

Loss Aversion in Territorial Disputes: An Empirical Test

*Pre-Analysis Plan**

Andi Zhou[†] H. E. Goemans[‡] Michael Weintraub[§]

December 11, 2019

*We thank Jack Levy.

[†]Department of Government, Harvard University, azhou@g.harvard.edu.

[‡]Department of Political Science, University of Rochester, henk.goemans@rochester.edu.

[§]School of Government, Universidad de los Andes, ml.weintraub@uniandes.edu.co.

Contents

1	Introduction	3
1.1	Motivation	4
2	Loss Aversion in IR	5
3	Theory, Method and Hypotheses	7
3.1	Theory	7
3.2	Method	11
3.3	Survey Questions and Hypotheses	11
3.4	Secondary hypotheses: dispositional and demographic heterogeneity	17
3.4.1	Social and international trust	17
3.4.2	Authoritarianism	18
3.4.3	Social dominance orientation	19
3.4.4	Demographic heterogeneity	19
4	Empirical analysis	20

1 Introduction

Since Daniel Kahneman and Amos Tversky introduced the idea that humans treat risk differently depending on whether they consider themselves ahead in the game or behind (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), international relations scholars have theorized extensively how the theory – known as *cumulative prospect theory* – can be applied to a variety of phenomena in world politics (Levy, 1997, 2003; McDermott, 2004). In the light of this sustained theoretical interest, there is surprisingly little systematic empirical evidence whether, when, and how prospect theory affects international politics or foreign policy. Without the opportunity to manipulate how a situation is framed in the minds of policymakers, citizens, and voters, researchers are ill-equipped to make robust claims about the impact of so-called framing effects in observational studies of real-world events.

To systematically evaluate the potential of prospect theory in helping us understand international politics, we explore territorial conflict. We employ experimental methods to measure framing effects on citizens' preferences in the context of territorial bargaining. Specifically, we field survey experiments on nationally representative samples in Chile and Argentina in which we manipulate the framing of a prompt about disputed territorial claims in Antarctica. We then measure respondents' propensity to choose between either a certain gain/loss or a lottery with the same expected payoff. In doing so, we introduce two major innovations into the application of prospect theory in international relations: 1) we run an experiment in a setting of an actual territorial dispute, and 2) we introduce an opponent, an essential element in international politics but missing from canonical research of prospect theory.

The evidence generated by this experiment can shed light not only on the psychological microfoundations of territorial contestation but also on broader questions of bargaining and international conflict.¹

¹In terms of the bargaining theory of war, risk tolerant leaders require a premium in bargaining, which can make war rational for completely-informed actors, even in the absence of commitment problems (Fearon, 1995).

1.1 Motivation

At its core, prospect theory hinges on the insight that humans feel the pain of loss more strongly than the joys of gain. This asymmetry between losses and gains means that actors assess outcomes in comparison to a “reference point,” often conceived of as the prior status quo. Ending up with \$50, for example, feels very different if one begins with \$100 than if one begins with nothing. Risk preferences, in particular, have been shown to be highly reference-dependent. Given a choice between a certain gain (e.g. \$50) and a lottery with the same expected payoff (e.g. flipping a coin with a \$100 gain for heads and zero gain for tails), people are more likely to choose the certain gain. When choosing between a certain loss and a lottery with the same expected negative payoff, however, people are more likely to choose the lottery. People are risk-averse when striving for gains, but more willing to take on risk when avoiding or recovering losses.

