Tying hands, sinking costs, and leader attributes: Supplementary appendix

May 25, 2018

Contents

1 Recruitment protocol

1.1 Recruitment letter to Knesset Members

הנך מוזמן להשתתף במחקר בחסות פרופסורים מאוניבריטאות פרינסטון, הרווארד, וסטנפורד שבארצות הברית.

פרויקט זה בוחן כיצד מקבלי החלטות חושבים על מדיניות חוץ. השאלון מיועד לנשים וגברים כאחד.

נהלי המחקר:!המחקר כרוך בהשלמת שאלון אמריקאי ויקח כ "! דקות להשלים. שאלון זה מופץ לכל חברי הכנסת בישראל בהווה או בשני העשורים האחרונים. בכוונתנו לערוך בעתיד מחקר השוואתי עם חברי הפרלמנט הבריטי ואליטות בארצות הברית.

סודיות: מחקר זה הינו אקדמי בלבד, ועל מנת לשמור על אמינותו המחקרית אנו מחויבים לשמר את האנונימיות של המשתתפים. התשובות שלך יהיו אנונימיות לחלוטין ללא אפשרות לקשר את זהותך לתשובות שאתה מספק.

להתחיל את השאלון, אנא הקש על הקישור הבא: [redacted]

אתה תתבקש להזין סיסמה בעלת ! ספרות. סיסמה זו הינה אישית,סודית, ונועדה לשימושך הפרטי בלבד.
הסיסמה שלך היא: XXXXX

בשל הפורמט של השאלון, אנו מבקשים שלא תענה על השאלון מהמכשיר הסלולרי שלך.(אם אתה מעדיף שנשלח לך עותק קשיח של השאלון באי מייל או בדואר, נשמח לעשות זאת!)

אנו מודים לך על השתתפותך במחקר זה. אנו יודעים שזמנך יקר מאוד, ומעריכים את תרומתך למחקר בנושאים חשובים אלו. נשמח לשתף אותך בתוצאות המחקר לאחר שנעבד את הנתונים.

אם יש לך שאלות אנא פנה לפרופ' קרן ירחי מילוא מאוניברסיטת פרינסטון באימייל:
[redacted] .

בברכה,

מייקל טומז, יהושע דויד קרצר, קרן ירחי מילוא,
אוניברסיטת פרינסטון !

Figure 1: Recruitment Letter

1.2 Recruitment procedures

We began the recruitment process by compiling a dataset of all 415 individuals who had served as members of the Parliament of Israel (i.e., the Knesset) from the beginning of the 14th Knesset in June 1996 through the 20th Knesset (the current Knesset) that was sworn in in March 2015. We compiled a data set that included the following information about our population:

1. full name
2. party affiliation while in Knesset
3. names of all Knesset committees on which (s)he served
4. number of terms served
5. whether (s)he served as a minister in the government, and if so, what portfolios (s)he held
6. whether (s)he was a member of the Cabinet

Contact information for our participants was obtained through a variety of channels, including the Secretary of the Knesset, the Knesset Channel, the different parties' leadership offices in the Knesset and other government offices where former Knesset members are currently employed. Email addresses for all current members of the Knesset were obtained through the Secretary of the Knesset. To verify whether the contact information we obtained was correct, we either called or emailed all the former Knesset members from the last twenty years and asked them if they would be interested in taking a "10 minute electronic survey by a team of professors from leading American Universities." 30.6% of the initial population was removed from the sampling frame at this stage, either because the members were deceased, were too sick to participate, or because their contact information was out of date and newer contact information could not be found. This process left us with a sample of 288 potential candidates to take our survey. This pool included all 120 current members of the Knesset along with 168 former members whose contact information was available.

In July 2015, we executed a soft launch of our on-line survey. The survey included a recruitment email, written in Hebrew (reproduced in Appendix §1.1), a link to our on-line survey, and an individual six-digit password that was pre-assigned to each member. In the following days, we emailed the invitation to all current and former members in our dataset. A few weeks later, we sent a reminder email to those who had not responded to the survey. We sent a third round of reminders a few weeks later. In between these rounds, we phoned former and current Knesset members or their assistants to remind them to take the survey. In early August, the Director of Academic Affairs at the Knesset, together with the Secretary of the Knesset, sent an email to all current Knesset members encouraging them to take the survey, repeating essentially the same information we provided in the introductory email.

