Interdependent Security Games under Behavioral Probability Weighting

Ashish R. Hota and Shreyas Sundaram*

Abstract

We consider a class of *interdependent security games* where the security risk experienced by a player depends on her own investment in security as well as the investments of other players. In contrast to much of the existing work that considers risk neutral players in such games, we investigate the impacts of behavioral probability weighting by players while making security investment decisions. This weighting captures the transformation of objective probabilities into perceived probabilities, which influence the decisions of individuals in uncertain environments. We show that the Nash equilibria that arise after incorporating probability weightings have much richer structural properties and equilibrium risk profiles than in risk neutral environments. We provide comprehensive discussions of these effects on the properties of equilibria and the social optimum when the players have homogeneous weighting parameters, including comparative statics results. We further characterize the existence and uniqueness of pure Nash equilibria in *Total Effort* games with heterogeneous players.

1 Introduction

Interdependent security games are a class of strategic games where multiple selfish players choose personal investments in security, and the security risk faced by a player depends on the investments of other players in the society [Laszka et al., 2014, Kunreuther and Heal, 2003]. These games serve as abstract frameworks that capture various forms of risk externalities that users face in networked environments. There is a large body of literature on this class of problems starting from the early works in Varian [2004] and Kunreuther and Heal [2003]; a comprehensive recent survey can be found in Laszka et al. [2014].

The risk faced by individuals in these settings is often manifested as the probability of a successful attack, and this probability is a function of the investment by the individual and the externalities due to the investments by other interacting individuals. The system-wide landscape of security investments in this setting will be a function of the decisions that individuals make under this notion of risk. Much of the work in interdependent security games considers players who are risk neutral, or are risk averse in the sense of classical expected utility theory [Laszka et al., 2014]. On the other hand, there is a rich literature in decision theory and behavioral economics which concerns itself with decision making under risk with findings that suggest consistent and significant deviations in human behavior from the predictions of classical expected utility theory [Camerer et al., 2011, Barberis, 2013]. One of the most important deviations is the way individuals perceive the probability of an uncertain outcome (e.g., cyber attack). In particular, empirical studies show that individuals tend to overweight small probabilities and underweight large probabilities. Thus, the objective (i.e., true) probabilities are typically transformed in a highly nonlinear fashion into

*The authors are with the School of Electrical and Computer Engineering, Purdue University. E-mail: {ahota,sundara2}@purdue.edu

perceived probabilities, which are then used for decision making [Kahneman and Tversky, 1979, Gonzalez and Wu, 1999].

While there have been some studies highlighting the significance of biases and irrationalities in human decision making in information security domains [Christin, 2011, Baddeley, 2011], theoretical analyses of relevant deviations from classical notions of rational behavior are scarce in the literature on interdependent security games.[1] Empirical investigations [Christin et al., 2012, Rosoff et al., 2013] are also limited in this context.

In this paper, our goal is to study the effects of behavioral probability weighting of players on their equilibrium strategies in three fundamental interdependent security game models: *Total Effort*, *Weakest Link* and *Best Shot* games. Each of these games captures a certain manifestation of risk (i.e., probability of successful attack) as a function of investment by the players, and is motivated by practical scenarios as discussed in [Grossklags et al., 2008a]. In the *Total Effort* game, the probability of a successful attack depends on the average of the security investments by the players. As an example, an attacker might want to slow down the transfer of a file in a peer-to-peer file sharing system, while the speed is determined by the sum of efforts of several participating machines. In the *Weakest Link* game, the society is only as secure as the least secure player, while in the *Best Shot* game, the player with the maximum investment must be successfully attacked for the attack on other players to be successful. Weakest link externalities are prevalent in computer security domains; successful breach of one subsystem often increases the vulnerability of other subsystems, such as by giving the attacker increasing access to otherwise restricted parts. Best shot externalities often arise in cyber systems with built in redundancies. In order to disrupt a certain functionality of the target system, the attack must successfully compromise all the entities that are responsible for maintaining that functionality.

These game-theoretic models were first introduced in [Varian, 2004], and were subsequently extended in [Grossklags et al., 2008a,b] to cases where players can also purchase (static) cyber insurance in addition to investing in security. All three models are instances where the nature of externalities is positive[2] as the investment by a player (weakly) improves the security of the everyone in the society (by reducing the attack probability). The fact that the players' utility functions in these games are coupled through the shared probability of successful attack motivates our focus on studying the effects of behavioral probability weighting.[3]

We model the nonlinear probability weightings of players using the weighting function due to Prelec [1998], whose properties we discuss in the next section. We first characterize the pure Nash equilibria (PNE) in *Total Effort* games; we compare the (structural) properties of the equilibrium under behavioral probability weighting to the equilibria that arise under risk neutral [Grossklags et al., 2008a] and classical expected utility maximization behavior [Johnson et al., 2011]. We then examine how the intensity of probability weighting affects the probability of successful attack at equilibrium. We carry out a similar analysis for the social welfare maximizing solution in the

[1] In a related class of security games known as Stackelberg security games with two players, one attacker and one defender, there have been recent studies [Yang et al., 2013, Jiang et al., 2013, Kar et al., 2015, Brown et al., 2014, Zhuang, 2014] that incorporate behavioral decision theoretic models, including prospect theory and quantal response equilibrium. However, this class of games is very different from interdependent security games [Laszka et al., 2014], which is the focus of the current work.

[2] Both positive and negative externalities have been studied in the literature. Negative externalities capture settings where more investment by others makes a player more vulnerable, and this is usually the case where the attack is targeted towards individuals who have invested less in security. Most of the literature in security games has focused on positive externalities [Laszka et al., 2014].

[3] There are also various behavioral characteristics that affect the perceived *values* of gains and losses [Kahneman and Tversky, 1979, Hota et al., 2014]. However, as the values of the gains and losses are not strategy-dependent in the games that we consider here, behavioral value functions would not affect the equilibria that arise.

Total Effort game, and Nash equilibria in *Weakest Link* and *Best Shot* games. Subsequently, we prove general existence and uniqueness results in *Total Effort* games when the probability weighting parameters and cost parameters are heterogeneous (player-specific) under certain conditions on the number of players.

2 Probability Weighting

As discussed in the previous section, our focus in this paper will be on understanding the effects of nonlinear weighting of objective probabilities by individuals while making decisions under risk. Such weightings have been comprehensively studied in the behavioral economics and psychology literature [Camerer et al., 2011], and have certain fundamental characteristics, including (i) *possibility effect*: overweighting of probabilities very close to 0, (ii) *certainty effect*: underweighting of probabilities very close to 1, and (iii) *diminishing sensitivity* from the end points 0 and 1. These characteristics are usually captured by an inverse S-shaped weighting function. After the initial advancements in the development of prospect theory and rank dependent utility theory, various parametric forms of weighting functions were proposed, most prominently in Tversky and Kahneman [1992], Gonzalez and Wu [1999], and Prelec [1998]. All of these parametric weighting functions exhibit the qualitative and analytical characteristics (i) to (iii) described above.

In this paper, we consider the one parameter probability weighting function due to Prelec.[4] If the objective (i.e., true) probability of an outcome is x, the weighting function is given by

$$w(x) = \exp(-(-\ln(x))^{\alpha}), \qquad x \in [0,1], \tag{1}$$

where exp() is the exponential function. The parameter $\alpha \in (0,1]$ controls the curvature of the weighting function as illustrated in Figure 1. For $\alpha = 1$, we have $w(x) = x$, i.e., the weighting function is linear. As α decreases away from 1, $w(x)$ becomes increasingly nonlinear, with an inverse S-shape. For smaller α, the function $w(x)$ has a sharper overweighting of low probabilities and underweighting of high probabilities.

