

# Network Defense and Behavioral Biases: An Experimental Study

Daniel Woods, Mustafa Abdallah, Saurabh Bagchi, Shreyas Sundaram, Timothy Cason\*

February 9, 2021

## 1 Abstract

How do people distribute defenses over a directed network attack graph, where they must defend a critical node? This question is of interest to computer scientists, information technology and security professionals. Decision-makers are often subject to behavioral biases that cause them to make sub-optimal defense decisions, which can prove especially costly if the critical node is an essential infrastructure. We posit that non-linear probability weighting is one bias that may lead to sub-optimal decision-making in this environment, and provide an experimental test. We find support for this conjecture, and also identify other empirically important forms of biases such as naive diversification and preferences over the spatial timing of the revelation of an overall successful defense. The latter preference is related to the concept of anticipatory feelings induced by the timing of the resolution of uncertainty.

## 2 Introduction

Economic resources spent on securing critical infrastructure from malicious actors are substantial and increasing, with worldwide expenditure estimated to exceed \$124 billion in 2019 (Gartner, 2018). Cybersecurity defense is becoming increasingly difficult, as systems are frequently connected to the outside world through the Internet, and attackers innovate many new methods of attack. The interaction of computers, networks, and physical processes (termed ‘Cyber-Physical Systems’, or CPS) has a wide variety of applications, such as manufacturing, transportation, medical care, power generation and water management (Lee, 2015), and has both practical and theoretical importance. Proposed CPS such as the ‘Internet of Things’ promise vast benefits and efficiencies, but at the cost of increased attack vectors and targets (see Alaba et al. (2017) and Humayed et al. (2017) for surveys). To realize the potential gains that these new technologies can provide, we must understand and maximize their security.

To reduce interference with their systems, institutions allocate a security budget and hire managers responsible for minimizing the probability of successful attacks on important assets and other vital parts of the infrastructure. Such decision-makers, however, are subject to behavioral biases that can lead to sub-optimal security decisions (Abdallah et al. (2019b), Abdallah et al. (2019a), Acquisti and Grossklags (2007)). Human decision-makers can exhibit many possible biases. The security decisions they face broadly involve probabilistic assessments across multiple assets and attack vectors, many featuring low individual likelihood. We therefore ex-ante focus on the possibility that people incorrectly weight the actual probability of attack and defense (Tversky and Kahneman, 1992). We ex-post find that people also exhibit locational and spreading biases in their defense resource allocations, due to the directional and compartmentalized nature of these systems. Given the immense size of global expenditures on cybersecurity, as well as successful attacks being potentially very damaging, it is important to understand the nature and magnitude of any biases that can lead to sub-optimal security decisions. Such insights on biases can then be applied by security professionals to reduce their impact.

---

\*This research was supported by grant CNS-1718637 from the National Science Foundation. Daniel Woods and Timothy Cason are with the Economics Department in the Krannert School of Management at Purdue University. Email: {woods104, cason}@purdue.edu. Mustafa Abdallah, Saurabh Bagchi, and Shreyas Sundaram are with the School of Electrical and Computer Engineering at Purdue University. Email: {abdalla0, sbagchi, sundara2}@purdue.edu. We thank the editor, two anonymous referees, and participants at the Economic Science Association and Jordan-Wabash conferences for valuable comments.

We focus on human biases as infrastructure security decisions have not yet been given over to algorithmic tools. They are still mostly made by human security managers (Paté-Cornell et al., 2018). Adoption of automated tools are stymied by legacy components in these interconnected systems, so instead managers use threat assessment tools which return the likely probability that individual components of the infrastructure will be breached (Jauhar et al., 2015). These probabilities must be interpreted by the human manager, which motivates our initial emphasis on non-linear probability weighting. Evidence also exists that security experts ignore more accurate algorithmic advice when available and instead rely more on their own expertise (Logg et al., 2019).

We model a security manager’s problem as allocating his budget over edges in a directed attack graph with the nodes representing various subsystems or components of the overall CPS. An example of a directed attack graph is shown in Figure 1. The manager’s goal is to prevent an attacker who starts at the red node on the left from reaching the critical green node on the far right. The inter-connectivity of different systems is represented by connections between nodes, and alternative paths to a given node represent different methods of attack. Allocating more of the security budget to a given edge increases the probability that an attack through that edge will be stopped. Such an ‘interdependency attack graph’ model is considered an appropriate abstraction of the decision environment a security professional faces in large-scale networked systems.<sup>1</sup> The probability of successful defense along an edge is weighted according to the manager’s probability weighting function. We use the common Prelec (1998) probability weighting function, but similar comparative statics can be obtained with any ‘inverse S-shaped’ weighting function. We assume the attacker is sophisticated and observes the manager’s allocation decision, and does not mis-weight probabilities. This reflects a ‘worst-case’ approach to security (discussed further in Section 3.1), and represents a necessary first step in investigating the impact of probability weighting and other biases on security expenditures.

The manager’s mis-weighting of probabilities can cause investment decisions to substantially diverge from optimal decisions based on objectively correct true probabilities, depending on network structure and the security production function. The security production function maps defense resources allocated to an edge to the probability that an attack along that edge will be stopped. Empirical evidence has shown probability weighting to be relatively non-linear on the aggregate subject level (Bleichrodt and Pinto, 2000), so the impact on security decisions could be substantial. Probability weighting is also heterogeneous across individuals (Tanaka et al. (2010), Bruhin et al. (2010)). Therefore, if probability weighting affects choices in this environment, individuals should exhibit heterogeneity in their sub-optimal security decisions.

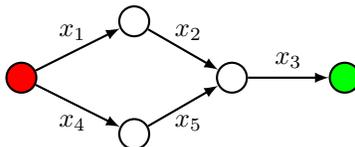


Figure 1: Example Directed Network Attack Graph

We seek to address the following research questions:

**Question 1:** What is the effect of probability weighting on security investments over a directed network graph?

**Question 2:** Is probability weighting an empirically relevant factor in human security decision-making?

**Question 3:** What other behavioral biases significantly affect decision-making in this environment?

To address Question 1, we numerically solve the security manager’s problem described above. In practical situations the relationship between investment spending and reductions in the probability of an attack is far from explicit to an outside observer. Moreover, investigations of successful breaches are often not revealed until months or years later. Furthermore, information on security investments is highly confidential for obvious reasons, making it difficult or impossible to obtain directly from firms. We therefore conduct an incentivized laboratory experiment to address Questions 2 and 3. We employ networks that cleanly identify the impact of non-linear probability weighting on security investment decisions, and the generated data also reveal other behavioral biases that exist in this environment.

<sup>1</sup>A non-exhaustive list of research considering the attack graph model from the Computer Security literature includes Sheyner and Wing (2003), Nguyen et al. (2010), Xie et al. (2010), Homer et al. (2013), and Hota et al. (2018). The length of this list and the ease in which it could be extended is indicative of the prominence that this literature places on the attack graph model.

Our experiment elicits separate measures of probability weighting outside the network defense problem to help address Question 2. One measure uses binary choices between lotteries which is relatively standard, and elicits probability weighting while controlling for the confound of utility curvature. The other measure is novel, and uses a similar network path framing to the network defense environment. This new measure reduces procedural variance relative to the main network defense task. It also exploits the irrelevance of utility curvature when there are only two outcomes to focus solely on probability weighting.

We find that the network-framed measure of non-linear probability weighting is statistically significantly correlated with all the network defense allocations situations we consider. However, this correlation exists even in cases where probability weighting should have no impact. This suggests that subjects may exhibit limited sophistication beyond probability weighting alone. We therefore conduct a cluster analysis to identify heterogeneous patterns of behavior not predicted by probability weighting. This identifies additional behavioral biases. The first is a form of ‘naive diversification’ (Benartzi and Thaler, 2001), where subjects have a tendency towards allocating their security budget evenly across the edges. The second is a preference for stopping the attacker earlier or later along the attack path. Stopping an attack earlier can be seen as reducing the anticipatory emotion of ‘dread’ (Loewenstein, 1987) while stopping it later can be seen as delaying the revelation of potentially bad news (e.g., see Caplin and Leahy (2004) for a strategic environment). Accounting for these additional biases, we continue to find some evidence that non-linear probability weighting influences subject behavior, as well as strong evidence for the additional biases. In our environment the additional biases seem especially naive, as edges are not different options with benefits beyond defending the critical node, and information on the attacker’s progress is not presented to the subjects sequentially. These inconsistencies possibly reflect a subject’s own mental model (e.g., of how an attack proceeds), but should be accounted for in future directed network decision environments.

This paper contributes to the theoretical literature on attack and defense games over networks of targets, most of which can be related to computer network security in some fashion.<sup>2</sup> Our attack graph environment is rather flexible, and can represent some of the strategic tensions present in alternative network environments. Instead of focusing on attack graph representations of these other environments (which can be quite complex), we utilize more parsimonious networks in order to specifically parse out the effect of probability weighting. We have the ‘security manager’ play against a sophisticated computerized attacker who moves after observing the manager’s allocation. Playing against a computer dampens socially related behavioral preferences.<sup>3</sup> It also removes the need for defenders to form beliefs about the attacker’s probability weighting. This allows us to more cleanly identify the empirical relevance of non-linear probability weighting in this spatial network defense environment. If probability weighting is important empirically, then future research should incorporate it into models to better understand the decisions of real-world decision-makers.

This paper also contributes to the experimental literature of attack and defense games in network environments.<sup>4</sup> One set of related experimental studies test ‘Network Disruption’ environments. McBride and Hewitt (2013) consider a problem where an attacker must select a node to remove from a partially obscured network, with the goal to remove as many edges as possible. Djawadi et al. (2019) consider an environment where the defender must both design the network structure as well as allocate defenses to nodes, with the goal of maintaining a network where all nodes are linked after an attack. Hoyer and Rosenkranz (2018) consider a similar but decentralized problem where each node is represented by a different player. Our environment differs from these Network Disruption games as we consider a directed attack graph network, i.e. the attacker must pass through the network to reach the critical node rather than remove a node to disrupt the network. Some other related experimental papers include ‘multi-battlefield’ attack and defense games, such as Deck and Sheremeta (2012), Chowdhury et al. (2013) and Kovenock et al. (2019). The most closely related of these types of papers is Chowdhury et al. (2016), who find experimental evidence for the bias of salience in a multi-battlefield contest, which induces sub-optimal allocations across battlefields. We are the first to investigate empirically the bias of probability weighting in networks and attack and defense games.

<sup>2</sup>A non-exhaustive list of related theory papers include Clark and Konrad (2007), Acemoglu et al. (2016), Dziubiński and Goyal (2013), Goyal and Vigier (2014), Dziubiński and Goyal (2017), Kovenock and Roberson (2018), and Bloch et al. (2020).

<sup>3</sup>Sheremeta (2019) posits that things such as inequality aversion, spite, regret aversion, guilt aversion, loss aversion (see also Chowdhury (2019)), overconfidence and other emotional responses could all be important factors in (non-networked) attack and defense games. Preferences and biases have not received substantial attention in the experimental or theoretical literature in these games, although it should be noted that Chowdhury et al. (2013) and Kovenock et al. (2019) both find that utility curvature does not appear to be an important factor in multi-target attack and defense games.

<sup>4</sup>See Kosfeld (2004) for a survey of network experiments more generally.

## 3 Theory and Hypotheses

### 3.1 Attacker Model

In order to describe the security manager’s (henceforth defender) problem, it is necessary to describe and justify the assumptions we make about the nature of the attacker that he faces. As our focus is on network defense by humans, in our main experimental task we automate the role of the attacker and describe their decision process to a human defender. We assume that the attacker observes the defender’s decision, has some fixed capability of attack, and linearly weights probabilities. While these assumptions may seem strong, they are consistent with a ‘worst-case’ approach, the motivation of which we now describe.

Due to the increasing inter-connectivity of cyber-physical systems to the outside world (e.g. through the internet), a defender faces a wide variety of possible attackers who can differ substantially in their resources, abilities and methods. The defender could undertake the challenging exercise of considering the attributes of all possible attackers, but this would involve many assumptions that the defender might get wrong. Instead, we assume that the defender takes a worst-case approach and defends against a sophisticated attacker, so that he can achieve a certain level of defense regardless of what type of attacker eventuates. The sophisticated attacker can be interpreted as the aggregate of all attackers perfectly colluding. They may also have the ability to observe the defender’s decision either through a period of monitoring or by using informants. Taking a worst-case approach is common in the security resource allocation literature (e.g. Yang et al. (2011), Nikoofal and Zhuang (2012), and Fielder et al. (2014)), as is the assumption that the attacker observes the defender’s allocation.<sup>5</sup>

### 3.2 Defender Model

The defender faces a network consisting of  $J$  total paths from the start node to the critical node, with each edge belonging to one or more of the  $J$  paths. The defender’s security decision is to allocate a security budget of  $B \in \mathbb{R}_{>0}$  units across the edges; this is represented by a vector  $x$  with  $N$  elements, where  $N$  is the number of edges. The edge defense function  $p(x_i)$  is a production technology that transforms the number of units allocated to edge  $i$  (denoted by  $x_i$ ) to the probability of stopping an attack (from the worst-case attacker) as it passes along edge  $i$ . We assume the defender has probability weighting from the one parameter model described in Prelec (1998), i.e.  $w(p(x_i); \alpha) = \exp[-(-\log(p(x_i)))^\alpha]$  with  $\alpha \in (0, 1]$ , although our findings hold with other ‘inverse-S’ shaped weighting functions (e.g., Tversky and Kahneman (1992)). For ease of notation we will frequently shorten  $w(p(x_i); \alpha)$  to  $w(p)$  or  $w(p(x_i))$ .

The defender gains a payoff of 1 if the critical node is not breached by the attacker, and gains a payoff of 0 if the attacker breaches the critical node. As the attacker observes the defender’s allocation and chooses the objectively most vulnerable path (i.e. the attacker has  $\alpha = 1$ ), the attacker’s action directly follows from a given allocation. However, the defender’s non-linear weighting of probabilities ( $\alpha < 1$ ) may cause him to have a different perception about which paths are the most vulnerable. Thus, the defender thinks the attacker will choose the path with the lowest *perceived* probability of successful defense (from the defender’s perspective, in accordance with their probability weighting parameter). The defender’s goal is to maximize his perceived probability of successfully defending the critical node, which is determined by his weakest perceived path. The defender’s optimization problem depends on the network structure, edge allocations, edge defense function  $p(x_i)$ , and his probability weighting parameter  $\alpha$ . We denote the defender’s overall perceived probability of defense along path  $j$  as  $f_j(x; \alpha)$ .

An attacker passing along an edge to reach a specific node is a separate and independent event from all other edges.<sup>6</sup> We assume the defender applies his weighting function to each probability individually before calculating the probability of overall defense along a path. The defender ranks the event of stopping an attack along a given edge higher than the event of an attack proceeding. Therefore, in accordance with Rank Dependent Utility (RDU) (Quiggin, 1982) and Cumulative Prospect Theory (CPT)

<sup>5</sup>For example, Bier et al. (2007), Modelo-Howard et al. (2008), Dighe et al. (2009), An et al. (2013), Hota et al. (2016), Nithyanand et al. (2016), Guan et al. (2017), Wu et al. (2018), and Leibowitz et al. (2019).

<sup>6</sup>The events are independent as each edge represents a unique layer of security that is unaffected by the events in other edges/layers of security. Breaches of other layers of security can affect whether a specific layer is encountered, but they do not change the probability that layer is compromised.

(Tversky and Kahneman, 1992), he applies his weighting function to the probability of stopping an attack along an edge ( $w(p)$ ), and considers the other event (the attack proceeding) to have a probability of  $1 - w(p)$ . Therefore, a path  $j$  with three edges has an overall perceived probability of defense of  $f_j(x; \alpha) = w(p(x_1)) + [1 - w(p(x_1))] [w(p(x_2)) + (1 - w(p(x_2)))w(p(x_3))]$ .<sup>7</sup> The defender’s constrained objective problem is presented in Equation 1.

$$\begin{aligned} \operatorname{argmax}_x \quad & \min\{f_1(x; \alpha), f_2(x; \alpha), \dots, f_J(x; \alpha)\} \\ \text{s.t.} \quad & x_i \geq 0, i = 1, 2, \dots, N \\ & \sum_i^N x_i \leq B \end{aligned} \tag{1}$$

We now consider the impact that non-linear probability weighting by a defender has on various network structures and defense production functions. We analyze the situation in a general setting, before considering the experimental design that we implement in the laboratory.

### 3.3 Common Edges

The described objective in Equation 1 is a straightforward constrained optimization problem. Unfortunately, the problem is analytically intractable and no closed-form solution exists. Consider our first type of network structure, presented in Figure 1. The key feature of this network is that one of the edges is common to both paths, while the other edges belong only to the top or bottom path. We denote  $x_3 = y$ , and assume that  $v = x_1 = x_2 = x_4 = x_5$ , an edge defense function of  $p(x_i) = 1 - e^{-\frac{x_i}{z}}$  (where  $z$  is some normalization parameter), and that  $v > 0$ ,  $y > 0$ . Even with these simplifications and assumptions, taking the first order conditions of the associated Lagrangian yields a set of equations that is intractable to solve for a closed form solution for either  $y$  or  $v$ .<sup>8</sup> Fortunately, it is possible to numerically solve the defender’s optimization problem. For example (and anticipating our experimental design), when  $z = 18.2$ ,  $B = 24$  and  $\alpha = 0.6$ , the optimal allocation is  $v = x_1 = x_2 = x_4 = x_5 = 1.26$  and  $y = x_3 = 18.96$ . Appendix 7.1 provides more analysis on how the numerical solution is calculated and whether the solution is unique.

The main trade-off in this type of network is the allocation to edges that are common to both paths or to edges that are only on one path. Consider taking a small amount  $\epsilon$  from the common edge  $x_3$  and placing it on a non-common edge. Placing the  $\epsilon$  only on one edge is non-optimal for any  $\alpha$ , as the sophisticated attacker will attack the weaker path, meaning  $\epsilon$  should be split across paths. This need to split over paths reduces the marginal impact of units allocated to the non-common edges on the overall probability of defense, making them relatively less attractive compared to the common edge. However, with non-linear probability weighting ( $\alpha < 1$ ), small probabilities are over-weighted, i.e. perceived to be higher than their actual probabilities. This increases the perceived impact of units placed on non-common edges, and can exceed the loss of having to split the allocation across more than one path. This makes expenditures on non-common edges more likely for those with non-linear probability weighting.

We can confirm this intuition numerically for a variety of edge defense functions. We mainly consider concave functions in our experiment, which have a natural interpretation of diminishing marginal returns of production.<sup>9</sup> In particular, consider the edge defense function from before ( $p(x_i) = 1 - e^{-\frac{x_i}{z}}$ ). Figure 2 plots the optimal amount to allocate to the common edge for different values of  $z$  and different levels of probability weighting  $\alpha$ . At  $\alpha = 1$  the optimal allocation is to place all  $B = 24$  units on the common edge. A defender with  $\alpha = 1$  will always place all of his units on the common edge for the exponential family

<sup>7</sup>This approach is similar to the concept of ‘folding back’ sequential prospects, as described in Epper and Fehr-Duda (2018) with regards to ‘process dependence’. The alternative (i.e.,  $f_j(x; \alpha) = w(p(x_1)) + [1 - w(p(x_1))] [p(x_2) + (1 - w(p(x_2)))p(x_3)]$ ) does not yield interesting comparative statics in  $\alpha$  due to the monotonicity of the probability weighting function, so we do not consider it further.

<sup>8</sup>Weighting the probability of a successful attack along an edge instead is analytically tractable as terms conveniently cancel, as shown in Abdallah et al. (2019b). However, this would be inconsistent with how events are ranked and weights are applied in RDU and CPT. Despite the lack of symmetry in the one parameter Prelec weighting function, the qualitative comparative statics presented in Abdallah et al. (2019b) have been numerically confirmed to hold in the current environment.

<sup>9</sup>Concavity and diminishing marginal returns is a common assumption in the computer security literature (e.g., Pal and Golubchik (2010), Boche et al. (2011), Sun et al. (2018), Feng et al. (2020))

of edge defense functions (Abdallah et al., 2019b). As  $\alpha$  decreases, i.e., the defender exhibits increasing levels of non-linear probability weighting, he places fewer units on the common edge (and more units on the non-common edges).

