Risk Analysis, Vol. 36, No. 8, 2016 DOI: 10.1111/risa.12547

Attacker Deterrence and Perceived Risk in a Stackelberg Security Game

Garret Ridinger,[1,2] Richard S. John,[3] Michael McBride,[1,2,*] and Nicholas Scurich[4,5,6]

In Stackelberg security games, a defender must allocate scarce resources to defend against a potential attacker. The optimal defense involves the randomization of scarce security resources, yet how attackers perceive the risk given randomized defense is not well understood. We conducted an experiment where attackers chose whether to attack or not attack targets protected by randomized defense schemes, the key treatment variable being whether the defender picks one target at random to guard or imperfectly guards all targets. The two schemes are expected-payoff equivalent, and when provided separately we found no effect of having one scheme or the other. Yet, when both are present, we found that subjects had a preference for the fixed scheme, a preference that cannot be reduced to differences in beliefs. Overall, our results suggest that understanding how individuals perceive risk is vital to understand the behavior of attackers.

KEY WORDS: Attack; perception; security

1. INTRODUCTION

In Stackelberg security games, a defender must allocate scarce resources to defend multiple targets against a potential attacker, and the attacker then chooses which target to attack.(1) The optimal defense generally randomizes resources across the targets to avoid being exploited by potential attackers under the assumption that attackers are risk neutral.[7] However, recent research finds that human subject attackers select options that fail to maximize their expected value, thereby allowing defenders to exploit this fact in implementing their strategies. (5–7) In order to select the best defense strategy, it is important for defenders of real-world targets to understand how attackers perceive risk.

This article uses a laboratory experiment to study the perceived risk of attackers in security games. Similar to Yang *et al.*,(7) each subject takes the role of attacker and decides which of five potentially guarded doors to attack or not to attack any of them. Both the number of guards and the potential payoffs were randomized across rounds by a computer. Our primary treatment variables are the number of guards (one or two) and two guarding schemes. The guarding schemes differ only in how they implemented the randomized strategy, allowing us to investigate the role of perceived risk

[1]Department of Economics, University of California, Irvine, CA, USA.
[2]Experimental Social Science Laboratory, University of California, Irvine, CA, USA.
[3]Department of Psychology, University of Southern California, Los Angeles, CA, USA.
[4]Department of Psychology and Social Behavior, University of California, Irvine, CA, USA.
[5]Department of Criminology, Law, and Society, University of California, Irvine, CA, USA.
[6]School of Law, University of California, Irvine, CA, USA.
*Address correspondence to Michael McBride, 3151 Social Science Plaza, Department of Economics, University of California, Irvine, California, 92697-5100, USA; tel: 949-824-7417; mcbride@uci.edu.

[7]Theoretical work on Stackelberg security games has important policy applications. For example, recent decision-making software has helped aid in allocating limited resources at the Los Angeles International Airport,(2) the U.S. Federal Air Marshals Service,(3) the U.S. Transportation Security,(4) and the U.S. Coast Guard.(5)

in attacker decision making. In Scheme F (fixed), individual guards randomly select a single door to guard. In Scheme R (rotating), individual guards randomly patrol all doors. In all treatment conditions, we randomly varied the cost and reward of attacking.

We find that subjects respond qualitatively as predicted to changes in both the payoffs and the probability of being caught, and that risk preferences were important in explaining deterrence. Deterrence rates (i.e., the proportion of times subjects chose not to attack any door) were similar in Scheme F and Scheme R, suggesting that subjects perceived the risks similarly in both schemes. However, when subjects could choose which scheme to attack, there was a clear preference for Scheme F. We ran two additional sessions to test the robustness of this result and identify potential causes. In one, we explicitly asked the subjects to report probabilistic beliefs they would get caught, thus allowing us to identify if the preference for F was due to a misperception in risk. In the other session, we changed the wording of the descriptions of Schemes F and R to check if the preference for F was due to an unintentional framing in our original script. We find that differences in beliefs across the schemes cannot fully explain the preference for F over R, and that this finding was robust to wording changes in the script. In short, our research identifies a behavioral regularity not easily explained by standard behavioral theories.

There are many potential ways that attacker-defender games can be modeled. Substantial focus on attacker and defender behavior in security games has centered on the Colonel Blotto game.[8–10] The attacker and defender each allocate resources over a number of battlefields simultaneously, and the resource allocation by both players influences the probability that one or the other wins a given battlefield. The Colonel Blotto game has been used to examine a number of security issues, including: suicide terrorism,[11] jamming radio communications,[12] and phishing attacks.[13] Experimental tests where players have unequal resources find evidence that subjects are sensitive to the asymmetry and the results show some support for the theoretical predictions.[14,15] It is important to note that the results from Chowdhury *et al.*[15] showed substantial variance in subject choices, suggesting that expected value maximization may not be the appropriate model for human behavior in these games. It is often assumed that players in Colonel Blotto games make choices simultaneously.[8] This is a key difference between the Blotto game literature and our article as we focus on instances where the defender commits to the strategy prior to the attacker's decision. Our study thus sheds light on situations where a potential criminal or terrorist may conduct surveillance to learn the defenders' strategy prior to making an attack choice.

Our research complements previous work that accounts for behavioral anomalies in Stackelberg security games. Pita *et al.*[6] introduced the Combined Observability and Bounded Rationality Assumption (COBRA) model, which assumes that subjects have an anchoring bias when evaluating probabilities and allows a chance that subjects will deviate from best responses. Yang *et al.*[7] introduce additional behavioral models that include prospect theory[17] and Quantal response.[18] Experimental results found that their Best Response to Quantal Response (BRQR) model preformed best overall compared to competing models like COBRA.[6] A recent real-world application assumed that potential attackers behave according to a Quantal Response (QR) model. [5] In the model, potential attackers can make mistakes when choosing best responses. Results revealed how defenders could improve their strategies by taking advantage of attackers' mistakes. This line of research shows that improved understanding of attacker decision making can lead to better expected outcomes for defenders.

The experiment closest to ours is Yang *et al.*,[7] which draws inspiration from the security problem at the Los Angeles International Airport. [2] Each subject played the role of attacker and could choose to enter one of eight doors (i.e., gates at the airport). Each door had different rewards and penalties as well as probabilities that the door was guarded. Our experiment differs from Yang *et al.*[7] in that subjects can always choose to not attack any of the doors. This allows us to investigate conditions under which subjects will be deterred from attacking. Our experiment also differs in that we examine how changes in the implementation of the randomized strategies influence attacker behavior.

While our objective is to examine deterrence and perceived risk in Stackelberg security games,

[8]A notable exception is Powell,[16] who analyzed a sequential Blotto game where a defender allocates resources prior to being attacked. In the subgame perfect equilibrium the defender minmaxes the attacker and the attacker chooses a best response that minimizes the defender's expected loss.

we focus exclusively on attacker behavior. [9] Similar to Yang *et al.*,[7] subjects in our experiment are making decisions under risk. In our experiment, attackers' decisions are in the context of a Stackelberg security game, which could differ from standard risk experiments where subjects select among a set of alternative gambles. Recent experiments examining deterrence and risk include Friesen[26] and DeAngelo and Charness,[27] Motivated by tax compliance, Friesen[26] found that subjects were more deterred by the severity of punishment compared to the probability of being caught. In DeAngelo and Charness,[27] subjects were asked to choose whether to speed or not. The authors found that deterrence increased in a regime with a high probability of being caught and a low fine compared to a regime with a low probability of being caught and a high fine. When subjects could vote for their preferred regime, they were less likely to be deterred if their favored regime was implemented. Instead of voting, our experiment includes treatments where subjects can select which guarding scheme they would like to attack. In addition, our schemes keep the probability and costs of being caught the same, allowing us to investigate the influence of different guarding schemes on the risk perception of potential attackers.

