ESSAYS ON EXPERIMENTAL ECONOMICS

by

Daniel Woods

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Economics

West Lafayette, Indiana

August 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Timothy Cason, Co-chair

Department of Economics

Dr. David Gill, Co-chair

Department of Economics

Dr. Evan Calford

Department of Economics

Dr. Collin Raymond

Department of Economics

Approved by:

Dr. Brian Roberson

ACKNOWLEDGMENTS

I gratefully acknowledge funding from the National Science Foundation (grant number CNS-1718637). I also thank the following people (in alphabetical order) for helpful comments/advice/support throughout this process: Evan Calford, Tim Cason, David Gill, Stanton Hudja, Collin Raymond, and Yaroslav Rosokha.

TABLE OF CONTENTS

LIST OF TABLES . . . 10

LIST OF FIGURES . . . 12

ABSTRACT . . . 14

1 IMPROVING EX-ANTE POWER ANALYSIS WITH QUANTAL RESPONSE SIMULATIONS . . . 15

- 1.1 Introduction . . . 15
- 1.2 Payoff Hills and Power Analysis . . . 18
  - 1.2.1 Payoff Hills . . . 18
  - 1.2.2 Effects of Payoff Hills on Power Analysis . . . 19
- 1.3 QR Simulations for Optimal Experimental Design . . . 23
  - 1.3.1 Describing the QR Simulation Power Analysis . . . 23
  - 1.3.2 Optimal Experimental Design . . . 24
  - 1.3.3 Analysis of Closed-form Expressions for Power . . . 26
    - Binary Actions . . . 27
    - Non-Binary Actions or Outcomes . . . 28
  - 1.3.4 Optimal Design and Replications . . . 29
- 1.4 Power Analysis Using Bayesian Persuasion as an Illustrative Example . . . 30
  - 1.4.1 Bayesian Persuasion Overview . . . 30
  - 1.4.2 Normal Distribution DGP . . . 31
    - Optimal Experimental Design . . . 32
  - 1.4.3 QR Framework DGP . . . 33
    - Motivation . . . 33
    - The QR Approach . . . 34
    - Optimal Experimental Design Revisited . . . 35
- 1.5 Ex-ante Power Analyses of Prominent Papers Using the QR Framework . . . 40
  - 1.5.1 Abeler et al. (2011) . . . 41

Ex-ante Power Analysis 42
Optimal Experimental Design 44
1.5.2 Charness and Dufwenberg (2011) 45
Ex-ante Power Analysis 45
Optimal Experimental Design 48
1.5.3 R. Chen and Chen (2011) 50
Ex-ante Power Analysis 51
Optimal Experimental Design 52
1.5.4 Fehr et al. (2013) 53
Ex-ante Power Analysis 56
Optimal Experimental Design 57
1.5.5 Ifcher and Zarghamee (2011) 57
Ex-ante Power Analysis 58
Optimal Experimental Design 59
1.5.6 Summary 60
1.6 Conclusion and Future Research 60

2 NETWORK DEFENSE AND BEHAVIORAL BIASES: AN EXPERIMENTAL STUDY 63
2.1 Introduction 63
2.2 Theory and Hypotheses 68
2.2.1 Attacker Model 68
2.2.2 Defender Model 69
2.2.3 Common Edges 71
Network Red 74
Network Orange 75
Network Yellow 75
2.2.4 Extraneous Edges 76
Network Blue 77
2.2.5 Unequal Path Lengths 77

Network Green . . . 78
2.3 Experimental Design . . . 81
2.3.1 Probability Weighting Elicitation . . . 81
2.3.2 Network Defense Tasks . . . 83
2.3.3 Procedures . . . 84
2.4 Results . . . 85
2.4.1 Network Attack Task . . . 85
2.4.2 Network Defense Tasks . . . 86
Summary . . . 86
Subject Consistency and Non-parametric Tests . . . 88
Cluster Analysis . . . 90
Regression Analysis . . . 92
2.5 Conclusion and Discussion . . . 96

3 BEHAVIORAL BANDITS: ANALYZING THE EXPLORATION VERSUS EXPLOITATION TRADE-OFF IN THE LAB . . . 100
3.1 Introduction . . . 100
3.2 Theory . . . 104
3.2.1 Discrete Implementation . . . 106
3.3 Experimental Design . . . 107
3.3.1 Treatments and Parameters . . . 107
3.3.2 Experiment . . . 109
3.3.3 Testable Hypotheses . . . 111
3.3.4 Behavioral Factors . . . 112
3.3.5 Procedures . . . 114
3.4 Results . . . 114
3.4.1 Hypothesis 1 . . . 116
3.4.2 Hypothesis 2 . . . 117
3.4.3 Hypothesis 3 . . . 119
3.5 Estimating Behavioral Factors . . . 121

3.5.1 Setup and Estimation . . . 121
3.5.2 Effects of Each Behavioral Factor . . . 124
3.5.3 Estimated Model Predictions . . . 126
3.6 Conclusion . . . 127

REFERENCES . . . 128

A CHAPTER 1 APPENDIX . . . 141
A.1 Example Stata and Python Code for a Normal DGP Simulation based Power Analysis . . . 141
A.1.1 Stata Code . . . 141
A.1.2 Discussion . . . 144
A.1.3 Python Code . . . 145
A.1.4 Python Code for QR Simulations of Bayesian Persuasion . . . 145
A.2 D-optimal Example With Treatment-specific Standard Deviation . . . 151
A.3 Differential Treatment Variance in Bayesian Persuasion . . . 151
A.4 Ex-ante Power Analysis of Omitted Papers . . . 152
A.4.1 De Clippel et al. (2014) . . . 152
Ex-ante Power Analysis . . . 152
Optimal Experimental Design . . . 153
A.4.2 Huck et al. (2011) . . . 154
Ex-ante Power Analysis . . . 158
Optimal Experimental Design . . . 159

B CHAPTER 2 APPENDIX . . . 160
B.1 Discussion of Numerical Solutions and Unique Solutions . . . 160
B.1.1 Common Edge Networks - Exponential Edge Defense Function . . . 160
B.1.2 Common Edge Network - Alternative Edge Defense Function . . . 163
B.1.3 Network Green . . . 166
B.1.4 $\alpha = 1$ with Exponential Edge Defense Function . . . 167
B.1.5 $\alpha = 0$ . . . 171

B.2 DOSE Procedures 171
B.2.1 Network Attack Task 171
B.2.2 Binary Lottery Task 172
B.3 Hypothesis 2 Robustness Tests 174
B.4 Binary Lottery Task Results 174
B.5 Individual Network Cluster Analysis 176
B.5.1 Network Red 176
B.5.2 Network Orange 177
B.5.3 Network Yellow 178
B.5.4 Network Blue 179
B.5.5 Network Green 180
B.6 Experiment Interface 181
B.7 Experiment Instructions 181
B.7.1 Overview 181
B.7.2 Binary Lottery Task 183
B.7.3 Network Attacker Task 183
B.7.4 Network Defense Tasks 186

C CHAPTER 3 APPENDIX 193
C.1 Theory Appendix 193
C.1.1 Continuous Time Predictions 193
C.1.2 Discrete Time Predictions 193
C.1.3 Payoff Hills and Behavioral Predictions 194
C.2 Empirical Appendix 197
C.2.1 Product Limit Estimator 197
C.2.2 Order Effects and Demographics 198
C.2.3 Censoring 200
C.3 Model Appendix 201
C.3.1 Robustness Checks 201
C.3.2 Effects for Each Treatment 202

C.4 Power Analysis Appendix 205
C.5 Instructions 209

VITA 214

LIST OF TABLES

1.1 Effects of Payoff Hill Asymmetry and Steepness on Power 23
1.2 Payoff Hills for Various Parameters in Abeler et al. (2011) 43
1.3 QR Simulations for Different Combinations of f in Abeler et al. (2011) 44
1.4 Probability of Actions by Treatment 48
1.5 Power for various parameters in Charness and Dufwenberg (2011) 49
1.6 Ex-ante Power Analysis for R. Chen and Chen (2011) 52
1.7 Ex-ante Experimental Design for R. Chen and Chen (2011). 53
1.8 Fehr et al. (2013) Power Analysis 57
1.9 Ifcher and Zarghamee (2011) Power Analysis 59
1.10 Required Total Number of Subjects for Various Approaches 60
2.1 Theoretical Predictions for Selected α with $B = 24$ 80
2.2 Spearman's ρ Correlation Table 89
2.3 Cluster Analysis 91
2.4 Tobit Regression Analysis 94
3.1 Values of p_0, $s\Delta$, and δ for each Treatment 107
3.2 Predictions for each Treatment in Discrete Time and Continuous Time 108
3.3 Summary Statistics for the Experiment 115
3.4 Estimates from the Structural Model 124
A.1 Effects of Unequal Variance 152
A.2 De Clippel et al. (2014) - Power of Preference Orderings at $N = 300$ 156
A.3 De Clippel et al. (2014) - Power of Selected Preference Orderings at $N = 50$ . . 157
A.4 Huck et al. (2011) QRE Power Analysis 158
B.1 One-sided p-values from Wilcoxon Signed Rank Tests of Hypothesis 2 for Different α Thresholds 174
B.2 Full Spearman's ρ Table. 175
B.3 Network Red Cluster Analysis 176
B.4 Network Orange Cluster Analysis 177
B.5 Network Yellow Cluster Analysis 178

B.6 Network Blue Cluster Analysis 179
B.7 Network Green Cluster Analysis 180
C.1 Tests of Order Effects in each Treatment 198
C.2 Demographic Robustness Checks 199
C.3 Estimates from Robustness Checks on the Structural Model 201

LIST OF FIGURES

1.1 Payoff Hills with Different Types of Asymmetries . . . 21

1.2 Example Normal DGP for Bayesian Persuasion . . . 32

1.3 Sender's Expected Profit for $\pi_R = 2$, $\pi_B = 1$, and $q^* = 0.5$ . . . 33

1.4 Sender's QRE DGP when $\lambda = 11.26$ . . . 35

1.5 Power Surface for QRE Treatment Priors . . . 37

1.6 Power Curves for π . . . 39

1.7 Charness and Dufwenberg (2011)'s '(5,7) Game' . . . 46

1.8 Game in Charness and Dufwenberg (2006) . . . 47

1.9 Effect of Increasing Step Size in the Minimum Effort Game . . . 54

2.1 Example Directed Network Attack Graph . . . 65

2.2 Allocation to Common Edge for $p(x_\mathrm{i}) = 1 - \exp(-x_\mathrm{i}/z)$ . . . 72

2.3 Allocation to Common Edge for $p(x_\mathrm{i}) = (x_\mathrm{i}/z)^b$ . . . 74

2.4 Network Graph with an Extraneous Edge (x_3) . . . 76

2.5 Network Graph with Unequal Path Lengths . . . 77

2.6 Allocation to Top (Long) minus Bottom (Short) Path for $p(x_\mathrm{i}) = 1 - \exp(-x_\mathrm{i}/z)$ 79

2.7 Network Attack Task Example . . . 82

2.8 CDF of Elicited α from the Network Attack Task . . . 86

2.9 CDFs of Average Behavior Across all Rounds in the Network Defense Tasks . . . 87

3.1 Example of the Experimental Interface . . . 110

3.2 Impact of Behavioral Factors on Experimentation . . . 112

3.3 Difference in Average Stopping Time in the Baseline Treatment and the Session's Other Treatment . . . 116

3.4 Mean Subject Stopping Times in Each Treatment . . . 119

3.5 Estimated Model Predictions for the Baseline Treatment as each Behavioral Factor is Varied . . . 125

3.6 Estimated Model Predictions for each Treatment using the Subset Approach . . . 126

B.1 Selected Numerical Solutions for Common Edge and $p(x_\mathrm{i}) = 1 - \exp(-x_\mathrm{i}/z)$ . . . 162

B.2 Selected Numerical Solutions for Common Edge and $p(x_\mathrm{i}) = (x_\mathrm{i}/z)^b$ . . . 165

B.3 Network Green - $\alpha = .55$, $B = 1.1$, $z = 2.3$ . . . 168

B.4 Network Green - $\alpha = .7$, $B = 369$, $z = 179$ 169

B.5 Network Green - $\alpha = .95$, $B = 1$, $z = 2$ 170

B.6 CDF of Elicited σ 174

B.7 CDF of Elicited α 174

C.1 Payoff Hills for each Treatment of the Experiment 194

C.2 Predictions for the High Prior Treatment as Different Behavioral Factors are Unilaterally Varied 195

C.3 Predictions for the Low Safe Action Treatment as Different Behavioral Factors are Unilaterally Varied 196

C.4 Predictions for the High Discount Factor Treatment as Different Behavioral Factors are Unilaterally Varied 196

C.5 Rate of Censoring in each Treatment under the Subset Approach 200

C.6 The Effect of Unilaterally Changing the CRRA Coefficient on Subset Approach Predictions from the Model 203

C.7 The Effect of Unilaterally Changing the Base Rate Neglect Parameter on Subset Approach Predictions from the Model 204

C.8 The Effect of Unilaterally Changing the Probability Weighting Parameter on Subset Approach Predictions from the Model 205

ABSTRACT

This thesis contains three chapters, each of which covers a different topic in experimental economics.

The first chapter investigates power and power analysis in economics experiments. Power is the probability of detecting an effect when a true effect exists, which is an important but under-considered concept in empirical research. Power analysis is the process of selecting the number of observations in order to avoid issues with low power. However, it is often not clear ex-ante what the required parameters for a power analysis, like the effect size and standard deviation, should be. This chapter considers the use of Quantal Choice/Response (QR) simulations for ex-ante power analysis, as it maps related data-sets into predictions for novel environments. The QR can also guide optimal design decisions, both ex-ante as well as ex-post for conceptual replication studies. This chapter demonstrates QR simulations on a wide variety of applications related to power analysis and experimental design.

The second chapter considers a question of interest to computer scientists, information technology and security professionals. How do people distribute defenses over a directed network attack graph, where they must defend a critical node? Decision-makers are often subject to behavioral biases that cause them to make sub-optimal defense decisions. Non-linear probability weighting is one bias that may lead to sub-optimal decision-making in this environment. An experimental test provides support for this conjecture, and also other empirically important biases such as naive diversification and preferences over the spatial timing of the revelation of an overall successful defense.

The third chapter analyzes how individuals resolve an exploration versus exploitation trade-off in a laboratory experiment. The experiment implements the single-agent exponential bandit model. The experiment finds that subjects respond in the predicted direction to changes in the prior belief, safe action, and discount factor. However, subjects also typically explore less than predicted. A structural model that incorporates risk preferences, base rate neglect/conservatism, and non-linear probability weighting explains the empirical findings well.

1. IMPROVING EX-ANTE POWER ANALYSIS WITH QUANTAL RESPONSE SIMULATIONS

1.1 Introduction

Statistical power, the likelihood of detecting an effect when a true effect exists, is an important consideration in any empirical research, both inside and outside of economics. Low power is problematic by definition, but even when an effect is detected in a low powered study, the finding is either likely to be false (Ioannidis, 2005), or the estimated effect size can be exaggerated or in the wrong direction (Gelman & Carlin, 2014). These issues contribute to the replication or credibility 'crisis' that affects a wide range of scientific disciplines. It also can lead to research efforts being misplaced, with theories or extensions being proposed on effects that do not actually exist. Power is frequently ignored in economics, although that is becoming less true over time.[1] As a result, low power is pervasive in empirical economics research, which is clearly a situation that needs to be addressed.[2] Unlike other empirical fields, experimental research (such as lab or field experiments) has more control over the power of a study. In particular, the number of observations in a study is a decision variable of the researcher, where increasing observations increases power. Despite this, economics experiments have also historically neglected the consideration of statistical power, and thus also have exhibited low power.[3]

Power analysis is the process of determining an appropriate sample size for an experiment. As a result of previous low power in the field, power analysis is rightfully becoming increasingly important in experimental economics. How a power analysis is often conducted is that estimates of the treatment effect size and standard deviations are obtained from previous studies, or inferred from introspection or beliefs. These parameters are then used in a closed-form representation of power (typically from a two-sample t-test) or a statisti-

[1]↑Ziliak and McCloskey (2004) report that only 4% of the papers published in the *American Economic Review* in the 1980s mention power, increasing to 8% in the 1990s.

[2]↑Ioannidis et al. (2017) report a median power of 18% over a wide variety of empirical economics fields.

[3]↑Zhang and Ortmann (2013) report 1 out of 95 papers in *Experimental Economics* from 2010-2012 mention statistical power, as well as a median power of 25% in experimental studies of the Dictator Game. For reference, a power of 80% is generally considered sufficient in experimental economics (Moffatt, 2016, pg. 22; List et al., 2011, pg. 448).

cal software package, which gives the required sample size. The question is, where should these parameters come from? The treatment effect size could be specified from the point predictions of the model in question. However, economic theory is typically more successful in predicting comparative statics than point estimates. In addition, economic theory is typically silent on the expected standard deviation, so it cannot provide guidance on that parameter.[4] A more suitable approach is to use past experimental data as a starting point for these parameters. However, sufficiently close or exhaustive experimental environments may not exist. Even with appropriate data, a novel experiment will differ from past experiments in meaningful ways. For example, a new treatment might be conducted, or different parts of a theory's parameter space might be explored. It is not clear ex-ante what the actual treatment effect size would be, or how much variability subjects would exhibit, in these new treatments.