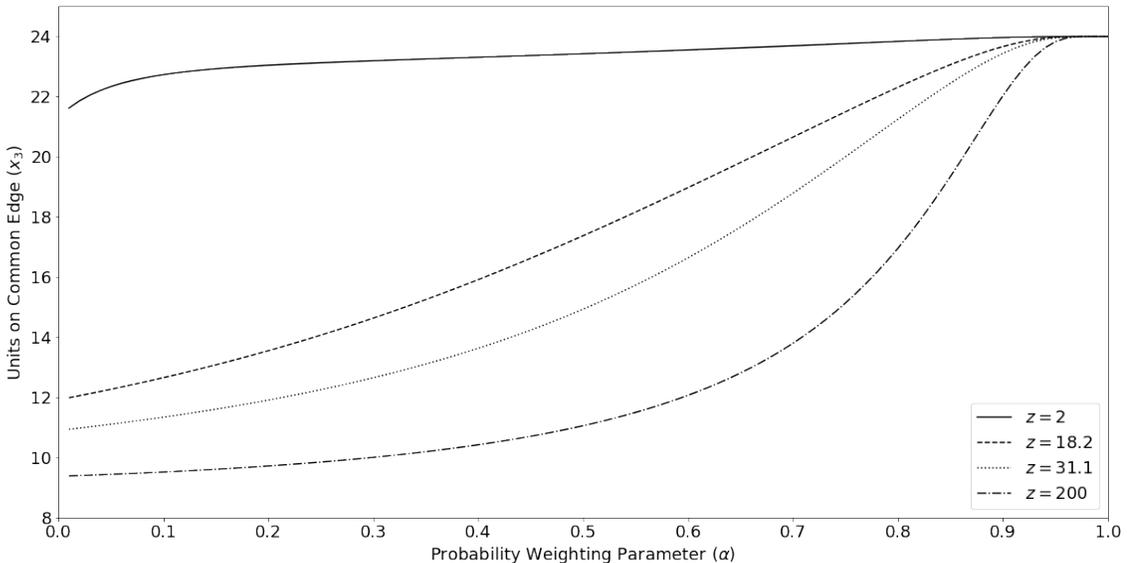


Figure 2: Allocation to Common Edge for  $p(x_i) = 1 - e^{-\frac{x_i}{z}}$

Consider next a non-exponential edge defense function  $p(x_i) = (\frac{x_i}{z})^b$ , where  $z$  is again a normalization factor and  $b \in (0, \infty)$ . If  $b < 1$ , this function is concave, if  $b = 1$  it is linear and if  $b > 1$  it is convex. Figure 3 illustrates that regardless of the convexity of the edge defense function, the amount allocated to the common edge decreases as  $\alpha$  decreases from 1. Note also that for concave functions of this form, it is no longer optimal for  $\alpha = 1$  defenders to place all of their allocation on the common edge. This is because the slope of the edge defense function for small values is sufficiently steeper than the slope of the function when all units are allocated to one edge. To see this, consider some  $p(x_i)$  and denote the number of units allocated to the non-common edges as  $v$ , and the number of units allocated to the common edge as  $y$ . Denoting the overall probability of a successful defense as  $F(v, y)$ , then:  $F(v, y) = p(\frac{v}{4}) + (1 - p(\frac{v}{4}))(p(\frac{v}{4}) + (1 - p(\frac{v}{4}))p(y))$ . Taking the first order conditions:  $\frac{\partial F(v, y)}{\partial v} = \frac{1}{2}p'(\frac{v}{4})[1 - p(\frac{v}{4}) - p(y) - p(\frac{v}{4})p(y)]$  and  $\frac{\partial F(v, y)}{\partial y} = p'(y)[1 - 2p(\frac{v}{4}) + p(\frac{v}{4})^2]$ . At the boundary solution corresponding to  $v = 0$  and  $y = B$ , if  $p(0) = 0$  the above expressions show that allocating all units to the common edge is optimal if  $p'(0)(1 - p(B)) \leq 2p'(B)$ , i.e., the marginal return to placing another unit on  $y$  exceeds that of  $v$  at the boundary. It follows that if the slope is sufficiently steep for small  $v$ 's (i.e.  $p'(0) > \frac{2p'(B)}{1-p(B)}$ ), then an  $\alpha = 1$  defender will allocate a strictly positive amount to non-common edges.<sup>10</sup>

These observations lead to our first testable hypotheses:

**Hypothesis 1** *The amount allocated to common edges (weakly) decreases as  $\alpha$  decreases from 1.*

**Hypothesis 2** *If  $p'(0) > \frac{2p'(B)}{1-p(B)}$  (such as for a concave power function), then a decision-maker with linear probability weighting ( $\alpha = 1$ ) will allocate a strictly positive amount to non-common edges.*

<sup>10</sup>Any  $\alpha \in (0, 1]$  defender is making a similar trade-off of  $\frac{\partial F(v, y)}{\partial v}$  against  $\frac{\partial F(v, y)}{\partial y}$ , either equating them if the solution is interior, or allocating to whichever is greater at the boundary. We do not present these first order conditions here as they are not as succinct due to the presence of  $w(p; \alpha)$ , although we do report the first order condition in Appendix 7.1. Where exactly the trade-off is resolved depends on  $\alpha$  as well as the specific functional form of  $p(x_i)$ . This is why the optimal allocation differs over  $\alpha$  for a given  $p(x_i)$ , as well as over different  $p(x_i)$  for a given  $\alpha$ . Both patterns are displayed in Figures 2 and 3.

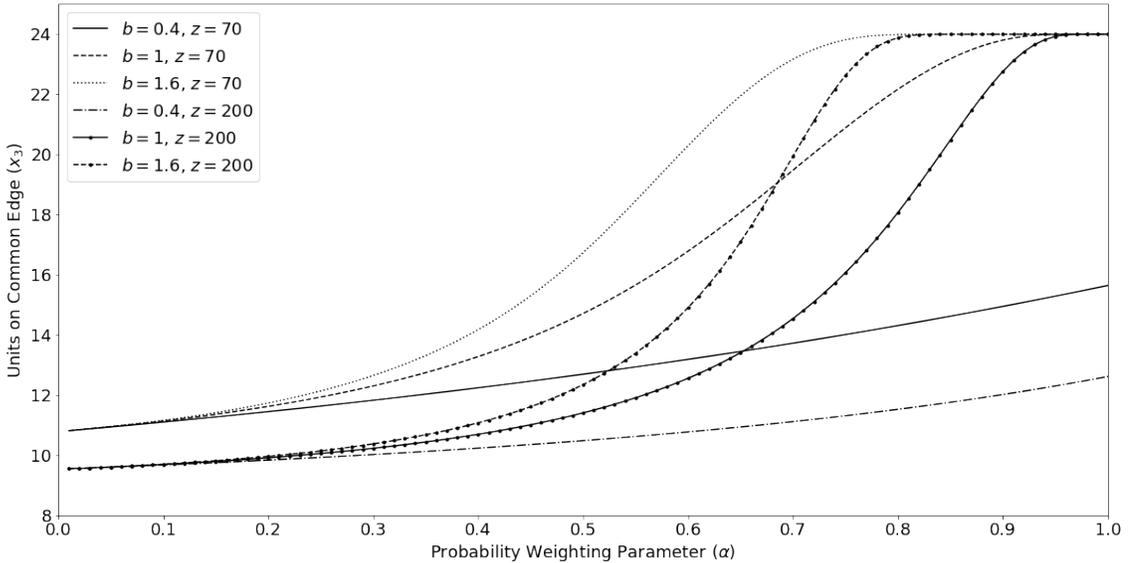


Figure 3: Allocation to Common Edge for  $p(x_i) = (\frac{x_i}{z})^b$

We now present the three color-coded networks from our experiment that are designed to explore these two hypotheses.

### 3.3.1 Network Red

Network Red employs the network structure presented earlier in Figure 1, and has an edge defense function of  $p(x_i) = 1 - e^{-\frac{x_i}{18.2}}$ .<sup>11</sup> According to Hypothesis 1, a defender with  $\alpha < 1$  will place less than 24 units on the common edge, and the amount placed on the common edge is decreasing as  $\alpha$  decreases from 1. For example, a defender with  $\alpha = 0.5$  will allocate  $x_3 = 17.36$ , and  $x_1 = x_2 = x_4 = x_5 = 1.66$ , while other  $\alpha$ 's are displayed graphically in Figure 2 by the line associated with  $z = 18.2$ .<sup>12</sup> According to Hypothesis 2, a defender with  $\alpha = 1$  would allocate  $x_3 = 24$ , and  $x_1 = x_2 = x_4 = x_5 = 0$ .

### 3.3.2 Network Orange

Network Orange also takes place on the network shown in Figure 1, but differs in having an edge defense function of  $p(x_i) = 1 - e^{-\frac{x_i}{31.1}}$ . The prediction for a defender with  $\alpha = 1$  remains unchanged from Network Red. Because  $p(x_i) \leq 0.46 \forall x_i \in [0, 24]$ , edge allocations in Network Orange mostly result in probabilities that a defender with  $\alpha < 1$  will overweight. Therefore, the predictions for a defender with a particular value of  $\alpha < 1$  will differ from Network Red. For example, a defender with  $\alpha = 0.5$  will now allocate  $x_3 = 14.92$ , and  $x_1 = x_2 = x_4 = x_5 = 2.27$ . The prediction for other  $\alpha$ 's is displayed in Figure 2 on the line associated with  $z = 31.1$ . The change in the edge defense function increases the separation of behavior between moderate to high levels of non-linear probability weighting, increasing our ability to detect differences between  $\alpha$  types.

### 3.3.3 Network Yellow

Network Yellow also takes place on the network shown in Figure 1. The edge defense function is now of a different concave functional form,  $p(x_i) = \frac{x_i^{0.4}}{70^{0.4}}$ . Unlike Networks Red and Orange, it is now optimal for a

<sup>11</sup>The normalization factor  $z = 18.2$  was chosen such that 1 unit allocated to an edge would yield a commonly overweighted probability ( $p = 0.05$ ), while 24 units allocated to an edge would yield a commonly underweighted probability ( $p = 0.73$ ).

<sup>12</sup>These numerical solutions are continuous, although subjects were restricted to discrete (integer-valued) allocations.

non-behavioral defender to allocate units to the non-common edges, in accordance with Hypothesis 2. In particular, a defender with  $\alpha = 1$  will allocate  $x_3 = 15.64$ , and  $x_1 = x_2 = x_4 = x_5 = 2.09$ , while a defender with  $\alpha = 0.5$  will allocate  $x_3 = 12.68$ , and  $x_1 = x_2 = x_4 = x_5 = 2.83$ . Predictions for other  $\alpha$ 's are presented in Figure 3, on the line associated with  $z = 70$ ,  $b = 0.4$ .

Networks Red, Orange, and Yellow are jointly designed to test Hypotheses 1 and 2. In all three of these networks, the amount allocated to the common edge should decrease as  $\alpha$  decreases, according to Hypothesis 1. In Networks Red and Orange, Hypothesis 2 predicts that those with  $\alpha = 1$  should place all 24 units on the common edge, while in Network Yellow, Hypothesis 2 predicts those with  $\alpha = 1$  should place less than 24 units on the common edge.

### 3.4 Extraneous Edges

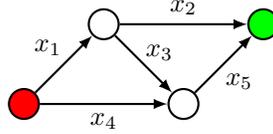


Figure 4: Network Graph with an Extraneous Edge ( $x_3$ )

Consider the network displayed in Figure 4. The new feature of this network is the edge denoted  $x_3$ , which creates a third possible path from the red node to the green node. In this network, a defender's overall perceived probability of defense is:

$$\begin{aligned}
 F(x; \alpha) = \min\{ & w(p(x_1); \alpha) + [1 - w(p(x_1); \alpha)] w(p(x_2); \alpha), \\
 & w(p(x_4); \alpha) + [1 - w(p(x_4); \alpha)] w(p(x_5); \alpha), \\
 & w(p(x_1); \alpha) + [1 - w(p(x_1); \alpha)] [w(p(x_3); \alpha) + (1 - w(p(x_3); \alpha))w(p(x_5); \alpha)] \}
 \end{aligned} \tag{2}$$

Call the possible paths as top (through  $x_1$  then  $x_2$ ), middle (through  $x_1$ , then  $x_3$ , then  $x_5$ ), and bottom (through  $x_4$  then  $x_5$ ). The optimal allocation will always equalize the perceived probability of successful defense for these three paths. Otherwise, the defender could increase utility by allocating an infinitesimal amount from a non-minimum path to the minimum path. Suppose  $x_1 = x_2 = x_4 = x_5 = \frac{B}{4}$ . The top, middle, and bottom paths all have the same perceived probability of successful defense at this allocation. Taking an infinitesimal  $\epsilon$  from any (or all) of these edges and placing it on  $x_3$  increases the perceived probability of defense of the middle path, but at the expense at the top and/or bottom path, which would now become the minimum path.

This solution of  $x_1 = x_2 = x_4 = x_5 = \frac{B}{4}$  and  $x_3 = 0$  is unique for any  $\alpha \in (0, 1)$  whenever the edge defense function has  $p'(x_i) > 0 \forall x_i$ . For  $\alpha = 1$  with the exponential defense function the solution is not unique, since any combination that allocates  $\frac{B}{2}$  to the top and bottom paths (which implies  $x_3 = 0$ ) is an optimal solution.<sup>13</sup> These results lead to our next testable Hypothesis:

**Hypothesis 3** *The amount allocated to extraneous edges is 0, and is invariant in  $\alpha$ .*

#### 3.4.1 Network Blue

Network Blue takes place on the network with an extraneous edge, as shown in Figure 4, with an edge defense function of  $p(x_i) = 1 - e^{-\frac{x_i}{18.2}}$ . The edge defense function for Network Blue (as well as the subsequent Network Green) is the same as Network Red, which reduces the number of different edge defense functions subjects have to consider in our within-subjects design. Network Blue is designed to test Hypothesis 3, as no subject with any  $\alpha \in (0, 1]$  should place any number of defense units on the extraneous edge labeled  $x_3$ . This network is useful to identify subjects with alternative behavioral biases.

<sup>13</sup>Further details are presented in Appendix 7.1.

### 3.5 Unequal Path Lengths

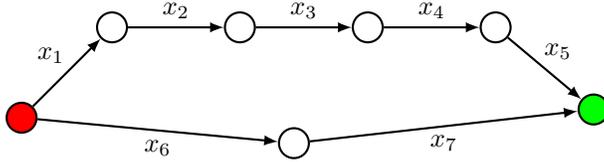


Figure 5: Network Graph with Unequal Path Lengths

Consider the network displayed in Figure 5. The key feature of this network is the different number of edges on each path. With an edge defense function of  $p(x_i) = 1 - e^{-\frac{x_i}{z}}$ , a defender with linear probability weighting should place half of his budget on each path. A defender with non-linear probability weighting, however, will place less of his budget on the top path (with more edges), and more of his budget on the bottom path (with fewer edges). To see the intuition behind this, consider a case where the defender starts with his allocation split equally across the two paths. Assume he spreads units allocated to a path equally across all edges in that path, and that the edge defense function yields a probability of less than  $\frac{1}{e}$  (i.e. the Prelec inflection point) on each edge along the top path but more than  $\frac{1}{e}$  on each edge along the bottom path. A defender with  $\alpha < 1$  over-weights small probabilities and perceives the small investments across the many edges along the top path as providing more protection that they actually do. Conversely, the  $\alpha < 1$  defender under-weights large probabilities and perceives investments across the two edges along the bottom path as providing less protection that they actually do. Consequently, such a defender should reallocate his investment to equalize his perceived probability of successful defense on the two paths, and this requires shifting some of the allocation from the top path to the bottom path.

This is also true for any combination of the top and bottom path edge probabilities that are above or below the inflection point of the probability weighting function. If both are above the inflection point, the defender perceives both paths as being weaker than they actually are, but perceives the bottom path as being relatively weaker due to the more extreme under-weighting of larger probabilities. Similar logic applies if both probabilities are below the inflection point, since the over-weighting is stronger for the smaller probabilities along the top path.

Again, the analytical solution proves intractable, but Figure 6 shows the numerical solutions considering the total relative allocations to edges in the top and bottom paths of the exponential edge defense functions of  $p(x_i) = 1 - e^{-\frac{x_i}{z}}$  for various values of  $z$ . The overall optimal allocation for  $\alpha \in (0, 1)$  occurs when equally spreading the total allocation to a path across each edge along a path, and this optimal allocation is unique. For example, for  $\alpha = .9$  the optimal allocation is  $x_1 = x_2 = x_3 = x_4 = x_5 = 2.23$ , and  $x_6 = x_7 = 6.43$ . For  $\alpha = 1$  the solution is not unique, as any solution that allocates  $\frac{B}{2}$  over the top and bottom paths is an optimal allocation. These results lead to our final testable hypothesis:

**Hypothesis 4** *The total amount allocated to a path with more edges decreases as  $\alpha$  decreases from 1.*

#### 3.5.1 Network Green

Network Green takes place on the network shown in Figure 5, again with an edge defense function of  $p(x_i) = 1 - e^{-\frac{x_i}{18.2}}$ . It is designed to test Hypothesis 4, as defenders with  $\alpha < 1$  should place fewer units on paths with fewer edges. For example, a defender with  $\alpha = 0.5$  would place  $x_1 = x_2 = x_3 = x_4 = x_5 = 0.964$  on each edge in the top path, and  $x_6 = x_7 = 9.59$  on each edge in the bottom path.

The most simple network that could address Hypothesis 4 is that of 2 edges for one path and 1 edge for the other path. However, we deliberately exaggerated the difference between the two paths in Network Green by having 5 edges on the top path and 2 edges on the bottom path. This results in increased separation in predicted behavior between subjects with different  $\alpha$ 's.

Table 1 summarizes the predictions for all five networks in the experiment for three levels of  $\alpha$ .

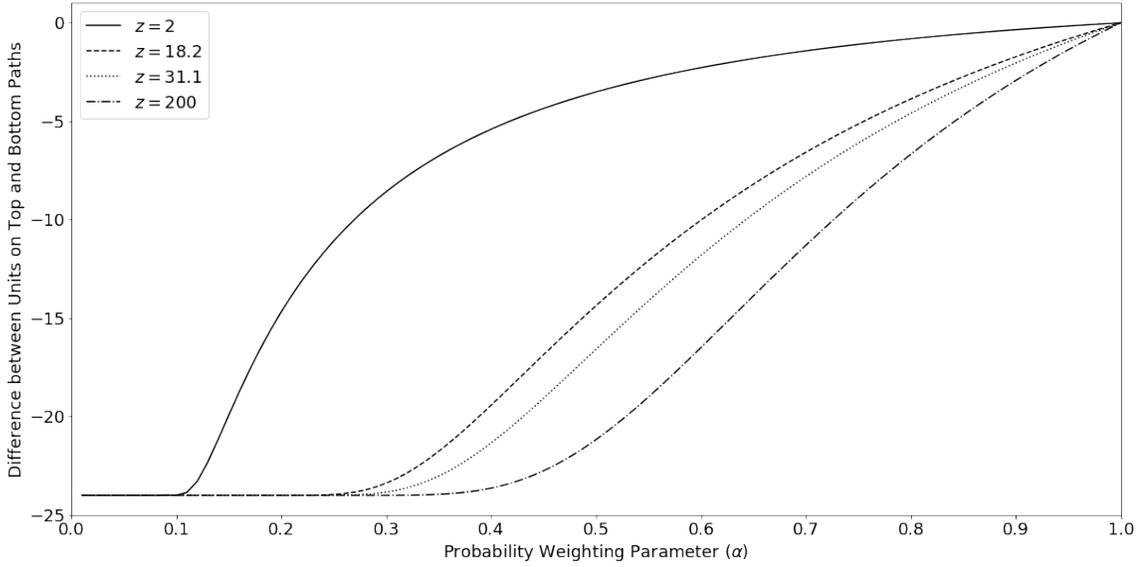
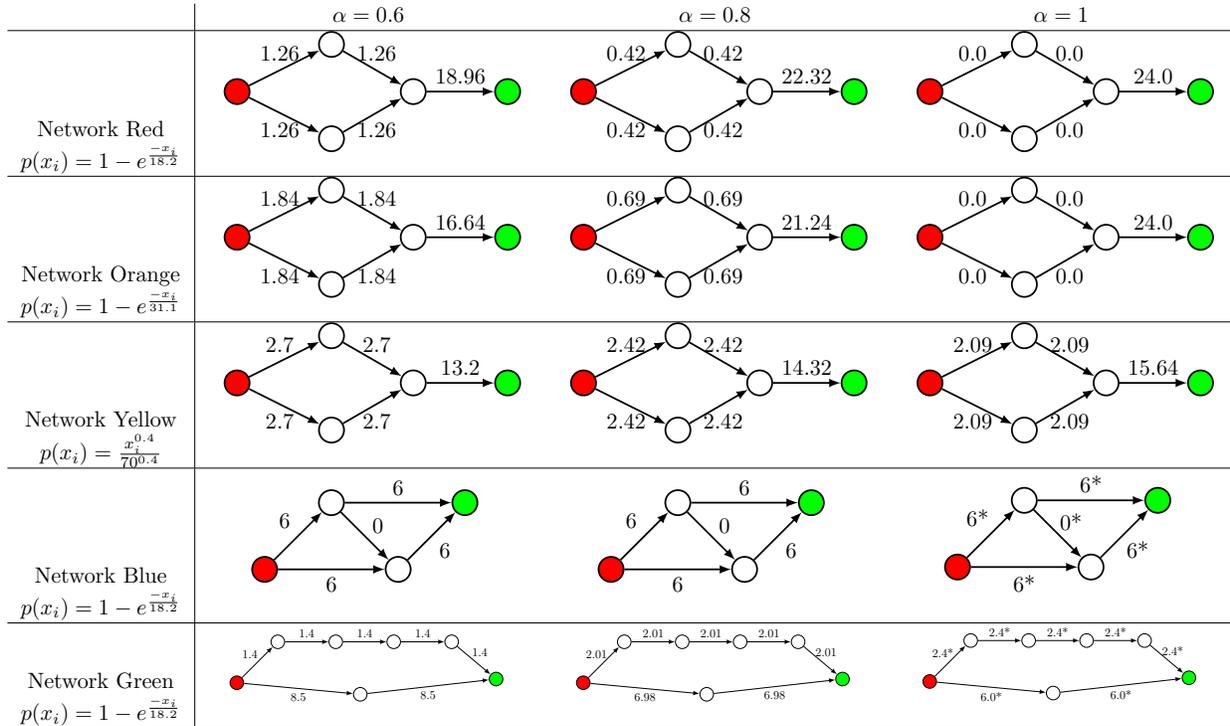


Figure 6: Allocation to Top (Long) minus Bottom (Short) Path for  $p(x_i) = 1 - e^{-\frac{x_i}{z}}$



\* Any combination that allocates  $\frac{B}{2}$  units to the top and bottom paths is optimal for  $\alpha = 1$ .

Table 1: Theoretical Predictions for Selected  $\alpha$  with  $B = 24$

## 4 Experimental Design

### 4.1 Probability Weighting Elicitation

Our main ex ante research question as well as our hypotheses focus on how the level of non-linear probability weighting affects security investment decisions. To directly relate subjects’ allocation behavior to probability weighting, we would like some external measure of probability weighting. In other words, we wish to have an accurate measure of  $\alpha$  that has good internal validity with the network security problem, while also not taking a substantial period of time away from the main Network Defense Task.