More importantly, experimental evidence has found that framing a set of outcomes in loss terms or gain terms can change subjects’ reference points, producing inconsistent preferences over the same set of outcomes, a violation of standard rational choice assumptions. Taking the same coin flip example, imagine that person A is asked to choose between accepting \$50 for sure or flipping a coin for \$100, and person B is handed \$100, and then asked to choose between giving *back* \$50 or flipping a coin to keep or forfeit all of it. In both scenarios, the outcome is the same: the person is choosing between getting \$50 with certainty or getting \$100 with 50 percent probability. Person A, however, is much more likely to accept the \$50 with certainty, while person B is much more likely to take the coin flip. The framing of the outcomes determines their reference points: person A’s reference point remains at \$0, so a certain outcome of \$50 feels like a gain; person B’s reference point, however, shifted to \$100 as soon as she was handed the money at the beginning of the game, so a certain outcome of \$50 feels like a loss.²

Tversky and Kahneman (1992) developed the original theory into *cumulative* prospect

²This is a paraphrasing of an experiment first presented in Kahneman and Tversky (1979, 273).

theory, which hinges on four core concepts: “1) reference dependence, 2) loss aversion, 3) diminishing sensitivity, and 4) probability weighting” (Barberis, 2013, 175). Cumulative prospect theory posits that people do not simply derive utility from wealth in a linearly increasing value function. Instead, a reference point divides the value function into two regions: the domain of losses, represented by a convex value function below the reference point, and the domain of gains, represented by a concave value function. Cumulative prospect theory showed that people are “much more sensitive to losses – even small losses – than to gains of the same magnitude” (Barberis, 2013, 175). This phenomenon, known as *loss aversion*, is reflected by positing a value function steeper in the domain of losses than in the domain of gain. Finally, “people do not weight outcomes by their objective probabilities . . . but rather by transformed probabilities or decision weights . . . [which are] applied to cumulative probabilities. . . . [This] leads the individual to overweight the *tails* of any distribution” (Barberis, 2013, 176).³ Below, we build on these core concepts to argue that people should be susceptible to similar cognitive biases when evaluating territorial disputes. In particular, loss framing will tend to set “extreme” reference points corresponding to their country’s farthest bounds of territorial aspiration, regardless of the actual status quo of territorial control on the ground (Kőszegi and Rabin, 2006), and the salience of an opponent could both shift the reference point and augment the slope of the value function, the latter leading to stronger framing effects (Tversky and Kahneman, 1992).

2 Loss Aversion in IR

IR scholars have noted the potential for loss-averse behavior to explain patterns of territorial conflict. Drawing on insights from evolutionary biology, Johnson and Toft (2014, 32) argue that “agents who are (or perceive themselves to be) territory owners will act like hawks—a strategy that may have been generally adaptive in the past but is often coun-

³As Levy (1997) puts it, “This means that except for small probabilities people tend to give more weight to the utility of a possible outcome than to its probability of occurrence.”

terproductive today, especially where perceptions of prior residency have been manipulated or blurred in history.” They point out that it is possible for more than one disputant to identify as the “owner” over a piece of territory, with ominous consequences: “[w]here both sides perceive themselves to be the territory[’s] resident, the problem looms especially large because each side may expect to win and expect the other side to back down, somewhat regardless of size and strength. Manipulation experiments in biology show that when two individuals are tricked into believing a particular territory belongs to both of them, fighting can be especially intense.”⁴

This latter circumstance—a dispute between parties who all believe the territory in question already belongs to them—appears to be the default mode of territorial disputation in the international system today. The rise of territorial integrity as a norm has delegitimized territorial “gain” or “expansion” as a valid foreign policy objective (Zacher, 2001). Thus, states that seek to contest a piece of territory face powerful incentives to justify their efforts in terms of territorial defense or recovery to rally both domestic and international support, regardless of their “true” motivations (if such a thing can even be identified).

Although some IR scholars have incorporated basic insights of prospect theory, no systematic evidence has been brought to bear to illustrate empirically its causal logic (Levy, 1997). There are compelling reasons to believe that loss aversion underlies much of the territorial conflict dynamics we observe today, yet we cannot assume that the voluminous evidence for loss aversion collected in psychology and behavioral economics will extrapolate cleanly to territory. One reason is that canonical experimental tests of prospect theory are all missing an *opponent*. Territorial disputes involve multiple actors by definition, and so one actor’s decisions about a territorial dispute will naturally take into consideration its relationships with the other actors involved.