Our article contributes to this line of research by using a controlled laboratory experiment to investigate how attackers respond to changes in payoffs and risk.[10] We recognize that the university students used as human subjects in our experiments may not manifest all the characteristics of real-life attackers. Nor do the relatively low stakes used in the experiment capture the high stakes in real-life attacker-defender settings. It is, of course, impossible to capture all features of real life for ethical and practical reasons. However, it is hoped that our findings will ultimately provide some practical benefit akin to those found in other experimental studies of attacker-defender games such as Shieh *et al.*[5] The applicability of our findings will depend on the extent to which the observed behavioral peculiarities are common to all humans.[11]

2. EXPERIMENTAL DESIGN

The experiment was programmed and conducted with the z-Tree experimental software package.[32] In each round, each subject was given an endowment of $8, told how many guards were at work, the monetary reward for success, and the fine for getting caught. Each subject then chose whether to attack or not attack. Attacking requires the subject to pay a cost of $3, which reflects the effort and resources put into planning and carrying out an attack. If a subject chose to attack, then she had to select one of five doors to enter. If the door attacked by the subject was not guarded, then the subject received a monetary reward. If the attacked door was guarded, then the subject incurred a punishment fine. At the end of the round, the subject was told the outcome of

[9] Recent research has considered the defender's behavior in sequential choice in security games. Bier *et al.*[19] found that defenders maximize deterrence by centrally allocating resources and publicly revealing their defensive allocations. Dighe *et al.*[20] show deterrence can be less costly using partial disclosure of defenses. Berman and Gavious[21] analyzed the optimal placement of resources in a network to respond to potential acts of terrorism, while Zhuang *et al.*[22] show in a multiperiod game that defenders can provide more cost-effective security by utilizing both secrecy and deception. Bier *et al.*[23] examine optimal allocation of defensive resources computationally. Shan and Zhuang[24] study credible and noncredible retaliation in preventing the smuggling of nuclear weapons, and Hausken and Zhuang[25] examine deterrence of terrorist attacks in a dynamic game. Our study focuses on how attackers respond rather than on the defender's optimal strategy.

[10] Although not investigating attacker behavior, Scurich and John[28] used survey evidence to examine public perception of two types of randomization schemes implemented at airport security. With the probability of detection the same, subjects were presented with a traditional scheme where everyone was searched and a randomized scheme were people were randomly selected to be searched . Participants rated the traditional scheme as fairer and safer, but less convenient compared to the randomized scheme. In this article, we show that how the randomization is implemented could be an important aspect in understanding potential attacker behavior.

[11] There are reasons to believe, however, that our subjects provide a meaningful sample. First, there is reason to believe that our student subjects have risk preferences sufficiently similar to other groups. Recent research on a criminal population, for example, has found that inmates respond similarly to risk and deterrence as student populations, although there is a higher percentage of inmates who tend to be risk seeking compared to students.[29] While the criminal population is likely to differ from terrorist populations, this result suggests that research in regards to risk and decision making may have applicability in other populations thought to differ greatly compared to our student population. Second, while individual data on terrorist decision making are scarce, evidence seems to suggest that planners of attacks take into account the expected costs and benefits in their decisions.[30,31] This suggests that actual attackers would respond to risks in similar ways as other subjects. Third, although our experiment lacks mundane realism, it does have high experimental realism. Subjects are highly involved in the task and are motivated to perform the task well.

that round. A new round would then begin, with 40 rounds in total.

The reward and punishment cost varied each round. In the High Stakes condition, the reward for a successful attack was \$10 and the punishment cost if caught was \$5. Under the Low Stakes condition, the reward for a successful attack was \$5 and the punishment if caught was \$2.50. These potential payments are revealed to the subject at the beginning of each round, so the subject understands the stakes when deciding to attack or not attack.

The probability of getting caught depended on how many guards were at work. Under the One Guard condition, a single guard was responsible for all five doors but had to split efforts across them. This corresponds to a 20% chance of getting caught if choosing to attack. With Two Guards, two guards were responsible for all five doors, and the corresponding chance of getting caught is 40%. The number of guards is revealed to the subject before the subject makes the choice for that round. If the subject chooses attack, the computer software uses a random number generator to select whether the subject was successful or caught, and the result is then revealed to the subject.

Thus, a single round consisted of one of four possible stakes-guards combinations (High Stakes-One Guard, High Stakes-Two Guards, Low Stakes-One Guard, Low Stakes-Two Guards). Each subject participated in 10 rounds for each of these combinations, the order being randomized at the subject level. Fig. 1(a) shows the decision tree for subject and Fig. 1(b) shows the expected values of the subject choices based on the different conditions.

This setup is admittedly oversimplified but captures the essence of many real-world attacker-defender setting. For example, at airports and ports and harbors, limited security personnel are tasked with ensuring safety at multiple potential targets (e.g., different gates at an airport), and those personnel must decide at which targets they will provide defensive presence. Indeed, predicted behavior in attacker-defender models such as the one we consider here are used to develop the optimal defensive strategies used for the Los Angeles International Airport, the U.S. Federal Air Marshals, the U.S. Coast Guard, and more (see footnote 1).

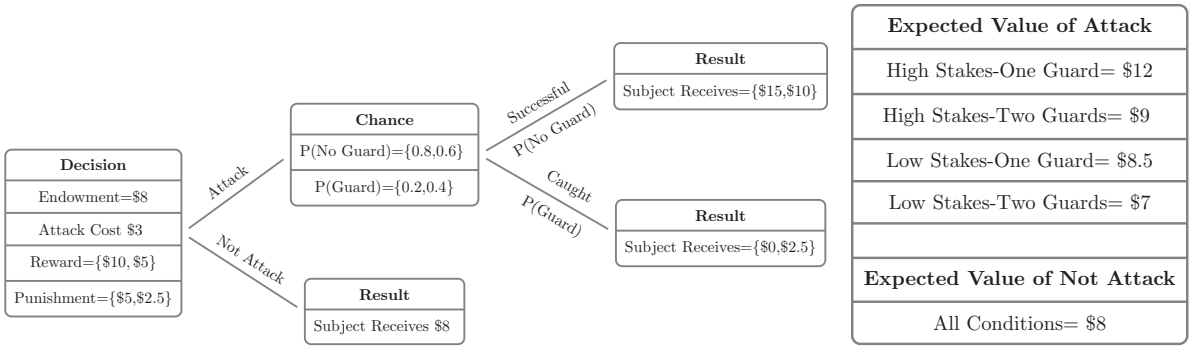
While this stakes-guards combination was varied within subject, the guarding scheme varied between subjects; that is, a subject only ever faced a single guarding scheme treatment condition. We consider three primary treatments, with two additional treatments for robustness. The three primary treatments were: Treatment F where subjects chose to attack Scheme F (fixed) or not attack, Treatment R where subjects chose to attack Scheme R (rotating) or not attack, and Treatment FR where subjects chose to attack Scheme F, attack Scheme R, or not attack. In any given round, the probability of getting caught under either Scheme F or Scheme R was identical and determined solely by the number of guards.

Under Scheme F, the guards were randomly assigned to guard exactly one of the five doors. The guards stayed at that door the entire time. If there was one guard, then one of the five doors was guarded. If two guards, then two of the five doors were guarded. Under Scheme R, guards randomly selected which doors to guard. The guards spent an equal amount of time guarding each door, and never guarded the same door at the same time. In both schemes, if a subject chose to attack and the door they selected was guarded, then they were caught and had to pay the punishment cost.

We also conducted two additional treatments to further investigate the behavior observed in the primary treatments. In the Treatment FR with Beliefs, after a subject made her attack choice, she was asked to select a door whether choosing attack or not attack. If the subject chose not to attack, then the subject was prompted to select the door she would have selected had she chosen to attack. After the door selection, but before the subject learned whether that door was guarded or not, she was asked the percent chance that she believed that the selection of that door would have been successful if she had attacked Scheme F and if she had attacked Scheme R. Both stated beliefs were incentivized using a proper scoring rule where the beliefs of an individual subject i are β_i^k, where $k \in \{F, R\}$ represents the belief of a successful attack for a given scheme. If the attack was successful, then subjects received $\$1 - \$1 * (1 - \beta_i^k)^2$. If caught, then subjects received $\$1 - \$1 * (\beta_i^k)^2$. To ensure understanding of the rule, subjects were presented with a table that gave numerical examples of potential payoffs that subjects would have received based on their stated belief and outcome.[12] The purpose of this treatment condition was to obtain data that could be correlated with choices to determine whether individual perceived risk predicts behavior.

