DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

Bilateral Conflict: An Experimental Study of Strategic Effectiveness and Equilibrium

Thomas R. Palfrey
California Institute of Technology

Charles A. Holt
University of Virginia

SOCIAL SCIENCE WORKING PAPER 1459

January 2022

Bilateral Conflict: An Experimental Study of Strategic Effectiveness and Equilibrium

Charles A. Holt and Thomas R. Palfrey*

January 17, 2022

Abstract

Bilateral conflict involves an attacker with several alternative attack methods and a defender who can take various actions to better respond to different types of attack. These situations have wide applicability to political, legal, and economic disputes, but are particularly challenging to study empirically because the payoffs are unknown. Moreover, each party has an incentive to behave unpredictably, so theoretical predictions are stochastic. This paper reports results of an experiment where the details of the environment are tightly controlled. The results sharply contradict the Nash equilibrium predictions about how the two parties' choice frequencies change in response to the relative effectiveness of alternative attack strategies. In contrast, nonparametric quantal response equilibrium predictions match the observed treatment effects. Estimation of the experimentally controlled payoff parameters across treatments accurately recovers the true values of those parameters with the logit quantal response equilibrium model but not with the Nash equilibrium model.

JEL Code: C92

Keywords: international conflict, terrorism deterrence, attacker-defender games, Nash equilibrium, quantal response equilibrium

---

*Charles Holt, Economics, University of Virginia, Charlottesville, VA 22904-4182, cah2k@virginia.edu and Thomas Palfrey, Division of the Humanities and Social Sciences, California Institute of Technology, trp@hss.caltech.edu. We would like to acknowledge the thoughtful suggestions provided by Clement Bohr, helpful discussions with Michael Gibilisco, and the research assistance provided by Nicholas Anderson, Georgia Beazley, Courtney Carpenter, Jessica Ewing, Maggie Isaacson, Yilin Huang, Sophia Lochner, Madison Smither, Erica Sprott, Allison Sullivan, Elin Woolf, and Max Zilvitis. Institutional Review Board approval for this research was obtained from the University of Virginia.

I. Introduction

Many settings that arise in the study of politics and international relations involve strategic conflict between two opposing sides with asymmetric roles, where the defender takes a position to fend off the attacker's attempts to exploit vulnerabilities. A country facing an imminent attack must consider how to deploy forces across alternative points of attack, as in the classic Colonel Blotto game (Roberson 2006), and the invader must choose how best to allocate attack assets. A government facing an active domestic or foreign terrorist group must decide how to allocate security forces across multiple potential targets such as stadiums, theaters, markets, airports, pipelines, and schools (Powell 2007a, Powell 2007b, Sandler and Lapan 1988).

Applications of attacker-defender games in political science are not limited to situations of violent conflict, but encompass a broader range of environments, including electoral competition. A vulnerable centrist incumbent politician facing a possible electoral challenge from the left or the right must choose which policy position to adopt to fend off alternative types of challenges (Ansolobehere et al. 2001). Candidates who must win both a primary election and a general election face a similar strategic positioning problem (Brady et al. 2007, Gerber and Morton 1998).

Even though both sides may have a rich array of available actions in these examples, it is useful to consider broad classifications to simplify analysis and isolate salient aspects of the problem for the design of a laboratory experiment. Military assaults may be either frontal or indirect, as with amphibious landings farther from the front. Candidate positions can be left or right. The games to be studied model the binary conflict as one where the attacker has two alternative strategies, while the defender two defensive strategies, one of which is more effective against a particular attack.

The payoffs for this class of games are shown in Table 1, with the payoff for the defender (row player) listed on the left side of each cell. The attacker (column player) chooses between actions a_1 and a_2, which could represent a terrorist attack at site 1 or site 2. The defender allocates defensive resources between these two alternate attacks. The defense is most effective by matching the attack, and the attack is most effective against a “mismatched” defense. The attacker’s payoff listed on the right side of each cell can be interpreted as the probability of a successful attack. This payoff is highest when the subscripts for the attack and defense actions do not match. For example, attack a_1 is most effective against defense d_2, where the positive differences $B-A$ represents the relative advantage of this attack against a mismatched defense compared to a matched defense.

Table 1. Attacker-Defender Game
with Choice Probabilities α for Attack a_1 and δ *for Defense* d_2
(defender payoff, attacker payoff)

	a_1 (α)	a_2 $(1-\alpha)$
d_1 $(1-\delta)$	$-A,\ A$	$-C,\ C$
d_2 (δ)	$-B,\ B$	$-D,\ D$

Besides the substantive importance of bilateral conflict situations, our study is also motivated by more general theoretical and behavioral questions, in particular questions about the usefulness of equilibrium models. In this regard, one key issue is the predictive value of Nash equilibrium – the workhorse theoretical paradigm for analysis of strategic interactions. The Nash equilibrium of these games, which is in mixed strategies, makes sharp predictions about how behavior should respond to changes in the relative payoffs of different combinations of attacker and defender choices. That is, if payoffs change, then the equilibrium choice frequencies also change.

Unfortunately, Nash predictions in binary conflict games can be highly counter-intuitive, which raises questions about its usefulness as a model. For example, one natural question that one might hope a model could answer involves predictions about how the probability of using an attack depends on its effectiveness. Consider the effect of raising or lowering the effectiveness of the first attacker strategy, a_1, so that the probability of success strictly increases against both defenses. It is reasonable to expect that the probability of action a_1 would increase, but Nash equilibrium predicts that such an increase in effectiveness can have no effect at all, or even worse, it can go in the opposite direction and lead to a reduction in the use of the enhanced attack strategy.

In spite of these implausible Nash predictions, there is a widely used modification of Nash equilibrium that incorporates a specific type of bounded rationality and generates intuitive predictions: *quantal response equilibrium*. (QRE, McKelvey and Palfrey, 1995; Goeree, et al., 2016, 2020). This approach, which has been applied to the analysis of both experimental and field data, is a stochastic generalization of Nash equilibrium: it maintains the equilibrium assumption of rational expectations (i.e., players' beliefs about the distribution of their opponents' actions are correct), but relaxes perfect rationality by incorporating payoff-responsive errors in the choice behavior of players. Optimal actions are not always chosen, rather actions with higher expected payoffs are chosen more frequently than actions with lower expected payoffs, as in logit and probit models. Large mistakes are less likely than small ones, which captures a range of behavioral phenomena such as distraction, miscalculation, or inattention.

The experiment reported here is designed to compare the predictive value of these two equilibrium models. The design involves two payoff manipulations of the effectiveness of attack a_1. The first manipulation increases its effectiveness by the same *absolute* amount against either defense, holding all other payoffs constant. That is, it compares the behavior in a game with payoff parameters (A, B, C, D) to the game with payoff parameters (A+γ, B+γ,C, D). Nash equilibrium predicts no change in attacker behavior, while QRE predicts an increased use of attack a_1. The second manipulation increases its effectiveness by the same *proportional* amount: the payoff parameters (A, B, C, D) become (πA, πB, C, D). Nash equilibrium predicts a reduction in the use of attack a_1, while QRE again predicts an increase in the presence of sufficient behavioral noise. The key result of the experiment is that attack frequencies are increasing in attack effectiveness, regardless of whether the change in effectiveness is proportional or absolute. This contrasts sharply with the unintuitive Nash predictions but is correctly explained by QRE.

The experiment thus provides some evidence about the relative predictive value of Nash equilibrium and QRE, in favor of the latter. Given that both approaches are being used for structural empirical work in political science, our findings may help inform researchers who face a choice of whether to analyze data through the lens of Nash equilibrium or QRE. Furthermore, for the specific application of modeling conflict situations, it raises a red flag about potential pitfalls of using Nash equilibrium to make predictions, and argues for developing predictions using QRE instead.

In addition to providing insights for this specific class of games and the potential for using QRE in structural estimation with field data, the experiment can also be seen as a textbook illustration of three advantages of laboratory experiments that make them especially valuable for theory-testing exercises like this one.

One advantage is obvious. In a laboratory study, the payoffs of the game matrix are *controlled* and hence (1) known to the experimenter and all players; (2) the same for all players; and (3) assigned exogenously in a manner that prevents unobserved nuisance variables from interfering with causal inference. In contrast, a comparable analysis of naturally occurring games would require estimation of all the relevant payoff parameters. Moreover, unobserved payoff heterogeneity creates further challenges for estimation with field data, and additional assumptions about the players' beliefs about the payoffs in the game and distribution of heterogeneity are required.

The second advantage concerns the *precision* of control. Because the experimenter is able to specify exact payoffs and subject payoff information, manipulation of payoffs to test comparative static predictions can be done with precision. In relevant naturally occurring data, for example athletic contests (Walker and Wooders 2001, Chiappori et al. 2002, Palacios-Huerta 2003), the effect of enhanced attack effectiveness on the *magnitude* of the payoff changes (for example, *B-A*) is difficult or impossible to control or estimate precisely.

The third advantage of a laboratory experiment over field data is *design*. In the laboratory, one can *choose* which payoff manipulations to employ in order to provide a powerful test of the hypotheses, e.g. when different theories make sharply opposing predictions. In contrast, with naturally occurring data, it is usually the case that the payoff manipulations are neither directly observable nor a choice variable for the researcher.

These advantages of *control, precision, and design* should not be interpreted as a call to replace the study of naturally occurring data with laboratory data. To the contrary, various data sources each have their specific advantageous features. Naturally occurring data, as well as data generated by field experiments, are typically analyzed in the specific context of a significant substantive or policy problem, enriched by important institutional details.

The next section presents the model and analyzes the Nash and QRE predictions for changes in the effectiveness of one attack option against either defense. Section III details the design of the laboratory experiment, where the treatments implement exogenous changes in the effectiveness of one of the attack actions. The main features of the data and statistical tests for the significance of treatment effects are presented in Section IV.

