

**“The First Image Reversed:
Can IGO Signals Influence Mass Political Attitudes?”**

by

David H. Bearce
david.bearce@colorado.edu

and

Thomas R. Cook
thomas.cook@colorado.edu

Department of Political Science
University of Colorado at Boulder

2nd draft (August 2015)

Acknowledgements: Earlier versions of this research were presented in the Government Department at the University of Texas at Austin in March 2015, at the MPSA Annual Conference in April 2015, and in the Political Science Department at the University of Colorado at Boulder in April 2015. We thank Bethany Albertson, Craig Louis Arceneaux, Terry Chapman, Stephen Chaudoin, Ryan Dawkins, Mike Findley, Jeff Harden, Moonhawk Kim, Amy Liu, Pat McDonald, Rachel Wellhausen, Jenny Wolak, and Scott Wolford for helpful comments and suggestions.

The field known as political *behavior* studies the attitudes (e.g. policy opinions) and activities (e.g. voting) of individuals. This focus at the individual level of analysis constitutes a large portion of the research in American Politics and a growing share of the research in other domestic contexts (i.e. Comparative Politics). But there is also an understudied international dimension related to political behavior. Certainly, scholars have long considered how individual attitudes and activities influence the formulation of foreign policy. But relatively little effort has been made to understand if international factors also influence individual-level attitudes and behavior (thus reversing the causal arrow).

This oversight is important because many important arguments in International Relations (IR) have components that include the latter. For example, within broader theories about the role of international institutions, scholars have argued that such institutions send signals to the mass public, legitimizing certain foreign policy options, helping to assure voters about the competence of the elected government, and identifying treaty violations (e.g. Mansfield, Milner and Rosendorff 2002; Dai 2007; Fang 2008; Simmons and Danner 2010; Chapman 2011). While such arguments have been advanced in all subfields of International Relations, they have scarcely been tested. This empirical neglect is problematic because while it is perhaps not hard to assume that signals can be sent through/by intergovernmental organizations (IGOs), it is much harder to assume that the mass public successfully receives these IGO signals. In this regard, we are not questioning whether the signals can be *sent*; instead, we question whether IGO signals are *received*.

This paper thus directly considers the behavioral micro-foundations of these IR arguments, asking if IGO signals can influence mass public attitudes. We first discuss why this assumption may be problematic: many individuals among the mass public lack the motivation and knowledge to receive and process international signals. And among those who are motivated and knowledgeable, pre-existing negative attitudes about IGOs may lead them to process the signal differently than

intended. This understanding leads to our general expectation that IGO signals should have a greater effect on more motivated and knowledgeable individuals and on those with more positive pre-existing attitudes about the relevant IGOs. We also expect that IGO signals should generally be weaker on economic issues than on security issues because of the difficulty that many non-elite individuals face in understanding complex economic issues and their comparative lack of familiarity with economic IGOs (compared to security IGOs).

Using an experimental design in which respondents are presented with two scenarios mirroring contemporary IR literature, an economic scenario consistent with Mansfield, Milner and Rosendorff (2002) and a security scenario that parallels Fang (2008) and Chapman (2011), we present evidence of how individuals within the mass public receive an IGO signal. In the economic scenario, we find a weak average receptiveness to the IGO signal (associated with participation in free trade agreements) that is stronger among more motivated individuals and those with positive pre-existing attitudes about the North American Free Trade Agreement (NAFTA). In the security scenario, we find a stronger average receptiveness to the IGO signal (associated with United Nations Security Council authorization) that is present even among less motivated and knowledgeable individuals and those with pre-existing negative attitudes about the United Nations.

I. The Argument

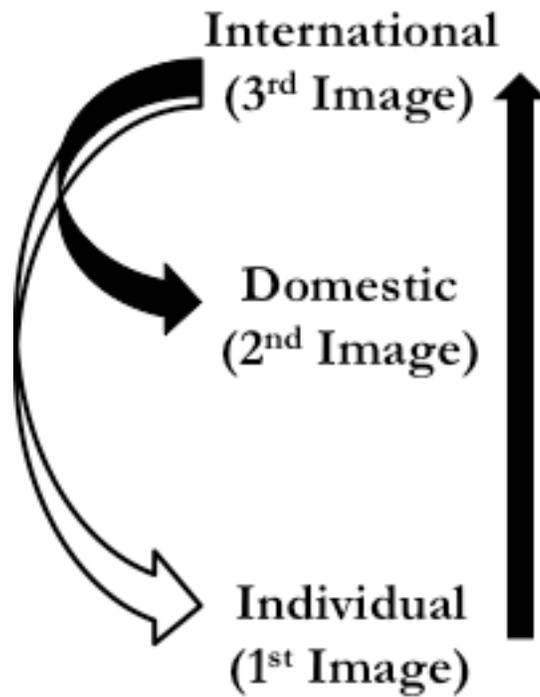
In this section, we first discuss prominent first image reversed arguments within the IR literature, noting the lack of empirical evidence to demonstrate their behavioral micro-foundations. We then consider why the mass public may have difficulty receiving IGO signals as proposed by these arguments. This understanding leads to our expectations that IGO signals will be stronger for more motivated and knowledgeable individuals and for citizens with more positive attitudes about

the relevant IGOs. We also expect that IGO signals should generally be stronger in a high-political issue-area like conflict/security than in a lower-political economic issue-area.

The First Image Reversed

Here we focus on a specific subset (described below) of the IR behavioral arguments offering a “first image reversed” logic. Waltz (1959) described the IR levels of analysis problem using the concepts of images. The first image was human behavior, or the individual level of analysis; the second was the internal structure of states, or the domestic level; and the third image was anarchy at the international level of analysis. But Waltz also offered a directional logic, as illustrated by the straight dark arrow on the right-hand side of Figure 1, working up from the first image towards his emphasis on anarchy at the international level (ibid, chapter VIII).

Figure 1: The First Image Reversed



Gourevitch (1978) later reversed this directional logic by discussing how the international system also affects the internal structure of states. His “second image reversed” logic is illustrated by the curved dark arrow on the left-hand side of Figure 1. In this paper, we consider a set of arguments that take this directional logic one step further, from the international to the individual level, suggesting a first image reversed consistent with the curved white arrow on the left-hand side of Figure 1. More specifically, we focus on arguments that concern individual attitudes within the *mass* public. There is another set of arguments about how international factors influence *elite* attitudes (e.g. Davis 2012), but these will not be part of our analysis in this paper.¹

A number of prominent arguments in IR include this first image reversed, proposing that international factors/outcomes influence attitudes within the mass public. For example, the “domestic audience cost” proposition, starting with Fearon (1994), argues that citizens can observe international defections and will punish governments who make international commitments but then fail to honor them. Audience cost arguments have subsequently been offered to explain international behavior on a variety of topics beyond just military crisis bargaining (the topic where Fearon introduced his audience cost proposition). Tests of these many audience cost arguments were usually conducted using observational macro-level data, but only recently have scholars begun to explore the behavioral micro-foundations of the audience cost proposition with somewhat mixed results (e.g. Tomz 2007; Levendusky and Horowitz 2012; Chaudoin 2014; Kertzer and Brutger 2015).

With a substantive focus on environmental agreements, Dai’s (2007) argument that the victims of non-compliance can monitor and help enforce international agreements also follows the

¹ Perhaps not surprisingly, the argument offered below suggests that it would be harder for less informed members of the mass public to receive international signals than the more informed members of the elite public. Thus, to the extent that we can empirically demonstrate that the mass public, under certain conditions, do successfully receive these international signals, we would expect that elite reception would be even less problematic. In this regard, mass public reception is the harder case, while elite reception represents the easier case.

first image reversed. The behavioral logic within this argument was assumed to be true: “*to the extent that they [citizens] can detect noncompliance*, potential victims of noncompliance provide a partial solution to the information problem in [international] treaty regimes” (ibid, 37 *emphasis added*). But the micro-foundations were not directly tested: can citizens actually detect [environmental] non-compliance, and do they adjust their support for compliant versus non-compliant governments?