For our second robustness treatment, called Treatment FR with Time, the descriptions of the

[12]Screenshots of the table viewed by subjects can be found in the Supporting Information.



(a) Subject Decision Tree (b) Expected Value by Condition

Fig. 1. Experimental parameters and conditions.

schemes were changed to give an explicit timing frame. If subjects chose to attack Scheme F or attack Scheme R, then they had 10 seconds to choose a door. In Scheme F, guards were randomly assigned to guard one of the five doors and stayed at the door for the entire 10 seconds. In Scheme R, guards visited every door once but in an order randomly determined by the computer such that they never guarded the same door at the same time. Each guard spent 2 seconds guarding each door and it took them zero seconds to switch to a different door. Subjects were caught based on whether the door was guarded at the exact time they selected that door. The purpose of this treatment was to test whether behavior found in Treatment FR was due to misunderstanding in the original FR script. The explicit timing description adds enhanced clarity.

To measure risk preferences we follow Caldara,[33] who modified the approach used by Eckel and Grossman.[34] Subjects were asked to select one of three lottery choices: (1) 100% chance to receive \$5, (2) 50% chance to receive \$7.00 and 50% chance to receive \$2.80, or (3) 50% chance to receive \$6.05 and 50% chance to receive \$4.00. If subjects exhibit constant relative risk aversion (CRRA), then subjects who have a CRRA coefficient: greater than 0.24 should choose option (1), less than −0.21 should prefer option (2), and between 0.24 and −0.21 should select option (2).[13] Subjects who chose option (1) were classified as risk averse. Subjects who chose option (2) were classified as risk loving. Subjects who chose option (3) were classified as risk neutral. We acknowledge that our type classification is imperfect as individuals who are slightly risk averse or risk loving could be classified as risk neutral. Despite this caveat, we show that our measure of risk preference correlates with subject behavior in the expected manner.

In all treatments, one round out of the 40 was chosen randomly for payment at the end of the experiment. The random payment of a single round helps to mitigate potential endowment effects that may link behavior across rounds and better captures the notion of one-shot decisions. At the end of each round, subjects were told the results and the payment they would receive if the current round were randomly selected at the end of the period. Subjects were not told how other players behaved in the experiment. For all previous rounds, subjects could view whether they had been caught and the payoff earned. After subjects finished the 40 rounds, but before they learned which round was selected for payment, subjects participated in the risk elicitation procedure. Subjects then completed a questionnaire that asked demographic questions.

Screenshots of the experiment software and instructions are provided in the Supporting Information available from the authors.

3. RESULTS

Our experiment was conducted with 300 students at the Experimental Social Science Laboratory (ESSL) at the University of California, Irvine, a large public university. Each subject participated in one of 10 experimental sessions that took place in a computer laboratory. Prior to each experimental session, students were randomly selected from the laboratory's subject pool and sent an email about the

[13]The CRRA values for the risk preference categories differ from Caldara,[33] who uses greater than 0.4 for risk averse and less than −0.4 for risk seeking.

Table I. Treatment Summary Statistics

	Not Attack	Attack Scheme F	Attack Scheme R	Number of Subjects	Average Earnings ($)
Treatment F	0.46	0.54		58	20.83
Treatment R	0.45		0.55	68	20.34
Treatment FR	0.43	0.41	0.17	60	20.16
Treatment FR with Beliefs	0.42	0.39	0.19	54	22.16
Treatment FR with Time	0.39	0.43	0.19	60	21.90

upcoming session. Students then signed up for the session via the laboratory's website. After arriving for the experiment, each subject was randomly assigned to a computer terminal. The subject then received the instruction via computer display and made choices using the mouse. After all subjects had completed their decisions, they lined up to receive their monetary earnings. All sessions lasted 40–60 minutes.[14] No subject participated in more than one session. In addition to earnings based on decisions described in the prior section, subjects were each given a $7 show-up payment.

Table I gives summary statistics for the different treatments. Overall, subjects were more likely to choose to attack Scheme F compared to Scheme R across all treatments. The minimum payment earned by subjects was $10 and the maximum payment was $31. The average payment across all treatments was $21.04. Table II provides summary statistics for independent variables that will be used in subsequent regression analysis. The variable Expected Payoff Max Not Attack is a dummy variable equal to 1 if choosing not attack maximizes expected value. This occurs in the Low Stakes-Two Guards condition. The variable Expected Payoff Difference is equal to the difference in expected value between attacking and not attacking.[15] The vast majority of subjects were classified as either risk neutral or risk averse by our risk elicitation procedure; only 18% of subjects were classified as risk seeking. The variables Scheme F and Scheme R belief are the subject's reported beliefs about the success of attacking Scheme F and Scheme R, respectively. The dummy variable Scheme F (R) more likely is equal to one if subjects reported that Scheme F (R) was more likely to be successful. If subjects reported that the likelihood of success was the same for both schemes, then the dummy variable Scheme F and R Equally Likely is equal to one. The variable Correct Beliefs is a dummy variable equal to one if subjects reported correct beliefs for both schemes. Overall, subjects on average reported a lower likelihood of success in attacking Scheme R compared to attacking Scheme F.

Table II. Summary Statistics: Independent Variables

	Mean	*SD*	Min	Max
Condition and Risk Variables				
High Stakes	0.50	0.50	0	1
Two Guards	0.50	0.50	0	1
Expected Payoff Max Not Attack	0.25	0.43	0	1
Expected Payoff Difference	1.38	1.85	−1	4
Risk Averse	0.38	0.49	0	1
Risk Neutral	0.44	0.50	0	1
Risk Loving	0.18	0.39	0	1
Self-Reported Variables				
Female	0.58	0.49	0	1
Age	20.30	2.10	17	30
Belief Variables				
Scheme R Belief	65.19	26.63	0	100
Scheme F Belief	70.24	25.38	0	100
Scheme R Belief One Guard	69.38	25.77	0	100
Scheme F Belief One Guard	75.76	23.79	0	100
Scheme R Belief Two Guards	60.97	26.82	0	100
Scheme F Belief Two Guards	64.66	25.71	0	100
Scheme R Belief More Likely	0.14	0.35	0	1
Scheme F Belief More Likely	0.32	0.46	0	1
Scheme F and R Equally Likely	0.54	0.50	0	1
Correct Beliefs	0.14	0.34	0	1

Result 1. *As predicted, deterrence is positively associated with risk aversion, negatively associated with risk loving, and positively associated with expected payoff.*

Table III gives results of Probit regressions examining the relationship between deterrence and

[14] Economics experiments typically last 1–1.5 hours and involve a similar number of decision. Boredom and fatigue can occur in these types of experiments, but we believe were less likely in our experiment because it was shorter (under an hour), and the decisions were fairly simple.

[15] For example, in the one guard and high stakes condition the expected value of attacking is $12. Since subjects get $8 for not attacking, Expected Payoff Difference takes on a value of 4 in this case.