Formal statistical estimation and model comparison between QRE and Nash is reported in Section V. This is done in two different ways. We first fit a parametric QRE model to the data. The resulting estimates of choice frequencies produce a close match to the data, while the Nash equilibrium choice frequencies are far from the data in all but one treatment. The second approach to model comparison, which is closely related to structural empirical studies, estimates the payoff parameters in a game experiment *as if* they were unknown. Because those payoff parameters are laboratory-controlled constants that are known with certainty, we can test whether the estimates based on the QRE and Nash models successfully recover those known values. The QRE model accurately recovers the concealed payoff parameters, while Nash equilibrium does not.

II. The Model, Nash Equilibrium, and QRE

A. The Model

Consider a zero-sum game, with two possible actions for the column player (attacker), a_1 and a_2, and two possible actions for the row player (defender), d_1 and d_2. All parameters, A, B, C, and D are positive, which are payoffs for each outcome, with the defender's payoff shown as negative on the left in each cell in Table 1 in the previous section. The assumption that the attacker action does better against the "wrong" defense (i.e., with mismatched subscript) is indicated by the two inequalities in (i) below. The game is only interesting when there are no dominated strategies, which is ensured by assuming that the highest attacker payoff for each decision is greater than the lowest payoff for the other one, as specified in (ii):

i) $B > A$ and $C > D$

ii) $C > A$ and $B > D$

Finally, without loss of generality, we assume that all four payoff parameters are positive.

B. Nash Equilibrium

The Nash equilibrium is unique and is in mixed strategies, i.e. a pair of choice probabilities between 0 and 1, one for each player. Let α denote the probability that Column chooses attack a_1, and let δ denote the probability that Row chooses defense d_2; these probabilities are shown in parentheses next to the corresponding action in Table 1. The attacker's expected payoffs for each action, as a function of the defender's mixed strategy, are: $E_{a_1}(\delta) = A(1-\delta) + B\delta$ and $E_{a_2}(\delta) = C(1-\delta) + D\delta,$ so the expected payoff difference is:

$$E_{a_1}(\delta) - E_{a_2}(\delta) = A - C + (C - D + B - A)\delta.$$

This expected payoff difference must be 0 in a mixed strategy Nash equilibrium, which determines the equilibrium probability δ^* for the defender:

(1) $\delta^* = \frac{C-A}{C-D+B-A}$ (Nash equilibrium probability of d_2).

Note that this expression has been conveniently organized so that the various parameter differences (B−A), (C−A), and (C–D) are positive by the initial inequality assumptions.

The equilibrium probability for the attacker is obtained similarly. The expected payoff difference for d_2 versus d_1, as a function of the attacker's mixed strategy, α, is:

$$E_{d_2}(\alpha) - E_{d_1}(\alpha) = (C - D) - (C - D + B - A)\alpha,$$

which yields:

(2) $\alpha^* = \frac{C-D}{C-D+B-A}$ (Nash equilibrium probability of a_1).

Now consider an exogenous change that increases the effectiveness of attack a_1, by raising *both A* and *B*. Notice that these parameters only appear as a difference (B–A) in the denominator of (2). This model, therefore, yields the unintuitive prediction that equal *absolute* increases in the effectiveness parameters A and B for attack a_1 will have no effect on the equilibrium probability of choosing that attack. This prediction is due to a countervailing increase in the equilibrium probability of defending against that attack. For example, suppose that attacks a_1 and a_2 correspond to passing and running in football, and defenses d_1 and d_2 correspond to defending against a pass or against a run respectively. In this case, the treatment change implies that a pass has enhanced effectiveness against either defense, with the enhancement being of the same absolute magnitude. The Nash equilibrium prediction, however, is that the more effective pass attack is not used any more than it was prior to the enhancement.

On the other hand, it is clear from the formula for α^* in (2) that unequal increases in both A and B do not necessarily leave the Nash choice rate for a_1 unchanged. In particular, consider the effects of *proportional* changes in A and B, to πA and πB, where $D/B < \pi < C/A$, which ensures that the inequality restrictions in (ii) are not violated. With $\pi > 1$, for example, the proportional increases in the effectiveness parameters A and B for attack a_1 actually cause an increase in the denominator of (2), so the Nash probability of attack a_1 would *decrease*, which is a clearly unintuitive prediction. The reason is that the equilibrium response of an increase in the use of defense d_1 *more than offsets* the attractiveness of attack a_1.

C. Quantal Response Equilibrium

With two decisions for each player, QRE is specified by a pair of choice probabilities, α and δ, as functions of (equilibrium) expected payoff differences, in a manner that depends on the "precision" of choice behavior, denoted by $\lambda > 0$. As precision is increased, behavior gradually approaches fully rational best replies as choice probabilities respond more sharply to expected payoff differences.

Formally, the Nash indifference equations that generated (1) and (2) are replaced by QRE choice probability equations. Let $Q(\Delta)$ be any continuous, non-negative, and strictly increasing *quantal response function* of the precision-weighted expected payoff difference, Δ, between actions a_1 and a_2 (or between d_2 and d_1):

$$(3)\quad \alpha \;=\; Q\big(\lambda[E_{a_1}(\delta) - E_{a_2}(\delta)]\big) \;=\; Q(\lambda[\,\boldsymbol{A} - C + (C - D + B - A)\delta\,])$$

$$(4)\quad \delta \;=\; Q\big(\lambda[E_{d_2}(\alpha) - E_{d_1}(\alpha)]\big) \;=\; Q(\lambda[\,C - D - (C - D + B - A)\alpha\,])$$

Q is assumed to satisfy:

iii) $\lim_{\Delta\to\infty} Q(\Delta) = 1$, and $Q(-\Delta) = 1 - Q(\Delta)$ for all Δ.

The limit condition ensures that choices are approximately deterministic for extreme payoff differences or high precision. The second part of (iii) requires that the choice probabilities are not biased for one attack action over the other, and (from continuity) implies that actions are chosen with equal probability when Δ is 0. In applied work, Q is often assumed to be a logit function, which will be used in the section on estimation.

A QRE is any pair of choice probabilities, α^* and δ^*, that solve equations (3) and (4). The α and δ probabilities on the right sides of these equations can be interpreted as representing "beliefs" about the other's strategy. Rational expectations requires that these beliefs match the true equilibrium choice probabilities on the left sides of (3) and (4).

Notice that the left equation in (3) can be rewritten in terms of the inverse of Q:

$$(5)\qquad \frac{Q^{-1}(\alpha)}{\lambda} \;=\; \boldsymbol{A} - C + (C - D + B - A)\delta\,.$$

As precision goes to infinity, the left side of (5) approaches 0, so the equation reduces to the standard indifference equation (1) for the Nash equilibrium defense strategy, δ^*. Similarly, (4) reduces to equation (2) for large λ. Thus, the mixed Nash equilibrium is the limiting case of QRE for any Q. Typically, different functional forms Q will provide different QRE outcomes for finite values of λ, but the key results about the effect of changing the effectiveness of a particular attack method that are nonparametric in the sense that these qualitative predictions of QRE hold for all Q.

QRE Predictions for Equal Absolute Changes in Attack Effectiveness Parameters

Even though equal absolute changes in the effectiveness parameters A and B will have no effect on the Nash equilibrium a_1 attack probability in (2), a quick inspection of the QRE equilibrium equations (3) and (4) indicates that equal changes in the A and B parameters will have an unambiguous effect on QRE predictions. There is an additional stand-alone A term on the right side of (3), which is highlighted in bold and directly affects the expected payoff difference for the attacker. With equal changes in A and B, this stand-alone term will generate a higher a_1 attack response for any given value of the defense choice probability, δ, on the right-hand side of (3). As a result, the QRE equilibrium a_1 attack rate, α, will *increase* when a constant is added to both A and B. This leads to Proposition 1.

Proposition 1. *Under assumptions* (i), (ii), and (iii), *if the effectiveness parameters A and B increase by equal absolute amounts* to $A + \gamma$ and $B + \gamma$, *then the Nash equilibrium probability of attack* a_1 *is unchanged, but the QRE probability* α *of this attack method is increased. The QRE probability of defending against this attack* $(1-\delta)$ *increases. This result holds for any quantal response function, Q and for any positive precision,* λ.

Proof. The comparative statics prediction for the effect of equal absolute shifts in A and B by the same amount γ can be derived by differentiating (3) and (4) with respect to γ under the assumption that $dA/d\gamma = dB/d\gamma = 1$, so there is no effect on the $(B-A)$ terms in these equations. The result of this differentiation is a pair of equations in $d\alpha/d\gamma$ and $d\delta/d\gamma$:

(6) $$\frac{d\alpha}{d\gamma} = \lambda Q'(\cdot)\frac{dA}{d\gamma} + \lambda Q'(\cdot)[C - D + B - A]\frac{d\delta}{d\gamma} \quad \text{and}$$

$$\frac{d\delta}{d\gamma} = -\lambda Q'(\cdot)[C - D + B - A]\frac{d\alpha}{d\gamma}$$

Then substituting the $d\delta/d\gamma$ expression on the right into the first equation in (6) and using the fact that $dA/d\gamma = 1$, one obtains:

$$(7)\quad \frac{d\alpha}{d\gamma} = \frac{\lambda Q'(\cdot)}{1+Q'(\cdot)Q'(\cdot)\lambda^2[C-D+B-A]^2} > 0 \quad \text{(equal absolute changes in } A \text{ and } B\text{)}.$$

The ratio in (7) is positive since the $Q'(\cdot)$ derivatives are positive, even though the arguments of these functions (not shown) are different. This implies that an equal absolute increase in the *A* and *B* parameters will increase the QRE a_1 attack probability, and vice versa, for any quantal response function *Q*. The second equation in (6) then implies that $\frac{d\delta}{d\gamma} < 0$, so the probability of defending against attack a_1 (i.e., 1-δ) increases.■

Although Proposition 1 holds under general conditions, it is illustrated in Figure 1 for specific sets of payoffs that will be described in more detail in the experiment design section. The figure shows the graph of the logit QRE as a function of λ (using a log scale) for two payoff profiles, $A = 2$, $B = 8$, $C = 7$, $D = 4$, (Low effectiveness) and $A = 6$, $B = 12$, $C = 7$, $D = 4$, (High effectiveness with γ=4). For all positive values of the precision λ, QRE predicts the "intuitive" pattern, with higher effectiveness resulting in higher a_1 attack proportions.