⁴See also Levy (1997, 93).

3 Theory, Method and Hypotheses

We briefly sketch our theoretical innovation, how we plan to test our main ideas, and the specific hypotheses.

3.1 Theory

Building on the core concepts of cumulative prospect theory briefly outlined above, we explore how the setting of an existing territorial dispute and the introduction of a hostile opponent affects choice.

We argue that both the framing and the setting of an actual territorial dispute will affect the location of the reference point. As Tversky and Kahneman (1981) demonstrated, framing an outcome as a gain or as a loss changes the reference point against which someone evaluates that outcome, giving rise to inconsistent preferences when the same outcomes are presented in gain terms or loss terms. In particular, preferences tend to be risk-averse in gain frame but risk-tolerant in loss frame. In the context of a territorial dispute, a disputed territory could be perceived either as territory to be gained beyond the state's frontier of undisputed control, or as *claimed* territory at risk of being lost or which was previously lost and ought to be regained. The former implies the on-the-ground status quo of territorial control as the reference point, while the latter implies the claim as the reference point, which we term the "aspirational" boundary. A territorial settlement framed as a gain would be evaluated against the status quo, whereas the same settlement framed as a loss would be evaluated against the aspiration.

A basic presentation of territorial conflict would locate the reference point at the status quo. However, Kahneman and Tversky (1979, 286, emphasis added) recognized that "there are situations in which gains and losses are coded relative to an *expectation or aspiration level* that differ from the status quo." Köszegi and Rabin (2006, 1134) propose that "a person's reference point is the probabilistic beliefs she held in the recent past about outcomes." This

perspective helps explain so-called endowment effects (Thaler, 1980), which Köszegi and Rabin (2006, 1142) interpret as the “disposition of subjects to believe that their current ownership status is indicative of their ensuing ownership status.” Mercer (2005, 5-6) argues that the potential of actors’ reference points to be set in an aspirational way is particularly relevant in international politics and specifically refers to North Korea’s claims over South Korea. In this sense, reference points may be forward-looking: rather than corresponding to what one possesses today, it may instead reflect what one aspires to or expects to attain in the future.

Reference points, however, can also be backward-looking, reflecting what one possessed in the past, or more importantly, what one *believes* oneself to have possessed in the past. Tversky and Kahneman (1981) gives the example of a gambler at a race track who has lost \$140 over the course of the afternoon and is considering a final long-shot bet to recoup the losses. Such backward-looking reference points are especially important for the study of territorial conflict, where irredentist narratives of “lost” territories and “historical” claims are ubiquitous. North Korea’s claim over South Korea may be an aspiration for the future, but the North Korean regime legitimizes the claim as a return to the past of a united Korea. Saddam Hussein similarly declared Kuwait a “lost province” of Iraq before invading in 1990, drawing on a longstanding narrative prevalent in Iraq that Kuwait’s independence was an accident of colonial history (Millett and Maslowski, 2012; MacFarquhar, 2002). Fang and Li (2019) find that presenting China’s territorial claims in historical terms elicits less willingness to compromise among Chinese survey respondents. In international politics, our expectations for the future are rooted in our understanding of the past. We therefore expect loss-framing to be a powerful way for states and other territorial identity entrepreneurs to shift citizens’ reference points past the status quo and out toward an aspirational boundary. Through sustained education, propaganda, and indoctrination of both overt and subtle varieties, states inculcate citizens with narratives and myths of loss that consolidate reference points around the claim rather than the status quo.

Additionally, our experimental design tests the impact of another salient feature of territorial disputes: the opponent. Raising the specter of a territorial opponent changes the scenario from a simple choice between a certain gain/loss and a gamble to a strategic interaction. To our knowledge, no experimental evidence exists on how the presence or identity of an opponent in a strategic interaction might impact reference-dependent risk preferences. This is striking, given that our ability to import prospect theory’s insights to international relations is limited if we cannot account for the *relations* part of international relations. In our experimental design, we test for interaction effects between opponent salience and frame manipulation. We anticipate that an opponent may affect reference-dependent preferences in two ways. First is through the slope of the value function, which we call the “s-curve” mechanism. Respondents may register the opponent’s loss as an additional gain and the opponent’s gain as an additional loss, steepening the slope of the utility function on both sides of the reference point and thus increasing the difference in utility between certain outcomes and lotteries with equivalent expected outcomes. We should then expect to see a wider gap between gain-frame and loss-frame responses.