Table III. Predicting Choice to Not Attack by Expected Payoffs and Risk

	(1) Not Attack	(2) Not Attack	(3) Not Attack	(4) Not Attack	(5) Not Attack	(6) Not Attack
Expected Payoff Max Not Attack	1.18***	0.51***	0.52***	0.52***	0.70***	1.24***
	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)	(0.09)
Expected Payoff Difference		−0.22***	−0.23***	−0.23***	−0.30***	−0.51***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Risk Averse			0.24***	0.23***	0.31*	
			(0.03)	(0.03)	(0.12)	
Risk Loving			−0.34***	−0.35***	−0.47**	
			(0.04)	(0.04)	(0.16)	
Female				−0.09***	−0.10	
				(0.03)	(0.11)	
Age				0.03***	0.05+	
				(0.01)	(0.03)	
Round				0.01***	0.01***	0.01***
				(0.00)	(0.00)	(0.00)
Intercept	−0.48***	−0.03	−0.06*	−0.82***	−1.17*	
	(0.01)	(0.02)	(0.03)	(0.13)	(0.57)	
N	12,000	12,000	12,000	12,000	12,000	11,600
Pseudo R^2	0.11	0.14	0.16	0.16		0.26
ρ					0.46	
Model χ^2	1,829.24	2,311.39	2,587.43	2,671.80	2,456.11	3,138.28

Notes: Standard errors in parentheses. Columns (1)–(5) are Probit regressions. Column (6) is a random effects Probit regression. Column (7) is a fixed effects Logit regression. $^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

risk. The Probit model is a latent variable model that is used for binary dependent variables assuming the error term follows a normal distribution.[35] Subjects who maximize expected payoffs should only choose to not attack in the two guard and low stakes condition. Results from column (1) show that deterrence was higher in the two guard and low stakes condition. This result, though not surprising, is an important benchmark; because subjects responded to the rewards and costs in predictable manners, the decision setting is one in which we can plausibly study both perception of risk and decisions under risk. However, the pseudo R^2 is low, suggesting that expected payoff maximization is not the whole story. Depending on the condition, the difference in expected payoff between attacking and not attacking changes. Column (2) includes a variable that captures the differences in expected payoffs between not attacking and attacking. Both Expected Payoff Max Not Attack and Expected Payoff Difference are significant. In addition, the overall fit of the regression improved, with the pseudo R^2 increasing from 0.11 to 0.14.

Column (3) includes the elicited risk preference. The estimated coefficients for risk-averse and risk-loving subjects have the expected signs, i.e., risk-loving (averse) subjects are less (more) likely to be deterred from attacking. Column (4) adds additional controls where the coefficients for female, age, and round are all significant.

Because subjects made repeated choices, it is unlikely that the independence assumption will be met. We confront this issue by using both random effects Probit and fixed effects Logit. The random effects Probit model allows us to account for random individual heterogeneity among subjects assuming there is no correlation between the individual error term and the independent variables. Column (5) gives the results using a random effects Probit regression. Except for the coefficient on round, the additional controls are no longer significant. Column (6) shows the results of a fixed effects Logit regression. Due to differencing the data, using the fixed effects Logit model allows us to control for any individual subject fixed effects that may be correlated with the independent variables. It is clear that, even after controlling for individual fixed heterogeneity, the subjects were influenced by both expected value maximization as well as changes in payoff differences.

In short, while subjects responded to changes in the expected payoffs, additional factors like the

Table IV. Fraction of Subjects Consistent with Expected Value Maximization

	Percent of Consistent Choices					
	100%	90%	80%	70%	60%	50%
Treatment F	0.00	0.12	0.31	0.60	0.79	0.86
Treatment R	0.00	0.10	0.31	0.52	0.68	0.91
Treatment FR	0.02	0.08	0.23	0.57	0.78	0.93
Treatment FR with Beliefs	0.00	0.02	0.19	0.55	0.81	0.94
Treatment FR with Time	0.00	0.12	0.33	0.72	0.83	0.87
Total	0.00	0.09	0.28	0.59	0.78	0.90

difference in payoffs and risk preferences also matter. This final point can be seen clearly in Table IV, which shows the fraction of subjects who are consistent with expected payoff maximization. For only 1 out of 300 subjects were all of the subject's choices consistent with expected payoff maximization.

Result 2. *The overall rate of deterrence was not influenced by the treatment schemes.*

Table V shows the effect of the treatment variables on deterrence. In the first column, both Treatment F and Treatment FR with Time are statistically different from Treatment FR, yet this significance is lost in column (2) when using a random effects Probit. In column (3), additional controls were added to the pooled Probit model. Column (4) shows that the significant difference of the treatment variables is lost when using a random effects Probit. The estimates indicate that overall deterrence was similar across the different treatments. Though perhaps unsurprising, this is a key result of the article and forms an important benchmark for comparison with the rest of the results.

One interest is whether subjects behaved differently in Treatment F compared to Treatment R. We fail to reject the hypotheses that deterrence rates were the same between the two treatments for each of the four conditions. Nonparametric results are as follows: (1) One Guard and High Stakes (Wilcoxon rank-sum test, $z = -0.418$, $p = 0.676$), (2) One Guard and Low Stakes (Wilcoxon rank-sum test, $z = 0.699$, $p = 0.485$), (3) Two Guards and High Stakes (Wilcoxon rank-sum test, $z = -0.478$, $p = 0.633$), and (4) Two Guards and Low Stakes (Wilcoxon signed-sum test, $z = 1.238$, $p = 0.216$).[16] Overall, deterrence was similar between Treatments F and R.

[16] All p-values have been rounded to the third decimal place. Unless otherwise noted, all tests are two-sided.

Adding additional alternatives could potentially influence subject choices. While overall deterrence rates are similar between Treatment F and Treatment FR, there was a statistical difference at the 10% level between the two treatments in the One Guard and High Stakes condition (Wilcoxon rank-sum test, $z = 1.893$, $p = 0.058$). No statistically significant difference was found in the remaining conditions. When comparing deterrence rates between Treatment R and Treatment FR, none of the conditions were significantly different from each other.

Result 3. *Under Treatment FR in which both F and R are available, subjects were more likely to attack Scheme F than Scheme R.*

Result 3 is a striking result, and the first result that was unanticipated. Under Treatment FR, there is on average a preference to attack Scheme F over Scheme R when choosing to attack. Looking at Fig. 2 we can see that this preference held over the three different FR treatments as well. Since subjects made repeated choices, statistical tests on the pooled data violate the independence assumption. To deal with this we let the subject averages be our unit of observation. In Treatment FR, subjects were overall more likely to choose to attack Scheme F compared to Scheme R (Wilcoxon signed-rank test, $z = 4.610$, $p = 0.000$).

In addition, from Fig. 3 we can see that Scheme F was chosen more often than Scheme R in all four conditions; that is, although Schemes F and R have the same probability of success, subjects prefer Scheme F when given the choice. Using subject averages as our unit of observation, we test whether subjects were more likely to attack Scheme F compared to Scheme R across the four conditions. The results of the nonparametric tests for Treatment FR are as follows: (1) One Guard and High Stakes (Wilcoxon signed-rank test, $z = 4.852$, $p = 0.000$), (2) One Guard and Low

Table V. Predicting Choice to Not Attack by Treatment

	(1) Not Attack	(2) Not Attack	(3) Not Attack	(4) Not Attack
Treatment F	0.08*	0.13	0.07+	0.11
	(0.04)	(0.14)	(0.04)	(0.18)
Treatment R	0.06	0.08	0.10**	0.14
	(0.04)	(0.14)	(0.04)	(0.17)
Treatment FR with Beliefs	−0.02	−0.05	−0.01	−0.05
	(0.04)	(0.15)	(0.04)	(0.18)
Treatment FR with Time	−0.10**	−0.10	−0.09*	−0.12
	(0.04)	(0.14)	(0.04)	(0.18)
Expected Payoff Max			0.52***	0.70***
Not Attack			(0.04)	(0.05)
Expected Payoff			−0.23***	−0.30***
Difference			(0.01)	(0.01)
Risk Averse			0.24***	0.33**
			(0.03)	(0.13)
Risk Loving			−0.33***	−0.44**
			(0.04)	(0.16)
Female			−0.08**	−0.09
			(0.03)	(0.12)
Age			0.04***	0.05*
			(0.01)	(0.03)
Round			0.01***	0.01***
			(0.00)	(0.00)
Intercept	−0.19***	−0.24*	−0.89***	−1.27*
	(0.03)	(0.10)	(0.13)	(0.57)
N	12,000	12,000	12,000	12,000
Pseudo R^2	0.00		0.16	
ρ		0.36		0.46
χ^2	29.89	3.46	2,700.96	2,457.35

Notes: Standard errors in parentheses. Columns (1) and (3) are Probit regressions. Columns (2) and (4) are random effect Probit regressions. $^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Stakes (Wilcoxon signed-rank test, $z = 3.629$, $p = 0.000$), (3) Two Guards and High Stakes (Wilcoxon signed-rank test, $z = 2.446$, $p = 0.015$), and (4) Two Guards and Low Stakes (Wilcoxon signed-rank test, $z = 1.747$, $p = 0.081$). While the Two Guards-Low Stakes condition was only significant at the 10% level, the number of subjects choosing to attack was low. Scheme F was statistically more likely to be chosen across all conditions.