Remark 1. *The probability weighting function $w(x)$ in (1) has the following properties.*

1. $w(0) = 0$, $w(1) = 1$, *and* $w(\frac{1}{e}) = \frac{1}{e}$.
2. $w(x)$ *is concave for* $x \in [0, \frac{1}{e}]$, *and convex for* $x \in [\frac{1}{e}, 1]$.
3. $w'(x)$ *attains its minimum at* $x = \frac{1}{e}$. *In other words,* $w''(x) = 0$ *at* $x = \frac{1}{e}$. *The minimum value of* $w'(x)$ *is* $w'(\frac{1}{e}) = \alpha$.
4. $w'(\epsilon) \to \infty$ *as* $\epsilon \to 0$, *and* $w'(1-\epsilon) \to \infty$ *as* $\epsilon \to 0$.

3 Interdependent Security Games

As discussed in the introduction, in interdependent security games, a player makes her security investment decision independently, while the probability of successful attack on the player depends on the strategies of other players. We denote the number of players by n, and denote the investment in security by player i as s_i, where $s_i \in [0,1]$. Following the conventional game theoretic notation,

[4] While we focus on the Prelec weighting function here, many of our results will also hold under a broader class of weighting functions with similar qualitative properties as the Prelec weighting function.

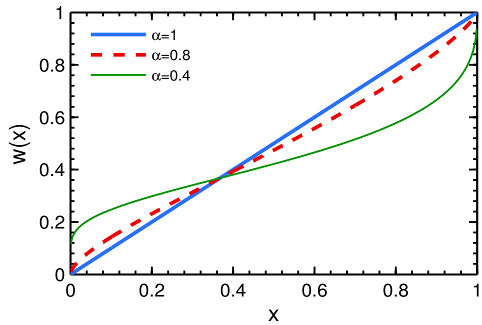

Figure 1: Shape of the probability weighting function. The quantity x is the objective probability of failure, and $w(x)$ is the corresponding perceived probability of failure.

we use s_{-i} to denote the investment profile of all players other than i. The formulation that we consider here has the following characteristics. The objective probability of a successful attack on the system is given by $f(s_i, s_{-i}) \in [0, 1]$, for some function f. Player i incurs a cost-per-unit of security investment of $b_i \in \mathbb{R}_{\geq 0}$, and if the system experiences a successful attack, player i incurs a loss of $L_i \in \mathbb{R}_{>0}$. The expected utility of a player (under the true probabilities of successful attack) is then

$$Eu_i = -L_i f(s_i, s_{-i}) - b_i s_i. \tag{2}$$

In settings that model positive externalities, as is the focus in the current paper, $f(s_i, s_{-i})$ is nonincreasing in both s_i and s_{-i}, and usually assumed to be convex in s_i for analytical tractability.

In this work, we consider three canonical models of interdependent security games with positive externalities, initially presented in [Grossklags et al., 2008a,b]. The models differ in the attack probability function $f(s_i, s_{-i})$ as described below.

- *Total Effort*: $f(s_i, s_{-i}) = 1 - \frac{1}{N}(\sum_{i=1}^{n} s_i)$.
- *Weakest Link*: $f(s_i, s_{-i}) = 1 - \min_{i=1}^{n} s_i$.
- *Best Shot*: $f(s_i, s_{-i}) = 1 - \max_{i=1}^{n} s_i$.

In [Grossklags et al., 2008a,b], the authors additionally considered the possibility of players investing in insurance to reduce the magnitude of loss. In order to isolate the effects of nonlinear probability weighting, we do not consider insurance here. To establish a baseline, the following proposition describes the main results from [Grossklags et al., 2008a] regarding the properties of Nash equilibria in the three security games defined above with homogeneous risk neutral players (without self-insurance). We will compare these results with the equilibria under nonlinear probability weighting in subsequent sections.

Proposition 1. *Consider a set of N risk neutral players with homogeneous cost parameters (b and L).*

1. *Total Effort: There is a unique symmetric PNE except for the special case where $\frac{Nb}{L} = 1$. If $\frac{Nb}{L} < 1$, then each player invests to fully protect herself, i.e., $s_i^* = 1, i \in \{1, 2, \ldots, N\}$. Otherwise if $\frac{Nb}{L} > 1$, $s_i^* = 0$.*

2. *Weakest Link: At any PNE, all players have identical security investment. If $\frac{b}{L} > 1$, then $s_i^* = 0$ for every player i. Otherwise, any investment $s_i^* \in [0, 1]$ can constitute a PNE.*

3. *Best Shot: If* $\frac{b}{L} > 1$, *then* $s_i^* = 0$ *for every player* i *at the PNE. Otherwise, there is no symmetric PNE, and at most one player has a nonzero investment of* 1, *while all other players free ride without making any security investment.*

We make two preliminary observations regarding the above equilibria in games with risk neutral players. First, in the *Total Effort* game, the best response of a player is independent of the decisions of other players and only depends on her cost parameters, i.e., the interdependence has no impact on her strategy. Secondly, in both *Total Effort* and *Best Shot* games, the PNE causes the system to be either fully secure from attack, or to be fully unprotected. We will show that under behavioral probability weighting, both the best response and the equilibria have much richer structural properties and vary more smoothly with the parameters of the game.

4 Total Effort Game with Probability Weighting: Homogeneous Players

First we characterize the pure Nash equilibria in a *Total Effort* security game when the number of players is sufficiently large and players are homogeneous in their weighting functions and cost parameters. Unlike the risk neutral case described in Proposition 1, the best response in this case will be potentially discontinuous in the strategies of other players. Furthermore, for a sufficiently large number of players, we show that there always exists an interior equilibrium. This equilibrium is not necessarily unique, and in fact, can coexist with an equilibrium where all players invest 1.

With probability weighting, the expected utility of player i under investment $s_i \in [0,1]$ is given by

$$\mathbb{E}u_i(s_i, s_{-i}) = -Lw\left(1 - \frac{s_i + \bar{s}_{-i}}{N}\right) - bs_i,$$

where $\bar{s}_{-i} = \sum_{j \neq i} s_j$ is the total investment in security by all players other than i. The function w is the Prelec weighting function defined in (1). The marginal utility is given by

$$\frac{\partial \mathbb{E}u_i}{\partial s_i} = \frac{L}{N} w'\left(1 - \frac{s_i + \bar{s}_{-i}}{N}\right) - b. \tag{3}$$

The solutions of $\frac{\partial \mathbb{E}u_i}{\partial s_i} = 0$ satisfy the first order condition of optimality, and are therefore candidate solutions for players' best responses and the PNE. Note that $\left(1 - \frac{s_i + \bar{s}_{-i}}{N}\right)$ is the objective attack probability faced by the players (without probability weighting). For a given strategy profile of other players, player i's strategy can change the objective attack probability in the interval $\mathcal{X}(\bar{s}_{-i}) := \left[1 - \frac{1+\bar{s}_{-i}}{N}, 1 - \frac{\bar{s}_{-i}}{N}\right]$. In other words, when the number of players is N, each player can directly change the probability of successful attack by at most $\frac{1}{N}$.

Recall from Remark 1 that the minimum value of $w'(x)$ for $x \in [0,1]$ is α. Therefore, if $\alpha > \frac{Nb}{L}$, from (3) we have $\frac{\partial \mathbb{E}u_i}{\partial s_i} > 0$ for $s_i \in [0,1]$, and investing 1 is the only best response of player i irrespective of the strategies of the other players. Therefore, the only PNE strategy profile for $\alpha > \frac{Nb}{L}$ is when each player invests 1 in security. Note that in the special case where $\alpha = 1$ (i.e., $w(x) = x$), this reduces to the risk-neutral strategy profile given in Proposition 1.