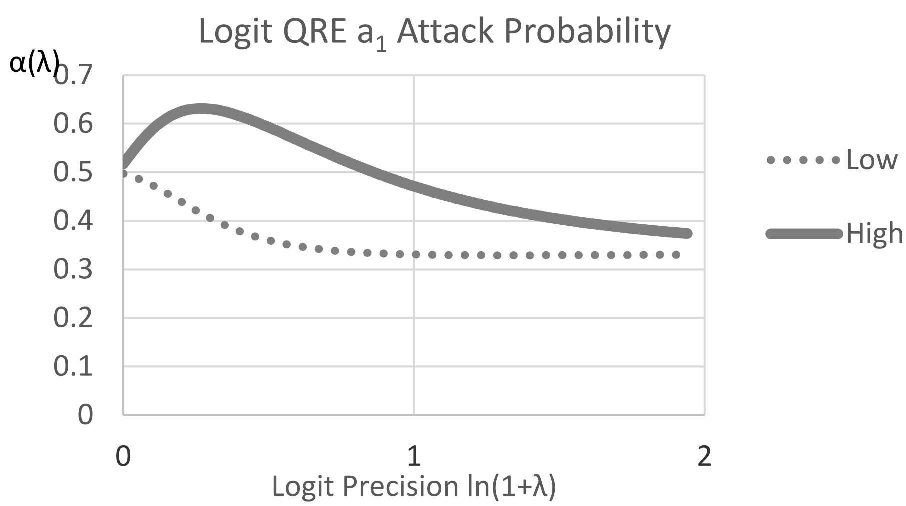

Figure 1. Locus of QRE a_1 Attack Frequencies as Precision Increases from Left to Right, with Attack Effectiveness Reduced in Equal Absolute Amounts from High (Solid) to Low (Dotted)

Notice that both curves in Figure 1 converge to the same Nash equilibrium mixed strategy of $\alpha = 0.33$ as precision increases on the right side of the figure, as implied by the theory. In contrast to the fact that they both converge to the same point, the QRE correspondence exhibits the intuitive tendency for a_1 attack proportions to be higher, *for all values of* λ for the High payoffs compared with the Low payoffs. Furthermore, the convergence is from above for both treatments: i.e., QRE predicts $\alpha > 0.33$ for both treatments, regardless of the precision parameter.

Figure 2 illustrates, in a different way, the equilibrium effects of equal absolute shifts in the attack effectiveness parameters A and B, by graphing the best response correspondences of the attacker and defender, and the analogous quantal response functions. The figure is constructed with α on the horizontal axis and δ on the vertical axis. For the Low payoffs ($A = 2$, $B = 8$, $C = 7$, $D = 4$), the attacker is indifferent between the two attack methods at $\delta = 0.55$, and so every mixture is a best response at that point, represented by the horizontal solid line at a height of $\delta = 0.55$. Similarly, the defender is indifferent between the two defense methods (and any mixture) when the a1 attack is used with probability 0.33, which results in a vertical defender best response line at that point. Therefore, the resulting best response graphs have sharp corners and intersect at $\alpha = 0.33$ and $\delta = 0.55$, the Nash equilibrium, marked by the solid diamond in the figure. Quantal response functions are the smooth curves shown in the figure for a logit specification ($\lambda = 0.7$). For the Low payoff profile, the attacker and defender quantal responses are shown as the curved upward sloping and downward sloping solid curves, respectively, which intersect at the Logit QRE slightly lower and to the right of the Nash equilibrium, marked by a solid square.

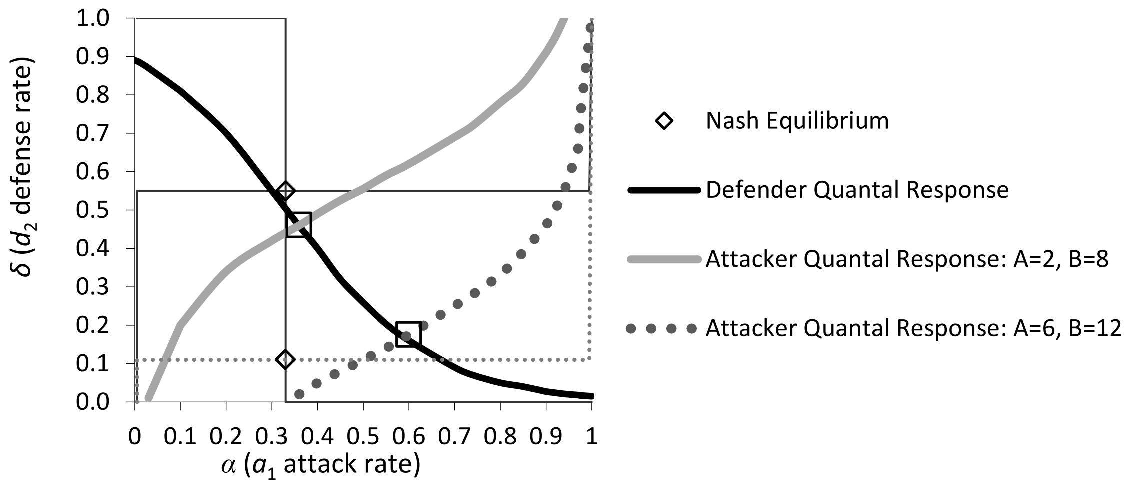

Figure 2. Effects of Equal Absolute Increases in Attack Effectiveness

Note: The mixed Nash equilibrium points (diamonds) are at intersections of sharp best response lines (straight or dotted). Quantal response equilibria (squares) are at intersections of the curved quantal response lines that were constructed for $\lambda = 0.7$.

Next consider the effect of equal absolute increases in A and B, from 2 and 8 to 6 and 12, respectively. The defender's quantal response function and best response curves depend only on the difference, B-A, and therefore both are invariant to equal *absolute* changes in A and B. However, we can see from equation (3) that the attacker's quantal response function (and best response correspondence) shifts down in response to equal absolute increases in A and B, since α depends on A directly, as well as $A-B$. The shifted quantal response function is shown in the figure by the by the dotted upward sloping curve. This shift moves the QRE intersection down and to the right, as indicated by the hollow diamond, so the QRE a_1 attack probability α increases as a result. Notice that this argument does not depend on the functional form of the quantal response function Q, and in this sense the comparative static predictions are non-parametric.

For the Nash equilibrium analysis, the increased effectiveness of a_1 causes the best response line to shift down (dotted horizontal line), but this does not change the Nash equilibrium probability of a_1 (marked by the hollow diamond). This invariance is due to the fact that the vertical solid line at α = 0.33 representing the indifference point for the defender is not affected by equal increases in A and B, so the Nash equilibrium value of α is unchanged at 0.33.

QRE Predictions for Equal Proportional Changes in Attack Effectiveness

As shown earlier, a *proportional increase* in attack effectiveness actually *decreases* the Nash equilibrium probability of this attack, which is a clearly unintuitive prediction. For very high levels of precision, the QRE equilibria must converge to Nash equilibria, so some QRE predictions would share this unintuitive prediction. However, this is not the case for a wide range of plausible lower precisions, as can be seen in Figure 3, which compares the logit QRE as a function of λ for two payoff profiles that are used in the experiment: $A = 3$, $B = 6$, $C = 7$, $D = 4$, (low effectiveness) and $A = 6$, $B = 12$, $C = 7$, $D = 4$, (high effectiveness). Except for very high values of λ, QRE predicts the "intuitive" pattern, with higher effectiveness resulting in higher a_1 attack proportions.

This effect of equal proportional changes in A and B shown for the example in Figure 3 is a general property of *any* QRE (not only logit):

Proposition 2. *Under assumptions* (i), (ii), and (iii), *if the a_1 attack effectiveness payoffs, A and B*, *increase by equal proportional amounts* (*to πA and πB, where $\pi > 1$*), *then for any quantal response function Q satisfying (iii), there exists a positive precision parameter λ_0 such that for* $0 < \lambda < \lambda_0$, *the* QRE *a_1 attack rate with high effectiveness is higher*, *while the analogous Nash attack rate is lower*. *Furthermore, the frequency of defending against a_1 increases in π for all values of λ.*

Proof. Replace A and B with πA and πB, respectively, in equations (3) and (4). Denoting $x = C–D + \pi(B−A)$, equations (3) and (4) reduce to:

$$\alpha = Q(\lambda[\pi A - C + x\delta]) \quad \text{and} \quad \delta = Q(\lambda[C - D - x\alpha])$$

Since the derivative of x with respect to γ is $B−A$, the total derivative of the QRE conditions with respect to π are:

$$\alpha' = \lambda Q'(\cdot)[A + \delta(B - A) + x\delta'] \quad \text{and} \quad \delta' = -\lambda Q'(\cdot)[\alpha(B - A) + x\alpha'].$$

These 2 equations can be solved for α' and δ'. Hence, the comparative statics effect of π on δ is:

$$\delta' = \frac{-\lambda Q'(\cdot)[\alpha(B-A) + x\lambda Q'(\cdot)[A+\delta(B-A)]\,]}{1+\lambda^2 Q'(\cdot)Q'(\cdot)x^2}$$

Even though the arguments of the Q' functions are omitted, Since $Q' > 0$, it follows that δ'<0, as stated in the proposition. Similarly, the effect of an increase of π on the attack rate α is $\alpha' = \lambda Q'(\cdot)[A + \delta(B-A) + x\delta']$. The first term, $\lambda Q'(\cdot)$ is strictly positive for all $\lambda > 0$ and the second term, $A + \delta(B-A) - x\delta'$, is also positive for sufficiently small λ. It is easily verified that the effect on the Nash equilibrium goes in the opposite direction.■