This chapter explores a potential alternative method of power analysis based on a Quantal Choice/Response framework (henceforth QR) of individual behavior over discrete options (McFadden, 1976). The QR, also known as a logistic regression, has been successfully applied to many fields and applications, such as computer science, biometrics, transportation, psychology, and other social sciences, to help explain or model choice behavior. The QR framework is not limited to single decision-maker problems, as the Quantal Response Equilibria (henceforth QRE) extends it to multiple interacting decision-makers (McKelvey & Palfrey, 1995). The QRE has proven successful in explaining subject behavior in strategic settings and is widely used in experimental economics and political science. The QR framework assumes a noise structure in which individuals choose actions in proportion to their relative payoffs, such that better actions are chosen more frequently. Relative payoffs can be presented graphically as a 'payoff hill', which I describe further in Section 1.2, and these hills intuitively capture subject behavior through the incentives that they face. The method of power analysis that I propose fits a structural QR framework to the most closely related previous experimental data-sets. Given the parameters from that model, the QR framework provides predictions for the probability that each possible action would be taken in the pro-

[4]↑Notable exceptions exist, such as price dispersion (Burdett & Judd, 1983) where predictions of variability have been experimentally tested (Cason et al., 2021).

posed new experiment. In other words, the QR approach takes what is known about subject behavior in previous studies, and maps that onto likely subject behavior in the new experiment that is yet to be conducted. From the predicted probabilities, simulated data-sets can be generated, and it is on these data-sets that power analysis is then conducted. As QR has proven successful in modeling choice behavior, with appropriate parameters inferred from past data the simulated data-sets should be a reasonable approximation of likely subject behavior. The more accurate the approximation of actual subject behavior is, the more accurate the ex-ante power analysis will be, meaning the study should be reasonably powered ex-post.

This chapter describes and exhibits the QR based simulation approach to power analysis. In Section 1.2 I describe the payoff hills which effectively underlie QR, and arguably contribute to its success. Payoff hills can help explain where empirical treatment effect sizes are likely to differ from theoretical point predictions, and what treatments are likely to exhibit high subject variation, both of which are important considerations for power analysis. In Section 1.3, I outline the QR simulation approach in more detail, and describe how it could be used to help design experiments. In particular, experiments are typically designed such that theoretical predictions are distantly spaced, i.e. the treatment effect size from point predictions is maximized. However, if theoretical predictions differ substantially from subject behavior, it should be that the likely empirical average behavior is distantly spaced instead. The QR framework provides estimates for likely behavior for a variety of possible design decisions. Design decisions with respect to power also need to consider the impact on likely subject variability. The QR framework also provides an estimate of likely subject variability for many possible design decisions. I confirm the importance of considering the treatment effect and standard deviations simultaneously using closed-form expressions of power, as it is not always optimal to maximize the treatment effect size. In Section 1.4, I demonstrate the application of the QR approach using a motivating example of an experiment in Bayesian Persuasion, and compare it to a more standard approach to power analysis. The Bayesian Persuasion environment is a non-abstract example where maximizing the theoretical treatment effect size is very poor for power. Finally, in Section 1.5 I apply the QR simulation approach to various high-profile papers. I conduct a thought exercise where I place myself in a similar situation as the original authors, that is, I largely ignore the data in the paper and rely on other sources to conduct the analysis. I find that the QR approach can be applied to a wide range of environments, sometimes with creative modifications to incorporate the operative channel of the treatment effect. However, there are exceptions, namely experiments where no explicit theory is provided that could explain a treatment difference. The ex-ante power analysis from the QR simulations compares favorably to an ex-post power analysis, with the caveat that ex-post analysis is an excessively high standard for ex-ante analysis to meet. I also suggest alternative parameters and treatments to increase power for each paper, which decreases the required number of subjects. Optimally selecting design parameters in such a way could also be conducted ex-post to guide a conceptual replication that is 'efficient', in that it minimizes subject costs.

1.2 Payoff Hills and Power Analysis

1.2.1 Payoff Hills

A payoff hill is a graphical representation of the level of payoffs arising from different actions, so-called as the optimum action is the 'peak' of the hill, while the decreases in payoffs from deviating away from the optimum form the 'sides' of the hill. As QR agents exhibit noisy behavior in proportion to their relative payoffs, the QR effectively incorporates the payoff hill, making it a useful visual aid to help explain why QR is successful in explaining choice behavior. Payoff hills are an important design factor that most experimenters are aware of, as they reflect the incentives subjects face to behave optimally. Much attention was brought to this topic by Harrison (1989), who proposes a metric of foregone expected income, which is the inverse of what I describe as a payoff hill. In what is now known as the 'flat-maximum critique', Harrison notes that in an experimental implementation of a first-price sealed bid auction (Cox et al., 1988), large deviations from the optimal action result in only in small decreases in expected payoffs. Therefore, reading too much into such deviations could be problematic. Following this logic, payoff hills that are 'flatter' are likely to exhibit more noisy behavior, because subjects suffer smaller payoff consequences for deviations from the optimal action. The flatness or steepness of a payoff hill can therefore influence the standard deviation, and thus have an impact on power analysis. The Harrison (1989) paper sparked a flurry of discussion, with multiple comments on the article published (e.g. Cox et al., 1992, Kagel and Roth, 1992, Merlo and Schotter, 1992, and Harrison, 1992). Of particular note, Friedman (1992) points out that the flat-maximum critique would predict deviations in both directions if the flatness was symmetrical about the optimal action. In order to predict consistent deviations in one direction (i.e. above or below the optimal action), it would need to be the case that the payoff hill is asymmetric with one side of the hill being flatter than the other. Subjects will be more likely to make deviations in the direction where they face smaller payoff consequences for doing so. Such directional errors can influence the treatment effect size, especially when the asymmetry of the payoff hill differs by treatment, and thus can affect power analysis. The following section investigates the impact of both the steepness/flatness and asymmetry of the payoff hill on the likely treatment effect size and standard deviations.

1.2.2 Effects of Payoff Hills on Power Analysis

Power analysis is largely driven by three parameters: the treatment effect size τ; the standard deviation of the first treatment σ_1; and the standard deviation of the second treatment σ_2. Due to its impact on likely subject behavior, payoff hills can thus affect all three of these parameters, which is now demonstrated with a constructed example.

For this example, subjects make their decision x over an action space of $x \in [0, 10]$. To consider asymmetry, I define the payoff hill over x to be a piece-wise Gaussian function about some maximal payoff $a = 1$ which occurs at point b, while allowing for the possibility for the variable c to be different on either side of the maximum point. The variable c represents the steepness of the payoff hill, as it determines the slope of the function when moving away from the optimal point, with lower values of c being steeper.[5] More explicitly, the payoff hill is:

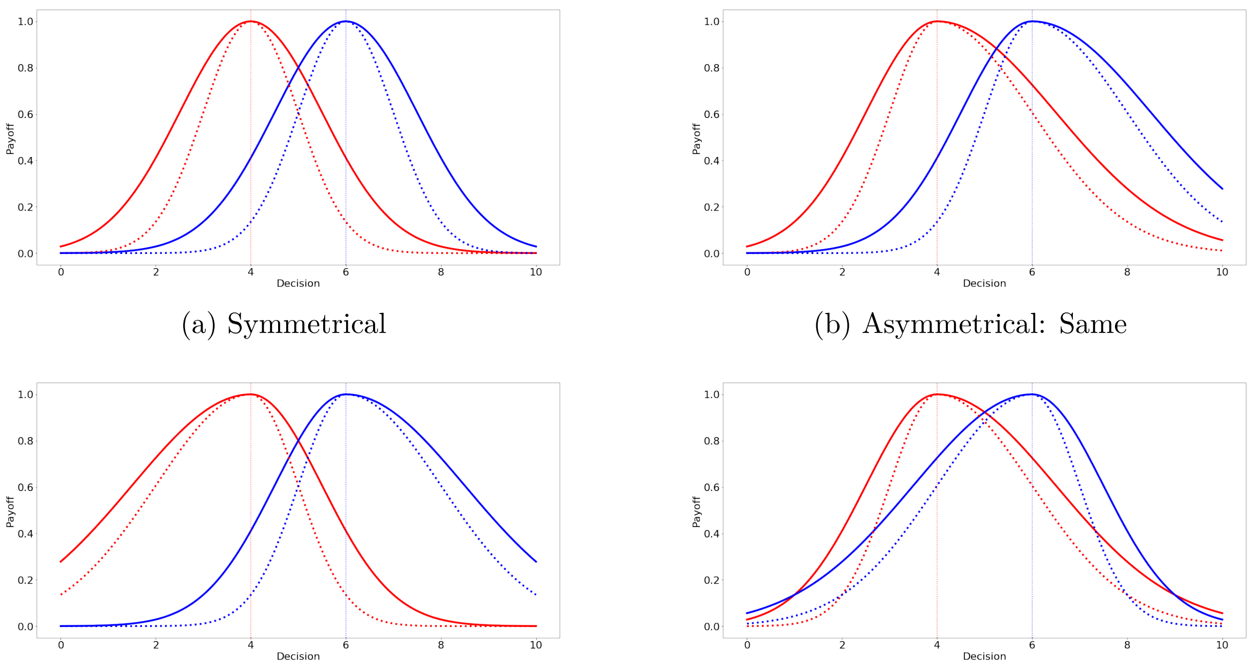
[5]↑The variable c is analogous to the standard deviation in the normal distribution, where lower standard deviations imply mass is distributed more 'tightly' about the mean.

$$\pi(x; a, b, c_{LHS}, c_{RHS}) = \begin{cases} a \exp(-\frac{(x-b)^2}{2c_{LHS}^2}), & \text{if } x < b \\ a \exp(-\frac{(x-b)^2}{2c_{RHS}^2}), & \text{otherwise} \end{cases}$$

Suppose there are two treatments, where the payoff maximizing action is $b_1 = 4$ in Treatment 1, and $b_2 = 6$ in Treatment 2. An initial estimate of the treatment effect size can be derived from the point predictions of standard theory, $\tau = b_2 - b_1 = 6 - 4 = 2$. However, this estimate is likely erroneous, as it is ignorant to the shape of the payoff hill. I consider four potential types of treatment payoff hills, a symmetrical payoff hill where $c_{LHS} = c_{RHS}$ for both treatments, and three different combinations of asymmetric payoff hills. The first type of asymmetric hill has the flatter side of the hill to the same side of the optimum value in both treatments, with $c_{1,LHS} = c_{2,LHS} < c_{1,RHS} = c_{2,RHS}$. The second type of asymmetric hill has the flatter side of the payoff hills facing away each other, with $c_{1,RHS} = c_{2,LHS} < c_{1,LHS} = c_{2,RHS}$. The third type of asymmetric hill has the flatter sides facing towards each other, with $c_{1,LHS} = c_{2,RHS} < c_{1,RHS} = c_{2,LHS}$. These four types of payoff hills can be considered in conjunction with their relative steepness as encapsulated by c. Diagrams of these four combinations of payoff hills for a 'baseline' hill with $c_{steep} = 1.5$ and $c_{shallow} = 2.5$, and a 'steeper' hill with $c_{steep} = 1.0$ and $c_{shallow} = 2.0$ are shown in Figure 1.1. Similar payoff hills can occur in more natural experiments, but for the purposes of this exercise I assume such payoff hills are constructed exactly in a single decision maker environment.[6]

I now analyze the impact that asymmetry of the payoff hills has on the likely treatment effect size. As mentioned previously, an immediate initial estimate of the treatment effect size would be $\tau = 2$, from the point predictions of theory. However, asymmetries in the hills will impact the likely empirical treatment effect size that is observed. This is most evident from the payoff hills displayed in Figures 1.1c and 1.1d. Subjects in Figure 1.1c are more likely to deviate from optimal behavior towards the edges of the action space, increasing the treatment effect size. Whereas those in Figure 1.1d will be deviating towards the center of the action space, decreasing the treatment effect size. In the limit, this distortion from the

[6]↑It is typically the case that the experimenter does not have this level of precise control over the payoff hill, however, it is also the case that it is rare that the experimenter has no control at all over the payoff hill.



(a) Symmetrical

(b) Asymmetrical: Same

(c) Asymmetrical: Away

(d) Asymmetrical: Towards

Figure 1.1. Payoff Hills with Different Types of Asymmetries

point prediction τ would diminish as both $c_{RHS} \to 0$ and $c_{LHS} \to 0$, or as $\lambda \to \infty$. Therefore, it is not only asymmetry that effects the treatment effect size, the steepness of the payoff hill also matters. Ignoring either of these payoff hill features can result in erroneous conclusions about power through incorrect estimates of the treatment effect size. The steepness of the payoff hill more directly impacts the standard deviation rather than the effect size, and the standard deviation is also a determinant of power. Therefore, when considering the impact of the payoff hill on power, these two parameters need to be considered in tandem.

In order to calculate power for the different combinations of asymmetry and steepness, τ, σ_1, and σ_2 need to be specified, which is where the QR framework comes in handy. I apply the QR framework to map the payoff hills into a probability of choosing a particular action using the logit choice rule $p(a_\mathrm{i}) = \frac{\mathrm{e}^{\lambda E \pi_{a_\mathrm{i}}}}{\sum_{a_\mathrm{j} \in A} \mathrm{e}^{\lambda E \pi_{a_\mathrm{j}}}}$ with an arbitrarily selected $\lambda = 1$.[7] Table 1.1 reports the output of a simulation based power analysis for various combinations of c_{steep} and $c_{shallow}$. Asymmetric payoff hills do affect the treatment effect size in the manner previously suggested, in that it is suppressed when the shallow sides face towards each other and typically amplified when the shallow sides face away. Increasing the steepness of the payoff hill (decreasing c) unambiguously decreases the standard deviation. There is also an interaction with the boundaries of the action space, as evidenced by the 'Same' column in Table 1.1 where the treatment effect size is suppressed, which could be lessened if the action space were extended or eliminated if the boundary were removed.[8] This interaction is potentially important as experiments typically have bounded action spaces. Finally, flatter payoff hills overall, as demonstrated in the last row of Table 1.1, suppress the treatment effect. This is due to subjects getting closer to indifference over all actions, which in the QR framework entails uniform random play, and thus no treatment effect. All of these potential factors arising from the structure of the payoff hills should be considered when evaluating power.

[7]↑I discretized the action space as 10000 evenly spaced points from 0 to 10 inclusive, and binned adjacent actions depending on their impact on payoffs.

[8]↑The bounded action space also explains the small deviations from $\tau = 2$ in the Symmetrical case.

Table 1.1. Effects of Payoff Hill Asymmetry and Steepness on Power

Parameters	Symmetrical	Asymmetrical: Same	Asymmetrical: Away	Asymmetrical: Towards
$c_{steep} = 1.5$ $c_{shallow} = 2.5$	$\tau = 1.944$ $\sigma_1 = 1.804$ $\sigma_2 = 1.806$ **Power**= 64.6%	$\tau = 1.604$ $\sigma_1 = 2.329$ $\sigma_2 = 2.072$ **Power**= 28.0%	$\tau = 2.147$ $\sigma_1 = 2.069$ $\sigma_2 = 2.072$ **Power**= 60.4%	$\tau = 1.058$ $\sigma_1 = 2.329$ $\sigma_2 = 2.333$ **Power**= 9.1%
$c_{steep} = 1.0$ $c_{shallow} = 2.5$	$\tau = 1.997$ $\sigma_1 = 1.221$ $\sigma_2 = 1.223$ **Power**= 97.1%	$\tau = 1.657$ $\sigma_1 = 2.068$ $\sigma_2 = 1.771$ **Power**= 41.3%	$\tau = 2.735$ $\sigma_1 = 1.769$ $\sigma_2 = 1.771$ **Power**= 95.1%	$\tau = 0.575$ $\sigma_1 = 2.068$ $\sigma_2 = 2.072$ **Power**= 3.5%
$c_{steep} = 1.5$ $c_{shallow} = 2.0$	$\tau = 1.944$ $\sigma_1 = 1.804$ $\sigma_2 = 1.806$ **Power**= 64.6%	$\tau = 1.796$ $\sigma_1 = 2.095$ $\sigma_2 = 1.974$ **Power**= 43.7%	$\tau = 2.150$ $\sigma_1 = 1.971$ $\sigma_2 = 1.974$ **Power**= 65.9%	$\tau = 1.441$ $\sigma_1 = 2.095$ $\sigma_2 = 2.098$ **Power**= 24.4%
$c_{steep} = 1.0$ $c_{shallow} = 2.0$	$\tau = 1.997$ $\sigma_1 = 1.221$ $\sigma_2 = 1.223$ **Power**= 97.1%	$\tau = 1.835$ $\sigma_1 = 1.831$ $\sigma_2 = 1.675$ **Power**= 61.4%	$\tau = 2.707$ $\sigma_1 = 1.673$ $\sigma_2 = 1.675$ **Power**= 96.8%	$\tau = 0.961$ $\sigma_1 = 1.831$ $\sigma_2 = 1.834$ **Power**= 12.7%
$c_{steep} = 2.0$ $c_{shallow} = 3.0$	$\tau = 1.629$ $\sigma_1 = 2.262$ $\sigma_2 = 2.265$ **Power**= 27.3%	$\tau = 1.215$ $\sigma_1 = 2.655$ $\sigma_2 = 2.404$ **Power**= 10.3%	$\tau = 1.385$ $\sigma_1 = 2.401$ $\sigma_2 = 2.404$ **Power**= 15.9%	$\tau = 1.042$ $\sigma_1 = 2.655$ $\sigma_2 = 2.660$ **Power**= 6.7%

Power is evaluated using a t-test at the $\alpha = 0.01$ level with $N = 15$ observations per treatment.