Now suppose $\alpha < \frac{Nb}{L}$. In this case, the first order condition $w'(x) = \frac{Nb}{L}$ has two distinct interior solutions corresponding to objective attack probabilities $X_1 < \frac{1}{e}$ and $X_2 > \frac{1}{e}$, as illustrated in Figure 2. It is easy to see that as the number of players N increases, $X_2 - X_1$ increases while $\frac{1}{N}$ decreases.

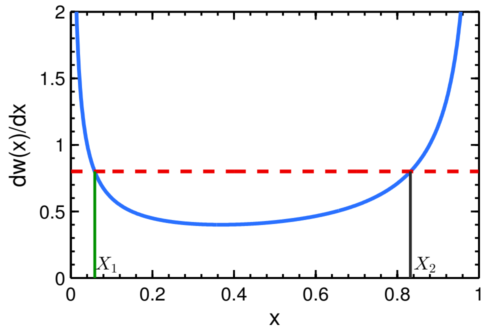

Figure 2: Interior solutions of $w'(x) = \frac{Nb}{L}$ are denoted by X_1 and X_2. In this example, $\frac{Nb}{L} = 0.8$ and is shown by the horizontal line. The weighting parameter is $\alpha = 0.4$.

In the following proposition, we characterize the PNE for sufficiently large N such that $X_2 - X_1 > \frac{1}{N}$. This condition implies that at a given strategy profile of other players, $\mathcal{X}(\bar{s}_{-i})$ does not simultaneously contain both X_1 and X_2. This makes the analysis more tractable, and is a reasonable assumption for networked environments where the number of players is large.

Proposition 2. *Consider a Total Effort security game with homogeneous players with probability weighting parameter $\alpha < \frac{Nb}{L}$. Let N be sufficiently large so that $X_2 > \frac{1}{N} + X_1$, where $X_j, j = 1, 2$ are solutions to the equation $w'(x) = \frac{Nb}{L}$. Then,*

1. *any strategy profile with $\left(1 - \frac{s_i + \bar{s}_{-i}}{N}\right) = X_2$ is a PNE,*
2. *any strategy profile with $\left(1 - \frac{s_i + \bar{s}_{-i}}{N}\right) = X_1$ is not a PNE, and*
3. *there exists a PNE with all players investing 1 if and only if*
   (a) $X_1 \geq \frac{1}{N}$, or
   (b) $X_1 < \frac{1}{N}$ and $w(\frac{1}{N}) > \frac{b}{L}$.

Proof. **Part 1.** Consider any strategy profile $\mathbf{s}^* = \{s_1^*, \ldots, s_N^*\}$ with $\left(1 - \frac{s_i^* + \bar{s}_{-i}^*}{N}\right) = X_2$. The best response of each player i is obtained by solving the following optimization problem:

$$\max_{x \in [0,1]} \quad -Lw\left(1 - \frac{x + \bar{s}_{-i}^*}{N}\right) - bx. \tag{4}$$

At $x = s_i^*$, the player satisfies the first order condition of optimality $\frac{\partial \mathbb{E} u_i}{\partial s_i} = \frac{L}{N} w'(X_2) - b = 0$. For any $x < s_i^*$, we have $X = 1 - \frac{x + \bar{s}_{-i}}{N} > X_2$ and $\frac{\partial \mathbb{E} u_i}{\partial s_i} = \frac{L}{N} w'(X) - b > 0$ since $X_2 > \frac{1}{e}$. As a result, no $x < s_i^*$ would satisfy the first order necessary condition of optimality. On the other hand, for any $x > s_i^*$, we have $X = 1 - \frac{x + \bar{s}_{-i}}{N} < X_2$. However, from our assumption that $X_2 > \frac{1}{N} + X_1$, we would have $X = 1 - \frac{x + \bar{s}_{-i}}{N} > X_1$. Therefore, $\frac{\partial \mathbb{E} u_i}{\partial s_i} < 0$ for any $x > s_i^*$. As a result, $x = s_i^*$ is the only candidate for optimal investment, and it also satisfies the second order sufficient condition since $w''(X_2) > 0$ as $X_2 > \frac{1}{e}$. This concludes the proof.

Part 2. Consider any strategy profile $\mathbf{s}^* = \{s_1^*, \ldots, s_N^*\}$ with $\left(1 - \frac{s_i^* + \bar{s}_{-i}^*}{N}\right) = X_1$. For any $x > s_i^*$, we have $X = 1 - \frac{x + \bar{s}_{-i}}{N} < X_1$. Since $X_1 < \frac{1}{e}$, $w'(X) > \frac{Nb}{L}$ (see Figure 2). Thus $\frac{\partial \mathbb{E} u_i}{\partial s_i} > 0$ for any $s_i > s_i^*$. As a result, $s_i = 1$ is also a candidate solution for the utility maximization problem of player i, along with s_i^*. However we show that for Prelec weighting functions, a player would always prefer to invest $s_i = 1$ over $s_i = s_i^*$. To simplify the notation, define $Y_1 = 1 - \frac{1+\bar{s}^*_{-i}}{N}$. Note that $X_1 - Y_1 = \frac{1}{N}(1 - s_1^*)$. Now we compute

$$\begin{aligned}
\mathbb{E}u_i(1, \mathbf{s}^*_{-\mathbf{i}}) - \mathbb{E}u_i(s_i^*, \mathbf{s}^*_{-\mathbf{i}}) &= -Lw(Y_1) - b + Lw(X_1) + bs_i^* \\
&= L(w(X_1) - w(Y_1)) - b(1 - s_i^*) \\
&= L(w(X_1) - w(Y_1)) - bN(X_1 - Y_1) \\
&= L(X_1 - Y_1)\left[\frac{w(X_1) - w(Y_1)}{(X_1 - Y_1)} - \frac{Nb}{L}\right] \\
&\geq L(X_1 - Y_1)\left[w'(X_1) - \frac{Nb}{L}\right] = 0,
\end{aligned}$$

where the inequality is due to the fact that $w(x)$ is concave for $x \in [0, \frac{1}{e}]$, and $Y_1 < X_1 < \frac{1}{e}$, with equality at $\alpha = 1$. Therefore, between the potential interior solution s_i^* that satisfies the first order condition and the boundary solution $s_i = 1$, the player will always prefer the boundary solution. Since $X_1 > 0$, there always exists a player with $s_i^* < 1$ which would prefer to invest 1, and therefore the strategy profile is not a PNE.

Part 3. Suppose $X_1 \geq \frac{1}{N}$ and all players other than player i are investing 1. Then player i's investment can vary the objective probability of successful attack in the range $\left[0, \frac{1}{N}\right]$, and in this region, $\frac{\partial \mathbb{E}u_i}{\partial s_i} > 0$. As a result, investing $s_i = 1$ is the only best response for player 1. Thus $\mathbf{s}^* = \{1, \ldots, 1\}$ is a PNE.

On the other hand, suppose $X_1 < \frac{1}{N}$. Consider a strategy profile where all players other than i are investing 1, i.e., $\mathbf{s}^*_{-\mathbf{i}} = \{s_1^* = 1, \ldots, s_{i-1}^* = 1, s_{i+1}^* = 1, \ldots, s_N^* = 1\}$. The following three strategies satisfy the first order necessary condition of optimality of the utility maximization problem (4), and thus are candidates for best responses: (i) $s_i^* = 1$ as $\frac{\partial \mathbb{E}u_i}{\partial s_i}|_{s_i^*=1} > 0$, (ii) $s_i^* = 1 - NX_1$ as $\frac{\partial \mathbb{E}u_i}{\partial s_i}|_{s_i^*=1-NX_1} = \frac{L}{N}w'(X_1) - b = 0$, and (iii) $s_i^* = 0$ as $\frac{\partial \mathbb{E}u_i}{\partial s_i}|_{s_i^*=0} = \frac{L}{N}w'(\frac{1}{N}) - b < 0$ (from Figure 2 and the fact that $X_1 < \frac{1}{N}$).