Result 4. *Under Treatment FR with Beliefs: (a) if a subject assigns different beliefs to Scheme F and Scheme R, then that subject is more likely to attack the scheme with higher expected payoff; and (b) subjects who assign the same beliefs to Schemes F and R are more likely to attack Scheme F than Scheme R.*

One reason that subjects may prefer Scheme F over Scheme R could be due to differing beliefs about the probabilities of being caught. In Treatment FR with Beliefs, subjects reported their beliefs about the success of attacking Scheme R and Scheme F. Subjects indicated that attacking Scheme F was more likely to be successful compared to Scheme R in the One Guard condition (Wilcoxon signed-rank test, $z = 3.725$, $p = 0.000$) and the Two Guard condition (Wilcoxon signed-rank test, $z = 2.452$, $p = 0.014$).

One potential issue is that eliciting beliefs can potentially change subject behavior. Comparing the two treatments, we fail to reject the hypotheses that the distributions of subjects choosing not attack (Wilcoxon rank-sum test, $z = -0.014$, $p = 0.989$), attack Scheme F (Wilcoxon rank-sum test, $z = 0.338$, $p = 0.735$), and attack Scheme R (Wilcoxon rank-sum test, $z = -0.103$, $p = 0.918$) are the same.

Fig. 4 shows the cumulative distribution of subject stated beliefs about being successful if attacking Scheme F and Scheme R by the number of guards. It is clear that there is a wide range of reported

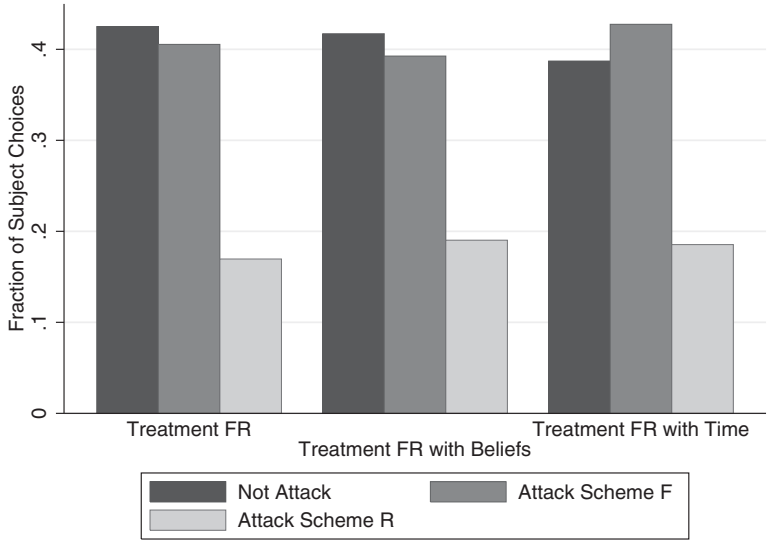


Fig. 2. Average attack choice by treatment.

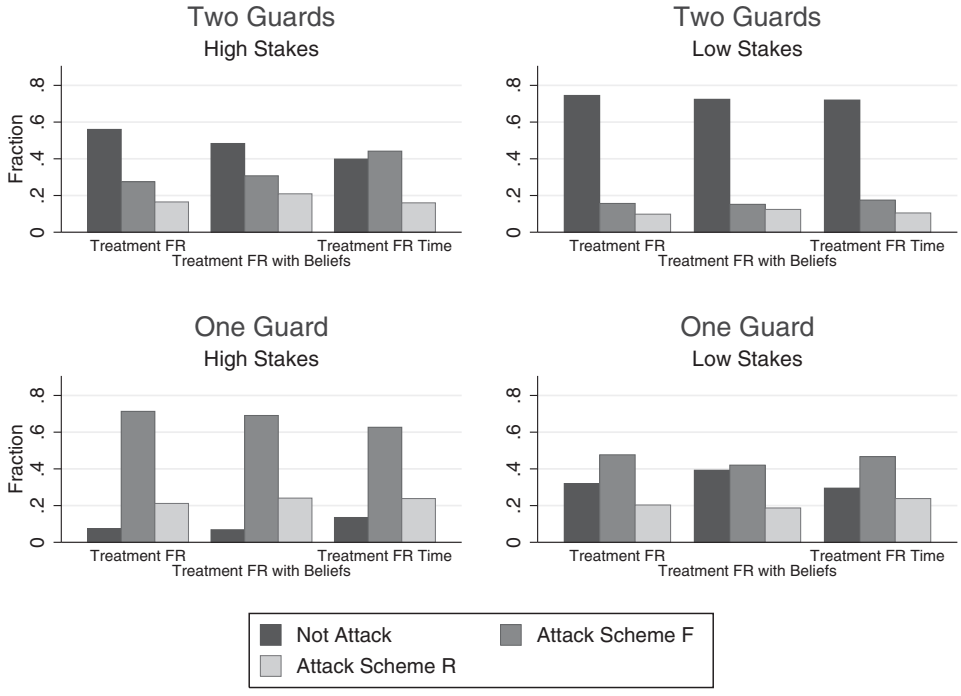


Fig. 3. Average attack choice by treatment and condition.

beliefs for both schemes and guard conditions. Overall, subjects indicated that an attack in Scheme F is more likely to be successful compared to an attack in Scheme R. Despite this, there was a great deal of heterogeneity in subject reported beliefs. [17]

Table VI gives results from a multinomial Probit regression predicting subjects' choices. The multinomial Probit framework is designed for settings in which there are three or more alternatives and those

[17]One potential worry is that some subjects reported a success probability of 20% and 40% which is the correct probability of failure. It could be that subjects misunderstood what they were asked. While this is a concern, the results in the article are robust to removing these subjects from the data.

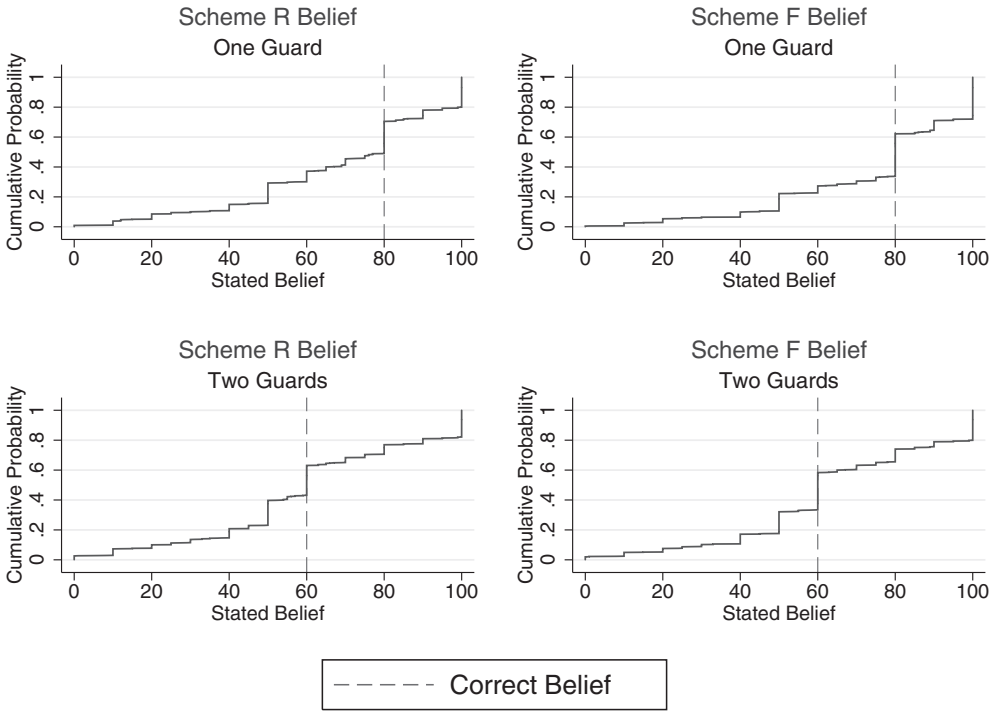


Fig. 4. Cumulative distribution of subjects stated beliefs of success by scheme and number of guards.