Likewise, Simmons and Danner (2010) argue that a government will join the International Criminal Court (ICC) to send a signal to its citizens “that there are boundaries in quelling future threats beyond which the government will not go” (ibid, 234). This argument is tested using observational macro-level data, but it has not yet been shown that citizens interpret ICC non(ratification) in the manner consistent with the argument that was advanced to explain it. Indeed, there may be other reasons why democracies without recent civil wars and non-democracies with recent civil wars are more likely to ratify the ICC, and why democracies with recent civil wars are less likely to ratify it (the macro-level evidence in support of their argument).²

Mansfield, Milner and Rosendorff (2002) propose that governments will sign international free trade agreements (or preferential trading arrangements [PTAs]) to solve an “informational problem that limits the head of state’s credibility. In the eyes of the voters, participating in a PTA bolsters the credibility of offers made by the executive to reduce trade barriers [and therefore lower domestic prices], and enhances the leader’s prospects of re-election” (ibid, 488). Assuming “that voters can observe this signal” (ibid, 489), Mansfield, Milner and Rosendorff test their argument using observational macro-level data: more democratic regimes are more likely to join PTAs. But can citizens actually observe and correctly interpret this signal from international trade agreements, especially when they do not know much macroeconomics and hold trade policy as a low salience political issue (Guisinger 2009)? As noted for the argument above, there may be reasons other than

² It should also be noted that Chapman and Chaudoin (2013) dispute this macro-level evidence, casting further doubt on its assumed micro-foundations.

domestic signaling to explain why more democratic regimes are more likely to join free trade arrangements (e.g. a basic ideological compatibility between political rights and economic openness).

Finally, Fang (2008) and Chapman (2011) argue that governments will seek authorization from a multilateral institution like the United Nations Security Council before engaging in the use of military force in order to bolster public support for this potentially costly foreign policy action. But even in this high political issue area, do voters successfully receive this international signal, especially when they are not well informed about foreign policy issues and are unfamiliar with the operation of multilateral security institutions? The behavioral micro-foundations of this argument were not tested at all by Fang (2008), although Chapman (2011, 121-6) does provide some experimental evidence using a small undergraduate sample.

Our point in this subsection is not to single out these specific scholars for not testing their behavioral micro-foundations. Instead, we seek to demonstrate that first image reversed arguments appear in widely-cited papers/books spread across all IR subfields. Yet there is also precious little evidence that the mass public actually receive the IGO signals implied by these arguments. As Pelc (2013, 629) recently observed on this point: “Such theories rest on a common set of highly specific assumptions about individual behavior. Yet strikingly, these often elaborate premises have been left largely unexamined.” We now turn to a consideration of these behavioral premises.

Mass Public Motivation and Knowledge

Society can be roughly divided into two groups. The elite public, which represents a minority of the population, is relatively interested in and knowledgeable about public policy, including foreign policy. Conversely, the mass public, or the majority of the population, has comparatively little interest in or knowledge about public policy, especially foreign policy. As Holsti (2004, 48) summarized about American society despite its access to information: “Surveys not only

repeatedly reveal that the public has a very thin veneer of factual knowledge about politics, economics, and geography but also indicate that most Americans are poorly informed about the specifics of conflicts, treaties, negotiations with other nations, characteristics of weapons systems, foreign leaders and the like.”

The key point here is obvious and not particularly original: since most citizens are not interested in international affairs, nor are they particularly motivated to learn more, they may not be looking for the IGO signal implied by the various arguments outlined above. And even when these less-informed citizens do observe the IGO signal, perhaps because government leaders and the popular press/media have highlighted this signal, they may have trouble understanding what it implies or means for policy (Converse 1964, Miller and Stokes 1963).

As Lupia and McCubbins (1998, chapter 2) wrote about the components of political knowledge, citizens need to know about causal relationships and the consequences of policy decisions. In the case of Mansfield, Milner and Rosendorff's (2002) argument, citizens not only need to know that their government has (or has not) joined a free trade agreement, they also need to understand what this means for domestic prices (e.g. free trade leads to reduced tariffs and lower domestic prices). Likewise, for Fang's (2008) and Chapman's (2011) argument, citizens need not only to observe that the UN Security Council has (or has not) endorsed the use of military force, they also need to understand what this means for the external popularity of this policy choice (e.g. a set of foreign governments with different interests have all accepted military force as being legitimate in this particular context).

The lack of knowledge and motivation to learn more about public policy bodes ill for the successful reception of IGO signals. But scholarship in political behavior has also demonstrated that poorly informed citizens are nonetheless often able to respond to political signals. “While information levels are indeed low and information about public issues unevenly distributed, citizens

are generally found to employ sensible decision rules, or heuristics, in acquiring and storing information about politics. These decision rules do, in fact, entail remaining ignorant on most issues most of the time, making inferences about the likely behavior of officeholders from very little information, and basing one's vote or other political actions on a relatively sparse set of signals about governmental activity" (Ferejohn 1990, 5).

As Popkin and Dimock (2000, 228) similarly concluded about foreign policy issues: "Ordinary Americans do not know many of the details of foreign policy and typically give limited attention to political debate on foreign issues, but their natural gut reasoning is not devoid of substance. We have shown the ordinary Americans can call on many predispositions and past experiences." Likewise, Hurwitz and Peffley (1987, 1114 *emphasis in original*) demonstrated an unexpected structure and organization to foreign policy attitudes among Americans, arguing that "a paucity of information does not *impede* structure and consistency; on the contrary, [instead] it *motivates* the development and employment of structure."

But this hidden structure in foreign policy attitudes, coupled with a lack of information, presents yet another problem for the successful reception of IGO signals, at least as intended by the policymakers who send them. Specifically, it suggests that individual preconceptions could change their interpretation of an IGO signal. For example, many citizens may have structured attitudes about international trade that run contrary to elite beliefs about trade lowering domestic prices. If citizens instead believe that international trade leads primarily to lost jobs in the national economy, then the IGO signal associated with membership in a free trade agreement may be viewed negatively, not positively as implied by Mansfield, Milner and Rosendorff (2002). Likewise, if citizens have structured, but poorly informed, attitudes about the United Nations as an incipient form of world government, usurping the sovereignty of member-states, then signals from this IGO

may be received with suspicion, rather than with appreciation as argued by Fang (2008) and Chapman (2011).

Theoretical Expectations

Based on the discussion above, we expect that the ability of individuals within the mass public to receive successfully an IGO signal should be conditioned by at least three factors: 1) their motivation to pay attention to the issue, 2) their knowledge about the IGO sending the signal, and 3) their pre-existing attitudes about the IGO in question. Stated as a general hypothesis, we expect that *the reception of the intended IGO signal should be stronger for more motivated and knowledgeable individuals and for those with more positive pre-existing attitudes about the relevant IGO.*

We deliberately phrase this as a general, rather than as an IGO-specific, hypothesis, because the relevant motivation, knowledge, and prior attitudes depend on the scenario and on the issue-area. Stated differently, international affairs “is a highly diffuse subject-matter domain” making it “difficult for individuals to pay attention to all matters pertaining to foreign affairs” (Iyengar 1990, 179). This means that one’s knowledge and beliefs about IGOs, as well as the motivation to learn about causal relationships, may differ at least on an individual basis across issue-areas. Consequently, the operational measures for individual motivation, knowledge, and prior attitudes must be specific to the experimental scenarios, which will be presented below.

We also hypothesize that *the reception of the intended IGO signal should be stronger in higher political issue-areas.* High politics refer to those issues related to national survival and the security of the state, including war, while low politics traditionally includes most social and economic issues. Our reasoning here is related to the concepts in our first hypothesis, namely motivation and knowledge. For issues that citizens perceive (perhaps incorrectly) as more directly and immediately related to national security (i.e. high politics), they arguably have a greater incentive to look for and to process the relevant IGO signal. Indeed, there may also be more media coverage of that signal and more

readily identifiable past experiences to draw upon for the use of heuristics. As Delli Carpini and Keeter (1996) showed, American citizens have greater political knowledge of higher salience foreign policy issues, especially those related to war and peace. But for lower salience issue-areas, including international trade, citizens may, *on average*, be less knowledgeable and less motivated to understand the arguably more complex economic relationships, leading to our cross-issue expectation that the reception of IGO signals should generally be weaker in lower political issue-areas.

II. The Evidence

Research Design

We explore these theoretical expectations using an experimental research design. Given that the various first image reversed arguments discussed in the previous section differ in their application (i.e. different issue-areas, different IGOs, etc.), they cannot all be tested in a single paper. Instead, we select two prominent scenarios from this literature, one where we expect the IGO signal to be relatively weak and another where there are stronger expectations. The first (and corresponding to the former) is an economic scenario drawn from Mansfield, Milner and Rosendorff (2002). The second (and corresponding to the latter) is a conflict/security scenario that fits with those considered by Fang (2008) and Chapman (2011). Our two specific scenarios will be described in more detail below.