From our analysis in Case 2, we know that the player would always prefer to invest 1 over investing $s_i = 1 - NX_1$. Therefore, a necessary and sufficient condition for all players investing 1 to be a PNE is $\mathbb{E}u_i(1, \mathbf{s}^*_{-\mathbf{i}}) > \mathbb{E}u_i(0, \mathbf{s}^*_{-\mathbf{i}})$. Since

$$\mathbb{E}u_i(1, \mathbf{s}^*_{-\mathbf{i}}) - \mathbb{E}u_i(0, \mathbf{s}^*_{-\mathbf{i}}) = -b + Lw(\frac{1}{N}),$$

we have the equivalent condition that $\mathbf{s}^* = \{1, \ldots, 1\}$ is a PNE if and only if $w(\frac{1}{N}) > \frac{b}{L}$. Otherwise, player i would achieve greater utility by investing 0, and therefore, all players investing 1 is not a PNE. □

Discussion: The above proposition completely characterizes the pure Nash equilibria in *Total Effort* security games with homogeneous nonlinear probability weighting for a sufficiently large number of players. It is instructive to compare this set of equilibria with the ones for risk neutral players given in Proposition 1. When $\frac{Nb}{L} > \alpha$, there always exists an interior PNE corresponding to a (true) probability of successful attack equal to X_2. As the number of players N increases, the probability of successful attack X_2 gradually increases to 1 (by the definition of X_2 and Figure 2). This is in contrast to the situation with risk neutral players ($\alpha = 1$), where the probability of successful attack jumps suddenly from 0 to 1 once the number of players increases beyond a certain point.

Secondly, under behavioral probability weighting, there can *also* be an equilibrium where all players invest 1 (under the conditions described in the third part of the above result). This is a consequence of overweighting of small probabilities; individual players with $\alpha < 1$ do not find it profitable to reduce their investments ("free-ride"), since the resulting small increase of attack probability is perceived to be much larger.

In the next subsection, we discuss how the attack probabilities at PNE are affected by the weighting parameter α, and then analyze the social optimum in the *Total Effort* game.

4.1 Comparative statics

Consider two *Total Effort* games where the parameters N, b and L are the same across the two games. The first game has homogeneous players with weighting parameter α_1, and the second game has homogeneous players with weighting parameter α_2, where $\alpha_1 < \alpha_2 < \frac{Nb}{L}$. By Proposition 2, both games have an interior PNE, with corresponding true probability of successful attack equal to X_2^1 and X_2^2, respectively. Note that for $i \in \{1, 2\}$, $X_2^i > \frac{1}{e}$, and is the solution to the equation $w_i'(x) = \frac{Nb}{L}$, where $w_i(x)$ be the Prelec function (1) with weighting parameter α_i, for $i \in \{1, 2\}$. As we illustrate in Figure 3 for $\alpha_1 = 0.4$ and $\alpha_2 = 0.8$, $w_1'(x)$ is initially smaller than $w_2'(x)$ as x starts to increase from $\frac{1}{e}$, until the quantity $x = \bar{X}$ (which depends on the values of α_1 and α_2) at which $w_1'(x) = w_2'(x)$. For $x > \bar{X}$, $w_1'(x) > w_2'(x)$. We first formally prove this observation via the following lemma and proposition.

Lemma 1. *The function*

$$g(x) = (-\ln(x))^{\alpha_2 - \alpha_1} \exp[(-\ln(x))^{\alpha_1} - (-\ln(x))^{\alpha_2}], \;\; \alpha_1 < \alpha_2 < 1,$$

is strictly decreasing in $x \in [\frac{1}{e}, 1]$.

Proof. We compute

$$\begin{aligned} g'(x) &= \exp((-\ln(x))^{\alpha_1} - (-\ln(x))^{\alpha_2}) \times \\ &\quad \Bigg[(\alpha_1(-\ln(x))^{\alpha_1 - 1} - \alpha_2(-\ln(x))^{\alpha_2 - 1}) \frac{-1}{x} (-\ln(x))^{\alpha_2 - \alpha_1} \\ &\qquad - (\alpha_2 - \alpha_1)(-\ln(x))^{\alpha_2 - \alpha_1 - 1} \frac{1}{x} \Bigg] \\ &= -\frac{1}{x} \exp((-\ln(x))^{\alpha_1} - (-\ln(x))^{\alpha_2})(-\ln(x))^{\alpha_2 - \alpha_1 - 1} \times \\ &\quad [(\alpha_1(-\ln(x))^{\alpha_1} - \alpha_1 - \alpha_2(-\ln(x))^{\alpha_2} + \alpha_2] \,. \end{aligned}$$

When $x > \frac{1}{e}$, $(-\ln(x)) < 1$ and thus $1 > (-\ln(x))^{\alpha_1} > (-\ln(x))^{\alpha_2}$. As a result, $1 - (-\ln(x))^{\alpha_1} < 1 - (-\ln(x))^{\alpha_2}$. Since $\alpha_1 < \alpha_2$, this implies $\alpha_1(-\ln(x))^{\alpha_1} - \alpha_1 > \alpha_2(-\ln(x))^{\alpha_2} - \alpha_2$. Therefore, $g'(x) < 0$ over $x \in [\frac{1}{e}, 1]$. □

Proposition 3. *Consider two Prelec weighting functions w_1 and w_2 with parameters α_1 and α_2, respectively and let $\alpha_1 < \alpha_2$. Then there exists a unique $\bar{X} > \frac{1}{e}$ such that (i) $w_1'(\bar{X}) = w_2'(\bar{X})$, ii) for $x \in [\frac{1}{e}, \bar{X}]$, $w_1'(x) < w_2'(x)$, and iii) for $x \in [\bar{X}, 1]$, $w_1'(x) > w_2'(x)$.*

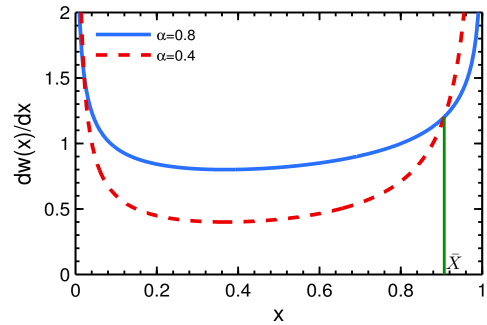

Figure 3: $w'(x)$ for two different weighting parameters, $\alpha = 0.4$ and $\alpha = 0.8$. At $x = \bar{X}$, $w'_1(x) = w'_2(x)$.

Proof. The first derivative of the Prelec weighting function is given by $w'(x) = w(x)\frac{\alpha}{x}(-\ln(x))^{\alpha-1}$. Therefore, if at a given x, $w'_1(x) = w'_2(x)$, we have

$$\begin{aligned} & w_1(x)\alpha_1(-\ln(x))^{\alpha_1} = w_2(x)\alpha_2(-\ln(x))^{\alpha_2} \\ \implies & \frac{\alpha_1}{\alpha_2} = \frac{w_2(x)(-\ln(x))^{\alpha_2}}{w_1(x)(-\ln(x))^{\alpha_1}} \\ \implies & \frac{\alpha_1}{\alpha_2} = (-\ln(x))^{\alpha_2-\alpha_1}\exp((-\ln(x))^{\alpha_1} - (-\ln(x))^{\alpha_2}) = g(x). \end{aligned}$$

From the definition of $g(x)$, $g(\frac{1}{e}) = 1 > \frac{\alpha_1}{\alpha_2}$. Furthermore, as $x \to 1$, $g'(x) \to -\infty$. As a result, $g(x)$ becomes smaller than $\frac{\alpha_1}{\alpha_2}$ for some $x < 1$. Thus there exists $\bar{X}$ at which $w'_1(x) = w'_2(x)$. The uniqueness of $\bar{X}$ follows from the strict monotonicity of $g(x)$ as proved in Lemma 1.

