Appendix to "The Domestic Sources of International Reputation"

Contents

1 Formal Model

In this section we formalize our theory. Following the experimental literature, we focus on the question: When country A observes country B take a costly action (e.g., fight in a crisis) how does it influence A's beliefs about whether B will stand firm if A challenges B in a crisis? We analyze the simplest possible model for understanding reputation in repeated crises. Others have introduced additional features into the model to match their substantive focus. For example, some randomly perturb B's payoff each period and allow B to engage in costless diplomacy (Sartori 2005). Others study in-crisis militarization (Renshon, Dafoe, and Huth 2018). Others still allow A to deploy forces in the first crisis, to increase the amount that A can learn about B's type (Slantchev 2011). Our results are consistent with the findings in these more complex models.

Set up

We study a two-period interaction between two actors—A and B.We depict the sequence of moves and payoffs in Table 1 in the manuscript. In the first period, B is faced with a crisis[1] and either stands firm (SF) or backs-down (BD). If B stands firm he enters a costly lottery (winning with p and losing otherwise) to take the issue contested in the crisis. If B backs down, B concedes the good at no cost.

The second period models an international militarized crisis between A and B over a contested foreign policy issue. In this crisis, A is given the opportunity to threaten B over the issue in dispute. If A makes no threat, B keeps the issue in dispute. If A makes a threat, B has an opportunity to stand firm (SF) or backdown.

Past studies have assumed that the first period is also a foreign policy crisis. We do not restrict ourselves to this assumption. Rather, we innovate by assuming that the first period can represent any setting where B faces a choice between backing down and standing firm, where standing firm yields potentially higher benefits, but is more costly and risky.

Most features of the payoff structure are similar to past models. A's payoffs depend on the outcome of the second period crisis. These payoffs include three terms: (1) The benefit for capturing the issue in dispute, 1. (2) An expectation about victory if A is forced to fight

[1]Possibly a domestic crisis or a foreign policy crisis instigated by a third-party.

for the issue in dispute, $(1-p \in (0,1))$. (3) A cost of fighting, $c_A > 0$.

B's payoff includes three equivalent terms: (1) The benefit for retaining the issue in dispute, $\pi > 0$. (2) An expectation about victory if B is forced to fight for the issue in dispute, $(p \in (0,1))$. (3) A cost if B chooses to take a costly action in either the first or second period: $c_1, c_2 \in [0,1]$. c_2 represents how resolved B is to stand firm and fight if threatened in a foreign policy crisis.[2] We assume that c_2 is constant across international crises and is known privately by B. To model uncertainty about B's resolve to stand firm in an international crisis, we assume Nature draws $c_2 \sim f()$ at the beginning of the game and shows it privately to B. For simplicity, we model the case where $f = U[0,1]$.

We introduce two concepts into the model that represent the two ways that domestic decision-making can differ from foreign policy decision-making: cost similarity and salience.

Cost similarity: After Nature draws c_2, Nature sets $c_1 = c_2$ with probability $\alpha \in 0, 1$. Nature draws an uncorrelated $c_1 \sim U[0,1]$ with probability $1-\alpha$. B knows its own exact values of c_1 and c_2, whereas A only knows the probability α.

Substantively, α represents the level of cost similarity between the first period choice and a foreign policy crisis. The classic international reputation model in which the first period is a foreign crisis is represented by $\alpha = 1$. In this case, the costs for standing firm in the first and second periods are assumed to be the same because they both involve human casualties, the chance of losing office and brutal treatment after exit, damage to national honor, etc. If the leader of B is sensitive to these sorts of costs in the first period, they will also be sensitive to them in the second.

Any $\alpha < 1$ implies that the first period is a domestic policy choice rather than a foreign crisis, but a higher α suggests a greater probability that the costs are still the same. When $\alpha = 0$, it indicates that B's costs for standing firm domestically in the first period are certainly uncorrelated with B's costs for fighting in the second period. In this case, if we learned B's exact value c_1, we would still have no additional information about c_2.

Salience: We model salience as a resolution parameter θ that amplifies B's costs and benefits for standing firm in the first period. Thus, B's utility from standing firm in the first period is $\theta(\pi p - c_1)$. When $\theta > 1$, it implies that the domestic choice is more salient than a foreign policy crisis. When $\theta < 1$, it implies that the domestic choice is less salient.

Definition The classic model of international reputation is a special case of the model presented here where: $\alpha = 1, \theta = 1$.

A strategy for A is a single choice: $s_2^A \in \{T, NT\} \, | s_1^B$. That is, A chooses to make a threat or not in the second period crisis. This choice is possibly conditional on B's strategy. A strategy for B is two choices $s^B(s_1 \in \{SF_1, BD_1\}, s_2 \in \{SF_2, BD_2\}) | f(), s^A$.

Summary of our results that are useful for our experiment

The following summarizes the features of the technical analysis that we use to generate predictions for our experiment. Following the literature on reputation for resolve, our basic question is: If B stands firm in the first period, how does this influence A's belief about how B will behave in the second period? We are not interested in a complete analysis of

[2] B's resolve to fight is based partly on its leader's individual psychological or dispositional characteristics.

the conditions that generate war and peace. Rather, our goal is to make precise predictions about beliefs that can be tested in an experimental setting.

In what follows, we answer this question with a focus on pure strategy semi-separating equilibria. These are not the only equilibria.[3] For ease of discussion, we refer to pure strategy semi-separating equilibria as separating equilibria.

To make claims about reputation, we must define what A is uncertain about. We argue that A wants to know the probability that B will stand firm if threatened. Since the second period is the last period, B's decision to stand firm if challenged is a simple cost-benefit analysis. B stands firm if $\pi p > c_2$ and not otherwise. A knows this is B's cost-benefit calculation. What A does not know is B's sensitivity to the costs of fighting relative to B's benefit from fighting: $pr(\pi p > c_2)$. Our core question relates to how A's beliefs about this probability changes across different stages of the game.

Define three belief-states that A can hold in a separating equilibrium. Define $y(prior) = pr(\pi p > c_2|f())$ as A's belief that B will stand firm if threatened in the second period crisis at the beginning of the game. This belief is based on the prior distribution $f(c)$.

Define $y(SF_1) = pr(\pi p > c_2|f(), \theta, \alpha, s_1 = SF)$ as A's belief that B will stand firm if threatened in the second period crisis after A observes B stand firm in the first period (in the separating equilibrium). We emphasize that the model's dynamics can change as a function of α, θ because we will later analyze shifts in A's beliefs as a function of these parameters.

Define $y(BD_1) = pr(\pi p > c_2|f(), \theta, \alpha, s_1 = SF)$ as A's belief that B will stand firm if threatened in the second period crisis after A observes B back down in the first period (in the separating equilibrium).

In our experimental analysis, we tell subjects that B has faced a variety of domestic and international choices. We then randomly vary whether B stood firm or backed down in each. To wit, the difference between these treatments is summarized by:

$$\mathcal{Q}|\theta, \alpha = y(SF_1) - y(BD_1) \tag{1}$$

The belief state represents the following version of our main question: Given that A observes B faced with an opportunity to stand firm over some issue, what is the relative difference in A's beliefs given B's choice to stand firm versus back down?

We say that B's decision to stand firm in the first period cultivates a reputation for resolve if $\mathcal{Q} > 0$. We will show that there is a separating equilibrium where it is possible to cultivate a reputation for resolve.

Expectation 1 *There are conditions where B can cultivate a reputation for resolve to fight in an international crisis based on B's domestic choice.*

Our theory is not only about if reputations can form. Rather, it is about how informative different domestic choices can be for cultivating a reputation for resolve in a foreign policy crisis. This is why we emphasize that $\mathcal{Q}$ is conditional on θ, α.

Our goal is to understand how $\mathcal{Q}$ varies with marginal increases in the dimensions along which domestic choices can vary. We will show that $\frac{\partial \alpha}{\partial \mathcal{Q}} > 0$ for all $\alpha \in (0, 1)$. This finding leads to the following expectation that we describe informally in the manuscript:

[3] We can generate substantively similar predictions if we analyze mixed strategy equilibria.

Expectation 2 *As the costs associated with standing firm when facing a domestic choice look increasingly like the costs associated with the choice to stand firm in a foreign crisis ($\alpha \to 1$), it becomes easier for B to exploit that domestic choice to cultivate a reputation for international resolve ($\mathcal{Q}$ increases).*

We will also analyze how B's capacity to learn shifts as a function of θ. We will show $\frac{\partial \theta}{\partial \mathcal{Q}} > 0$. This finding leads to the following expectation that we describe informally in the manuscript:

Expectation 3 *As the domestic choice increases in salience ($\theta \to \infty$), it is easier for B to exploit that domestic choice to cultivate a reputation for international resolve ($\mathcal{Q}$ increases).*

The partial derivatives help us understand how these two dimensions alter A's capacity to learn from B's behavior in absolute terms. However, we are also interested in how informative certain domestic choices are relative to international crises. Thus, we define

$$\mathcal{Q}_{int} = \mathcal{Q}|\alpha = 1, \theta = 1 \tag{2}$$

In the manuscript, we use Figure 1 to show that it is theoretically possible that B's domestic choices can be more informative about how B will behave in a future foreign policy crisis than B's choice in a prior foreign policy crisis. That is, there exists $\mathcal{Q}(\theta > 1, \alpha < 1) > \mathcal{Q}_{int}$

Expectation 4 *It is possible that a past domestic choice to stand firm will tell us more than a prior foreign policy crisis about whether B will stand firm in a future foreign policy crisis.*

Statement of Separating Equilibrium

To begin, we report the conditions for the separating equilibrium.

Proposition 1.1 *If*

$$y(SF_1) > \frac{1}{p + c_A} > y(BD_1) \tag{3}$$

$$\theta > \frac{\pi}{1 - p\pi} \tag{4}$$

is satisfied, then there is a separating PBE where all types play pure strategies. In it, $s_2^A = T|BD_1, s_2^A = NT|SF_1$.