In addition to the on-line survey, we created identical hard-copy versions of our survey. In mid-August we sent those who had not responded to our survey a reminder email and attached an electronic copy of our survey that could be opened in Microsoft Word. Respondents were given the option of either faxing or emailing the completed survey back to us. That same code was the only

identifying information on the paper copies of the survey, allowing us to track completion among our sample population. Members of our research team also traveled to the Knesset on four separate occasions to invite current members to participate.

The entire recruitment process was done in Hebrew. Two Hebrew-speaking research assistants and one member of the research team who is a native Hebrew speaker corresponded with the members of the Knesset or their assistants. Participants were informed that there would receive no financial reward for taking the survey, but that we would be happy to share with them the results of the survey. Moreover, participants were promised full anonymity: with the exception of the research team, participants were assured that identifiable information would not be released or reported.

1.3 Participant verification protocol

We took several steps to increase our confidence that the current and former decision-makers participated in the study rather than members of their staff. First, in the introductory email we explicitly indicated that the questionnaire should be fielded by the decision-maker himself, and not by members of his or her staff. We explained that the code we provided to access the on-line survey was personal, and should not be shared with others. Importantly, we did not offer any material incentives for filling out the survey, to dissuade decision-makers and assistants for taking the survey for those material reasons.

Second, in the survey itself we asked the participants to enter their complete date of birth. This allowed us to compare this information with the date of the decision-makers in official Knesset records. Third, for the 75% of our sample consisting of former Knesset members, a Hebrew-speaking research assistant and one of the authors were both in touch with the decision-maker directly via phone or email, and confirmed with him/her that they were the ones taking the survey. Anecdotally, our research team found that many of our participants were quite eager for the opportunity to opine on issues of foreign policy to an outside audience.

In the case of some current Knesset members, after receiving approval from their parliamentary assistant, a Hebrew-speaking research assistant from our team or one of the authors gave the Knesset members the survey directly and picked it up from them within a two-hour window. However, some Knesset members wished to maintain their anonymity and thus were not in direct contact with the research team.

Finally, although we follow best practices, as is always the case with elite experiments, we should note that decision-makers who wished to "cheat" and delegate their participation to others could have probably found ways to do so. However, the combination of the types of questions asked in the survey, the absence of material compensation for survey completion, our explicit request the

survey not be filled out by others, and the enthusiastic response to our survey from most of the decision-makers who took the survey leave us confident that the vast majority of them participated directly.

2 Study instrumentation

2.1 Experimental protocol (translated to English)

Here is the situation:

- Your country — Israel — is involved in a dispute with Country B, a strong military dictatorship.
- The dispute began with a collision between an Israeli shipping vessel and a ship registered to Country B.
- During the collision, injuries were reported on both sides.
- Additionally, both countries maintain that their ship was carrying sensitive military technology, and are suspicious of the motives of the other side, leading to a tense standoff at sea.
- Currently, because of the remote location, the public is not aware of the incident.

[*Outcome 1 (Baseline)*] Given the information available, what is your best estimate about whether Country B will stand firm in this dispute, ranging from 0% to 100%?

[NEW SCREEN]

Now we would like to ask you a question about a different, alternative version of the scenario you just read. Suppose the basic details remain the same:

- Israel is involved in a dispute with a dictatorship with a strong military, Country B.
- The dispute began with a collision between an Israeli shipping vessel and a ship registered to Country B. During the collision, injuries were reported on both sides.
- Both countries maintain that their ship was carrying sensitive military technology, and are suspicious of the motives of the other side, leading to a tense standoff at sea.
- Currently, because of the remote location, the public is not aware of the incident.

But this time, suppose that...

×

[Tying Hands]: The President of Country B has issued a public statement through the news media warning that they will "do whatever it takes" to win this dispute.

[Sinking Costs]: Country B has mobilized their military and sent additional gunboats to the location of the dispute at sea.