In order to prove the second and third parts of the lemma, it suffices to show that $w''_1(\bar{X}) > w''_2(\bar{X})$. Therefore, we compute $w''(x)$, which after some calculations yields

$$w''(x) = \frac{w'(x)}{-x\ln(x)}[1 + \ln(x) + \alpha((-\ln(x))^\alpha - 1)].$$

From the previous discussion, $\alpha_1(-\ln(x))^{\alpha_1} - \alpha_1 > \alpha_2(-\ln(x))^{\alpha_2} - \alpha_2$, and $w''(x) > 0$ for $x > \frac{1}{e}$. Therefore at $\bar{X}$, $w''_1(\bar{X}) > w''_2(\bar{X})$. This concludes the proof. □

The above proposition (illustrated in Figure 3) shows that if $\frac{Nb}{L} > w'(\bar{X})$, then we have $\bar{X} < X_2^1 < X_2^2$ whenever $\alpha_1 < \alpha_2$. On the other hand, when $\alpha_2 < \frac{Nb}{L} < w'(\bar{X})$, we have $\bar{X} > X_2^1 > X_2^2$. Since the attack probability at the interior equilibrium is equal to X_2^i, we have the following result.

Proposition 4. *Consider two Total Effort security games with homogeneous players. The players have weighting parameter α_1 in the first game and α_2 in the second game with $\alpha_1 < \alpha_2$. Let $\bar{X}$ be the intersection point defined in Proposition 3 for α_1 and α_2. If $\frac{Nb}{L} > w'(\bar{X})$, then the true probability of successful attack at the interior PNE is larger in the game where players have weighting parameter α_2. Similarly, if $\frac{Nb}{L} < w'(\bar{X})$, then the investments by players with weighting parameter α_1 results in higher probability of successful attack at the interior PNE. If $\frac{Nb}{L} = w'(\bar{X})$, both games have identical attack probability at the interior PNE.*

Discussion: The above result shows that when the attack probability is close to 1, the players view security investments to be highly beneficial in terms of reducing the perceived attack probabilities (due to significant underweighting of probabilities closer to 1). This effect is more pronounced in players with a smaller α (as shown in Figure 1), and as a result, the attack probability at the PNE is smaller compared to a game where players have higher α. On the other hand, when the quantity $\frac{Nb}{L}$ is smaller, the attack probability at the interior solution is more moderate and the players with smaller α do not find the perceived reduction of attack probability beneficial enough to make a high investment compared to players with larger α. Therefore, the nature of probability weighting plays a key role in shaping the attack probability at the PNE.

4.2 Social Optimum

We define the social welfare of the *Total Effort* game as the sum of the utilities of the individual players, as is commonly defined in the game theory literature [Nisan et al., 2007]. Formally, for a given strategy profile $\mathbf{s}$, the social welfare function is defined as

$$\Psi(\mathbf{s}) = -NLw\left(1 - \frac{\sum_{i=1}^{N} s_i}{N}\right) - b\sum_{i=1}^{N} s_i. \tag{5}$$

Noting that the social welfare function only depends on the aggregate investment, we denote with some abuse of notation

$$\Psi(\bar{s}) = -N[Lw(1-\bar{s}) + b\bar{s}], \tag{6}$$

where $\bar{s}$ is the average security investment by the players. As a result, the social welfare optimization performed by the social planner is independent of the number of players in the system. In fact, the optimal solution to the problem of the central planner is the same as the optimal investment when there is a single player in the game. In the following result, we discuss how the optimal investment under a central planner depends on the weighting parameter α and the cost parameters b and L. Subsequently, the following result will also be helpful in characterizing the nature of Nash equilibria in *Weakest Link* and *Best Shot* games.

The marginal of the social welfare function (6) is

$$\frac{1}{N}\frac{\partial \Psi}{\partial s} = Lw'(1-s) - b, \tag{7}$$

where we use s as the average security investment instead of $\bar{s}$ for notational convenience. Similar to the discussion prior to Proposition 2, if $\frac{b}{L} < \alpha$, then $\frac{\partial \Psi}{\partial s} > 0$ for $s \in [0, 1]$, and as a result, the optimal investment is 1. Therefore we focus on the case where $\frac{b}{L} > \alpha$; in this case, there are two solutions to $w'(x) = \frac{b}{L}$, denoted as $X_1 < \frac{1}{e}$ and $X_2 > \frac{1}{e}$ as before. We will need the following lemma for our analysis.

Lemma 2. *Let z be such that $w'(z) = \frac{w(z)}{z}$. Then i) z is unique, ii) $z > \frac{1}{e}$ and iii) for $x > z$, $w'(x) > \frac{w(x)}{x}$.*

Proof. For the Prelec weighting function,

$$w'(x) = w(x)\frac{\alpha}{x}(-\ln(x))^{\alpha-1}.$$

At any z with $w'(z) = \frac{w(z)}{z}$, we must have $\alpha(-\ln(z))^{\alpha-1} = 1$. Since $\alpha < 1$, we must have $-\ln(z) < 1$ or equivalently $z > \frac{1}{e}$. Furthermore, $(-\ln(x))^{\alpha-1}$ is strictly increasing in x for $\alpha < 1$. As a result, there is a unique $x = z$ at which $\alpha(-\ln(z))^{\alpha-1} = 1$, and for $x > z$, $w'(x) > \frac{w(x)}{x}$. Similarly, $w'(x) < \frac{w(x)}{x}$ for $x \in \left[\frac{1}{e}, z\right]$. □

Proposition 5. *Let z be as defined in Lemma 2.*

1. *If $\frac{b}{L} < w'(z)$, the socially optimal average investment is 1.*

2. *Otherwise, the socially optimal average investment is $1 - X_2$, where $w'(X_2) = \frac{b}{L}$.*

Proof. Since $\frac{b}{L}$ is finite, from (7) we have $\frac{\partial \Psi}{\partial s} > 0$ at $s = 0$ and therefore investing 0 in security is not a utility maximizer. So we have three candidate solutions for utility maximization, $s_1 = 1 - X_1$, $s_2 = 1 - X_2$ or $s_3 = 1$.

From the analysis in Part 2 of Proposition 2 with $\bar{s}^*_{-i} = 0$ and $N = 1$, we have $\mathbb{E}u(1) > \mathbb{E}u(1 - X_1)$. Therefore, between the potential interior solution $s_1 = 1 - X_1$ that satisfies the first order condition and the boundary solution $s_3 = 1$, the player will always prefer the boundary solution.

Now, to compare the utilities at the solutions s_2 and 1, we compute

$$\begin{aligned}\frac{1}{N}(\Psi(1) - \Psi(s_2)) &= Lw(1 - s_2) - b(1 - s_2) \\ &= L(1 - s_2)\left[\frac{w(1 - s_2)}{1 - s_2} - \frac{b}{L}\right] \\ &= L(1 - s_2)\left[\frac{w(1 - s_2)}{1 - s_2} - w'(1 - s_2)\right].\end{aligned}$$

From our discussion in Lemma 2, if $\frac{b}{L} = w'(1 - s_2) < w'(z)$, the player would prefer to invest 1, and otherwise will prefer to invest s_2. □

5 Weakest Link and Best Shot Games

The Nash equilibrium strategies in *Weakest Link* and *Best Shot* games have very special properties, as the security level of the entire system is determined by the investment of a single player (the one with the smallest and largest investment, respectively). As a result, the levels of investment at the equilibria often depend on the optimal investment by a single user (i.e., investment in a game with $N = 1$), which we analyzed in the previous subsection. We first characterize the PNE in *Weakest Link* games.