The experimental treatment for each scenario is the randomized exposure to an IGO signal. In the economic scenario, the IGO signal concerns the government's (non)participation in free trade agreements (FTAs). In the conflict/security scenario, the IGO signal concerns (non)authorization from the United Nations Security Council (UNSC). Since both scenarios are presented to all respondents, we also randomize the order of the two scenarios; thus, about half of our sample received the economic scenario first, while the other half were first presented with the

conflict/security scenario. *Before* the randomized scenarios, we ask a series of questions about issue motivation, IGO knowledge, and pre-existing attitudes to help assure that our measures for these moderating, or conditioning, variables are exogenous to the questions used for our attitudinal dependent variable, which come *after* the randomized scenarios (followed by the standard demographic queries).

Our experiment can thus be divided into three sequential blocs. The first provides the information for the hypothesized moderating variables. The second presents our scenarios, randomized by order and across treatments. The third obtains demographic information to assure balance across the randomized treatments.

To read these scenarios and answer the related questions, we recruited more than 1200 voting-age Americans using Mechanical Turk (MTurk) in July 2015. While we recognize that our Mechanical Turk sample is not nationally representative, Berinsky, Huber, and Lenz (2012, 352) demonstrate that “the demographic characteristics of domestic MTurk users are more representative and diverse than the corresponding student and convenience samples typically used in experimental political science studies.” In addition, they show “that the estimates of average treatment effects are similar in the MTurk and original [other] samples.” So we judge our experimental sample to be a valid one for testing the effect of IGO signals on those within the *mass* public (there will be more discussion on this point later). It is worth noting that a convenience sample (from MTurk or from an undergraduate pool) would not necessarily be valid for testing the effect of IGO signals on members of the *elite* public, but that smaller societal group is not our focus in this paper.

Since our two scenarios are not directly parallel (nor could they be made completely so),³ we present them separately: first the economic scenario and then the conflict/security scenario. After

³ To preview, our scenarios are constructed to be parallel in terms of three dimensions: 1) their length as measured by word count, 2) the structure of their IGO treatments, and 3) in the operationalization of the attitudinal dependent variable.

these sequential presentations, we will make some careful comparisons between the two issue-areas. Within each scenario, we also present the results for two samples: first using the full sample and then restricted to the sub-sample passing the treatment check. We make this dual presentation in each scenario to address the concern by Barabas and Jerit (2010) that survey experiments often exaggerate real-world treatment effects because, in the real-world, many citizens simply do not receive the treatment equivalent. We thus begin with the full sample, *including those who failed the manipulation check*, to guard against exaggerated results. We then present the same for only those who passed the treatment check (about 75% in each scenario). We view these latter results as an upper limit to our treatment effects with all respondents having received their intended treatment.

The Economic Scenario

Our economic scenario is drawn from the argument advanced by Mansfield, Milner and Rosendorff (2002) that was discussed earlier. While it concerns the IGO signal sent by the government's (non)participation in free trade agreements (FTAs), this economic scenario is ultimately about inflation (or domestic price stability) with trade policy as an intervening variable. Described simply, a rise in prices, or inflation, could be due to some exogenous shock that is beyond the government's control or it could be due to trade protection, which increases the price for imported goods. Given a general preference within society for lower inflation, democratic governments have an incentive to sign/join free trade agreements in order to demonstrate to potential voters that a rise in prices, when this event occurs, is not due to bad policy, which the government can control.

We judge this economic scenario to be a relatively complicated one, requiring potential voters to understand the relationship between trade policy and domestic prices and to be able to correctly observe the government's (non)participation in FTAs. Of course, a democratic

government has the incentive to explain this relationship and to highlight their FTA participation to potential voters. But even with this incentive, do potential voters receive and understand this international signal? To explore this possibility, we asked the respondents in our MTurk sample to read the following short vignette, which included one of three randomized treatments.

Inflation can be defined as a rise in the price that must be paid for a set of goods and services. Many different factors can affect the inflation rate. Since many goods are imported from other countries, the price of imports goods can increase due to government-imposed tariffs, defined as a tax imposed on foreign producers who want to sell their goods in this country. This tax tends to raise the price of imported goods, which can contribute to higher inflation. In addition, inflation can increase due to unexpected economic events. For example, a rise in world energy prices can also cause inflation in this country. Unfortunately, it may be difficult for the government to prevent these international events from increasing domestic prices. Recently, there has been a substantial rise in inflation. In response, the government presented a revised economic policy, which includes a tighter monetary policy (146 words)

1. *[and a commitment not to sign any new free trade agreements to reduce the tariffs on imported goods.]*
2. *[and a commitment to sign new free trade agreements to reduce the tariffs on imported goods.]*
3. *[.]*

The first treatment identifies the situation where the government intends not to sign any new FTAs (*No FTA*), while the second treatment indicates the opposite intention (*FTA*). The third treatment serves as a baseline, giving the respondent no information about the government's FTA intentions. Our randomization of these three treatments was effective, thus achieving balanced subsamples across the treatment groups for a set of basic demographic characteristics as shown in Table 1. For a range of tests, there are no statistically significant differences among these three subsamples, so our regression results will not include these demographic factors as control variables since their inclusion has no effect on the treatment coefficients that will be presented and discussed below.

Table 1: Descriptive Statistics by Treatment Group in Economic Scenario

Treatment:	1. <i>No FTA</i>	2. <i>FTA</i>	3. <i>Baseline</i>
<i>Education</i>	4.16 (1.28)	4.18 (1.27)	4.24 (1.26)

<i>Age</i>	35.00 (12.49)	34.73 (12.78)	35.59 (12.57)
<i>White Race</i>	0.76 (0.43)	0.78 (0.42)	0.79 (0.41)
<i>Conservative Ideology</i>	3.32 (1.67)	3.30 (1.69)	3.50 (1.70)
<i>Partisan Democrat</i>	4.67 (1.72)	4.71 (1.72)	4.49 (1.73)
<i>Working Full Time</i>	0.52 (0.50)	0.50 (0.50)	0.52 (0.50)
<i>Income</i>	6.05 (3.72)	5.68 (3.28)	6.00 (3.65)

Cell entries include the mean value with the standard deviation in parentheses.

Following the randomized treatment, we asked respondents the following question: *Based on the information provided, would you oppose or support the government's revised economic policy?* Their response (*Strongly oppose / Oppose / Slightly Oppose / Uncertain / Slightly support / Support / Strongly support*) serves as our attitudinal dependent variable for the regression results reported in Table 2. Although our dependent variable is ordinal (with 7 categories), we present our results with OLS coefficients. There are no statistically significant differences between the OLS coefficients and those coming from either an ordered logit or probit, but the OLS coefficients are more straightforward to interpret. As mentioned earlier, we first present our results using the full sample and then again using only the sub-sample passing the treatment check. As would be expected, the first set of results is weaker compared to the latter.

Table 2: Estimates of Support for the Government's Economic Policy
Using the Full Sample.

Model:	2.1	2.2	2.3	2.4	2.5
<i>No FTA</i>	-0.11 (0.11)				
<i>FTA</i>	0.13† (0.09)	0.24** (0.09)	0.12 (0.13)	0.08 (0.20)	-0.83*** (0.25)

<i>Baseline</i>		0.11 (0.11)	0.11 (0.11)	0.10 (0.11)	0.12 (0.11)
<i>Inflation Importance</i>			-0.10** (0.05)		
<i>FTA x Inflation Importance</i>			0.09† (0.06)		
<i>NAFTA Purpose Correct</i>				0.06 (1.14)	
<i>FTA x NAFTA Purpose Correct</i>				0.18 (0.20)	
<i>NAFTA Favorability</i>					0.08† (0.05)
<i>FTA x NAFTA Favorability</i>					0.49*** (0.14)
F	4.01**	4.01**	3.73***	2.66**	18.03***
N	1271	1271	1269	1269	1271

OLS estimates with robust standard errors clustered on the state.

* $p < .10$, ** $p < .05$, *** $p < .01$ (two-tailed) and † $p < .10$ (one-tailed).

In model 2.1 in Table 2, we present the coefficients for *No FTA* and *FTA* with the no-information *Baseline* as the omitted category. Both coefficients take on the expected sign: negative for *No FTA* and positive for *FTA*. But when compared to the baseline (non)treatment, the effect of the *No FTA* treatment is substantively small (-0.11) and statistically insignificant. The effect of the *FTA* treatment is also substantively small (0.13) but statistically significant when compared to the baseline (non)treatment using a one-tailed.

