

The second set of concerns addressed by our design was projection: although we deliberately told respondents that the scenario was not about their own country, we were still concerned some would simply project their own state into the scenario. Fixing a number of these characteristics thus helps to minimize concerns about confounding as well the potential bias that may result from positive illusions were subjects to perceive the vignette as being about their own country (Brown 1986).⁶ Finally, specifying a number of different dimensions helps to avoid “putting our thumb on the scale”: if respondents are only provided information on one dimension, it is far more likely for them to weight that dimension heavily. We thus described each country's military, economy, geographic size, and so on, seeking to avoid demand effects for our treatment by ensuring the two countries slightly differed on multiple dimensions. Participants then were administered our primary outcome measures: how likely they thought it was that country A would stand firm in the crisis and, if the dispute

escalated into war, how likely they thought it was that country A would win.⁷

In one of the studies, respondents were also administered a set of secondary measures (reproduced in fig. 3). We divided our respondents in two, with half being reminded of their answer to the “resolve in a crisis” outcome and asked questions about the crisis scenario and half being reminded of their answer to the “effectiveness in war” outcome and asked questions about the war scenario.⁸ These secondary measures specifically test the five sets of mechanisms—perceptions of domestic constraints, the credibility of democratic threats, dovish preferences, democratic selectivity, and military effectiveness—implicated by the existing literature on democratic foreign policy.

Elite sample: The Israeli Knesset

Our elite sample consists of current and former members of the Israeli Knesset, who bring several important advantages

6. See app. sec. D.1 for a discussion on how we distinguish reputations from related concepts such as identity or signaling.

7. In the two Israeli public samples, we included a manipulation check asking participants to recall which regime type condition they were in. In the Israel I sample, 85% correctly recalled their treatment condition; in Israel II, the passage rate was 90%. In the analysis below, we include all respondents regardless of whether they passed the manipulation check, although the results do not substantively change if those participants who failed the manipulation check are excluded.

8. In addition to this between-subjects manipulation, we included an order manipulation to avoid potential downstream effects, and we estimate within-subject effects for two of the domestic constraint mechanisms, described in greater detail in app. sec. B.3.

to the study of democratic reputations in international conflict. First, the most direct way to achieve our goal of examining decision makers' beliefs about democracies in war and peace is to sample from a population that has wrestled with those issues outside the lab. This is, after all, what is unique about studying leaders rather than the public. "Use of force" decisions are ubiquitous in Israel and highly salient for Israeli decision makers: during our leaders' time in office (from 1996 onward), Israel was involved in 16 MIDs. Our participants were involved in many of these cases. More broadly, since earlier experiences of Israeli conflict might have shaped their beliefs, we note that Israel has been involved in 128 documented MIDs since the state's inception, as well as seven militarized compellent threat episodes.

Second, because of the structure of Israel's parliamentary system, the vast majority of the executive branch (e.g., the prime minister and ministers in the security cabinet) are also elected members of the Knesset (the legislative branch). Thus, the Israeli Knesset (unlike, e.g., the US Congress) is composed of policy makers who are directly involved in use-of-force decisions. Because of political norms and relatively short election cycles, it is common for former members of the executive branch to later become members of the opposition in the Knesset; conversely, nearly all current members of the executive branch were at some point in their career members of the opposition in the Knesset. Thus, even the Knesset members in our sample who are currently part of the opposition either have been members of the executive branch in the past or are likely candidates to become so in the future. Put differently, by sampling current and former members of the Knesset, we are also effectively sampling current, former, and potentially future members of the executive branch. Indeed, since our study was fielded, over a quarter of the sitting members of Knesset (MKs) who were backbenchers when they took our study have been promoted to minister.

Our elite survey was fielded July–October 2015. Of 288 potential subjects, 89 participated, leaving us with a 31% response rate. The Knesset sample is the same one used for several other experiments and is described in greater detail in Yarhi-Milo et al. (2018, table 2, 2161). Our recruitment procedures are described in appendix section A.2, while appendix section A.3 describes our protocol to increase our confidence that the MKs themselves participated in the study, rather than their staff. As discussed in Yarhi-Milo et al. (2018), 25% were current members; the rest (75%) were former MKs. Many of our participants had experience in IR-relevant contexts: 64% had active combat experience, and 67% had experience serving as members of the Knesset's Foreign Affairs and Defense Committee. They also had considerable political experience: on average, our participants had served three terms in Par-

liament, and some had served as many as nine terms. While 58% of the Knesset subjects had never served as a minister at the time of taking the study, 29% had been at least a deputy minister, and fully 12% of our sample was in our highest category of elite experience, such that our participants include individuals who had served as cabinet members or above.

We show in previous work (Yarhi-Milo et al. 2018, app. 3) that our leader sample is fairly representative of the universe of Israeli political leaders from the time frame we examined, although not surprisingly, our analysis reveals that current members of the Knesset were less likely to participate than former members. As is evident, this is an extremely unusual sample: our least "elite" participant is a member of parliament; our most, a prime minister. In fact, even in many elite studies of decision-making in IR, subjects are often far removed from the actual decision makers of primary interest to IR theories. Renshon (2015), for example, uses political and military leaders drawn from a midcareer training program at Harvard Kennedy School, while Alatas et al. (2009) use Indonesian civil servants, Hafner-Burton et al. (2014) use "policy elites" (including civil servants, corporate executives, former members of Congress, and US trade negotiators), and Mintz, Redd, and Vedlitz (2006) use Air Force officers. While more elite than college freshmen, to be sure, the samples used are still somewhat removed from the dictators, presidents, leaders of the military and foreign ministry, trusted advisors, and generals who are the primary decision makers in most interstate conflicts.⁹ This serves as a reminder that the use of quasi-elite subjects, while interesting and helpful, does not completely obviate the necessity of extrapolating from one population to another. The research design we employ here is thus perhaps most similar to Findley et al. (2017), who field experiments about foreign aid on paired samples of Ugandan parliamentarians and members of the mass public.

Mass public samples

While the Knesset sample is valuable both theoretically and methodologically—if we care about democratic stereotypes in crises and wars, we surely especially care about those in the eyes of leaders in a democracy frequently engaged in both—we are also interested in ordinary citizens' beliefs about democracies across a range of countries. We supplement our elite sample by fielding our experiment in six mass public samples across four democracies in very different foreign policy contexts.

9. A focus on elites is more common within the literature on foreign policy attitudes, but those works have a similarly expansive view of "elites" (e.g., business executives) and were focused on broad foreign policy attitudes, rather than dynamic judgment and decision-making. See, e.g., Holsti and Rosenau (1988).