B plays $s^B(SF_1, SF_2)$ if $c_2 < \pi p$ and $c_1 < \pi p + \frac{c_2 + \pi(1-p)}{\theta}$. B plays $s^B(BD_1, SF_2)$ if $c_2 < \pi p$ and $c_1 > \pi p + \frac{c_2 + \pi(1-p)}{\theta}$. B plays $s^B(SF_1, BD_2)$ if $c_2 > \pi p$ and $c_1 < \pi(p + \frac{1}{\theta})$. B plays $s^B(BD_1, BD_2)$ if $c_2 > \pi p$ and $c_1 > \pi(p + \frac{1}{\theta})$.

We solve for this PBE working backwards. If A threatens in the second period crisis, then B will backdown if $\pi p < c_2$ and not otherwise.

Working backwards, if A makes a threat, it gets $y(1 - p - c_A) + (1 - y)$. Here y represents A's belief that B will stand firm if threatened given some history of the game and each players' on path strategy in the separating equilibrium. If A does not make a threat, it gets 0. Using these two values, A makes a threat if $y < \frac{1}{p+c_A}$ and not otherwise. In a separating equilibrium, it must be the case that this inequality is satisfied if A observes B back down in the first period (yielding $y(BD_1)$) but not if A observes B stand firm (yielding $y(SF_1)$). This gives us the conditions for inequality 3.

Finally, we derive B's choice to stand firm or back down in the first period. Given A's response, B's value for standing firm is $\theta(\pi p - c_1) + \pi$. B's benefit from backing down or standing firm depends on the instant payoff B receives as well as the strategic implications for the second period crisis.

There are two cases to consider. If $p\pi < c_2$, B will back down in the second period crisis if threatened. In this case, B's value from backing down in the first period is 0. Thus, for types $p\pi < c_2$, B stands firm in the first period if $\theta(\pi p - c_1) + \pi > 0 \implies c_1 < \pi(p + 1/\theta)$. B backs down otherwise. This condition bounds B's strategy in the equilibrium as desired.

Notice that this creates an incentive for some types to overstate their resolve to stand firm in the second period by standing firm in the first period. Specifically, consider the group $\pi p < c_1 < \pi(p + 1/\theta), p\pi < c_2$. These types would not stand firm in the second period crisis if threatened. Furthermore, these incur a net loss if they stand firm in the first period ($\pi p < c_1$). Thus, if A could credibly promise to make a threat no matter what B did, then this type would back down in both periods. But this type stands firm in the first period and incurs a net loss. The reason is that standing firm induces A to play NT in the second period. The amount this type gains from taking the second issue is enough to compensate it for paying the first period cost of standing firm.

If instead, $p\pi > c_2$, B will stand firm in the second period crisis if threatened. In this case, B's value for backing down in the first period is $\pi p - c_2$. Thus, for types $p\pi > c_2$, B stands firm in the first period if $\theta(\pi p - c_1) + \pi > \pi p - c_2 \implies c_1 < \pi p + \frac{c_2 + \pi(1-p)}{\theta}$. B backs down in the first period otherwise. This condition bounds B's strategy in the equilibrium as desired.

A's beliefs in the separating equilibrium

We compute A's beliefs in the separating equilibrium.

Lemma 1.2 *In the separating equilibrium, A's beliefs that B will stand form in the second period are:*

$$y(prior) = \pi p \tag{5}$$

$$y(SF_1) = \alpha \frac{\pi p}{\pi(p + \frac{1}{\theta})} + (1-\alpha) \frac{\pi p \left(\pi(p + \frac{1}{\theta}) - \frac{\pi p}{2\theta}\right)}{\pi(p + \frac{1}{\theta}) - \frac{(\pi p)^2}{2\theta}} \tag{6}$$

$$y(BD_1) = \alpha 0 + (1-\alpha) \frac{\pi p \left(1 - \pi(p + \frac{1}{\theta}) + \frac{\pi p}{2\theta}\right)}{1 - \pi(p + \frac{1}{\theta}) + \frac{(\pi p)^2}{2\theta}} \tag{7}$$

We explicitly separate out A's beliefs as a function of whether B's first and second period costs are correlated (which happens with probability α) or not $(1-\alpha)$. We can interpret A's posterior beliefs as follows. There is a α probability that the costs of standing firm are correlated ($c_1 = c_2$), so the term multiplied by α represents A's expectation that follows from what different types of B would do in this case. There is a $1-\alpha$ probability that the costs of standing firm are uncorrelated. The term multiplied by $1-\alpha$ represents A's expectation that follows from what different types of B would do in this case. We analyze A's prior, then consider these two cases in turn.

Since c_2 is drawn from $U \sim [0,1]$, A's prior is simply $\int_{-\infty}^{\pi p} f()dc_2 = \pi p$

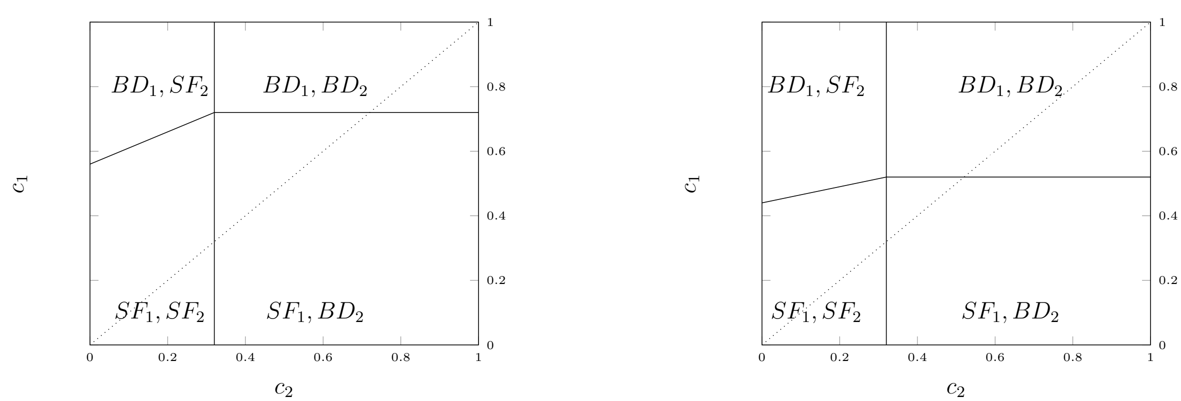
Turning to A's posterior beliefs, we begin by considering the case where Nature assigns $c_1 = c_2$. This happens with probability α. In this case, types $c_2 \geq \pi p$ will stand firm if threatened in the second period. The remaining types $c_2 < \pi p$ will not stand firm if threatened in the second period. Within this group there are two types that are separated by the partition at $c_2 = \pi(p + 1/\theta)$. Those who satisfy $\pi p < c_1 = c_2 < \pi(p+1/\theta)$ stand firm in the first period even though they are unwilling to stand firm in the second. Those in the region $\pi(p+1/\theta) < c_1 = c_2$ back down in the first period accepting that B will threaten them, and they will back down a second time.

Thus, we can compute A's beliefs if A observes B back down in the first period $y(BD_1)$ in the case that $c_1 = c_2$. Notice that in this case where $c_1 = c_2$, no types who back down in the first period are willing to stand firm in the second. It follows that if $c_1 = c_2$, then $y(BD_1) = 0$. This gives us the $\alpha 0$ term in $y(BD_1)$. Clearly, this implies that so long as $f()$ supports types that are willing to back down, backing down decreases A's confidence that B will stand firm in the second period crisis: $y(prior) - y(BD_1) < 0 | c_1 = c_2$.

We can also compute A's beliefs if A observes B stand firm in the first period $y(SF_1)$ in the case that $c_1 = c_2$. If B backs down, A can now be certain that if $c_1 = c_2$, then B's costs must satisfy $c_1 = c_2 < \pi(p + \frac{1}{\theta})$. We can compute the proportion of types that will stand firm in the second period crisis given the number of types that stood firm in the first period as $\frac{\int_{-\infty}^{\pi p} f()dc_2}{\int_{-\infty}^{\pi(p+1/\theta)} f()dc_2}$. This solves for $\frac{\pi p}{\pi(p+\frac{1}{\theta})}$ as desired. It follows that so long as $f()$ supports values of c from $[\pi p, \pi(p + 1\theta)$, B's choice to stand firm in the first period increases A's confidence that B will stand firm in the second period crisis: $y(prior) - y(SF_1) > 0 | c_1 = c_2$.

Figure 1: A's on-path actions in separating equilibrium

(a) Assumes high θ (b) Assumes low θ



Each plot represents A's on path actions in both periods under the assumption that c_1 is drawn i.i.d from c_2 (this happens with pr. $1-\alpha$).

These conditions also illustrate an insight that is critical for our analysis of θ later on. Notice that as $\theta \to \infty$, that $y(SF_1) \to 1$. The reason is that as θ grows larger, B is more sensitive to the consequences of what happens in the first period. Fewer types are willing to pay the cost for standing firm in the first period because they care more about the outcome in first period than the second. In the limit, the first issue is so important that B does not care about the consequences of the second period crisis at all and does not factor in A's strategic response into B's first period decision. In this case, A can draw a complete inference from SF_1.

We now consider the case that c_1, c_2 are uncorrelated and drawn i.i.d from $f()$. This case occurs with probability $1-\alpha$. We visualize each B's best reply to A's strategy in Figure 1. Along the x-axis are possible values of c_2. Along the y-axis are possible values of c_1. The dotted line marks the case where $c_1 = c_2$. We partition the figure into four quadrants based on the strategy that A plays for draws of c_1, c_2. We derived these partitions in our discussion of B's preferences in the equilibrium analysis.