Substantively, we are more interested in directly comparing the *FTA* treatment to the *No FTA* treatment. So in model 2.2 (Table 2), we use the latter as our omitted category, thus comparing both the baseline (non)treatment and *FTA* treatment to the situation where the respondent is explicitly told that the government does not intend to sign any free trade agreements. Again, we observe no statistically significant difference between *Baseline* and *No FTA* (as the omitted

category). But we observe a statistically significant effect for the *FTA* treatment when compared to the omitted No FTA treatment (0.24**). Given that our dependent variable is coded across a 6-point range (*Strongly oppose* =1 and *Strongly support* =7), this coefficient could be interpreted to indicate that the *FTA* treatment makes the respondents about 4 percent more supportive of the government's economic policy ($0.04=0.24/6.00$) when compared to the No FTA treatment.

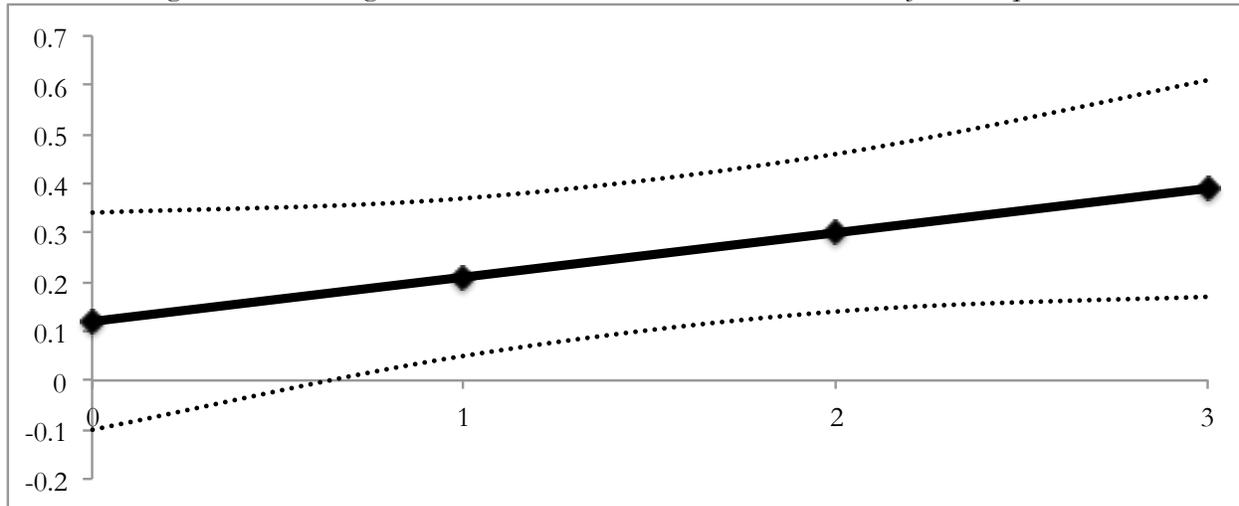
As discussed earlier, we expect that the FTA treatment effect should be stronger for more motivated individuals. We measured motivation in this issue-area by asking all respondents to rank four macroeconomic issues (inflation, unemployment, the budget deficit, and the trade deficit) in terms of their personal importance.⁴ From this information, we created an ordinal variable labeled *Inflation Importance* indicating where inflation stood within this rank (=0 for the least important and =3 for the most important). We thus identify respondents with a higher value as being potentially more motivated to understand the relationship between FTA (non)participation and domestic prices. This issue-specific motivational indicator is then interacted with the FTA treatment ($FTA \times Inflation\ Importance$) and added, along with the *Inflation Importance* constitutive term, to the specification in model 2.3 (Table 2).

With this specification, the FTA treatment effect comes from the linear combination of the *FTA* constitutive term and the $FTA \times Inflation\ Importance$ interaction terms, both of which are positively signed. Figure 2 graphs this *FTA* marginal effect for the different values of *Inflation Importance*. For those who consider inflation to be the least important macroeconomic issue, the FTA treatment has no statistically significant effect. But this effect grows and becomes statistically significant for all other values of *Inflation Importance*. For those who consider inflation to be the most important macroeconomic issue, the marginal effect of the FTA treatment is 0.39***, which is about

⁴ Within our sample, unemployment was ranked as the top macroeconomic issue with the trade deficit at the bottom. Inflation and the budget deficit were ranked in the middle with the budget deficit slightly ahead of inflation.

two-thirds greater than the un-moderated FTA effect observed in model 2.2 and indicating about 7 percent greater support for the government's economic policy.

Figure 2: The marginal effect of *FTA* for different values of *Inflation Importance*.



Y axis is the marginal effect of *FTA*.

X axis is the value for *Inflation Importance*.

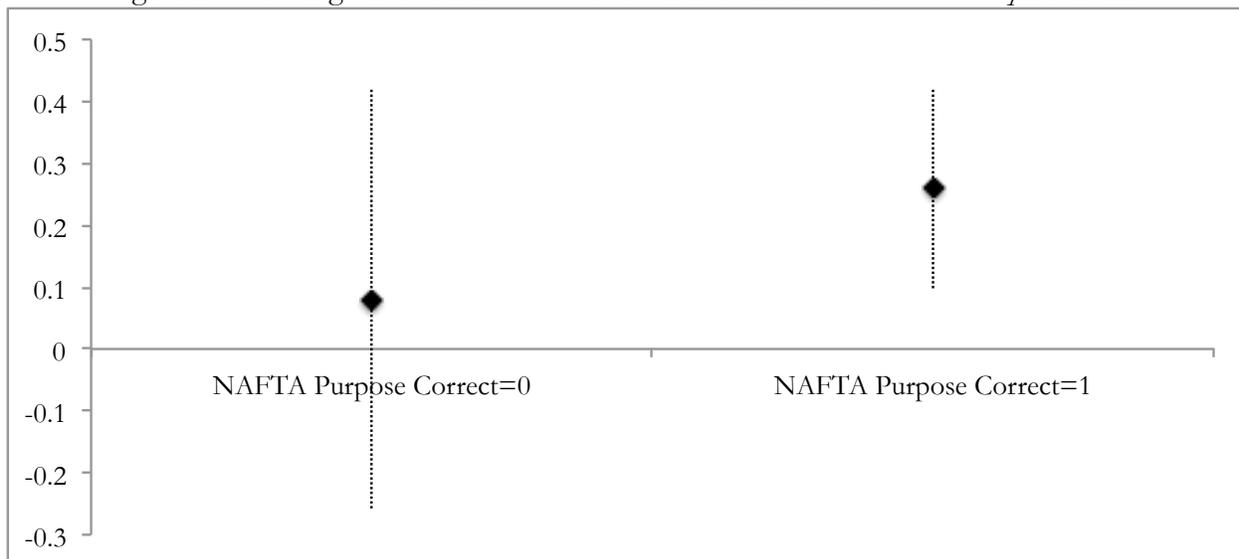
Dashed lines indicate 90 percent confidence intervals.

As also discussed earlier, we expect that this FTA treatment should be stronger for individuals who are more knowledgeable about this type of international economic agreement. Given that our sample consists of American respondents, we focus on their knowledge about the North American Free Trade Agreement (NAFTA), asking them if they can correctly identify the primary purpose of this IGO (*to reduce trade barriers/ to regulate immigration/ to stabilize exchange rates*). We then created a dummy variable labeled *NAFTA Purpose Correct* coded as 1 for all respondents who correctly identified reducing trade barriers as the primary purpose from this set.⁵

⁵ About 85 percent of our respondents are able to correctly identify reducing trade barriers as the primary purpose of NAFTA. This high percentage came as a surprise and may be due to respondents being able to read about NAFTA on-line as they answer our questions. It may also be due to the word “trade” being part of the response and also part of this FTA’s name. In the next

This moderating variable is interacted with *FTA* and added to the specification in model 2.4 (Table 2). The results, also presented in Figure 3, show that for those who are unable to identify correctly the primary purpose of NAFTA, there is no statistically significant FTA treatment effect; the *FTA* marginal effect is only 0.08 with a very large standard error (0.20). However, the marginal effect of *FTA* is larger and statistically significant ($0.08 + 0.18 = 0.26^{***}$) for respondents who are able to correctly identify the primary purpose of NAFTA, suggesting that some basic knowledge of this IGO's purpose may be a necessary pre-condition to achieve a significant FTA treatment effect.