Many ways exist to elicit an individual’s probability weighting parameter. Typically researchers control for or simultaneously elicit risk preferences (taken here to mean the curvature of the utility function) when measuring probability weighting. This is because the specific range of probability weighting parameters that are consistent with a decision depends on the level of utility curvature assumed, and vice versa. However, in the defender’s problem considered here, utility curvature does not play a role as there are only two payoff outcomes. The defender either successfully defends the critical node, or does not. This means that the defender always wants to maximize their (perceived) probability of the high payoff outcome, which is invariant to utility curvature. Therefore, for the Network Defense Task we are not concerned about risk preferences, other than to parse out their effect to obtain an accurate measure of probability weighting.

With that in mind, we employ a new Network *Attack* Task as a way to measure probability weighting. In this task, we have subjects swap roles, i.e., they encounter a simplified version of this network environment in the role of an attacker against a computerized defender. Not only does this elicitation task reduce the procedural variance with respect to the main defense task, it also exploits the irrelevance of utility curvature in situations with two outcomes.

Consider the network in Figure 7, where the attacker’s goal is to successfully compromise the critical node, by choosing the top or bottom path to attack. The attacker receives 3000 points for compromising the critical node, and 0 points otherwise, meaning there are only two payoff outcomes. The numbers given on each edge represent the probability of a successful attack along this edge. Because the subject plays the role of an attacker (who ranks a successful attack along an edge higher than an unsuccessful attack) in this preliminary task, he weights the probability of successful attack along an edge when making his decision.

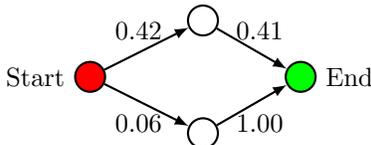


Figure 7: Network Attack Task Example

An attacker with  $\alpha = 1$  should choose the top path, which has a greater probability of overall success than the bottom path ( $0.42 \times 0.41 = .1722 > 0.06 = 0.06 \times 1.00$ ). However, an attacker with  $\alpha < 1$  may instead prefer to attack along the bottom path, due to over-weighting 0.06 and under-weighting 0.41 and 0.42. Assuming that the attacker applies his probability weighting function to each individual probability and then calculates the probability of success of each path, then the attacker would choose the top path if  $\alpha > 0.597$  and the bottom path if  $\alpha < 0.597$ . By asking for multiple responses with different probabilities (which imply different  $\alpha$  cutoffs),  $\alpha$  can be bounded.

Using a dynamic bisection or staircase method could recover increasingly tight bounds on  $\alpha$ , assuming subjects respond without error. Of course, subjects typically exhibit some level of noise in their decisions. Any mistake, especially early on in the elicitation procedure, would cause a bisection method to never be able to recover the subject’s true  $\alpha$ . A dynamic method that allows for the subject to make some errors is the Dynamically Optimized Sequential Experimentation (DOSE) method, as described in Chapman et al. (2018). In DOSE, the most informative question is asked based on the current Bayesian update of the subject’s parameters. The subject’s response is then used to update the current belief that the subject is of a given type, and this is then used to ask another question. The DOSE process recovers from errors as specific  $\alpha$  types are not ruled out completely as being the subject’s true type after an inconsistent response. Therefore, a subject’s consistent future responses can raise the procedure’s belief of the true type, and

adapt future questions accordingly. DOSE always asks the most informative question given the current belief distribution over types, meaning that fewer questions are required for an accurate measure. A full description of the DOSE procedure that was implemented for this task is presented in Appendix 7.2.

One potential concern with the Network Attack Task is that calculating the probability of a successful attack along a path is simply a case of multiplying the probabilities along the path.<sup>14</sup> Subjects may instead perform this step before applying their subjective probability weighting, instead of after as we have assumed. We therefore take two steps to make it more difficult for a subject to trivially multiply along paths. First, we avoid using probabilities that are more easily multiplied together (such as those that end in multiples of 0 or 5). Second, we did not allow subjects to use writing utensils or calculators during this task.

In our analyses we focus on  $\alpha$  estimates from this Network Attack Task because it reduces procedural variance from the main network defense tasks and focuses solely on probability weighting. We also measured  $\alpha$  using binary lottery choices derived from Multiple Price Lists (MPL) used to measure probability weighting (e.g., Tanaka et al. (2010), Bruhin et al. (2010)). Subjects choose between two lotteries one at a time (i.e., consider one row of an MPL in isolation), again using the DOSE procedure, also estimating an additional risk preference parameter. Details are also presented in Appendix 7.2.

## 4.2 Network Defense Tasks

For the Network Defense Tasks, subjects had a 24 ‘defense unit’ budget to use each period. These defense units could be allocated in integer amounts across edges. Defense units not used in one period did not roll over to the next period (i.e., this was a ‘use it or lose it’ situation). Subjects could submit a defense allocation of less than 24 units, but the software would prompt them to confirm they actually wanted to submit such an allocation.<sup>15</sup> Subjects chose the number of defense units to allocate to an edge using a dropdown menu that automatically updated the possible options based on the remaining number of units available. The initial value of this dropdown menu was not a number, meaning subjects had to make a selection for each edge, even if the desired allocation was zero. An example of the interface is shown in Appendix 7.6. Subjects play each of the five different networks 10 consecutive times to allow for some feedback and learning. The ordering of these five blocks was varied randomly across subjects.

## 4.3 Procedures

The experiments were conducted at the Vernon Smith Experimental Economics Laboratory (VSEEL). In total, 91 subjects participated, all students at Purdue University recruited from a subject database using ORSEE (Greiner, 2015).<sup>16</sup> Subjects received a packet of written instructions, some of which were printed on color paper that aligned with the color of the Network Defense Task.<sup>17</sup> Subjects were instructed to refer to specific instructions when the software (implemented in oTree (Chen et al., 2016)) prompted them to do so. Subjects participated in the Binary Lottery Task first, followed by the Network Attacker Task. During these first two tasks, as noted above subjects were not allowed to use calculators or writing utensils, and this was strictly enforced. Subjects then completed the colored Network Defense Tasks in an order that was varied randomly and unique to each subject. Subjects could request a calculator and pen from the experimenter during the Network Defense Tasks, due to the increased computational difficulty of these tasks. To simplify probability calculations, the instructions included a table for every network indicating how allocated defense resources mapped numerically into defense likelihood for any edge.

All payoffs were denoted in experimental points, with 350 points = \$1.00. Subjects received 3000 points in a round for successfully reaching the end node in the Network Attacker Game, and 1500 points for successfully preventing the computerized attacker from reaching the end node in a Network Defense Task. One round from each task was randomly selected for payment at the end of the experiment. Subjects were able to proceed through the tasks at their own pace, with most taking between 30-90 minutes (about 45

<sup>14</sup>Another potential issue for both tasks is that subjects may not understand this point at all, instead of finding it simple. The number of such subjects should be limited due to our subject pool being drawn from a university student population.

<sup>15</sup>In only 5 of the 4550 total decisions did subjects allocate less than all 24 units.

<sup>16</sup>Due to an error with the software, decision times were not recorded for 4 subjects. For consistency, we present our results only considering the remaining 87 subjects. Where the inclusion of decision times is not necessary, our results do not substantially change if the dropped observations are included.

<sup>17</sup>These instructions are available in Appendix 7.7.

minutes on average) and earning an average of \$20.10. To ensure subjects had read the instructions carefully, before each Network Defense Task subjects were asked to report the probability of two randomly selected rows (one from 1-12, one from 13-24) of the edge defense function for that task, and were paid an additional 50 points if they answered correctly.

## 5 Results

We begin the results with an overview of the probability weighting ( $\alpha$ ) elicitation from the Network Attack Task. We then consider the consistency of subject behavior between and within the Network Attack Task and Network Defense Tasks, including non-parametric tests of our Hypotheses. We then present a cluster analysis to broadly summarize the heterogeneity in the strategies that subjects employ. This identifies other possible biases that subjects exhibit. Finally, we present a regression analysis on key defense allocations that controls for the identified biases and other important factors like cognitive ability and decision time.

### 5.1 Network Attack Task

The main purpose of the Network Attack Task is to obtain an estimate of an individual subject’s probability weighting, parameterized by  $\alpha$ . A useful comparison point for our results comes from Bruhin et al. (2010), who estimate a finite mixture model on certainty equivalents for lotteries elicited over many Multiple Price Lists. They find evidence for two groups, with approximately 20% of subjects exhibiting near linear probability weighting, and the remaining 80% of subjects exhibiting non-linear probability weighting.

Figure 8 presents the CDF for the subjects’ elicited  $\alpha$ ’s from the Network Attack Task. Considerable heterogeneity exists in the degree of non-linear probability weighting, so our Hypotheses predict heterogeneity in the Network Defense Tasks as well. Considering the quintiles of the distribution, we have 20% of subjects with  $\alpha \geq 0.95$ , 20% with  $0.90 \leq \alpha < 0.95$ , 20% with  $0.80 \leq \alpha < 0.90$ , 20% with  $0.64 \leq \alpha < 0.80$ , and finally the bottom quintile with  $\alpha < 0.64$ . This suggests the presence of both relatively linear and non-linear probability weighting groups. Our results are in line with the 20% of subjects exhibiting linear weighting as in Bruhin et al. (2010), albeit with our second highest quintile being somewhat linear as well.

**Result 1** *Considerable heterogeneity exists in the inferred  $\alpha$  from the Network Attack Task. The quintile cutoff points are  $\alpha = 0.64$ ,  $\alpha = 0.80$ ,  $\alpha = 0.90$ ,  $\alpha = 0.95$ .*

### 5.2 Network Defense Tasks

#### 5.2.1 Summary

Figure 9 presents the CDF’s of mean defense allocations for key subject decisions in each of the five Network Defense Tasks. This also indicates substantial heterogeneity in subject behavior. In the Red and Orange networks, 26.4% and 16.1% of subjects respectively allocate all units to the common edge, and this fraction of subjects decreases to 13.8% in Network Yellow. This suggests that a relatively small proportion of subjects exhibit behavior consistent with an  $\alpha$  near 1. However, about 40% of subjects in all three of these networks allocate less defense to the common edge than can be justified even for very low levels of  $\alpha$ , suggesting a role for additional behavioral biases. About one-third of subjects allocate no units to the extraneous edge in Network Blue, in accordance with Hypothesis 3, while 46.0% allocate more than 1 unit on average to the extraneous edge. This again suggests a role for other behavioral biases. The Network Green CDF indicates that 19.5% of subjects allocate equal amounts to the top and bottom paths, consistent with  $\alpha = 1$ , and 25.3% of subjects allocate less units to the top path, consistent with  $\alpha < 1$ . However, over half of the subjects allocate more to the top path (and many quite substantially so), which is the opposite of what Hypotheses 4 predicts for  $\alpha < 1$ . Overall, the CDFs provide some casual evidence in support of probability weighting playing a role in subject behavior, but that other biases appear to influence behavior as well.

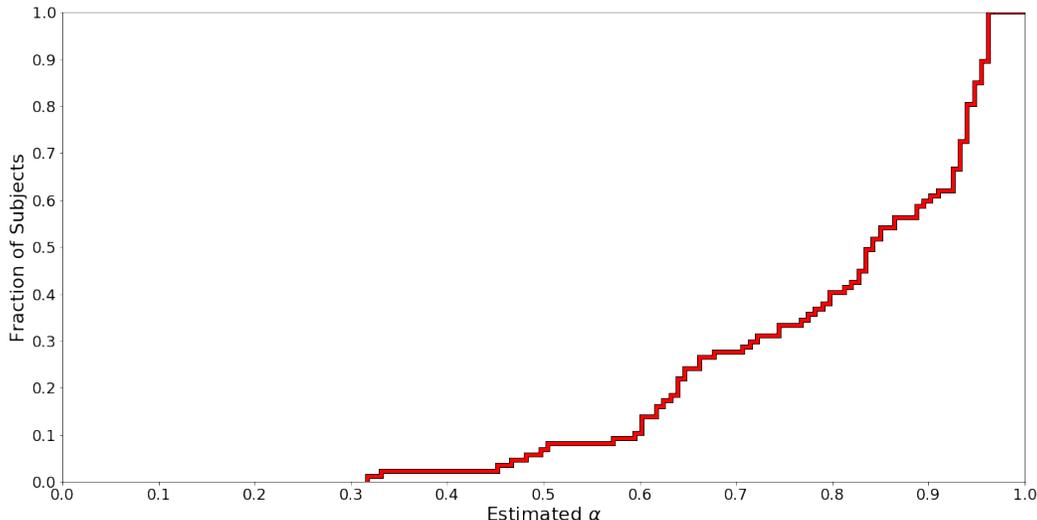


Figure 8: CDF of elicited  $\alpha$  from the Network Attack Task

### 5.2.2 Subject Consistency and Non-parametric Tests

We first consider individual subject consistency between and within Network Defense and Attack Tasks. Recall that our measure of  $\alpha$  is derived from the Network Attack Task, which we use to test our Hypotheses in the Network Defense Task. Table 2 presents the non-parametric Spearman’s  $\rho$  between the elicited probability weighting ( $\alpha$ ) and decision noise ( $\lambda$ ) from the Network Attack Task and key average subject behavior in each of the Network Defense Tasks.<sup>18,19</sup> The decision noise parameter  $\lambda$  is estimated by the logit function, a commonly used structure in Quantal Response Equilibria (McKelvey and Palfrey, 1995).<sup>20</sup> A higher  $\lambda$  is consistent with less noisy behavior, meaning subjects choose their payoff maximizing action more frequently. Table 2 indicates consistency within the Network Defense Tasks as correlations are strongly statistically significant in all but one of the pairwise comparisons between these tasks. We also observe consistency between behavior in the Network Attack and Defense Tasks, with the leftmost column of Table 2 reporting statistically significant correlations in all but one comparison.

We now consider non-parametric tests of Hypotheses 1, 3, and 4, all of which are included in the leftmost column of Table 2. The elicited  $\alpha$  from the Network Attack Task is strongly and significantly correlated with defense in all Network Defense Tasks except Network Yellow.<sup>21</sup> The correlations of  $\alpha$  with the common edge networks are positive, consistent with Hypothesis 1. The negative correlation in Network Green indicates that subjects with estimated  $\alpha$ ’s closer to 1 tend to place less defense on the top path. This is the opposite of Hypothesis 4. The negative correlation in Network Blue for defense resources placed on the extraneous edge provides evidence against Hypothesis 3. This suggests that our external measure of  $\alpha$  from the Network Attack Task also captures some element of cognitive ability. This interpretation is consistent with the strong correlation of decision noise ( $\lambda$ ) with probability weighting ( $\alpha$ ).

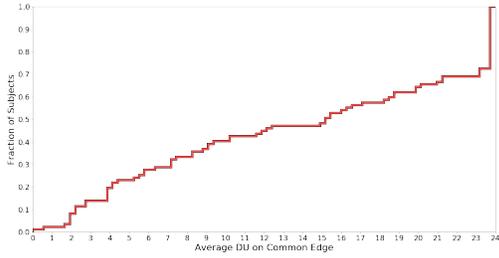
To conduct a non-parametric test of Hypothesis 2, we use the Wilcoxon signed-rank test. In particular, we compare the paired observations of an individual subject’s average allocation to the common edge in Network Yellow to Networks Red and Orange. According to Hypothesis 2, those with an  $\alpha$  close to 1 should

<sup>18</sup>We consider the same analysis including the Binary Lottery Task in Appendix 7.4. The elicited  $\alpha$ ’s of these tasks are not correlated ( $\rho = 0.166$ ,  $p = 0.117$ ), suggesting the procedural differences are important, or that cognitive ability may play a role.

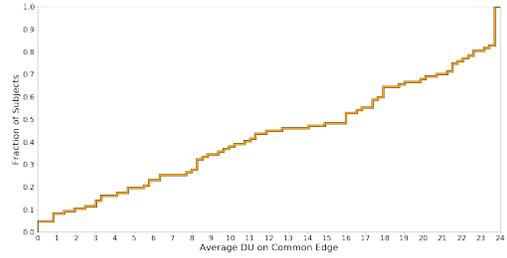
<sup>19</sup>Unless otherwise stated, all p-values and statistical tests are two-sided.

<sup>20</sup> $Prob(Top) = \frac{1}{1+e^{-\lambda(U(Top)-U(Bottom))}}$  if  $U(Top) \geq U(Bottom)$ ,  $Prob(Top) = \frac{1}{1+e^{-\lambda(U(Bottom)-U(Top))}}$  otherwise, where  $U(Top)$  is the weighted then compounded probability of successful attack multiplied by the payoff from a successful attack.

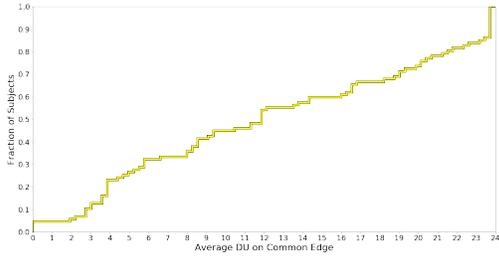
<sup>21</sup>The lack of a significant correlation in Network Yellow is not necessarily surprising, due to the deliberate reduction of the separation of  $\alpha$  types in this network to evaluate Hypothesis 2.



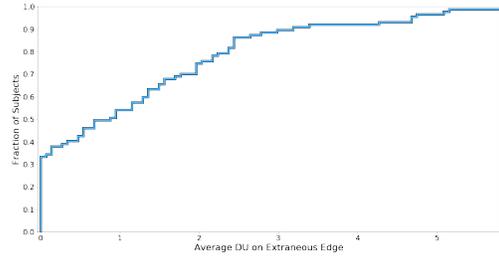
(a) Network Red - Common Edge



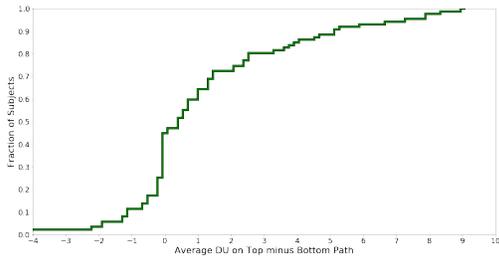
(b) Network Orange - Common Edge



(c) Network Yellow - Common Edge



(d) Network Blue - Extraneous Edge



(e) Network Green - Top minus Bottom

Figure 9: CDFs of per-subject average behavior across all rounds in the Network Defense Tasks

exhibit a particularly pronounced decrease in this key allocation for Network Yellow. As we have a clear directional theoretical prediction we report one-sided p-values for this test. Considering subjects with an estimated  $\alpha \geq .9$  (i.e. the 40th percentile closest to the linear  $\alpha = 1$ ), we find a statistically significant decrease at the five percent level when testing Network Red against Yellow ( $p = 0.031$ ), as well as when we test Network Orange against Yellow ( $p = 0.017$ ). We find no similar difference in subjects with an estimated  $\alpha < .9$  ( $p = 0.451$  and  $p = 0.447$  respectively). These results are robust at the five percent level for any  $\alpha$  cutoff  $\in [.86, .95]$  (see Appendix 7.3).

$\lambda$	$\alpha$	$\lambda$	Red Common Edge	Orange Common Edge	Yellow Common Edge	Blue Extra Edge
Red Common Edge	$\rho = 0.764^{***}$	$\rho = 1$	$\rho = 1$			
Orange Common Edge	$\rho = 0.260^{**}$	$\rho = 0.151$	$\rho = 0.654^{***}$	$\rho = 1$		
Yellow Common Edge	$\rho = 0.072$	$\rho = -0.058$	$\rho = 0.622^{***}$	$\rho = 0.646^{***}$	$\rho = 1$	
Blue Extra Edge	$\rho = -0.286^{***}$	$\rho = -0.230^{**}$	$\rho = -0.352^{***}$	$\rho = -0.308^{***}$	$\rho = -0.112$	$\rho = 1$
Green Top-Bottom	$\rho = -0.255^{**}$	$\rho = -0.139$	$\rho = -0.241^{**}$	$\rho = -0.292^{***}$	$\rho = -0.237^{**}$	$\rho = 0.241^{**}$

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Table 2: Spearman's  $\rho$  Correlation Table

**Result 2** *The probability weighting parameter  $\alpha$  is positively correlated with allocations to the common edge in the Red and Orange Networks, consistent with Hypothesis 1. Subjects with  $\alpha \geq 0.9$  allocate less to the common edge in Network Yellow as compared to Networks Red and Orange, consistent with Hypothesis 2.  $\alpha$  is negatively correlated with allocations to the extra edge in Network Blue and the top path in Network Green, inconsistent with Hypotheses 3 and 4.*

These initial results should be considered more as a guide to the analysis rather than an exhaustive test of our hypotheses. For instance, the Spearman correlation is a bivariate measure that does not control for any other possible biases and observed characteristics of the subject. We therefore conduct a cluster analysis in the following subsection to identify additional biases. We then conduct a regression analysis that controls for the identified biases and other factors like cognitive ability.

### 5.2.3 Cluster Analysis

The previous subsection documents that probability weighting is associated with defense misallocation in this network defense environment, but not always in the manner originally hypothesized. Other behavioral biases also appear important, but these biases are not clear ex-ante. It is also likely that subjects exhibit heterogeneity in these biases and that these biases may interact, making them difficult to predict or otherwise identify. One way to summarize general patterns of subject behavior is using a cluster analysis. In particular, we use the method of Affinity Propagation (Frey and Dueck, 2007), which endogenously determines the number of clusters. We cluster at the session level, i.e., a subject’s average behavior across individual networks, as we consider their behavior to be related across tasks.<sup>22</sup>

Table 3 presents the ‘exemplar’ of each of the 10 clusters, summarizing an individual subject’s behavior that is the most representative of that cluster. The leftmost column presents the percentage of subjects represented by that cluster, alongside a descriptive name to aid exposition.