1.3 QR Simulations for Optimal Experimental Design

1.3.1 Describing the QR Simulation Power Analysis

A simulation approach specifies a data-generating process (DGP), which can then used to generate many simulated data-sets upon which power analysis can be conducted. The QR is an appropriate choice for a DGP of subject behavior, as it has been successful in predicting a wide range of experimental environments, and thus is likely to resemble the true DGP. The specific DGP used in a QR simulation is the logit function: $p(a_{\mathrm{i}}) = \frac{\mathrm{e}^{\lambda EU_{a_{\mathrm{i}}}}}{\sum_{a_{\mathrm{j}} \in A} \mathrm{e}^{\lambda EU_{a_{\mathrm{j}}}}}$, where $EU_{a_{\mathrm{i}}}$ is the expected utility or payoff of choosing action i, and λ is a decision precision or noise parameter. The expected utility of each action, $EU_{a_{\mathrm{i}}}$, will depend on the environment in question, and will be a function of the parameters in the experimental design (e.g. monetary payoffs, treatments) as well as any other utility relevant parameters (e.g. risk aversion, other-regarding preferences). In a single decision-maker problem, it is relatively straightforward to calculate $EU_{a_{\mathrm{i}}}$ for each action. In a strategic setting, calculating $EU_{a_{\mathrm{i}}}$ is more involved as it now additionally depends on other player's actions. The other player's actions are also probabilistic in a similar manner, i.e. by the logit choice function based on their expected payoffs. In the strategic case the QRE must be computed, which is a fixed point where each player's logit choice rule is a best response to all other player's logit choice rules.

For the QR approach, the main parameter the researcher has to specify is λ, a decision precision parameter that represents how frequently subjects play 'better' actions. Reasonable estimates of λ should be obtained by structurally fitting a QR model of the logit function to data from the most closely related previous studies. Additional parameters may need to be specified if the theory requires them (e.g. a hypothesis about reciprocity may require an estimate of other-regarding preferences). These parameters should also be calibrated from previous experiments in the same manner as λ.[9] In the event that no suitable data is available, a small pilot could be conducted and the parameters estimated from that, or other parameters that are considered acceptable by the literature could be used. The parameters inferred from the previous data are then substituted into the QR logit choice rule for the new experimental environment in question. To finalize the DGP, the statistical test that will be used in the experiment needs to be considered. For example, if the final statistical test will cluster its standard errors, then the DGP should have something in it to justify the need for clustering, otherwise the power analysis may be misleading. There are various ways this could be implemented, for example, an individual or session level λ or other parameter could be drawn. This finalizes the specification of the DGP, from which simulated data-sets can be generated. Power is then calculated by performing the desired statistical test on each simulated data-set, and counting the number of times the null hypothesis is rejected. A simulation approach is particularly helpful for more complicated statistical tests that do not have closed-form expressions for their power.

1.3.2 Optimal Experimental Design

There are many considerations that go into an experimental design, which include the number and levels of treatments, whether to vary treatments between- or within-subjects,

[9]↑The calibration will likely be imperfect considering that it was probably not the original study's intention to structurally fit such parameters. Instead, the aim is to obtain reasonable estimates and improve upon subjective guesses of parameters ex-ante.

whether to use the strategy method, have fixed groups or random re-matching, what payoffs should occur from outcomes, and so forth and so on. In the strictest possible sense, power analysis determines only one design decision, the number of subjects to have in each treatment, when other design decisions are held fixed. However, any design decision could be evaluated and selected on the basis of power. Optimal experimental design refers to making design decisions with respect to maximizing some metric. I focus on the metric of power, and the design decisions of selecting treatment levels and environmental parameters that are fixed by treatment (e.g. payoffs, probabilities, etc.).[10] Traditionally, these parameters would be chosen to increase the difference in theoretical predictions, which usually requires exploring the furthest reaches of the parameter space.[11] This rule of thumb is often successful for two reasons. Firstly, if the effect size is linear in the treatment level, by 'D-optimal design' (Moffatt, 2016, ch. 14) it is always optimal to space the treatment levels as far apart as feasibly possible.[12] Secondly, maximizing the theoretical treatment effect size usually maximizes the actual treatment effect size, which is beneficial for power. However, on the first point, it is not always D-optimal to maximize the difference in treatment levels if the relationship is non-linear, which is effectively the case if the standard deviations depend on the treatment level.[13] On the second point, Section 1.2 already established that asymmetric payoff hills can result in different theoretical and actual effect sizes, so it is not immediately clear that maximizing one maximizes the other. However, even if it does, the impact that the treatment level or any other parameter has on the standard deviation needs to be taken into account. It could be the case that the standard deviation increases substantially, meaning a given parameter change might not increase power despite increasing the effect size. The rule of thumb proves reasonably successful in a lot of environments, otherwise it would not

[10]↑It should be noted that the strategy method unambiguously increases power over direct response, and doing treatments within-subject is very beneficial for power if there is substantial individual heterogeneity.

[11]↑For example, from classic textbooks on experimental design: *"A related aspect of calibration is the use of a design in which the predictions of alternative theories are cleanly separated."* (Davis & Holt, 1993, pg. 28), *"Use widely separated levels to sharpen the contrasts."* (Friedman & Sunder, 1994, pg. 31).

[12]↑D-optimal design maximizes the determinant of the information matrix implied by maximizing the log-likelihood function. This minimizes the confidence intervals on estimated parameters, which is similar to maximizing power.

[13]↑For an example with D-optimal design where the standard deviations depend on the treatment level, see Appendix A.2.

be accepted as such. However, QR simulations can help identify the environments where the described issues with the rule of thumb could be present.

A simulation approach is ideal for optimal experimental design, as it does not require analytical solutions for power. Instead, power is determined by simply calculating the proportion of data-sets where the null was rejected, as a true effect is known to exist in the simulation DGP.[14] Using a QR simulation in particular provides reasonable predictions for how treatment levels and other parameters will affect both the treatment effect size and treatment-specific standard deviation. This means that optimal parameters can be selected considering both the likely treatment effect size and standard deviations simultaneously and in a realistic manner.

Note that power is not the only metric that could be used for optimal experimental design. For example, the researcher may want to identify strategies from subject behavior in a probabilistic environment where they may not observe their behavior in all relevant states or for enough periods. A simulation approach would count the number of subjects that meet the criterion, a metric which could be maximized. Another metric could be the accuracy of estimated parameters compared to their true parameter, which is known in a simulation. This metric can guide experimental design through selecting the most effective questions to identify certain types, as well as confirm the likely identifiability of a structural model given a set of questions and responses.

1.3.3 Analysis of Closed-form Expressions for Power

The purpose of this section is to confirm that it is not always optimal to maximize the actual treatment effect size when it comes to power, and that standard deviations need to be jointly considered. Restricting the analysis to tests with closed-form expressions for power removes the need for simulations.

[14]↑ A good example of using (non-QR) simulations to select experiment parameters to improve power is Rutström and Wilcox (2009), who aimed to distinguish between two competing models of subject behavior.

Binary Actions

Binary actions are common in economics experiments, particularly as it is often considered good design to have the most simplified environment possible that still tests the conjecture. The random nature of an individual binary decision is a Bernoulli trial with probability p, and multiple observed decisions makes up a Binomial distribution. A common statistical test used for binary data is a test of proportions, either the exact Binomial probability test or the large-sample approximation Z-test of proportions, either of which can be a one-sample (i.e. point prediction) or two-sample (i.e. treatment effect) test. The standard error of the observed proportions is determined by the sample estimate of p, and is $\sigma = \sqrt{\frac{p(1-p)}{N}}$. To decrease σ, one could either increase N, or have p be close to zero or one. The two-sampled, two-sided power has a closed form solution: $\Phi\{\frac{(p_2-p_1)-c-z_{1-\alpha/2}\sigma_P}{\sigma_D}\}+\Phi\{\frac{-(p_2-p_1)-c-z_{1-\alpha/2}\sigma_P}{\sigma_D}\}$, where c is a normal approximation continuity correction ($c = 2/n$ when $n_1 = n_2 = n/2$), $\sigma_P = \sqrt{\bar{p}(1-\bar{p})(1/n_1 + 1/n_2)}$ is the pooled standard deviation and $\sigma_D = \sqrt{p_1(1-p_1)/n_1 + p_2(1-p_2)/n_2}$ is the standard deviation of the difference between proportions (StataCorp, 2013). Power can be increased by increasing $p_2 - p_1$ (i.e. increase the treatment effect), or by moving p_1 or p_2 closer to one or zero (i.e. decrease the standard deviation of either sample), or increasing n_1 and n_2 (again, decreasing the standard deviation). In terms of choosing an optimal treatment intensity, if possible the optimum would be to set $p_1 = 0$ and $p_2 = 1$, maximizing the treatment effect size and minimizing the standard deviation. However, p_1 and p_2 are provided by the model, and are subject to both theoretical and practical constraints, in that it might not be possible to choose parameters in such a way to achieve both $p_1 = 0$ and $p_2 = 1$. Suppose, for whatever reason, that the maximum possible effect size is $p_2 - p_1 = 0.25$. If this were the case, power would be maximized at $p = 0$, $q = .25$ or $p = 0.75$, $q = 1$, but not at any equivalent interior combination of the same magnitude (e.g. $p = 0.25$, $q = 0.5$). This is an example of where it is not sufficient to only consider the treatment effect size when choosing parameters, as the standard deviation should be considered as well. It is even the case that smaller treatment effects can be desired over larger ones due to the standard error. One example with a large difference in treatment effect sizes would be the selection of $p_{1A} = 0.0$ and $p_{2A} = 0.3$ over

$p_{1B} = 0.3$ and $p_{2B} = 0.7$. Experiment A has a smaller treatment effect (0.3 vs 0.4), but has slightly improved power (84.8% vs 83.2%). Another example with a larger difference in power is $p_{1C} = 0.0$ and $p_{2C} = 0.29$ over $p_{1D} = 0.35$ and $p_{2D} = 0.65$. Experiment C has a slightly smaller treatment effect (0.29 vs 0.3), but has substantially improved power (82.8% vs 54.4%). So even in this relatively simple binary environment (where the standard deviation is determined by the given p), we still must be cognizant of the fact that maximizing treatment effect size alone will not necessarily maximize power.

Non-Binary Actions or Outcomes

Sometimes the parameter of interest is the average outcome that results from binary actions (e.g. election outcomes from a group of binary voters), or that the action set is larger than 2 decisions, and therefore no longer binary. In these cases, we use tests of the differences in means, and the most commonly used test for this is the t-test. The functional form for power of the two-sided version of this test is: $\Phi\{\frac{(\mu_2-\mu_1)}{\sigma_D} - z_{1-\alpha/2}\} + \Phi\{-\frac{(\mu_2-\mu_1)}{\sigma_D} - z_{1-\alpha/2}\}$, where $\sigma_D = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$. This is similar to the functional form for the z-test of proportions considered in the previous section, except now the mean of a treatment is decoupled from the standard deviation. Increasing power could be achieved by increasing the treatment effect size, decreasing either of the standard deviations, or increasing either of the sample sizes. However, changing the treatment effect size in an experiment without influencing the standard deviation is unlikely. It is possible that increasing the treatment effect size could actually decrease power, if the standard deviation also increases substantially, due to whatever experimental change induced the increase in the treatment effect size. To construct an example, I hold Treatment 2 constant at $\mu_2 = 0$, and $\sigma_2 = 1$, and let Treatment 1 have $\mu_1 \in [0, 1]$ and $\sigma_1 = f(\mu_1)$. With $\mu_2 = 0$, the treatment effect size is just the value of μ_1. An example of where maximizing μ_1 does not maximize power is when $f(\mu_1) = 2\mu_1^2 + \mu_1 + 1$, where setting $\mu_1 = 1$ yields a power of 27.8%, while power is maximized at $\mu_1 = 0.714$, with a power of 29.9%. More extreme examples could be arbitrarily constructed, but this again illustrates that maximizing the treatment effect size alone is not necessarily sufficient to maximize power.

1.3.4 Optimal Design and Replications

Replication in economics is an important but under-produced form of research, both inside and outside experimental economics (Duvendack et al., 2017; Maniadis et al., 2014). There have been recent laudable efforts and proposals to increase the provision of this crucial research.[15] Some proposals include increasing the benefit of independent replications by offering co-authorship (Butera et al., 2020), or including the number of replications as a metric for an author's research quality (Maniadis et al., 2015). Camerer et al. (2019) describe two types of replications, direct and conceptual. A direct replication is where the same exact experiment is conducted with as few procedural differences as feasibly possible. A conceptual replication is where the same hypothesis as the original study is tested but using different methods or parameters. The bulk of the encouragement in this area is with regards to direct replication, but conceptual replication is just as important, particularly in experiments that test an explicit economic model. A conceptual rather than direct replication provides more insight on whether the underlying model and its proposed operative channel is an accurate description of behavior, or what extensions or caveats to the model may be required. The suggestions to incentivize more replications tend to focus on increasing the benefits of replication, rather than lowering the costs. I propose the use of QR based simulations to design conceptual replications that economize on the required number of observations. The model parameters and treatment levels can be altered from the original study to maximize power, which minimizes the required number of subjects and thus the expenditure in terms of both time and money. I call a conceptual replication that explicitly minimizes costs an 'efficient replication'. A QR simulation is particularly well suited for efficient replication, as an appropriate data-set with which to calibrate the required parameters already exists.

[15]↑For example, the establishment of the *Journal of the Experimental Science Association*, which aims in part to publish under-represented experimental research like replications, and the Experimental Economics Replication Project (Camerer et al., 2016), which conducts high-powered replication studies on multiple high-profile papers.

1.4 Power Analysis Using Bayesian Persuasion as an Illustrative Example

1.4.1 Bayesian Persuasion Overview

Bayesian Persuasion (Kamenica & Gentzkow, 2011) is a model of information provision with commitment that has formed the basis for a substantial amount of research.[16] Bayesian Persuasion has been applied to a wide variety of areas, such as finance, optimal feedback, medical testing and research, insurance, and many others too numerous to list here (Kamenica, 2019). Despite this, there have been relatively few experimental studies of Bayesian Persuasion.[17] I use a hypothetical experimental test of Bayesian Persuasion as a motivating example to illustrate the QR approach to power analysis, as well as its potential benefits relative to a more standard approach.

In Bayesian Persuasion, there is a Sender and a Receiver who are paired together. The Receiver has a true unobserved state, either Red or Blue, with a common initial prior p that their state is Red. The Sender designs an information structure that can condition on the Receiver's true state, sending one message that either says r or b. The Sender chooses x and y, where $Pr(b|Blue) = x$ and $Pr(r|Red) = y$. The Receiver observes x and y as well the generated message, updates his belief according to Bayes Rule, and then decides whether to choose Red or Blue. The Receiver gets a payoff of π_m if he chooses the same color as his true state, and π_0 otherwise, with $\pi_m > \pi_0$. The Sender gets a payoff of π_R if the Receiver chooses Red, and π_B if the Receiver chooses Blue, with $\pi_R > \pi_B$. The Sender's optimal strategy is to set the information structure in such a way as to minimally update the Receiver's beliefs such that the Receiver wants to choose Red. If messages are straightforward (i.e. a message is seen as a suggested action), then this involves setting $y = 1$ and $x = \frac{1-\frac{p}{p^*}}{1-p}$, where p^* is the Buyer's reservation threshold (in the current symmetrical setup $p^* = .5$).

Suppose the hypothesis in question is whether Sender's behavior (i.e. their choice of x) changes in response to p. I first consider a power analysis approach of perturbing behavior about the theoretical predictions.

[16] ↑Google Scholar reports more than 1600 citations.

[17] ↑I am aware of the following papers: Q. Nguyen (2017), Au and Li (2018), Fréchette et al. (2018), and W. Wu and Ye (2019).

1.4.2 Normal Distribution DGP

As mentioned previously, there is not a wealth of previous experimental data to help guide a power analysis. For this current example, consider a between-subjects treatment where $p = \frac{1}{5}$ or $p = \frac{1}{3}$. The optimal x's given those p's are $x^* = \frac{3}{4}$ and $x^* = \frac{1}{2}$ respectively. I assume only a one-shot interaction between the Sender and the Receiver for this exercise. A power analysis using a standard rule of thumb would require a treatment effect size, τ, and the level of noise in an individual's behavior in each treatment, σ_0 and σ_1, to be specified. For now, suppose there is no strong a-priori reason to believe that the noise should differ between treatment, so $\sigma_0 = \sigma_1 = \sigma$, and I also assume no heterogeneity.[18] The data-generating process (DGP) is relatively simple: $Y_{iT} = \alpha + \tau T + \epsilon_i$, where $\epsilon_i \sim N(0, \sigma^2)$. If $p = \frac{1}{3}$ is considered $T = 0$, then $\alpha = .5$ and $\tau = \frac{1}{4}$ follows from the theory. All that remains is to specify a σ. Typically this should be chosen based on the most relevant previous experiments. However, given that data is not currently available, and this is just an illustrative example, I set it to $\sigma = 0.5$.[19] Given the assumed normal DGP, the Stata code to yield the required sample size for 80% power for a t-test is: `sampsi .5 .75, sd(.5) p(.8)`, which yields a sample size of $n = 63$, so 63 Sender observations for each treatment.

This is also a good opportunity to show how power analysis could instead be done by generating multiple simulated data-sets of size n from the DGP, conducting the statistical test on each data-set and counting how many times the null is rejected, which gives a simulated measure of power. The code is provided in Appendix A.1. The general idea is that each data-set is assigned an additional value/column called 'power_block', with which Stata can conduct the statistical test considering only the data within one power block at a time.

The simulation approach using the code in Appendix A.1 returns a power of 79%, which is very close to the 80% power specified in the `sampsi` command. The small deviation is due to the random nature of the simulations, but this is of no practical concern.

[18]↑The impact of differential variance in this environment is presented in Appendix A.3. In short, more subjects should be allocated to the noisier treatment, with the ratio of subject allocation equal to the ratio of standard deviations (List et al., 2011, pg. 447). However, ex-ante it is not necessarily clear which treatment will have greater variance, something a QR simulation could address.

[19]↑So chosen such that it is sufficiently noisy so that the exercise is not trivial (i.e. excessive power for low n).