Proposition 6. *Consider a Weakest Link game with homogeneous players having probability weighting parameter $\alpha \in (0, 1]$. Then at any PNE, all players have identical investment. If $\frac{b}{L} > w'(z)$, where z is as defined in Lemma 2, then there is a continuum of pure Nash equilibria with attack probability greater than or equal to X_2. When $\frac{b}{L} < w'(z)$, then there are additional equilibria (including the ones in the previous case) with attack probabilities less than $\frac{1}{e}$.*

Proof. The first part of the statement is easy to see, as no player prefers to invest more than the current minimum level of investment. In fact, if at a given strategy profile, no player can improve her utility by unilaterally deviating to a lower investment level, then the strategy profile is a PNE.

When $\frac{b}{L} > w'(z)$, our result in Proposition 5 states that a single player investing in isolation would prefer to invest $s^* = 1 - X_2$ where $X_2 > \frac{1}{e}$ is the interior solution to the first order condition $w'(x) = \frac{b}{L}$. Now suppose all players have identical security investment $s \leq s^*$, i.e., the objective attack probability $X \geq X_2$. Since for each player $w'(x) > \frac{b}{L}$ for $x > X_2$, no player would unilaterally deviate to make a lower investment. Therefore, any investment less than s^* by all the players would result in a PNE.

When $\frac{b}{L} < w'(z)$, the optimal investment when $N = 1$ is 1 since $\mathbb{E}u(1) > \mathbb{E}u(1-X_2)$. Therefore, some very low attack probabilities (such as with investment $s = 1 - \epsilon, \epsilon \to 0$ where $\mathbb{E}u(s) > \mathbb{E}u(1-X_2)$) can be supported at a PNE, in addition to the set of equilibria with attack probabilities greater than X_2. □

Discussion: The main differences of the above result compared to the equilibria in *Weakest Link* games without probability weighting (Proposition 1) are twofold. First, for large enough $\frac{b}{L}$, the only possible equilibrium in the risk neutral case is when all players invest 0, while under probability weighting, there is a range of possible equilibrium investments. Secondly, for smaller values of $\frac{b}{L}(< 1)$, any investment level by the players can give rise to a PNE for risk neutral players, while that is no longer the case with behavioral probability weighting. However, at the social optimum of *Weakest Link* games, the investment level chosen by the central planner will coincide with the optimal investment stated in Proposition 5.

Proposition 7. *Consider a Best Shot game with players having identical cost and weighting parameters. Then at any PNE, at most one player can have a nonzero investment with investment level according to Proposition 5, and the rest of the players invest zero.*

Discussion: Recall that the structure of the PNE was similar in homogeneous *Best Shot* games without probability weighting, with at most one player having nonzero investment. However, the nonzero investment level was at one of the boundary points, either 0 or 1, and as a result, the equilibrium was either entirely protected or vulnerable (Proposition 1). With probability weighting, the nonzero equilibrium investment is one of the interior solutions when $\frac{b}{L} > w'(z)$, and the investment level gradually decreases as $\frac{b}{L}$ increases. Finally, note that the social optimum solution coincides with the PNE in *Best Shot* games.

6 Total Effort Game with Heterogeneous Players

In this section, we consider *Total Effort* games with player-specific weighting parameters α_i and cost parameters $\frac{b_i}{L_i}$. We first prove the existence of a PNE in games with a sufficiently large number of players, and subsequently give a constructive proof of uniqueness of PNE.

In the rest of the analysis, we denote the weighting function of player i as $w_i(x)$, and denote the solutions of $w_i'(x) = \frac{Nb_i}{L_i}$, if any, as $X_1^i < \frac{1}{e}$ and $X_2^i > \frac{1}{e}$, respectively. Now we are ready to state our assumptions on the number of players, which are sufficient conditions for our results to hold.

Assumption 1. *Let the number of players N be large enough such that for every player i,*

1. $\alpha_i < \frac{Nb_i}{L_i}$,
2. $X_2^i - X_1^i > \frac{1}{N}$,
3. $X_1^i < \frac{1}{N}$ *and*
4. $w_i(\frac{1}{N}) < \frac{b_i}{L_i}$.

The above conditions are guaranteed to be simultaneously satisfied for a sufficiently large number of players due to the properties of the weighting functions. Recall from our discussion in Proposition 2 that for homogeneous players, the last two of the above assumptions are necessary and sufficient to ensure that all players investing 1 is not an equilibrium.

Proposition 8. *Consider a Total Effort game where player i has player-specific weighting parameter α_i and cost ratio $\frac{b_i}{L_i}$. Let the number of players satisfy Assumption 1. Then there exists a PNE of the game.*

Proof. We show that the best response of a player is unique, continuous and is an affine decreasing function of the total security investment by other players. The result will then follow from Brouwer's fixed point theorem. Our analysis of the best response of a player holds for every player i, and therefore we drop the subscript in the rest of proof.

Let $y \in [0, N-1]$ be the total security investment by all other players. For a given y, the investment of the player can only change the true probability of successful attack within the interval $\mathcal{X}(y) := \left[1 - \frac{1+y}{N}, 1 - \frac{y}{N}\right]$, the length of which is $\frac{1}{N}$. Under Assumption 1, this interval might fall into one of four different cases. As y increases from 0, we gradually go through each of the cases, starting from either Case 1 or Case 2.

Case 1: $X_2 < 1 - \frac{1+y}{N}$

In this case, the interval $\mathcal{X}(y)$ lies to the right of X_2. Therefore, for any attack probability $x \in \mathcal{X}(y)$, $\frac{L}{N}w'(x) > b$ (from Figure 2). Thus, $\frac{\partial \mathbb{E}u}{\partial s}|_{s=x} > 0$ and consequently, $b(y) = 1$ this case.

Case 2: $1 - \frac{1+y}{N} \leq X_2 \leq 1 - \frac{y}{N}$

In this case, $X_2 \in \mathcal{X}(y)$, and therefore, the player has a feasible investment strategy $s^* = N(1 - X_2) - y$ at which the first order condition is satisfied with equality. By identical arguments as in Part 1 of Proposition 2, the player is only going to invest at the interior solution s^*. The second requirement of Assumption 1 ensures that $X_1 \notin \mathcal{X}(y)$, and as result, the utility function remains concave, and therefore the best response is unique for any given y.

Since the optimal solution s^* must have the property that $1 - \frac{s^*+y}{N} = X_2$, it is continuous and linearly decreasing in y, with boundary values at 1 and 0 respectively for $y = N(1 - X_2) - 1$ and $y = N(1 - X_2)$.

Case 3: $X_1 < 1 - \frac{1+y}{N}$ and $X_2 > 1 - \frac{y}{N}$

In this case, the interval $\mathcal{X}(y)$ lies in the region between X_1 and X_2. Therefore, for any objective failure probability $x \in \mathcal{X}(y)$, $\frac{\partial \mathbb{E}u}{\partial s} < 0$. As a result, $b(y) = 0$.