Outcome 2 (Treatment)] Given the information available, what is your best estimate about whether Country B will stand firm in this dispute, ranging from 0% to 100%?

2.2 Individual difference measures

2.2.1 Military assertiveness

1. The best way to ensure peace is through military strength
2. The use of force generally makes problems worse
3. Rather than simply reacting to our enemies, it's better for us to strike first.

All scaled from 1 (strongly agree) to 5 (strongly disagree). Items 1 and 3 reverse-coded such that higher scores indicated higher level of military assertiveness.

2.2.2 International trust

1. Some people say that Israel can trust other nations, while others think that Israel can't be too careful in dealing with other nations. Where would you place yourself on this scale from 1 (Israel can count on other countries) to 7 (Israel cannot count on other countries)?

Reverse-scored such that higher scores indicate more higher levels of trust.

2.2.3 Right-wing ideology

1. There is much talk of "left" and "right" in politics. How would you rate yourself on a left-right scale, from 1 (right) to 7 (left)?

Reverse-scored such that higher scores indicate more right-wing ideology.

3 Supplementary analysis

3.1 Representativeness and survey non-response

There are two ways of thinking about the representativeness of our elite sample. The first asks how our participants compare to the complete population of individuals who served in the Knesset from 1996 to the present. The second asks how our participants compare to our sampling frame, a different group than the complete population because it does not include members who had passed away, were too sick to participate, or for whom we were unable to acquire up to date contact information. Thus, whereas the first quantity explores whether our participants look like the universe of Knesset members in this time period, the second explores survey non-response. We explore both questions in Table 1 below, which presents a set of linear probability models comparing our participants to the universe of Knesset members from 1996-2015 (models 1-2) and to only those Knesset members who had been sent the survey (models 3-4).[1] The results show that unsurprisingly, current members of the Knesset were less likely to participate in the survey than former members, but that interestingly, our participants are not significantly less "elite", as measured by the proportion of respondents with experience as deputy ministers, or as cabinet members or higher. If anything, our sample is slightly more experienced than the universe of decision-makers, though the number of terms in office did not significantly predict survey response.

Israeli legislative politics features a number of characteristics that makes calculating an summary partisan representativeness score somewhat complex, including a high degree of fragmentation, parties frequently splintering and forming, frequent party switching, and the presence of Arab and religious parties that cannot be cleanly positioned on a unidimensional partisan space. As a result, after coding all parties in the Knesset from 1996-2015 as being either left, center, right, Arab, religious, or none of the above, we look at partisan representativeness in two different ways. First, we focus on the representation of MKs from Arab parties (e.g. Hadash, Ra'am, Ta'al, Balad, the Joint Arab list, etc.), and religious parties (e.g. Shas, United Torah, Tkuma, Jewish Home, etc.), which are orthogonal to the left-right spectrum and thus are separated from the main analyses. We find that although both Arab and religious MKs are included in the sample, they are both slightly underrepresented: MKs from Arab parties make up approximately 7% of the population of MKs in this time range, but only 3% of the sample; MKs from religious parties make up approximately 20% of the population of MKs in this time frame, but only 7% of the sample.[2] Second, we take the

[1] For a similar strategy to test for balance on observables, see Table 2 of Blattman (2009).

[2] Given the presence of party switching, we code affiliation for Arab and religious parties here using a simple binary decision rule, in which an MK is classified as having a religious or Arab affiliation if they represented a religious or Arab party at any point in time during the time period under investigation. Thus, an MK who served a term as a member of a religious party, but then switched to a non-religious party in a subsequent term would retain their

Table 1: Gauging the representativeness of the sample

	Compared to...			
	All Knesset members		Sampling frame	
	(1)	(2)	(3)	(4)
Current member	−0.043	−0.049	−0.210***	−0.184***
	(0.045)	(0.057)	(0.054)	(0.065)
Highest level of experience:				
...Deputy minister	0.017	0.044	0.035	0.079
	(0.054)	(0.071)	(0.072)	(0.088)
...Cabinet member or higher	−0.044	−0.098	−0.075	−0.096
	(0.076)	(0.098)	(0.093)	(0.114)
Male	0.025	0.081	0.072	0.097
	(0.053)	(0.063)	(0.067)	(0.076)
Terms in office	0.011	0.021	0.008	0.013
	(0.012)	(0.016)	(0.015)	(0.018)
Left-right party membership		−0.070**		−0.063
		(0.030)		(0.038)
Constant	0.177***	0.312***	0.320***	0.436***
	(0.054)	(0.087)	(0.070)	(0.108)
N	415	295	288	225
R^2	0.007	0.043	0.063	0.080