While fielding our experiment on democratic publics allows us to answer questions of theoretical relevance to debates about democratic reputations, the pairing of elite and mass public samples across multiple countries provides several additional benefits. The first is that it lets us speak to debates about differences between elites and masses in IR. To that end, two of the samples are drawn from the Israeli general public. Including mass samples from the same country as the elite survey provides leverage on the selection of leaders from the general population, thereby giving us insight into the dimensions on which they differ and those on which they resemble their compatriots. The second is the value inherent in any replication, increasing confidence in the overall research program and generalizability of our findings. This is particularly germane for the study of democracies in IR, given Narang and Staniland's (2018, 413) call for IR scholars to "disaggregate democracies" and test our theories about democratic foreign policy across a wider range of democracies, a point echoed by Levin and Trager (2019). We thus replicate our findings in four other national samples in three other democracies (South Korea, the United States, and United Kingdom), selected because they vary along multiple dimensions, including region, political system, relative military capabilities, and the salience of foreign policy. Altogether, our four democracies vary in a variety of ways: both Israel and South Korea, for example, face regional security threats and have some form of mandatory military service, while the United States and United Kingdom do not; both Israel and the United Kingdom are parliamentary democracies, whereas the United States and South Korea are presidential systems. More systematically, the polity scores of the countries we fielded our studies in range from 5 (the United States) to 8 (South Korea and the United Kingdom; Marshall and Gurr 2019), and one of our samples comes from a country (South Korea) that has transitioned to democracy relatively recently (within the last 30 years). A third and final benefit is that surveys on mass samples suffer from fewer space constraints, enabling us to study democratic reputations in greater depth.

The two Israeli public samples were fielded in September–October 2015 and January 2016, respectively, by iPanel, an Israeli polling firm that has been used effectively by other recent surveys and experiments (e.g., Manekin, Grossman, and Mitts 2015). Both samples are representative of the Israeli Jewish population and stratified on the basis of gender, age, living area, and education.¹⁰ The third and fourth samples are

national samples in the United Kingdom and were fielded in May and June 2018 by Survey Sampling International. The fifth sample is a national sample in South Korea fielded by the Korean polling firm Embrain in May 2018. The British and Korean samples were both stratified on the basis of gender, age, and location. The final sample was fielded on a nationally diverse, although not nationally representative, sample in the United States, through Amazon Mechanical Turk in June 2015. In addition to standard demographic data, participants in each survey completed questionnaires capturing a variety of political orientations; the list of orientations varied on the basis of the country, but in the Israeli case it included military assertiveness, political ideology, stance on the Arab-Israeli conflict, and international trust.¹¹ Descriptive statistics for these six samples, totaling nearly 9,000 respondents, are presented in table 1. Our combination of samples, while obviously not comprehensive of all democracies, captures a wide range of democratic security environments, across a wide range of regions (North America, Western Europe, the Middle East, and East Asia) and in a manner not usually addressed with original data in a single study.

RESULTS

We present our results in two stages. We begin by looking at the content of democratic reputations in crises and war (hypotheses 1 and 2), focusing in depth on the results from one of our Israeli public samples, which includes both our primary outcome measures and the detailed battery of secondary measures. We then get at the question of the prevalence of democratic reputations (hypothesis 3) by analyzing the results of the primary outcome measures in six other samples we fielded: one on an elite sample of members of the Israeli Knesset and five other mass public samples in Israel, the Republic of Korea, the United Kingdom, and the United States.

The content of democratic reputations

Figures 4A and 4B visualize the content of democratic reputations in crises and war in a nationally representative sample of the Israeli public; our primary outcome measure is in the top row of each panel, and a series of secondary outcome measures are shown in the subsequent rows. Each row depicts the bootstrapped distribution of average treatment effects of democracy. All outcome measures have been recoded so as to be easily interpretable on a probability scale.¹²

We begin with our primary outcome measures. The top rows of figures 4A and 4B show that for our respondents,

10. Our focus on the Israeli Jewish population is due entirely to logistical constraints, specifically the inability of online polling companies in Israel to provide anything close to a representative sample of the minority Israeli Arab population.

11. See app. sec. B.2 for the complete instrumentation.

12. The results also hold when we replicate this analysis on a non-probability scale.

Table 1. Mass Public Samples

	Israel I	Israel II	Korea	UK I	UK II	USA
Male	.53	.52	.54	.44	.44	.52
Age:						
<25	.13	.13	.14	.11	.16	.18
25–34	.27	.26	.17	.25	.24	.45
35–44	.20	.20	.21	.23	.23	.21
45–54	.17	.17	.33	.22	.19	.09
>55	.24	.24	.15	.18	.18	.07
Education:						
High school or less	.35	.27	.20	.54	.55	.11
Some college	.22	.22	.16	.11	.11	.26
College/university	.27	.31	.54	.25	.24	.51
Postgraduate	.15	.19	.10	.10	.10	.12
Ideology	.61	.59	.52	.52	.50	.39
Military assertiveness	.58	.56				.39
International trust	.32					.35
Arab-Israeli conflict	.63	.63				
Political interest			.45	.47	.47	.56
Interest in foreign policy			.40	.51	.52	.54
N	1,599	1,111	1,797	1,105	1,186	2,057

Note. Total $N = 8,555$.