Clusters 1 and 2 appear largely consistent with an  $\alpha = 1$ , meaning that approximately 17% of subjects exhibit linear probability weighting through their network defense decisions. This is close to the 20% as reported in Bruhin et al. (2010).

The cluster analysis identifies three additional biases, which along with probability weighting can describe the behavior of the cluster exemplars. The first bias is that of naive diversification: when subjects are given  $n$  options to invest in, they have a tendency towards investing  $1/n$  units to each option (Benartzi and Thaler, 2001). Note that this is especially *naive* naive diversification, as the edges do not represent different assets, just different ways to protect the same asset (the critical node). Naive diversification explains Cluster 4 particularly well, but can also explain situations where less units are placed on the common edge that can be justified by probability weighting alone. A defender with  $\alpha < 1$  as well as some mild preference towards evening out his allocation on the common and non common edges would place even less on the common edge than his level of  $\alpha$  would predict. Some level of naive diversification clearly explains non-zero allocations to the extraneous edge in Network Blue. Naive diversification can also explain the tendency for subjects to place more units on the top rather than the bottom path in Network Green; as the top path has more edges, a  $1/n$  heuristic would place more units overall on the top path.

The second and third biases are related to each other, and we term them early or late revelation of the overall outcome. Early revelation means that subjects try to stop the attack as soon as possible, and thus allocate more to edges nearer to the start node on the left. Clusters 6 and 10 are good examples of early revelation. Late revelation is the opposite, referring to subjects that allocate more units to edges nearer to the critical node, as exemplified by Cluster 9. Early revelation can explain an excessively low allocation to the common edge, in a manner similar to naive diversification except that more units are placed on the front two non-common edges instead of equally to all non-common edges. Late revelation can explain the failure of some subjects to reduce their common edge allocation in Network Yellow. Note that, like the naive diversification bias, this is an especially naive preference as the outcome is revealed immediately after the allocation decision is made, and importantly, all at once. In the experiment there is no animation that sequentially displays the attacker’s progress. Therefore, holding anticipatory emotions such as dread over

<sup>22</sup>We also cluster at the individual network task level in an alternative estimation presented in Appendix 7.5. That analysis identifies similar patterns of behavior.

Cluster Name	Network Red	Network Orange	Network Yellow	Network Blue	Network Green
C1: Near Optimal $\alpha = 1$ - Late Revelation 12.6%					
C2: Near Optimal $\alpha = 1$ - Early Revelation 4.6%					
C3: Late Revelation - Some Diversification 4.6%					
C4: Naive Diversification 16.1%					
C5: $\alpha < 1$ - Some Diversification 12.6%					
C6: Early Revelation - Some Diversification 9.2%					
C7: $\alpha < 1$ - Mild Diversification 16.1%					
C8: Naive Diversification and Late Revelation 5.7%					
C9: Late Revelation - mild Diversification 13.8%					
C10: Early Revelation - mild Diversification 4.6%					

Table 3: Cluster Analysis. Left column displays the percentage of subjects classified in each cluster. Numbers on edges denote average DU (out of 24 total) allocated by the exemplar subject for that cluster.

a period of time is minimized within an attack.<sup>23</sup> The concepts of early and late revelation are related to the literature on anticipatory utility with regards to the revelation of uncertainty (e.g., Loewenstein (1987), Caplin and Leahy (2001)).

#### 5.2.4 Regression Analysis

The cluster analysis identifies additional biases that may interact with probability weighting and influence subject behavior. To address more directly our original hypotheses regarding the behavioral implications of probability weighting, we account for these other biases by including appropriate measures in a regression analysis. In addition to these control variables, we include additional independent variables to investigate systematically how they influence behavior.

Ex-ante we did not anticipate the additional biases, and therefore did not specifically design separate elicitation tasks to identify them. Fortunately, our Blue and Green networks allow us to measure subjects' naive diversification and early/late revelation preferences, which can then be used as controls when considering behavior in other networks.

Our main naive diversification measure is calculated from a subject's allocation to the extraneous edge in Network Blue. Specifically, we calculate each individual's average allocation to this extra edge. However, this measure clearly does not work for Network Blue, as it is based on behavior in this network. Therefore, in order to obtain a measure of naive diversification for Network Blue, we use behavior from Network Green. A fully naive individual would allocate  $\frac{24}{7} = 3.4$  units to each edge in Network Green, so we calculate the average absolute distance of each edge from this equal spread. A fully naive individual would have a measure of 0, and the most extreme optimal allocation of 12 units to one top and bottom edge would have a measure of  $\frac{3.4 \times 5 + 8.6 \times 2}{7} = 4.88$ . We then multiply this measure by  $-1$ , so that the comparative static is comparable with the measure based on Network Blue, which has naive individuals having a higher (rather than lower) value of this measure.

For early/late revelation we consider the Blue and Green networks without the common edge, as expressing this preference is far less costly in these networks. Furthermore, early and late revelation should not impact allocations to the dependent variable in the regressions for Networks Blue and Green, so we omit this particular independent variable for those networks. The early revelation measure is based on the ratio of units allocated to the nearest two edges to the start node, and the nearest two edges to the critical node, averaged for the Blue and Green networks.

The regressions also add variables to account for cognitive ability, as the results from Table 2 suggest that  $\alpha$  may be picking up some measure of cognitive ability.<sup>24</sup> We include information self-reported by subjects in a post-experiment survey, such as field of study, a high GPA, and whether a subject is a graduate student. We also include decision times and time spent with the instructions, as this may be correlated with subjects' understanding. Finally, we note that gender is of particular interest ex-ante based on previous observations that women tend to exhibit greater non-linear probability weighting on average than men (Fehr-Duda et al. (2006), Bruhin et al. (2010), Fehr-Duda et al. (2011)).

Table 4 reports a series of censored tobit regressions, with the dependent variable for each network corresponding to the key summary statistics shown earlier in Figure 9: the allocation to the common edge for Networks Red, Orange, and Yellow, the allocation to the extraneous edge in Network Blue, and the difference in allocations to the top and bottom paths for Network Green. The regressions are censored at 0 to 24 for all networks except Network Green, which is censored -24 to 24.

The top row shows the effect that our measure of probability weighting ( $\alpha$ , estimated from the Attacker task) has after controlling for the identified additional biases and the other subject characteristics. Consistent with Hypothesis 1, amounts allocated to the common edge in Networks Red, Orange and Yellow is increasing in  $\alpha$ . These coefficients are statistically significant in Networks Orange and Yellow, but not in Network Red. These results provide partial further evidence in support of Hypothesis 1 when combined with our correlation results. According to Hypothesis 3,  $\alpha$  should not have an effect on the amount allocated to the extra edge in Network Blue. After controlling for naive diversification, we no longer find a statistically significant affect of  $\alpha$  on this allocation. Finally, Hypothesis 4 predicts that increasing  $\alpha$  should increase the amount allocated to

<sup>23</sup>It is of course possible that subjects are playing out the attack process in their imagination, while reading the outcomes sequentially.

<sup>24</sup>Choi et al. (2018) reports evidence suggesting a correlation between cognitive ability and probability weighting.

	Red Common Edge	Orange Common Edge	Yellow Common Edge	Blue Extra Edge	Green Top-Bottom
$\alpha$ (Attacker Task)	15.89 (1.28)	19.06* (1.66)	27.86*** (2.73)	-3.928 (-1.19)	-4.341** (-2.24)
$\mu$ (Attacker Task)	-0.0110 (-0.30)	-0.0220 (-0.65)	-0.0823*** (-2.78)	-0.0116 (-1.12)	0.00745 (1.30)
Naive Diversification †	-3.395*** (-3.13)	-2.830*** (-2.81)	-1.713** (-2.04)	1.425*** (4.10)	0.578*** (3.40)
Early Revelation ‡	-11.24*** (-3.65)	-9.604*** (-3.34)	-12.74*** (-5.32)		
Time Spent on Decision	-0.0119 (-1.46)	-0.0220** (-2.05)	-0.00924 (-1.21)	0.00133 (0.44)	-0.000939 (-0.32)
Total Time Spent on Instructions	-0.00492 (-0.57)	-0.00598 (-0.74)	-0.0109 (-1.61)	-0.00604** (-2.53)	0.000438 (0.33)
Age	-0.144 (-0.31)	0.492 (1.13)	0.615* (1.65)	0.0103 (0.08)	0.0646 (0.89)
Born in USA	-1.107 (-0.35)	-4.795* (-1.66)	-2.621 (-1.09)	1.081 (1.28)	-0.898* (-1.86)
Period	0.188** (2.18)	0.491*** (4.28)	0.0911 (1.11)	-0.246*** (-5.15)	0.0110 (0.22)
Male	3.908 (1.25)	3.978 (1.39)	1.763 (0.73)	-1.351 (-1.58)	-0.200 (-0.41)
Economics Major	8.190 (1.34)	3.155 (0.56)	3.147 (0.66)	0.114 (0.07)	1.167 (1.23)
Engineering Major	9.388** (2.03)	7.358* (1.72)	3.273 (0.92)	1.523 (1.15)	-1.887*** (-2.63)
Science Major	2.574 (0.56)	3.318 (0.77)	1.476 (0.41)	1.443 (1.14)	-0.500 (-0.70)
Management Major	-3.095 (-0.63)	-0.754 (-0.17)	-3.492 (-0.92)	-0.0188 (-0.01)	-1.034 (-1.34)
GPA > 3.5	2.108 (0.70)	-6.652** (-2.38)	1.431 (0.61)	0.514 (0.62)	-0.479 (-1.02)
Graduate Student	3.447 (0.81)	-3.169 (-0.81)	-4.232 (-1.27)	-1.117 (-0.94)	0.0124 (0.02)
Constant	8.911 (0.56)	-2.588 (-0.17)	-9.053 (-0.70)	10.54** (2.49)	3.125 (1.24)
Observations	870	870	870	870	870

$t$  statistics in parentheses

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

† Generated from Blue for Red, Orange, and Yellow, and from Green for Blue.

‡ † Generated from the average of Blue and Green.

Table 4: Tobit Regression Analysis

the top path in Network Green. While there is a statistically significant affect of  $\alpha$ , it is not in the direction predicted by Hypothesis 4. This is surprising because the regression controls for naive diversification, which should account for some subjects' tendency to allocate relatively more to the top path than the bottom.

**Result 3** *After controlling for other biases,  $\alpha$  is a statistically significant predictor of behavior in Networks Orange and Yellow (evidence in support of Hypothesis 1), and not in Network Blue (evidence in support of Hypothesis 3).  $\alpha$  is a statistically significant predictor of behavior in Network Green, but in the opposite direction than predicted (evidence against Hypothesis 4).*

We now consider the impact of naive diversification, which is predicted to decrease the amount allocated to the common edge, increase the amount allocated to the extra edge in Network Blue, and increase the amount allocated to paths with more edges (i.e. the top path in Network Green). Table 4 shows that naive diversification has a negative and significant impact on the amount allocated to the common edge in all three common edge networks. Naive diversification also has a positive effect on the number of units allocated to the extra edge in Network Blue, and to the top path in Network Green, all as predicted.

**Result 4** *A higher level of preference for naive diversification is correlated with a lower allocation to the common edge in Networks Red, Orange, and Yellow. It is also correlated with a higher allocation to the extra edge in Network Blue, and the longer top path in Network Green.*

Finally we consider early/late revelation, which only impacts the dependent variable for the common edge networks. Early revelation is predicted to decrease the amount allocated to the (late) common edge. The results show that a preference for early revelation has a strong and highly significant negative effect on the amount allocated to the common edge in all common edge networks.

**Result 5** *A higher preference for early revelation, measured using Networks Blue and Green, is correlated with a lower allocation to the common edge in Networks Red, Orange, and Yellow.*

The other independent variables include a period variable to capture the time trend, which suggests that some learning occurs as subjects gain experience with a particular network. The only other independent variable that is statistically significant over more than two networks is whether the student was an Engineering major, which has a statistically significant effect in the direction of optimal behavior in three networks. This suggests that cognitive ability or mathematical sophistication could promote better understanding and performance in this network defense problem.

## 6 Conclusion and Discussion

Cybersecurity and network defense is becoming increasingly important for economic, social, and even political activity. Both the financial and non-pecuniary costs of successful cyberattacks can be substantial, and thus it is important to minimize their likelihood. We investigate how behavioral biases, in particular probability weighting, could lead to sub-optimal defense allocations. We modeled the situation as a directed network graph, to capture in a simple way some trade-offs that security professionals face. Probability weighting has differing effects on various network structures and defense functions, which generates testable hypotheses. We found that a separately elicited measure of probability weighting ( $\alpha$ ) has a statistically significant correlation with key defense allocations in most Network Defense Tasks, including a network where probability weighting is predicted to have no effect. Motivated by this finding, we used a cluster analysis to identify additional biases that could also influence defense behavior. We identify preferences for naive diversification and for earlier or later revelation of attack outcomes. Controlling for these biases and other subject characteristics, we find evidence that probability weighting has predictive power in this environment, as do preferences for both naive diversification and early/late revelation.

An important question is how applicable are the findings from our student subject pool for security experts. We are not excessively concerned about this for several reasons. Firstly, a security expert may exhibit ‘other-evaluation’ (Curley et al., 1986). In the event of a successful attack, a security expert must justify his decision to others within his or her organization. If these other individuals exhibit biases, the expert may allocate in accordance to these biases to more easily justify their decision post-attack. Secondly, even if security experts exhibit fewer or weaker biases, given the magnitude of potential losses even very small biases could have a large impact on welfare. Finally, the empirical evidence on differences in behavior between students and experts is weak. Fr chet te (2015) conducts a survey of experiments that considered behavior of students compared to experts in a wide variety of professions. This survey reports only one out of thirteen considered studies found that professionals make decisions more closely in line with standard economic theory. Considering security professionals specifically, Mersinas et al. (2016) find that while security professionals do calculate expected values better than students, they also exhibit systematic biases such as ambiguity aversion and framing effects. We therefore consider the findings from our student subject pool to be sufficiently informative and useful to be taken seriously by cyber-security researchers.

Another important question is how robust our findings are to learning. We find some evidence of learning in our networks, suggesting that biases may reduce over time as subjects receive feedback and become more familiar with the task. The question is whether these biases vanish in the long run, or whether they persist. The ten rounds used for each network environment is likely insufficient for subjects to fully learn given the complexity of the environment. The number of rounds was a practical constraint, trading-off time spent in the lab against the overall number of network structures. It would be interesting to see how behavior evolves over longer repetitions of play, but that is beyond the scope of this paper.

There are many possible avenues for future research. First, theoretical work could incorporate the additional biases into a model over directed networks. This would be a very challenging endeavor. For example, consider observing an allocation of 2 on each non-common edge and 16 on the common edge in Network Red. A wide variety of situations could be consistent with this allocation, such as  $\alpha \approx 0.66$ , or any  $\alpha \in [.66, 1]$  with some level of naive diversification, or an  $\alpha < .66$  but with some preference for late revelation, or  $\alpha = 1$  having mild diversification preferences interacting with a stronger preference for late revelation, and so on. Adding to this complexity, it is not clear how the additional biases should be defined across different types of networks, or how they should interact with each other. For example, consider a subject who consistently places 2 units on the extraneous edge in Network Blue. This subject clearly has some preference for diversification, but what does that imply for his decision in Network Red? Many possibilities

exist. He could be facing a minimum constraint of 2 units per edge to satisfy a diversification preference, or he could allocate 2 units to each edge initially and allocate the remaining 14 units according to his weighting parameter  $\alpha$  (either disregarding or regarding the 2 units already allocated). Or he could be willing to give up a small amount in terms of perceived probability from his optimal strategy in order to more evenly spread his allocation, etc. Although there are many different ways that this could be modeled over different networks, the literature currently offers no guidance for explicit functional forms over directed networks to discipline these modeling decisions.

A second line of future research could incorporate strategic considerations by having human decision-makers interact with each other, in either roles of attacker and a defender, or multiple defenders on the same network defending the same or different critical nodes. For example, it may be in a defender's best interest to allocate his resources differently if he believes the attacker to have  $\alpha < 1$ . Alternative network structures in both the Network Defense and Network Attack Tasks could also be worth investigating, particularly in light of the identified naive diversification and early/late revelation biases. Third, it is not clear why  $\alpha$  has a significant impact of behavior in Network Green in the opposite direction that is predicted. It may be the case that our elicitation of  $\alpha$  is only picking up on cognitive ability. Future research could investigate why results from Network Green are anomalous, perhaps with an alternative elicitation of  $\alpha$  or naive diversification or additional controls of cognitive ability. Finally, the effect of probability weighting in more standard attack and defense games has not yet received much attention. Given the empirical relevance of it in the current environment, this may prove to be an interesting avenue to explore.

## References

- Mustafa Abdallah, Parinaz Naghizadeh, Timothy Cason, Saurabh Bagchi, and Shreyas Sundaram. Protecting Assets with Heterogeneous Valuations under Behavioral Probability Weighting. In *2019 IEEE Conference on Decision and Control (CDC)*, pages 5374–5379, 2019a.
- Mustafa Abdallah, Parinaz Naghizadeh, Ashish R. Hota, Timothy Cason, Saurabh Bagchi, and Shreyas Sundaram. The Impacts of Behavioral Probability Weighting on Security Investments in Interdependent Systems. In *2019 American Control Conference (ACC)*, pages 5260–5265, Philadelphia, 2019b.
- Daron Acemoglu, Azarakhsh Malekian, and Asu Ozdaglar. Network security and contagion. *Journal of Economic Theory*, 166:536–585, 11 2016. ISSN 10957235. doi: 10.1016/j.jet.2016.09.009.
- Alessandro Acquisti and Jens Grossklags. What can behavioral economics teach us about privacy. In Alessandro Acquisti, Stefanos Gritzalis, Costos Lambrinoudakis, and Sabrina di Vimercati, editors, *Digital privacy: theory, technologies and practices*, chapter 18, pages 363–377. Auerbach Publications, 2007.
- Fadele Ayotunde Alaba, Mazliza Othman, Ibrahim Abaker Targio Hashem, and Faiz Alotaibi. Internet of Things security: A survey. *Journal of Network and Computer Applications*, 88:10–28, 6 2017. ISSN 1084-8045. doi: 10.1016/J.JNCA.2017.04.002. URL <https://www.sciencedirect.com/science/article/pii/S1084804517301455#bib27>.
- Bo An, Matthew Brown, Yevgeniy Vorobeychik, and Milind Tambe. Security games with surveillance cost and optimal timing of attack execution. In *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 223–230, 2013.
- Shlomo Benartzi and Richard H. Thaler. Naive Diversification Strategies in Defined Contribution Savings Plans. *The American Economic Review*, 91(1):79–98, 3 2001. URL <https://www.jstor.org/stable/2677899>.
- Vicki Bier, Santiago Oliveros, and Larry Samuelson. Choosing what to protect: Strategic defensive allocation against an unknown attacker. *Journal of Public Economic Theory*, 9(4):563–587, 2007.
- Han Bleichrodt and Jose Luis Pinto. A Parameter-Free Elicitation of the Probability Weighting Function in Medical Decision Analysis. *Management Science*, 46(11):1485–1496, 11 2000. ISSN 0025-1909. doi: 10.1287/mnsc.46.11.1485.12086. URL <http://pubsonline.informs.org/doi/abs/10.1287/mnsc.46.11.1485.12086>.
- Francis Bloch, Bhaskar Dutta, and Marcin Dziubinski. A game of hide and seek in networks. 1 2020. URL <http://arxiv.org/abs/2001.03132>.
- Holger Boche, Siddharth Naik, and Tansu Alpcan. Characterization of convex and concave resource allocation problems in interference coupled wireless systems. *IEEE Transactions on Signal Processing*, 59(5):2382–2394, 2011.
- Adrian Bruhin, Helga Fehr-Duda, and Thomas Epper. Risk and Rationality: Uncovering Heterogeneity in Probability Distortion. *Econometrica*, 78(4):1375–1412, 7 2010. ISSN 0012-9682. doi: 10.3982/ECTA7139. URL <http://doi.wiley.com/10.3982/ECTA7139>.
- Michael Callen, Mohammad Isaqzadeh, James D. Long, and Charles Sprenger. Violence and Risk Preference: Experimental Evidence from Afghanistan. *American Economic Review*, 104(1):123–148, 1 2014. ISSN 0002-8282. doi: 10.1257/aer.104.1.123. URL <http://pubs.aeaweb.org/doi/10.1257/aer.104.1.123>.
- Andrew Caplin and John Leahy. Psychological Expected Utility Theory and Anticipatory Feelings. *The Quarterly Journal of Economics*, 116(1):55–79, 2 2001. doi: 10.1162/003355301556347. URL <https://academic.oup.com/qje/article-lookup/doi/10.1162/003355301556347>.
- Andrew Caplin and John Leahy. The Supply of Information by a Concerned Expert. *The Economic Journal*, 114(497):487–505, 7 2004. doi: 10.1111/j.0013-0133.2004.0228a.x. URL <https://academic.oup.com/ej/article/114/497/487-505/5085695>.