Figure 3: The marginal effect of *FTA* for different values of *NAFTA Purpose Correct*.



Y axis is the marginal effect of *FTA*.

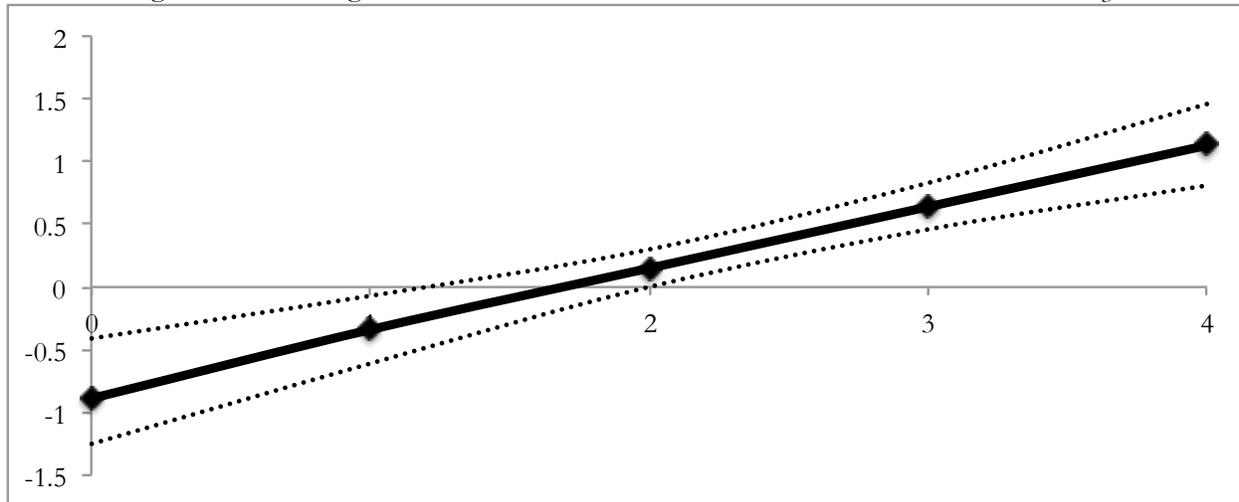
Dashed lines indicate 90 percent confidence intervals.

Finally, we also expect that this positive FTA treatment effect may only be present for citizens who have positive pre-existing attitudes about this type of IGO. To capture the latter concept, we asked (in advance) for their opinion of the North American Free Trade Agreement

version of this experiment, we will find another more satisfying way to measure knowledge about FTAs.

(*very unfavorable/ unfavorable/ no opinion/ favorable/ very favorable*). From their response, we create a five level ordinal variable (=0 if “very unfavorable” and =4 if “very favorable”) labeled *NAFTA Favorability*, which is interacted with *FTA* and added to the specification in model 2.5 (Table 2).⁶

Figure 4: The marginal effect of *FTA* for different values of *NAFTA Favorability*.



Y axis is the marginal effect of *FTA*.

X axis is the value for *NAFTA Favorability*.

Dashed lines indicate 90 percent confidence intervals.

The results illustrated in Figure 4 show a very strong moderating effect for *NAFTA Favorability*. For those who are the most favorably disposed towards this most prominent American FTA, the marginal effect of *FTA* is 1.13***, which is almost five times larger than the same marginal effect observed in model 2.2 (0.24**) and indicating about 19 percent greater support for the government’s economic policy. This marginal effect remains positive and statistically significant (0.15*) for respondents who have no opinion of NAFTA (*NAFTA Favorability* = 2) due to the tight confidence intervals. But it is interesting to observe that for Americans who have a pre-existing *unfavorable* opinion of NAFTA, the FTA treatment has the opposite effect: it *decreases* support for the

⁶ The modal respondent in our sample has “no opinion” concerning NAFTA, but there are more respondents who are favorable than unfavorable.

government's policy to fight inflation. Indeed, for those who are the most unfavorably disposed towards this IGO (*NAFTA Favorability* = 0), the marginal effect of the FTA treatment is -0.83***, indicating about 14 percent *less* support for the government's economic policy.

In focusing on these full sample results from Table 2, it is also important to know how many respondents correctly understood/recognized their randomized treatment. To this end, we asked them - *after* the attitudinal dependent variable question – to identify which of the following was a component of the government's economic policy (*a commitment not to enter new free trade agreements / a commitment to enter new free trade agreements / the scenario did not mention free trade agreements*). 74 percent of the respondents were able to answer this question correctly, which could be read to indicate that our necessarily short scenario/treatments were not too subtle. But since 26 percent failed to pass this treatment check, our economic scenario certainly does not correspond to Kinder's (2007, 157) concern about exaggerated power from an experiment that is "carried out in such a way that virtually everyone receives the message." A full quarter of our respondents did not receive the message, so we now offer a parallel set of results using only the subsample that passed the treatment check to assess how much stronger these effects might be if everyone had indeed received the intended message. These results are presented in Table 3.

Table 3: Estimates of Support for the Government's Economic Policy
Using Subsample Passing Treatment Check.

Model:	3.1	3.2	3.3	3.4	3.5
<i>Baseline</i>		0.08 (0.13)	0.09 (0.12)	0.08 (0.12)	0.08 (0.12)
<i>No FTA</i>	-0.08 (0.13)				
<i>FTA</i>	0.20** (0.10)	0.29*** (0.11)	0.08 (0.16)	0.26 (0.27)	-1.24*** (0.28)
<i>Inflation Importance</i>			-0.13** (0.05)		

<i>FTA x Inflation Importance</i>			0.15* (0.09)		
<i>NAFTA Purpose Correct</i>				0.05 (0.21)	
<i>FTA x NAFTA Purpose Correct</i>				0.03 (0.26)	
<i>NAFTA Favorability</i>					-0.01 (0.05)
<i>FTA x NAFTA Favorability</i>					0.70*** (0.12)
F	4.38**	4.38**	3.22**	2.33*	18.03***
N	932	932	930	931	932

OLS estimates with robust standard errors clustered on the state.

* $p < .10$, ** $p < .05$, *** $p < .01$ (two-tailed) and † $p < .10$ (one-tailed).

Not surprisingly, the un-moderated FTA treatment effects are stronger in this restricted sample, comparing the *FTA* coefficients in model 3.1 in Table 3 to model 2.1 in Table 2 (0.20** versus 0.13†) and the same in model 3.2 to model 2.2 (0.29*** versus 0.24**). But the conditional FTA effect related to the importance of inflation also becomes more pronounced, comparing the *FTA x Inflation Importance* coefficient in model 3.3 to the same in model 2.3 (0.15* versus 0.09†). Likewise, the conditional FTA effect related to NAFTA attitudes also becomes starker, comparing the larger negative *FTA* constitutive term coefficient in model 3.5 to model 2.5 (-1.24*** versus 0.83***) and the larger positive *FTA x NAFTA Favorability* interaction coefficient in model 3.5 to the same in model 2.5 (0.70*** versus 0.49***).

The Conflict Scenario

Our conflict scenario comes from the argument advanced by Fang (2008) and Chapman (2011) that governments will seek authorization from a multilateral institution like the United

Nations Security Council (UNSC) before engaging in the use of military force in order to bolster public support for this potentially costly foreign policy action. As compared to the economic scenario, we view our conflict scenario to be less complicated for those in the mass public to understand, although it does require the individual to have some basic knowledge of the United Nations (UN) system.

To explore whether UNSC authorization makes American individuals more supportive of US military force, we asked our MTurk respondents to read the following vignette, which includes one of three randomized treatments.

A country in a different region of the globe has been invaded. The invaded country is a US ally, while the invading country is not a US ally. Furthermore, the invaded country is rich in energy resources and has been an important energy supplier to many countries. This invasion puts those energy resources at potential risk. The invading country has no known nuclear weapons, and its military assets are only about 1/10 the size of US conventional military assets. The US government is considering a unilateral military response to secure these energy supplies. This planned military response would include both airstrikes against the invading country's military assets and the potential deployment of up to 50,000 ground forces (117 words)

1. [*, and did not receive a supportive resolution from the United Nations Security Council.*]
2. [*, and received a supportive resolution from the United Nations Security Council.*]
3. [*.*]

The first treatment identifies a situation where there is no UNSC authorization (No UNSC), while the second indicates that the UNSC has authorized US military action (UNSC). Parallel to the economic scenario, the third conflict treatment provides no information about UNSC authorization to serve as a baseline (*Baseline*).