Table VI. Predicting Attack Choices by Beliefs: Multinomial Probit Regression

	(1) Not Attack	(2) Attack Scheme F
Scheme R Belief	−0.02***	−0.03***
	(0.00)	(0.00)
Scheme F Belief	0.02***	0.04***
	(0.00)	(0.00)
Expected Payoff	0.12	−0.07
	(0.17)	(0.17)
Expected Payoff Difference	−0.36***	0.09*
	(0.04)	(0.04)
Risk Averse	0.35**	0.04
	(0.12)	(0.12)
Risk Loving	−0.41**	−0.26*
	(0.13)	(0.13)
Female	0.13	0.21+
	(0.12)	(0.11)
Age	0.08**	0.01
	(0.03)	(0.03)
Round	0.01*	−0.00
	(0.00)	(0.00)
Intercept	−0.89	−0.33
	(0.61)	(0.61)
Predicted Probability	0.41	0.40

Notes: Standard errors in parentheses. The base outcome variable is Attack Scheme R. + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

alternatives do not have a natural ordering,[35] such as the case in Treatment FR where subjects can choose to attack F, attack R, or not attack. The results in both columns are compared to the base variable attack Scheme R. Looking at column (2) it is clear that subjects who reported higher Scheme F beliefs of success were more likely to attack Scheme F compared to Scheme R. Similarly, higher Scheme R beliefs were associated with a lower probability of attacking Scheme F.[18] The predicted probabilities of the model are similar to the actual probabilities (see Supporting Information), suggesting that the multinomial Probit regression estimates fit the data reasonably well.

While beliefs seem to influence the choices that subjects make, even when subjects reported the correct probability for both Scheme F and Scheme R, subjects still prefer to attack Scheme F. This can be seen in Fig. 5, which shows the average attack choices broken down by stated beliefs. Not surprisingly, subjects who indicated that Scheme F (R) was more likely to be successful were more likely to

[18]The regression results hold for a number of additional belief measures including using squared deviations from the correct probability, and differences between the two stated beliefs.

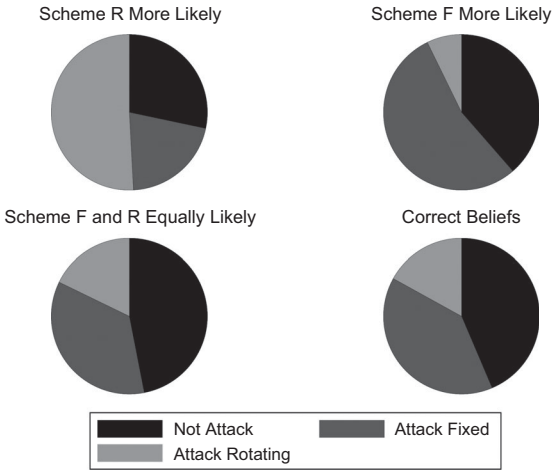


Fig. 5. Attack choices by stated beliefs.

attack Scheme F (R). Interestingly, subjects who reported that Scheme F and R were equally likely to be successful still preferred to attack Scheme F over Scheme R. This pattern holds when we restrict the sample to include only subjects who indicated correct beliefs for both schemes. A similar pattern occurs when attack choices are broken down by conditions (see Supporting Information). Clearly, even when subjects stated correct beliefs for both schemes, they still preferred to attack Scheme F compared to Scheme R.

Result 5. *The preference for Scheme F was robust to changes in the description of the scheme.*

One potential explanation for Results 3 and 4 is that the wording used in Treatment FR unintentionally conveyed to subjects a greater sense of risk associated with Scheme R than Scheme F. To test for this possibility, we conducted the Treatment FR with Time treatment that altered the description of the schemes. Table VII presents a multinomial Probit regression addressing how attack choices were influenced by treatment. From columns (1) and (2), we see that subjects were more likely to attack Scheme F and Scheme R in Treatment FR with Time compared to Treatment FR (see the Supporting Information for predicted probabilities). The preference for Scheme F is thus robust to changes in the description. This finding lends support to the conclusion that this behavioral puzzle is not a simple artifact of our basic experiment design but may extend to other attacker-defender settings, a fact we discuss more below.

Table VII. Predicting Attack Choices by Treatment: Multinomial Probit Regression

	(1) Attack Scheme F	(2) Attack Scheme R
Treatment FR with Beliefs	−0.02	0.11
	(0.06)	(0.07)
Treatment FR with Time	0.11^{+}	0.14^{*}
	(0.06)	(0.06)
Expected Payoff Max	-0.61^{***}	-0.50^{***}
Not Attack	(0.08)	(0.09)
Expected Payoff Difference	0.35^{***}	0.25^{***}
	(0.02)	(0.02)
Risk Averse	-0.19^{***}	-0.33^{***}
	(0.05)	(0.06)
Risk Loving	0.62^{***}	0.33^{***}
	(0.07)	(0.07)
Female	0.21^{***}	0.04
	(0.05)	(0.06)
Age	-0.05^{***}	-0.07^{***}
	(0.01)	(0.01)
Round	-0.01^{***}	-0.01^{*}
	(0.00)	(0.00)
Intercept	0.73^{**}	0.73^{*}
	(0.25)	(0.28)
Predicted Probability	0.41	0.19

Notes: Standard errors in parentheses. The base outcome variable is Not Attack. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

4. DISCUSSION

As Becker[36] argues, individuals are likely to choose to engage in criminal activity if the expected utility of doing so is greater than the expected utility of abstaining. Utility is a mathematical construction that intends to represent a decisionmaker's preferences over alternatives, and this function accounts for the benefits and costs of available options. In this framework, potential attackers can be deterred by changes in the costs and benefits of attacking as well as the probability that they could be caught. How sensitive an individual is to changes in benefits and costs depends on the properties and shape of the utility function. In our experiment, deterrence was heavily influenced by changes in the potential payoffs and the probability of being caught. However, as commonly found in experiments, the subjects do not act to maximize expected payoffs, nor do they act to maximize expected utility that incorporates only risk preferences and beliefs. Another factor or factors are also in operation.

Other models of decision making have been proposed to account for various possible factors that can explain deviations from the expected

utility model. Two prominent ones are prospect theory[17,37] and quantal response.[18] In prospect theory, probabilities are weighted according to a probability weighting function and individuals have a reference point that determines whether they view potential outcomes as gains or losses. In the quantal response equilibrium, players can make errors when choosing pure strategies. These errors are less likely to occur the more costly the error is. Yet, neither prospect theory nor quantal response predict any difference between Scheme F and Scheme R. In neither of these theories does a decisionmaker who assesses the expected payoff and risk of one alternative to be the same as another alternative have any reason to prefer one option over another, yet that is what we find in our study.