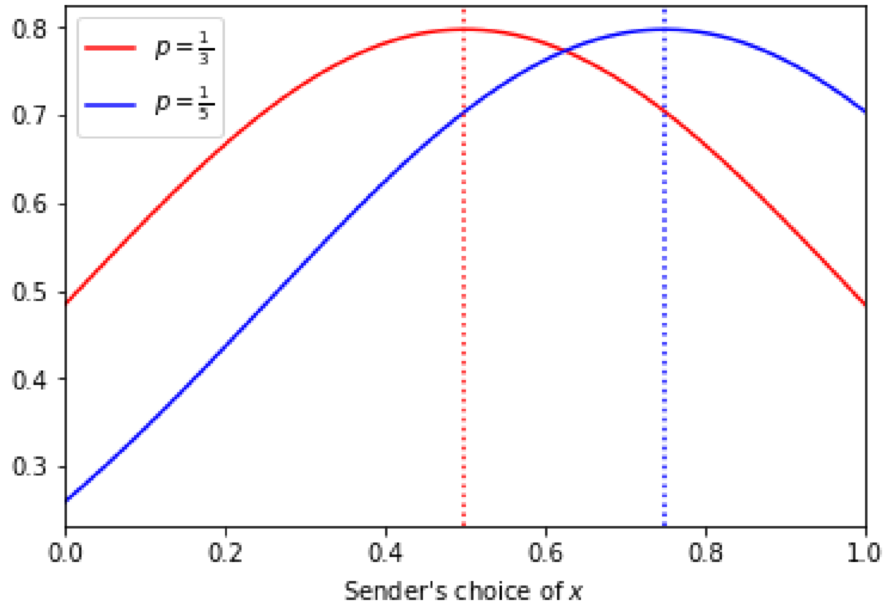


Figure 1.2. Example Normal DGP for Bayesian Persuasion

Optimal Experimental Design

The initial treatment priors of $p = \frac{1}{3}$ and $p = \frac{1}{5}$ were chosen arbitrarily, but they could instead be chosen to maximize power, or equivalently minimize the required number of subjects to obtain a power of 80%. A simple way to do so would be to consider the parameters that go into the rule of thumb for a t-test, $n = \frac{12.35\sigma^2}{\tau^2}$. To lower the required number of subjects, either the treatment effect size should be increased, or the variance should be decreased. Absent an explicit model for variance, it is not clear how much of a change in the initial prior would change the variance. However, theory does provide an estimate on how the treatment effect would change, so in the absence of anything else I change τ.

Following the above, optimal experimental design calls for increasing the distance between the treatment priors as much as possible. For example, for the alternative priors of $p = \frac{1}{10}$ and $p = \frac{2}{5}$, theory predicts Sender behavior of $x^* = 0.89(2dp)$ and $x^* = 0.33(2dp)$. Therefore, the treatment effect size is now $\tau = 0.56$. By the rule of thumb, the required sample size is 13 observations for each treatment. This is substantially less than the 63 observations for each treatment when $\tau = 0.25$.

1.4.3 QR Framework DGP

Motivation

The previous subsection reports a more standard approach for power analysis in a simple one shot experiment, and explores optimal design through the choice of treatment levels. However, ex-ante it is not clear what the standard deviation should be, both in general as well as by treatment. As for the treatment effect itself, the theory should provide an estimate for an appropriate magnitude. Of course, subject behavior may not align with the theory depending on the structure of the payoff hills, as discussed in Section 1.2. But in what direction should the theoretical treatment effect be adjusted, and by how much?

To illustrate the motivation for using the QR framework, consider the expected profits of a Sender where $\pi_R = 2$ and $\pi_B = 1$, given that a Receiver will choose Red if $q \geq p^*$, which is presented in Figure 1.3. In other words, Figure 1.3 represents the Sender's payoff hill over his available actions.

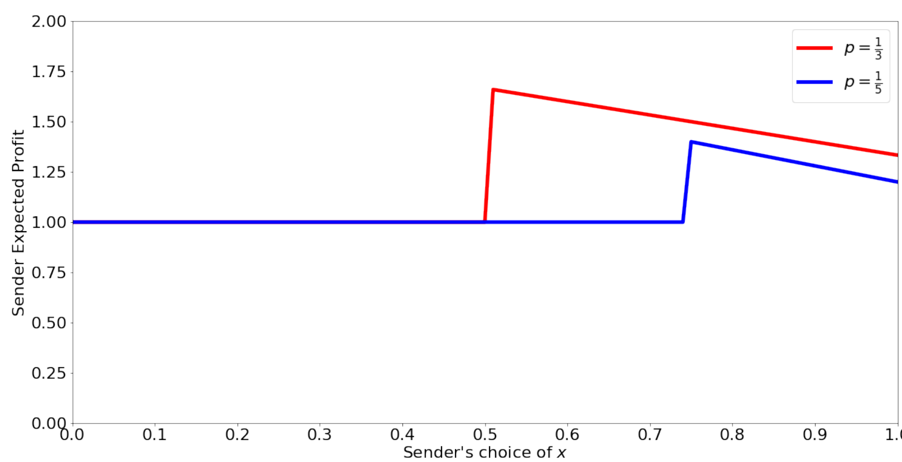


Figure 1.3. Sender's Expected Profit for $\pi_R = 2$, $\pi_B = 1$, and $q^* = 0.5$

Note two main features of the payoff hill in Figure 1.3. The first is that it is always more profitable to choose an x that is higher than the level required to update the Receiver's beliefs above their reservation threshold, than it is to choose an x lower than this level. Therefore, Senders are more likely to choose an x that is higher than the theoretical point prediction, relative to an x that is lower than the theoretical point prediction. In other words, the payoff hills are asymmetric, slightly above the equilibrium value the slope is quite shallow, whereas slightly below the equilibrium value it is very steep. Deviating to a slightly lower x is very costly in terms of expected payoff, while deviating to a slightly higher x is not very costly. Such directional errors provide an adjustment to the likely treatment effect size. In this case the treatment effect would likely differ but not excessively so, as both treatments exhibit likely deviations in the same direction, much like the example payoff hill in Figure 1.1b. Secondly, note that the relative payoff peak is much higher in the red (high-prior) treatment than in the blue (low-prior) treatment. In the blue treatment the overall likelihood of deviating from the optimal action would be greater, as there is less relative incentive to choose that action. Therefore, subject behavior in the blue treatment should exhibit relatively more noise and thus a higher standard deviation than the red treatment. It is well-established that payoff hills should be considered when designing an experiment, as discussed in Section 1.2. The question is, what effect will the payoff hill have on the magnitudes on the treatment effect and treatment specific subject variance? Despite the fact that payoff hills are considered to be a facet of good experimental design, they have currently have no bearing on power analysis except through the researcher's subjective guess.

The QR Approach

Providing some explicit but tractable structure to bring payoff hills into power analysis is the goal of a QR simulation approach. Because the Bayesian Persuasion environment has two players, a QRE needs to be calculated. In this environment, this is not too complicated, as the game is sequential so it can be solved by backwards induction. For each possible belief the state is Red, q, the Receiver could hold after receiving a signal r given x, the expected payoff of choosing Red and Blue is calculated. Then, the logit choice rule given λ calculates the probabilities of choosing Red or Blue for each possible q. Given those Receiver decision probabilities, the Sender's expected profit is calculated for each possible x that could be chosen, taking into account p and the probability of sending r or b messages. From the Sender's expected profits, the logit choice rule is applied to calculate the probabilities of choosing each possible x.

The remaining step is to specify a λ. For the purposes of this example, I select $\lambda_S = \lambda_R = 11.26$ so that a similar number of subjects is required to attain 80% power as the normal DGP ($n = 63$).[20] Given this λ, the QRE implies a DGP of Sender behavior as is graphically presented in Figure 1.4. Something that should be noted is that the information provision x has increased under a QRE relative to standard theory. This is because in the QRE, when the Receiver is indifferent he chooses either action with 50% probability, whereas the standard theory assumes the Receiver would choose Red. Therefore, in a QRE, a Sender optimally provides more information such that the Receiver chooses the Sender's preferred action with a sufficiently high probability given the trade-off implied by the signal structure. This is probably a more realistic assumption about actual human behavior.

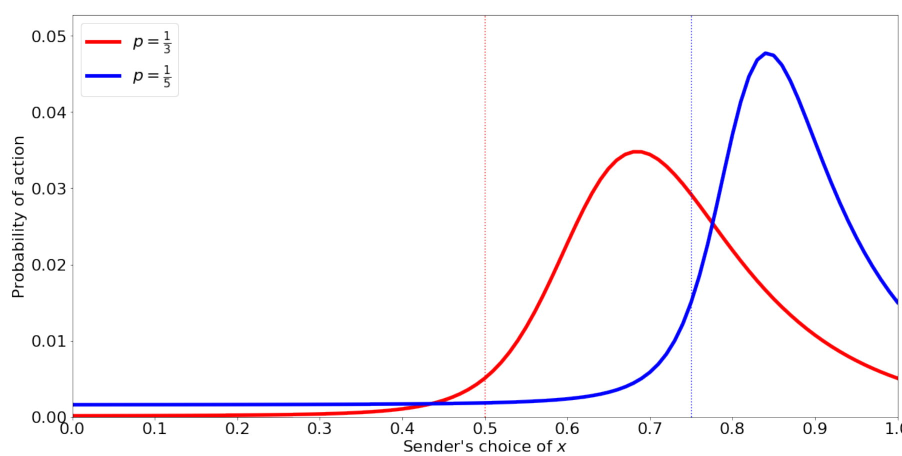


Figure 1.4. Sender's QRE DGP when $\lambda = 11.26$

Optimal Experimental Design Revisited

In the previous section, more optimal treatment levels were selected from the perspective of power by increasing the theoretical treatment effect size. These were treatments with priors of $p = \frac{1}{10}$ and $p = \frac{2}{5}$ with predicted Sender behavior of $x^* = 0.89(2dp)$ and $x^* = 0.33(2dp)$. Using QRE simulations with $\lambda = 11.26$ and $n = 63$, the statistical power is only 7%, i.e. power is substantially reduced at these 'optimal' treatment levels. To obtain a

[20]In practice λ should be estimated from the closest experimental data available.

power of 80%, the design would require $n > 4000$.[21] What went wrong with the previous approach to optimal experimental design? Especially given that it coincides with the common practice of increasing the theoretical treatment effect size. As emphasized in Section 1.3, it is because the likely actual treatment effect size, as well as treatment-specific standard deviations, were not considered. Maximizing the separation of likely behavior needs to be traded-off against the likely standard deviation. As the QRE provides an explicit prediction for both of these parameters for any treatment level, it can be used to help revolve this trade-off. In particular, the proposed $p = \frac{1}{10}$ treatment substantially increases noise, which is why the statistical power decreased by so much. In a Bayesian Persuasion environment, Senders with a low prior face a low likelihood that the Receiver's true state coincides with the Sender's preferred action. Therefore, regardless of the level of information x sent, it is unlikely that the Sender will successfully persuade the Receiver to choose Red. As a result, the Sender's expected payoff hill is quite flat, and thus their behavior is likely to be very noisy, increasing the standard deviation. In a QRE, a perfectly flat payoff hill implies random uniform play. Therefore, flatter hills would move the average x closer to 0.5, which would impact the treatment effect size. Both of these factors contribute to the substantially decreased power of these 'optimal' treatment levels.

I now use QRE simulations to select the optimal treatment levels, i.e. the optimal priors (p and q) for each treatment. I assume $p > q$, restrict $p, q \in [.01, 0.4]$, and only let p and q be integer percentages (as would be likely in an actual experiment). I use a grid search and calculate the power for each possible combination of p and q, and select the pair with the highest power. A 3-D surface of power over the different combinations of p and q is shown in Figure 1.5. The power surface is multi-peaked, with the primary peak having a maximum power of 97.1% at $p = 0.25$ and $q = 0.01$, and the secondary peak having a maximum power of 91.9% at $p = 0.4$ and $q = 0.24$. The primary peak seems counter-intuitive, given the discussion about low-priors being very noisy. However, behavior with a very flat payoff hill in a QRE environment entails near-uniform random play, so the mean would be close to 0.5. As the t-test is a test of means, and as the mean of the QRE distribution for $p = 0.25$ is $x = 0.787$, having a mean of 0.5 is quite powerful for detecting differences (more-so than the

[21] ↑The code starts to hit RAM limitations above $n = 4000$, which yields a power of 75%.

mean at $p = 0.4$, 0.619). However, this conclusion would be erroneous unless the researcher was explicitly testing for a subject's indifference when the payoff hill is very flat. This is an extreme insight from the QRE model, but the design goal is to test the Bayesian Persuasion environment. A more sensible conclusion would arise from the restriction that the mean x in the p treatment is less than the mean x in the q treatment (as the theory predicts if $p > q$), which would only yield the secondary peak, suggesting optimal treatment priors of $p = 0.4$ and $q = 0.24$. Such a design requires only 13 Sender observations for each treatment to reach a power of 80%, substantially less than the 63 in the original design.

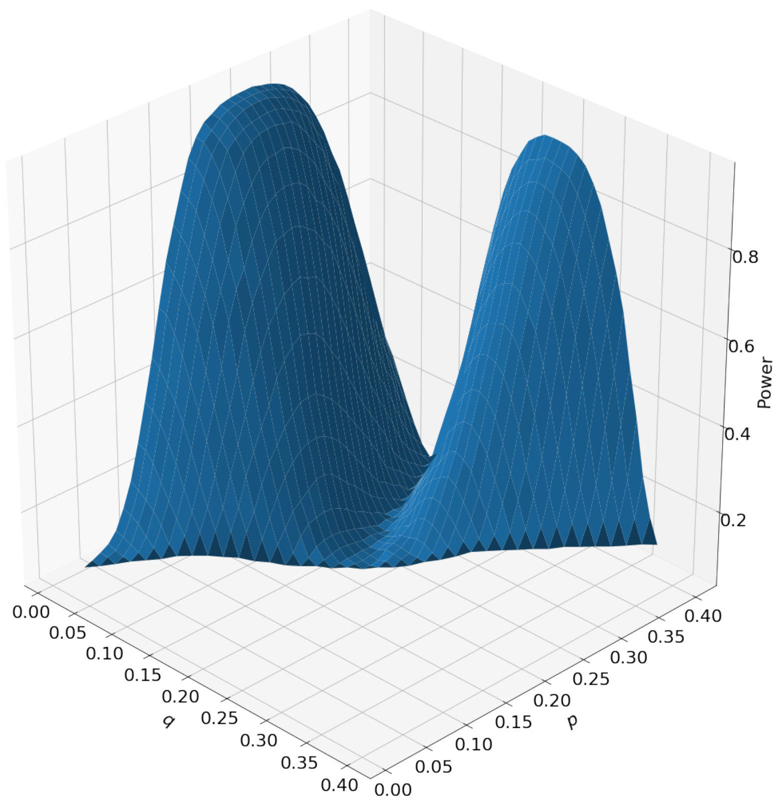


Figure 1.5. Power Surface for QRE Treatment Priors

An advantage of the QR approach is that it provides a structure that maps any given treatment levels into estimates of the likely treatment effect size and standard deviations. In addition, the QR approach can also do this for any treatment invariant parameters that affect expected utility. For example, the payoff that a Sender gets for persuading the Receiver to choose Red, π_R, or the payoff that a Receiver gets for his guess matching his true type, π_m, could also be optimally selected in a similar manner as the treatment level. Increasing π_R would steepen the Sender's payoff hill, and similarly increasing π_m would steepen the Receiver's payoff hill. The former has a direct effect on Sender behavior, as they are now more incentivized to choose their optimal action. The latter has an indirect effect on Sender behavior, as Receivers would now more frequently guess the color with the highest belief, which the Sender would respond to. Increasing these payoffs while holding everything else fixed is equivalent to increasing that particular type's λ. A power curve is presented in Figure 1.6 for increases in π_R and π_m. The resulting curve is as expected, in that increasing π increases the statistical power, and this effect is more pronounced for the direct effect on the Sender.

The power curves suggests that π_m and π_R should be increased as much as possible relative to π_B and π_0. There are practical constraints to this, one caveat might be that subjects need to earn a particular amount even in the worst case scenario (so that they do not get too upset), which would cap how large π could practically be relative to the worst-case payoff.

To summarize, a standard approach to power analysis does not explicitly consider the relative payoff hills that subjects face. This can lead to erroneous assumptions about the direction and magnitude of subject deviations from theory, as well as subject variability by treatment. By incorporating the QR framework into simulation based power analyses, we can more accurately identify likely treatment effects and standard deviations, which can lead to more accurate ex-ante power analysis. Furthermore, we can use the QR framework to guide design decisions, such as selection of optimal treatment levels or other fixed parameters, in order to maximize power.