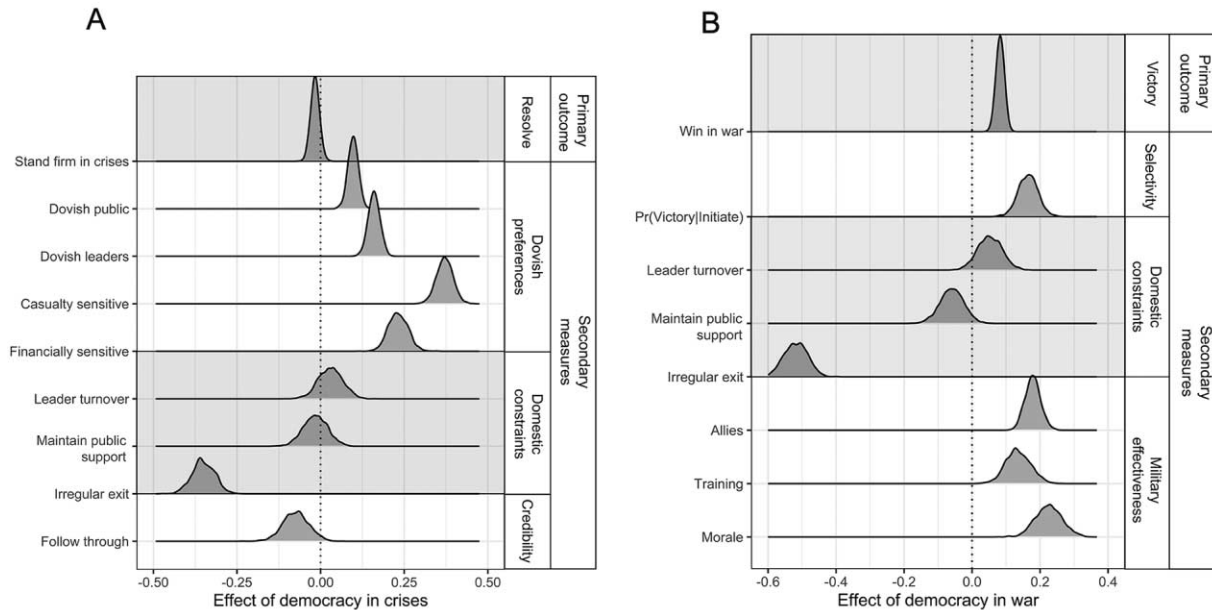


Figure 4. Democratic reputations in crises (A) and wars (B) in a nationally representative sample of the Israeli public; our primary outcome measure, the likelihood of standing firm (A) or winning the war (B), is shown in the top row, and a series of secondary outcome measures are shown in the subsequent rows. Each row depicts the bootstrapped distribution of treatment effects of democracy. All outcome measures have been recoded to be interpretable on a probability scale. The results suggest democratic reputations in crises are often unfavorable (democracies are seen as slightly less likely to stand firm, more sensitive to the costs of fighting, and less likely to follow through on their threats), while reputations in war are very favorable (democracies are seen as more likely to win the wars they fight, selective about the wars they fight, and higher in military effectiveness).

democracies have favorable reputations in war, but not in crises, providing evidence in favor of hypotheses 1a, 2b, and 3a. In a scenario that controls for a wide range of features, knowing that an actor is democratic causes respondents to see it as 8 percentage points more likely to win the war than if it were nondemocratic. Democracies thus appear to have a sizable reputation for military victory. However, even if respondents perceive democracies to be at an advantage in war, they do not perceive this same advantage to extend to crises: democracies are seen as 2 percentage points less likely to stand firm in the crisis than nondemocracies. Our respondents do not perceive democracies to be at a significant disadvantage in crises, but the favorable reputations democracies possess in battle do not appear to travel to the bargaining table.

To better understand the content of democratic reputations, we turn to the secondary outcome measures, beginning with crises in figure 4A. As noted above, the IR literature has two competing images of democracies in crises, each of which implicates a different set of mechanisms. One is an optimistic account offered by the domestic constraint and democratic credibility literatures, which argues that the prospect of losing public support or being thrown out of office altogether incentivizes democratic leaders to stand firm, bolstered by democracies' ability to send more credible threats. Another is a pessimistic account invoking dovish preferences, which argues that democratic leaders and publics should have reputations for being more sensitive to the human and financial costs of fighting and less interested in fighting in the first place.

Consistent with democracies lacking reputations for resolve, the results from our secondary outcome measures much stronger and more consistent support for the dovish preference mechanisms than for the domestic constraint and democratic credibility mechanisms. We find very strong evidence that democracies have reputations for dovish preferences. Knowing that an actor is democratic causes respondents to see its public as 10 percentage points more likely to believe that force should only be used as a last resort and to see its leaders as 16 percentage points more likely to do so. Democracies have particularly strong reputations for sensitivity to the costs of war: democracies are seen as 37 percentage points more sensitive to casualties and 23 percentage points more sensitive to the financial costs of fighting than nondemocracies. However, we find only mixed evidence in favor of the domestic constraint mechanisms: the effect of backing down in the crisis on the leader losing power is only 3 percentage points higher in democracies than nondemocracies, and the effect of losing the war on the regime maintaining public support is only 2 percentage point lower in democracies than nondemocracies. The sole domestic constraint mechanism we find evidence consistent with involves the perceived likelihood of irregular exit

after backing down: knowing the actor is democratic causes respondents to perceive the leader as 35 percentage points less likely to be removed from office through coups, exile, or death. We also find strong evidence against democratic credibility theory. We find that democracies have reputations for lacking credibility in their threats: democratic leaders who issue threats are seen as 7 percentage points less likely to follow through than nondemocratic leaders. In general, then, regime type is seen as informative in predicting crisis dynamics, in a manner that paints a relatively negative picture: democratic reputations in crises are more consistent with the pessimistic predictions of theories of dovish preferences and cost sensitivity than the optimistic predictions of theories of democratic credibility or domestic constraint.

Figure 4B conducts a similar analysis for wars. Here, the IR literature points to two classes of mechanisms in particular: one concerning democratic selectivity, and the other military effectiveness. We find strong evidence of democratic reputations in war consistent with both sets of mechanisms. Democracies appear to have reputations for being smart about which wars they fight: knowing the actor is democratic causes respondents to see it as 16 percentage points more likely to win if it had initiated the conflict than if that same actor were nondemocratic. Similarly, we find consistent evidence that democracies have strong reputations for military effectiveness: knowing that the actor is democratic causes respondents to perceive it to be 18 percentage points more likely that other countries will come to the actor's defense, 13 percentage points more likely that the morale of the soldiers is high, and 22 percentage points more likely that the soldiers are well trained.

Similar to figure 4A, however, we find only mixed evidence that democracies have reputations for domestic constraint: the effect of losing the war on the leader losing power is only 5 percentage points higher in democracies, and the effect of losing the war on the regime maintaining public support is only 6 percentage points lower in democracies. As before, we do see a very large effect of regime type on the perceived likelihood of irregular exit after defeat: knowing the actor is democratic causes respondents to perceive the leader as 52 percentage points less likely to be removed from office through coups, exile, or death. In general, then, these findings show strong evidence that democratic reputations exist, but unlike in crises, democratic reputations in war are more positive, consistent with theories of democratic selectivity and military effectiveness. The question about why democracies have asymmetric reputations—stronger in war than in crises—is one we turn to below.