The independence of draws alters A's inferences in two ways. First, consider the case where A observes B stand firm in the first period. In the correlated model, B could rule out the possibility that $c_2 > \pi(p + 1/\theta)$ because $c_1 = c_2$. But in the uncorrelated model, c_2 is drawn independently from c_1. It follows that A's posterior belief $y(SF_1)$ in the case that c_1 and c_2 are uncorrelated is $\frac{\int_{-\infty}^{\pi p} \int_{-\infty}^{\pi p + \frac{c_2 + \pi(1-p)}{\theta}} f() dc_1 dc_2}{\int_{-\infty}^{\pi p} \int_{-\infty}^{\pi p + \frac{c_2 + \pi(1-p)}{\theta}} f() dc_1 dc_2 + \int_{\pi p}^{\infty} \int_{-\infty}^{\pi(p + \frac{1}{\theta}} f() dc_1 dc_2}$. This solves for the value multiplied by $1-\alpha$ as desired.

Second, consider the case where A observes B back down in the first period. In the correlated model, types that would back down in the first period would certainly back down in the second. However, it is possible that B has a high c_1 and a low c_2 and so B will play BD_1, SF_2. It follows that A's posterior belief $y(BD_1)$ in the case that c_1 and c_2 are uncorrelated is $\frac{\int_{-\infty}^{\pi p} \int_{\pi p + \frac{c_2 + \pi(1-p)}{\theta}}^{\infty} f() dc_1 dc_2}{\int_{-\infty}^{\pi p} \int_{\pi p + \frac{c_2 + \pi(1-p)}{\theta}}^{\infty} f() dc_1 dc_2 + \int_{\pi p}^{\infty} \int_{\pi(p + \frac{1}{\theta})}^{\infty} f() dc_1 dc_2}$. This solves for the value multiplied by $1-\alpha$ as desired.

A closer look at the results from the uncorrelated case shows that if we could support a separating equilibrium, then A's posterior beliefs would satisfy $y(BD_1) \geq y(prior) \geq y(SF_1)|\alpha = 0$. That is, if A observed B stand firm in the first period, A would be less confident that B would stand firm in the second period than A was in the prior state. However, if A observed B back down in the first period, A would be more confident that B would stand firm in the second period relative to the prior state.

This result follows from the introduction of types whose preferences for standing firm vary across issues. In the correlated model, if $c_1 > \pi p$, then $c_2 > \pi p$. Thus, if B's choice had no strategic implications, B would make the same choice in both periods. In the uncorrelated model, this is not the case. There are now types that would prefer to back down in the first period but stand firm in the second. Types $c_1 > \pi p$, $c_2 < \pi p$ are the least likely to succumb to incentives to misrepresent. The reason is that the incentive to misrepresent is driven by B's preferences to avoid paying c_2 (because A plays no threat if B plays SF_1) by paying c_1. Since these types face the smallest incentive to misrepresent, they are more likely to conform to the strategy that matches their preferences for each individual period. It follows that when A observes B back down, A increases its confidence that B is likely to stand firm in the next period.

This result also provides an interesting insight about salience (θ) that will help us compute our marginal effects in the next section. As θ increases, the incentive to pay a cost in the first period for a strategic advantage in the second diminishes for all types. Thus, all types that incur a net first period loss $\pi p < c_1$ revert to BD_1, and the first period is totally uninformative. It follows that in the case that c_1 and c_2 are uncorrelated, that as $\theta \to \infty$, $y(SF_1) \to y(prior)$ from below, and $y(BD_1) \to y(prior)$ from above.

Deriving our quantities of interest for experimental testing

We can use these beliefs to compute the quantity of interest. Starting with the case of international reputation: $\mathcal{Q}_{int} = \frac{\pi p}{\pi(p+1)}$ Then also the general case that covers every domestic issue: $\mathcal{Q}(\theta, \alpha) = \alpha \frac{\pi p}{\pi(p+\frac{1}{\theta})} - (1-\alpha) \frac{(\pi p)^2(1-\pi p)}{2\theta\left(\pi(p+\frac{1}{\theta})-\frac{(\pi p)^2}{2\theta}\right)\left(1-\pi(p+\frac{1}{\theta})+\frac{(\pi p)^2}{2\theta}\right)}$

When $\mathcal{Q}$ is positive, it means that B can cultivate a reputation for resolve in an international crisis. In our experiment, we are interested in the relative informational content of choices in domestic and international situations. We argue that these situations vary along two dimensions: salience (θ) and cost similarity (α). Thus, we compute the partial derivative: $\frac{\partial \alpha}{\partial \mathcal{Q}_1} = \frac{\pi p}{\pi(p+\frac{1}{\theta})} + \frac{\pi p^2(1+\pi p)}{2\theta\left(\pi(p+\frac{1}{\theta})-\frac{(\pi p)^2}{2\theta}\right)\left(1-\pi(p+\frac{1}{\theta})+\frac{(\pi p)^2}{2\theta}\right)} > 0.$

This implies that holding all the other parameters of the model constant, as the costs in the first period become more similar to the second, it is easier for B to cultivate a reputation based on its first period choice. In the limit, as $\alpha \to 1$, $\mathcal{Q} \to \frac{\pi p}{\pi(p+\frac{1}{\theta})}$.

This follows intuitively from the analysis above. We saw that in the separating equilibrium that A always draws an inference $y(SF_1) > y(prior) > y(BD_1)|\alpha = 1$ and $y(SF_1) < y(prior) < y(BD_1)|\alpha = 0$. It follows that as we place more weight on α, we get a larger difference between $y(SF_1) - y(BD_1)$.

Second, we compute the partial derivative: $\frac{\partial \theta}{\partial \mathcal{Q}_1} = \alpha \frac{p}{(p\theta+1)^2} + (1-\alpha) \frac{2p^2\pi(1-p\pi)(4(\pi+\theta^2 p-p^2\pi\theta^2-p^2\pi^2)+p^4\pi^3)}{(p^2(-\pi)+2p\theta+2)^2(\pi(p^2\pi-2)+\theta(2-2p\pi))^2}$.

Both fractions are positive so long as $4(\pi+\theta^2 p-p^2\pi\theta^2-p^2\pi^2)+p^4\pi^3 > 0$. This must be the case given our assumption $\theta > \frac{\pi}{1-\pi p}$. This implies that holding all the other parameters of the model constant, as the first period situation becomes more salient, it is easier for B to cultivate a reputation based on its first period choice. In the limit, as $\theta \to \infty$, $\mathcal{Q} \to \alpha$.

This follows intuitively from the analysis above. Recall that when c_1, c_2 are correlated and A observed B play SF_1, that $y(SF_1)|\alpha = 1$ was increasing in θ. In the limit, when $\theta \to \infty$, A inferred that $y(SF_1)|\alpha = 1 \to 1$. It follows that increasing θ allows for stronger inferences in the case that costs are correlated and B plays SF_1.

We also saw in the case where c_1, c_2 were uncorrelated, that for any set of parameters $y(SF_1) < y(prior) < y(BD_1)|\alpha = 0$. That is, observing B stand firm (back down) in the first period reduces (increases) A's confidence that B will stand firm in the second period crisis. However, as $\theta \to \infty$, then $y(SF_1) \to y(prior), y(BD_1) \to y(prior)$. It follows that the negative influence on learning is lessened as θ grows larger. This leaves us only with the positive learning effects that occur in the case that c_1, c_2 are correlated.

Does variation in the probability of success have an independent affect on $\mathcal{Q}$?

Our theory makes claims about how changes in cost similarity (α) and salience (θ) impact the amount of learning that happens in a separating equilibrium ($\mathcal{Q}$). One might wonder if these effects are mediated by the probability of success (p). We showed in our partial derivative analysis that they are not. That is, we could show that the partial derivatives for α, θ were positive given any value of p. This implies that p does not confound our core predictions.

Still one might wonder: does a change in p also exert an independent affect on $\mathcal{Q}$? To see if this is the case, we solve for the partial of $\mathcal{Q}$ with respect to p:

$$\{(1+\theta p)^2(2+2\theta p-\pi p^2)^2(2\theta+\pi^2p^2-2\pi(1+\theta p))^2\}$$
$$/$$
$$\{\theta(2\pi p(1+\theta p)^2(-2\pi^3(-2+p)p^3+\pi^4p^5-4\theta(2+\theta p)+4\pi(2+5\theta p+2\theta^2p^2)-2\pi^2p(6-\theta(-8+p)p+2\theta^2p^2))+$$
$$a(\pi^6p^8+16\theta^2(1+\theta p)^2-2\pi^5p^6(5+6\theta p+\theta^2p^2)+4\pi^4p^4(4+p+8\theta p+\theta(3+4\theta)p^2+\theta^2p^3)+$$
$$8\pi\theta(-4+(2-12\theta)p+3(1-4\theta)\theta p^2+2(1-2\theta)\theta^2p^3+\theta^3p^4)-4\pi^2(-4+(4-16\theta)p+6(1-4\theta)\theta p^2+$$
$$4(1-4\theta)\theta^2p^3-\theta^2(1-6\theta+4\theta^2)p^4+4\theta^4p^5)+4\pi^3p^2(-2-8\theta^2p^3-\theta^3(-4+p)p^3+2\theta^4p^4-\theta p(4+7p))))\}$$

Whether this result is positive or negative is a function of θ. This implies two facts. First, it means that we cannot make monotonic predictions about how p affects $\mathcal{Q}$. Second, the amount that the probability of success affects $\mathcal{Q}$ is inextricably linked to the salience of the first period issue.

2 Experiments

2.1 Description of Ethical Practices

This section summarizes the ethical practices in each of our experiments.

Consent: Participants viewed a consent form (included in the survey instruments) and clicked a button to consent.

Deception: There was no deception. All scenarios were clearly labeled as hypothetical.

Confidentiality: Responses were anonymous, and our dataset contains no identifying information.

Harm and impact: We do not foresee any possible harm or lasting impact on respondents.

Compensation: We paid Lucid one dollar for each participant who passed the attention checks and completed the survey. According to Lucid's website, "Lucid's partnering companies find research participants from a diverse array of sources, many of which are double opt-in panels. These companies invite participants to partake in research opportunities through emails, push notifications, in-app pop-ups, or through offerwalls of engagement opportunities. These companies incentivize their users to participate in opportunities often by sharing the revenue earned for a survey's complete."