*p < .1; **p < .05; ***p < .01

members who were exclusively associated with parties on the main ideological spectrum (left, center, and right), and calculate a left-right ideology score averaging across their terms in office within the sampling period, such that the lowest score (1) is assigned to an MK who consistently represented a left-wing party, and the highest score (3) is assigned to an MK who consistently represented a right-wing one. So, for example, an MK that served two terms in a left-wing party and a third term in a centrist party would be coded as being slightly more conservative (1.33) than an MK that served all of their terms in a left-wing party (1), and much less conservative than MK who served their terms in right-wing parties (3). Models 2 and 4 in Table 1 show that our sample is slightly more left-wing than the population of MKs in this time period as a whole, but interestingly, that this skew is not due to non-response bias, in that left-leaning MKs were not significantly more likely to participate in the survey than right-leaning ones. The partisan difference thus appears to stem from the probability of entering the sampling frame rather than the probability of response.

religious classification for the purpose of this analysis, and would be excluded from the left-right ideology analysis discussed below.

3.2 Generalizability outside the Israeli context

As we note in the main text, our choice to study how leaders understand costly signals using Israeli leaders is valuable for a number of reasons, both theoretically and methodologically. First, because of the Israeli parliamentary system, the majority of the executive branch are also elected members of the Knesset, such that the Knesset is comprised of policy makers who are directly involved in decisions about the use of force. Second, because of Israel's relatively short election cycles, it is common for former members of the executive to later become members of the opposition in the Knesset, such that nearly all of Israel's current Executive branch members were at one point members of the opposition. Thus, even those participants in our study who are currently opposition members have either been members of the Executive in the past, or are likely to be so in the future, such that we are effectively sampling current, former, and potentially future members of the Executive. Third, if the goal of our project is to examine how decision-makers perceive costly signals in foreign policy crises, it makes the most sense to do so in a country where these questions about war and peace are salient and realistic, where decision-makers have had to wrestle with exactly these kinds of issues — this, after all, is what is unique about studying elite decision-makers rather than college sophomores. Israel is a country where these "use of force" decisions are extremely common, rendering these questions particularly salient. Indeed, during the time frame in which our participants were in office (from 1996 onwards), Israel was involved in 16 Militarized Interstate Disputes (MIDs – Jones, Bremer and Singer, 1996); many our participants were involved in the decision-making process for many of these incidents. And, since we can expect that earlier experiences of Israeli conflict might have shaped leaders' beliefs, we note that Israel has been involved in 128 MIDs since the state's founding, and seven militarized compellent threats (Sechser, 2011).

Nonetheless, the value of the Israeli case does not preclude the need for thoughtful reflection concerning the generalizability of our findings. Typically this question is understood as implicating the tradeoff between internal and external validity. As Imbens (2010, 403) notes, the statistical literature has generally "emphasized internal validity over external validity, with the view that a credible estimate of the average effect for a subpopulation is preferred to an estimate of the average for the target population with little credibility." In political science, McDermott (2011) takes the same view, noting that one needs to know whether an effect exists at all before one can meaningfully ask whether it travels to other contexts. In a sense, though, the question is more subtle than how the tradeoff between internal and external validity is often understood by experimental critics: in their important critique of whether survey experiments lack external validity, for example, Barabas and Jerit (2010) show that one weakness of survey experiments is that they exaggerate treatment effects, because participants outside of the survey context are often less likely to receive the treatment in

the first place. In our case, however, the unique salience of foreign policy crises in Israel actually *mitigates* this aspect of external validity concerns!