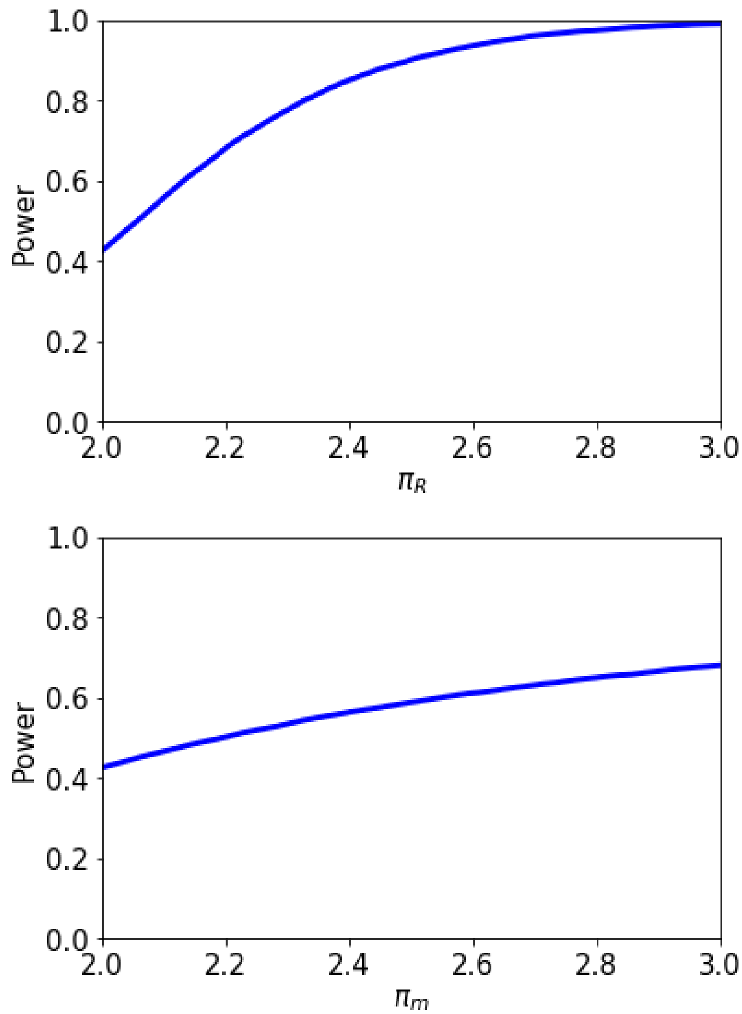


Figure 1.6. Power Curves for π

1.5 Ex-ante Power Analyses of Prominent Papers Using the QR Framework

I demonstrate the application of the QR framework using a thought experiment where I place myself in a similar position as authors of previous papers, and conduct an ex-ante power analysis. In this thought experiment, I take one of the following three approaches. In the first approach, the data-set from the original experiment is ignored and only data from previous related experiments are used. In the second approach a small artificial 'pilot' data-set is sampled from the original experiment. In the third approach, QR parameters are selected where subject behavior is consistent with the spirit of the model in question, which is particularly useful in situations where a structural model proves unidentifiable. All of these situations are similar to the ones a researcher might find themselves in. I only access data that is roughly contemporaneous or previous to the published article and is publicly available.[22] The previous experiments I will use for this exercise come from the ones selected by the Experimental Economics Replication Project (EERP) (Camerer et al., 2016). These are good candidates as they represent a broad range of economics experiments from high-impact journals, and narrows the exercise to the specific 'main' hypotheses selected by the EERP. It should be noted that the QR framework cannot be applied to all of these papers. In particular, there needs to be explicit theory specified that incorporates the channel through which the treatment effect is thought to operate. However, oftentimes small modifications or simplifying assumptions are sufficient to introduce an operative treatment channel when none is explicitly specified. These modifications should not be considered a suggestion of the correct way to model the channel in question, which is far beyond the scope of this exercise. Instead, they are presented as a way to create reasonable predictions of subject behavior ex-ante in environments that would otherwise prove intractable to a simulation approach. The point of the QR framework is to try and improve upon what is currently standard (i.e. subjective extrapolation from previous data, or no ex-ante power analysis at all), rather than provide exact predictions, which is an impossible task ex-ante. I conduct ex-ante power analyses using the QR approach on seven of the papers in the EERP. I present five of

[22]↑I define 'publicly available' data as data made available with the journal article or on another website. It is highly likely that authors would be willing to share other data upon request, but technically this requires data transfer agreements to be in compliance with some Human Ethics Committees.

the more interesting cases in the following sections, and the remaining two in Appendices A.4.1 and A.4.2. In addition, where applicable I provide suggestions on what parameters could be changed to increase power. This demonstrates using QR simulations for optimal experimental design, and is also in the spirit of an efficient conceptual replication, as more optimal parameters could be selected to improve power and reduce subject costs.

1.5.1 Abeler et al. (2011)

Abeler et al. (2011) use a real effort task to test a model of expectations-based reference-dependence (e.g. Kőszegi and Rabin, 2006). In the first stage, they have subjects complete a set number of tasks where they counted the number of zeros in a table. In the second stage, it is announced they can complete as many of these tasks as they want at a piece rate w per task, but that this would only be paid to them 50% of the time. Otherwise, they would be paid a fixed sum f that is independent of the number of tasks they had completed. The treatment is to vary f, to either be 3 euros or 7 euros. According to expectations-based reference-dependence, this should make subjects more likely to stop at the point where they have performed enough tasks to earn f in the eventuality they are paid for their tasks. This is because earning more than f from the tasks introduces psychologically costly losses if f eventuates, and vice-versa for earning less than f from the tasks. A standard model of effort provision would predict no differences in f, rather subjects would perform the task up until the point that their marginal cost of performing that task exceeds half of the piece rate.

Ex-ante Power Analysis

An expectations-based reference-dependent agent in this environment faces the following piece-wise utility function:

$$U(\mathrm{e}; w, f) = \begin{cases} \frac{w\mathrm{e}+f}{2} - c(\mathrm{e}) + \frac{1}{2}\eta[\frac{1}{2}\lambda(w\mathrm{e} - f)] + \frac{1}{2}\eta[\frac{1}{2}(f - w\mathrm{e})], & \text{if } w\mathrm{e} < f \\ \frac{w\mathrm{e}+f}{2} - c(\mathrm{e}) + \frac{1}{2}\eta[\frac{1}{2}(w\mathrm{e} - f)] + \frac{1}{2}\eta[\frac{1}{2}\lambda(f - w\mathrm{e})], & \text{otherwise} \end{cases},$$

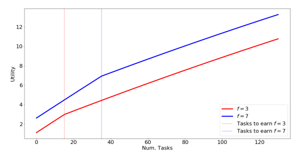
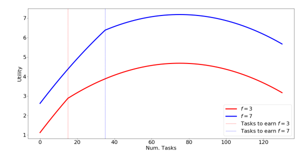
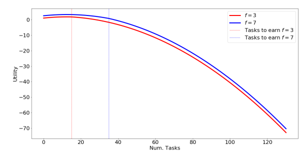
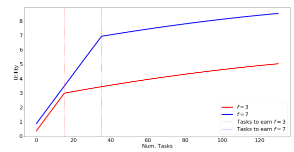
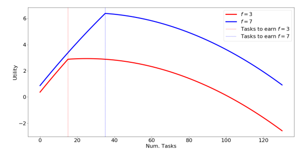
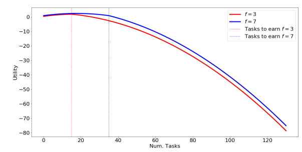
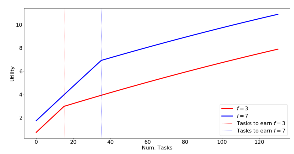
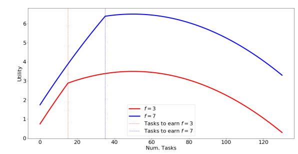
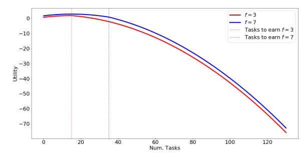
where e is the number of tasks completed, $\eta \geq 0$ is the sensitivity to reference-dependence, $\lambda > 1$ is loss aversion (not to be confused with λ_{QR}), and $c(\mathrm{e})$ is the cost of performing e tasks.

There are two implementation issues for this environment. Firstly, the variables η and λ are not separately identifiable, as multiple combinations of the two could explain the same decision. Secondly, this particular environment was relatively novel at the time, and there does not appear to be a wealth of previous experimental evidence with which the authors could draw upon. For example, data from a similar real-effort task experiment might make it possible to obtain a reasonable estimate of $c(\mathrm{e})$. Given these limitations, I specify reasonable values for the parameters given the goal of testing a model of reference-dependence. For $c(\mathrm{e})$ I assume $c(\mathrm{e}; a) = a\mathrm{e}^2$, which encapsulates increasing costs in a simple one-parameter model. I consider what the payoff hills will look like for some various combinations of η, λ, and a, presented in Table 1.2.

Table 1.2 highlights that it is important that the costs do not outweigh the reference-dependent and loss averse parts of the utility function.[23] If costs are too low, then subjects will want to complete as many tasks as they can, and reference-dependence will only have a small effect. If costs are too high, then subjects will not want to complete many tasks at all, and again, reference-dependence will have little effect. An intermediate cost parameter can still overpower the effect of reference-dependence depending on the parameters, as the center column of Table 1.2 indicates. Two of the three specifications for $a = 0.001$ result in

[23]↑It is not possible to exactly control the cost of the real effort task, however, it is possible to indirectly influence it, by making the task more or less difficult or unpleasant.

Table 1.2. Payoff Hills for Various Parameters in Abeler et al. (2011)

	$a = 0.0001$	$a = 0.001$	$a = 0.01$
$\eta = 0.5$ $\lambda = 2$			
$\eta = 1.5$ $\lambda = 2$			
$\eta = 0.5$ $\lambda = 3$			

the same effort point prediction for both treatments, whereas the specification in the middle has a point prediction consistent with reference-dependence.

All of the given combinations would predict some treatment effect in f, but given the focus on reference-dependence, I select $\eta = 1.5$, $\lambda = 2$ and $a = 0.001$ (i.e. the center graph of Table 1.2). The parameter of λ_{QR} remains to be specified, which is estimated from a small artificial pilot of 10 subjects per treatment.[24] This estimation yields an estimate of $\lambda_{QRE} = 0.741$, which then implies $\tau = 10.297$, $\sigma_{f=3} = 25.039$, and $\sigma_{f=7} = 22.216$. These values are based on the number of tasks completed, which are paid at a piece rate of 0.20 euros. The comparison for τ from the point prediction is therefore $\frac{7-3}{0.20} = 20$ tasks, meaning the QR predicts a treatment effect size of approximately half of that implied by the point prediction. According to a t-test (which is equivalent to an OLS regression with treatment as a dummy variable utilized in the actual paper), with these parameters 83 subjects per treatment are required to reach a power of 80% at the 5% level.[25]

[24]↑Note, it would not possible to fit the other parameters using this pilot data-set, as η, λ, and $c(\mathrm{e})$ are not separately identifiable, especially on such a small data-set.
[25]↑For the remainder of this section the required number of subjects refers to the number of subjects to obtain 80% power at the 5% level unless otherwise specified.

Table 1.3. QR Simulations for Different Combinations of f in Abeler et al. (2011)

	$f_{high} = 6.60$	$f_{high} = 7.00$	$f_{high} = 7.40$
$f_{low} = 2.60$	$\tau = 9.930$ $\sigma_{low} = 25.296$ $\sigma_{high} = 22.502$ $Power = 75.9\%$	$\tau = 11.088$ $\sigma_{low} = 25.296$ $\sigma_{high} = 22.216$ $Power = 86.5\%$	$\tau = 12.268$ $\sigma_{low} = 25.296$ $\sigma_{high} = 21.934$ $Power = 91.4\%$
$f_{low} = 3.00$	$\tau = 9.139$ $\sigma_{low} = 25.039$ $\sigma_{high} = 22.502$ $Power = 70.4\%$	$\tau = 10.297$ $\sigma_{low} = 25.039$ $\sigma_{high} = 22.216$ $Power = 80.2\%$	$\tau = 11.478$ $\sigma_{low} = 25.039$ $\sigma_{high} = 21.934$ $Power = 88.8\%$
$f_{low} = 3.40$	$\tau = 8.293$ $\sigma_{low} = 24.769$ $\sigma_{high} = 22.502$ $Power = 62.2\%$	$\tau = 9.451$ $\sigma_{low} = 24.769$ $\sigma_{high} = 22.216$ $Power = 72.8\%$	$\tau = 10.632$ $\sigma_{low} = 24.769$ $\sigma_{high} = 21.934$ $Power = 84.2\%$

Optimal Experimental Design

The most obvious candidate to increase power would be to increase the distance between the f's in each treatment, assuming standard deviations do not substantially rise when doing so. However, there are practical constraints on f, for example, if f was quite high, it might induce loss-averse subjects to stay in the lab doing the task for excessive periods of time. In addition, the values of f should be 'clean' to present to subjects (e.g. $f = 2.79$ is not great, while $f = 2.80$ would be fine), and in multiples of w to align with the reference-dependent predictions (a preference to match $f = we$). There may also be a design goal of keeping each treatment's reference-dependent prediction sufficiently different from the standard prediction (here, given w and a, $e^* = 50$). I therefore implement a restricted grid search around f's that were (presumably) carefully selected by the original authors. In particular, the grid search is over combinations of $f_{low} \in \{2.60, 3.00, 3.40\}$ and $f_{high} \in \{6.60, 7.00, 7.40\}$, which is presented in Table 1.3.

Table 1.3 confirms that increasing the distance between the f's increases power. This is despite the fact that decreasing f_{low} increases the standard deviation of that treatment, as the increase in treatment effect size is clearly offsetting this. The overall optimal configuration is $f_{low} = 2.60$ and $f_{high} = 7.40$, which would require 59 observations per treatment, less than the 83 observations required for the original treatment parameters.

1.5.2 Charness and Dufwenberg (2011)

Charness and Dufwenberg (2011) experimentally investigate the impact of communication through free-form messages could have on behavior in a hidden information principal/agent game, as displayed in Figure 1.7. Player A is the principal, and player B is the agent. The agent has a type, either low or high, which determines their suitability for a difficult but high paying task. If a low agent attempts the difficult task, they always fail, but if a high agent attempts the difficult task, they only fail with $p = \frac{1}{6}$. The agent's type is known to them when making their decision, and is never fully revealed to the principal if the type is low. Given this arrangement, the principal must decide whether they want to hire the agent and let them decide what difficulty task to undertake, or collect some outside option instead. The particular treatment effect is whether allowing B's to send a free-form message to their paired A discourages low B's from choosing Roll (or in other words, undertaking the unsuitable high difficulty task).

Ex-ante Power Analysis

The effect of communication is often not explicitly incorporated into theory, but Charness and Dufwenberg (2011) suggest two possible channels, a cost of lying and guilt from blame. I incorporate a simple cost of lying model supplemented by other-regarding preferences, as this is independent of the complications of first- and second-order beliefs present in guilt models. Both types of players have the monetary payoffs of their counterpart enter into their utility function, weighted relative to their own payoff by the parameter α. In other words, A's utility function is $U_A(x_A, x_B) = (1 - \alpha)x_A + \alpha x_B$, where x_A and x_B are the monetary payoffs of players A and B respectively. I assume that when given the opportunity, B sends a message indicating that they will choose Don't when they are the low type, and incurs some disutility k if they do not follow through on this, to reflect a cost of lying. I denote the two treatments as M (Messages) and NM (No Messages), with $k_M > k_{NM} = 0$. So B's

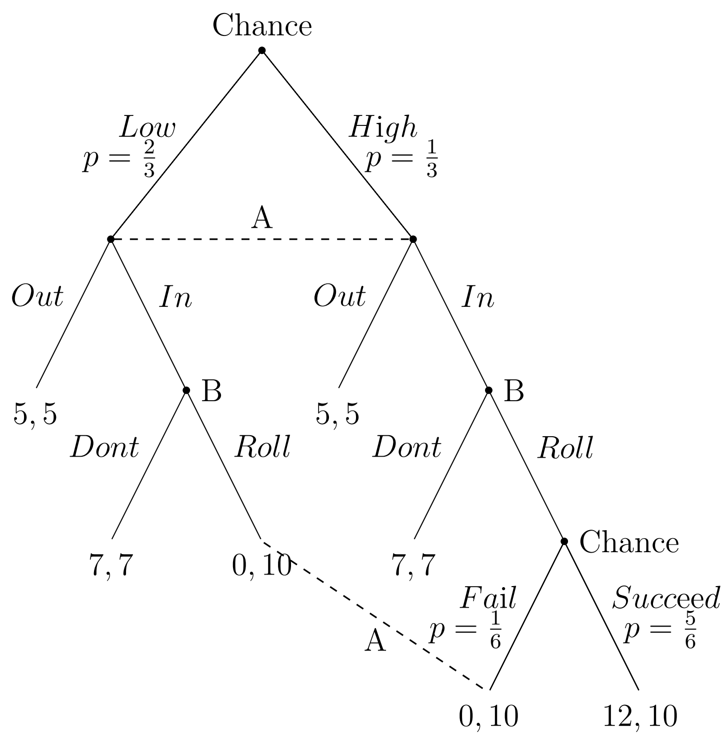


Figure 1.7. Charness and Dufwenberg (2011)'s '(5,7) Game'

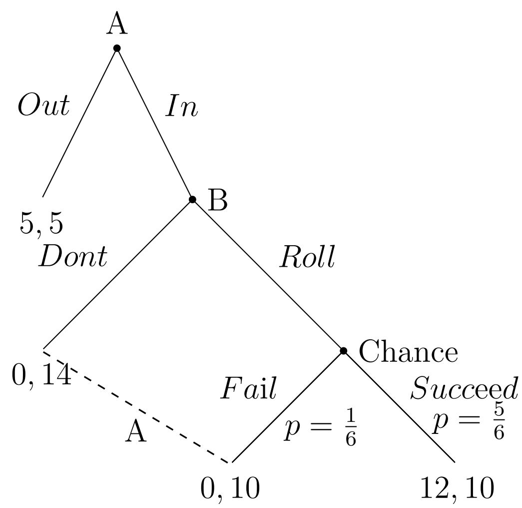


Figure 1.8. Game in Charness and Dufwenberg (2006)

utility function if they have told a lie is: $U_B(x_A, x_B) = (1 - \alpha)x_B + \alpha x_A - k_M$, and if they have not told a lie: $U_B(x_A, x_B) = (1 - \alpha)x_B + \alpha x_A$. For any given α, k_M, and λ, the QRE provides predictions for the probability of a low B choosing Roll, a high B choosing Roll, and A choosing In.

The next step is that the values of α, k_M, and λ need to be specified. The best approach would be to use a closely related study. Fortunately, the game considered in Charness and Dufwenberg (2006) (presented in Figure 1.8) is a good candidate for this purpose. It has similar forms of communication, subject pools, payoff magnitudes, and a hidden action. I assume other-regarding preferences in the same manner as before, and that B faces the same disutility from lying, except that in this game he now promises to always Roll, and thus faces disutility when choosing Don't. I specify a log likelihood function that depends on the data observed in Charness and Dufwenberg (2006) and the parameters α, k_M, and λ. Using this function, I use maximum likelihood estimation techniques to fit values of the parameters and obtain $\alpha = 0.18$, $k_M = 2.15$, and $\lambda = 4.41$.