- Jonathan Chapman, Erik Snowberg, Stephanie Wang, and Colin Camerer. Loss Attitudes in the U.S. Population: Evidence from Dynamically Optimized Sequential Experimentation (DOSE). Technical report, National Bureau of Economic Research, Cambridge, MA, 9 2018. URL <http://www.nber.org/papers/w25072.pdf>.
- Daniel L. Chen, Martin Schonger, and Chris Wickens. oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97, 3 2016. ISSN 2214-6350. doi: 10.1016/J.JBEF.2015.12.001. URL <https://www.sciencedirect.com/science/article/pii/S2214635016000101?via%3Dihub>.
- Syngjoo Choi, Jeongbin Kim, Eungik Lee, and Jungmin Lee. Probability Weighting and Cognitive Ability. 2018.
- Subhasish M Chowdhury. The Attack and Defense Mechanisms - Perspectives from Behavioral Economics and Game Theory. *Behavioral and Brain Sciences*, 42:e121, 2019. doi: 10.1017/S0140525X19000815.
- Subhasish M. Chowdhury, Dan Kovenock, and Roman M. Sheremeta. An experimental investigation of Colonel Blotto games. *Economic Theory*, 52(3):833–861, 2013. ISSN 09382259. doi: 10.1007/s00199-011-0670-2.
- Subhasish M Chowdhury, Dan Kovenock, David Rojo Arjona, and Nathaniel T Wilcox. Focality and Asymmetry in Multi-battle Contests. 2016. URL [https://digitalcommons.chapman.edu/esi\\_working\\_papers/194/](https://digitalcommons.chapman.edu/esi_working_papers/194/).
- Derek J. Clark and Kai A. Konrad. Asymmetric Conflict: Weakest Link against Best Shot. *Journal of Conflict Resolution*, 51(3):457–469, 6 2007. doi: 10.1177/0022002707300320. URL <http://journals.sagepub.com/doi/10.1177/0022002707300320>.
- Shawn P Curley, J Frank Yates, and Richard A Abrams. Psychological Sources of Ambiguity Avoidance. *Organizational Behavior and Human Decision Processes*, 38(2):230–256, 1986.
- Cary Deck and Roman M. Sheremeta. Fight or Flight?: Defending against Sequential Attacks in the Game of Siege. *Journal of Conflict Resolution*, 56(6):1069–1088, 12 2012. doi: 10.1177/0022002712438355. URL <http://journals.sagepub.com/doi/10.1177/0022002712438355>.
- Nikhil S Dighe, Jun Zhuang, and Vicki M Bier. Secrecy in defensive allocations as a strategy for achieving more cost-effective attacker deterrence. *International Journal of Performability Engineering*, 5(1):31–43, 2009.
- Behnud Mir Djawadi, Angelika Endres, Britta Hoyer, and Sonja Recker. Network formation and disruption - An experiment are equilibrium networks too complex? *Journal of Economic Behavior and Organization*, 157:708–734, 1 2019. ISSN 01672681. doi: 10.1016/j.jebo.2018.11.004.
- Marcin Dziubiński and Sanjeev Goyal. Network design and defence. *Games and Economic Behavior*, 79(1): 30–43, 5 2013. doi: 10.1016/j.geb.2012.12.007.
- Marcin Dziubiński and Sanjeev Goyal. How do you defend a network? *Theoretical Economics*, 12(1):331–376, 1 2017. ISSN 1555-7561. doi: 10.3982/te2088.
- Thomas Epper and Helga Fehr-Duda. The Missing Link: Unifying Risk Taking and Time Discounting. 2018.
- Helga Fehr-Duda, Manuele de Gennaro, and Renate Schubert. Gender, Financial Risk, and Probability Weights. *Theory and Decision*, 60(2-3):283–313, 5 2006. doi: 10.1007/s11238-005-4590-0. URL <http://link.springer.com/10.1007/s11238-005-4590-0>.
- Helga Fehr-Duda, Thomas Epper, Adrian Bruhin, and Renate Schubert. Risk and rationality: The effects of mood and decision rules on probability weighting. *Journal of Economic Behavior & Organization*, 78(1-2): 14–24, 4 2011. ISSN 0167-2681. doi: 10.1016/J.JEBO.2010.12.004. URL <https://www.sciencedirect.com/science/article/pii/S0167268111000035#bib0025>.

- Shaohan Feng, Zehui Xiong, Dusit Niyato, Ping Wang, Shaun Shuxun Wang, and Xuemin Sherman Shen. Joint pricing and security investment in cloud security service market with user interdependency. *IEEE Transactions on Services Computing*, 2020.
- Andrew Fielder, Emmanouil Panaousis, Pasquale Malacaria, Chris Hankin, and Fabrizio Smeraldi. Game theory meets information security management. In *International Information Security Conference (IFIP)*, pages 15–29, 2014.
- Guillaume R. Fréchette. Experiments: Professionals Versus Students. In Guillaume Fréchette and A. Schotter, editors, *Handbook of Experimental Economic Methodology*, chapter 17, pages 360–390. Oxford University Press, 2015.
- Brendan J. Frey and Delbert Dueck. Clustering by passing messages between data points. *Science*, 315, 2007. URL <https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.121.3145>.
- Gartner. Gartner Forecasts Worldwide Information Security Spending to Exceed \$124 Billion in 2019, 2018. URL <https://www.gartner.com/en/newsroom/press-releases/2018-08-15-gartner-forecasts-worldwide-information-security-spending-to-exceed-124-billion-in-2019>.
- Jacob K. Goeree, Charles A. Holt, and Thomas R. Palfrey. Risk averse behavior in generalized matching pennies games. *Games and Economic Behavior*, 45(1):97–113, 10 2003. ISSN 0899-8256. doi: 10.1016/S0899-8256(03)00052-6. URL <https://www.sciencedirect.com/science/article/pii/S0899825603000526>.
- Sanjeev Goyal and Adrien Vigier. Attack, Defence, and Contagion in Networks. *The Review of Economic Studies*, 81(4):1518–1542, 10 2014. doi: 10.1093/restud/rdu013. URL <https://academic.oup.com/restud/article-lookup/doi/10.1093/restud/rdu013>.
- Ben Greiner. Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1):114–125, 7 2015. ISSN 2199-6776. doi: 10.1007/s40881-015-0004-4. URL <http://link.springer.com/10.1007/s40881-015-0004-4>.
- Peiqiu Guan, Meilin He, Jun Zhuang, and Stephen C Hora. Modeling a multitarget attacker–defender game with budget constraints. *Decision Analysis*, 14(2):87–107, 2017.
- Charles A Holt and Susan K Laury. Risk Aversion and Incentive Effects. *American Economic Review*, 92(5):1644–1655, 11 2002. ISSN 0002-8282. doi: 10.1257/000282802762024700. URL <http://pubs.aeaweb.org/doi/10.1257/000282802762024700>.
- John Homer, Su Zhang, Xinming Ou, David Schmidt, Yanhui Du, S. Raj Rajagopalan, and Anoop Singhal. Aggregating vulnerability metrics in enterprise networks using attack graphs. *Journal of Computer Security*, 21(4):561–597, 2013. doi: 10.3233/JCS-130475.
- Ashish R. Hota, Abraham A. Clements, Shreyas Sundaram, and Saurabh Bagchi. Optimal and Game-Theoretic Deployment of Security Investments in Interdependent Assets. pages 101–113. Springer, 2016. doi: 10.1007/978-3-319-47413-7\_{\\_}6. URL [http://link.springer.com/10.1007/978-3-319-47413-7\\_6](http://link.springer.com/10.1007/978-3-319-47413-7_6).
- Ashish R. Hota, Abraham A. Clements, Saurabh Bagchi, and Shreyas Sundaram. A Game-Theoretic Framework for Securing Interdependent Assets in Networks. In Stefan Rass and Stefan Schauer, editors, *Game Theory for Security and Risk Management: From Theory to Practice*, pages 157–184. Springer, 2018. doi: 10.1007/978-3-319-75268-6\_{\\_}7. URL [http://link.springer.com/10.1007/978-3-319-75268-6\\_7](http://link.springer.com/10.1007/978-3-319-75268-6_7).
- Britta Hoyer and Stephanie Rosenkranz. Determinants of Equilibrium Selection in Network Formation: An Experiment. *Games*, 9(4):89, 11 2018. ISSN 2073-4336. doi: 10.3390/g9040089. URL <http://www.mdpi.com/2073-4336/9/4/89>.
- Abdulmalik Humayed, Jingqiang Lin, Fengjun Li, and Bo Luo. Cyber-Physical Systems Security—A Survey. *IEEE Internet of Things Journal*, 4(6):1802–1831, 12 2017. ISSN 2327-4662. doi: 10.1109/JIOT.2017.2703172. URL <http://ieeexplore.ieee.org/document/7924372/>.

- Sumeet Jauhar, Binbin Chen, William G. Temple, Xinshu Dong, Zbigniew Kalbarczyk, William H. Sanders, and David M. Nicol. Model-Based Cybersecurity Assessment with NESCOR Smart Grid Failure Scenarios. In *2015 IEEE 21st Pacific Rim International Symposium on Dependable Computing (PRDC)*. IEEE, 11 2015. ISBN 978-1-4673-9376-8. doi: 10.1109/PRDC.2015.37.
- Michael Kosfeld. Economic Networks in the Laboratory: A survey. *Review of Network Economics*, 3(1), 2004.
- Dan Kovenock and Brian Roberson. The Optimal Defense of Networks of Targets. *Economic Inquiry*, 56(4): 2195–2211, 10 2018. doi: 10.1111/ecin.12565. URL <http://doi.wiley.com/10.1111/ecin.12565>.
- Dan Kovenock, Brian Roberson, and Roman M. Sheremeta. The attack and defense of weakest-link networks. *Public Choice*, 179(3-4):175–194, 6 2019. ISSN 15737101. doi: 10.1007/s11127-018-0618-1.
- Edward Lee. The Past, Present and Future of Cyber-Physical Systems: A Focus on Models. *Sensors*, 15(3): 4837–4869, 2 2015. ISSN 1424-8220. doi: 10.3390/s150304837. URL <http://www.mdpi.com/1424-8220/15/3/4837>.
- Hemi Leibowitz, Ania M Piotrowska, George Danezis, and Amir Herzberg. No Right to Remain Silent: Isolating Malicious Mixes. In *28th USENIX Security Symposium (USENIX Security 19)*, pages 1841–1858, Santa Clara, CA, 8 2019. USENIX Association. ISBN 978-1-939133-06-9. URL <https://www.usenix.org/conference/usenixsecurity19/presentation/leibowitz>.
- George Loewenstein. Anticipation and the Valuation of Delayed Consumption. *The Economic Journal*, 97(387):666, 9 1987. doi: 10.2307/2232929. URL <https://academic.oup.com/ej/article/97/387/666-684/5190020>.
- Jennifer M. Logg, Julia A. Minson, and Don A. Moore. Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151:90–103, 3 2019. ISSN 07495978. doi: 10.1016/j.obhdp.2018.12.005.
- Michael McBride and David Hewitt. The enemy you can't see: An investigation of the disruption of dark networks. *Journal of Economic Behavior & Organization*, 93:32–50, 9 2013. ISSN 01672681. doi: 10.1016/j.jebo.2013.07.004. URL <https://linkinghub.elsevier.com/retrieve/pii/S0167268113001704>.
- Richard D. McKelvey and Thomas R. Palfrey. Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior*, 10(1):6–38, 7 1995. ISSN 0899-8256. doi: 10.1006/GAME.1995.1023. URL <https://www.sciencedirect.com/science/article/pii/S0899825685710238>.
- Konstantinos Mersinas, Bjoern Hartig, Keith M. Martin, and Andrew Seltzer. Are information security professionals expected value maximizers?: An experiment and survey based test. *Journal of Cybersecurity*, 2(1):57–70, 12 2016. doi: 10.1093/cybsec/tyw009.
- Gaspar Modelo-Howard, Saurabh Bagchi, and Guy Lebanon. Determining Placement of Intrusion Detectors for a Distributed Application through Bayesian Network Modeling. In *11th International Symposium on Research in Attacks, Intrusions and Defenses (RAID)*, pages 271–290, 2008.
- Kien C. Nguyen, Tansu Alpcan, and Tamer Basar. Stochastic Games for Security in Networks with Interdependent Nodes. 3 2010. URL <http://arxiv.org/abs/1003.2440>.
- Mohammad E Nikoofal and Jun Zhuang. Robust allocation of a defensive budget considering an attacker's private information. *Risk Analysis: An International Journal*, 32(5):930–943, 2012.
- Rishab Nithyanand, Oleksii Starov, Adva Zair, Phillipa Gill, and Michael Schapira. Measuring and mitigating AS-level adversaries against Tor. In *Network & Distributed System Security Symposium (NDSS)*, 2016.
- Ranjan Pal and Leana Golubchik. Analyzing self-defense investments in internet security under cyber-insurance coverage. In *2010 IEEE 30th International Conference on Distributed Computing Systems*, pages 339–347. IEEE, 2010.

- M. Elisabeth Paté-Cornell, Marshall Kuypers, Matthew Smith, and Philip Keller. Cyber Risk Management for Critical Infrastructure: A Risk Analysis Model and Three Case Studies. *Risk Analysis*, 38(2):226–241, 2 2018. ISSN 15396924. doi: 10.1111/risa.12844.
- Drazen Prelec. The Probability Weighting Function. *Econometrica*, 66(3):497, 5 1998. ISSN 00129682. doi: 10.2307/2998573. URL <https://www.jstor.org/stable/2998573?origin=crossref>.
- John Quiggin. A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4):323–343, 12 1982. ISSN 0167-2681. doi: 10.1016/0167-2681(82)90008-7. URL <https://www.sciencedirect.com/science/article/pii/0167268182900087>.
- Roman M. Sheremeta. The attack and defense games. *Behavioral and Brain Sciences*, 42:e140, 8 2019. ISSN 0140-525X. doi: 10.1017/S0140525X19000931. URL [https://www.cambridge.org/core/product/identifier/S0140525X19000931/type/journal\\_article](https://www.cambridge.org/core/product/identifier/S0140525X19000931/type/journal_article).
- Oleg Sheyner and Jeannette Wing. Tools for generating and analyzing attack graphs. In *International Symposium on Formal Methods for Components and Objects (FMCO)*, pages 344–371. Springer, 2003. doi: 10.1007/978-3-540-30101-1{-}17.
- Xiaofang Sun, Chao Shen, Tsung-Hui Chang, and Zhangdui Zhong. Joint resource allocation and trajectory design for UAV-aided wireless physical layer security. In *2018 IEEE Globecom Workshops (GC Wkshps)*, pages 1–6. IEEE, 2018.
- Tomomi Tanaka, Colin F Camerer, and Quang Nguyen. Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam. *American Economic Review*, 100(1):557–571, 3 2010. ISSN 0002-8282. doi: 10.1257/aer.100.1.557. URL <http://pubs.aeaweb.org/doi/10.1257/aer.100.1.557>.
- Amos Tversky and Daniel Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323, 10 1992. ISSN 0895-5646. doi: 10.1007/BF00122574. URL <http://link.springer.com/10.1007/BF00122574>.
- Di Wu, Hui Xiao, and Rui Peng. Object defense with preventive strike and false targets. *Reliability Engineering & System Safety*, 169:76–80, 2018.
- Peng Xie, Jason H. Li, Xinming Ou, Peng Liu, and Renato Levy. Using Bayesian networks for cyber security analysis. In *Proceedings of the International Conference on Dependable Systems and Networks (DSN)*, pages 211–220, 2010. ISBN 9781424475018. doi: 10.1109/DSN.2010.5544924.
- Rong Yang, Christopher Kiekintveld, Fernando Ordonez, Milind Tambe, and Richard John. Improving resource allocation strategy against human adversaries in security games. In *Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI)*, 2011.

## 7 Appendices

### 7.1 Discussion of Numerical Solutions and Unique Solutions

#### 7.1.1 Common Edge Networks - Exponential Edge Defense Function

As stated in the text, the defender's problem is one of constrained optimization. For the exponential defense function, denoting the amount allocated to non-common edges as  $x$ , the amount allocated to the common edge as  $y$ , and also denoting  $\alpha$  as  $a$ , his objective function with the budget constraint substituted for  $x$  is the following:

$$\begin{aligned} & \left(1 - e^{-\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a}\right) \left(\left(1 - e^{-\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a}\right) e^{-\left(-\log\left(1 - e^{-\frac{y}{z}}\right)\right)^a}\right. \\ & \left. + e^{-\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a}\right) + e^{-\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a} \end{aligned} \quad (3)$$

Differentiating with respect to  $y$  yields the following first order condition:

$$\begin{aligned} & -\frac{a\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a e^{-\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a} e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\left(\left(1 - e^{-\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a}\right) e^{-\left(-\log\left(1 - e^{-\frac{y}{z}}\right)\right)^a} + e^{-\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a}\right)}{4z\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)} \\ & + \frac{a\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a e^{-\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a} e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}}{4z\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)} \\ & + \left(1 - e^{-\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a}\right) \left(-\frac{a\left(-\log\left(1 - e^{-\frac{y}{z}}\right)\right)^a e^{-\left(-\log\left(1 - e^{-\frac{y}{z}}\right)\right)^a} e^{-\frac{y}{z}}\left(1 - e^{-\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a}\right)}{z\left(1 - e^{-\frac{y}{z}}\right)\log\left(1 - e^{-\frac{y}{z}}\right)}\right) \\ & - e^{-\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a} \\ & + \frac{a\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a e^{-\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a} e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}}{4z\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)} \\ & - \frac{a\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a e^{-\left(-\log\left(1 - e^{-\frac{y}{z}}\right)\right)^a} e^{-\left(-\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\right)^a} e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}}{4z\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)\log\left(1 - e^{-\frac{1}{z}\left(\frac{B}{4} - \frac{y}{4}\right)}\right)} \right) = 0 \end{aligned} \quad (4)$$

Clearly it is practically infeasible to obtain a closed form solution for  $y$  in terms of  $B$ ,  $z$ , and  $a$ . However, it is trivial to compute the above first order condition for a given  $y$ ,  $B$ ,  $z$ , and  $a$ , meaning it is feasible to numerically calculate the optimal allocations. Figure 10 presents such an analysis graphically for selected combinations of  $B$ ,  $z$ , and  $a$ .

The top left graph in Figure 10 is an example of where the budget is very high, which makes the first order condition (in blue) undefined for large  $y$  due to the term  $\log\left(1 - e^{-\frac{B-y}{4z}}\right)$ . The inner term tends to zero as  $B - y$  increases, which then implies  $\log(0)$  which is mathematically impossible. As the first order condition does not cross (but approaches) zero before it becomes undefined, we turn to a numerical approach that maximizes the perceived probability of overall defense (in black). For this case the solution is not unique

as multiple allocations can obtain a perceived probability of what is computationally indistinguishable from one due to machine precision. We consider this an appropriate prediction as our computer is operating on levels of precision far in excess of human subjects. We find numerically that the excess budget issue occurs when  $B > 37.42z$ , which is independent of  $\alpha$  as  $w_p(1; \alpha) = 1 \quad \forall \alpha \in (0, 1]$ . This constraint on the budget is not relevant for the parameters we used in the experiment ( $B = 24, z = 18.2, z = 31.1$ ).

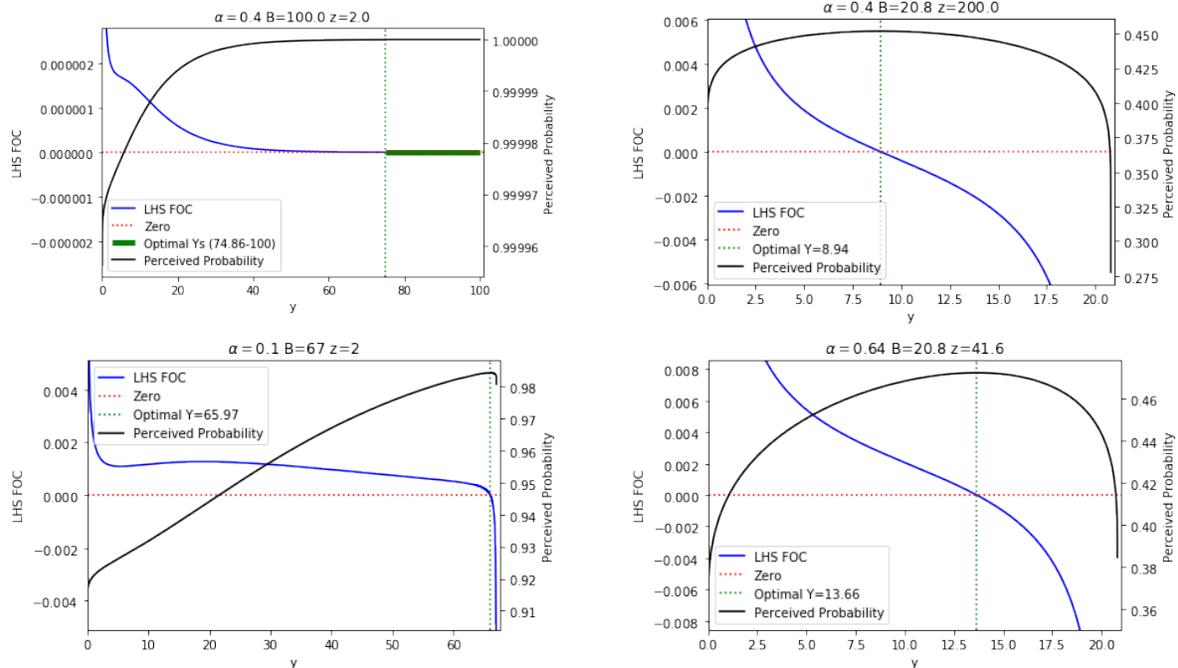


Figure 10: Selected Numerical Solutions for Common Edge and  $p(x_i) = 1 - e^{-\frac{x_i}{z}}$

Several general patterns are evident from the graphs in Figure 10. Firstly, the first order condition is always positive as  $y$  tends to zero. Secondly, the first order condition is either negative or undefined as  $y$  tends to  $B$ . Finally, the first order condition is often always decreasing in  $y$ , but not always. For example, the graph in the bottom left corner shows a slight increase in the first order condition for the range about  $y \in [5, 20]$ . Because the first order condition is not always decreasing the optimal solution is not always uniquely defined by the first order condition since it is possible for it to cross zero more than once. However, this is uncommon for common edge networks. To investigate this we conducted a grid search of 10 equally spaced points over the parameters  $\alpha \in [.2, .99]$ ,  $B \in [1, 1000]$  and  $z \in [2, 200]$ , while imposing the condition  $B < 37.42z$  as we have already established that if the budget is too large the solution is non-unique. This results in 7740 combinations of parameters. We confirm our first and second identified patterns, that for all of the combinations of parameters we consider the first order condition is positive when  $y$  tends to 0 and negative when  $y$  tends to  $B$ . We also find that in the vast majority (7729/7740) of our parameters the first order condition is always decreasing, and so we can be generally comfortable in assuming the solution is unique for most combinations of parameter sets. As for the 11 combinations of parameters where the first order condition is not always decreasing, they are characterized by low levels of  $\alpha$  (typically  $\alpha < .25$ ). For these combinations of parameters, we check their uniqueness in two ways. Firstly, we see whether the first order condition ever becomes positive again after first crossing zero. Nine of the eleven combinations pass this test, while the remaining two are a precision error, confirmed graphically as well as by requiring the subsequent observed positive first order condition to be over a small threshold. Secondly, we maximize the overall perceived probability of defense over a fine grid of  $y$ , and see if this set has more than one point in it. All eleven combinations pass this test. Based on this analysis, for the range of parameters we consider, we can be reasonably confident that the solutions are unique in the common edge network with the exponential edge defense function.