IRB approval numbers:
Yale: 2000030502
UCSD: 2000030502
Penn State: STUDY00017730

2.2 Main Study Additional Tables and Figures

Main Study AMCE Estimates

OLS estimates of Average Marginal Component Effects (AMCEs) with robust standard errors clustered at the respondent level are presented below (Table 1).

Table 1: AMCEs for Domestic Politics and Resolve Conjoint

	Six-Point Outcome		Binary Outcome	
	No Controls	Controls	No Controls	Controls
Coup (vs. Non-violent Ascent)	0.40***	0.41***	0.11***	0.11***
	(0.03)	(0.03)	(0.01)	(0.01)
Repression (vs. Allowed Protests)	0.83***	0.83***	0.24***	0.24***
	(0.03)	(0.03)	(0.01)	(0.01)
Gave Speech (vs. Cancelled)	0.03	0.03	0.02*	0.02*
	(0.03)	(0.03)	(0.01)	(0.01)
Inflexible Domestically (vs. Compromised)	0.24***	0.24***	0.07***	0.07***
	(0.03)	(0.03)	(0.01)	(0.01)
Fought in Past Crisis (vs. Backed Down)	0.78***	0.77***	0.25***	0.25***
	(0.03)	(0.03)	(0.01)	(0.01)
Age		−0.00***		−0.00***
		(0.00)		(0.00)
Male		−0.14***		−0.05***
		(0.04)		(0.01)
USD 14,999 or less		−0.00		−0.05
		(0.09)		(0.03)
USD 15,000-19,999		−0.04		−0.03
		(0.11)		(0.04)
USD 20,000-24,999		−0.06		−0.04
		(0.12)		(0.04)
USD 25,000-29,999		−0.01		−0.03
		(0.12)		(0.04)
USD 30,000-34,999		−0.02		−0.03
		(0.11)		(0.03)
USD 35,000-39,999		0.06		0.03
		(0.11)		(0.04)
USD 40,000-44,999		0.11		0.03
		(0.11)		(0.04)
USD 45,000-49,999		0.24*		0.01
		(0.12)		(0.04)
USD 50,000-54,999		0.09		0.03
		(0.11)		(0.04)
USD 55,000-59,999		−0.12		−0.06
		(0.13)		(0.04)
USD 60,000-64,999		−0.07		−0.05
		(0.13)		(0.04)
USD 65,000-69,999		0.13		0.01
		(0.14)		(0.05)
USD 70,000-74,999		0.10		0.02
		(0.13)		(0.04)
USD 75,000-79,999		0.23		0.06
		(0.13)		(0.05)
USD 80,000-84,999		−0.05		−0.04
		(0.14)		(0.05)
USD 85,000-89,999		−0.15		−0.06
		(0.21)		(0.07)
USD 90,000-94,999		−0.15		−0.10
		(0.16)		(0.05)
USD 95,000-99,999		−0.22		−0.09
		(0.13)		(0.05)
USD 125,000-149,999		0.36**		0.10*
		(0.13)		(0.04)
USD 150,000-174,999		0.16		0.04
		(0.14)		(0.05)

	Six-Point Outcome		Binary Outcome	
	No Controls	Controls	No Controls	Controls
USD 175,000-199,999		−0.10		−0.02
		(0.15)		(0.05)
USD 200,000-249,999		0.07		0.02
		(0.19)		(0.05)
USD 250,000 and above		0.08		0.03
		(0.20)		(0.06)
USD prefer no answer		−0.06		−0.03
		(0.10)		(0.04)
Ethnicity: Indian		−0.18		−0.03
		(0.25)		(0.09)
Ethnicity: Chinese		−0.27		−0.08
		(0.21)		(0.07)
Ethnicity: Filipino		−0.08		0.00
		(0.25)		(0.08)
Ethnicity: Japanese		0.04		0.03
		(0.25)		(0.10)
Ethnicity: Korean		−0.57		−0.17
		(0.36)		(0.10)
Ethnicity: Other Asian		−0.10		0.01
		(0.20)		(0.09)
Ethnicity: Vietnamese		−0.02		0.07
		(0.20)		(0.08)
Ethnicity: Black		−0.23		−0.07
		(0.16)		(0.06)
Ethnicity: Guamanian		−0.54**		−0.11
		(0.18)		(0.06)
Ethnicity: Hawaiian		−0.77***		−0.28***
		(0.18)		(0.06)
Ethnicity: Other Pacific Island		−0.11		0.05
		(0.29)		(0.12)
Ethnicity: Prefer not to answer		−0.25		−0.10
		(0.18)		(0.06)
Ethnicity: Other		−0.03		−0.03
		(0.18)		(0.06)
Ethnicity: White		−0.15		−0.04
		(0.16)		(0.05)
Some HS or less		−0.33**		−0.06
		(0.12)		(0.04)
High school		−0.10		−0.02
		(0.07)		(0.02)
Some College		−0.01		0.01
		(0.07)		(0.02)
Vocational		−0.07		−0.04
		(0.14)		(0.04)
Bachelor's		−0.01		−0.00
		(0.07)		(0.02)
Master's or professional		0.11		0.03
		(0.08)		(0.03)
Doctorate		0.14		0.02
		(0.14)		(0.04)
None of above (edu)		−0.26		−0.05
		(0.28)		(0.07)
Independent		−0.09		−0.02
		(0.06)		(0.02)
Republican		−0.10*		−0.04**
		(0.04)		(0.01)
Intercept	2.57***	3.05***	0.22***	0.40***
	(0.04)	(0.19)	(0.01)	(0.07)
Num. obs.	9302	9302	9302	9302
N Clusters	1878	1878	1878	1878

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$ Robust standard errors clustered at the subject level in parentheses

Main Study Linear Hypothesis Tests

To assess the difference in effect sizes for the domestic and international behavior attributes, we turn to linear hypothesis tests. Table 2 presents the linear hypothesis tests we use to assess Hypothesis 2, that highly salient and similar choices (in our experiment, coups and protest crackdowns) will have a larger effect than choices that are lower along these two dimensions (in our experiment, standing firm on domestic policy and giving a speech). Using both the binary and six-point dependent variables, these tests show that each of the high cost similarity and salience behaviors had a statistically significantly larger effect than the moderate and low cost similarity and salience behaviors.

Table 2: Linear Hypothesis Tests for Hypothesis 2

	Binary Resolve DV	Estimate	SE	P-Val	CI Low	CI High	N	N clus.
1	Repressed = Proceeded with Speech	0.22	0.01	0.00	0.19	0.24	9302	1878
2	Repressed = Refused to Compromise on Dom. Policy	0.17	0.01	0.00	0.14	0.20	9302	1878
3	Coup = Proceeded with Speech	0.09	0.01	0.00	0.06	0.12	9302	1878
4	Coup = Refused to Compromise on Dom. Policy	0.04	0.01	0.00	0.02	0.07	9302	1878

	6-Point Resolve DV	Estimate	SE	P-Val	CI Low	CI High	N	N clus.
1	Repressed = Proceeded with Speech	0.80	0.04	0.00	0.72	0.88	9302	1878
2	Repressed = Refused to Compromise on Dom. Policy	0.59	0.04	0.00	0.51	0.67	9302	1878
3	Coup = Proceeded with Speech	0.37	0.04	0.00	0.29	0.45	9302	1878
4	Coup = Refused to Compromise on Dom. Policy	0.17	0.04	0.00	0.09	0.24	9302	1878

Note: Estimates are equal to the coefficient on the left-hand side of the equation less the coefficient on the right-hand side of the equation.

Table 3 presents the linear hypothesis tests we use to assess Conjecture 1, that some domestic choices can influence international observers' estimates of a leader's willingness to fight in an international crisis as much as past international crisis behavior. For both the binary and six-point dependent variables, these tests show that the effect of repressing protesters was statistically indistinguishable from the effect of standing firm in a prior international crisis. Table 3 also presents the linear hypothesis tests we use to assess Conjecture 2, that the cumulative effect of the domestic actions we expected to matter is as large or larger than past crisis behavior. Using both our six-point and binary DVs, we find that the cumulative effect of the domestic actions was statistically significantly larger than the effect of standing firm in a prior international crisis.

Table 3: Linear Hypothesis Tests for Conjectures 1 and 2

	Binary Resolve DV	Estimate	SE	P-Val	CI Low	CI High	N	N clus.
1	Stood Firm = Repressed	0.02	0.02	0.29	-0.01	0.05	9302	1878
2	Stood Firm = Cumulative Effect of Dom. Actions	-0.16	0.02	0.00	-0.21	-0.12	9302	1878