I apply the estimated parameters to the QRE of the current game, which provides a prediction for the probabilities for each action in each treatment, summarized in Table 1.4. With these probabilities, I simulate a data-set of size n, where n is the candidate total subjects needed.[26] It should be noted that the current experimental design calls for a direct response method, so power needs to be considered not only in terms of the treatment effect that is likely to eventuate, but the number of subjects that will actually undertake the relevant decision (i.e. low type B). In order for that to occur, the B's type must be assigned as low by nature, but it must also be the case that the type A player chooses In. Therefore, it is important that A's differential behavior between treatments is modeled given B's cost of lying, despite the fact that I am not directly testing any hypotheses on A's behavior. Due to the direct response method, I simulate each A/B pair sequentially as the game tree indicates, i.e. I draw B's type, A makes their decision, and if that decision is In, B then makes their decision given their type, which is then added to the data-set. I generate 10,000 'power blocks' of data-sets with n total subjects, so $n/2$ subjects per treatment, or $n/4$ pairs of A and B per treatment. On each simulated data-set, I conduct the statistical test on low B decisions, which in this case is a two-proportion z-test, and count how many times the null hypothesis is rejected. I then adjust n in multiples of 4, until the lowest n that has a power of greater than 80% is obtained. This process suggests a total number of subjects of 388, or 97 pairs of A and B per treatment.

Table 1.4. Probability of Actions by Treatment

	Messages	No Messages
Prob. A In	0.74	0.60
Prob. Low B Don't	0.64	0.34
Prob. High B Don't	0.21	0.21

Optimal Experimental Design

I now consider if there are any parameters that could be changed that would increase power. A related question of the optimal treatment intensity is not relevant in this environment, as it is not immediately obvious how the intensity of communication could be varied, and the impact that would have.[27] I instead turn to the other parameters that are fixed in both treatments. I consider that practical constraints exist in the form that numbers

[26] ↑I assume that each treatment has equal sample sizes, but this assumption could be relaxed.
[27] ↑Perhaps if the message is only sent with a certain probability?

should be easily communicable to subjects (e.g. $p = \frac{1}{6}$ is fine, but $p = .17258013$ is not), and that payoffs should not move too far (e.g. increasing one payoff to \$100 may increase power substantially, but would not be affordable). The theory and motivation of the paper also indicates that some parameters should also change together. For example, a low B choosing Roll should have the same payoffs as a high B choosing Roll but failing, or otherwise Roll is not a hidden action.

With that in mind, I consider either raising or lowering each parameter to the nearest 'clean number' (i.e. integer for payoffs, nearest tenth for probabilities). Table 1.5 presents the power for raising or lowering one parameter in isolation, holding the other parameters fixed at the levels that were used in the experiment. Table 1.5 also presents the results of a grid search over every possible combination of the given parameters. The largest increases in power come when lowering the payoff of the outside option or lowering B's payoff in the event that a Roll fails. Both of these factors encourage A to more frequently choose In, which increases power as it increases the number of low B's that actually get to make a decision.[28] Changing only the most effective parameter in isolation, the outside option, would reduce the total number of required subjects to 320 to reach a power of 80%. The grid search set of parameters would require even fewer subjects again, for a total of 256.

Table 1.5. Power for various parameters in Charness and Dufwenberg (2011)

Parameter	Lower	Baseline	Higher
Prob. low $\in \{0.6, \frac{2}{3}, 0.7\}$	78.8%	80.3%	80.5%
Out payoff $\in \{4, 5, 6\}$	87.6%	"	66.6%
Don't payoff $\in \{6, 7, 8\}$	60.4%	"	79.9%
A fail payoff $\in \{0, 1\}$	N/A	"	80.3%
B fail payoff $\in \{9, 10, 11\}$	86.1%	"	63.0%
A succeed payoff $\in \{11, 12, 13\}$	78.4%	"	81.8%
Prob. Succeed $\in \{0.8, \frac{5}{6}, 0.9\}$	79.4%	"	81.8%
Grid Search	0.7,4,6,0,9,13,0.9		93.7%

[28]↑This illustrates the advantage of using the strategy method, as a B decision for the case they are low would be observed for every B, regardless of A's decision and the random draw. However, it may be the case that using the strategy method could change behavior relative to the direct response method (Brandts & Charness, 2000; Casari & Cason, 2009).

1.5.3 R. Chen and Chen (2011)

R. Chen and Chen (2011) experimentally test the role of group identity on coordination in a two-player minimum effort game. In particular, they induce group identity in an 'enhanced' treatment by not just assigning a group, but reinforcing that identity by having the group help answer a paid question about identifying paintings. They posit that group identity can increase the concern that others have about the payoffs of those in their 'in-group'. The increase in other-regarding preferences for in-group people can change the cost threshold in which a QRE would predict low or high levels of effort in a minimum effort game.

R. Chen and Chen (2011) provide the following explicit utility function for a two-player minimum effort game:

$$U_{\mathrm{i}}(x_{\mathrm{i}}, x_{\mathrm{j}}; \alpha_{\mathrm{j}}) = a \min x_{\mathrm{i}}, x_{\mathrm{j}} - c[(1 - \alpha_{\mathrm{j}})x_{\mathrm{i}} + \alpha_{\mathrm{j}} x_{\mathrm{j}}],$$

where a is the benefit of an additional unit of minimum effort, c is the (non-refundable) cost of effort, and α_{j} is the relative weighting that i placed on j's payoff, which can differ if the other is in the out-group O, or the in-group I (α_O, α_I).

Ex-ante Power Analysis

R. Chen and Chen (2011) use the $a = 1$, $c = 0.75$ environment from Goeree and Holt (2005), where subjects can choose efforts of $x \in [110, 170]$ in steps of 0.01 (i.e. they can choose from 110, 110.01, 110.02, ..., 169.98, 169.99, 170). They then calculate the QRE for different levels of α given the $\lambda = 0.125$ reported in Goeree and Holt (2005). This is what the QRE approach would have done had it not already been conducted, as Goeree and Holt (2005) is clearly the most closely related environment from which an estimate of λ would be obtained from. The only thing that remains is to specify α_O and α_I, as the hypothesis is whether effort levels differ substantially if the counterpart is in the same group, or in a different group. This would usually be difficult, as there are many different ways of inducing group identities, and this would likely be sensitive to the method used as well as the subject pool.[29] Fortunately, Y. Chen and Li (2009) conduct a near-identical task to induce group identity on a similar subject pool. They obtain estimates of α_I and α_O from tasks relating to social preferences and simple reciprocity games in the style of Charness and Rabin (2002). I take the parameters $\alpha_I = 0.474$ and $\alpha_O = 0.323$ from the Y. Chen and Li (2009) paper.

The statistical test used in this experiment is a panel data regression with clustering at the session level and individual random effects. If the statistical test controls for such things, then the generated data-set should justify the use of these controls, in order to obtain a reasonable estimate of power. To justify session level clustering, I draw a $\lambda \sim U(0.1, 0.15)$ for each session.[30] As for the individual random effect, which is modeled as a normal draw from $N(0, \sigma_\mathrm{e})$, I run a panel regression on the $c = 0.75$ data-set from Goeree and Holt (2005), and obtain an estimate of $\sigma_\mathrm{e} = 15$. In the simulation, in each period the subject's action is drawn from the session level QRE distribution, and then the individual's random effect is applied. After this, their effort decision is rounded to the nearest possible valid action.

This particular experimental design calls for exact session sizes of 12 subjects, and so the minimum 'step' for power consideration is 12, of which I will refer to as a session. Table 1.6

[29] ↑For example, see the replication report in Camerer et al. (2016) for this paper. The same group identity task seems to have not induced much group identity in a different subject pool, which may be because subjects did not interact as much with the chat task.

[30] ↑Although, drawing a session-level α would also be a valid approach.

reports that 6 sessions per treatment is required for a power of 80%. This means that 120 subjects would be required in total.

Table 1.6. Ex-ante Power Analysis for R. Chen and Chen (2011)

Sessions Per Treatment	Power
2	57%
3	65%
4	74%
5	78%
6	81%

Optimal Experimental Design

The first thing that should be noted is that increasing α_I or decreasing α_O would increase power, however, it is not clear how specific changes in the group identity induction would change the α parameters. So, if there is a better way to induce stronger group identity then it should be implemented, but that this is not something the QR framework can address without some estimate of the induced α. The remaining parameters that could vary are a and c. I focus on the cost c, as $a = 1$ is convenient to explain to subjects. I consider only small changes in c, to $c = 0.7$ and $c = 0.8$, as $c = 0.75$ was originally chosen as it tended to result in low effort in experiments without group identity, and the authors posited that that the introduction of group identity might overcome this. Another design consideration is the set of permissible effort levels that can be chosen. I consider the range of actions, in that I expand and reduce the action space by 10 effort points. Additionally I consider making the action space coarser, by considering integer steps and steps of 10, instead of the original step size of 0.01. In this exercise, I hold all other parameters at the same original level. The output of this exercise is summarized in Table 1.7.

What Table 1.7 reveals is that increasing c increases power, as it decreases effort in the outgroup treatment by more than the ingroup treatment.[31] It also reveals that increasing the range of permissible actions also increases power, which is because the higher efforts will

[31]↑This effect would not be monotonic though. For example, increasing $c \geq 1$ would result in a decreased treatment effect size, as both treatments would tend to choose the lowest possible effort.

Table 1.7. Ex-ante Experimental Design for R. Chen and Chen (2011).

	Baseline	$c = 0.7$	$c = 0.8$	$x \in [115, 165]$	$x \in [105, 175]$	x step= 1	x step= 10
τ	6.61	5.12	7.29	4.03	9.94	6.04	4.06
σ_O	18.32	18.16	18.23	16.18	20.28	18.48	19.62
σ_I	17.74	17.56	17.97	15.90	19.22	18.04	19.66
Power at 2	57%	52%	61%	38%	76%	51%	38%
Power at 3	65%	60%	75%	37%	86%	62%	39%
Power at 4	74%	63%	75%	48%	92%	66%	49%
Power at 5	78%	70%	88%	52%	97%	72%	54%

τ is the empirical average treatment effect size by power block for the 5 session data-set.
σ_O is the empirical average standard deviation of the outgroup's effort levels by power block for the 5 session data-set.
σ_I is the empirical average standard deviation of the ingroup's effort levels by power block for the 5 session data-set.
'Power at X' is the simulated power for the X sessions per treatment data-set.

be more likely to be played by in-group players (and vice versa), which in the QRE can have a reinforcing effect. The increase in the treatment effect dominates the increase in standard deviation brought about by the additional noise due to having more available actions. Finally, increasing the coarseness of the strategy space decreases power, so this should be avoided. Figure 1.9 illustrates the effect of the action space coarseness, where increasing the size of the step causes substantially more low effort decisions for both treatments, compressing the treatment effect. Of the possible design changes, increasing the range of actions would be the most beneficial for power. Utilizing the increased strategy space design would require only 3 sessions per treatment to attain 80% power, less than the 6 sessions per treatment required under the original parameters.

1.5.4 Fehr et al. (2013)

Fehr et al. (2013) consider the effects of delegation in a principal agent game where there are a number of projects, one of which must be selected to be implemented. There is a known 'outside option' project, while the states of the all of other projects are unknown, but can be revealed through costly effort. The principal is initially endowed with the decision right on what project to implement. If the principal retains this decision right, the agent can provide a recommendation for which project to do, while the principal makes his decision based on his own knowledge and/or the recommendation. If the principal delegates this decision right, the roles are reversed, the principal can provide a recommendation but it is the agent who

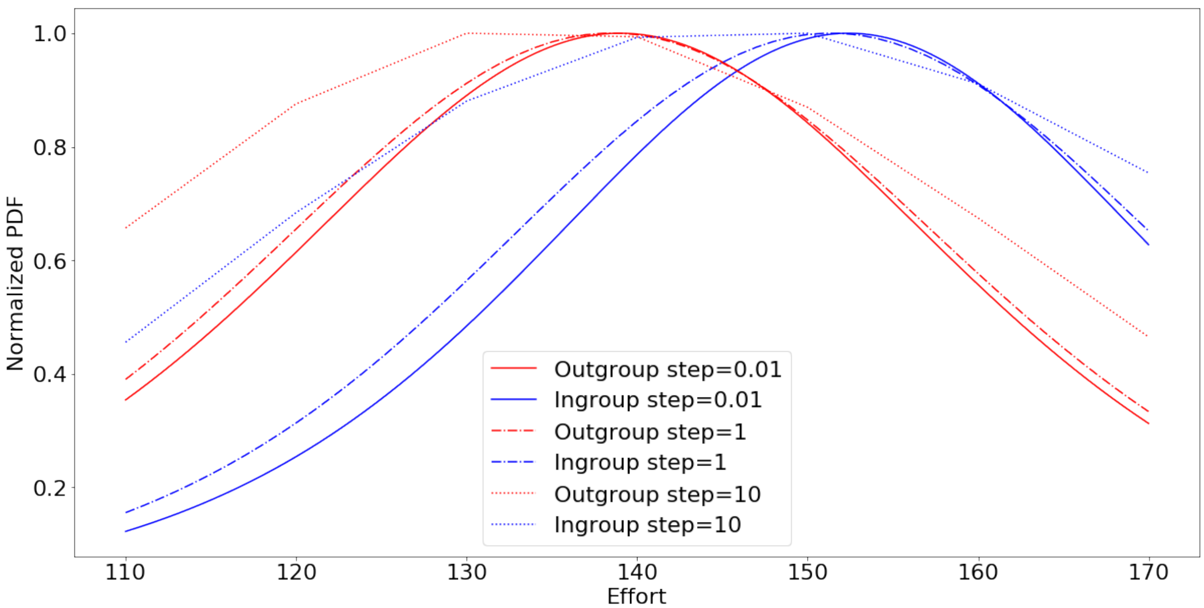


Figure 1.9. Effect of Increasing Step Size in the Minimum Effort Game

makes the final decision based on the agent's knowledge and/or the recommendation. The game starts with the principal deciding whether to delegate or not, which is followed by simultaneous effort decisions by both players. The effort level chosen by an individual is the probability that the true state of all projects will be revealed to that individual. The final stage is the controlling party deciding which project to implement, given their own information (if the true state was revealed to them), and the recommendation from the other party. In this particular setup, there are 36 projects, 1 of which is the outside option which pays out 10 to each subject. Of the remaining 35, 33 of the projects pay out 0 to each subject. Of the remaining 2 non-zero projects, one of these projects is preferred by the principal, and the other is preferred by the agent (although both of these projects are preferred to the known outside option). The cost of effort function is $c(\mathrm{e_i}) = 25\mathrm{e_i}^2$, where $\mathrm{e_i} \in [0, 1]$, and the true state is revealed to i with probability $p(\mathrm{e_i}) = \mathrm{e_i}$. Fehr et al. (2013) consider a range of payoff values for projects, such that the principal is predicted to either want to delegate or not. However, they find that delegation is substantially lower than predicted.

Ex-ante Power Analysis

Given the low delegation rates from the initial experiments, Fehr et al. (2013) conduct additional treatments to test whether the low delegation rates were driven by a non-pecuniary cost of being overruled. They propose a treatment where whichever party that is not in control cannot exert any effort, and therefore cannot provide a recommendation that could be overruled. This is the treatment effect that the EERP considers, that the delegation rates in this treatment are higher than in an environment with similar Nash equilibrium returns to delegation but where there is the possibility of being overruled. Because this experiment was designed after observing the initial results, I can place myself in the exact position of the authors, by restricting myself to only consider their baseline set of data.

I fit a QRE to the baseline data with one modification to the model, a disutility D of being overruled. I assume that a subject is overruled when they do not have control over the final decision, but that they have full information (i.e. the true state is revealed to them) and they provide a recommendation which is not then followed. This changes effort levels in equilibrium, as well as the delegation decision. Fitting this model yields $D = 41.22$ and $\lambda = 0.216$. The disutility parameter is substantial, and implies that being overruled is disliked so much as to more than wipe out even the highest pecuniary incentive. I then apply these parameters to a QRE of the two additional treatments, and conduct a power analysis on the generated data-sets. Because the conducted statistical test incorporates a learning trend, I increment λ linearly by period, such that the estimated $\lambda = 0.216$ is obtained halfway into the experiment. As a higher λ is associated with less noisy play, this is an appropriate way to model a learning effect. In addition, the statistical test clusters at the individual level. I justify clustering at this level by having an individual specific disutility of being overruled drawn from $N(41.22, 10)$. The results are presented in Table 1.8, and suggest that 5 groups per treatment (each group had 10 subjects in it, so a total number of 100 subjects), would be sufficient to reach 80% power.

Table 1.8. Fehr et al. (2013) Power Analysis

Num. Groups per Treatment	Power
$N = 3$	64.4%
$N = 4$	77.7%
$N = 5$	84.7%
$N = 6$	88.0%

Optimal Experimental Design

The two additional treatments were carefully and explicitly designed so that the Nash equilibrium returns to delegation were very similar. Because the two environments are procedurally different, the payoffs for different outcomes are necessarily different between the two treatments. Changing these payoffs would mechanically induce a difference in delegation rates quite trivially (due to the equilibrium returns to delegation), but it would not be due to the mechanism that is proposed. This particular hypothesis is not a good application of the QRE approach to optimal experimental design. Therefore, I do not suggest alternative parameters to increase power for this paper.

1.5.5 Ifcher and Zarghamee (2011)

Ifcher and Zarghamee (2011) conduct an experiment where time preferences are elicited directly after viewing a video clip designed to induce a positive mood. As compared to a neutral video clip, they find that subjects exhibit more patience (i.e. discount the future less) when in the induced positive mood.