As before, our randomization across these three treatments was effective as demonstrated by the demographically balanced sub-samples shown in Table 4. Using a range of different tests, there are no statistically significant differences among the three treatment groups for any of the demographic variables listed in the first column. We thus omit any demographic control variables as they would have no effect on the treatment coefficients of interest here.

Table 4: Descriptive Statistics by Treatment Group in Conflict Scenario

Treatment:	1. <i>No UNSC</i>	2. <i>UNSC</i>	3. <i>Baseline</i>
<i>Education</i>	4.17 (1.26)	4.20 (1.29)	4.20 (1.26)
<i>Age</i>	34.83 (12.85)	34.52 (11.90)	35.99 (13.05)
<i>White Race</i>	0.78 (0.42)	0.80 (0.40)	0.75 (0.43)
<i>Conservative Ideology</i>	3.36 (1.66)	3.44 (1.72)	3.32 (1.69)
<i>Partisan Democrat</i>	4.69 (1.66)	4.55 (1.74)	4.63 (1.77)
<i>Working Full Time</i>	0.48 (0.50)	0.52 (0.50)	0.54 (0.50)
<i>Income</i>	5.95 (3.63)	5.82 (3.45)	5.98 (3.59)

Cell entries include the mean value with the standard deviation in parentheses.

After the randomized treatment, we asked respondents the following question: *Based on the information provided, would you oppose or support the government's decision to use military force abroad?* Their response (*Strongly oppose / Oppose / Slightly Oppose / Uncertain / Slightly support / Support / Strongly support*) serves as our attitudinal dependent variable. Once again, we report our estimates of individual-level support for a US military response using OLS coefficients that allow for a more straightforward interpretation of the results. But there would be no statistically significant differences if we reported ordered logit/probit coefficients.

As before, we first present our results using the full sample of respondents, including those who failed the treatment check (25%), to guard against exaggerated treatment effects. These full sample results for our conflict scenario are shown in Table 5. The first model (5.1) includes the first two treatments (*NO UNSC* and *UNSC*), treating *Baseline* as the omitted category. Both of these treatments take on the expected sign (negative for *No UNSC* and positive for *UNSC*) and are statistically different from the no-information baseline treatment.

But as before, we are more interested in a direct comparison between explicitly receiving versus not receiving UN Security Council authorization. So in model 5.2, we treat *No UNSC* as the omitted category, adding the baseline treatment to the specification. The results show a large positive coefficient for UNSC (0.49 ***) that could be interpreted, given our six level dependent variable, to indicate that the *UNSC* treatment makes the respondents about 8 percent more supportive of the government's economic policy ($0.08=0.49/6.00$) when compared to the *No UNSC* treatment. Both in terms of statistical and substantive significance, this represents a stronger IGO signal than was observed in the economic scenario, consistent with our second hypothesis.

Table 5: Estimates of Support for the Government's Military Policy
Using the Full Sample.

Model:	5.1	5.2	5.3	5.4	5.5
<i>NoUNSC</i>	-0.28** (0.12)				
<i>UNSC</i>	0.21* (0.11)	0.49*** (0.11)	0.58*** (0.14)	0.81*** (0.25)	0.37 (0.30)
<i>Baseline</i>		0.28** (0.12)	0.27** (0.13)	0.27** (0.13)	0.28** (0.12)
<i>Energy Importance</i>			0.09† (0.06)		
<i>UNSC</i> × <i>Energy Importance</i>			-0.09 (0.09)		
<i>UN Purpose Correct</i>				0.12 (0.21)	
<i>UNSC</i> × <i>UN Purpose Correct</i>				-0.36† (0.25)	
<i>UN Favorability</i>					-0.11** (0.05)
<i>UNSC</i> × <i>UN Favorability</i>					0.05 (0.11)
F	9.87***	9.87***	5.80***	5.26**	5.19***

N	1270	1270	1268	1269	1270
---	------	------	------	------	------

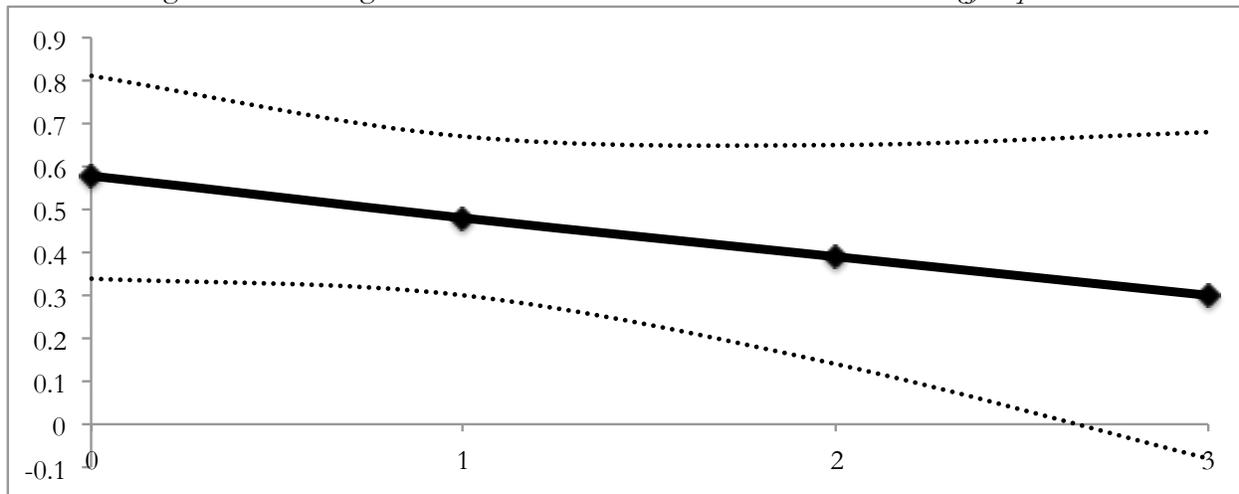
OLS estimates with robust standard errors clustered on the state.

* $p < .10$, ** $p < .05$, *** $p < .01$ (two-tailed) and † $p < .10$ (one-tailed).

We next explore whether this UNSC signal is stronger for more motivated individuals. In this issue-area, we measured an individual's motivation to look for and consider the UNSC authorization signal by asking them to rank four possible justifications for the use of military force (*to stop the spread of Islamic fundamentalism/ to secure energy supplies/ to promote human rights/ to insure the territorial integrity of other countries*).⁷ Since our conflict scenario concerned the security of energy resources, we used the respondents ranking of the second justification above to create an ordinal variable labeled *Energy Importance* (=0 for the least important and =3 for the most important among this set). This issue-specific motivational indicator is then interacted with the UNSC treatment ($UNSC \times Energy\ Importance$) and added, along with the *Energy Importance* constitutive term, to the specification in model 5.3 (Table 5).

As shown by the statistically insignificant $UNSC \times Energy\ Importance$ interaction term, there is no real moderating effect for the UNSC treatment based on this variable (*Energy Importance*). As illustrated in Figure 5, there is a statistically significant marginal effect for *UNSC* at all values of *Energy Importance*, except for the highest value where there are also the fewest observations (only 7%). While statistically insignificant, it is interesting to note the negative sign on this interaction variable, which is the opposite of our expectation. We expected a stronger and more positive UNSC signal for those who cared more about energy security as a justification for military force; instead, we find a somewhat weaker UNSC signal, both substantively and significantly.

⁷ The modal respondent ranked securing energy supplies as the lowest priority. Promoting human rights was ranked as the highest priority, followed by the territorial integrity of other states, and stopping the spread of Islamic fundamentalism.

Figure 5: The marginal effect of *UNSC* for different values of *Energy Importance*.

Y axis is the marginal effect of *UNSC*.

X axis is the value for *Energy Importance*.