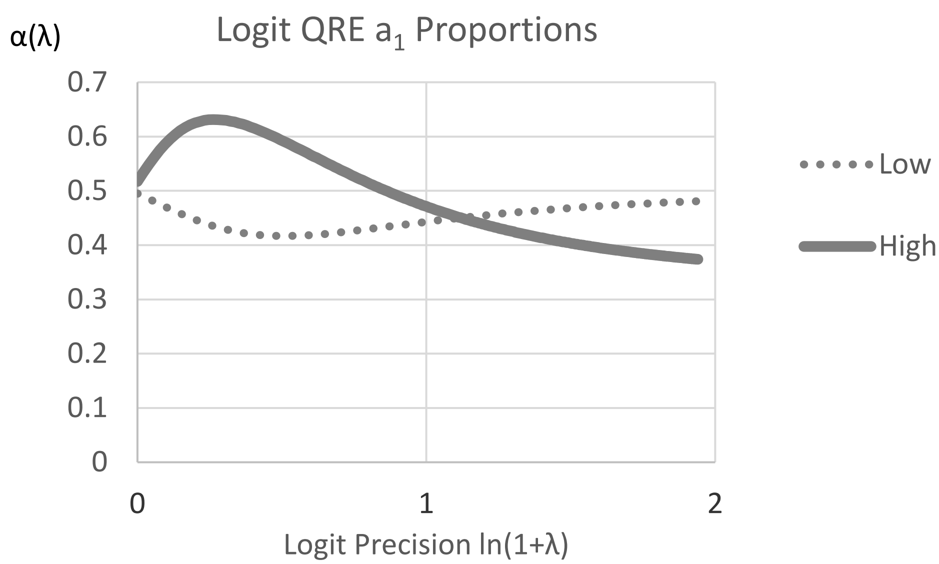

Figure 3. QRE Locus Predicted a_1 Attack Proportions as Precision Increases from Left to Right, with Attack Effectiveness Reduced in Equal Proportions from High (Solid Line) to Low (Dotted)

Figure 4 shows the intuition behind the much different Nash and logit QRE changes in attacker and defender behavior resulting from an equal proportional increase in both A and B.[1] The change induces a leftward shift of the downward sloping defense quantal response function (to the dark dotted curve) and a rightward shift of the upward-sloping attacker quantal response function (to the light dotted curve). The combined effect of both shifts is to reduce the QRE d_2 defense rate, δ, and *increase* in the QRE a_1 attack

[1] The quantal response curves in the figure correspond to the value of λ (0.7) estimated from our data.

rate. In contrast, the Nash predictions for this treatment change are indicated by the switch from the solid, straight-line best responses in Figure 4 to the dotted-line best responses, resulting in a *decrease* in the Nash equilibrium a_1 attack rate from 0.50 to 0.33.

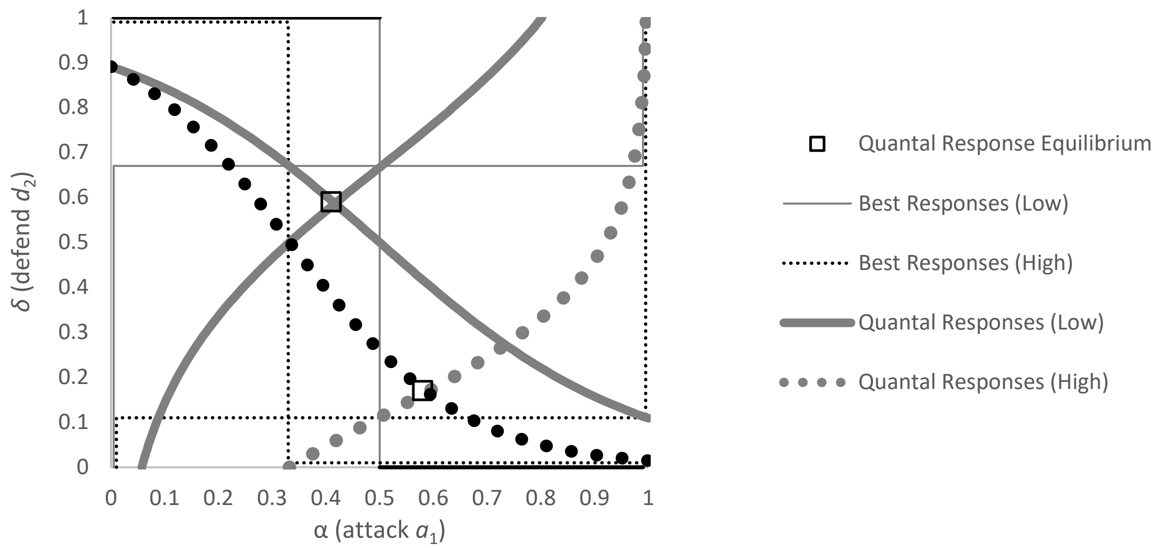

Figure 4. Effects of Equal Proportional Increases in *A* and *B*

III. Experimental Design and Procedures

The experimental design, summarized in Table 2 uses the three payoff functions analyzed in the previous section, which allows a comparison of the effect of absolute versus proportional changes in the effectiveness of attack a_1. The Low Effectiveness treatment in the bottom row of the table corresponds to relatively low attacker payoffs for attack a_1, and the High treatment corresponds to high attacker payoffs for a_1.

Table 2. Treatments, Predictions, and Data Averages for Equal Absolute or Equal Proportional Changes in the Effectiveness of Attack a_1

	Equal Absolute Changes		Equal Proportional Changes	
	A, B, C, D	Nash a_1 Rate	A, B, C, D	Nash a_1 Rate
High Effectiveness of Attack a_1	6, 12, 7, 4	0.33	6, 12, 7, 4	0.33
Low Effectiveness of Attack a_1	2, 8, 7, 4	0.33	3, 6, 7, 4	0.50

The design on the left side of the table involves *equal absolute changes* in the two attack effectiveness parameters, A (against defense d_1) and B (against defense d_2). The High Effectiveness treatment is shown in the top row of Table 2, with a parameter set for A, B, C, D of 6, 12, 7, 4. The low absolute effectiveness treatment (henceforth called LowAb) is shown in the bottom row on the left side, with a reduction of both A and B by 4 in each case, i.e. from 6 and 12 to 2 and 8, respectively, resulting in no change of the Nash equilibrium a_1 attack rate of 0.33. The low proportional effectiveness treatment (henceforth called LowProp) is shown in the bottom row on the right side, with both A and B reduced by 50% from 6 and 12 in the top row to 3 and 6, respectively, resulting in an *increase* in the Nash equilibrium a_1 attack rate from 0.33 to 0.50

The experiment was conducted with payoffs (in dollars) reduced by a factor of 10, e.g. from 6 to 0.60, etc. We also added a fixed amount of 1.20 to all defender payoffs to avoid negative payoffs. These adjustments have no effect on the Nash equilibrium predictions in (1) and (2). We will continue to discuss the treatments in terms of integer amounts, e.g. 6, 12, 7, 4, which now refer to 10-cent payoff units.

We conducted 4 sessions (48 subjects) with the equal absolute change design on the left side of Table 2, and another 4 sessions (48 subjects) with the equal proportional change on the right side. Each session consisted of 20 rounds for each of the two attack effectiveness treatments (High and either LowAb or LowProp). Roles stayed the same (attacker or defender), and matchings were fixed, both within each 20-round part and across parts.[2] Both treatment orders (Low-High and High-Low) were used, with Low-High in half of the sessions and High-Low in the other half of the sessions. Subjects were recruited from the University of Virginia student population and were paid $6 plus all earnings. Total earnings averaged $30 for a session that lasted about 45 minutes, which included the reading of instructions aloud. The experiment was run with the V*e*conlab matrix game software. The game was presented with minimal context, i.e. with decisions labeled Top or Bottom for the row player, and Left or Right for the column player. Verbatim instructions from one of the treatments are provided in the Appendix.

[2] Fixed matching does not invite repeated game effects, since the payoffs are constant-sum.

IV. Results

Figure 5 shows the time series of a_1 attack frequency averages ($\bar{\alpha}$) for the first paired treatment that changes the attack effectiveness parameters A and B by equal absolute amounts. Recall that this change does not affect the Nash equilibrium prediction for the a_1 attack rate, which is 0.33 for both treatments, as indicated by the horizontal dark line at that level. The solid lines that connect data average points in the figure are averages for High treatment, and the dashed lines are for LowAb treatment. The averages for sessions with High in the first 20 rounds are displayed in gray, and averages for the reverse order are displayed in black. For example, the black dashed line on the left shows that the observed a_1 frequencies center around the Nash prediction of 0.33 with low effectiveness for the first 20 rounds. In contrast, the black solid line on the right indicates that a_1 attack frequencies almost doubled in the final 20 rounds after the treatment switch *for the same subjects and the same Nash prediction*. Average a_1 attack rates are obviously greater in the High treatment, with an overall average of 0.59 (for both treatment orders), than in the Low treatment, with an overall rate of 0.33. This difference holds true for pairwise comparisons in all rounds but one. Moreover, there are no clear sequence or trend effects in terms of average attack frequencies.