Case 4: $1 - \frac{1+y}{N} \leq X_1 \leq 1 - \frac{y}{N}$

In this case, there are three candidate solutions for utility maximization, $s = 1$, $s = N(1-X_1)-y$ and $s = 0$, analogous to the candidate solutions in Part 3 of Proposition 2. We have $X = 1 - \frac{s+y}{N}$ as the objective failure probability resulting from the strategies of the players. From an identical analysis as in Part 2 of Proposition 2 with $\bar{s}_{-i} = y$, we conclude that the player would always prefer to invest 1 over investing $s^* = N(1 - X_1) - y$. This leads to the possibility that the best response might have a discontinuous jump from 0 to 1 at some value of y in this region. However, we show that under the third and fourth conditions of Assumption 1, the player would always prefer to invest 0 over investing 1.

We compute

$$\begin{aligned}
\mathbb{E}u(1, y) - \mathbb{E}u(0, y) &= L\left[w(1 - \frac{y}{N}) - w(1 - \frac{1+y}{N})\right] - b \\
&= L\left[w(\lambda + \frac{1}{N}) - w(\lambda)\right] - b \\
&\leq L\left[w(\frac{1}{N}) - w(0)\right] - b,
\end{aligned}$$

where $\lambda = 1 - \frac{1+y}{N}$. The last inequality follows because the function $h(\lambda) \triangleq w(\lambda + \frac{1}{N}) - w(\lambda)$ is a strictly decreasing function of λ for $\lambda \in [0, X_1]$. Indeed, $h'(\lambda) = w'(\lambda + \frac{1}{N}) - w'(\lambda) < 0$, as

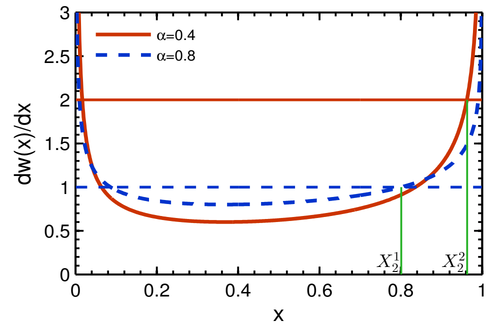

Figure 4: First order conditions for heterogeneous players. The horizontal lines represent the quantities $\frac{Nb_i}{L_i}$, while the curved lines represent $w_i'(x)$, which depends on the player specific weighting parameter α_i. The quantity X_2^i represents the largest solution to the equation $w_i'(x) = \frac{Nb_i}{L_i}$.

$w'(\lambda) \geq w'(X_1) = \frac{Nb}{L}$, and $w'(\lambda + \frac{1}{N}) < \frac{Nb}{L}$, as $X_1 < \lambda + \frac{1}{N} < X_2$. Therefore, if $w(\frac{1}{N}) < \frac{b}{L}$, the player would always prefer to invest 0 over investing 1, regardless of the value of y. Furthermore, since $X_1 < \frac{1}{N}$, and $1 - \frac{y}{N} = \frac{1}{N}$ at $y = N - 1$, the best response remains at 0 in this region of y.

Combining the analysis in all four cases together, the best response of any player is unique and continuous in the strategies of the other players, regardless of the value of $\alpha_i \in (0, 1]$ and $\frac{b_i}{L_i}$. In addition, the strategy space of each player is $[0, 1]$, which is compact and convex. Therefore, a PNE always exists by Brouwer's fixed point theorem. □

In the following result, we establish the uniqueness of the PNE.

Proposition 9. *Consider a Total Effort game where each player i has player-specific weighting parameter α_i and cost ratio $\frac{b_i}{L_i}$. Let the number of players N satisfy Assumption 1. Then all PNE have the same objective probability of successful attack.*

Proof. Without loss of generality, let players be ordered such that $X_2^1 \leq X_2^2 \leq X_2^3 \leq \ldots \leq X_2^N$, where X_2^i is the largest solution to $w_i'(x) = \frac{Nb_i}{L_i}$; note that such a solution is guaranteed by the first requirement of Assumption 1. We present a numerical illustration in Figure 4. Note that this ordering does not necessarily mean that the corresponding α_i's or cost ratios form a monotonic sequence. Under Assumption 1, no objective attack probability $X < X_2^1$ would be a PNE, since there would always exist a player with positive investment who would prefer to reduce her investment.

Now suppose there are two PNEs with different corresponding probabilities of successful attack. Consider the strategy profile with the smaller attack probability, denoted X^*. Note that we ruled out the possibility of $X^* < X_2^1$ above. There are two exhaustive cases: either $X_2^l < X^* < X_2^{l+1}$ for some player l, or $X_2^l = X^*$ for some player l.

Let $X_2^l < X^* < X_2^{l+1}$ for some player l. By the definition of the quantities X_2^i, we have $w_i'(X^*) < \frac{Nb_i}{L_i}$ for $i \in \{l+1, \ldots, N\}$, and therefore, $s_i^* = 0$ for $i \in \{l+1, \ldots, N\}$. Similarly, $w_i'(X^*) > \frac{Nb_i}{L_i}$ and $s_i^* = 1$ for $i \in \{1, \ldots, l\}$. In this case, $X^* = 1 - \frac{l}{N}$. Now at the second PNE with objective attack probability $Y^* > X^*$, the players in $\{1, \ldots, l\}$ would continue to invest 1, with the possibility of more players investing nonzero amounts if $Y^* \geq X_2^{l+1}$. But then the objective attack probability $X = 1 - \frac{\sum_i s_i}{N}$ would decrease from X^*, contradicting the assumption that $Y^* > X^*$.

The proof of the case where $X^* = X_2^l$ for some player l follows identical arguments, and therefore we omit its presentation. □

7 Discussion and Conclusion

In this paper, we studied a class of interdependent security games where the players exhibit certain behavioral attributes vis-a-vis their perception of attack probabilities while making security investment decisions. In particular, we considered the parametric form of probability weighting function proposed by Prelec [1998], which is commonly used in behavioral decision theory settings, including prospect theory. We analyzed the properties of pure Nash equilibria (PNE) in three canonical interdependent security game models, (i) *Total Effort*, (ii) *Weakest Link*, and (iii) *Best Shot* games.

We first considered the *Total Effort* game with players having homogeneous weighting functions and cost parameters, and characterized the PNE strategies under a sufficiently large number of players. The equilibria with nonlinear probability weightings have much richer structural properties than the corresponding equilibria for risk-neutral players. There are only two types of equilibria with risk neutral players; one where the probability of successful attack is 1 (completely undefended), and the other where the probability is 0 (completely defended) with an abrupt transition to the latter as the number of players increase. However under behavioral probability weighting, there exist interior equilibria where the attack probability lies between 0 and 1. Furthermore, the equilibrium attack probability gradually increases to 1 with respect to certain cost parameters and the number of players. In addition to the interior equilibrium, there might coexist equilibria where players invest to fully secure themselves. In these equilibria, overweighting of low probabilities disincentivizes individuals from reducing their investments, since the perceived increase in attack probability due to reduced investment is much larger.

We also obtained interesting comparative statics results on the effect of the weighting parameter on the magnitude of the attack probability at equilibrium. If the probability of successful attack is sufficiently high, then players whose weighting functions are closer to linear prefer to invest relatively less in security, while players who exhibit a large underweighting of probabilities closer to 1 (certainty effect) prefer to invest more. This is due to the fact that the perceived reduction in attack probability is larger for the latter players. However, if the attack probability is only moderately high, we observe the opposite behavior; the attack probability at the equilibrium with highly nonlinear weighting functions is larger compared to the attack probability at equilibrium with players who more accurately perceive the true probabilities.

We subsequently analyzed the social welfare maximizing investment profiles in *Total Effort* games, which also had implications for the equilibria in *Weakest Link* and *Best Shot* games. In *Weakest Link* games, there often arise a multitude of equilibria with a continuum of attack probabilities, while in *Best Shot* games, at most one player makes a nonzero investment at any PNE. The investment levels at the equilibria in both these games have a more smooth variation in the game parameters compared to the investments by risk neutral players.