The results of our experiment suggest that aspects of the decision-making scenario not accounted for by standard expected utility maximization, prospect theory, or quantal response behavior may affect behavior. One possibility is that decision making could be influenced by whether options are evaluated simultaneously or in isolation.[38] For example, when attackers face two similar choices to attack they may view them as more distinct when evaluating them at the same time compared to when they evaluated them separately. As a result, certain attributes that may have not been as important to a potential attacker may be more salient when comparing the two options, leading to what is called a distinction bias.[38] This result has a practical implication for implementing defense at real-world targets. For example, consider an attacker that can choose to target a gate at one of two different airports, where one airport is known to use a fixed guarding scheme and the other is known to use a rotating guarding scheme. Our findings suggest that the airport with the fixed scheme has a higher chance of being attacked. The key issue is that when different airports (or ports, etc.) use different guarding schemes, then the chance of one airport being attacked depends partly on the security decisions of the other airports.

That respondents were equally likely to attack in Treatment F and Treatment R, but preferred to attack Scheme F over Scheme R in Treatment FR, could potentially be interpreted as a type of preference reversal paradox. In Treatment FR, while respondents chose whether to attack Scheme F, attack Scheme R, or not attack, the Scheme F and Scheme R choices may be similar to a matching response mode. That is, in Treatment FR respondents chose between two gambles, Scheme F and Scheme R, and a sure thing (not attack). Research on the preference reversal effect[39,40] suggests that the probability of success (or failure) is a prominent cue when choosing between two gambles. While the probability of success is defined to be the same for Scheme F and Scheme R, 32% of the time subjects reported that Scheme F was more likely to be successful, and 14% of the time reported Scheme R was more likely to be successful. If respondents were more concerned about the likelihood of success in Treatment FR, and Scheme F (R) was judged to have a greater chance of success, then we would expect the greater frequency of subjects selecting Scheme F (R). But in Treatment F and Treatment R, respondents choose between attacking or not attacking, which involves a comparison between a single gamble and a sure thing. In these attack versus not attack treatments, respondents could be more likely to focus on the outcomes; that is, the sure thing outcome (not attack) versus the possible outcomes resulting from an attack. Unlike Treatment FR, which requires comparison of two gambles, respondents are placed in the situation of evaluating the attack option. While respondents are choosing between the gamble and the sure thing, they are implicitly evaluating the relative payoffs of the two options, attack versus not attack. It is possible that Treatment F and Treatment R trigger an evaluation mode in which the subjective probability of success or failure receives much less weight than in Treatment FR. Of course, since the potential payoffs are the same in Treatment F and Treatment R, one would predict no difference in the likelihood of an attack if respondents are focused on payoffs.

There exist two issues with the preference reversal paradox explanation. First, if subjects are placing more weight on the probability of success in Treatment FR compared to Treatment F and Treatment R, then we would expect the fraction of subjects choosing not attack in Treatment FR to be different compared to Treatment F and Treatment R. However, these rates were quite similar. Second, while the subjective beliefs may have been important for some subjects, 54% of the time subjects reported that both Scheme F and Scheme R were equally likely to be successful. Despite viewing the schemes as having the same probability of success, Scheme F was still preferred over Scheme R. This suggests that there exists an additional attribute or attributes that are important in explaining the observed preference.

Combining expected utility and the idea of distinction bias we can model individuals as having lexicographic preferences. Assume that there are a set of options for an attacker to choose from, including the choice to not attack. Define the set of options as $A = \{a_1, a_2, \ldots\ldots, a_n\}$. Each element of the set $a_{i \in A}$ contains a set of attributes $X = \{x_1, x_2, \ldots\ldots, x_m\}$. For simplicity assume that each option a_i contains two attributes: x_1^i and x_2^i.

Definition 1. *Option a_i is lexicographically preferred to option a_j if and only if one of the following two conditions are met:*

(1) $x_1^i > x_1^j$

(2) $x_1^i = x_1^j$ *and* $x_2^i \geq x_2^j$

In our context, x_1^i could represent the expected utility from the material payoffs of attacking or not attacking and x_2^i could represent the perceived risk of choosing that option. Perceived risk may capture differences in the behavioral response between the two schemes. When asked to discuss their choices at the end of the experiment, a number of subjects reported that they felt safer choosing Scheme F. Examples of subjects' responses include: "It seemed the most safe," "I would rather have chosen to attack something that doesn't move, because I don't think I could take the emotional consequence of being caught based on chance of rotation," and "I felt safest with stationary guards." Since deterrence rates were similar between the two schemes when subjects could not choose, it is possible that Scheme F only feels different to subjects when compared to Scheme R. When subjects can choose between targets with different schemes, the movement of the guards may become more salient, leading them to perceive Scheme F as safer. Our model differs from distinction bias because it is only when options are the same in expected utility that the second attribute matters. As a result, it is possible that attacking behavior would be similar when Scheme F and Scheme R are presented alone. However, if some individuals have lexicographic preferences then when subjects can choose between the schemes they may prefer one scheme over the other.

Possible support for our interpretation may be found in dual-process theory, which suggests that people have two systems (System I and System II) of reasoning that may influence decision making.[41–44] Relying on emotion and intuition, System I tends to be fast and automatic. System II operates more slowly, focusing on logic and rule-based reasoning. In one study, Gelder *et al.*[45] showed that priming subjects in tasks focusing on a specific system prior to making risk judgments led to that system being more predictive of risky choice. It is possible that when comparing the two schemes, subjects feel better about choosing Scheme F compared to Scheme R, and this intuitive feeling could be what is driving the observed behavior. Indeed, research on the affect heuristic finds that intuitive (System I) feelings about an alternative guide an individual's judgment, including the assessment of the riskiness of that alternative.[46,47] Though the affect heuristic may be at work in our study, we note that such feelings in our study are at work independently of the risk assessment; remember that individuals who report identical risk and expected payoffs for both Scheme F and Scheme R still prefer Scheme F to Scheme R. This suggests that the affect heuristic could be at work in our experiment but with additional nuance. Future research is needed to verify whether this theory or some other theory best accounts for the observed behavior and to rule out potential alternative explanations.

5. CONCLUSION

In the experiment, deterrence was clearly influenced by the probability of success, payoffs, and risk preferences. The results from this study suggest that the assumption that attackers are expected value maximizers may be too strong. As argued in Shieh *et al.*,[5] Pita *et al.*,[6] and Yang *et al.*,[7] real-world decision-making software could be improved by incorporating more realistic assumptions of adversary behavior.

The results from this study suggest that when given the choice the majority of people would prefer to attack Scheme F compared to Scheme R. Interestingly, despite this preference, there was little difference in deterrence rates when subjects could only choose to attack Scheme F or only choose to attack Scheme R. Even when subjects held correct beliefs about the probability of being caught for the two schemes, there was still a preference to attack Scheme F. One potential explanation for this result is that subjects may be prone to a distinction bias.[38] In isolation, the two schemes may seem quite similar to subjects. However, when viewed jointly the differences between the two schemes may become more salient to subjects, leading to differences in observed choices. Similarly, subjects may have lexicographic

preferences. As assumed in our model, subjects may be concerned primarily about the expected utility of their choices. This predicts that the deterrence rates for both schemes in isolation should be similar. However, when presented with both Scheme F and R the expected utility for each scheme does not differ. In this case, subjects choose the option that gives them the lowest perceived risk, which appeared to be Scheme F for the majority of subjects.

This study demonstrates the importance of considering the risk perception of adversaries. The perceived risk of attackers can influence both beliefs about the success of an attack and the selection of targets. Even if potential targets are similar in both risk and payoffs, the way guards protect those targets can influence which target is likely to be attacked. Accounting for adversary risk perception may improve security and minimize the expected costs caused from attacks by both criminals and terrorists.

ACKNOWLEDGMENTS

We would like to thank participants for their helpful comments and suggestions at the 2014 Economic Science Association (ESA) International Meeting in Hawaii. We acknowledge financial support from Army Research Office Award No. W911NF-11-1-0332. McBride also acknowledges financial support from Air Force Office of Scientific Research Award No. FA9550-10-1-0569.