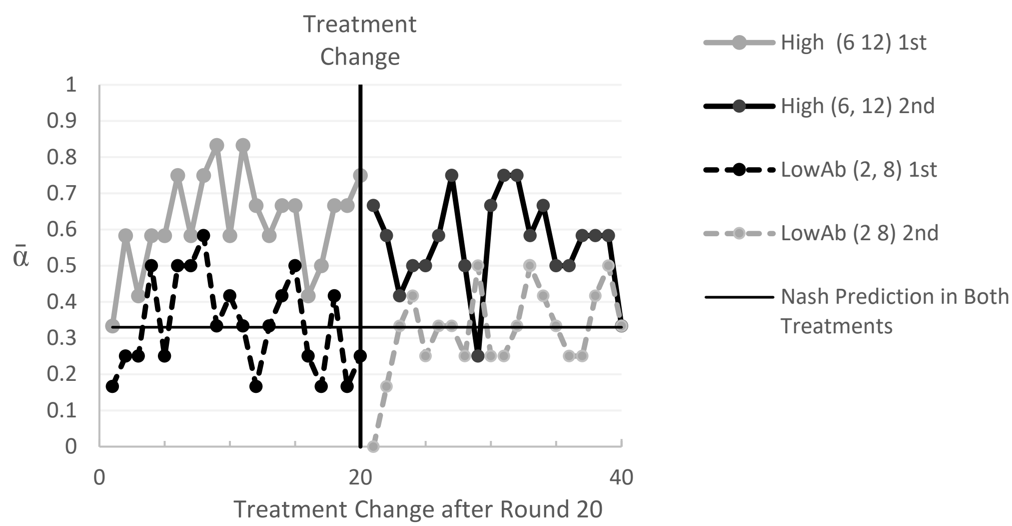

Figure 5. Average a_1 Attack Frequencies, $\bar{\alpha}$, by Round, with Equal Absolute Changes: Solid Lines for HighAb (6, 12), Dashed Lines for LowAb(2, 8)

Figure 6 shows a similar plot of a_1 frequencies for the case of equal proportional changes in the *A* and *B* effectiveness parameters for this attack method. The treatment change from the LowProp treatment (3, 6) to the High treatment (6, 12) reduces the Nash equilibrium value of α from 0.5 (horizontal line on left side of Figure 6) to 0.33 (horizontal line on right), or vice versa for the opposite sequencing of the treatments. Data averages for the 20 rounds in the LowProp treatment are shown by a dotted line on the left side of Figure 6, regardless of whether it was done first or second, and the 20 rounds of data with the High treatment are displayed on the right. Observed a_1 attack frequencies increase from 0.49 to 0.63 as *A* and *B* increase proportionally, which is the opposite of the Nash equilibrium downward shift at the midpoint of the figure.

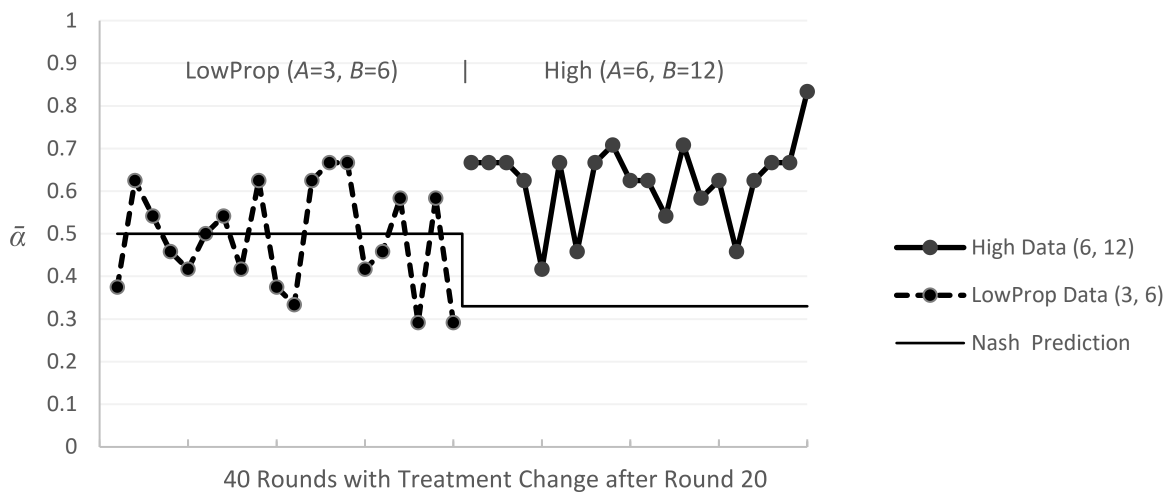

Figure 6. Average a_1 Attack Frequency by Round, with Equal Proportional Changes: Dashed Line for LowProp (3, 6), Solid Line for High (6, 12)

If the enhanced effectiveness of the a_1 attack is anticipated by the defenders, then the incidence of the matched d_1 defense should rise and the incidence for d_2 should fall. Table 3 shows a breakdown of the overall frequencies of both attack (a_1) and defense (d_2) decisions, with a row of treatment averages in bold and a row of Nash predictions in italics. First consider the left side of the table, for the equal absolute change, from the LowAb treatment with $A = 2$ and $B = 8$ to the High treatment with $A = 6$ and $B = 12$. When the *A* and *B* effectiveness parameters for attack a_1 increase by equal amounts in this manner, the targeted defense d_1 against this attack is predicted to increase, and the propensity δ for the other defense d_2 is predicted to decline sharply from 0.56 to 0.11, as shown in the bottom row on the left. The corresponding increase in $1-\delta$ is what causes the Nash equilibrium value of α for attack a_1 to stay constant at 0.33, even in the presence of the doubled effectiveness of this attack method. The data show a less dramatic decline in the d_2 frequency, from 0.41 to 0.19 percent, coupled with the a_1 frequency that almost doubles, from 0.33 to 0.59 percent, which leads to our first result.

Result 1a: An equal absolute increase in effectiveness parameters for attack a_1 (from 2 and 8 to 6 and 12) causes a significant and large (nearly doubled) increase in the observed frequency of this attack, despite the Nash prediction of no change.

Table 3. Treatments, Nash Predictions, and Observed Data Proportions

	Equal Absolute Change		Equal Proportional Change	
Treatment Payoffs	LowAb	HighAb	LowProp	HighProp
(A, B, C, D):	(2 8 7 4)	(6 12 7 4)	(3 6 7 4)	(6 12 7 4)
Observed Attack Rate α:	**0.33**	**0.59**	**0.49**	**0.63**
Nash prediction:	*0.33*	*0.33*	*0.50*	*0.33*
Observed Defense Rate δ :	**0.41**	**0.19**	**0.66**	**0.17**
Nash prediction:	*0.56*	*0.11*	*0.67*	*0.11*

Support: The observed overall treatment difference is about 26 percentage points (0.5917 – 0.3292 = 0.2625). The a_1 frequency increased in the High treatment for all but 6 of 24 attacker-defender pairs, with 3 relatively small reversals and 3 ties with equal attack rates in each treatment. Since each pair interacted with each other in all 20 rounds, we use an exact matched-pairs test to calculate the proportion of permutations of treatment labels that yield a treatment difference as great as or greater than the observed difference of 0.2625. The resulting p-value is 0.0004 for a 2-tailed test and 0.0002 for a 1-tailed test, so the null hypothesis of no effect is rejected at the 1% level.

The increase in a_1 attack rates in the HighAb treatment is consistent with the observation that the d_1 defense against this attack was observed less frequently than predicted in the Nash equilibrium. The increase in a_1 attack effectiveness resulted in a decline in the d_2 defense frequency from 0.41 to 0.19, which is a substantially smaller decline than the predicted decline from 0.56 to 0.11, as shown in the bottom row of Table 2, left side, a difference that is highly significant:

Result 1b: An equal <u>*absolute*</u> *increase in effectiveness parameters for attack* a_1 *(from 2 and 8 to 6 and 12) results in a tendency to over-defend against the other attack* a_2 *relative to the Nash prediction of a sharper reduction in* d_2 *defense rates.*

Support: The observed overall reduction in the d_2 defense from 0.41 to 0.19 is about 22 percentage points, which is much less than the predicted 45 percentage point reduction from 0.56 to 0.11. The d_2 frequency decreased by less than 45 points for all but 4 of 24 attacker-defender pairs, as can be verified by considering the d_2 defense rate changes shown in the rows of the first page of the Data Appendix. This difference is significant at $p < 0.001$ for a 2-tailed binomial test of the null hypothesis that the reduction in d_2 defense rates is equally likely to be greater than or less than the Nash prediction.

Next consider the results for the equal proportional change with doubled A and B parameters, shown on the right side of Table 3. The Nash equilibrium value of α no longer stays the same, but rather, declines from 0.50 to 0.33. In contrast to this predicted decline, the attack frequency α *increased* from 0.49 to 0.63 percent, a 14 point change, yielding an a_1 attack frequency that is nearly double (30 percentage points higher than) the Nash equilibrium frequency of 0.33. The observed increase is significant for a 2-tailed matched pairs permutation test (p = 0.034). This is summarized as our second result:

Result 2a: An equal <u>*proportional*</u> *increase in effectiveness parameters for attack* a_1 *(from 3 and 6 to 6 and 12) causes a significant increase in the observed frequency for this attack, despite the Nash prediction that the frequency will decline.*

The final result pertains to the defender responses:

Result 2b: An equal *<u>proportional</u>* *increase in effectiveness parameters for attack a_1 (from 3 and 6 to 6 and 12) does not result in a clear tendency to over-defend against the other attack a_2 relative to the Nash prediction.*

Support: The overall observed reduction in the d_2 defense from 0.66 to 0.17 is close to the predicted 56 percentage point reduction (from 0.67 to 0.11) in a Nash equilibrium. The d_2 frequency decreased by less than the predicted 56 points for 16 of 24 fixed attacker-defender pairs and by more than 56 points for 8 of 24 pairs. This difference is not significant, with $p = 0.152$ for a 2-tail binomial test of the null hypothesis that the reduction in d_2 defense rates is equally likely to be greater than or less than the Nash prediction.

Finally, note that the choice frequencies in the third column of Table 3 for both attackers and defenders in the LowProp treatment are almost exactly equal to the Nash predictions (and a null hypothesis of no difference cannot be rejected by standard statistical tests). Some intuition for this result is suggested by Figure 4, where the solid quantal response curves for the LowProp treatment intersect at a point that is relatively close the Nash intersection of the straight best response lines for this treatment. This suggests that Nash predictions may be more accurate when Nash and QRE predictions are close, and less accurate otherwise, which leads into the next section where we characterize the properties of QRE in binary conflict games.