The prevalence of democratic reputations

The above analysis provides a nuanced look at democratic reputations in crises and war but prioritizes depth over

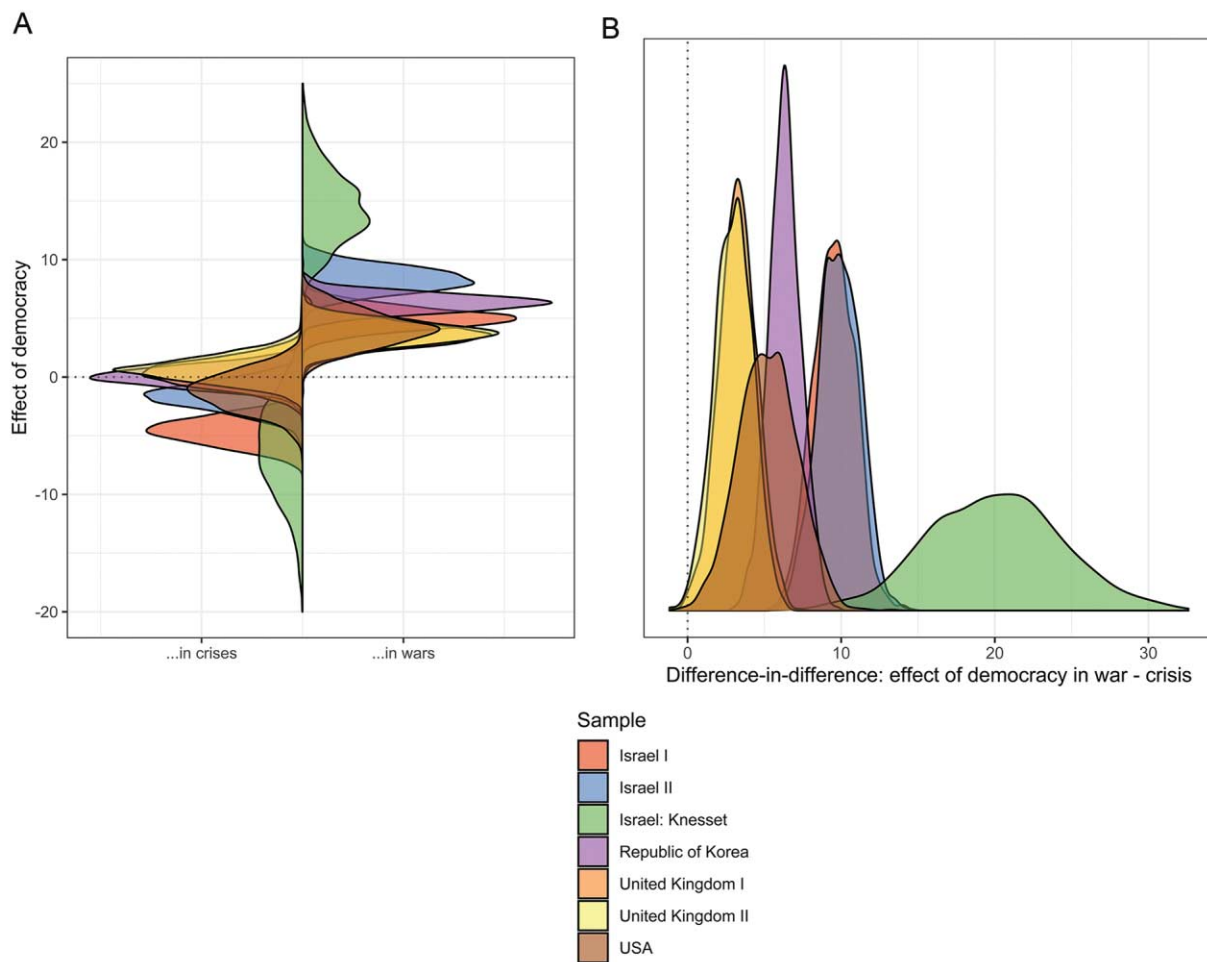


Figure 5. A, Distributions of average treatment effects (calculated using $B = 1,500$ bootstraps) show democracies do not have a reputation for standing firm in crises (and, in the Knesset and one of the Israeli public samples, are believed to be significantly more likely to back down) but always have a reputation for winning wars. B, Bootstrapped difference-in-difference results show that democratic reputations are significantly more favorable in wars than in crises across all seven samples. Note that the greater variance in the Knesset sample is due to its smaller sample size; see appendix section C.5 for results with downsampled public distributions that feature a similar spread.

breadth, studying democratic reputations for a wide range of different considerations in crises and war (finding that democratic reputations exist and that they are generally more favorable in war than in crises) but in a single sample in a single country. To provide additional breadth, we therefore replicate our primary outcomes in six other samples across four different democracies, allowing us to shed greater light on hypothesis 4 on the prevalence of democratic reputations across mass samples and between mass and elite populations.

Figure 5A displays bootstrapped density distributions of the average treatment effects for our primary outcome measures across all seven samples. Four patterns are evident. First, across all seven samples, respondents never perceived democracies as possessing an advantage in crises. In two of the samples (the Knesset sample and one of the Israeli public samples), we find democracies are in fact seen as significantly less likely to stand firm than dictatorships (by 5.9 and 4.6 percentage

points, respectively).¹³ On the whole, then, democracies either lack a reputation for resolve in crises in the eyes of our respondents or have a reputation for displaying less resolve than nondemocratic states. Second, across all seven samples, respondents perceived democracies as possessing a significant advantage in wars; the effect was particularly pronounced among our Knesset respondents (14.4 percentage points) but is positive and statistically significant in all of the other samples as well.

Third, as figure 5B shows, the difference in difference between democracies and dictatorships in crises versus wars is statistically significant across each of our seven samples: in

13. In app. sec. C.6, we show that these effect sizes are comparable to those from large-N studies of democracy and victory in crisis bargaining drawn from the MID data. Full results for Knesset sample are depicted in table 5 in app. sec. C.1. Results for the Israeli public are reproduced in app. sec. C.3.

each country, democratic reputations in war were seen as significantly more favorable than democratic reputations in crises. Altogether, our results show that democracies have asymmetric reputations among democratic leaders and their publics, who see democracies as having significant advantages in full-scale interstate wars but no corresponding advantage in crises. This begs both a practical question—How might we interpret the relative magnitude of the effects in the different domains?—as well as a theoretical one related to how the two parts of our experiment connect and what light they shed on theoretical questions related to the study of reputations.

On the practical interpretation of the results, we can benchmark our outcomes through a comparison with extant (observational) work on related subjects. While direct comparisons to observational work are difficult, appendix section C.6 suggests that the effect sizes we find for democracies in crises in our experiment (a maximum value of 6%) are similar to those found in one of the canonical large-N studies of this question (Schultz 2001).