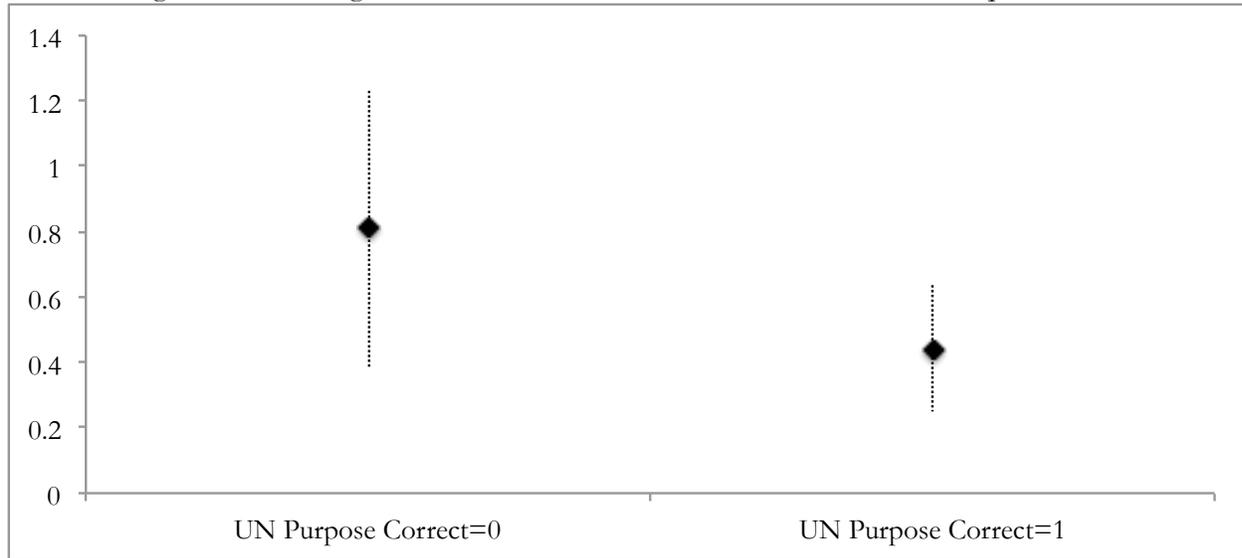
Dashed lines indicate 90 percent confidence intervals.

At the risk of post hoc theorizing, this weak negative effect may stem from the possibility that those who do view energy security as a higher priority are less motivated to look for UNSC authorization *because they believe military force would be justified even without it*. The weakly significant positive sign on the *Energy Importance* constitutive term (measuring the effect of greater energy supply importance for those who did not receive the UNSC treatment) is consistent with this possibility. If this is true, then it suggests that our issue-specific motivation indicator is a very poor measure of the concept. Accordingly, in the next version of this experiment, we will find another more satisfying way to measure an individual's motivation to look for the UNSC authorization signal.

We next consider to what extent knowledge about the United Nations conditions the UNSC signal. As done for NAFTA in the economic scenario, we simply asked respondents if they could correctly identify the primary purpose of the United Nations (*to help ensure peace and stability/ to facilitate international trade and commerce/ to fund development in needy countries*). We then created a dummy variable labeled *UN Purpose Correct* coded as 1 for all respondents (90%) who correctly identified

peace and security as the primary purpose from this set. This variable is then interacted with *UNSC* and added to the specification in model 5.4 (Table 5).

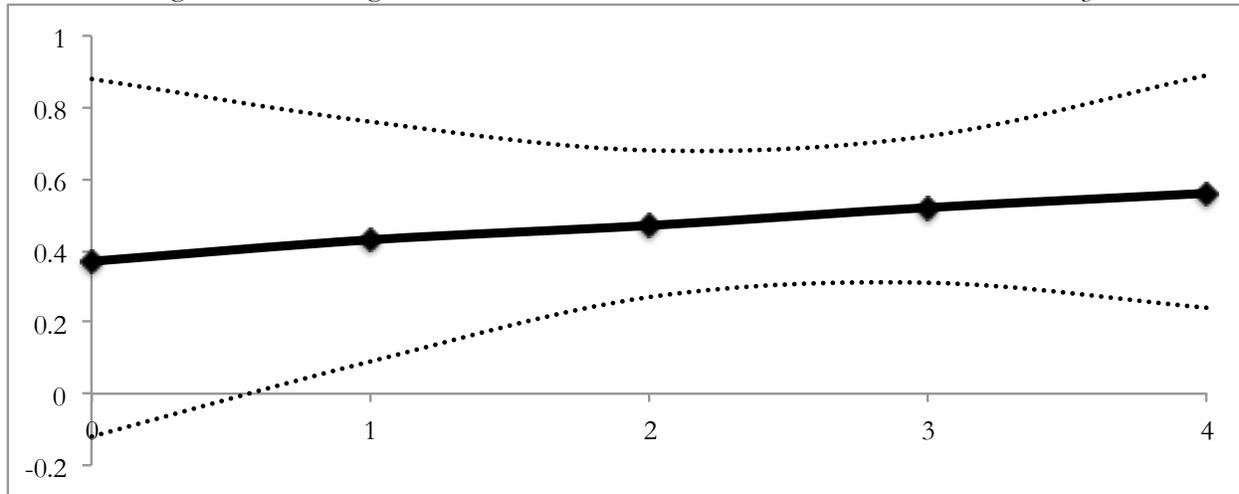
Figure 6: The marginal effect of *UNSC* for different values of *UN Purpose Correct*.



Y axis is the marginal effect of *UNSC*.

Dashed lines indicate 90 percent confidence intervals.

As illustrated in Figure 6, we find a statistically significant positive UNSC treatment for both groups, although this treatment effect is somewhat larger for the minority who were unable to correct identify the primary purpose of the UN, suggesting less knowledge of this particular IGO. We have no plausible post hoc explanation for this unexpected result except that it may be an artifact in this sample of American respondents. The weakly significant negative coefficient for $UNSC \times UN\ Purpose\ Correct$ observed above in model 5.4 in Table 5 will lose significance in model 6.4 in Table 6 below when we use a sample that includes only those respondents who passed the treatment check in this conflict scenario.

Figure 7: The marginal effect of *UNSC* for different values of *UN Favorability*.

Y axis is the marginal effect of *UNSC*.

X axis is the value for *UN Favorability*.

Dashed lines indicate 90 percent confidence intervals.

Finally, we explore to what extent the UNSC signal depends on the individual's pre-existing attitudes about the UN. Parallel to our measure in the first scenario, we asked respondents (in advance of the randomized treatment to assure exogeneity) for their opinion of the United Nations (*very unfavorable/ unfavorable/ no opinion/ favorable/ very favorable*). From their response, we created a five level ordinal variable (=0 if "very unfavorable" and =4 if "very favorable") labeled *UN Favorability*, which is interacted with *UNSC* and added to the specification in model 5.5 (Table 5).⁸ The results show no significant conditioning effect for *UNSC* based on attitudes about the UN. As illustrated in Figure 7, there is a statistically significant positive UNSC treatment effect for all values of *UN Favorability*, except for the lowest value (*UN Favorability* =0) where there are also the fewest observations (only 6%).

As a post hoc explanation for why we did not observe the same moderating effect for UN favorability in the conflict scenario as we did for NAFTA favorability in the economic scenario, we

⁸ The modal respondent in our sample has a "favorable" opinion concerning the United Nations with 26% having an unfavorable opinion and 20% having "no opinion."

offer the following possibility. It may be the case that the factors leading certain Americans to view the Security Council less favorably – the inability of the United States to control this multilateral forum and the fact that it often frustrates American foreign policy (e.g. its non-authorization before 2003 Iraq war) – are the very same factors that make it more useful as a forum to legitimize the use of military force. Its authorization is a rare event and even those who are unfavorably disposed to this international institution can recognize what it means when foreign powers with different strategic interests accept military force as legitimate in a particular context.

As done for the economic scenario, we also present the results using only the subsample that passed our treatment check for the conflict scenario.⁹ 75 percent of our respondents in the conflict scenario provided the correction answer, a number that closely corresponds the treatment check for the economic scenario (74%). This fact is important because it means that differences in the IGO treatment effect between the two issue-areas cannot easily be explained by differences in signal strength: almost the same percentage of American citizens successfully received the IGO signal in both scenarios.

Table 6: Estimates of Support for the Government’s Military Policy
Using Subsample Passing Treatment Check.

Model:	6.1	6.2	6.3	6.4	6.5
<i>Baseline</i>		0.32** (0.13)	0.31** (0.13)	0.31** (0.13)	0.31** (0.12)
<i>NoUNSC</i>	-0.32** (0.13)				
<i>UNSC</i>	0.34** (0.13)	0.65*** (0.12)	0.75*** (0.20)	1.11*** (0.41)	0.58* (0.33)
<i>Energy Importance</i>			0.11* (0.06)		
<i>UNSC x</i>			-0.11		

⁹ After the attitudinal dependent variable question, respondents were asked to identify which of the following best describes the government’s planned military response: *the action was not supported by a UN resolution / the action was supported by a UN resolution / the scenario did not mention the UN.*

<i>Energy Importance</i>			(0.12)		
<i>UN Purpose Correct</i>				0.31† (0.23)	
<i>UNSC × UN Purpose Correct</i>				-0.51 (0.42)	
<i>UN Favorability</i>					-0.17*** (0.05)
<i>UNSC × UN Favorability</i>					0.03 (0.13)
F	14.55***	14.55***	5.80***	8.28***	9.26***
N	951	951	951	950	951

OLS estimates with robust standard errors clustered on the state.