V. Estimation and Model Comparison of QRE and Nash

In this section, we begin with a standard estimation of the logit responsiveness parameter, λ, from data generated by controlled laboratory experiments. Every non-negative value of λ implies a profile of six QRE choice frequencies (for 2 player roles and 3 treatments). Because each such QRE profile is uniquely associated with a single value of λ in this class of games, it is straightforward to obtain a maximum likelihood estimate of λ given the data (Goeree et al. 2016, Ch. 6).

Our second method of model comparison is less standard in the analysis of experimental data and more commonly used in structural estimation of game theoretic models in empirical studies based on historical data or on field observations. Such data sets typically do not have direct measurements of the key structural parameters of the game theoretic model used as the theoretical foundation for the empirical study, which in our case would be the exact payoffs in the binary conflict game matrix. Such an empirical analysis would requires estimation of these parameters, under maintained assumptions about the equilibrium behavior of the agents. One can conduct the same kind of estimation exercise with experimental data by concealing these (known) payoff parameters and treating them as if they were unknown. For example, Merlo and Palfrey (2018) call this method of model validation "concealed parameter recovery" and used it to compare the performance of different behavioral models of voter turnout using experimental data. Here we use this method by concealing the two game payoff parameters (C and D) that are constant across our treatments and then compare how well the QRE and Nash models recover the revealed parameters.

A. Logit QRE Estimation

The previous analysis was based on qualitative data comparisons with patterns predicted by a Nash equilibrium or with general QRE comparative static predictions that hold non-parametrically. In this section, we report estimation results based on the parametric logit specification. The estimated value of λ was constrained to be the same for all treatments. The quantal response functions (3) and (4) were rewritten as ratios of exponential expressions using the logit form: $Q(\lambda\Delta) = \frac{1}{1+\exp(-\lambda\Delta)}$, which resulted in two equations in the equilibrium attack and defense rates, α and δ, for a given logit precision λ. The equilibrium values of α and δ for each treatment were calculated numerically for all possible values of λ. The likelihood function is constructed as a product of each predicted choice probability (for each treatment) raised to a power that equals the number of times that choice was observed in the data. Since the QRE predictions are depend on the logit precision λ, the likelihood function is evaluated for each λ to determine that it is maximized at $\lambda = 0.70$.[3]

The maximum likelihood estimate of λ was used to construct the quantal response lines shown earlier in Figures 2 and 4 for the paired treatment changes. In each case, the fitted QRE choice predictions are at the intersections of the decreasing defense quantal response curve with the solid and dotted attack (increasing) quantal response curves. In both cases, the QRE intersections move to the right, correctly predicting the observed increase in the α_1 attack rate after the treatment change documented in the previous section. These QRE predictions contrast sharply with the Nash attack rate predictions (of no change for equal absolute increases in A and B, and of a reduction for equal proportional increases).

Figure 7 shows a unified picture of QRE predictions (squares at intersections of quantal response lines) and dark data dots that indicate the data averages for the various treatments. The data dot in the lower right part shows the averages for the High treatment that was run twice, once paired with the LowAb treatment and once with the LowProp treatment. The QRE (like Nash) predictions are consistent with the sharp reduction in the d_2 frequency in the High treatment, but QRE correctly accounts for the salient feature of the data, the increase in the a_1 attack frequency, which is the opposite of the Nash prediction of no change (from LowAb) or a decrease (from LowProp). The QRE intuition is that the enhanced effectiveness of the a_1 attack causes the attacker's upward sloping quantal response curves to shift to the right, as shown by the upward sloping dotted line.

[3] This process is described in more detail, with examples, in Goeree et al. (2016), chapter 4. Since there is twice as much data for the A=6, B=12 treatments, we divided the data counts and number of observations for this treatment by 2 in the likelihood function to avoid over-weighting this treatment.

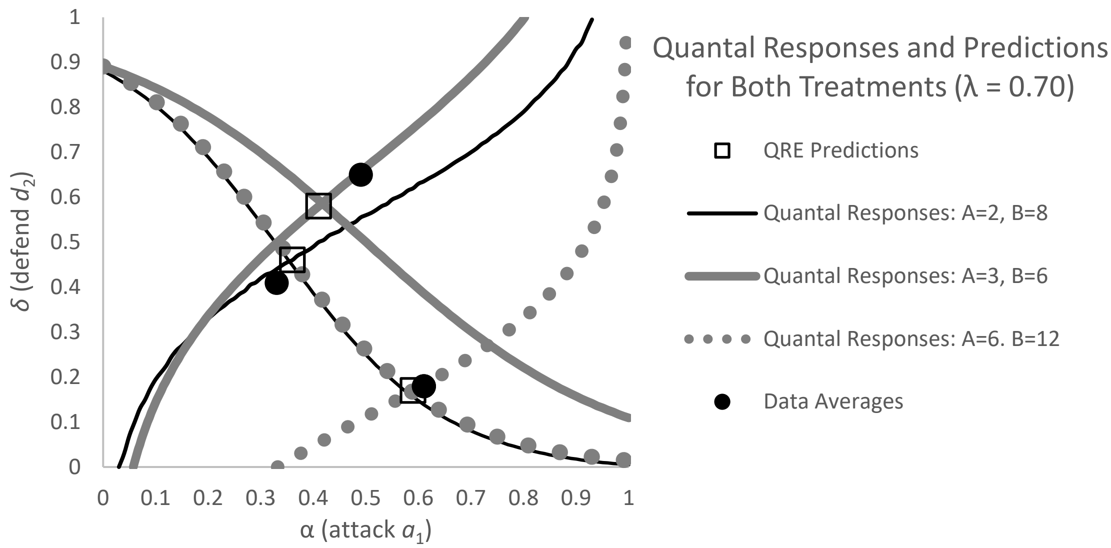

Figure 7. Effects of Equal Absolute and Equal Proportional Increases in *A* and *B*

Table 4 presents a numerical comparison of the treatment data averages with the theoretical predictions. The data points in the top row, in bold, are listed as (α, δ) pairs for a_1 and d_2 choice rates. The middle and bottom rows show analogous QRE and Nash point predictions by treatment. For the equal absolute change treatments on the left side of the table, the fitted logit QRE predictions in the middle row imply that the a_1 attack frequency should increase from 0.36 in the LowAb treatment ($A = 2$, $B = 8$) to 0.59 in the HighAb treatment ($A = 6$, $B = 12$). This strong treatment effect in attack rates is inconsistent with the Nash prediction of no effect, i.e. an unchanged attack rate of 0.33 in the bottom row, left side. A similar pattern can be observed for attack rates in response to equal proportional increases in attack effectiveness (from LowProp to HighProp) on the right side of Table 4. This treatment change increased the a_1 attack frequency observed in the data from 0.49 to 0.63 (top row, right side), a change that was reasonably well predicted by logit QRE, from a predicted attack frequency increase from 0.41 to 0.59, in stark contrast to the Nash prediction (bottom row, right side) of a decline from 0.50 to 0.33.

Table 4. Data, Nash, and Logit QRE: (a_1 Attack Rate, d_2 Defense Rate)

	Equal Absolute Change		Equal Proportional Change	
Treatment Condition:	LowAb	HighAb	LowProp	HighProp
Data Avg.:	**(0.33, 0.41)**	**(0.59, 0.19)**	**(0.49, 0.66)**	**(0.63, 0.17)**
Logit QRE for $\lambda = 0.7$*:*	(0.*36, 0.46*)	(0.*59, 0.17*)	(0.*41, 0.58*)	(0.*59,0.17*)
Nash Equilibrium:	(0.*33, 0.56*)	(0.*33, 0.11*)	(0.*50, 0.67*)	(0.*33, 0.11*)

Turning next to the defense behavior, the Nash equilibrium over-predicts the reduction in d_2 frequencies with equal absolute changes, which fell from 0.41 to 0.19. This observed d_2 defense frequency of 0.19 after the treatment change (top row, left side) is close to and not significantly different from the logit QRE prediction of 0.17 (middle row). But the observed 0.19 defense frequency rate in the HighAb treatment is somewhat greater than the 0.11 Nash prediction of 0.11 shown below it in the bottom row, left side, although this latter difference is not significant.[4] As was the case with equal absolute changes in A and B, there is a sharp reduction in d_2 rates from an equal *proportional* increase in A and B, from 0.66 to 0.17. This observed reduction in d_2 defense rates is also over-predicted by Nash, which predicts a reduction to 0.11, but not by QRE, which predicts the reduction to 0.17 that was observed. As before, neither of these differences between observed and either Nash or QRE predictions are statistically significant.

B. Concealed Payoff Parameter Estimation

The method used above for QRE estimation and comparison with Nash equilibrium was only possible because, in a laboratory setting, the payoff parameters of the games for which the data is generated are known, fixed numbers. This contrasts with structural estimation methods used in applied field settings, for example in the analysis

[4] Even though statistical tests of attack-effectiveness treatment effects have been provided, we include several more tests here for descriptive purposes. First consider whether the observed post-treatment defense rate of $\delta = 0.19$ is different from the QRE prediction of a 0.17 in the HighAb treatment. The far right column of the first page of the Data Appendix can be used to document that 11 of the 24 subjects in defender roles had d_2 defense rates above QRE prediction (0.17), and the other 13 had lower defense rates. This difference is not significant at standard levels with a binomial test. For comparison, 15 the 24 defenders had defense rates above the Nash prediction of 0.11, with 9 defense rates below, which is again not significant.

of auction data, where payoff parameters like prize values are unknown to the econometrician and need to be estimated using the error structure of the model.

In this subsection, we follow this traditional field-data approach by ignoring the fact that the payoff parameters in the experiment are known, and instead blinding ourselves to those parameters that are common to all treatments. Then we directly estimate those payoff parameters for both QRE and Nash equilibrium models. Since the true values of the estimated parameters are known with certainty, we can then conduct a simple statistical test, for each of the two models, to see whether the concealed parameters are accurately recovered by the estimation.