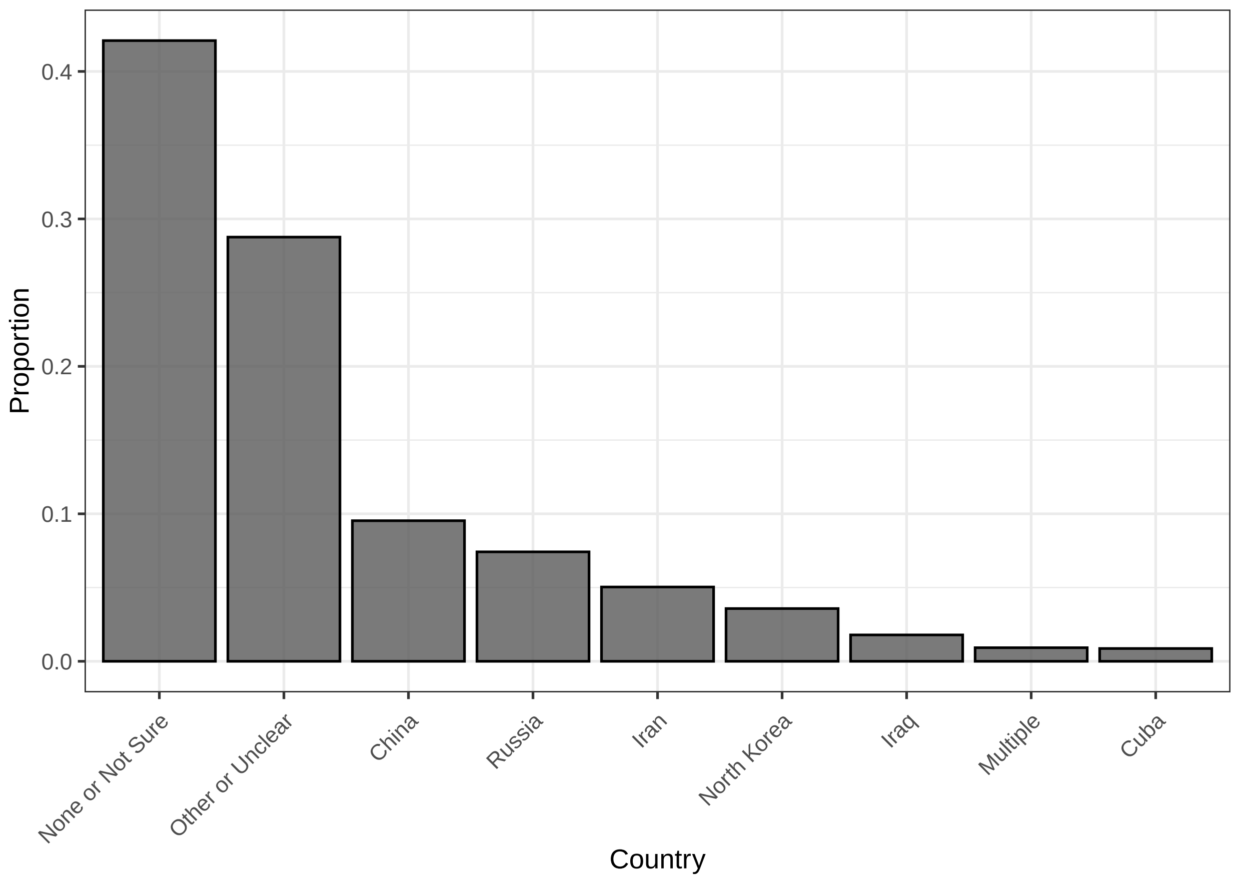
	6-Point Resolve DV	Estimate	SE	P-Val	CI Low	CI High	N	N clus.
1	Stood Firm = Repressed	-0.05	0.04	0.21	-0.13	0.03	9302	1878
2	Stood Firm = Cumulative Effect of Dom. Actions	-0.69	0.06	0.00	-0.81	-0.58	9302	1878

Note: Estimates are equal to the coefficient on the left-hand side of the equation less the coefficient on the right-hand side of the equation.

Main Study Post-Conjoint Questions

Subjects were asked two post-conjoint questions. The first, which all subjects saw, asks whether the fictional country of Arcadia described in our experiment reminded them of any real country, and if so, which. The results of this question are reported in Figure 2. The results suggest that the vignette did not overwhelmingly remind subjects of a real country, much less the same one.

Figure 2: Did Arcadia remind you of a real country?



Subjects were next randomly assigned to answer one of four questions before the survey terminated. The first examined the logic of domestic repression as an indicator of resolve. The survey asked subjects to indicate their agreement with three statements expressing distinct pathways through which repression could influence beliefs about resolve. One focused on the leader's disposition and corresponds to the logic of the theory presented in this paper. The other two focused on domestic political constraints and domestic political incentives. The results are presented in Figure 3 and show that the logic underpinning our theory received high levels of agreement.

The second randomly assigned post-conjoint question asked subjects to indicate their agreement with three statements expressing distinct pathways through which coups could influence beliefs about resolve. Again, one focused on the leader's disposition and corresponds to the logic of the theory presented in this paper. The other two focused on domestic political constraints and domestic political incentives. The results are presented in Figure 4 and, again, show that the logic underpinning our theory received high levels of agreement.

Figure 3: Logic of Domestic Repression as an Indicator of Resolve

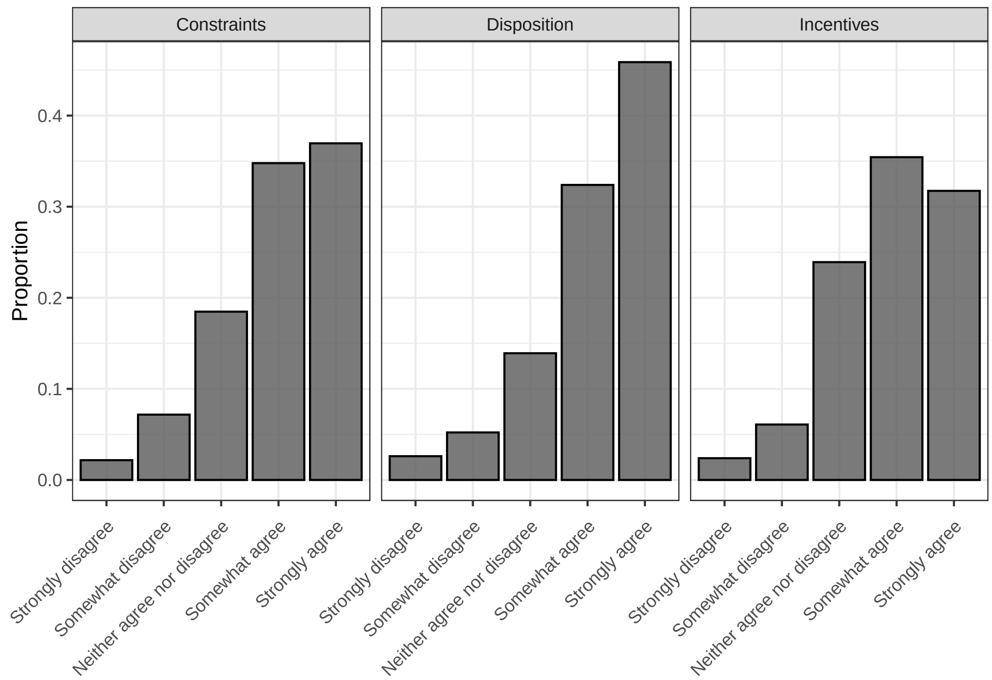
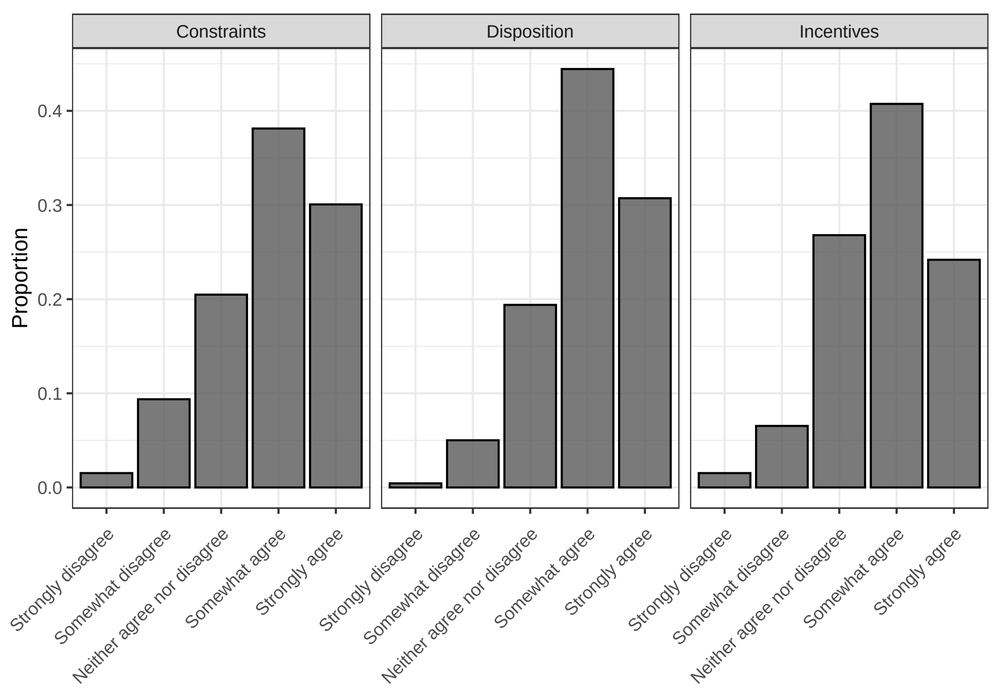
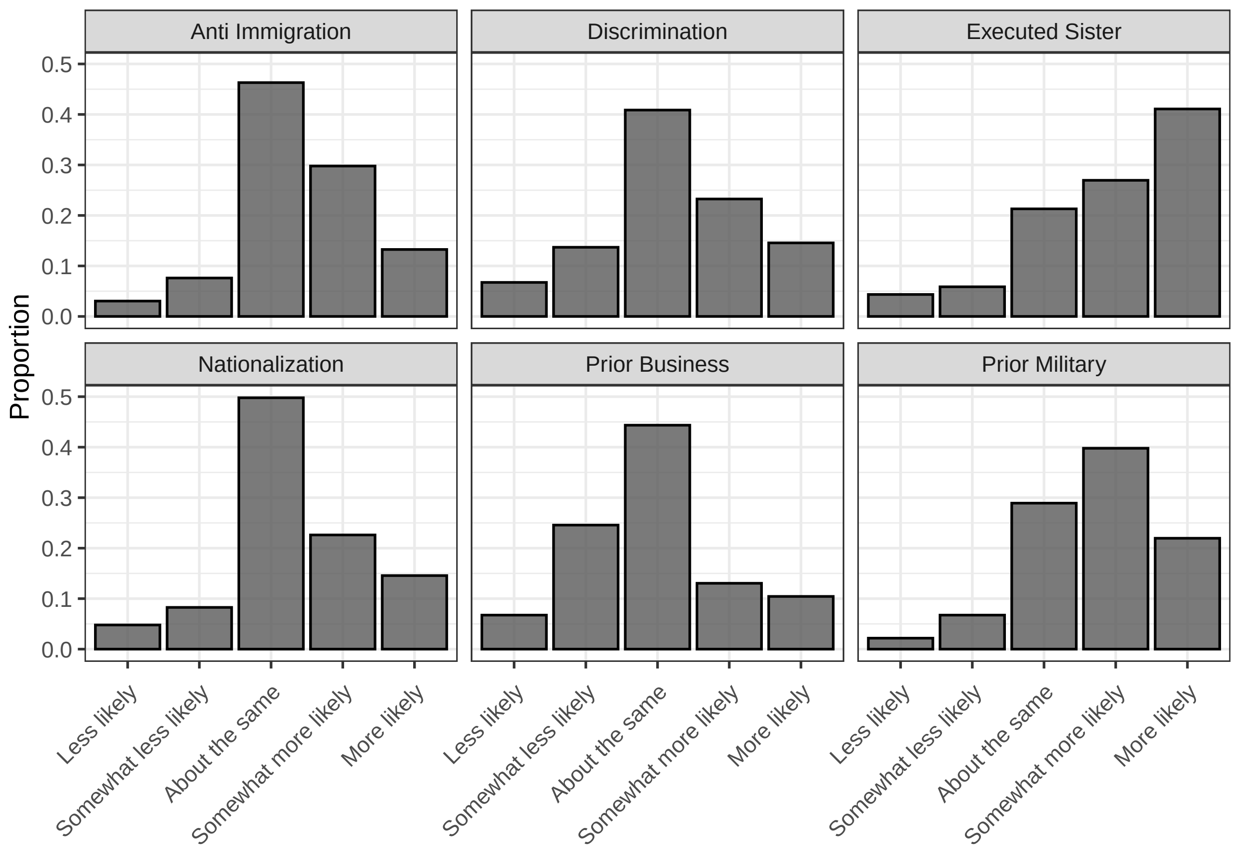