Ex-ante Power Analysis

The challenge in this environment is to specify an explicit functional form with which positive affect can operate on time preferences. A brief foray into the experimental economics literature that induces positive affect in a similar way yields some papers with other-regarding preferences and strategic environments (Capra, 2004; Drouvelis & Grosskopf, 2016; Kirchsteiger et al., 2006). I relate other-regarding preferences with time preferences by assuming people see the present-self and the future-self as being different entities, as posited by research in psychology (Pronin & Ross, 2006). I therefore assume that actions towards the future-self are similar to actions towards other people in the present. In particular, I assume that subjects exhibit quasi-hyperbolic discounting, $D(y,t) = y\beta_{\mathrm{i}}\mathrm{e}^{-rt}$ if $t > 0$, where the present bias β_{i} can differ depending on what mood the subject is in.

Ifcher and Zarghamee (2011) elicit time preferences by asking for what value of $\$p$ would make a subject indifferent between receiving $\$p$ today and receiving $\$m$ in t days time. They consider 30 combinations of m and t, and incentivize the reporting of the indifference level of p using a discrete BDM mechanism. In this mechanism, balls labeled $b \in \{1, 2, ..., m+1\}$ are placed into an urn and then a ball is drawn at random.[32] If the ball drawn b is less than or equal to p, then the subject gets paid m in t days time, whereas if $b > p$, the subject gets paid $\$b$ today. The discretization of the BDM mechanism means that ranges of p result in the same payoff outcome, so I discretize the subject's p decisions as the mid-points of these ranges.[33]

A good place to obtain a reasonable estimates for β_N (neutral mood), r and λ would be from Benhabib et al. (2010), who elicit time preferences in a very similar manner, just without any mood induction. However, this data-set is not currently publicly available, and would not give an estimate for β_G (induced good mood) anyway. Instead, I assume a small initial pilot is conducted, and fit the model to that. I simulate such a pilot by taking 10 subjects per treatment from the full data-set. For the artificial pilot, the estimated parameters are $\beta_G = 0.814$, $\beta_N = 0.701$, $r = 0.002$, and $\lambda = 5.633$. Finally, the statistical test that is

[32]↑In the case where $m+1$ is not an integer, the final ball is the next highest integer above m.

[33]↑For the unbounded above upper range, I add the difference between the midpoint and lower-bound of the other ranges to the lower-bound of this range.

conducted is a regression clustered at the individual level. In order to justify individual level clustering, I assume that the induced mood has differential effects on each subject, such that $\beta_\text{i} \sim U[\beta - 0.025, \beta + 0.025]$. Table 1.9 presents the results of this QRE simulation power analysis. The required number of subjects per treatment to reach 80% power is 34 subjects per treatment, so 68 subjects in total.

Table 1.9. Ifcher and Zarghamee (2011) Power Analysis

Subjects per Treatment	Power
$N = 30$	74.1%
$N = 31$	74.5%
$N = 32$	77.0%
$N = 33$	78.2%
$N = 34$	80.1%
$N = 35$	81.8%

Optimal Experimental Design

The obvious initial change here would be to induce a bad mood rather than a neutral one, which has been shown to change other-regarding preferences in strategic settings (Drouvelis & Grosskopf, 2016). I assume a similar decrease from β_N to β_B as the increase from β_N to β_G, and set $\beta_B = 0.60$. This increases power substantially, for example at $N = 34$ power increases from 80.1% to 100%. Such a design change would then require only $N = 8$ subjects per treatment to reach a power of 80%. Another design change could be to reduce the discretization of the BDM mechanism to provide a richer action space with which subjects could exhibit their true preferences. I specifically consider adding additional balls such that $b \in \{0.1, 0.2, 0.3, ...m + 0.9, m + 1\}$, substantially decreasing the coarseness of the effective strategy space in p. However, the effect on power is small, driven by small decreases in the standard deviation due to the finer space. The only minor improvement of power justifies the original experimental design, as it may be easier to explain the BDM mechanism with discrete integer ball drawing.

1.5.6 Summary

Table 1.10 presents a summary of the total number of required subjects by the ex-ante QR simulation approach using the same parameters that were selected in the experiment, as well as if 'optimal' parameters in terms of power were selected. It also presents the total number of subjects that were actually used in the original experiments, and an ex-post approach that uses the same power analysis method as the EERP (except for 80% power). Table 1.10 shows that the QR simulation approach is typically more conservative than the actual experiments. This is not necessarily surprising given that these papers are from a time when power analysis was less common. It is not exactly clear from the listed papers how these subject numbers were determined. The QR framework also compares somewhat favorably to the required number of subjects from an ex-post approach, where all data are known. This is a very high bar to clear, so having the QR framework be anywhere near those numbers is promising, and suggests the QR framework can be a reasonable way to improve ex-ante power analysis. Finally, in some environments there exists room to increase power through changes in the experimental design, as indicated by the lower subject numbers required when optimal parameters are selected.

Table 1.10. Required Total Number of Subjects for Various Approaches

Paper	Ex-ante QR (Original Params.)	Ex-ante QR (Optimal Params.)	Ex-ante Original	Ex-post* (Camerer et al., 2016)
Abeler et al. (2011)	166	118	120	237
Charness and Dufwenberg (2011)	320	256	162	192
R. Chen and Chen (2011)	120	72	72	125
De Clippel et al. (2014)	200	40	158	115
Fehr et al. (2013)	100	N/A	60	73
Huck et al. (2011)	60	N/A	120	114
Ifcher and Zarghamee (2011)	68	16	58	98

*Using the z-approximation formula for 80% power, rather than the 90% used in the EERP.

1.6 Conclusion and Future Research

Power analysis is an important, but until recently under-considered, factor in experimental economics research. Power analysis should be the sole determinant of the number of subjects to be used in an experimental study. However, the correct way to conduct a power analysis is only clear when enough data already exists in the same treatments that are to be conducted, a situation a researcher only finds themselves in when conducting a replication. Instead, the researcher looks to the most closely related previous experiments, and must extrapolate from there what behavior might look like in the novel environments that they are considering. The question is how this extrapolation is done, as there is substantial room for subjectivity in this process, which can reduce the accuracy of an ex-ante power analysis. To address this issue, I propose the use of a QR simulation based approach to power analysis. The QR framework assumes that subjects make proportional errors in line with the relative payoff of each action, or in other words, their payoff hill. I demonstrate that the steepness of the payoff hill as well as potential asymmetry can influence the parameters used in a power analysis, and that assuming normally distributed noise about theoretical point predictions can result in under-powered studies. The QR simulation approach to power analysis uses structural techniques on previous but related data-sets to calculate appropriate estimates of the QR noise parameter λ, as well as any additional parameters that are required for the operative channel of the treatment effect to work. With those parameters, the QR framework provides predictions for behavior in novel environments. Those predictions can then be used to generate simulated data-sets, with which power analyses can be conducted even in situations where a closed-form solution for power does not exist.

I demonstrate the use of the QR simulation approach using a motivating example of Bayesian Persuasion. That particular experimental environment shows the potential failings of a more standard rule of thumb due to asymmetric payoff hills that vary in steepness by treatment. It is not always the case that theoretical treatment differences should be maximized when considering the objective of maximizing power. I then use the QR framework to conduct ex-ante power analyses on various high-profile papers. From an ex-ante perspective, QR simulations usually perform well, getting reasonably close to the objectively correct ex-post approach. For selected papers, I provide an ex-ante suggestion for the optimal selection of parameters or treatments that would improve power. This results in a reasonable decrease in the number of required subjects, potentially saving both time and money. Such a process can also be conducted ex-post, when considering replication studies. Instead of the more standard direct replication, the researcher could conduct a conceptual replication with the goal of minimizing the cost of the replication, an ‘efficient’ replication. Lowering the costs of replication could result in more provision of this currently under-provided research.

Future research should take the following messages from this research. Firstly, an ex-ante power analysis should be included to justify the number of subjects in any experimental paper, and ideally this should be pre-registered. The profession as a whole should discuss phasing this in so that in the future a clear ex-ante power analysis plan would be a requirement for publication. Secondly, the QR framework is applicable to a wide variety of experimental environments, and can not only be used for power analysis in novel environments, but it can also be used as a tool to guide experimental design both in terms of power as well as any other conceivable metric. As for future research on this particular methodological topic, further confirmation that this approach provides reasonable predictions from an ex-ante perspective would be helpful. In addition, conceptual replications in the spirit of being efficient from a power perspective would be an important contribution to supplement more traditional direct replications.

2. NETWORK DEFENSE AND BEHAVIORAL BIASES: AN EXPERIMENTAL STUDY

with Mustafa Abdallah, Saurabh Bagchi, Shreyas Sundaram, and Timothy Cason

A version of this chapter has previously been published in the Experimental Economics Journal. Citation: Woods, D., Abdallah, M., Bagchi, S., Sundaram, S., and Cason, T. (2021). Network defense and behavioral biases: an experimental study. *Experimental Economics*, https://doi.org/10.1007/s10683-021-09714-x

2.1 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, Naghizadeh, Hota, et al., 2019; Abdallah, Naghizadeh, Cason, et al., 2019; 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 & 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.

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 2.1. 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.[1] The probability of successful defense along an edge

[1]↑A non-exhaustive list of research considering the attack graph model from the Computer Security literature includes Sheyner and Wing (2003), K. C. 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.

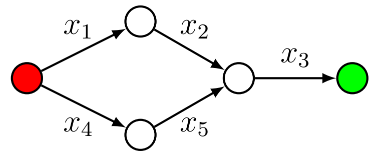


Figure 2.1. Example Directed Network Attack Graph

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 2.2.1), 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 & 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.

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.

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 & 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 research 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.[2] 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.[3] 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 research also contributes to the experimental literature of attack and defense games in network environments.[4] 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

[2]↑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).

[3]↑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.

[4]↑See Kosfeld (2004) for a survey of network experiments more generally.

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.

2.2 Theory and Hypotheses

2.2.1 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.[5]

2.2.2 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 $p(x_\mathrm{i})$ 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. $w(p(x_\mathrm{i});\alpha) = \exp[-(-\log(p(x_\mathrm{i})))^\alpha]$ 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 $w(p(x_\mathrm{i});\alpha)$ to $w(p)$ or $w(p(x_\mathrm{i}))$.

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 $p(x_\mathrm{i})$, and his

[5]↑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), D. Wu et al. (2018), and Leibowitz et al. (2019).

probability weighting parameter α. We denote the defender's overall perceived probability of defense along path j as $f_{\mathrm{j}}(x;\alpha)$.

An attacker passing along an edge to reach a specific node is a separate and independent event from all other edges.[6] 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) (Tversky & 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 $f_{\mathrm{j}}(x;\alpha) = w(p(x_1)) + [1 - w(p(x_1))]\,[w(p(x_2)) + (1 - w(p(x_2)))w(p(x_3))]$.[7] The defender's constrained objective problem is presented in Equation 2.1.

$$\begin{aligned} \underset{x}{\operatorname{argmax}} \quad & \min\{f_1(x;\alpha), f_2(x;\alpha), \ldots, f_J(x;\alpha)\} \\ s.t. \quad & x_{\mathrm{i}} \geq 0, \mathrm{i} = 1, 2, \ldots, N \\ & \sum_{\mathrm{i}}^{N} x_{\mathrm{i}} \leq B \end{aligned} \tag{2.1}$$

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.

[6]↑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.

[7]↑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., $f_{\mathrm{j}}(x;\alpha) = w(p(x_1) + [1 - p(x_1)]\,[p(x_2) + (1 - p(x_2))p(x_3)]))$ does not yield interesting comparative statics in α due to the monotonicity of the probability weighting function, so we do not consider it further.

2.2.3 Common Edges

The described objective in Equation 2.1 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 2.1. 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 $p(x_\mathrm{i}) = 1 - \mathrm{e}^{\frac{-x_\mathrm{i}}{z}}$ (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.[8] 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 B.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 ϵ from the common edge x_3 and placing it on a non-common edge. Placing the ϵ only on one edge is non-optimal for any α, as the sophisticated attacker will attack the weaker path, meaning ϵ 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.

[8]↑Weighting the probability of a successful attack along an edge instead is analytically tractable as terms conveniently cancel, as shown in Abdallah, Naghizadeh, Hota, et al. (2019). 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, Naghizadeh, Hota, et al. (2019) have been numerically confirmed to hold in the current environment.

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.[9] In particular, consider the edge defense function from before ($p(x_i) = 1 - e^{\frac{-x_i}{z}}$). Figure 2.2 plots the optimal amount to allocate to the common edge for different values of z and different levels of probability weighting α. 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 of edge defense functions (Abdallah, Naghizadeh, Hota, et al., 2019). As α 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).

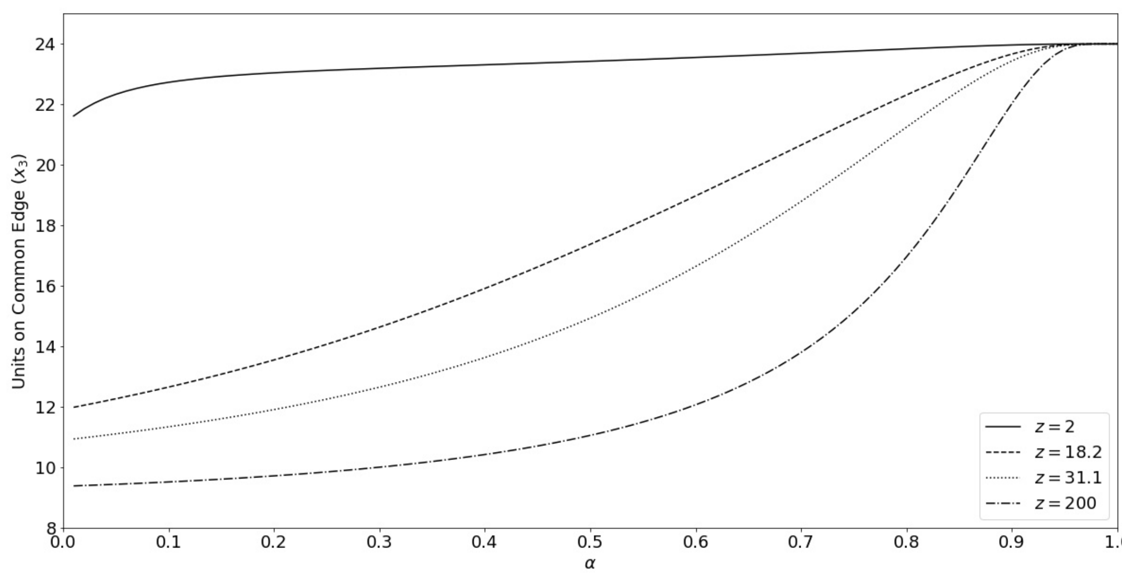


Figure 2.2. Allocation to Common Edge for $p(x_i) = 1 - e^{\frac{-x_i}{z}}$

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 2.3 illustrates that regardless of the convexity of the edge defense function, the amount allocated to the common edge decreases as α decreases

[9]↑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)

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 $p(x_{\mathrm{i}})$ 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.[10]

These observations lead to our first testable hypotheses:

Hypothesis 1. *The amount allocated to common edges (weakly) decreases as α 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.*

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

[10]↑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 B.1. Where exactly the trade-off is resolved depends on α as well as the specific functional form of $p(x_{\mathrm{i}})$. This is why the optimal allocation differs over α for a given $p(x_{\mathrm{i}})$, as well as over different $p(x_{\mathrm{i}})$ for a given α. Both patterns are displayed in Figures 2.2 and 2.3.

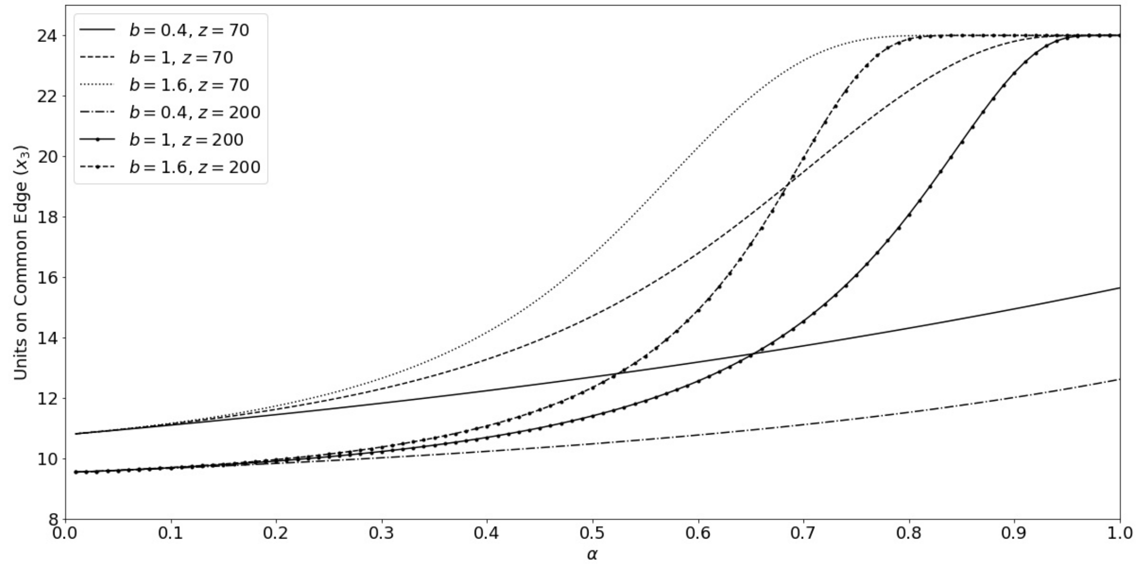


Figure 2.3. Allocation to Common Edge for $p(x_\mathrm{i}) = (\frac{x_\mathrm{i}}{z})^b$

Network Red

Network Red employs the network structure presented earlier in Figure 2.1, and has an edge defense function of $p(x_\mathrm{i}) = 1 - \mathrm{e}^{\frac{-x_\mathrm{i}}{18.2}}$.[11] 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 α 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 α's are displayed graphically in Figure 2.2 by the line associated with $z = 18.2$.[12] According to Hypothesis 2, a defender with $\alpha = 1$ would allocate $x_3 = 24$, and $x_1 = x_2 = x_4 = x_5 = 0$.