REFERENCES

1. Kiekintveld C, Jain M, Tsai J, Pita J, Ordonez F, Tambe M. Computing optimal randomized resource allocations for massive security games. Pp. 689–697 in Decker S, Castelfranchi C (eds). Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009), Budapest, Hungary.
2. Pita J, Jain M, Mareki J, Ordonez F, Portway C, Tambe M, Western C, Paruchuri P, Kraus S. Deployed armor protection: The application of a game theoretical model for security at the Los Angeles International airport. Pp. 125–132 in Padgham P (eds). Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2008), Budapest, Hungary, 2008.
3. Tsai J, Rathi S, Kiekintveld C, Ordonez F, Tambe M. IRIS—A tool for strategic security allocation in transportation networks. Pp. 37–44 in Decker S, Sichman J, Sierra C, and Castelfranchi C (eds). Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009), Budapest, Hungary.
4. Pita J, Tambe M, Kiekintveld C, Cullen S, Steigerwald E. GUARDS—Game theoretic security allocation on a national scale. Pp. 37–45 in Tumer, Yolum, Sonenberg and Stone (eds). Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems — Innovative Applications Track (AAMAS 2011), Taipei, Taiwan.
5. Shieh E, An B, Yang R, Tambe M, Baldwin C, DiRenzo J, Maule B, Meyer G. PROTECT: A deployed game the theoretic system to protect the ports of United States. In Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011), Conitzer, Winikoff, Padgham, and van der Hoek (eds.), Valencia, Spain.
6. Pita J, Jain M, Tambe M, Ordonez F, Kraus S. Robust solutions to Stackelberg games: Addressing bounded rationality and limited observations in human cognition. Artificial Intelligence, 2010; 174:1142–1171.
7. Yang R, Kiekintveld C, Ordonez F, Tambe M, John R. Improving resource allocation strategy against human adversaries in security games. Pp 458-465 in International Joint Conference on Artificial Intelligence (IJCAI-11), Walsh (ed), Menlo Park, California, USA.
8. Emile Borel. La theorie du jeu les equations integrales a noyau symetrique. Comptes Rendus del Academic, 1921; 173:1304–1308. English translation by Savage L: The theory of play and integral equations with skew symmetric kernals. Econometrica, 1953; 21:97–100.
9. Shubik M, Weber RJ. Systems defense games: Colonel Blotto, command and control. Navel Research Logistics Quarterly, 1981; 28:281–287.
10. Roberson B. The Colonel Blotto game. Economic Theory, 2006; 29:1–34.
11. Arce DG, Kovenock D, Roberson B. Weakest-link attacker-defender games with multiple attack technologies. Navel Research Logistics Quarterly, 2012; 59:457–469.
12. Wu Y, Wang B, Liu KR. Optimal power allocation strategy against jamming attacks using the Colonel Blotto game. Global Telecommunications Conference, 2009; 1–5.
13. Chia PH, Chuang J. Colonel Blotto in the phishing war. Decision and Game Theory for Security: Lecture Notes in Computer Science, 2011; 7037:201–218.
14. Avrahami J, Yaakov Kareev. Do the weak stand a change? Distribution of resources in a competitive environment. Cognitive Science, 2009; 33:940–950.
15. Chowdhury SM, Kovenock D, Sheremeta RM. An experimental investigation of Colonel Blotto games. Economic Theory, 2013; 52:833–861.
16. Powell R. Sequential, nonzero-sum "Blotto": Allocating defensive resources prior to attack. Games and Economic Behavior, 2009; 67:611–615.
17. Kahneman D, Tversky A. An analysis of decision under risk. Econometrica, 1979; 47:263–291.
18. McKelvey RD, Palfrey TR. Quantal response equilibria for normal form games. Games and Economic Behavior, 1995; 10:6–38.
19. Bier V, Oliveros S, Samuelson L. Choosing what to protect: Strategic defensive allocation against an unknown attacker. Journal of Public Economic Theory, 2007; 9:563–587.
20. Dighe NS, Zhuang J, Bier V. Secrecy in defensive allocations as a strategy for achieving more cost-effective attacker deterrence. International Journal of Performability Engineering, 2009; 5:31–43.
21. Berman O, Gavious A. Location of terror response facilities: A game between state and terrorist. European Journal of Operational Research, 2007; 177:1113–1133.
22. Zhuang J, Bier V, Alagoz O. Modeling secrecy and deception in a multiple-period attack-defender signaling game. European Journal of Operational Research, 2010; 203:409–418.
23. Bier V, Haphuriwat N, Menoyo J, Zimmerman R, Culpen AM. Optimal resource allocation for defense of targets based on differing measures of attractiveness. Risk Analysis, 2008; 28:763–770.
24. Shan XG, Zhuang J. Modeling credible retaliation threats in deterring the smuggling of nuclear weapons using partial inspection—A three-stage game. Decision Analysis, 2014; 11:43–62.

25. Hausken K, Zhuang J. The timing and deterrence of terrorist attacks due to exogenous dynamics. Journal of the Operational Research Society, 2012; 63:726–735.
26. Friesen L. Certainty of punishment versus severity of punishment: An experimental investigation. Southern Economic Journal, 2012; 79:399–421.
27. DeAngelo G, Charness G. Deterrence, expected cost, uncertainty and voting: Experimental evidence. Journal of Risk and Uncertainty, 2012; 44:73–100.
28. Scurich N, John RS. Perceptions of randomized security schedules. Risk Analysis, 2014; 34:765–770.
29. Khadjavi M. Deterrence works for criminals. European Journal of Law and Economics, 2015; doi:10.1007/s10657-015-9483-2.
30. Braithwaite A, Foster DM, Sobek DA. Ballots, bargains, and bombs: Terrorist targeting of spoiler opportunities. International Interactions: Empirical and Theoretical Research in International Relations, 2010; 36:294–305.
31. Putra IE, Sukabdi ZA. Basic concept and reasons behind the emergence of religious terror activities in Indonesia: An inside view. Asian Journal of Social Psychology, 2013; 16:83–91.
32. Fischbacher U. z-tree: Zurich toolbox for ready-made economic experiments. Experimental Economics, 2007; 10:171–178.
33. Michael C. Bidding behavior in pay-to-bid auctions: An experimental study. Department of Economics, University of California, Irvine Working Paper, 2013.
34. Eckel CC, Grossman PJ. Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. Journal of Economic Behavior & Organization, 2008; 68:1–17.
35. Wooldrige JM. Econometric Analysis of Cross Section and Panel Data. Cambridge, MA: USA MIT Press, 2010.
36. Gary B. Crime and punishment: An economic approach. Journal of Political Economy, 1968; 76:169–217.
37. Tversky A, Kahneman D. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty, 1992; 5:297–323.
38. Hsee CK, Zhang J. Distinction bias: Misprediction and mischoice due to joint evaluation. Journal of Personality and Social Psychology, 2004; 86:680–695.
39. Tversky A, Sattath S, Slovic P. Contingent weighting in judgment and choice. Psychological Review, 1988; 95:371–384.
40. Tversky A, Slovic P, Kahneman D. The causes of preference reversal. American Economic Review, 1990; 80:204–217.
41. Epstein S. Integration of the cognitive and the psychodynamic unconscious. American Psychologist, 1994; 49:709–724.
42. Stanovich K, West RF. Individual differences in reasoning: Implications for the rationality debate? Behavioral and Brain Sciences, 2000; 23:645–726.
43. Kahneman D. A perspective on judgment and choice. American Psychologist, 2003; 58:697–720.
44. Slovic P, Finucane ML, Peters E, MacGregor DF. Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. Risk Analysis, 2004; 24:311–322.
45. van Gelder JL, de Vries RE, van der Pligt J. Evaluating a dual-process model of risk: Affect and cognition as determinants of risky choice. Journal of Behavioral Decision Making, 2009; 22:45–61.
46. Slovic P, Peters E. Risk perception and affect. Current Directions in Psychological Science, 2006; 15:322–325.
47. Slovic P, Finucane ML, Peters E, MacGregor DG. The affect heuristic. European Journal of Operational Research, 2007; 177:1333–1352.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's website:

Online Supplement: Explaining the pattern in Fig. 2 approximations for estimating change in life expectancy attributable to air pollution in relation to multiple causes of death.