Second, identifying an opponent may shift the reference point itself from the status quo to the aspirational boundary, effectively serving as an alternative mechanism for loss-framing. We call this the “reference-point” mechanism. When the opponent is viewed as threatening, then the territorial dispute becomes a matter of territorial *defense*—in particular, defense of the *claim* against the encroachment or usurpation of the opponent. The aspirational boundary thus becomes the reference point. When respondents are presented with such an opponent, their responses should become more risk-tolerant in both gain- and loss-frame conditions. The strength of the reference point shift should be proportional to the degree to which the opponent is perceived to be a threat. In our experimental design, we are able to vary the identity of the opponent to test for this effect. If the shift due to hostility against the opponent is strong enough, ceiling effects may even narrow the gap between gain- and loss-frame responses.

Since these two mechanisms are not mutually exclusive, we can arrange our predictions on the two-by-two table shown in Table 1 below, which has the reference-point effect on the horizontal axis and the S-curve effect on the vertical axis.⁵ If both effects are weak, responses would look no different than if no opponent is identified in the scenario. If the S-curve effect dominates, then responses will be more strongly risk-averse in gain frame and more strongly risk-tolerant in loss frame. If the reference point effect dominates, then responses in both gain and loss frame will become more risk-tolerant, but the gap between gain- and loss-frame responses will stay constant (unless the shift is strong enough that ceiling effects begin to close the gap). Finally, if both effects are strong, then responses should be strongly and similarly risk-tolerant in both gain- and loss-frame conditions.

Table 1: Predictions of the opponent effect on risk preferences

		Reference-point effect	
		Weak	Strong
S-curve effect	Weak	No difference from no-opponent conditions	Increased risk tolerance in both gain and loss frame
	Strong	Amplified framing effect (more risk averse in gain frame and risk tolerant in loss frame)	Strongly and similarly risk tolerant in both gain and loss frame

In the context of territorial disputes, whether the status quo or the aspiration serves as the reference point for a disputant becomes a question of paramount importance. This empirical question has important implications for the behavior of states in a territorial dispute and for the application of the typical bargaining model, where actors are risk-neutral or risk averse (Fearon, 1995). Butler (2007) demonstrates formally the implications of different reference points for models of coercive bargaining, showing that when actors adopt “extreme” reference points (such as those that conform to territorial aspirations), the conflict space becomes very wide, barring extremely high costs of conflict. This study is a first step toward testing these

⁵We thank Dustin Tingley for suggesting this framework for our hypotheses.

formal implications empirically.

3.2 Method

We plan to administer the survey experiment in two different contexts: in Chile and in Argentina, both on nationally representative samples of approximately 3,000 respondents. Chile and Argentina are ideal places to administer this survey because of their competing territorial claims in Antarctica. Because so many disputed territories around the world are highly politically salient, it is difficult to find a case where framing the dispute in terms of territory to be gained (rather than territory already possessed “to be lost”) is credible for the population of a disputant state. The Antarctic claims are, however, well-suited for this role. On the one hand, Antarctica’s remoteness, lack of permanent population, and uneventful political history (largely owing to an international treaty that set Antarctica aside as a scientific preserve (Teller, 2014)) should keep its salience low in the public consciousness of both countries. On the other hand, both Chilean and Argentine law mandate official maps to depict the Antarctic claim as part of their respective national territories, so public awareness of the issue should be broad enough that presenting the issue as a political question should be credible.