Ultimately, the best means of exploring how our results might generalize outside the Israeli context is to address the question empirically, fielding the study in other democratic and nondemocratic regimes – a task we encourage interested readers to pursue. Short of that, however, we can address the question theoretically. We thus think through how our results might generalize outside the Israeli context by following best practices suggested by many scholars of comparative politics and qualitative work in IR (George and Bennett, 2005) by focusing on *within-case* variation: although Israel may differ from other countries in important ways, we can exploit within-country variation along the dimensions in which Israel differs to provide some perspective on how one might expect results to differ in other contexts. In particular, there are at least three features of the Israeli context that could potentially affect our ability to draw inferences about leaders in other countries. The first concerns the widespread prevalence of military service in Israel: unlike in many other countries, all Jewish Israeli citizens are required by law to serve in the IDF, many of whom experience combat firsthand. As noted in the main paper, although compulsory service in the IDF means we do not have sufficient variation in our sample of decision-makers to estimate heterogeneous treatment effects with respect to military service, there is considerable variation in combat experience. In the main paper, we exploit this variation to estimate whether there are heterogeneous treatment effects with respect to combat experience, which also serves as a proxy for whether our results would hold in countries where leaders might have less military experience. Yet as the main results show, participants' combat experience does not appear to affect how they update based on different types of costly signals.

The second concerns the relatively hardline nature of Israeli politics on security issues, which may make our participants consistently more hawkish than decision-makers in other countries. Here, we do find significant evidence of treatment heterogeneity in our sample, with hawks more sensitive to sunk cost signals of military mobilization, but less persuaded by verbal threats than doves. However, if the expectation is that Israeli leaders are unusually high in military assertiveness compared to leaders in other countries, this should make it *harder* to find an effect in Israel than in countries where a wider range of attitudes towards military force are represented.

The third concerns the high salience of foreign policy and security issues in Israeli politics, which may make our decision-makers more experienced and knowledgeable about assessing the credibility of threats and decisions about the use of force. As noted in the main text, this actually increases the validity of our participants' responses, since we are asking them to engage in the very behavior in which they are more likely to have expertise. Nonetheless, in the main paper we exploit variation

in foreign affairs experience among our participants to estimate heterogeneous treatment effects. Interestingly, at least when we measure foreign affairs experience formally (that is, in terms of serving on the Foreign Affairs and Defense Committee) we find no evidence that more experienced participants interpret signals differently than less experienced ones, although it is possible that in countries where foreign policy is less salient, formal foreign affairs experience might matter more. Thus, although future scholarship should extend our findings by carrying out similar studies in other countries, when we focus on the three dimensions above, we fail to find evidence of heterogeneous treatment effects.

3.3 Leader characteristics and robustness checks

Tables 2-5 present a series of supplementary analyses for the leader-level analyses from the main text. Due to the small sample size, and intercorrelation between a number of our measures of experience, models 1-6 of Table 2 present a series of regression models estimating the effect of each leader experience variable separately on the perceived efficacy of public threats; the complete model is presented in model 7. Table 3 does the same for the effects of leader experiences on the perceived efficacy of military mobilization. In both cases, the leader experience variables appear to be relatively unimportant in explaining variation in how leaders calculate credibility, while the effect of leader orientations is substantively stronger, particularly military assertiveness and international trust in Table 2. Indeed, a pair of Wald tests fails to find evidence of a significant reduction in model fit if the experiential variables are dropped altogether, apart from age in the military mobilization treatment, which implicates a variety of non-experiential mechanisms as well.[3]

We also carry out several additional robustness checks. First, to show that the null effects of leader experiences aren't simply due to post-treatment bias induced by estimating the effects of experiences and orientations simultaneously, Tables 4-5 re-estimate the models from Tables 2-3, this time omitting the leader orientation variables. Importantly, the leader-level variables continue to lack statistical significance; a Wald test comparing model 7 from Table 3 with model 7 from Table 5 confirms dropping orientations from the model produces a significant reduction in fit ($F = 4.917$, $p < 0.007$), while a Wald test comparing model 7 from Table 2 with model 7 from Table 4 fails to find evidence of a significant reduction in fit ($F = 2.026$, $p < 0.13$), although the sample size is small in both cases.