### 7.1.2 Common Edge Network - Alternative Edge Defense Function

We now consider the edge defense function  $p(x_i) = (\frac{x_i}{z})^b$ . The objective function is as follows:

$$e^{-(-b \log(y) + b \log(z))^a} + 2e^{-(-\log((\frac{B}{4z} - \frac{y}{4z}))^b))^a} - 2e^{-(-\log((\frac{B}{4z} - \frac{y}{4z}))^b))^a} e^{-(-b \log(y) + b \log(z))^a} - e^{-2(-\log((\frac{B}{4z} - \frac{y}{4z}))^b))^a} + e^{-2(-\log((\frac{B}{4z} - \frac{y}{4z}))^b))^a} e^{-(-b \log(y) + b \log(z))^a} \quad (5)$$

With the associated first order condition:

$$\begin{aligned} & \frac{ab \left( -\log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right) \right)^a e^{-(-\log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right))^a}}{4 \left( \frac{B}{4} - \frac{y}{4} \right) \log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right)} \left( \left( 1 \right. \right. \\ & \left. \left. - e^{-(-\log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right))^a} \right) e^{-(-\log \left( \left( \frac{y}{z} \right)^b \right))^a} + e^{-(-\log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right))^a} \right) \right) \\ & + \frac{ab \left( -\log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right) \right)^a e^{-(-\log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right))^a}}{4 \left( \frac{B}{4} - \frac{y}{4} \right) \log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right)} \\ & + \left( 1 - e^{-(-\log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right))^a} \right) \left( \frac{ab \left( -\log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right) \right)^a e^{-(-\log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right))^a}}{4 \left( \frac{B}{4} - \frac{y}{4} \right) \log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right)} \right. \\ & \left. - \frac{ab \left( -\log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right) \right)^a e^{-(-\log \left( \left( \frac{y}{z} \right)^b \right))^a}}{4 \left( \frac{B}{4} - \frac{y}{4} \right) \log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right)} e^{-(-\log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right))^a} \right. \\ & \left. - \frac{ab \left( -\log \left( \left( \frac{y}{z} \right)^b \right) \right)^a}{y \log \left( \left( \frac{y}{z} \right)^b \right)} \left( 1 - e^{-(-\log \left( \left( \frac{1}{z} \left( \frac{B}{4} - \frac{y}{4} \right) \right)^b \right))^a} \right) e^{-(-\log \left( \left( \frac{y}{z} \right)^b \right))^a} \right) = 0 \end{aligned} \quad (6)$$

This first order condition is also analytically intractable, but very numerically tractable. We conduct the same exercise as for the exponential edge defense function, except we have an additional parameter  $b$  which we will keep between 0 and 1 as this is the particularly interesting case for this style of function (and  $b = 0.4$  was actually used in the experiment). We also restrict  $B < Z$ , as the edge defense function is not a proper probability function otherwise ( $B > Z$  would imply a probability greater than one). Figure 11 presents a graphical analysis for a selection of parameter combinations.

The patterns identified with the exponential edge defense function also hold here. The main difference is that in this environment it is possible for the first order condition to be very non-decreasing instead of just slightly, as shown in the top left graph. We consider a grid search of combinations of 20 equally spaced points of the parameters  $z \in [2, 200]$ ,  $\alpha \in [.4, .99]$ ,  $B \in [1, 200]$  and  $b \in [.1, .9]$ , which yields 83600 combinations. We confirm that the first order condition is always positive as  $y$  tends to 0, but we find 27 instances of cases where the first order condition is not negative as  $y$  tends to  $B$ . These are characterized by high values of  $\alpha$ ,  $B$ s,  $z$ s, and  $b$ s, which are leading to an optimum value too close to  $B$  to observe the first order condition go negative before becoming undefined. For all 27 of these combinations, a numerical analysis of the maximum perceived probability confirms a unique optimum, as we cannot confirm uniqueness from the first order condition crossing zero. As expected, substantially more combinations (23456) exhibited a non-decreasing first order condition than in the exponential defense function case. For all of these combinations we confirm that the solution is unique in the same way as the exponential defense function, by observing whether the first order condition crosses the zero line more than once and by numerical maximization of perceived probability. Therefore, for the range of values we consider (which span the ones considered in the experiment), we can be confident that the solution is unique.

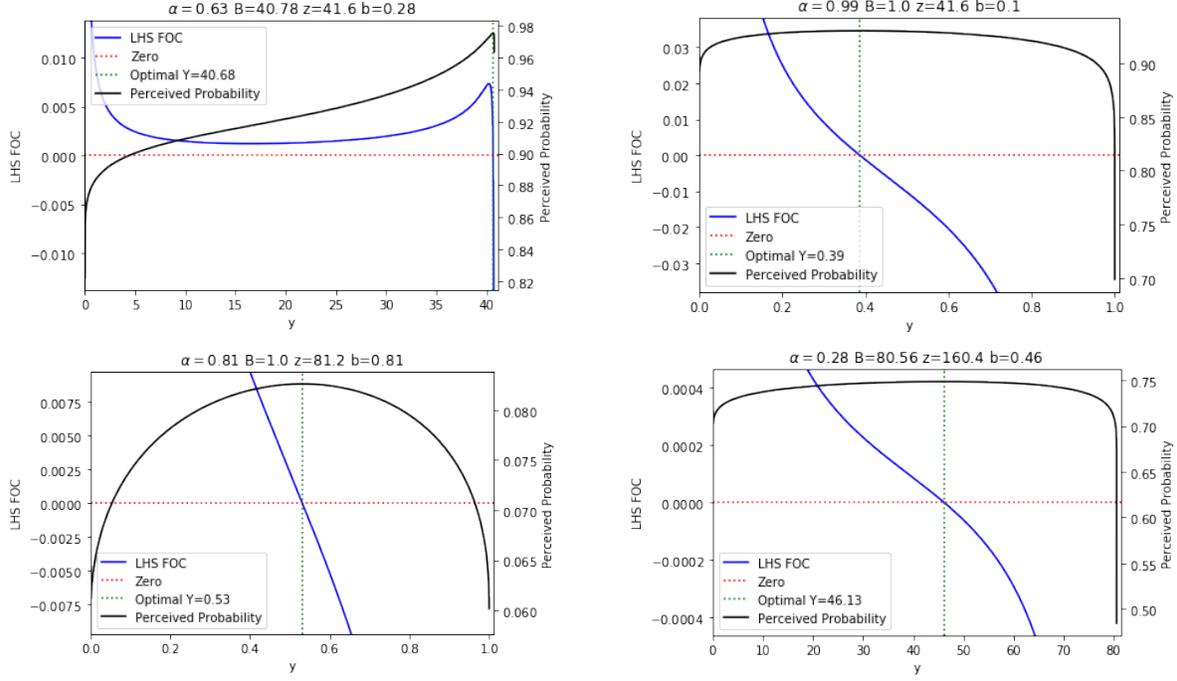


Figure 11: Selected Numerical Solutions for Common Edge and  $p(x_i) = \left(\frac{x_i}{z}\right)^b$

### 7.1.3 Network Green

We denote  $x$  as the number of units allocated to each edge in the path with 5 edges, and  $y$  as the number of units allocated to each edge in the path with 2 edges. The objective function contains a minimum operator. In order to express the objective function in a manner we can differentiate we can substitute in the constraints. The budget constraint is trivial to substitute, however there is also another implied constraint that each path should have the same perceived probability of defense. Substituting in this implied constraint would remove the need for the minimum operator, however this proves analytically intractable, as the following equation cannot be solved for  $x$  (or  $y$  for the other possible substitutable route) to obtain a substitutable closed form solution.

$$\begin{aligned}
& 5e^{-\left(\frac{x}{z}-\log\left(e^{\frac{x}{z}}-1\right)\right)^a} - 10e^{-2\left(\frac{x}{z}-\log\left(e^{\frac{x}{z}}-1\right)\right)^a} + 10e^{-3\left(\frac{x}{z}-\log\left(e^{\frac{x}{z}}-1\right)\right)^a} - 5e^{-4\left(\frac{x}{z}-\log\left(e^{\frac{x}{z}}-1\right)\right)^a} \\
& + e^{-5\left(\frac{x}{z}-\log\left(e^{\frac{x}{z}}-1\right)\right)^a} - 2e^{-\left(-\log\left(-e^{\frac{1}{2z}(-B+5x)}+1\right)\right)^a} + e^{-2\left(-\log\left(-e^{\frac{1}{2z}(-B+5x)}+1\right)\right)^a} = 0
\end{aligned} \tag{7}$$

An alternative route would be to form the associated Lagrangian. The two first order conditions (with respect to  $x$  and  $\mathcal{L}$ ) can be obtained, however the same situation occurs where it is not analytically possible to substitute one of these first order conditions into the other. We can, however, conduct a similar exercise as before, except this time considering both first order conditions simultaneously. That is, we can easily calculate the value for both first order conditions for any combination of  $y$ ,  $\mathcal{L}$ ,  $\alpha$ ,  $B$ , and  $z$ , and find the combinations of  $y$  and  $\mathcal{L}$  where both first order conditions are zero. For both ease of visual display as well as for a computer based minimizer we fold both first order conditions in one expression that should be minimized:  $F = FOC_y^2 + FOC_{\mathcal{L}}^2$ . This expression should equal zero when both first order conditions are zero, and be greater than zero otherwise. We present selected combinations of parameters in Figures 12, 13, and 14, which represent the three types of patterns that generally emerge. The top graph in each figure is a heatmap<sup>25</sup> of the expression representing both first order conditions, while the four following graphs are each first order condition holding one parameter fixed at the optimum value. The latter is to support the heatmap analysis, for example, the heatmap in Figure 13 looks rather flat along the  $\mathcal{L}$  dimension, but the

<sup>25</sup>Note, the heatmap is truncated for high values of  $F$  so we can more easily focus on values where it is close to zero.

first order conditions graphs confirm that there is some movement along that dimension, and that both of the first order conditions are zero at the optimal point. It should also be noted that the perceived probability is overlaid on the heatmap, and that it is single peaked at the point where the top and bottom paths are equal in terms of perceived probability. To confirm the uniqueness of the solution for a wide range of parameters, we check whether the perceived probability is in fact single peaked at the maximum (rather than having a non-unique flat maximum). We conduct an equally spaced grid search of size 40 along the parameter space of  $\alpha \in [0.4, .99]$ ,  $B \in [1, 1000]$ , and  $z \in [2, 200]$ , with the restriction  $2B < 37.42Z$ . This yields 55400 combinations, all of which are single peaked at a maximum in terms of perceived probability. We conclude that for the range of parameters we consider, we can be confident that the solution is unique.

#### 7.1.4 $\alpha = 1$ with Exponential Edge Defense Function

Consider a simple two path network with two edges along each edge, and no common edges (e.g. Figure 7). With an exponential defense function, and denoting the first edge as  $x$  and the second edge as  $y$ , the probability of a successful defense along a path is  $1 - e^{-\frac{x}{z}} + e^{-\frac{x}{z}}(1 - e^{-\frac{y}{z}}) = 1 - e^{-\frac{x+y}{z}}$ . Note that this is invariant to any  $x$  and  $y$  allocations given a fixed total  $T$ , where  $x + y = T$ . Therefore, any allocation that assigns a total of  $\frac{B}{2}$  to each path is optimal for  $\alpha = 1$  and  $p(x_i) = 1 - e^{-\frac{x_i}{z}}$ , and thus there is not a unique solution for this case. By the same logic, this can be shown for any number of edges along a path, the only requirement is that the total allocation along a path is split evenly. This is not the case for a network with one common edge, as with an exponential defense function an  $\alpha = 1$  type would allocate all their units to the common edge and thus the solution would be unique. However, if there were two common edges, then any allocation that allocates all of their units across those two non-common edges would be optimal for  $\alpha = 1$  due to the exponential defense function by the same logic as the case with the paths, and thus there would not be a unique solution.

#### 7.1.5 $\alpha = 0$

In the Prelec probability weighting function,  $w(p(x_i); \alpha) = \exp[-(-\log(p(x_i)))^\alpha]$ ,  $\alpha = 0$  is a limiting case where every probability is perceived as  $\frac{1}{e} = .368$ . In our model, this means that the defender perceives every probability of a successful defense along an edge as  $\frac{1}{e}$ . The  $\alpha = 0$  case is not very realistic, but solutions to any network structure and edge defense functions can be easily described. If one assumes that  $w(0) = w(1) = \frac{1}{e}$ , then literally any allocation is an optimal solution, so therefore there are no unique solutions. This is true for any network that could be specified. If instead one assumes that  $w(0) = 0$  and  $w(\epsilon) = \frac{1}{e}$ , then any combination that allocates  $x_i \geq \epsilon \forall i$  is an optimal solution, with additional possible solutions where  $x_i = 0$  if  $i$  is an extraneous edge. For either assumption about the nature of  $\alpha = 0$  there is no unique solution for any type of network.