Finally, we analyzed *Total Effort* games where players have heterogeneous cost and weighting parameters, and established the existence of PNE and uniqueness of the corresponding attack probability when the number of players is sufficiently large. We leave a more comprehensive discussion on the effects of heterogeneity in weighting parameters for future work, in addition to several other future directions that we discuss below.

Future Work: There are several directions in which this line of investigation can be extended.

Cyber Insurance: The dichotomy between investing in security (to potentially reduce likelihood of attack) and purchasing cyber insurance (to decrease the magnitude of loss) has received considerable attention among information security researchers [Böhme and Schwartz, 2010, Pal et al., 2014, Naghizadeh and Liu, 2014]. However, in practice, the market for cyber insurance has seen limited growth despite growing awareness in the industry about various security risks. Further analysis of behavioral risk preferences (such as probability weighting, loss aversion and reference dependence) of decision makers in the context of cyber insurance could potentially uncover important phenomena which is not easily captured in models that only consider the classical expected utility maximization framework. The work in [Johnson et al., 2011] is a step in this direction, where the authors investigate the strategies of risk averse players (with concave utility functions) in *Weakest Link* security games in the presence of market insurance.

Network structure: In this paper, we have only considered extreme forms of network effects between players, as only the average, the highest, or the lowest investment levels decide the overall failure probabilities. Analyzing the effects of probability weighting and other forms of deviations from classical expected utility maximization behavior in models that consider richer networked environments [Lelarge and Bolot, 2008b,a, La, 2015, Schwartz and Sastry, 2014] is a challenging future direction.

Inefficiency of equilibria: Selfish behavior by users often leads to reduced welfare and increased attack probability at equilibrium in interdependent security games with positive externalities. While there is prior work in the literature on price of anarchy [Jiang et al., 2011] and price of uncertainty [Grossklags et al., 2010] of the current class of security games, investigating the effects of (behavioral) risk preferences of users, including probability weighting, on the inefficiency of equilibria remains an important avenue for future research.

Experimental investigations: Finally, the results obtained in this paper compliments and further motivates experimental/empirical investigations of human decision making in the context of information security and privacy.

References

Michelle Baddeley. Information security: Lessons from behavioural economics. In *Security and Human Behavior*, 2011.

Nicholas C Barberis. Thirty years of prospect theory in economics: A review and assessment. *Journal of Economic Perspectives*, 27(1):173–96, 2013.

Rainer Böhme and Galina Schwartz. Modeling cyber-insurance: Towards a unifying framework. In *Workshop on the Economics of Information Security (WEIS)*, 2010.

Matthew Brown, William B Haskell, and Milind Tambe. Addressing scalability and robustness in security games with multiple boundedly rational adversaries. In *Decision and Game Theory for Security*, pages 23–42. Springer, 2014.

Colin F Camerer, George Loewenstein, and Matthew Rabin. *Advances in behavioral economics*. Princeton University Press, 2011.

Nicolas Christin. Network security games: combining game theory, behavioral economics, and network measurements. In *Decision and Game Theory for Security*, pages 4–6. Springer, 2011.

Nicolas Christin, Serge Egelman, Timothy Vidas, and Jens Grossklags. Its all about the Benjamins: An empirical study on incentivizing users to ignore security advice. In *Financial Cryptography and Data Security*, pages 16–30. Springer, 2012.

Richard Gonzalez and George Wu. On the shape of the probability weighting function. *Cognitive psychology*, 38(1): 129–166, 1999.

Jens Grossklags, Nicolas Christin, and John Chuang. Secure or insure?: a game-theoretic analysis of information security games. In *Proceedings of the 17th international conference on World Wide Web*, pages 209–218. ACM, 2008a.

Jens Grossklags, Nicolas Christin, and John Chuang. Security and insurance management in networks with heterogeneous agents. In *Proceedings of the 9th ACM conference on Electronic Commerce*, pages 160–169. ACM, 2008b.

Jens Grossklags, Benjamin Johnson, and Nicolas Christin. The price of uncertainty in security games. In *Economics of Information Security and Privacy*, pages 9–32. Springer, 2010.

Ashish R Hota, Siddharth Garg, and Shreyas Sundaram. Fragility of the commons under prospect-theoretic risk attitudes. *arXiv preprint arXiv:1408.5951*, 2014.

Albert Xin Jiang, Thanh H Nguyen, Milind Tambe, and Ariel D Procaccia. Monotonic maximin: A robust Stackelberg solution against boundedly rational followers. In *Decision and Game Theory for Security*, pages 119–139. Springer, 2013.

Libin Jiang, Venkat Anantharam, and Jean Walrand. How bad are selfish investments in network security? *Networking, IEEE/ACM Transactions on*, 19(2):549–560, 2011.

Benjamin Johnson, Rainer Böhme, and Jens Grossklags. Security games with market insurance. In *Decision and Game Theory for Security*, pages 117–130. Springer, 2011.

Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, pages 263–291, 1979.

Debarun Kar, Fei Fang, Francesco Delle Fave, Nicole Sintov, and Milind Tambe. A game of thrones: When human behavior models compete in repeated Stackelberg security games. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 1381–1390, 2015.

Howard Kunreuther and Geoffrey Heal. Interdependent security. *Journal of risk and uncertainty*, 26(2-3):231–249, 2003.

Richard J La. Interdependent security with strategic agents and cascades of infection. *IEEE Transactions on Networking (To appear)*, 2015.

Aron Laszka, Mark Felegyhazi, and Levente Buttyan. A survey of interdependent information security games. *ACM Computing Surveys (CSUR)*, 47(2):23:1–23:38, 2014.

Marc Lelarge and Jean Bolot. A local mean field analysis of security investments in networks. In *Proceedings of the 3rd international workshop on economics of networked systems*, pages 25–30. ACM, 2008a.

Marc Lelarge and Jean Bolot. Network externalities and the deployment of security features and protocols in the Internet. *ACM SIGMETRICS Performance Evaluation Review*, 36(1):37–48, 2008b.

Parinaz Naghizadeh and Mingyan Liu. Voluntary participation in cyber-insurance markets. In *Workshop on the Economics of Information Security (WEIS)*, 2014.

Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V Vazirani. *Algorithmic game theory*. Cambridge University Press, 2007.

Ravindra Pal, Leana Golubchik, Konstantinos Psounis, and Pan Hui. Will cyber-insurance improve network security? A market analysis. In *INFOCOM, 2014 Proceedings IEEE*, pages 235–243. IEEE, 2014.

Drazen Prelec. The probability weighting function. *Econometrica*, pages 497–527, 1998.

Heather Rosoff, Jinshu Cui, and Richard S John. Heuristics and biases in cyber security dilemmas. *Environment Systems and Decisions*, 33(4):517–529, 2013.

Galina A Schwartz and S Shankar Sastry. Cyber-insurance framework for large scale interdependent networks. In *Proceedings of the 3rd international conference on High confidence networked systems*, pages 145–154. ACM, 2014.

Amos Tversky and Daniel Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323, 1992.

Hal Varian. System reliability and free riding. In *Economics of information security*, pages 1–15. Springer, 2004.

Rong Yang, Christopher Kiekintveld, Fernando Ordóñez, Milind Tambe, and Richard John. Improving resource allocation strategies against human adversaries in security games: An extended study. *Artificial Intelligence*, 195: 440–469, 2013.

Jun Zhuang. Modeling attacker-defender games with risk preference. *Current Research Project Synopses. Paper 69*, 2014. URL http://research.create.usc.edu/current_synopses/69.