There are two parameters that we fixed throughout the four treatments of the experiment: *C*=7 and *D*=4. We use maximum likelihood to estimate these two parameters using the QRE model with λ fixed at 0.7, yielding $\hat{C}_{QRE}$ and $\widehat{D}_{QRE}$. For comparison, we also use maximum likelihood to estimate these two parameters using the Nash mixed strategy equilibrium model, yielding $\hat{C}_{Nash}$ and $\widehat{D}_{Nash}$ We then conduct a likelihood ratio test for each of the two models, QRE and Nash, with *C*=7 and *D*=4 as the null hypothesis. The other parameters in the game, *A* and *B*, vary across the three treatments, and are not estimated. Both QRE and Nash models are identified because we have three different (*A*, *B*) pairs in our treatment, each of which generates different QRE and Nash equilibrium strategy profiles for all values of *C* and *D* that satisfy inequalities (i) and (ii) in section II.

The results of the estimation are given in Table 5. We conduct a likelihood ratio test for both the QRE and the Nash models, where the test statistic is twice the likelihood ratio and is chi-square distributed with two degrees of freedom. For the QRE model, we obtain an estimate of $\hat{C} = 7.1$, $\widehat{D} = 3.7$, which are not significantly different from (7, 4) at the 5% significance level (χ^2=4.602). For the Nash model, we obtain an estimate of $\hat{C} = 7.6$, $\widehat{D} = 3.1$, and the test statistic (χ^2=69.94) is sufficiently high to strongly reject the null of (7, 4) at the 1% level (*p*<0.001). The bottom line is that the QRE model successfully recovers the concealed parameters, *C* and *D*, while the Nash model does not.

Table 5. Concealed Parameter Estimation (a_1 Attack Rate, d_2 Defense Rate)

Treatment:	**LowAb**	**High**	**LowProp**	**−logLikelihood**
Attack and Defense Data:	(0.33, 0.41)	(0.61, 0.18)	(0.49, 0.66)	
QRE estimates:				
with $C = 7, D = 4$	(0.*36, 0.46*)	(0.*59, 0.17*)	(0.*41, 0.58*)	1832.675
with $\hat{C} = 7.1, \hat{D} = 3.7$	(0.*38, 0.47*)	(0.60*, 0.18*)	(0.*45, 0.59*)	1830.374
Nash estimates:				
with $C = 7, D = 4$	(0.*33, 0.56*)	(0.*33, 0.11*)	(0.*50, 0.67*)	1924.214
with $\hat{C} = 7.6, \hat{D} = 3.1$	(0.43*, 0.53*)	(0.*43, 0.15*)	(0.60*, 0.61*)	1889.242

The difference between the performances of the two models in recovering the payoff parameters is also reflected in the difference between the constrained and unconstrained estimates of the mixed strategy attack and defend frequencies in the three treatments. The average difference in the estimate for the attack rate α is 0.10 for the Nash model and 0.02 for the QRE model, while the average difference in the defense rate δ estimate is 0.04 for the Nash model and 0.01 for the QRE model.

VI. Conclusions

The experiment reported here was motivated in three ways. First, the games we study capture some of the key strategic elements of bilateral conflict, a classic and important problem that arises in several areas of political science. Second, the workhorse theoretical paradigm for studying these models, Nash equilibrium, provides unintuitive predictions about how behavior responds to changes in underlying payoffs in the game. Third, from the perspective of behavioral game theory, quantal response equilibrium, which modifies Nash equilibrium by injecting payoff-responsive errors into choice behavior, yields more intuitive predictions about the effects of such changes. In particular, the effect of these errors is that defenders respond only partially to the increased effectiveness of an attack method. This partial responsiveness for defenders tends to reinforce the incentive for attackers to increase use of the attack method with enhanced effectiveness, resulting in intuitive attack patterns that differ from the Nash predictions.

The laboratory experiment implements two distinct changes in the effectiveness of one of the attacker's two strategies, equal *absolute* or *proportional* increases in the effectiveness of that strategy. In the first case, Nash equilibrium predicts no change of attacker behavior. In the second case, Nash equilibrium predicts that the improved strategy is used *less often*. Both predictions are sharply contradicted by the observed increase in a_1 attack rates following a switch from low to high effectiveness for that action. In contrast, the observed data are consistent with QRE.

One might naturally wonder whether it would have been possible to explain any data pattern with this model, but this is not the case. The discipline that ensures the *empirical content* for QRE in our setup derives from the requirement that the quantal response function in equations (3) and (4) are smooth and payoff responsive, in the sense that action choice probabilities are ordered by equilibrium expected payoffs. Better actions are chosen more often that worse actions. The results are nonparametric in the sense that they do not depend on the functional form of the quantal response functions, only its ordinal properties. Another concern is that our QRE estimation allows for an additional parameter, the logit precision, so we may be overfitting the data. Overfitting is avoided by considering decisions made by different subjects in different treatments, restricting the free parameter to be constant across subjects and treatments. One could get a much better, or perhaps perfect, fit by allowing for subject-specific or treatment-specific parameters, but this would be counter to the spirit of the study.

The use of Nash equilibrium is the standard approach for the structural estimation of game theoretic models using naturally occurring data that arise in applied microeconomics. Similarly, the Nash paradigm is used in political science for empirical analysis in a wide range of substantively important areas: legislative organization (Diermeier et al. 2003, Diermeier et al. 2005); judicial behavior (Iaryczower and Shum 2012); crisis bargaining in international conflicts (Crisman-Cox and Gibilisco 2018, Lewis and Schultz 2003, Signorino 1999, 2003); the causes of civil war (Gibilisco and Montero forthcoming); committee decision making (Lopez-Moctezuma 2019). As some of those studies have shown, QRE is a feasible and useful alternative to Nash equilibrium as a theoretical foundation for structural modeling. Our experimental findings suggest that in some empirical applications, when facing a choice between using Nash equilibrium and QRE, , the latter approach, if feasible, might be the better choice.

As a relatively novel methodological contribution, this paper used concealed parameter recovery as a method of validation of the Nash and QRE models. This method is closely related to such structural empirical studies, and our analysis demonstrates how this technique can be applied to experimental data: structural payoff parameters that are experimentally controlled and known with certainty to the experimenter and the subjects, are treated as unobserved ("concealed"), estimated ("recovered") from the observed choices, and compared to the true values of those parameters. In the application to bilateral conflict games, we find that QRE provides accurate parameter recovery, while estimation based on the Nash equilibrium model fails.

We close with two comments that are suggestive of some directions one might want to extend this study. First, while the study shows that payoff responsive errors can go a long way in explaining qualitative effects of payoff changes in these games, there are other behavioral factors that are also likely to play a role. Players in bilateral conflict games face risk and strategic uncertainty, which suggests that aversion to risk or ambiguity could be important factors in addition to bounded rationality. Payoff inequities could also produce systematic effects on the choice behavior of both players, and this could interact with the equilibrium effects. QRE provides a natural structural framework the estimation of the effects of such factors, which would be an interesting extension of our analysis. Second, we chose the simplest possible game for modeling bilateral conflict – a zero-sum two-by-two game where players have complete information about the payoffs. The real world is more complex than this, and it would be worthwhile to explore more general versions of these simple games that capture other relevant factors in the arenas international conflict and political competition. Asymmetric information and non-zero-sum payoffs would seem to be especially promising avenues to pursue.

References

Ansolabehere, Stephen, James M. Snyder, Jr. and Charles Stewart, III (2001) "Candidate Positioning in U.S. House Elections," *American Journal of Political Science*, 45(January): 136-159.

Brady, David W., Hahrie Han, and Jeremy C. Pope (2007) "Primary Elections and Candidate Ideology: Out of Step with the Primary Electorate?", *Legislative Studies Quarterly,* 32(February): 79-105.

Carrillo, Juan and Thomas R. Palfrey (2009) "The Compromise Game: Two-sided Adverse Selection in the Laboratory, *American Economic Journal: Microeconomics*, 2009 1(February):151-81.

Chiappori, Pierre-Andre; Levitt, Steve and Groseclose, Timothy (2002) "Testing Mixed Strategy Equilibria When Players Are Heterogeneous: The Case of Penalty Kicks in Soccer," *American Economic Review*, 92(4), 1138-1151.

Crisman-Cox, Casey and Michael Gibilisco (2018) "Audience Costs and the Dynamics of War and Peace," *American Journal of Political Science*, 62(3): 566-580.

Diermeier, Daniel, Hulya Eraslan, and Antonio Merlo (2003) "A Structural Model of Government Formation," *Econometrica*, 71(January): 27-70.

Diermeier, Daniel, Michael Keane, and Antonio Merlo (2005) "A Political Economy Model of Congressional Careers," *American Economic Review*, 95(1): 347-373.

Fearon, James D. (1994) "Signaling versus the Balance of Power and Interests An Empirical Test of a Crisis Bargaining Model." *Journal of Conflict Resolution*, 38:(June), 236-69.

Fey, Mark and K. W. Ramsay (2007) "Mutual Optimism and War", *American Journal of Political Science*, 51: 738-54.

Gerber, Elisabeth R. and Rebecca B. Morton (1998) "Primary Election Systems and Representation", *Journal of Law, Economics, and Organization*, 14(October): 304-24.

Gibilisco, Michael and Sergio Montero (forthcoming), "Do Major-Power Interventions Encourage the Onset of Civil Conflict? A Structural Analysis", *Journal of Politics*.

Goeree, Jacob K., Charles A. Holt, and Thomas R. Palfrey (2005) "Regular Quantal Response Equilibrium," *Experimental Economics*, 8: 347-367.

Goeree, Jacob K., Charles A. Holt, and Thomas R. Palfrey (2016) *Quantal Response Equilibrium: A Stochastic Theory of Games,* Princeton, NJ: Princeton University Press.

Goeree, Jacob K., Charles A. Holt, and Thomas R. Palfrey (2020) "Stochastic Game Theory for Social Science: A Primer on Quantal Response Equilibrium," in M. Capra, R. Croson, T. Rosenblatt, and M. Rigdon eds. *The Handbook of Experimental Game Theory*, Edward Elgar: Cheltenham, pp. 8-47.