Figure 4: Logic of Coups as an Indicator of Resolve



A third randomly assigned post-conjoint question asked subjects to consider how different domestic behaviors or attributes *would have* influenced their assessments of the Arcadian leader's resolve. The behaviors were: passing anti-immigration reforms, nationalizing a private oil concern, serving two decades in the military, having business experience, rolling back anti-discrimination protections for minorities and women, and executing a sister following a personal dispute. The results, presented in Figure 5, suggest the prior military experience and executing a family member are mostly likely to have an effect on reputation for resolve.

Figure 5: Influence of Other Domestic Behaviors



A final group of respondents were randomly assigned to answer an open-ended response question. Subjects were asked: "Did information about the Arcadian leader's actions at home influence whether you thought the Arcadian leader would use military force against the U.S.?" Below are a selection of responses that reflect the logic underpinning our theory:

- "Yes, how a leader behaves at home and how much his ego controls his actions are indicators of how he will react in other situations."
- "If this leader wouldn't use police force on his /her protesters, hopefully they also wouldn't use military force against others.
- "It's like what happened in Beijing & the protest as well as how China treats Taiwan."
- "Yes, because there were other actions the leader took instead of negotiating."
- "Showed he was violent [sic] and not afraid to take on challenges"
- "Yes. He did not consider the loss of life among civilians in his country. No compassion."

Main Study Demographics

Table 4 presents demographic data for our sample in percentage terms. Full sample includes subjects recruited to the survey, regardless of whether they ultimately consented or passed the pre-treatment attention check. Attentive refers to the sub-sample that consented and passed the attention check. Census refers to the analogous percentages according to the U.S. Census.[4]

Table 4: Demographic Composition of Sample (Percent)

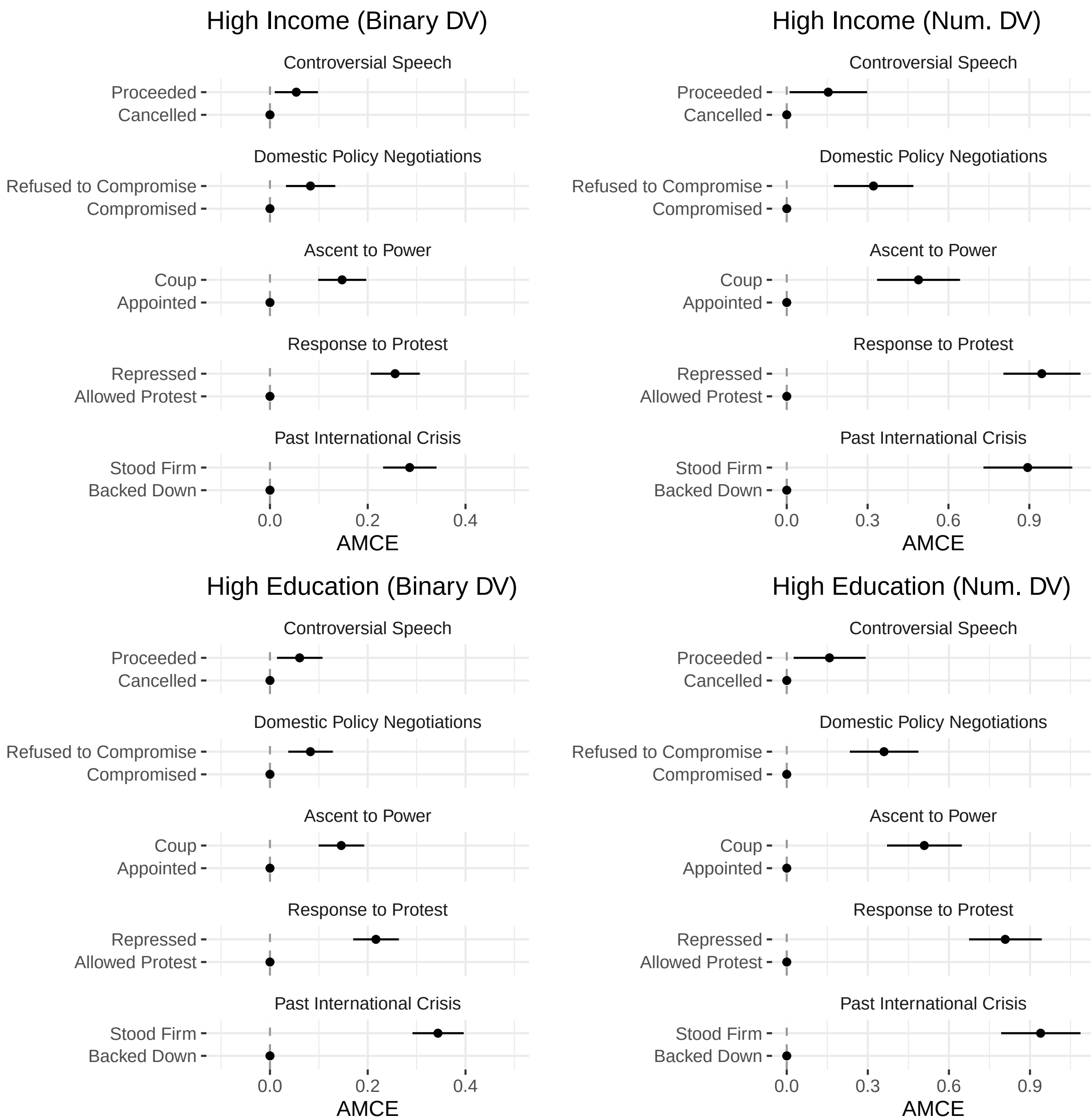
	Category	Full Sample	Attentive	Census
1	Female	52.40	51.50	50.8
2	18-19	5.30	4.30	3.4
3	20-24	12.80	9.60	8.4
4	25-29	9.70	8.30	9.1
5	30-34	11.70	10.30	8.7
6	35-39	10.70	9.90	8.5
7	40-44	8.00	8.10	7.9
8	45-49	4.80	4.80	7.9
9	50-54	6.10	6.70	8
10	55-59	8.20	9.60	8.4
11	60 and older	22.80	28.60	29.4
12	Less than $14,999	15.20	12.60	9.8
13	$15,000 to $24,999	11.00	11.10	8.3
14	$25,000 to $34,999	10.50	11.10	8.4
15	$35,000 to $49,999	11.90	12.90	11.9
16	$50,000 to $74,999	15.50	17.40	17.4
17	$75,000 to $99,999	9.20	9.60	12.8
18	$100,000 to $149,999	9.70	10.20	15.7
19	$150,000 to $199,999	3.30	3.60	7.2
20	$200,000 and above	7.30	3.40	8.5
21	Prefer not to answer	4.20	3.30	NA
22	White	67.10	72.80	76.3
23	Black	14.40	11.60	13.4
24	American Indian	1.40	1.60	1.3
25	Asian	6.30	6.00	5.9
26	Pacific Islander	0.40	0.20	0.2
27	Other	6.20	4.60	NA
28	Hispanic			18.5

[4]Census data comes from https://www.census.gov/quickfacts/fact/table/US/PST045219 and https://data.census.gov/cedsci/table?q=age&tid=ACSST1Y2019.S0101.

Results among Subjects with Elite-like Demographics (Main Study)

Figure 6 plots the average marginal component effects of our main experiment among subjects with demographic characteristics associated with elites, i.e., high incomes and educational attainment. High income is defined as greater than $100,000 per year (N = 300). High education is defined as earning a master's, doctorate, or professional degree (N = 334).

Figure 6: AMCEs among High Income and High Education Subjects



2.3 Supplementary Study Additional Tables and Figures

Supplementary Study AMCE Estimates

OLS estimates of the supplementary study Average Marginal Component Effects (AMCEs) with robust standard errors clustered at the respondent level are presented below (Table 5). The table presents estimates for model specifications using the six-point and binary outcome variables, with and without demographic controls.