Our results across the domains of crisis and war suggest answers to larger theoretical questions about the nature of reputations and, in particular, the extent to which they “transfer” from the context in which they were generated. Hypotheses 2 and 3 are about the same set of actors (democracies) in a broadly similar domain (international conflict). Yet, the qualities required in actual war fighting are not identical to those required in compellence episodes. A simplistic view would see reputations as effectively transferring if reputations in wars and crises were the same for democracies and as evidence against the ability to “redirect” reputations if results were of a different sign (or one set of results was null and the other was not). However, a more nuanced view would be to consider the comparison between our observed results and a counterfactual in which only one of those two reputations existed. In appendix section C.7, we use sequential g-estimation (Acharya, Blackwell, and Sen 2016) to estimate a series of average controlled direct effect models where we look at the direct effect of democracy on war outcomes, controlling for the “resolve in crises” mediator without inducing posttreatment bias. We find that our treatment effect slightly increases in magnitude once you control for the crisis mediator, implying that democratic reputations in war would be even stronger if not for their perceived disadvantage in crises. We find similar evidence in a simulation approach in appendix section C.7, showing that respondents’ beliefs about democratic advantages in war simply swamp their concerns about democratic disadvantages in crises. Altogether, this provides some limited support for the notion that reputations “transfer” or carry over to other domains, even if that effect is through suppressing a positive reputation in another context.

Fourth, focusing specifically on the Israeli samples, our results reveal a general similarity in the judgments displayed by our elite decision makers in the Knesset and the mass public they represent. As is starkly illustrated in figure 5, both our elite and our public samples espoused democratic pessimism in crises and democratic triumphalism in wars: democracies were believed to be less likely to stand firm in disputes and more likely to win conflicts that escalated to the use of force.

This congruence is notable given the intensity of debates about the extent to which elite decision makers differ from members of the mass public (e.g., Hafner-Burton, Hughes, and Victor 2013; Kertzer 2022; Linde and Vis 2017; Mintz et al. 2006). In both cases, any difference between the three samples was in magnitude not direction: leaders were both slightly more pessimistic about the odds for democracies in crises and also a bit more optimistic about their chances in war. While the congruence between our samples should be interpreted with caution, our results remind us that differences between elites and ordinary citizens should not be overstated without a theory as to why we might expect characteristics of elites to moderate the impact of treatment effects (Kertzer 2022; Renshon 2015; Yarhi-Milo et al. 2018). Indeed, supplementary analyses find relatively little sign of heterogeneous treatment effects within our Israeli samples: current MKs behave similarly to former MKs, more hawkish MKs behave similarly to more dovish MKs, and so on.¹⁴

CONCLUSION

If we take leaders at their word, actors in international politics frequently use regime type to draw inferences about states’ future behavior. Democracies have reputations, which affect both how other actors treat them and how they themselves behave. Yet, although there is a robust literature on reputation in IR, it has traditionally avoided conceptualizing reputations as adhering to regime types. And, although there is vibrant literature exploring the ways that democracy matters in international politics, it largely focuses on the direct effects of domestic institutions on state behavior, rather than examining this perceptual pathway. We sought to place democratic reputations in crises and war under the microscope, fielding a survey experiment on an unusually elite sample of past and present Israeli MKs, thereby providing direct evidence on the existence of democratic reputations in the eyes of democratic

14. More formally, we estimate a series of simple regression models in which we regress each dependent variable (either resolve or military effectiveness) on the democracy treatment, the respondent-level characteristic under investigation, and the interaction between the two. The interactions lack statistical significance (on resolve: $p < .13$ for current MK status, $p < .40$ for military assertiveness; on military effectiveness: $p < .71$ for current MK status, $p < .41$ for military assertiveness).

leaders. We also field studies on six public samples in four democracies to gain further insight into the content and prevalence of these reputations.

Across all seven samples, we find that democracies have asymmetric reputations that are consistently more favorable in war than in crises. This is true both in terms of our primary outcomes of interest—whether democracies are seen as standing firm at the bargaining table or winning on the battlefield—as well as a host of secondary measures: democratic leaders and publics are seen as more dovish in their preferences, and less credible in their threats, but democracies are seen as more selective in the wars they fight, with better trained and more effective militaries. We find that these reputations are pervasive: across six samples and four different democracies, similar patterns of beliefs emerge. Related to the question of pervasiveness, we also find strikingly similar findings between our elite and mass samples, offering one more data point demonstrating that we should not uncritically assume that leaders and the public will differ dramatically.

Our results suggest a number of implications. First, our finding that democratic leaders and publics perceive themselves as having no advantage over autocracies in signaling resolve in crises suggests that perhaps democratic leaders might attempt costlier signals than they would otherwise to compensate for their perception of weakness in “contests of will.” Second, our finding that democratic leaders and their publics see themselves as more likely than other regimes to win wars suggests that we might see observe overconfidence and increased risk acceptance among those groups in the lead-up to war. Finally, if reputations for resolve and war winning adhere to regime type—as our findings indicate—we should expect democratic leaders to show systematic differences in their willingness to engage in conflict with democracies versus autocracies. Indeed, in this sense, our findings provide micro-foundational support for the democratic peace’s central tenet that democracies are significantly less likely to fight each other.

And while our results show relatively stark differences in democratic reputations across different domains, there is plenty of scope for future research to examine the causes of this gulf. Our findings show that differences exist across the domains of crises and war fighting and that the (positive) reputations for war fighting that democracies possess swamp their less positive reputations for winning crises. Moreover, our results suggest that democracies’ reputations in war would be even stronger if not for their perceived disadvantage in crises. This, in turn, is suggestive of reputations in one domain carrying over to affect reputations in another, in line with work by Wiegand (2011) on the transferability of reputations. Future work might probe this finding through designs that assess the causal interrelationship between reputations in different domains.

Our conceptual framework—built on the notion that reputations may be “owned” by actors other than states and be “about” things other than resolve and built not just through a tabulation of past actions—also carries implications for future research. If reputations are formed through an interaction between past actions and other factors (such as discourse), it suggests the opportunity for experiments that manipulate more than just whether an actor stood firm in the last interaction. And while our experiments focused on the impact of domestic institutions on international reputations, future work might take up implications suggested by Schultz (2005) and examine how interactions between competing parties domestically might affect the dynamics we find in our results. Relatedly, a fruitful avenue for extension would be to consider how leader- and party-level reputations interact with regime-level ones (Kertzer et al. 2021).

Our framework and results also suggest a new twist on the relationship between domestic politics and IR. Fearon (1998) suggests two ways in which domestic politics may matter, either through causing states to pursue suboptimal foreign policies or when differences in political institutions are causally relevant in explaining different foreign policy choices. Our work suggests a third possibility: that domestic politics might matter by causally affecting *beliefs* about foreign policy choices, which in turn might have effects at the international level. Future work unpacking this question might begin by linking reputations to foreign policy choices made by the actor holding the reputational belief (reversing the common method of explaining A’s choices by reference to B’s past behavior).