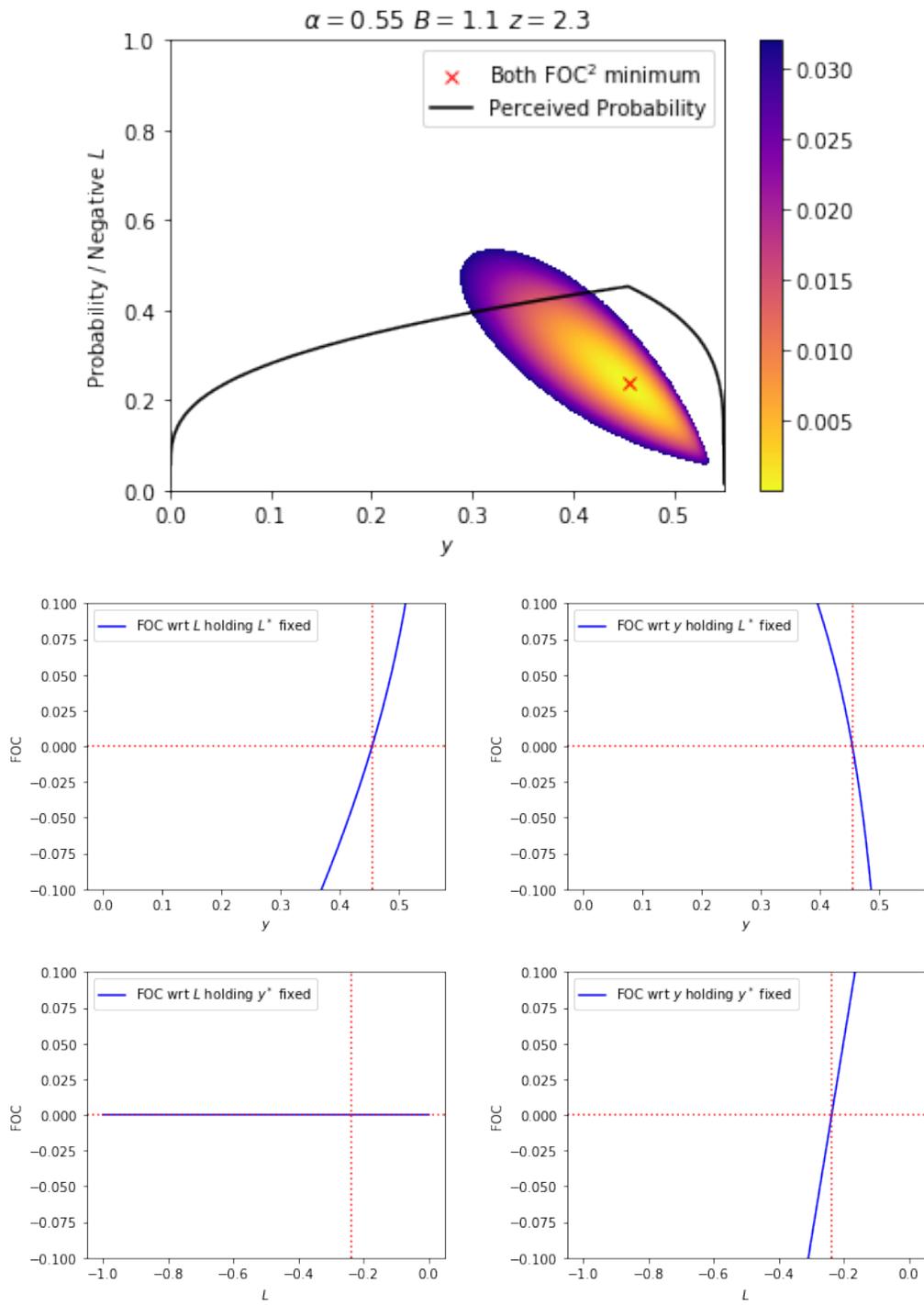


Figure 12: Network Green -  $\alpha = .55$ ,  $B = 1.1$ ,  $z = 2.3$

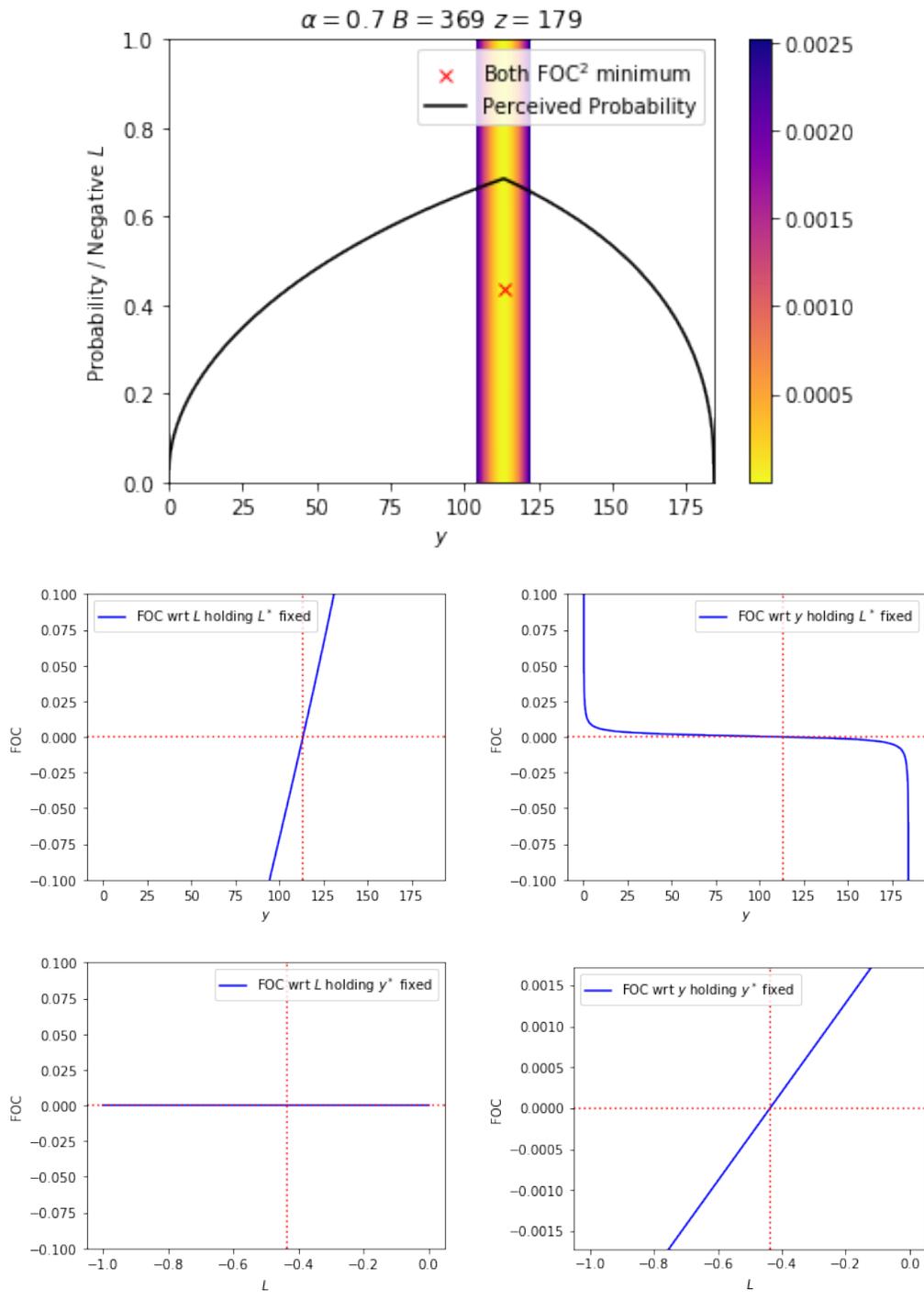


Figure 13: Network Green -  $\alpha = .7$ ,  $B = 369$ ,  $z = 179$

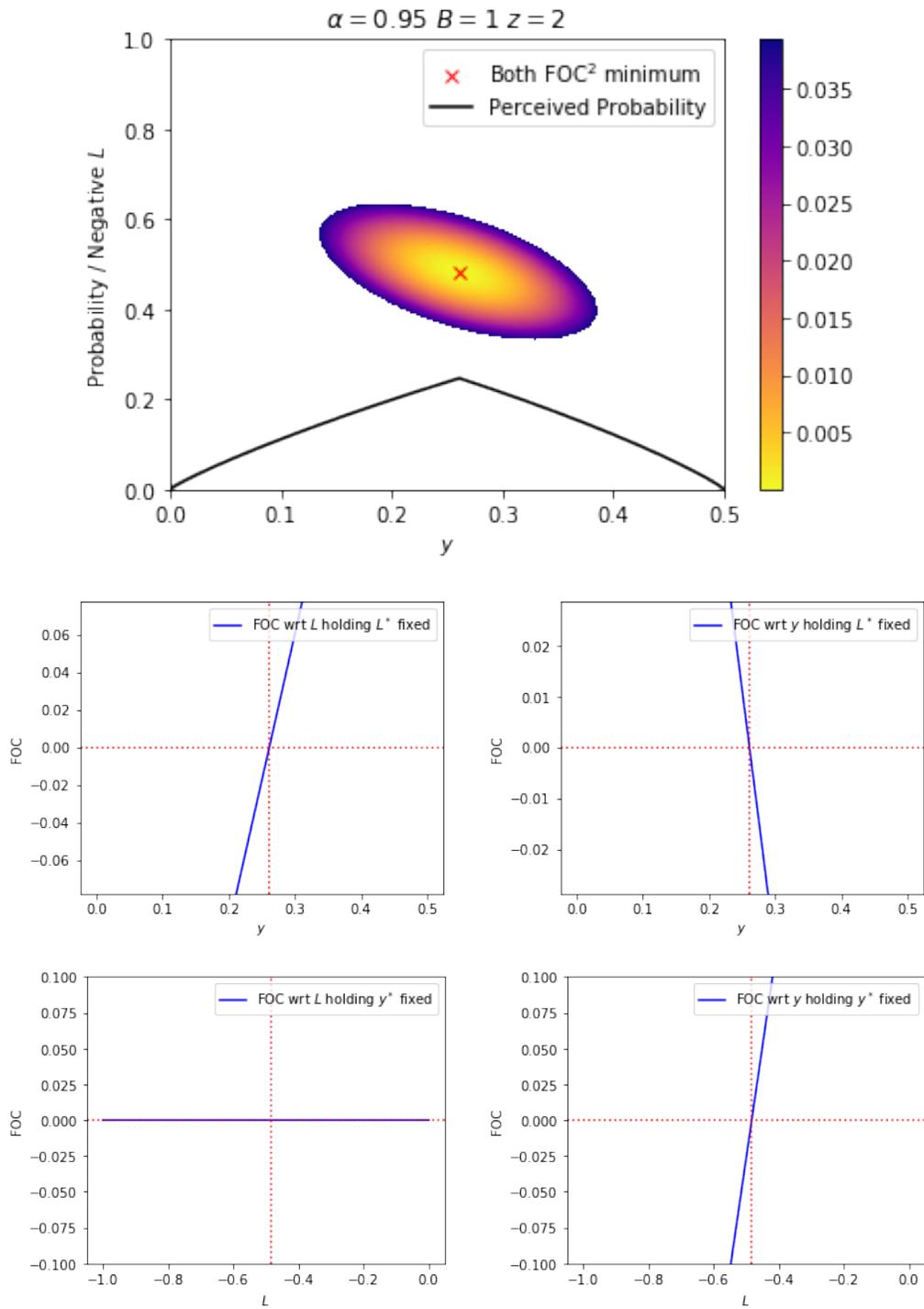


Figure 14: Network Green -  $\alpha = .95, B = 1, z = 2$

## 7.2 DOSE Procedures

### 7.2.1 Network Attack Task

We build a question bank by forming two lists of probabilities. List one consists of: [0.03, 0.08, 0.12, 0.16, 0.21, 0.25, 0.29, 0.34, 0.38, 0.42, 0.47, 0.51, 0.56, 0.6, 0.64, 0.69, 0.73, 0.77, 0.82, 0.86, 0.9, 0.95, 1] and list two is: [0.06, 0.11, 0.15, 0.19, 0.24, 0.28, 0.32, 0.37, 0.41, 0.45, 0.5, 0.54, 0.59, 0.63, 0.67, 0.72, 0.76, 0.8, 0.85, 0.89, 0.93, 0.98]. These numbers were chosen in an attempt to increase the complexity of multiplying probabilities along a path. Paths were created for each possible combination of one probability from list one and one probability from list two, and then all possible combinations of 2 different paths were possible questions. This yielded a question bank of possible 255530 questions. To reduce the question bank to more relevant questions, we only chose questions where subjects with  $\alpha \in \{.4, .5, .6, .7, .8, .9, .95\}$  would make different responses from a subject with  $\alpha = 1$ . This yields a question bank with 1397 questions.

To capture noise or errors in decision making, we assume that subjects best respond according to a logit function, so the probability of choosing the top path is:  $Prob(O_T = p_1, p_2) = \frac{1}{1 + \exp(-\lambda(U(O_T) - U(O_B)))}$  if  $U(O_T) > U(O_B)$  (and one minus this probability otherwise).

DOSE performs Bayesian updating over the likelihood of a subject being a specific type. Here, a type consists of  $\alpha$  and  $\lambda$ . We consider 20 equally spaced points for  $\alpha \in [.4, 1]$ , and  $\lambda \in [0, 100]$ , and form types based on all possible combinations of these. A finer grid would be better, however due to computational limitations we use the grid space of 20.

We then begin the DOSE procedure. We assume an initial uniform prior over types, and then calculate the Kullback-Leibler (KL) divergence that results between the posterior and the prior for each response to each question in the question bank. The specific functional form we use is presented as Equation (3) in Chapman et al. (2018). The question that has the highest KL is selected to be asked (and then removed from the question bank, so that it is not asked again). In order to reduce DOSE’s preference for repeatedly asking the same or similar questions (in order to get a better estimate of  $\lambda$ , which is of secondary interest), we remove all questions that share one of the same paths as the previously asked question. Then, for each possible response to the asked question, the prior is updated. The process then continues separately for each of the new priors, so that subjects that respond differently to the questions are asked a different sequence of questions. The order of questions are recorded in a tree-like structure, so that a software environment can ask the appropriate question given the response, without needing to calculate the KL for each individual each time. Due to computational limitations, we get a dynamic ordering of questions for 15 responses. We ask 20 questions, with the first 15 being dynamically generated and the last 5 being manually chosen by the experimenter (and the same for all subjects).

### 7.2.2 Binary Lottery Task

We built a question bank from the rows of the Multiple Price Lists found in Tanaka et al. (2010), Callen et al. (2014), Bruhin et al. (2010), and Holt and Laury (2002). Because these are in various currencies (and at various points of time), we normalized the highest payoff in each paper to be 1, and scaled the other payoffs accordingly. We use only Series 1 and 2 from Tanaka et al. (2010), and censor Series 1 at the 220 row, due to the large relative scale differences of the latter rows of Series 1. We then scale all payoffs up by a factor of 1472 (experimental points), to bring the average payoffs in line with our other tasks. This process gives us a question bank of 287 questions. We then remove questions with strictly dominated options (common in MPL’s). We assume a utility function of an option as  $U(p_1, d_1, p_2, d_2) = w_p(p_1)d_1^\sigma + (1 - w_p(p_1))d_2^\sigma$  if  $d_1 > d_2$  and  $U(p_1, d_1, p_2, d_2) = w_p(p_2)d_2^\sigma + (1 - w_p(p_2))d_1^\sigma$  otherwise, and iterating through perfectly responding agents with all combinations of  $\alpha \in [.4, 1]$  and  $\sigma \in [.2, 1.7]$  in a linespace of 12. If there are options in which none of the combinations of agents chose a particular option, then this question is deleted, leaving us with a bank of 163 questions.

To capture noise or errors in decision making, we assume that subjects best respond according to a logit function, so the probability of choosing Option A is:  $Prob(O_A = p_1, d_1, p_2, d_2) = \frac{1}{1 + \exp(-\lambda(U(O_A) - U(O_B)))}$  if  $U(O_A) > U(O_B)$  (and one minus this probability otherwise). We correct for the rescaling of payoffs relative to  $\lambda$  that risk aversion causes by re-normalizing the payoffs in the same manner as Goeree et al. (2003).

We then calibrate the upper bound for  $\lambda$ , where higher  $\lambda$  means subjects are getting closer to perfectly best responding. We start with a low  $\lambda$ , go through the logit responses of the aforementioned  $\alpha$  and  $\sigma$

types, and for each question in the bank, record the probability of choosing an option. If fewer than 80% of responses are either below 10% or above 90% (i.e. close to best responding), then  $\lambda$  is increased slightly. This process continues until the 80% threshold is reached. The ending  $\lambda_m$  is the upper bound for the DOSE procedure.

DOSE performs Bayesian updating over the likelihood of a subject being a specific type. Here, a type has its own  $\alpha$ ,  $\sigma$  and  $\lambda$ . We consider 20 equally spaced points for  $\alpha \in [.4, 1]$ ,  $\sigma \in [.2, 1.7]$ , and  $\lambda \in [.01, \lambda_m]$ , and form types based on all possible combinations of these. A finer grid would be better, however due to computational limitations we use the grid space of 20.

We then begin the DOSE procedure. We assume an initial uniform prior over types, and then calculate the Kullback-Leibler (KL) divergence that results between the posterior and the prior for each response to each question in the question bank. The specific functional form we use is presented as Equation (3) in Chapman et al. (2018). The question that has the highest KL is selected to be asked (and then removed from the question bank, so that it is not asked again). Then, for each possible response to that question, the prior is updated. The process then continues separately for each of the new priors, so that subjects that respond differently to the questions are asked a different sequence of questions. The order of questions are recorded in a tree-like structure, so that a software environment can ask the appropriate question given the response, without needing to calculate the KL for each individual each time. Due to computational limitations, we get a dynamic ordering of questions for 16 responses. We ask 20 questions, with the first 16 being dynamically generated and the last 4 being manually chosen by the experimenter (and the same for all subjects).

### 7.3 Hypothesis 2 Robustness Tests

	.8	.81	.82	.83	.84	.85	.86	.87	.88	.89	.9	.91	.92	.93	.94	.95	.96	.97
Red versus Yellow if $\alpha \geq$	.029	.03	.029	.036	.021	.055	.042	.026	.018	.021	.031	.021	.013	.013	.034	.017	.143	.222
Red versus Yellow if $\alpha <$	.324	.359	.362	.434	.337	.435	.445	.488	.45	.468	.451	.493	.455	.472	.38	.416	.167	.135
Orange versus Yellow if $\alpha \geq$	.022	.027	.023	.03	.051	.055	.033	.035	.037	.029	.017	.007	.006	.009	.011	.012	.096	.176
Orange versus Yellow if $\alpha <$	.479	.425	.478	.425	.302	.26	.361	.331	.306	.373	.447	.468	.421	.438	.349	.311	.112	.098

Table 5: One-sided P-values from Wilcoxon Signed Rank Tests of Hypothesis 2 for Different  $\alpha$  Thresholds

### 7.4 Binary Lottery Task Results

We assume a specific functional form for a lottery  $L = (x_1, p; x_2)$  where  $x_1 > x_2$  of:  $U(x_1, x_2) = w(p)x_1^\sigma + (1 - w(p))x_2^\sigma$ . Figures 15 and 16 present the CDFs of the estimated  $\sigma$  and  $\alpha$  from this measure.

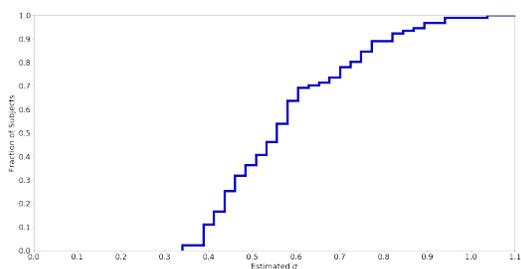


Figure 15: CDF of elicited  $\sigma$

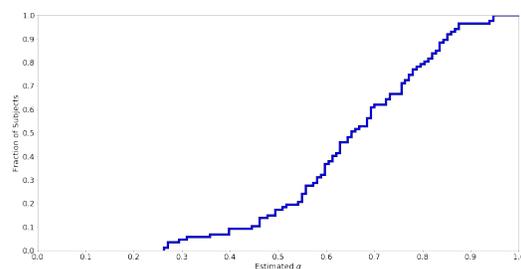


Figure 16: CDF of elicited  $\alpha$

Table 6 presents the Spearman's  $\rho$  for all combinations of the elicited parameters and average subject behavior. As discussed, there is a modest, marginally insignificant correlation between the two measures of  $\alpha$ .

	$\alpha$ (AT)	$\lambda$ (AT)	$\alpha$ (LT)	$\sigma$ (LT)	$\lambda$ (LT)	Red	Orange	Yellow	Blue
$\lambda$ (AT)	<b><math>\rho = .765</math></b> <b><math>p &lt; .001</math></b>	$\rho = 1$							
$\alpha$ (LT)	$\rho = .166$ $p = .117$	$\rho = .175$ $p = .098$	$\rho = 1$						
$\sigma$ (LT)	$\rho = .088$ $p = .406$	$\rho = .125$ $p = .238$	<b><math>\rho = .381</math></b> <b><math>p &lt; .001</math></b>	$\rho = 1$					
$\lambda$ (LT)	$\rho = .108$ $p = .308$	$\rho = .015$ $p = .886$	$\rho = -.105$ $p = .321$	<b><math>\rho = -.397</math></b> <b><math>p &lt; .001</math></b>	$\rho = 1$				
Red	<b><math>\rho = .289</math></b> <b><math>p = .005</math></b>	$\rho = .173$ $p = .101$	$\rho = -.014$ $p = .894$	$\rho = -.109$ $p = .306$	$\rho = .108$ $p = .308$	$\rho = 1$			
Orange	<b><math>\rho = .253</math></b> <b><math>p = .016</math></b>	$\rho = .078$ $p = .463$	$\rho = .056$ $p = .597$	$\rho = -.195$ $p = .064$	$\rho = .148$ $p = .161$	<b><math>\rho = .676</math></b> <b><math>p &lt; .001</math></b>	$\rho = 1$		
Yellow	$\rho = .079$ $p = .455$	$\rho = -.041$ $p = .696$	$\rho = .007$ $p = .946$	$\rho = -.200$ $p = .058$	$\rho = .073$ $p = .491$	<b><math>\rho = .601</math></b> <b><math>p &lt; .001</math></b>	<b><math>\rho = .617</math></b> <b><math>p &lt; .001</math></b>	$\rho = 1$	
Blue	<b><math>\rho = -.289</math></b> <b><math>p = .005</math></b>	<b><math>\rho = -.218</math></b> <b><math>p = .038</math></b>	$\rho = .106$ $p = .317$	$\rho = .123$ $p = .246$	$\rho = -.200$ $p = .058$	<b><math>\rho = -.380</math></b> <b><math>p &lt; .001</math></b>	<b><math>\rho = -.339</math></b> <b><math>p = .001</math></b>	$\rho = -.115$ $p = .276$	$\rho = 1$
Green	<b><math>\rho = -.271</math></b> <b><math>p = .009</math></b>	$\rho = -.146$ $p = .168$	$\rho = -.016$ $p = .876$	$\rho = .083$ $p = .437$	$\rho = -.010$ $p = .927$	<b><math>\rho = -.280</math></b> <b><math>p = .007</math></b>	<b><math>\rho = -.322</math></b> <b><math>p = .002</math></b>	<b><math>\rho = -.256</math></b> <b><math>p = .014</math></b>	<b><math>\rho = .290</math></b> <b><math>p = .005</math></b>

Table 6: Full Spearman's  $\rho$  Table (in bold if  $p < .05$ , in italics if  $.05 < p < .1$ )

(AT=Network Attack Task, LT= Binary Lottery Task)

## 7.5 Individual Network Cluster Analysis

### 7.5.1 Network Red

Cluster Name % of subjects in Cluster	Average Edge Allocation
Naive Diversification 14.9%	
Near Optimal $\alpha = 1$ 51.7%	
Early Revelation - Some Diversification 13.8%	
$\alpha < 1$ - Some Diversification 19.5%	

Table 7: Network Red Cluster Analysis

### 7.5.2 Network Orange

Cluster Name % of subjects in Cluster	Average Edge Allocation
$\alpha < 1$ - Some Diversification 23.0%	
Near Optimal $\alpha = 1$ 31.0%	
Naive Diversification 16.1%	
$\alpha < 1$ or Mild Diversification 23.0%	
Early Revelation - Mild Diversification 6.9%	

Table 8: Network Orange Cluster Analysis

### 7.5.3 Network Yellow

Cluster Name % of subjects in Cluster	Average Edge Allocation
$\alpha < 1$ - Some Diversification and Early Revelation 11.5%	
Early Revelation - Mild Diversification 8.0%	
Near Optimal $\alpha = 1$ - Some Late Revelation 19.5%	
Naive Diversification 24.1%	
$\alpha < 1$ - Some Diversification 16.1%	
Late Revelation 20.7%	

Table 9: Network Yellow Cluster Analysis

### 7.5.4 Network Blue

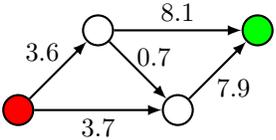
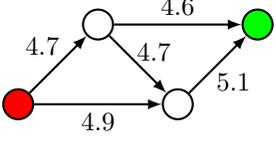
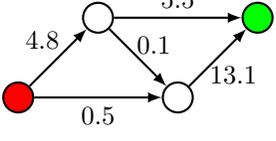
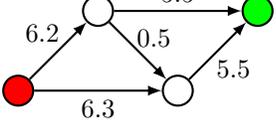
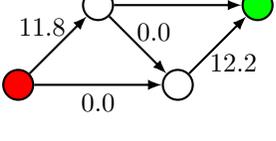
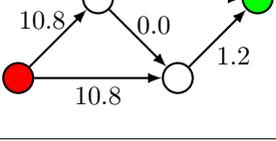
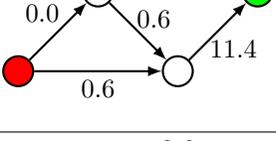
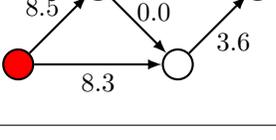
Cluster Name % of subjects in Cluster	Average Edge Allocation
Late Revelation - Mild Diversification 24.1%	
Naive Diversification 12.6%	
'False Common Edge' Heuristic 3.4%	
Mild Diversification and Mild Early Revelation 18.4%	
Near Optimal 2.3%	
Near Optimal - Early Revelation 9.2%	
Late Revelation - Very Mild Diversification 20.7%	
Near Optimal - Some Early Revelation 9.2%	

Table 10: Network Blue Cluster Analysis

### 7.5.5 Network Green

Cluster Name % of subjects in Cluster	Average Edge Allocation
Some Late Revelation - Very Mild Diversification 9.2%	
Some Diversification - Mild Late Revelation 12.6%	
$\alpha \approx .975$ or $\alpha < .975$ with Diversification 23.0%	
Sometimes Early Revelation 2.3%	
Naive Diversification 11.5%	
Near Optimal - Early Revelation 8.0%	
Near Optimal - Some Early Revelation 6.9%	
Late Revelation - Very Mild Diversification 11.5%	
Near Optimal - Some Early Revelation 14.9%	

Table 11: Network Green Cluster Analysis

## 7.6 Experiment Interface

### Decision - Task **Blue**

Start

End

Edge Urn: 100A, 0D  
Probability of D: 0%

Edge Urn: 100A, 0D  
Probability of D: 0%

Edge Urn: 76A, 24D  
Probability of D: 24%

Edge Urn: 72A, 28D  
Probability of D: 28%

Edge Urn: 100A, 0D  
Probability of D: 0%

Total Amount Currently Allocated: 11/24

Number of Previous Successes: 0/0

Units Allocated to an edge	Balls in Edge Urn (% chance D)
0	100 Attack balls, 0 Defend Balls (0%)
1	95 Attack balls, 5 Defend Balls (5%)
2	90 Attack balls, 10 Defend Balls (10%)
3	85 Attack balls, 15 Defend Balls (15%)
4	80 Attack balls, 20 Defend Balls (20%)

## 7.7 Experiment Instructions

### 7.7.1 Overview

#### Introduction

This experiment is a study of decision making. The amount of money you earn depends partly on the decisions that you make and thus you should read the instructions carefully. The money you earn will be paid privately to you, in cash, at the end of the experiment. A research foundation has provided the funds for this study. Please put away your cell phones and other distracting devices for the duration of the experiment.