Second, we seek to adjudicate between two different explanations for the divergent results we see between hawks and doves in the results. The account offered in the main text holds that hawks and doves differ from one another in how they interpret the signals. An alternative account would be that hawkish and dovish leaders interpret signals similarly, but have differing baseline expectations of how likely other actors are to stand firm in crises - those who update more may simply do so because they have more diffuse priors. We find this explanation less plausible, since it cannot explain why each type of costly signal has different effects, but the within-subject component of the experimental design lets us test this account directly. We calculate the strength of each participant's priors from by calculating the distance of their baseline resolve estimate from 50%. Importantly, the strength

[3]For public threats, a Wald test comparing the unrestricted model from model 7 of Table 2 with a restricted model without the experiential variables (combat experience, terms in office, foreign affairs experience, elite experience, and age) fails to find evidence of a significant reduction in model fit ($F = 0.85$, $p < 0.52$); an equivalent test for mobilization using model 7 of Table 3 suggests a significant reduction in model fit ($F = 2.29$, $p < 0.07$), but this is due to the significant effect of age, rather than any of the other experiential variables, and if we replicate the analysis while also retaining age, the reduction in fit is no longer significant ($F = 1.03$, $p < 0.41$).

Table 2: Effect of leader characteristics on perceived efficacy of public threats

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Experience							
Combat experience	−6.663						−4.966
	(4.586)						(5.246)
Military experience, no combat		6.519					
		(4.679)					
Terms			0.465				−0.246
			(1.140)				(1.557)
Foreign affairs experience				−4.255			−5.676
				(5.700)			(6.289)
Elite experience					6.992		10.375
					(6.855)		(9.469)
Age						−0.221	−0.135
						(0.217)	(0.259)
Orientations							
Military assertiveness	−36.061*	−37.782**	−38.009*	−34.410*	−38.687**	−38.433**	−38.031*
	(18.456)	(18.514)	(19.253)	(18.780)	(18.673)	(18.644)	(19.944)
International trust	6.777	6.237	5.667	9.587	5.250	9.860	8.045
	(9.071)	(9.109)	(10.376)	(9.426)	(9.328)	(9.268)	(10.746)
Ideology	7.522	7.185	6.975	10.462	4.548	10.559	10.450
	(12.722)	(12.748)	(12.921)	(13.642)	(13.004)	(13.230)	(14.487)
Male	6.232	6.439	3.819	5.561	3.929	6.930	8.546
	(6.894)	(6.944)	(7.045)	(7.034)	(6.834)	(7.257)	(7.602)
Constant	23.588*	18.437	21.953*	19.282	23.023*	30.976*	29.770
	(12.559)	(12.662)	(12.899)	(12.738)	(12.655)	(15.952)	(17.899)
N	42	42	43	43	43	43	42
R^2	0.212	0.208	0.169	0.178	0.188	0.188	0.263

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

Table 3: Effect of leader characteristics on perceived efficacy of military mobilization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Experience							
Combat experience	−1.972						−1.136
	(4.696)						(4.501)
Military experience, no combat		5.973					
		(5.167)					
Terms in office			−1.910*				−2.730*
			(0.987)				(1.425)
Foreign affairs experience				−0.408			4.463
				(4.369)			(4.531)
Elite experience					−4.341		5.399
					(6.038)		(7.548)
Age						−0.557***	−0.372
						(0.201)	(0.227)
Orientations							
Military assertiveness	40.243**	40.280**	41.022**	38.120**	36.766**	52.176***	52.519***
	(16.097)	(15.598)	(15.144)	(16.065)	(15.762)	(15.271)	(16.129)
International trust	22.135**	22.282**	21.198**	19.189**	18.983**	21.511***	25.635***
	(9.013)	(8.861)	(8.286)	(8.888)	(8.553)	(7.887)	(8.786)
Ideology	−16.479	−18.628	−16.260	−16.305	−14.001	−25.823**	−25.442**
	(12.303)	(12.027)	(11.225)	(11.793)	(12.149)	(11.245)	(12.288)
Male	−2.889	−0.572	−3.466	−3.055	−2.643	−2.940	−4.083
	(5.848)	(6.094)	(5.305)	(5.563)	(5.555)	(5.061)	(5.510)
Constant	−15.759*	−19.725**	−11.289	−14.209	−13.796	14.252	6.274
	(9.124)	(9.403)	(8.609)	(8.897)	(8.852)	(13.055)	(14.253)
N	41	41	43	43	43	43	41
R^2	0.218	0.243	0.263	0.188	0.199	0.328	0.407