* $p < .10$, ** $p < .05$, *** $p < .01$ (two-tailed) and † $p < .10$ (one-tailed).

Table 6 presents the same sequence of models using this restricted sample and, once again, the IGO treatment effects appear even stronger. By dropping those who did not successfully receive their IGO signal, the deliberately un-moderated UNSC treatment effect increases by about 60 percent in the first specification (from 0.21* in model 5.1 to 0.34** in model 6.1) and by 33 percent in the second specification (from 0.49** in model 5.2 to 0.65*** in model 6.2). And even when we deliberately try to moderate the UNSC treatment with interaction variables, it remains strong regardless of attitudes about energy importance in model 6.3, knowledge about the UN's purpose in model 6.4, and attitudes about UN favorability in model 6.5. The *UNSC* interaction term is statistically insignificant in all three specifications.

3. Discussion

After presenting our argument about why many members of the mass public might have difficulty receiving a signal sent by/through an IGO, we advanced two related hypotheses. The first was that the reception of the intended IGO signal should be stronger for more motivated and

knowledgeable individuals and for those with more positive pre-existing attitudes about the relevant IGO. In the economic scenario, we found that the IGO signal (associated with FTA participation) was stronger for individuals who ranked inflation as a more important macroeconomic priority, thus making them potentially more motivated to make the necessary connection between trade protection and domestic prices. We also found that the intended FTA signal was stronger for individuals with more positive pre-existing attitudes about NAFTA. For those with negative pre-existing attitudes about NAFTA, the FTA signal had an unintended effect, making respondents *less* supportive of the government's economic policy.

In the conflict scenario, we found that the IGO signal (about UN Security Council authorization) had the intended effect *regardless* of respondent issue motivation, UN knowledge, and pre-existing attitudes about the UN. None of our moderating variables showed a statistically significant effect on the UNSC signal in the conflict scenario, unlike what was observed in the economic scenario. Indeed, these cross-issue results - which cannot be explained by experimental treatment strength differences given that the same percentage of respondents passed our treatment check in both scenarios - are consistent with our second hypothesis that the reception of the intended IGO signal should be stronger in higher political issue-areas.

The obvious advantage in our experimental research design was that the regression coefficients represent treatment effects. Our randomly assigned treatments were clearly exogenous to our measures of government support and policy preferences. But the potential disadvantage in using an experimental research design is external validity, or the extent to which our results apply to other individuals and other situations. We now conclude with a discussion of this issue.

With regards to the first dimension – do the results apply to other individuals? – we argue for external validity. The IR arguments that we considered in this paper concerned attitudes *within the mass public*, and our experimental sample also came from this societal group even if it was not a

nationally representative sample. Thus, the bigger concern about external validity concerns the second dimension: do the results apply to other situations? Obviously, both of our scenarios were hypothetical and were not felt by the respondents as actual real-world events even if they resembled the latter by design. So does the obvious hypothetical character of our scenarios tend to over-state or under-state the true IGO signaling effect?

In response to this question, we argue in favor of the latter: if anything, our experimental IGO results may be *under-stated* relative to their signaling effect in parallel real-world events. As evidence on this point, we compared the strong IGO signal found in our conflict scenario to the observed change in US public support for the use of military force to expel Iraq from Kuwait right before and after UNSC Resolution 678 on November 29, 1990. Data from Gallup suggests that public support increased by 27 percent,¹⁰ which represents a larger substantive effect that was observed in our parallel, but hypothetical, conflict scenario. The obvious comparison problem is that one cannot be sure that UNSC Resolution 67, which authorized the use of military force, caused all of this observed change in American public opinion; indeed, this is precisely why one should prefer experimental evidence to observational data when trying to assess a causal effect. But the effect size for the observational data is larger than the effect size shown in our experiment. On this basis, it is hard to argue that the hypothetical nature of our experimental scenarios tends to overstate the IGO signal.

¹⁰ <http://www.gallup.com/poll/1963/amer...orthwhile.aspx>.

References

- Barabas, Jason, and Jennifer Jerit. 2010. "Are Survey Experiments Externally Valid?" *American Political Science Review* 104 (May): 226-42.
- Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. 2012. "Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk." *Political Analysis* 20.3: 351-368.
- Chapman, Terrence L. 2012. *Securing approval: domestic politics and multilateral authorization for war*. University of Chicago Press.
- Chapman, Terrence L., and Stephen Chaudoin. 2013. "Ratification Patterns and the International Criminal Court1." *International Studies Quarterly* 57.2: 400-409.
- Chaudoin, Stephen. 2014. "Promises or policies? An experimental analysis of international agreements and audience reactions." *International Organization* 68.01: 235-256.
- Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, edited by David E. Apter, pp. 206-61. New York: Free Press of Glencoe.
- Dai, Xinyuan. 2007. *International institutions and national policies*. Cambridge University Press.
- Davis, Christina L. 2012. *Why adjudicate? enforcing trade rules in the WTO*. Princeton University Press.
- Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know about Politics and Why It Matters*. New Haven, CT: Yale University Press.
- Fang, Songying. 2008. "The informational role of international institutions and domestic politics." *American Journal of Political Science* 52.2: 304-321.
- Fearon, James D. 1994. "Domestic political audiences and the escalation of international disputes." *American Political Science Review* 88.03: 577-592.
- Ferejohn, John A. 1990. "Information and the Electoral Process." in Ferejohn, John A., and James H. Kuklinski. *Information and democratic processes*. University of Illinois Press.
- Gourevitch, Peter. 1978. "The second image reversed: the international sources of domestic politics." *International Organization* 32.04: 881-912.
- Holsti, Ole R. 2004. *Public opinion and American foreign policy*. Ann Arbor: University of Michigan Press.
- Hurwitz, Jon, and Mark Peffley. 1987. "How are foreign policy attitudes structured? A hierarchical model." *American Political Science Review* 81 (04): 1099-1120.
- Iyengar, Shanto. 1990. "Shortcuts to Political Knowledge: The Role of Selective Attention and Accessibility." in Ferejohn, John A., and James H. Kuklinski. *Information and democratic processes*. University of Illinois Press.

- Kertzer, Joshua D. and Ryan Brutger. 2015. "Decomposing Audience Costs: Bringing the Audience Back into Audience Cost Theory." *American Journal of Political Science*.
- Kinder, Donald R. 2007. "Curmudgeonly Advice." *Journal of Communication* 57 (1): 155-162.
- Levendusky, Matthew S., and Michael C. Horowitz. 2012. "When backing down is the right decision: Partisanship, new information, and audience costs." *The Journal of Politics* 74.02: 323-338.
- Lupia, Arthur, and Mathew D. McCubbins. 1998. *The democratic dilemma. Can Citizens learn what they need to know*. Cambridge University Press.
- Miller, Warren E., and Donald E. Stokes. 1963. "Constituency influence in Congress." *American political science review* 57(01): 45-56.
- Pelc, Krzysztof J. 2013. "Googling the WTO: what search-engine data tell us about the political economy of institutions." *International Organization* 67.03: 629-655.
- Popkin, Samuel L. and Michael A. Dimock. 2000. "Knowledge, Trust, and International Reasoning." In Lupia, Arthur, Mathew D. McCubbins, and Samuel L. Popkin, eds. *Elements of reason: Cognition, choice, and the bounds of rationality*. Cambridge University Press.
- Mansfield, Edward D., Helen V. Milner, and B. Peter Rosendorff. 2002. "Why democracies cooperate more: Electoral control and international trade agreements." *International Organization* 56.03: 477-513.
- Simmons, Beth A., and Allison Danner. 2010. "Credible commitments and the international criminal court." *International Organization* 64.02: 225-256.
- Tomz, Michael. 2007. "Domestic audience costs in international relations: An experimental approach." *International Organization* 61.04: 821-840.
- Waltz, Kenneth N. 1959. *Man, the state, and war: a theoretical analysis*. Columbia University Press.