Our experiment is a 2-by-3 factorial design, with gain and loss frame conditions on one dimension and three conditions on the other that vary the salience and identity of the opponent. In each of the six conditions, respondents will be presented with a brief prompt introducing the territorial dispute and then are asked to choose between two policy options: accepting a certain gain or loss, or pursuing a lottery.⁶

3.3 Survey Questions and Hypotheses

In the first of the three opponent conditions, we do not mention any specific opponent in either the prompt or the policy options. Respondents in all experimental conditions receive

⁶We adopted language to mirror as closely as possible the language in Tversky and Kahneman (1981).

the same initial prompt:

As you may know, a large piece of territory in Antarctica is disputed between Chile/Argentina and several other countries. Imagine that the following two options are being considered to resolve the dispute:

Respondents in the gain frame condition see the following two options:

- A. Accept an agreement in which Chile/Argentina obtains 1/3 of the disputed territory in Antarctica.*
- B. Pursue a policy where Chile/Argentina has a 1/3 chance of obtaining the entire disputed territory in Antarctica and a 2/3 chance of obtaining none of this territory.*

Which option do you prefer?

Loss-frame respondents see the following options instead:

- A. Accept an agreement in which Chile/Argentina loses 2/3 of the disputed territory in Antarctica.*
- B. Pursue a policy where Chile/Argentina has a 1/3 chance of keeping the entire disputed territory in Antarctica and a 2/3 chance of losing all of this territory.*

Which option do you prefer?

As in Kahneman and Tversky's original experiments, the option pairs in our experiment yield the same expected utility, but one pair presents the choice in terms of "obtaining"

territory, while the other casts the choice in terms of “losing” territory. We expect the predictions of prospect theory to hold: respondents given the gain-frame version of the question (“obtaining” territory) will be more likely to choose option A, and those given the loss-frame version (“losing” territory) will more likely choose option B. The obvious null is no framing effects.

H1: Respondents in the loss frame will be more likely to choose the gamble than respondents in the gain frame.

We will test *H1* in the following linear probability framework.

$$\Pr(Y_i = B) = \beta_0 + \beta_1 T_i + \gamma \mathbf{X}_i + \epsilon_i \quad (1)$$

where Y_i is the respondent’s policy choice—here the choice for the gamble— T_i is a binary treatment variable taking a value of 1 in loss frame and 0 in gain frame, and \mathbf{X}_i is a vector of demographic and dispositional covariates.

To more carefully pin down any effects, we introduce two auxiliary hypotheses here for later reference. For loss aversion to hold, only *Auxiliary H1b* needs to find support.

Auxiliary H1a: Respondents exposed to the gain frame will be more likely to choose the certain gain/loss option over the lottery.

Auxiliary H1b: Respondents exposed to the loss frame will be more likely to choose the lottery over the certain gain/loss option.

In the second condition, we introduce the dyadic opponent: Argentina when the survey is run in Chile, and Chile when the survey is run in Argentina. The initial prompt is nearly identical, save the mentions of the opponent:

*As you may know, a large piece of territory on Antarctica is disputed between Chile **and Argentina** /Argentina **and Chile**. Imagine that the following two options are being considered to resolve the dispute:*

- A. Accept an agreement in which Chile/Argentina obtains 1/3 of the territory **disputed with Argentina/Chile in Antarctica** [loses 2/3 of the territory **disputed with Argentina/Chile in Antarctica**].
- B. Pursue a policy where Chile/Argentina has a 1/3 chance of obtaining the entire territory **disputed with Argentina/Chile in Antarctica** and a 2/3 chance of obtaining none of this territory [1/3 chance of keeping the entire territory **disputed with Argentina/Chile in Antarctica** and a 2/3 chance of losing all of this territory].

Which option do you prefer?

In section 3.1 above, we briefly sketched out two logics for how the introduction of a dyadic opponent could affect choice. Following the first logic, the value function becomes a steeper S-curve and by the second logic the reference point becomes extreme. The S-curve logic laid out above proposes that framing effects should become *augmented*. This logic thus again proposes *H1* and its auxiliary hypotheses. In addition, the S-curve logic proposes that this effect should be stronger when an opponent is introduced when compared to the absence of such an opponent.