The nature of the samples also suggests fruitful avenues for replications and extensions. For example, our experimental design fixes a number of attributes of the countries involved in the dispute. While those design decisions served the purpose of enabling more precise identification and minimized concerns about confounding, future research might relax some of those constraints. It could be, for example, that the democratic reputations we identify here are modified by specific aspects of the country such as its size and power or the nature of its opponent. And while our subjects were drawn from a variety of different democratic societies, future research should examine similar questions in nondemocratic societies. Such a comparison—between perceptions of democratic citizens and their nondemocratic counterparts—will significantly improve our understanding of the strategic interaction between different regimes. If autocracies shared the beliefs we find evidence of in our studies, they might be less likely to initiate conflicts for fear of escalation to war, which they see themselves at a disadvantage in; conversely, if they hold different beliefs about democratic prowess in war, they might be correspondingly more risk acceptant. Relatedly, further study might engage

with second-order beliefs: What do democratic leaders perceive to be the beliefs of autocratic leaders and vice versa? Finally, future research on the role of regime type in international security could formulate and test domain-specific claims about the effects of democracy in contexts such as counterterrorism, counterinsurgency, and low-intensity conflicts.

ACKNOWLEDGMENTS

This is one of several joint articles by the authors; the ordering of names reflects a principle of rotation. We gratefully acknowledge the support of the Weatherhead Center for International Affairs and the Institute for Quantitative Social Science at Harvard University; Perry Abdulkadir, Lotem Bas-san, Alice Hu, Aaron Miller, and Michal Sagiv for invaluable research assistance; Mike Tomz, Jessica Weeks, Roni Milo, Shirley Avrahami, and Yardena Miller at the Knesset—without whom the elite survey would not have been possible—Matt Blackwell, Matt Fuhrmann, Guy Grossman, Mike Horowitz, Luke Keele, Debbie Larson, Christoph Mikulaschek, Katy Powers, Elizabeth Saunders, Ken Schultz, Todd Sechser, Jake Shapiro, and Teppei Yamamoto; and participants at the Leaders and Military Conflict workshop at the Peace Science Society Annual Meeting and American Political Science Association 2016; and audiences at Columbia, Dartmouth, Georgetown, Princeton, the University of Chicago, the University of Virginia, Texas A&M, Yale, and Stanford for helpful feedback at various stages of the project. Finally, we thank the Knesset members who took the time to participate in the survey.

REFERENCES

- Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2016. "Explaining Causal Findings without Bias: Detecting and Assessing Direct Effects." *American Political Science Review* 110 (3): 512–29.
- Alatas, Vivi, Lisa Cameron, Ananish Chaudhuri, Nisvan Erkal, and Lata Gangadharan. 2009. "Subject Pool Effects in a Corruption Experiment." *Experimental Economics* 12 (1): 113–32.
- Baum, Matthew A., and Philip B. K. Potter. 2015. *War and Democratic Constraint: How The Public Influences Foreign Policy*. Princeton, NJ: Princeton University Press.
- Brown, Jonathon D. 1986. "Evaluations of Self and Others: Self-Enhancement Biases in Social Judgments." *Social Cognition* 4 (4): 353–76.
- Brutger, Ryan, and Joshua D. Kertzer. 2018. "A Dispositional Theory of Reputation Costs." *International Organization* 72 (3): 693–724.
- Brutger, Ryan, Joshua D. Kertzer, Jonathan Renshon, Dustin Tingley, and Chagai M. Weiss. 2022. "Abstraction and Detail in Experimental Design." *American Journal of Political Science*. <https://doi.org/10.1111/ajps.12710>.
- Bueno de Mesquita, Bruce, and David Lalman. 1992. *War and Reason*. New York: Cambridge University Press.
- Bueno de Mesquita, Bruce, James D. Morrow, Randolph M. Siverson, and Alastair Smith. 1999. "An Institutional Explanation of the Democratic Peace." *American Political Science Review* 93 (4): 791–808.
- Bush, Sarah Sunn, and Pär Zetterberg. 2021. "Gender Quotas and International Reputation." *American Journal of Political Science* 65 (2): 326–41.
- Carnevale, Jessica J., Yoel Inbar, and Jennifer S. Lerner. 2011. "Individual Differences in Need for Cognition and Decision-Making Competence among Leaders." *Personality and Individual Differences* 51 (3): 274–78.
- Caverley, Jonathan D. 2010. "Power and Democratic Weakness: Neo-conservatism and Neoclassical Realism." *Millennium* 38 (3): 593–614.
- Chan, Steve, and William Safran. 2006. "Public Opinion as a Constraint against War: Democracies' Responses to Operation Iraqi Freedom." *Foreign Policy Analysis* 2 (2): 137–56.
- Clinton, Bill. 1993. "Address by President Bill Clinton to the UN General Assembly." <https://2009-2017.state.gov/p/io/potusunga/207375.htm>.
- Copeland, Dale C. 1997. "Do Reputations Matter?" *Security Studies* 7 (1): 33–71.
- Dafoe, Allan, Jonathan Renshon, and Paul Huth. 2014. "Reputation and Status as Motives for War." *Annual Review of Political Science* 17:371–93.
- Dafoe, Allan, Baobao Zhang, and Devin Caughey. 2018. "Information Equivalence in Survey Experiments." *Political Analysis* 26 (4): 399–416.
- Debs, Alexandre, and Hein E. Goemans. 2010. "Regime Type, the Fate of Leaders, and War." *American Political Science Review* 104 (3): 430–45.
- Desch, Michael C. 2008. *Power and Military Effectiveness: The Fallacy of Democratic Triumphalism*. Baltimore: Johns Hopkins University Press.
- Doyle, Michael W. 1986. "Liberalism and World Politics." *American Political Science Review* 80 (4): 1151–69.
- Eden, Anthony. 1962. *The Eden Memoirs: Facing the Dictators*. London: Cassell.
- Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88 (3): 577–92.
- Fearon, James D. 1998. "Domestic Politics, Foreign Policy, and Theories of International Relations." *Annual Review of Political Science* 1:289–313.
- Findley, Michael G., Adam S. Harris, Helen V. Milner, and Daniel L. Nielson. 2017. "Who Controls Foreign Aid? Elite versus Public Perceptions of Donor Influence in Aid-Dependent Uganda." *International Organization* 71 (4): 633–63.
- Foa, Roberto Stefan, and Yascha Mounk. 2016. "The Danger of Deconsolidation: The Democratic Disconnect." *Journal of Democracy* 27 (3): 5–17.
- Gaddis, John Lewis. 2011. *George F. Kennan: An American Life*. New York: Norton.
- Gartner, Scott Sigmund, and Gary M. Segura. 1998. "War, Casualties, and Public Opinion." *Journal of Conflict Resolution* 42 (3): 278–300.
- Gelpi, Christopher, and Joseph M. Grieco. 2001. "Attracting Trouble: Democracy, Leadership Tenure, and the Targeting of Militarized Challenges, 1918–1992." *Journal of Conflict Resolution* 45 (6): 794–817.
- Gelpi, Christopher F., and Michael Griesdorf. 2001. "Winners or Losers? Democracies in International Crisis, 1918–94." *American Political Science Review* 95 (3): 633–47.
- Gibler, Douglas M. 2008. "The Costs of Reneging: Reputation and Alliance Formation." *Journal of Conflict Resolution* 52 (3): 426–54.
- Goemans, Hein E. 2008. "Which Way Out? The Manner and Consequences of Losing Office." *Journal of Conflict Resolution* 52 (6): 771–94.
- Guisinger, Alexandra, and Elizabeth N. Saunders. 2017. "Mapping the Boundaries of Elite Cues: How Elites Shape Mass Opinion across International Issues." *International Studies Quarterly* 61 (2): 425–41.
- Guisinger, Alexandra, and Alastair Smith. 2002. "Honest Threats: The Interaction of Reputation and Political Institutions in International Crises." *Journal of Conflict Resolution* 46 (2): 175–200.