Iaryczower, Matias and Matthew Shum (2012) "The Value of Information in the Court. Get it Right, Keep it Tight," *American Economic Review*, 102 (February): 202-237.

Lewis, Jeffrey B. and Kenneth A. Schultz (2003) "Revealing Preferences: Empirical Estimation of a Crisis Bargaining Game with Incomplete Information," *Political Analysis*, 11(4): 345-365.

Lopez-Moctezuma, Gabriel (2019) "Sequential Deliberation in Collective Decision Making: The Case of the FOMC", *Working Paper*, California Institute of Technology.

McKelvey, Richard D. and Thomas R. Palfrey (1995) "Quantal Response Equilibria in Normal Form Games," *Games and Economic Behavior*, 10: 6-38.

Merlo, Antonio and Thomas R. Palfrey (2018), "External Validation of Voter Turnout Models by Concealed Parameter Recovery," *Public Choice*, 176 (1-2): 297-314.

Palacios-Huerta, Ignacio (2003) "Professionals Play Minimax," *Review of Economic Studies* 70:2, 395-415.

Powell, Robert (2007a) "Defending Against Terrorist Attacks with Limited Resources," *American Political Science Review*, 101(3): 527-541.

Powell, Robert (2007b) “Allocating Defensive Resources with Private Information about Vulnerability,” *American Political Science Review*, 101(4): 799-809.

Ramsay, Kristopher W. (2017) “Information, Uncertainty, and War,” *Annual Reviews of Political Science*, 20: 505-527.

Roberson, Brian (2006) “The Colonel Blotto Game,” *Economic Theory*, 29: 1-24.

Signorino, Curt (1999) “Strategic Interaction and the Statistical Analysis of International Conflict,” *American Political Science Review*. 93(2):279-98.

Signorino, Curt (2003) “Structure and Uncertainty in Discrete Choice Models,” *Political Analysis*. 11(4):316-344.

Sandler, Todd and Harvey E. Lapan (1988) “The Calculus of Dissent: An Analysis of Terrorists’ Choice of Targets,” *Synthese*. 76:245-261.

Walker, Mark, and John Wooders (2001) “Minimax Play at Wimbledon.” *American Economic Review*, *91*(5): 1521-1538.

For Online Publication: Supplemental Data Appendix

Average Attack and Defense Rates for Equal Absolute Increases in Attack Effectiveness

Session	First Treatment	Second Treatment	Attacker ID	Decision Rates $A = 2, B = 8$		Decision Rates $A = 6, B = 12$	
				a_1 rate	d_2 rate	a_1 rate	d_2 rate
prqr2	$A = 2, B = 8$	$A = 6, B = 12$	7	0.2	0	0.7	0.25
prqr2	$A = 2, B = 8$	$A = 6, B = 12$	8	0.15	0.6	1	0
prqr2	$A = 2, B = 8$	$A = 6, B = 12$	9	0.55	0.45	0.25	0.1
prqr2	$A = 2, B = 8$	$A = 6, B = 12$	10	0.1	0.65	0.1	0.75
prqr2	$A = 2, B = 8$	$A = 6, B = 12$	11	0.3	0.2	0.45	0.3
prqr2	$A = 2, B = 8$	$A = 6, B = 12$	12	0	0	0.45	0.2
prqr4	$A = 2, B = 8$	$A = 6, B = 12$	7	0.45	0.5	0.6	0.3
prqr4	$A = 2, B = 8$	$A = 6, B = 12$	8	0.35	0.55	0.95	0
prqr4	$A = 2, B = 8$	$A = 6, B = 12$	9	0.4	0.45	0.4	0.15
prqr4	$A = 2, B = 8$	$A = 6, B = 12$	10	0.55	0.6	0.95	0
prqr4	$A = 2, B = 8$	$A = 6, B = 12$	11	0.55	0.4	0.3	0.25
prqr4	$A = 2, B = 8$	$A = 6, B = 12$	12	0.45	0.35	0.6	0.3
prqs1	$A = 6, B = 12$	$A = 2, B = 8$	6	0.15	0.4	1	0
prqs1	$A = 6, B = 12$	$A = 2, B = 8$	7	0.40	0.45	0.65	0.1
prqs1	$A = 6, B = 12$	$A = 2, B = 8$	8	0.35	0.3	0.6	0.15
prqs1	$A = 6, B = 12$	$A = 2, B = 8$	9	0.45	0.3	0.6	0.3
prqs1	$A = 6, B = 12$	$A = 2, B = 8$	10	0.1	0.65	0.1	0.6
prqs3	$A = 6, B = 12$	$A = 2, B = 8$	8	0.4	0.05	0.95	0
prqs3	$A = 6, B = 12$	$A = 2, B = 8$	9	0.35	0.45	0.6	0.15
prqs3	$A = 6, B = 12$	$A = 2, B = 8$	10	0.3	0.45	0.5	0.1
prqs3	$A = 6, B = 12$	$A = 2, B = 8$	11	0.4	0.45	0.3	0.15
prqs3	$A = 6, B = 12$	$A = 2, B = 8$	12	0.3	0.65	0.65	0.2
prqs3	$A = 6, B = 12$	$A = 2, B = 8$	13	0.2	0.35	0.95	0
prqs3	$A = 6, B = 12$	$A = 2, B = 8$	14	0.45	0.6	0.55	0.2

Average Attack and Defense Rates for Equal Proportional Increases in Attack Effectiveness

Session	First Treatment	Second Treatment	Attacker ID	Decision Rates $A = 3, B = 6$		Decision Rates $A = 6, B = 12$	
				a_1 rate	d_2 rate	a_1 rate	d_2 rate
prqu4	$A = 3, B = 6$	$A = 6, B = 12$	7	0.5	0.55	0.65	0.2
prqu4	$A = 3, B = 6$	$A = 6, B = 12$	8	0.45	0.55	0.4	0.15
prqu4	$A = 3, B = 6$	$A = 6, B = 12$	9	0.45	0.85	0.5	0.2
prqu4	$A = 3, B = 6$	$A = 6, B = 12$	10	0.45	0.45	0.3	0.2
prqu4	$A = 3, B = 6$	$A = 6, B = 12$	11	0.15	0.8	0.7	0.1
prqu4	$A = 3, B = 6$	$A = 6, B = 12$	12	0.55	0.45	0.6	0.25
prqu6	$A = 3, B = 6$	$A = 6, B = 12$	7	0.55	0.85	1	0.05
prqu6	$A = 3, B = 6$	$A = 6, B = 12$	8	0.55	0.45	0.85	0.2
prqu6	$A = 3, B = 6$	$A = 6, B = 12$	9	0.65	0.45	0.55	0
prqu6	$A = 3, B = 6$	$A = 6, B = 12$	10	0.65	0.6	0.6	0.1
prqu6	$A = 3, B = 6$	$A = 6, B = 12$	11	0.55	0.75	0.2	0.1
prqu6	$A = 3, B = 6$	$A = 6, B = 12$	12	0.45	0.55	0.7	0.05
prqu2	$A = 6, B = 12$	$A = 3, B = 6$	7	0.4	0.5	0.75	0.3
prqu2	$A = 6, B = 12$	$A = 3, B = 6$	8	0.1	0.85	1	0.1
prqu2	$A = 6, B = 12$	$A = 3, B = 6$	9	0.75	0.85	0.65	0.35
prqu2	$A = 6, B = 12$	$A = 3, B = 6$	10	0.5	0.8	0.35	0.05
prqu2	$A = 6, B = 12$	$A = 3, B = 6$	11	0.55	0.8	0.5	0.4
prqu2	$A = 6, B = 12$	$A = 3, B = 6$	12	0.2	0.75	0.7	0.3
prqu5	$A = 6, B = 12$	$A = 3, B = 6$	7	0.5	0.7	0.9	0.15
prqu5	$A = 6, B = 12$	$A = 3, B = 6$	8	0.4	0.55	0.6	0.4
prqu5	$A = 6, B = 12$	$A = 3, B = 6$	9	0.65	0.65	0.7	0.1
prqu5	$A = 6, B = 12$	$A = 3, B = 6$	10	0.75	0.7	0.65	0.05
prqu5	$A = 6, B = 12$	$A = 3, B = 6$	11	0.6	0.65	0.4	0.2
prqu5	$A = 6, B = 12$	$A = 3, B = 6$	12	0.4	0.75	0.75	0.15

For Online Publication: Instructions Appendix

- **Rounds and Matchings**: The experiment consists of a number of **rounds**. Note: You will be matched with the **same** person in all rounds.
- **Interdependence:** Your earnings are determined by the decisions that you and the other person make.
- **Roles:** In each pair of people, one person will be designated as the "row" player and the other will be the "column" player. You will be a **column player** (or) **row player** in all rounds.

Continue with Instructions

- The column player will press either the **Left** or the **Right** button. The row player will choose **Top** or **Bottom**. These choices determine which part of the matrix is relevant (Top Left, Top Right, Bottom Left, Bottom Right). In each cell, the row player's payoff is shown in blue and the column player's payoff is shown in red.

Payoff Matrix (Row, Column)

	Left	**Right**
Top	**$0.80, $0.40**	**$0.40, $0.80**
Bottom	**$0.50, $0.70**	**$1.00, $0.20**

- If you are a row player, your decision buttons will be on the left side of the payoff table, and if you are a column player, your decision buttons will be above the table.

Continue

- **Matchings:** Please remember that you will be matched with the **same** person in all rounds.
- **Earnings:** Your earnings are determined by the choices that you and the other person make in the round. You begin with a fixed payment of **$0**, and earnings will be added to this amount (losses, if the game has negative payoffs, will be subtracted). Your total earnings will be displayed in a cumulative earnings column on the page that follows.
- **Rounds:** There will be **20 rounds** in this part of the experiment, and you are matched with the same person in all rounds.