S-curve H2: The framing effect of *H1* is stronger for respondents in the dyadic opponent condition than for respondents in the non-opponent condition.

The argument that the introduction of a dyadic opponent shifts the reference point leads to a different hypotheses. We expect that introducing an opponent would shift an additional number of respondents' reference points toward the aspirational boundary in both gain and loss frames, raising the proportion of risk-tolerant responses in both groups without necessarily changing the difference between gain- and loss-frame responses. (If the reference-point shift is strong enough, or if loss-frame responses are already overwhelmingly risk-tolerant responses, then the gap between gain- and loss-frame responses may close due to ceiling effects in the loss-frame condition.)

Reference H3: Among both gain-frame and loss-frame respondents, those primed with an opponent will be more likely to choose the lottery over the certain gain/loss than respondents who were not primed with an opponent.

Auxiliary H3a: In the opponent condition, respondents in loss frame will be more likely to choose the lottery over the certain gain/loss than respondents in gain frame.

Strong Auxiliary H3b: In the opponent condition, respondents in the gain frame will prefer the lottery over the certain gain/loss.

If both S-curve and reference effects are present, then responses will be both more risk-tolerant and more similar across gain and loss frames, since reference points will be shifted toward the aspirational boundary *and* loss aversion will be intensified. In other words, respondents in the gain frame will not choose differently than respondents in the loss frame.

Combined H4: When respondents are presented with an opponent, respondents will strongly prefer the lottery over the certain gain/loss in both gain and loss frame, and the framing effect of *H1* will be attenuated.

In the third condition, we introduce the United Kingdom as the opponent in both surveys, since the United Kingdom's Antarctic claims overlap with those of both Chile and Argentina. The fact that we are running parallel surveys in Chile and Argentina affords a valuable opportunity to test the effect of historical rivalry on risk preferences in territorial bargaining. Argentina's defeat in the 1982 Falklands War against the United Kingdom remains politically salient in Argentine public discourse, whereas Chile has not had a similar history of hostility against the United Kingdom. The wording of the prompt and policy options is as follows:

As you may know, a large piece of territory on Antarctica is disputed between Chile/Argentina **and the United Kingdom**. Imagine that the following two options are being considered to resolve the dispute:

- A. Accept an agreement in which Chile/Argentina obtains 1/3 of the territory **disputed with the United Kingdom in Antarctica** [loses 2/3 of the territory **disputed with the United Kingdom in Antarctica**].
- B. Pursue a policy where Chile/Argentina has a 1/3 chance of obtaining the entire territory **disputed with the United Kingdom in Antarctica** and a 2/3 chance of obtaining none of this territory [1/3 chance of keeping the entire territory **disputed with the United Kingdom in Antarctica** and a 2/3 chance of losing all of this territory].

Which option do you prefer?

We expect that Chilean respondents will respond to this treatment similarly to the second condition in which Argentina was the opponent. In other words, we expect support for either *S-curve H2* or *Reference H3*, *Auxiliary H3a*, and *H3b*. We can examine how the experience of the 1982 Falklands war, and the salience of the United Kingdom as an opponent, affects choices in Argentina. Loosely speaking, if *S-curve H2* holds in Chile, we expect it to hold even more strongly in Argentina. (There may, however, be ceiling effects.)

S-curve H5: The amplified framing effect of *H2* will be stronger in Argentina than in Chile when the opponent is identified as the United Kingdom.

Mutatis mutandis, if the introduction of a dyadic opponent with a relatively recent history of armed conflict shifts the reference point, we expect support for *Reference H3* and *Auxiliary H3a* and *H3b* to be stronger in Argentina than in Chile (though, again, there may be ceiling effects in the loss-frame condition).

Reference H6: The reference point shifting effect of *H3* will be stronger in Argentina than in Chile when the opponent is identified as the United Kingdom.