Table 5: AMCEs for Domestic Politics and Resolve Conjoint

	Six-Point Outcome		Binary Outcome	
	No Controls	Controls	No Controls	Controls
Ethical CT (vs. Met Demands)	0.24***	0.22***	0.07**	0.07**
	(0.07)	(0.07)	(0.02)	(0.02)
Unethical CT (vs. Met Demands)	0.58***	0.59***	0.16***	0.16***
	(0.07)	(0.07)	(0.02)	(0.02)
Repression (vs. Allowed Protests)	0.56***	0.58***	0.15***	0.16***
	(0.06)	(0.06)	(0.02)	(0.02)
Gave Speech (vs. Cancelled)	0.07	0.07	0.03	0.03
	(0.05)	(0.05)	(0.02)	(0.02)
Inflexible Domestically (vs. Compromised)	0.18***	0.17**	0.05**	0.05*
	(0.05)	(0.05)	(0.02)	(0.02)
Fought in Past Crisis (vs. Backed Down)	0.74***	0.74***	0.25***	0.25***
	(0.06)	(0.06)	(0.02)	(0.02)
Age		−0.01*		−0.00**
		(0.00)		(0.00)
Male		0.00		0.01
		(0.08)		(0.03)
USD 14,999 or less		−0.14		−0.04
		(0.18)		(0.06)
USD 15,000-19,999		−0.11		−0.05
		(0.23)		(0.08)
USD 20,000-24,999		0.06		0.02
		(0.23)		(0.07)
USD 25,000-29,999		−0.11		−0.04
		(0.21)		(0.07)
USD 30,000-34,999		−0.16		−0.04
		(0.19)		(0.06)
USD 35,000-39,999		−0.24		−0.01
		(0.27)		(0.09)
USD 40,000-44,999		0.06		−0.01
		(0.23)		(0.08)
USD 45,000-49,999		−0.19		−0.10
		(0.22)		(0.08)
USD 50,000-54,999		0.10		0.05
		(0.23)		(0.08)
USD 55,000-59,999		−0.25		−0.08
		(0.20)		(0.07)
USD 60,000-64,999		0.36		0.07
		(0.23)		(0.08)
USD 65,000-69,999		−0.14		−0.04
		(0.52)		(0.13)
USD 70,000-74,999		0.13		0.06
		(0.25)		(0.07)
USD 75,000-79,999		−0.30		−0.18*
		(0.31)		(0.09)
USD 80,000-84,999		0.10		0.01
		(0.43)		(0.15)
USD 85,000-89,999		0.20		0.05
		(0.32)		(0.11)
USD 95,000-99,999		−0.19		−0.01
		(0.23)		(0.07)
USD 125,000-149,999		−0.09		−0.02
		(0.22)		(0.06)

	Six-Point Outcome		Binary Outcome	
	No Controls	Controls	No Controls	Controls
USD 150,000-174,999		0.58*		0.14*
		(0.24)		(0.08)
USD 175,000-199,999		−0.41		−0.04
		(0.45)		(0.13)
USD 200,000-249,999		0.17		0.05
		(0.29)		(0.09)
USD 250,000 and above		0.51*		0.11
		(0.29)		(0.09)
USD prefer no answer		0.12		0.03
		(0.21)		(0.07)
Ethnicity: Indian		−0.19		0.12
		(0.48)		(0.14)
Ethnicity: Chinese		−0.57		−0.03
		(0.49)		(0.14)
Ethnicity: Filipino		0.37		0.25*
		(0.39)		(0.08)
Ethnicity: Japanese		−0.98		−0.11
		(0.55)		(0.20)
Ethnicity: Other Asian		−1.38		−0.35
		(0.99)		(0.25)
Ethnicity: Vietnamese		−0.94*		−0.20
		(0.31)		(0.15)
Ethnicity: Black		−0.43		−0.02
		(0.29)		(0.06)
Ethnicity: Prefer not to answer		−0.09		0.01
		(0.36)		(0.12)
Ethnicity: Other		−0.09		0.06
		(0.31)		(0.07)
Ethnicity: White		−0.24		0.06
		(0.28)		(0.05)
Some HS or less		−0.22		−0.02
		(0.23)		(0.09)
High school		−0.26*		−0.03
		(0.15)		(0.05)
Some College		−0.04		0.03
		(0.16)		(0.05)
Vocational		0.28		0.08
		(0.31)		(0.09)
Bachelor's		−0.07		0.02
		(0.15)		(0.05)
Master's or professional		−0.07		−0.01
		(0.17)		(0.05)
Doctorate		0.18		0.10
		(0.31)		(0.09)
None of above (edu)		−0.85		−0.39*
		(0.63)		(0.14)
Independent		−0.15		−0.04
		(0.11)		(0.04)
Republican		−0.12		−0.01
		(0.09)		(0.03)
Intercept	2.82***	3.49**	0.29***	0.36*
	(0.09)	(0.35)	(0.03)	(0.08)
Num. obs.	2429	2429	2429	2429
N Clusters	486	486	486	486

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.1$ Robust standard errors clustered at the subject level in parentheses

Supplementary Study Linear Hypothesis Tests

As with the main study, we use linear hypothesis tests to assess the difference in effect sizes for the leader behaviors. Table 6 presents the linear hypothesis tests comparing domestic choices that are higher in cost similarity and/or salience with those that are lower. Despite the smaller sample size in the supplementary study, using both the binary and six-point dependent variables, these tests show that each of the high cost similarity and salience behaviors (protest response and unethical terrorism response) had a statistically significantly larger effect than the moderate and low cost similarity and salience behaviors (domestic bargaining and controversial speech).

Table 6: Linear Hypothesis Tests for Hypothesis 2, Supp. Study

	Binary Resolve DV	Estimate	SE	P-Val	CI Low	CI High	N	N clus.
1	Unethical CT = Proceeded with Speech	0.13	0.03	0.00	0.08	0.19	2429	486
2	Ethical CT = Proceeded with Speech	0.04	0.03	0.16	-0.02	0.10	2429	486
3	Repress Protests = Proceeded with Speech	0.12	0.03	0.00	0.07	0.18	2429	486
4	Unethical CT = Inflexible on Dom. Policy	0.11	0.03	0.00	0.06	0.17	2429	486
5	Ethical CT = Inflexible on Dom. Policy	0.02	0.03	0.45	-0.04	0.08	2429	486
6	Repress Protests = Inflexible on Dom. Policy	0.10	0.03	0.00	0.05	0.16	2429	486

	6-Point Resolve DV	Estimate	SE	P-Val	CI Low	CI High	N	N clus.
1	Unethical CT = Proceeded with Speech	0.52	0.09	0.00	0.34	0.69	2429	486
2	Ethical CT = Proceeded with Speech	0.17	0.09	0.05	0.00	0.34	2429	486
3	Repress Protests = Proceeded with Speech	0.49	0.08	0.00	0.33	0.65	2429	486
4	Unethical CT = Inflexible on Dom. Policy	0.41	0.09	0.00	0.23	0.58	2429	486
5	Ethical CT = Inflexible on Dom. Policy	0.06	0.08	0.44	-0.10	0.23	2429	486
6	Repress Protests = Inflexible on Dom. Policy	0.38	0.08	0.00	0.23	0.53	2429	486

Note: Estimates are equal to the coefficient on the left-hand side of the equation less the coefficient on the right-hand side of the equation.

Table 7 presents linear hypothesis tests comparing protest and terror response with past international crisis behavior. This test speaks to Conjecture 1, that a single domestic action could have as large a reputational impact as past crisis behavior. In contrast to the main experiment (and pilot) the effect of standing firm in a past international crisis in our supplementary study seems to have been larger than the effect of any single domestic action.

Table 7: Linear Hypothesis Tests for Conjecture 1, Supp. Study

	Binary Resolve DV	Estimate	SE	P-Val	CI Low	CI High	N	N clus.
1	Stood Firm in Past Crisis = Unethical CT	0.09	0.03	0.01	0.02	0.15	2429	486
2	Stood Firm in Past Crisis = Ethical CT	0.18	0.03	0.00	0.11	0.24	2429	486
3	Stood Firm in Past Crisis = Repress Protests	0.10	0.03	0.00	0.04	0.15	2429	486

	6-Point Resolve DV	Estimate	SE	P-Val	CI Low	CI High	N	N clus.
1	Stood Firm in Past Crisis = Unethical CT	0.16	0.09	0.07	-0.01	0.33	2429	486
2	Stood Firm in Past Crisis = Ethical CT	0.50	0.09	0.00	0.33	0.67	2429	486
3	Stood Firm in Past Crisis = Repress Protests	0.18	0.08	0.02	0.03	0.34	2429	486

Note: Estimates are equal to the coefficient on the left-hand side of the equation less the coefficient on the right-hand side of the equation.

Table 8 presents linear hypothesis tests comparing the cumulative effect of the domestic actions that had a statistically significant effect with past international crisis behavior. This speaks to Conjecture 2. As in the main experiment, the cumulative effect of standing firm in domestic situations is as large or larger than the effect of standing firm in an international crisis. When including the unethical rather than ethical response to terrorism in the analysis, the cumulative effect is larger for both our dependent variables.