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

Table 4: Robustness check: no evidence of post-treatment bias for threats

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Combat experience	−5.963						−6.233
	(4.687)						(5.348)
Military experience, no combat		4.931					
		(4.785)					
Terms in office			0.706				0.078
			(0.959)				(1.454)
Foreign affairs experience				−3.707			−8.236
				(4.851)			(5.928)
Elite experience					5.649		8.972
					(6.160)		(9.227)
Age						−0.049	0.012
						(0.199)	(0.258)
Male							7.982
							(7.844)
Constant	12.000***	6.536**	5.891	10.769**	6.439**	11.107	7.237
	(3.714)	(2.826)	(3.759)	(4.091)	(2.867)	(12.279)	(13.253)
N	43	43	45	45	45	45	43
R^2	0.038	0.025	0.012	0.013	0.019	0.001	0.124

*p < .1; **p < .05; ***p < .01

Table 5: Robustness check: no evidence of post-treatment bias for mobilization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Combat experience	0.606						2.024
	(4.602)						(5.072)
Military experience, no combat		3.191					
		(4.900)					
Terms in office			−1.376				−1.724
			(1.042)				(1.603)
Foreign affairs experience				2.366			7.646
				(4.261)			(5.118)
Elite experience					−4.784		−1.408
					(5.913)		(8.260)
Age						−0.305	−0.242
						(0.200)	(0.254)
Male							−0.192
							(6.012)
Constant	6.357*	5.900**	10.697***	5.313	7.960***	25.714**	20.976
	(3.734)	(2.538)	(3.563)	(3.399)	(2.482)	(12.535)	(14.944)
N	41	41	44	44	44	44	41
R^2	0.0004	0.011	0.040	0.007	0.015	0.053	0.125

*p < .1; **p < .05; ***p < .01

of participants' priors and their levels of hawkishness are independent of one another ($r = -0.025$); if we define the threshold between hawks and doves by mean-splitting hawkishness, hawks have an average baseline level of certainty of 17.38, and doves of 17.48. Alternately, rather than focus on the strength or diffusion of participants' priors more generally, we can simply look at the average baseline estimate offered by hawks and doves (e.g. perhaps hawks assume foes are more resolute than doves do?). Here too, though the correlation between participants' baseline assessments and their level of hawkishness is relatively weak ($r = 0.07$); mean-splitting hawkishness, it appears that hawks offer slightly higher estimates than doves (56.4 vs 51.9), but the difference is not significant: in a bivariate regression model of baseline estimates on hawkishness, the associated p-value for the hawkishness coefficient is $p < 0.51$. These tests thus offer further evidence that the results for hawkishness are not driven by differing priors, but rather, differing interpretations of the signals themselves.

Finally, given the effects of analogical reasoning (Khong, 1992) and the outsized role that a series of formative historical events — e.g. the six-day war in 1967, the Yom Kippur war in 1973 – play in Israeli political culture, we tested for generational effects (Kertzer, 1983). Were individuals old enough to remember the 1973 war, for example — a conflict in which Israel was surprised by military mobilization — more likely to take military mobilization seriously? Our results suggest the opposite, in that younger MKs update more in response to military mobilization than older ones. Given, however, the relatively small sample size (relatively few respondents in our sample were born after 1973, and because of longitudinal changes in the composition of the Knesset over time, younger MKs differ from older ones in a number of ways), we lack the data to test this hypothesis conclusively.