To more precisely specify how the postulated extreme shift in the reference point in the third condition affects choices, we propose two auxiliary hypotheses.

Auxiliary H6a: Respondents in the gain frame in Argentina will be even less likely to choose the certain option than respondents in the gain frame in Chile.

Auxiliary H6b: Respondents in the loss frame in Argentina will be even more likely to choose the lottery than respondents in the loss frame in Chile.

Finally, we consider the combined effects of the two mechanisms and how they differ across Argentina and Chile in the last opponent condition.

Combined H7: When the opponent is identified as the United Kingdom, respondents in Argentina will prefer the lottery over the certain option in both gain and loss frame, and the preference will be stronger than in Chile.

3.4 Secondary hypotheses: dispositional and demographic heterogeneity

Both surveys also include measures of four dispositional traits to test for heterogeneous treatment effects: social trust, international trust, authoritarianism, and social dominance orientation. The secondary hypotheses related to these dispositional measures are as follows:

3.4.1 Social and international trust

We expect respondents with low levels of social trust and international trust to react more strongly to opponent-priming. Under the S-curve logic, low social and international trust individuals should value gains at the expense of an opponent more greatly and feel losses to an opponent more acutely than high-trust individuals do. Under the reference point-shifting argument, priming with an opponent should shift low-trust individuals' reference

points toward the aspirational boundary more effectively than for high-trust individuals, since low-trust individuals should more readily ascribe malign intentions to an opponent and seek to defend their country's claims against that opponent.

S-curve H8: Low social trust individuals will display stronger *H2* effects than high social trust individuals.

Reference H9: Low social trust individuals will display stronger *Reference H3* and *Auxiliary H3* effects than high social trust individuals.

Combined H10: Low social trust individuals will display stronger *Combined H4* effects than high social trust individuals.

S-curve H11: Low international trust individuals will display stronger *H2* effects than high social trust individuals.

Reference H12: Low international trust individuals will display stronger *Reference H3* and *Auxiliary H3* effects than high international trust individuals.

Combined H13: Low international trust individuals will display stronger *Combined H4* effects than high international trust individuals.

3.4.2 Authoritarianism

We expect high-authoritarianism individuals to be less easily manipulated into gain-frame than low-authoritarianism individuals. High-authoritarianism individuals should be less likely to discount or question the legitimacy of their country's territorial claims, and thus should be more prone to treating the aspirational boundary as their reference point regardless

of the framing language. We also expect them to react more strongly to an opponent than low-authoritarianism individuals, since high-authoritarianism individuals would view territorial threats from an opponent to be a challenge to their own state's authority.

H14: Individuals displaying higher levels of authoritarianism will be more risk-tolerant in gain-frame than low-authoritarianism individuals and will display weaker *H1* effects than low-authoritarianism individuals.

H15: Individuals displaying higher levels of authoritarianism will display stronger *Reference H3* and *Auxiliary H3* effects than low-authoritarianism individuals.

3.4.3 Social dominance orientation

We expect high social dominance orientation (SDO) individuals to display stronger reference-point shifts due to priming with an opponent than low-SDO individuals due to their beliefs in in-group superiority. High-SDO respondents should be more likely to believe that their country can and should defend its territorial claims against opponents, and thus be more likely to have their reference points shifted to the aspirational boundary.

H16: Individuals displaying higher levels of social dominance orientation (SDO) will display stronger *Reference H3* and *Auxiliary H3* effects than low-SDO individuals.

3.4.4 Demographic heterogeneity

Finally, we anticipate two potential sources of heterogeneity in gender and partisan identification. We expect men to choose the lottery over the certain option at higher rates than women overall, in accordance with a well-established literature showing that men tend to be more risk-tolerant than women (e.g. Croson and Gneezy, 2009). We also anticipate that

men will react more strongly than women to the opponent conditions, especially when the opponent is hostile given the logic of male coalitional aggression (Lopez et al., 2011).