- Hafner-Burton, Emilie M., Alex Hughes, and David G. Victor. 2013. "The Cognitive Revolution and the Political Psychology of Elite Decision Making." *Perspectives on Politics* 11 (2): 368–86.
- Hafner-Burton, Emilie M., Brad L. LeVeck, David G. Victor, and James H. Fowler. 2014. "Decision Maker Preferences for International Legal Cooperation." *International Organization* 68 (4): 845–76.
- Hall, Peter A. 2013. "Anxiety about Democracy: Why Now?" SSRC Democracy Papers.
- Hayes, Jarrod. 2012. "Securitization, Social Identity, and Democratic Security: Nixon, India, and the Ties That Bind." *International Organization* 66 (1): 63–93.
- Hilton, James L., and William von Hippel. 1996. "Stereotypes." *Annual Review of Psychology* 47:237–71.
- Holsti, Ole R., and James N. Rosenau. 1988. "The Domestic and Foreign Policy Beliefs of American Leaders." *Journal of Conflict Resolution* 32 (2): 248–94.
- Huntington, Samuel P. 1991. *The Third Wave: Democratization in the Late Twentieth Century*. Norman: University of Oklahoma Press.
- Huth, Paul. 1988. "Extended Deterrence and the Outbreak of War." *American Political Science Review* 82 (2): 423–43.
- Hyde, Susan D., and Elizabeth N. Saunders. 2020. "Recapturing Regime Type in International Relations: Leaders, Institutions, and Agency Space." *International Organization* 74 (2): 363–95.
- Jervis, Robert. 2002. "Signaling and Perception: Drawing Inferences and Projecting Images." In Kristen Monroe, ed., *Political Psychology*. Hillsdale, NJ: Erlbaum, 293–312.
- Jervis, Robert, Keren Yarhi-Milo, and Don Casler. 2021. "Redefining the Debate over Reputation and Credibility in International Security: Promises and Limits of New Scholarship." *World Politics* 73 (1): 167–203.
- Johns, Robert, and Graeme A. M. Davies. 2019. "Civilian Casualties and Public Support for Military Action: Experimental Evidence." *Journal of Conflict Resolution* 63 (1): 251–81.
- Kertzer, Joshua D. 2022. "Re-assessing Elite-Public Gaps in Political Behavior." *American Journal of Political Science* 66 (3): 539–53.
- Kertzer, Joshua D., Deborah Jordan Brooks, and Stephen G. Brooks. 2021. "Do Partisan Types Stop at the Water's Edge?" *Journal of Politics* 83 (4): 1764–82.
- Kertzer, Joshua D., and Ryan Brutger. 2016. "Decomposing Audience Costs: Bringing the Audience Back into Audience Cost Theory." *American Journal of Political Science* 60 (1): 234–49.
- Kertzer, Joshua D., and Jonathan Renshon. 2022. "Experiments and Surveys on Political Elites." *Annual Review of Political Science* 25:529–50.
- Kertzer, Joshua D., Jonathan Renshon, and Keren Yarhi-Milo. 2021. "How Do Observers Assess Resolve?" *British Journal of Political Science* 51 (1): 308–30.
- Kertzer, Joshua D., and Thomas Zeitzoff. 2017. "A Bottom-Up Theory of Public Opinion about Foreign Policy." *American Journal of Political Science* 61 (3): 543–58.
- Kreps, Sarah, Elizabeth N. Saunders, and Kenneth Schultz. 2018. "The Ratification Premium: Hawks, Doves, and Arms Control." *World Politics* 70 (4): 479–514.
- Lake, David A. 1992. "Powerful Pacifists: Democratic States and War." *American Political Science Review* 86 (1): 24–37.
- Levin, Dov H., and Robert F. Trager. 2019. "Things You Can See from There You Can't See from Here: Blind Spots in the American Perspective in IR and Their Effects." *Journal of Global Security Studies* 4 (3): 345–57.
- Linde, Jona, and Barbara Vis. 2017. "Do Politicians Take Risks Like the Rest of Us? An Experimental Test of Prospect Theory under MPs." *Political Psychology* 38 (1): 101–17.
- Lupu, Noam. 2016. *Party Brands in Crisis: Partisanship, Brand Dilution, and the Breakdown of Political Parties in Latin America*. Cambridge: Cambridge University Press.
- Manekin, Devorah, Guy Grossman, and Tamar Mitts. 2015. "The Domestic Sources of Territorial Conflict: The Case of Israel." Working paper.
- Marshall, Monty G., and Ted Robert Gurr. 2019. "Polity5: Political Regime Characteristics and Transitions, 1800–2018." Online data set.
- Martin, Lisa L. 2000. *Democratic Commitments: Legislatures and International Cooperation*. Princeton, NJ: Princeton University Press.
- Mattes, Michaela, and Mariana Rodriguez. 2014. "Autocracies and International Cooperation." *International Studies Quarterly* 58 (3): 527–38.
- Mercer, Jonathan. 1996. *Reputation and International Politics*. Ithaca, NY: Cornell University Press.
- Mintz, Alex, and Nehemia Geva. 1993. "Why Don't Democracies Fight Each Other? An Experimental Study." *Journal of Conflict Resolution* 37 (3): 484–503.
- Mintz, Alex, Steven B. Redd, and Arnold Vedlitz. 2006. "Can We Generalize from Student Experiments to the Real World in Political Science, Military Affairs, and International Relations?" *Journal of Conflict Resolution* 50 (5): 757–76.
- Mueller, John E. 1971. "Trends in Popular Support for the Wars in Korea and Vietnam." *American Political Science Review* 65 (2): 358–75.
- Narang, Vipin, and Paul Staniland. 2018. "Democratic Accountability and Foreign Security Policy: Theory and Evidence from India." *Security Studies* 27 (3): 410–47.
- Page, Benjamin I., and Marshall M. Bouton. 2007. *The Foreign Policy Disconnect: What Americans Want from Our Leaders but Don't Get*. Chicago: University of Chicago Press.
- Pape, Robert A. 2003. "The Strategic Logic of Suicide Terrorism." *American Political Science Review* 97 (3): 343–61.
- Post, Abigail S., and Paromita Sen. 2020. "Why Can't a Woman Be More like a Man? Female Leaders in Crisis Bargaining." *International Interactions* 46 (1): 1–27.
- Press, Daryl Grayson. 2005. *Calculating Credibility: How Leaders Assess Military Threats*. Ithaca, NY: Cornell University Press.
- Reiter, Dan, and Allan C. Stam. 2002. *Democracies at War*. Princeton, NJ: Princeton University Press.
- Renshon, Jonathan. 2015. "Losing Face and Sinking Costs: Experimental Evidence on the Judgment of Political and Military Leaders." *International Organization* 69 (3): 659–95.
- Renshon, Jonathan, Allan Dafoe, and Paul Huth. 2018. "Leader Influence and Reputation Formation in World Politics." *American Journal of Political Science* 62 (2): 325–39.
- Renshon, Jonathan, and Arthur Spirling. 2015. "Modeling 'Effectiveness' in International Relations." *Journal of Conflict Resolution* 59 (2): 207–38.
- Risse-Kappen, Thomas. 1995. "Democratic Peace—Warlike Democracies? A Social Constructivist Interpretation of the Liberal Argument." *European Journal of International Relations* 1 (4): 491–517.
- Rose, Richard, William Mishler, and Christian Haerpfer. 1998. *Democracy and Its Alternatives: Understanding Post-Communist Societies*. Baltimore: Johns Hopkins University Press.
- Schelling, Thomas C. 1966. *Arms and Influence*. New Haven, CT: Yale University Press.
- Schultz, Kenneth A. 1999. "Do Democratic Institutions Constrain or Inform? Contrasting Two Institutional Perspectives on Democracy and War." *International Organization* 53 (2): 233–66.
- Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy*. New York: Cambridge University Press.
- Schultz, Kenneth A. 2005. "The Politics of Risking Peace: Do Hawks or Doves Deliver the Olive Branch?" *International Organization* 59 (1): 1–38.
- Sechser, Todd S. 2018. "Reputations and Signaling in Coercive Bargaining." *Journal of Conflict Resolution* 62 (2): 318–45.