Table 8: Linear Hypothesis Tests for Conjecture 1, Cumulative Domestic Effects, Supp. Study

	Binary Resolve DV	Estimate	SE	P-Val	CI Low	CI High	N	N clus.
1	Stood Firm = Dom. Actions (Unethical CT)	-0.12	0.04	0.00	-0.20	-0.04	2429	486
2	Stood Firm = Dom. Actions (Ethical CT)	-0.03	0.04	0.57	-0.11	0.06	2420	486

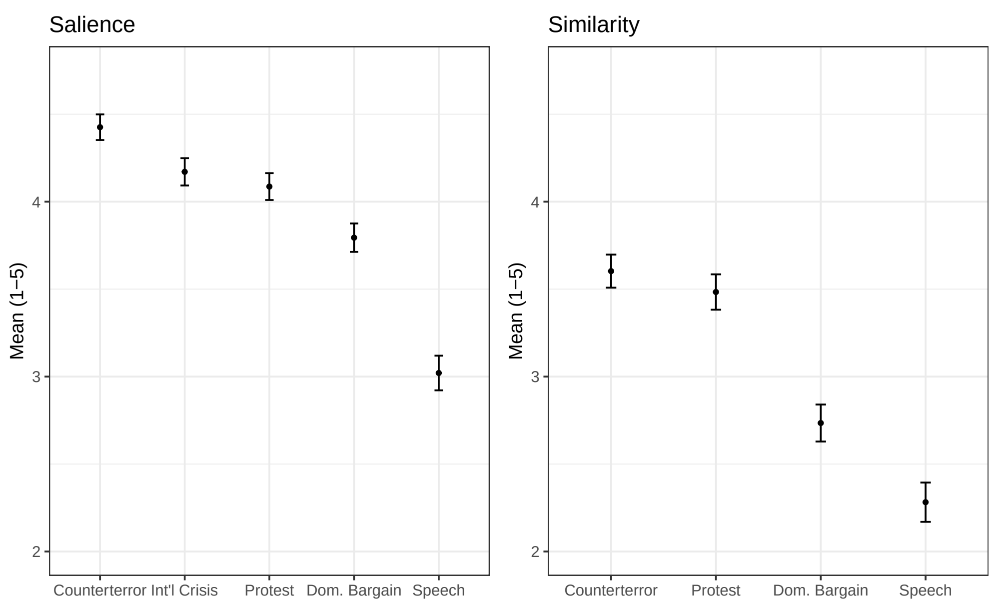
	6-Point Resolve DV	Estimate	SE	P-Val	CILow	CI High	N	N clus.
1	Stood Firm = Dom. Actions (Unethical CT)	-0.58	0.11	0.00	-0.80	-0.36	2429	486
2	Stood Firm = Dom. Actions (Ethical CT)	-0.24	0.12	0.04	-0.47	-0.01	2429	486

Note: Estimates are equal to the coefficient on the left-hand side of the equation less the coefficient on the right-hand side of the equation.

Supplementary Study Salience and Similarity Questions

Figure 7 presents the results for a post-conjoint questions about the salience and cost similarity of the choices that the hypothetical Chinese leader made in the conjoint.

Figure 7: Subject Rating of Salience and Cost Similarity of Leader Choices



Note: Following the conjoint, subjects were asked to rate the salience or 'importance' of each each conjoint attribute. Subjects could select 'Not at all important,' 'Slightly important,' 'Moderately important,' 'Very important,' or 'Extremely important.' Means calculated by translating responses to a 1-5 scale, with '1' representing 'Not at all important' and '5' representing 'Extremely important.' Following the conjoint, subjects were also asked to rate the cost similarity of standing firm in each domestic conjoint scenario to standing firm in an international crisis. Subjects could select 'Not at all similar,' 'Slightly similar,' 'Moderately similar,' 'Very similar,' or 'Extremely similar.' Means calculated by translating responses to a 1-5 scale, with '1' representing 'Not at all similar' and '5' representing 'Extremely similar.'

Crosstabs for Salience and Similarity Ratings

The tables below show the how subjects' ratings of the cost similarity and salience of different leader choices covaried in the supplementary study. Though there is a positive correlation, the tables show that it is far from perfect. This adds confidence to our theoretical claim that cost similarity and salience are and should be treated as distinct concepts.

Table 9: Similarity and Salience: Terrorism Response

	(1) Not similar	(2) Slightly similar	(3) Mod. Similar	(4) Very similar	(5) Ext. Similar
(1) Not at all important	2				
(2) Slightly important		5	7	3	1
(3) Moderately important	2	6	26	9	3
(4) Very important	5	13	50	47	16
(5) Extremely important	10	27	59	107	88

Table 10: Similarity and Salience: Protest Response

	(1) Not similar	(2) Slightly similar	(3) Mod. Similar	(4) Very similar	(5) Ext. Similar
(1) Not at all important	1		2	1	
(2) Slightly important	2	4	3	2	5
(3) Moderately important	4	13	41	24	8
(4) Very important	15	25	60	78	22
(5) Extremely important	10	16	32	54	64

Table 11: Similarity and Salience: Domestic Bargaining

	(1) Not similar	(2) Slightly similar	(3) Mod. Similar	(4) Very similar	(5) Ext. Similar
(1) Not at all important	1	1	2	1	
(2) Slightly important	6	5	14	5	2
(3) Moderately important	32	36	52	14	3
(4) Very important	38	40	59	52	7
(5) Extremely important	22	11	40	20	23

Table 12: Similarity and Salience: Controversial Speech

	(1) Not similar	(2) Slightly similar	(3) Mod. Similar	(4) Very similar	(5) Ext. Similar
(1) Not at all important	30	3	5	2	
(2) Slightly important	65	30	18	3	1
(3) Moderately important	63	42	51	20	5
(4) Very important	19	19	22	21	8
(5) Extremely important	10	3	13	18	15

2.4 Pilot Study

In August 2020 we fielded a Pilot Study ($N = 307$). Subjects were recruited through Lucid. The Pilot Study was nearly identical to the Main Study presented in the manuscript. The only major difference between the pilot and main study was that instead of the"controversial speech" attribute included in the main experiment, the pilot included a treatment about the general behavior of the leader. Leaders were randomly described as either "erratic" or "stable." The results of the pilot, presented in tabular form below, were substantively quite similar to the main experiment.

Table 13: AMCE for Domestic Politcs and Resolve Pilot Conjoint

	Outcome: Likely to Use Force (Binary)			
	Attentive	Attentive w/Controls	Full	Full w/Controls
Coup (vs. Non-violent Ascent)	0.08*	0.07*	0.08**	0.07**
	(0.03)	(0.03)	(0.03)	(0.03)
Repression (vs. Allowed Protests)	0.18***	0.18***	0.15***	0.15***
	(0.03)	(0.03)	(0.03)	(0.03)
Stable (vs. Erratic)	−0.02	−0.04	−0.01	−0.02
	(0.03)	(0.03)	(0.02)	(0.02)
Inflexible Domestically (vs. Did Compromise)	0.07*	0.06*	0.05	0.04
	(0.03)	(0.03)	(0.02)	(0.03)
Fought in Past Crisis (vs. Backed Down)	0.16***	0.15***	0.15***	0.14***
	(0.03)	(0.03)	(0.03)	(0.03)
Age		−0.00		−0.00*
		(0.00)		(0.00)
Male		−0.02		0.01
		(0.05)		(0.04)
USD 14,999 or less		−0.01		−0.08
		(0.10)		(0.08)
USD 15,000-19,999		−0.13		−0.00
		(0.12)		(0.11)
USD 20,000-24,999		−0.05		−0.09
		(0.13)		(0.11)
USD 25,000-29,999		0.08		0.10
		(0.12)		(0.10)
USD 30,000-34,999		−0.03		−0.03
		(0.14)		(0.11)
USD 35,000-39,999		−0.08		−0.04
		(0.11)		(0.09)
USD 40,000-44,999		−0.09		0.01
		(0.12)		(0.13)
USD 45,000-49,999		−0.15		−0.08
		(0.12)		(0.10)
USD 50,000-54,999		0.18		0.18
		(0.14)		(0.11)
USD 55,000-59,999		0.01		0.04
		(0.10)		(0.10)
USD 60,000-64,999		−0.02		0.07
		(0.13)		(0.14)
USD 65,000-69,999		−0.19		−0.18
		(0.19)		(0.18)
USD 70,000-74,999		−0.03		0.05
		(0.16)		(0.11)
USD 75,000-79,999		−0.49***		−0.26
		(0.09)		(0.28)
USD 80,000-84,999		−0.22		−0.08
		(0.17)		(0.14)
USD 85,000-89,999		−0.07		−0.24
		(0.29)		(0.29)
USD 90,000-94,999		0.42		0.15
		(0.18)		(0.27)
USD 95,000-99,999		−0.19		−0.11
		(0.11)		(0.12)
USD 125,000-149,999		0.03		0.10
		(0.14)		(0.11)

	Outcome: Likely to Use Force (Binary)			
	Attentive	Attentive w/Controls	Full	Full w/Controls
USD 150,000-174,999		0.31***		0.05
		(0.08)		(0.23)
USD 175,000-199,999		−0.37*		−0.07
		(0.14)		(0.15)
USD 200,000-249,999		−0.09		0.02
		(0.14)		(0.14)
USD 250,000 and above		−0.06		−0.02
		(0.14)		(0.10)
USD prefer no answer		−0.12		−0.11
		(0.13)		(0.10)
Ethnicity: Chinese		−0.07		0.05
		(0.18)		(0.16)
Ethnicity: Filipino		−0.41		−0.28
		(0.17)		(0.13)
Ethnicity: Japanese				0.08
				(0.13)
Ethnicity: Korean		0.37*		0.42**
		(0.15)		(0.14)
Ethnicity: Vietnamese		−0.53		−0.18
		(0.52)		(0.34)
Ethnicity: Other Asian		−0.37*		−0.15
		(0.15)		(0.11)
Ethnicity: Black		−0.13		−0.00
		(0.12)		(0.10)
Ethnicity: Guamanian		−0.47**		−0.30*
		(0.14)		(0.12)
Ethnicity: White		−0.22		−0.05
		(0.10)		(0.09)
Ethnicity: Other		−0.22		−0.19
		(0.13)		(0.11)
Ethnicity: Prefer no answer		−0.21		−0.06
		(0.15)		(0.12)
Some HS or less		−0.17		−0.09
		(0.16)		(0.11)
High school		0.00		0.02
		(0.08)		(0.07)
Some College		−0.08		−0.06
		(0.07)		(0.06)
Vocational		0.10		0.04
		(0.11)		(0.09)
Bachelor's		−0.01		−0.03
		(0.09)		(0.07)
Master's or professional		0.05		0.03
		(0.10)		(0.08)
Doctorate		0.07		0.07
		(0.18)		(0.11)
None of above (edu)		−0.18		0.00
		(0.16)		(0.18)
Independent		0.15*		0.03
		(0.07)		(0.05)
Republican		−0.04		−0.01
		(0.05)		(0.04)
Intercept	0.32***	0.76***	0.37***	0.58***
	(0.04)	(0.18)	(0.04)	(0.15)
Num. obs.	980	980	1535	1535
N Clusters	196	196	307	307

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$ Robust standard errors clustered at the subject level in parentheses