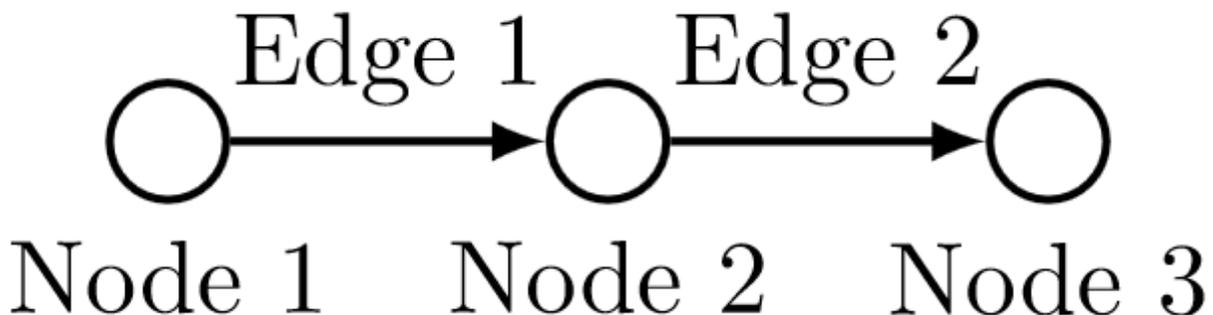
In this experiment, you will participate in Task 1, and then 6 additional colored tasks. The Task 1 instructions are given on a separate piece of paper. For the additional colored tasks, you have been given instructions printed on different colored paper. The color of the paper coincides to the name of the task. Only read the relevant instructions when the computer prompts you to do so. The tasks are independent meaning the decisions and payoffs from one do not affect the decisions and payoffs from the other. **Some of these tasks are similar, but you should take care when reading similar instructions to see what is different about the new task.**

Please do not attempt to communicate with other participants in the room during the experiment. If you have a question as you read through the instructions at any time during the experiment, please raise your hand and an experimenter will come by to answer it in private.

You cannot use a pen or a calculator until after you have completed Task White (which is the second task). If you want to use either a pen or a calculator after Task White, please raise your hand and an experimenter will bring you one.

Your earnings in this task are denominated in experimental dollars (called points in the software), which will be exchanged at a rate of 350 experimental dollars = 1 U.S. dollar at the end of the experiment.

The colored instructions use some terms which you may not be already familiar with, Nodes and Edges. The following figure is designed to illustrate these concepts:



A Node is a position, while Edges describe how you can go between these positions. A single Edge connects two Nodes. The arrow indicates the direction of the Edge, you can only go between Nodes in the direction of the Edge. In the given figure, you can go from Node 1 to Node 2 using Edge 1, and from Node 2 to Node 3 using Edge 2. Note, you cannot go from Node 2 to Node 1, or from Node 3 to Node 2, or from Node 3 to Node 1, as there are no Edges that connect those Nodes in that direction.

### 7.7.2 Binary Lottery Task

#### Task 1

Task 1 is divided into 20 decision “periods.” You will be paid for this task based on your decision in one of the periods, which will be randomly selected. Each decision you make is therefore important because it has a chance of determining the amount of money you earn.

In each period, you will be asked to choose between 2 Options, Option A and Option B.

Each Option has 20 balls in an Urn. These balls are colored Red or Blue. One ball will be drawn from the Urn of the Option you choose. Each Option has a payoff in points if a Red ball is drawn, and a payoff

in points if a Blue ball is drawn. You choose an Option by clicking on it, and then clicking on the Submit button.

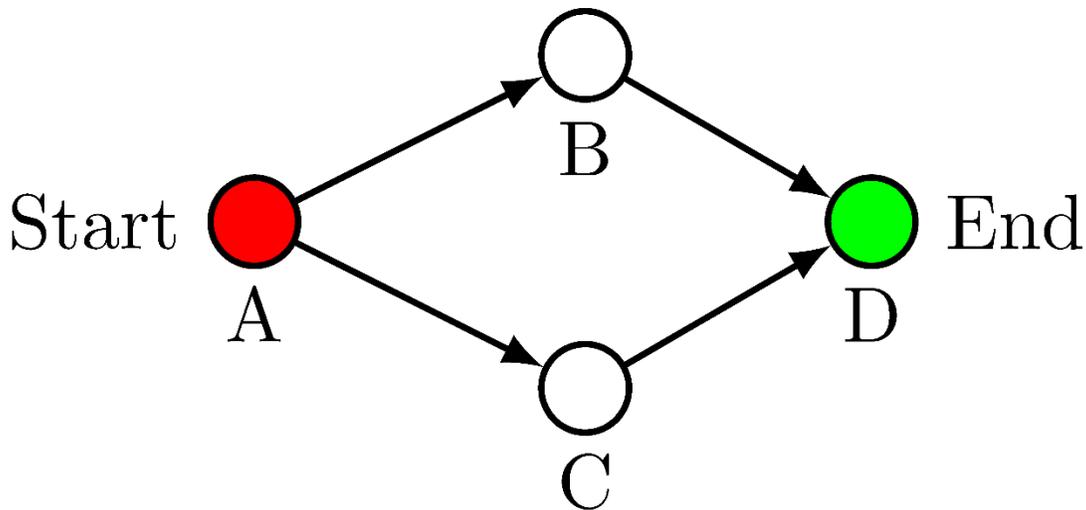
The ball for your chosen Option is not drawn until the end of the 20 decision periods. At the end of Task 1 you will be shown the randomly selected period, the Options that were available in that period, the Option you chose, the ball that was drawn from the chosen Option, and your final payoff for Task 1.

### 7.7.3 Network Attacker Task

#### Task White

Task White is divided into 20 decision “periods.” You will be paid for this task based on your decision in one of the periods, which will be randomly chosen. Each decision you make is therefore important because it has a chance of determining the amount of money you earn.

There are two roles in this task, Attacker and Defender. **You will be playing in the role of the Attacker against a computerized Defender. As an Attacker, your objective is to capture the node labelled End in the figure below.**



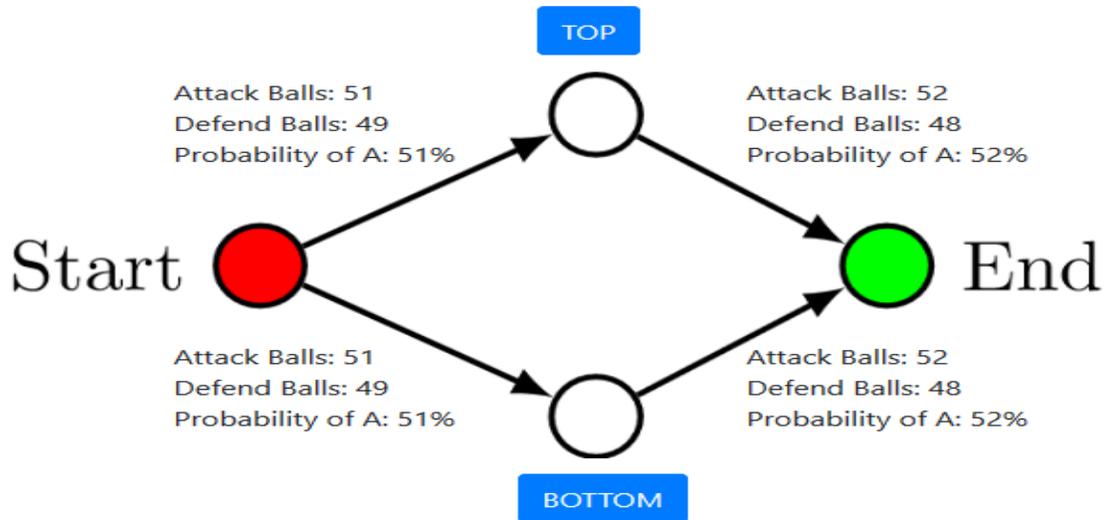
You start at Node A, labelled Start, and must decide whether to attack along the ‘Top’ path (from Node A to Node B, then Node B to Node D), or the ‘Bottom’ path (from Node A to Node C, then Node C to Node D). You can only attack along one of these paths in a period.

The probability that an attack on a Node along an Edge is successful is given by an ‘Urn’ attached to that Edge. This Urn contains Attack balls and Defend balls, and has 100 balls in total. When your chosen path takes you along an Edge, a ball is randomly drawn from that Edge’s Urn. If this ball is an Attack ball, the attack succeeds, and you capture the subsequent node. If this ball is a Defend ball, the attack fails, and your attack for this period is over. The ball contents of the Urns on each Edge are set by the computer, and differ from period to period.

#### **Example:**

In this example the Top and Bottom paths are the same. If the Top path is selected, then a ball would be drawn from the first Edge Urn with 51 Attack and 49 Defend balls. If a Defend ball is drawn (this occurs with 49% probability), then the attack on the Top Node will fail and the attack for the period is over. If an Attack ball is drawn (occurs with 51% probability), then the attack on the Top Node will succeed and the attack for the period continues. Then, a ball would be drawn from the second Edge Urn on the top path with 52 Attack balls and 48 Defend balls. If a Defend ball is drawn (occurs with 48% probability), then the attack on the End Node will fail and the attack from the period is over. If an Attack ball is drawn (occurs with 52% probability), then the attack on the End Node will succeed, and the overall attack for this period will be successful.

#### **Earnings**



If you succeed and reach the End node, you will receive 3000 experimental dollars for that period. If you fail and do not reach the End node, you will receive 0 experimental dollars for that period. You do not receive additional payment for capturing nodes other than the End node. At the end of the experiment, one period will be randomly selected for your payment from this task.

**Summary**

- You are an Attacker playing against a computerized Defender.
- Your goal is to capture the End node.
- You decide whether to attack along the top path or the bottom path.
- The probability of a successful attack on a Node is determined by the Edge Urn.
  - If an Attack ball is drawn, the attack on that Node succeeds, if a Defend ball is drawn, the attack for the period fails
- If you capture the End Node you will earn 3000 for that period.
- If you do not capture the End Node, you will earn 0 for that period
  - That is, Nodes other than the End Node are all worth 0

**7.7.4 Network Defense Tasks**

*Note: In the interests of space, the square parentheses [LIKE THIS] below indicate which parts of the instructions are common to all tasks, and which parts are unique to certain tasks. Subjects received an instructions packet with separate instructions (with both the repeated and unique parts) printed on colored paper associated with each task.*

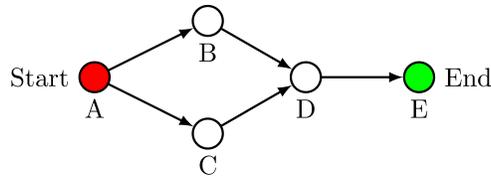
[ALL TASKS:]

**Task [COLOR]**

Task [COLOR] is divided into 10 decision “periods.” You will be paid for this task based on your decision in one of the periods, which will be randomly chosen. Each decision you make is therefore important because it has a chance of determining the amount of money you earn.

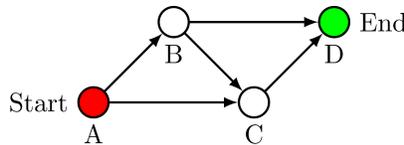
There are two roles in this task, Attacker and Defender. **You will be playing in the role of the Defender against a computerized Attacker. As a Defender, your objective is to prevent the Attacker from capturing the node labelled End in the figure below.**

[TASKS RED, ORANGE, and YELLOW:]



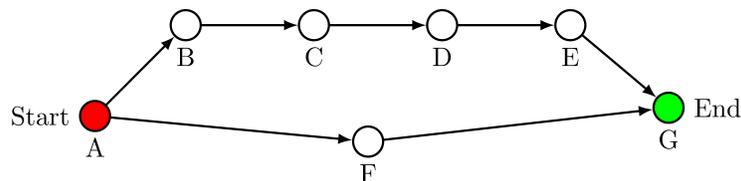
The Attacker starts at Node A, labelled Start, and can attempt to capture any other connected Node via the Edges in the direction of the arrows. For example, from Node A the attacker can attempt to capture Node B or Node C. If the Attacker captures the Node, it moves to that Node, and can then attempt to capture another Node in the direction of the arrows. For example, from Node A, if the Attacker captures Node B it will move there. Since no arrows connect Node B to Node C, and it cannot move in the opposite direction of the arrows, the attacker cannot attempt to capture Node C. If it has captured Node B, the Attacker only has one choice of Node to attack, Node D.

[TASK BLUE:]



The Attacker starts at Node A, labelled Start, and can attempt to capture any other connected Node via the Edges in the direction of the arrows. For example, from Node A the attacker can attempt to capture Node B or Node C. If the Attacker captures the Node, it moves to that Node, and can then attempt to capture another Node in the direction of the arrows. For example, from Node A, if the Attacker captures Node C it will move there. Since no arrows connect Node C to Node B, as it cannot move in the opposite direction of the arrows, the attacker cannot attempt to capture Node B. If it has captured Node C, the Attacker only has one choice of Node to attack, Node D. If the attacker captures Node B, it has two choices of Nodes to attack, Node C and Node D.

[TASK GREEN:]



The Attacker starts at Node A, labelled Start, and can attempt to capture any other connected Node via the Edges in the direction of the arrows. For example, from Node A the attacker can attempt to capture Node B or Node F. If the Attacker captures the Node, it moves to that Node, and can then attempt to capture another Node in the direction of the arrows. For example, from Node A, if the Attacker captures Node B it will move there. Since no arrows connect Node B to Node F, and it cannot move in the opposite direction of the arrows, the attacker cannot attempt to capture Node F. Node B is also not connected to any other Node except Node C. Therefore, if it has captured Node B, the Attacker only has one choice of Node to attack, Node D<sup>26</sup>.

[ALL TASKS:]

The probability that an attack on a Node along an Edge is successful is given by an ‘Urn’ attached to that Edge. This Urn contains Attack balls and Defend balls, and has 100 balls in total. When the Attacker’s path takes them along an Edge, a ball is randomly drawn from that Edge’s Urn. If this ball is an Attack ball, the attack succeeds, and the Attacker captures the subsequent node. If this ball is a Defend ball, the

<sup>26</sup>This should read ‘Node C’ instead of ‘Node D’. We present the instructions in their original form with typos included for replication purposes.

defense succeeds, and the attack for this period is over. The ball contents of the Urns on each Edge are set by you, in the role of the Defender.

You will have 24 units of defense in each period that you can allocate along each Edge to defend the Nodes. Each unit of defense allocated to an Edge increases the number of Defend balls and decreases the number of Attack balls in that Edge Urn. In other words, each unit of defense increases the probability of a successful defense if a Node is attacked along that Edge.

The table on the following page gives the probability that a Node will be successfully defended if it is attacked using an Edge with a given number of defense units.

**These Edge Urns and defense probabilities may or may not be the same between different Tasks. You should re-read this table on each subsequent set of instructions you receive.**

[TASKS RED, BLUE, GREEN:]

Number of Defense Units Allocated to an Edge	Balls in Edge Urn (% chance of successful defense)
0	100 Attack balls, 0 Defend Balls (0%)
1	95 Attack balls, 5 Defend balls (5%)
2	90 Attack balls, 10 Defend balls (10%)
3	85 Attack balls, 15 Defend balls (15%)
4	80 Attack balls, 20 Defend balls (20%)
5	76 Attack balls, 24 Defend balls (24%)
6	72 Attack balls, 28 Defend balls (28%)
7	68 Attack balls, 32 Defend balls (32%)
8	64 Attack balls, 36 Defend balls (36%)
9	61 Attack balls, 39 Defend balls (39%)
10	58 Attack balls, 42 Defend balls (42%)
11	55 Attack balls, 45 Defend balls (45%)
12	52 Attack balls, 48 Defend balls (48%)
13	49 Attack balls, 51 Defend balls (51%)
14	46 Attack balls, 54 Defend balls (54%)
15	44 Attack balls, 56 Defend balls (56%)
16	42 Attack balls, 58 Defend balls (58%)
17	39 Attack balls, 61 Defend balls (61%)
18	37 Attack balls, 63 Defend balls (63%)
19	35 Attack balls, 65 Defend balls (65%)
20	33 Attack balls, 67 Defend balls (67%)
21	32 Attack balls, 68 Defend balls (68%)
22	30 Attack balls, 70 Defend balls (70%)
23	28 Attack balls, 72 Defend balls (72%)
24	27 Attack balls, 73 Defend balls (73%)

For example, if a Node is attacked through an Edge that has 6 defense units allocated to it, the Edge Urn would have 72 Attack balls and 28 Defend balls. One ball will be drawn at random to determine the outcome, and so there is a 28 out of 100 chance (since there are 100 total balls) that the defense will succeed. If instead the Edge had 1 defense unit, then the Edge Urn would have 95 Attack balls and 5 Defend balls. In this case, the chances of a successful defense of the Node would be 5 out of 100.

[TASK ORANGE:]

Number of Defense Units Allocated to an Edge	Balls in Edge Urn (% chance of successful defense)
0	100 Attack balls, 0 Defend Balls (0%)
1	97 Attack balls, 3 Defend balls (3%)
2	94 Attack balls, 6 Defend balls (6%)
3	91 Attack balls, 9 Defend balls (9%)
4	88 Attack balls, 12 Defend balls (12%)
5	85 Attack balls, 15 Defend balls (15%)
6	82 Attack balls, 18 Defend balls (18%)
7	80 Attack balls, 20 Defend balls (20%)
8	77 Attack balls, 23 Defend balls (23%)
9	75 Attack balls, 25 Defend balls (25%)
10	73 Attack balls, 27 Defend balls (27%)
11	70 Attack balls, 30 Defend balls (30%)
12	68 Attack balls, 32 Defend balls (32%)
13	66 Attack balls, 34 Defend balls (34%)
14	64 Attack balls, 36 Defend balls (36%)
15	62 Attack balls, 38 Defend balls (38%)
16	60 Attack balls, 40 Defend balls (40%)
17	58 Attack balls, 42 Defend balls (42%)
18	56 Attack balls, 44 Defend balls (44%)
19	54 Attack balls, 46 Defend balls (46%)
20	53 Attack balls, 47 Defend balls (47%)
21	51 Attack balls, 49 Defend balls (49%)
22	49 Attack balls, 51 Defend balls (51%)
23	48 Attack balls, 52 Defend balls (52%)
24	46 Attack balls, 54 Defend balls (54%)

For example, if a Node is attacked through an Edge that has 6 defense units allocated to it, the Edge Urn would have 82 Attack balls and 18 Defend balls. One ball will be drawn at random to determine the outcome, and so there is an 18 out of 100 chance (since there are 100 total balls) that the defense will succeed. If instead the Edge had 1 defense unit, then the Edge Urn would have 97 Attack balls and 3 Defend balls. In this case, the chances of a successful defense of the Node would be 3 out of 100.

[TASK YELLOW:]

Number of Defense Units Allocated to an Edge	Balls in Edge Urn (% chance of successful defense)
0	100 Attack balls, 0 Defend Balls (0%)
1	82 Attack balls, 18 Defend balls (18%)
2	76 Attack balls, 24 Defend balls (24%)
3	72 Attack balls, 28 Defend balls (28%)
4	68 Attack balls, 32 Defend balls (32%)
5	65 Attack balls, 35 Defend balls (35%)
6	63 Attack balls, 37 Defend balls (37%)
7	60 Attack balls, 40 Defend balls (40%)
8	58 Attack balls, 42 Defend balls (42%)
9	56 Attack balls, 44 Defend balls (44%)
10	54 Attack balls, 46 Defend balls (46%)
11	52 Attack balls, 48 Defend balls (48%)
12	51 Attack balls, 49 Defend balls (49%)
13	49 Attack balls, 51 Defend balls (51%)
14	47 Attack balls, 53 Defend balls (53%)
15	46 Attack balls, 54 Defend balls (54%)
16	45 Attack balls, 55 Defend balls (55%)
17	43 Attack balls, 57 Defend balls (57%)
18	42 Attack balls, 58 Defend balls (58%)
19	41 Attack balls, 59 Defend balls (59%)
20	39 Attack balls, 61 Defend balls (61%)
21	38 Attack balls, 62 Defend balls (62%)
22	37 Attack balls, 63 Defend balls (63%)
23	36 Attack balls, 64 Defend balls (64%)
24	35 Attack balls, 65 Defend balls (65%)

For example, if a Node is attacked through an Edge that has 6 defense units allocated to it, the Edge Urn would have 63 Attack balls and 37 Defend balls. One ball will be drawn at random to determine the outcome, and so there is a 37 out of 100 chance (since there are 100 total balls) that the defense will succeed. If instead the Edge had 1 defense unit, then the Edge Urn would have 82 Attack balls and 18 Defend balls. In this case, the chances of a successful defense of the Node would be 18 out of 100.

[ALL TASKS:]

You can allocate your 24 defense units across the Edges in any pattern you wish. Defense units that are not allocated do not carry over to later periods. **You must choose a number of ‘DU’ (for Defense Units) for each Edge, even if you are allocating zero Defense Units, in order for the Next button to appear.** Once you have finished the allocation, you can finalize it by clicking on the Next button, at which point the computerized Attacker will begin.

The computerized Attacker will always attack along the path from the Start Node A to the End Node [TASKS RED, ORANGE, YELLOW]E [TASK BLUE]D [TASK GREEN]G that has the lowest probability of successful defense. If two or more paths have equally low probabilities of successful defense, then the Attacker will randomly choose one of the tied paths. Once a successful defense occurs (that is, if a Defend ball is drawn), or the Attacker captures the End node, the Attacker will stop. You will then be shown the outcome of the attack. The path the Attacker took will be represented with red arrows, and Nodes that were captured will appear red. Nodes that were not attacked at all, or a Node that was attacked but successfully defended, will appear green. If the End Node is green, then you prevented the Attacker and achieved your goal.

### Earnings

If you succeed and stop the Attacker before they reach the End node, you will receive 1500 experimental

dollars for that period. If you fail and the Attacker reaches the End node, you will receive 0 experimental dollars for that period. At the end of the experiment, one period will be randomly selected for your payment from this task.

### **Summary**

- You are a Defender playing against a computerized Attacker.
- Your goal is to stop the Attacker from capturing the End node.
- You have 24 defense units in each period to allocate across Edges.
- Defense units increase the number of Defend balls and decrease the number of Attack balls in that Edge Urn.
- The Attacker can only attack in the direction of the arrows from the Start Node.
- If a Defend ball is ever drawn, the Attacker will stop attacking, and you will earn 1500 for that period.
- If the Attacker captures the End Node, you will earn 0 for that period
- The Attacker will always choose the path to the End Node that has the lowest probability of successful defense. In the case of ties, the Attacker will randomly choose one of the tied paths.