- Sheffer, Lior, Peter John Loewen, Stuart Soroka, Stefaan Walgrave, and Tamir Sheafer. 2018. "Nonrepresentative Representatives: An Experimental Study of the Decision Making of Elected Politicians." *American Political Science Review* 112 (2): 302–21.
- Snyder, Glenn H., and Paul Diesing. 1977. *Conflict among Nations: Bargaining, Decision Making, and System Structure in International Crises*. Princeton, NJ: Princeton University Press.
- Snyder, James M., and Michael M. Ting. 2002. "An Informational Rationale for Political Parties." *American Journal of Political Science* 46 (1): 90–110.
- Tajfel, Henri. 1963. "Stereotypes." *Race* 5 (2): 3–14.
- Tang, Shiping. 2005. "Reputation, Cult of Reputation, and International Conflict." *Security Studies* 14 (1): 34–62.
- Taylor, Shelley E. 1981. "A Categorization Approach to Stereotyping." In David L. Hamilton, ed., *Cognitive Processes in Stereotyping and Intergroup Behavior*. Hillsdale, NJ: Erlbaum, 83–114.
- Tomz, Michael, Jessica L. P. Weeks, and Keren Yarhi-Milo. 2020. "Public Opinion and Decisions about Military Force in Democracies." *International Organization* 74 (1): 119–43.
- Tomz, Michael R., and Jessica L. P. Weeks. 2013. "Public Opinion and the Democratic Peace." *American Political Science Review* 107 (4): 849–65.
- Valentino, Benjamin A., Paul K. Huth, and Sarah E. Croco. 2010. "Bear Any Burden? How Democracies Minimize the Costs of War." *Journal of Politics* 72 (2): 528–44.
- Weeks, Jessica L. 2008. "Autocratic Audience Costs: Regime Type and Signaling Resolve." *International Organization* 62 (1): 35–64.
- Weeks, Jessica L. 2012. "Strongmen and Straw Men: Authoritarian Regimes and the Initiation of International Conflict." *American Political Science Review* 106 (2): 326–47.
- Weeks, Jessica L. P. 2014. *Dictators at War and Peace*. Ithaca, NY: Cornell University Press.
- Weisiger, Alex, and Keren Yarhi-Milo. 2015. "Revisiting Reputation: How Past Actions Matter in International Politics." *International Organization* 69 (2): 473–95.
- Wiegand, Krista E. 2011. "Militarized Territorial Disputes: States' Attempts to Transfer Reputation for Resolve." *Journal of Peace Research* 48 (1): 101–13.
- Yarhi-Milo, Keren. 2014. *Knowing the Adversary: Leaders, Intelligence, and Assessment of Intentions in International Relations*. Princeton, NJ: Princeton University Press.
- Yarhi-Milo, Keren. 2018. *Who Fights for Reputation in International Politics? Leaders, Resolve, and the Use of Force*. Princeton, NJ: Princeton University Press.
- Yarhi-Milo, Keren, Joshua D. Kertzer, and Jonathan Renshon. 2018. "Tying Hands, Sinking Costs, and Leader Attributes." *Journal of Conflict Resolution* 62 (10): 2150–79.
- Zaller, John R. 1992. *The Nature and Origins of Mass Public Opinion*. Cambridge: Cambridge University Press.