H17: Men will choose the lottery over the certain option at higher rates than women in both the gain and loss frames.

H18: Men will display stronger *H2*, *H3*, and *H4* effects than women when primed with an opponent, especially a hostile one.

We expect right-wing and right-leaning partisans to choose the lottery over the certain option more frequently than left-wing and left-leaning partisans across both gain and loss frames, and we expect the increase in risk tolerance to be greater in gain frame, in accordance with our hypothesis for high-authoritarianism individuals. We also expect right partisans to react more strongly to the opponent conditions, in accordance with high-SDO and low-social and international trust individuals.

H19: Right-wing and right-leaning partisans will be more risk-tolerant in both the gain and loss frames, show smaller *H1* effects, and show stronger *H2*, *H3*, and *H4* effects than left-wing and left-leaning partisans.

4 Empirical analysis

To test the full suite of secondary hypotheses, we plan to estimate the following linear probability model:

$$\Pr(Y_i = B) = \beta_0 + \beta_1 T_i + \beta_2 Neighbor_i + \beta_3 T_i * Neighbor_i + \beta_4 UK_i + \beta_5 T_i * UK_i + \gamma^T \mathbf{X}_i + \delta^T \mathbf{X}_i * T_i + \epsilon_i \quad (2)$$

where Y_i is the respondent's policy choice, T_i is a binary treatment variable taking value 1 in the loss frame and 0 in gain frame, $Neighbor$ is a binary variable taking value 1 when

either Argentina or Chile is presented as the opponent, UK is a binary variable taking value 1 when the United Kingdom is presented as the opponent, and \mathbf{X}_i is a vector of demographic and dispositional covariates.

References

- Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. *The Journal of Economic Perspectives* 27(1), 173–195.
- Butler, C. K. (2007, April). Prospect Theory and Coercive Bargaining. *Journal of Conflict Resolution* 51(2), 227–250.
- Croson, R. and U. Gneezy (2009, May). Gender Differences in Preferences. *Journal of Economic Literature* 47(2), 448–474.
- Fang, S. and X. Li (2019, August). Historical Ownership and Territorial Disputes. *The Journal of Politics*, 706047.
- Fearon, J. D. (1995). Rationalist explanations for war. *International Organization* 49(3), 379–414.
- Johnson, D. D. and M. D. Toft (2014, January). Grounds for War: The Evolution of Territorial Conflict. *International Security* 38(3), 7–38.
- Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47(2), 263–291.
- Köszegi, B. and M. Rabin (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics* 121(4), 1133–1165.
- Levy, J. S. (1997). Prospect theory, rational choice, and international relations. *International Studies Quarterly* 41(1), 87–112.
- Levy, J. S. (2003). Applications of Prospect Theory to Political Science. *Synthese* 135(2), 215–241.
- Lopez, A. C., R. McDermott, and M. B. Petersen (2011). States in Mind: Evolution, Coalitional Psychology, and International Politics. *International Security* 36(2), 48–83.

- MacFarquhar, N. (2002, March). Mideast Turmoil: Baghdad; Baghdad-Kuwait Accord – Support Is Rebuff to Bush’s Efforts. *The New York Times*.
- McDermott, R. (2004). Prospect Theory in Political Science: Gains and Losses from the First Decade. *Political Psychology* 25(2), 289–312.
- Mercer, J. (2005). Prospect theory and political science. *Annual Review of Political Science* 8(1), 1–21.
- Millett, A. R. and P. Maslowski (2012). *For the Common Defense*. Simon and Schuster.
- Teller, M. (2014, June). Why everyone wants a piece of Antarctica. *BBC News*.
- Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization* 1, 39. Last updated - 2013-02-24.
- Tversky, A. and D. Kahneman (1981, January). The Framing of Decisions and the Psychology of Choice. *Science* 211.
- Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5(4), 297–323.
- Zacher, M. W. (2001). The Territorial Integrity Norm: International Boundaries and the Use of Force. *International Organization* 55(2), 215–250.