4 Israeli public sample

Table 6: Israeli Public Sample

	Public sample
Male	53%
Education:	
No High School degree	2%
High School degree	33%
Some college	23%
College degree	27%
Masters degree	14%
Doctoral degree	2%
Military Experience:	
Did not serve	22%
Served, no active combat	50%
Combat experience	28%
Location:	
Jerusalem	10%
Tel Aviv	12%
Central Zone	18%
Haifa	15%
Northern Region	11%
Southern Region	12%
Lowland	10%
Sharon Area	8%
Yehuda and Shomron	3%
Religiosity:	
Secular	60%
Traditional	19%
Religious	15%
Orthodox	5%
Birth Country:	
Israel	81%
Former USSR	10%
Other	9%
	Mean (SD)
Age	41.8 (14.6)
Military Assertiveness	0.58 (0.19)
Right Wing Ideology	0.61 (0.25)
Hawkishness (Arab-Israeli conflict)	0.63 (0.25)
International Trust	0.32 (0.27)

4.1 Elite-public differences in political orientations

In addition to standard demographic data, subjects completed questionnaires relating to military assertiveness (described in the main text), political ideology, hawkishness (with respect to the Arab-

Israeli conflict) and international trust. Our political ideology item asks subjects to classify themselves along a single dimensions from "left" to "right" in politics, while the "Hawkishness" measure asks subjects to do the same specifically with respect to the Israeli-Arab conflict. As before, our international trust measure is adapted from the work of Brewer et al. (2004), which finds that generalized trust of other countries in the international system helps to structure beliefs about the foreign policy arena.

Table 7 compares our samples along several dimensions, all scaled from 0 (min response) to 1 (max response). First, we note that — within the leader sample — current and former Knesset members do not differ very much on measured covariates. In fact, the only statistically significant difference between the two subsamples of leaders is that current leaders are younger (by about 12 years) compared to former leaders. The differences along the ideational dimensions are all rather small, and none of them are statistically significant.

Perhaps unsurprisingly, our sample of leaders is significantly older (by about 20 years) than respondents from the Israeli public. Some of this is driven by the former Knesset members, who outnumber the younger, current members of Parliament in our sample. In fact, current Knesset members in our sample average 52 years (64 years for former members), far closer to the average age of the Israeli public sample (42).

Our Israeli leaders are notably less conservative, less hawkish with respect to the Arab-Israeli conflict and more trusting of the international community than the general public. With respect to left-right political ideology, Knesset members averaged a score of 0.45, a score that places them nearly an entire standard deviation *less conservative* than the mean score in the Israeli public sample. With respect to the Arab-Israeli conflict, our leaders were an entire standard deviation below the mean for the public sample.

	Knesset			Public	Elite–Public Gap
	Current	Former	Overall		
Age	52.2	64.4	61.4	41.8	**19.6**
Military Assertiveness	0.61	0.61	0.61	0.58	0.03
Right Wing Ideology	0.47	0.44	0.45	0.61	**–0.16**
Hawkishness (Arab/Israeli)	0.44	0.37	0.39	0.63	**–0.24**
International Trust	0.37	0.41	0.40	0.32	**0.08**

Table 7: Sample Comparison: Statistically significant differences in means between public sample and leader (overall) sample depicted in **bold**. p-values calculated via Wilcoxon rank-sum tests.

References

Barabas, Jason and Jennifer Jerit. 2010. "Are Survey Experiments Externally Valid?" *American Political Science Review* 104(2):226–242.

Blattman, Christopher. 2009. "From violence to voting: War and political participation in Uganda." *American Political Science Review* 103(02):231–247.

Brewer, Paul R, Kimberly Gross, Sean Aday and Lars Willnat. 2004. "International trust and public opinion about world affairs." *American Journal of Political Science* 48(1):93–109.

George, Alexander L. and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.

Imbens, Guido W. 2010. "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)." *Journal of Economic Literature* 48(2):399–423.

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):163–213.

Kertzer, David I. 1983. "Generation as a Sociological Problem." *Annual Review of Sociology* 9:125–149.

Khong, Yuen Foong. 1992. *Analogies at War*. Princeton, NJ: Princeton University Press.

McDermott, Rose. 2011. Internal and External Validity. In *Handbook of Experimental Political Science*, ed. James N. Druckman, Donald P. Green, James H. Kuklinski and Arthur Lupia. Cambridge University Press pp. 27–40.

Sechser, Todd. 2011. "Militarized Compellent Threats, 1918-2001." *Conflict Management and Peace Science* 28(4):377–401.