[11]↑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 ($p = 0.73$).

[12]↑These numerical solutions are continuous, although subjects were restricted to discrete (integer-valued) allocations.

Network Orange

Network Orange also takes place on the network shown in Figure 2.1, 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 α's is displayed in Figure 2.2 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 α types.

Network Yellow

Network Yellow also takes place on the network shown in Figure 2.1. 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 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 α's are presented in Figure 2.3, 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 α 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.

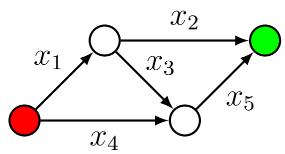


Figure 2.4. Network Graph with an Extraneous Edge (x_3)

2.2.4 Extraneous Edges

Consider the network displayed in Figure 2.4. 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:

$$\begin{aligned} F(x;\alpha) = \min\{&w(p(x_1);\alpha) + [1 - w(p(x_1);\alpha)]\, w(p(x_2);\alpha), \\ &w(p(x_4);\alpha) + [1 - w(p(x_4);\alpha)]\, w(p(x_5);\alpha), \\ &w(p(x_1);\alpha) + [1 - w(p(x_1);\alpha)]\, [w(p(x_3);\alpha) + (1 - w(p(x_3);\alpha))w(p(x_5);\alpha)]\} \end{aligned} \tag{2.2}$$

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 ϵ 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 $p(x_\mathrm{i}) > 0\ \forall x_\mathrm{i}$. 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.[13] These results lead to our next testable Hypothesis:

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

[13]↑Further details are presented in Appendix B.1.

Network Blue

Network Blue takes place on the network with an extraneous edge, as shown in Figure 2.4, 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.

2.2.5 Unequal Path Lengths

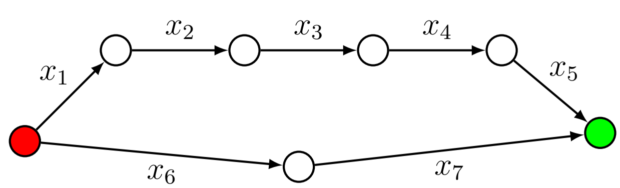


Figure 2.5. Network Graph with Unequal Path Lengths

Consider the network displayed in Figure 2.5. 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 2.6 shows the numerical solutions considering the total relative allocations to edges in the top and bottom paths of the exponential edge defense functions of $p(x_{\mathrm{i}}) = 1 - \mathrm{e}^{\frac{-x_{\mathrm{i}}}{z}}$ 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 α decreases from 1.*

Network Green

Network Green takes place on the network shown in Figure 2.5, again with an edge defense function of $p(x_{\mathrm{i}}) = 1 - \mathrm{e}^{\frac{-x_{\mathrm{i}}}{18.2}}$. 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.

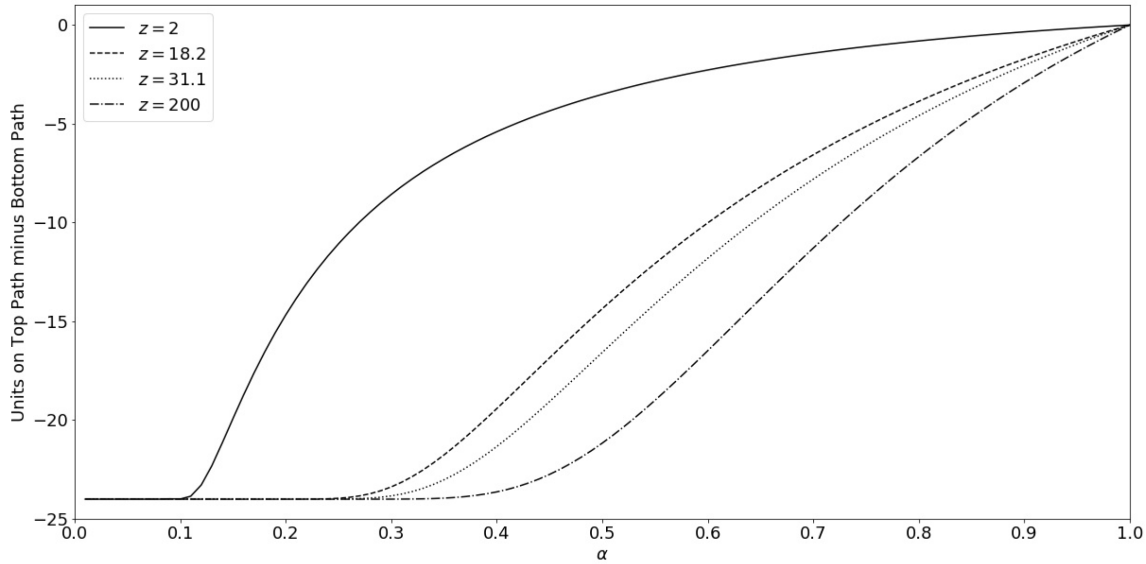
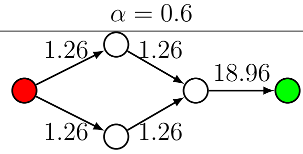
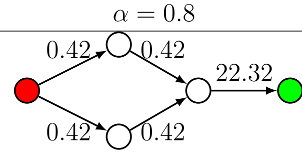
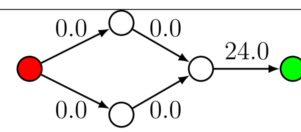
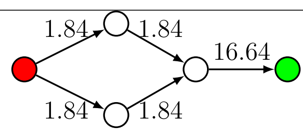
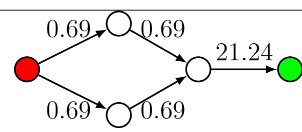
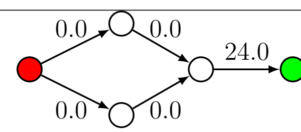
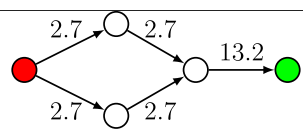
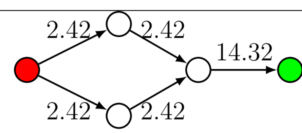
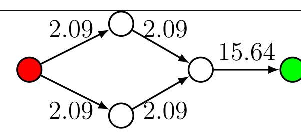
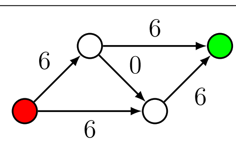
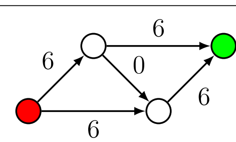
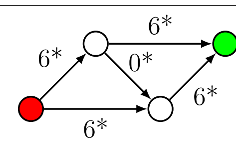
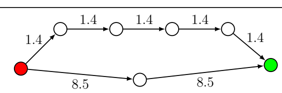
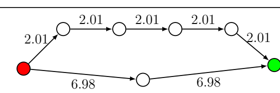
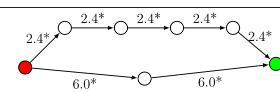


Figure 2.6. Allocation to Top (Long) minus Bottom (Short) Path for $p(x_\mathrm{i}) = 1 - \mathrm{e}^{\frac{-x_\mathrm{i}}{z}}$

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 α's.

Table 2.1 summarizes the predictions for all five networks in the experiment for three levels of α.

Table 2.1. Theoretical Predictions for Selected α with $B = 24$

	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 1$
Network Red $p(x_i) = 1 - e^{\frac{-x_i}{18.2}}$			
Network Orange $p(x_i) = 1 - e^{\frac{-x_i}{31.1}}$			
Network Yellow $p(x_i) = \frac{x_i^{0.4}}{70^{0.4}}$			
Network Blue $p(x_i) = 1 - e^{\frac{-x_i}{18.2}}$			
Network Green $p(x_i) = 1 - e^{\frac{-x_i}{18.2}}$			

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

2.3 Experimental Design

2.3.1 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 α 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 2.7, 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.

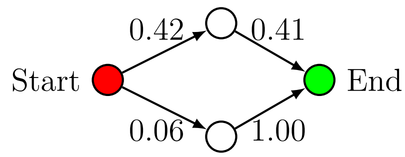


Figure 2.7. Network Attack Task Example

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 α cutoffs), α can be bounded.

Using a dynamic bisection or staircase method could recover increasingly tight bounds on α, 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 α. 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 α 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 B.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.[14] 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 α 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 α 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 B.2.

2.3.2 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.[15] 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

[14]↑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.

[15]↑In only 5 of the 4550 total decisions did subjects allocate less than all 24 units.

the interface is shown in Appendix B.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.

2.3.3 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).[16] 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.[17] Subjects were instructed to refer to specific instructions when the software (implemented in oTree (D. L. 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 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

[16]↑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.

[17]↑These instructions are available in Appendix B.7.

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.

2.4 Results

We begin the results with an overview of the probability weighting (α) 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.

2.4.1 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 α. 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 2.8 presents the CDF for the subjects' elicited α'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 α from the Network Attack Task. The quintile cutoff points are $\alpha = 0.64$, $\alpha = 0.80$, $\alpha = 0.90$, $\alpha = 0.95$.*

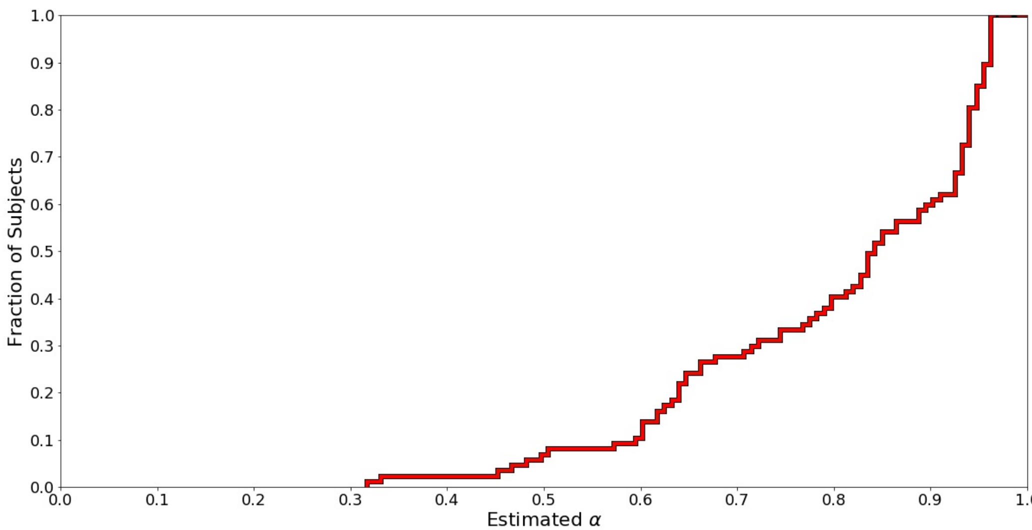
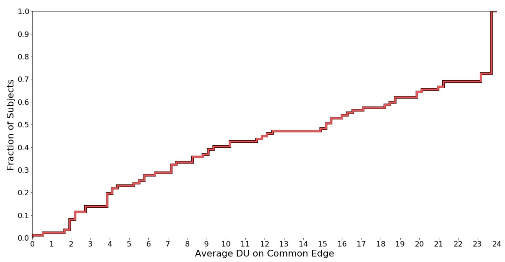


Figure 2.8. CDF of Elicited α from the Network Attack Task

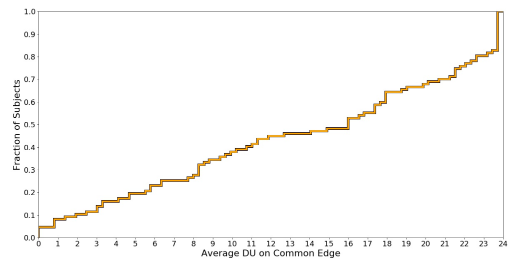
2.4.2 Network Defense Tasks

Summary

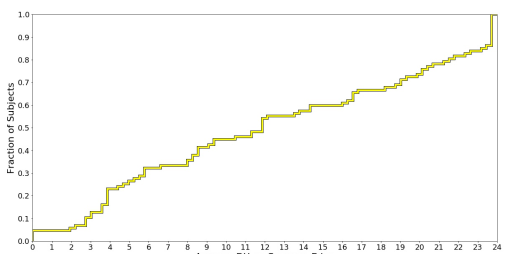
Figure 2.9 presents the CDF's of mean defense allocations for key subject decisions in each of the five Network Defense Tasks. This also indicates substantial heterogeneity in subject behavior. In the Red and Orange networks, 26.4% and 16.1% of subjects respectively allocate all units to the common edge, 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 α 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 α, 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.



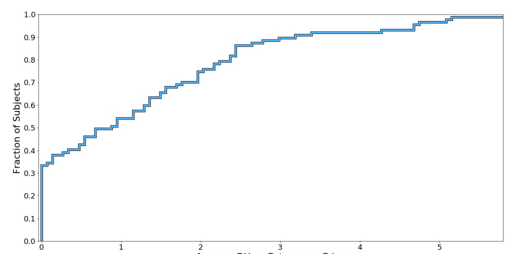
(a) Network Red - Common Edge



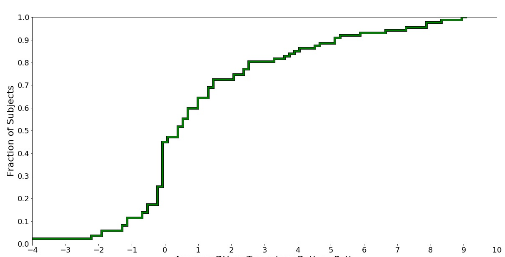
(b) Network Orange - Common Edge



(c) Network Yellow - Common Edge



(d) Network Blue - Extraneous Edge



(e) Network Green - Top minus Bottom

Figure 2.9. CDFs of Average Behavior Across all Rounds in the Network Defense Tasks

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 α is derived from the Network Attack Task, which we use to test our Hypotheses in the Network Defense Task. Table 2.2 presents the non-parametric Spearman's ρ between the elicited probability weighting (α) and decision noise (λ) from the Network Attack Task and key average subject behavior in each of the Network Defense Tasks.[18,19] The decision noise parameter λ is estimated by the logit function, a commonly used structure in Quantal Response Equilibria (McKelvey & Palfrey, 1995).[20] A higher λ is consistent with less noisy behavior, meaning subjects choose their payoff maximizing action more frequently. Table 2.2 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 2.2 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 2.2. The elicited α from the Network Attack Task is strongly and significantly correlated with defense in all Network Defense Tasks except Network Yellow.[21] The correlations of α with the common edge networks are positive, consistent with Hypothesis 1. The negative correlation in Network Green indicates that subjects with estimated α'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 α from the Network Attack Task also captures some element of cognitive abil-

[18]↑We consider the same analysis including the Binary Lottery Task in Appendix B.4. The elicited α'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.

[19]↑Unless otherwise stated, all p-values and statistical tests are two-sided.

[20]↑$Prob(Top) = \frac{1}{1+e^{-\lambda(U(Top)-U(Bottom))}}$ 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.

[21]↑The lack of a significant correlation in Network Yellow is not necessarily surprising, due to the deliberate reduction of the separation of α types in this network to evaluate Hypothesis 2.

ity. This interpretation is consistent with the strong correlation of decision noise (λ) with probability weighting (α).

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 α close to 1 should 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 α cutoff $\in [.86, .95]$ (see Appendix B.3).

Table 2.2. Spearman's ρ Correlation Table

	α	λ	Red Common Edge	Orange Common Edge	Yellow Common Edge	Blue Extra Edge
λ	$\rho = 0.764^{***}$	$\rho = 1$				
Red Common Edge	$\rho = 0.260^{**}$	$\rho = 0.151$	$\rho = 1$			
Orange Common Edge	$\rho = 0.226^{**}$	$\rho = 0.056$	$\rho = 0.654^{***}$	$\rho = 1$		
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	$\rho = -0.255^{**}$	$\rho = -0.139$	$\rho = -0.241^{**}$	$\rho = -0.292^{***}$	$\rho = -0.237^{**}$	$\rho = 0.241^{**}$

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

Result 2. *The probability weighting parameter α 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. α 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.

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 & 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.[22]

Table 2.3 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 & 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

[22]↑We also cluster at the individual network task level in an alternative estimation presented in Appendix B.5. That analysis identifies similar patterns of behavior.

Table 2.3. Cluster Analysis

Cluster Name	Network Red	Network Orange	Network Yellow	Network Blue	Network Green
C1: Near Optimal $\alpha = 1$ - Late Revelation 12.6%	0.0 0.4 23.3 0.0 0.3	0.0 0.3 23.6 0.1 0.0	0.9 0.9 20.4 0.9 0.9	12.0 0.0 0.0 12.0 0.0	0.1 0.8 1.7 0.2 9.2 2.4 9.6
C2: Near Optimal $\alpha = 1$ - Early Revelation 4.6%	0.0 0.1 23.8 0.0 0.1	0.0 0.0 24.0 0.0 0.0	1.0 1.0 20.0 1.0 1.0	0.0 12.0 0.0 0.0 12.0	0.1 0.1 0.0 12.1 0.0 11.6 0.1
C3: Late Revelation - Some Diversification 4.6%	0.4 2.2 19.0 0.7 1.7	2.7 2.9 14.1 1.7 2.6	0.0 0.6 22.7 0.0 0.7	10.6 1.3 2.1 9.8 0.2	0.0 0.0 0.7 0.0 11.7 0.8 10.8
C4: Naive Diversification 16.1%	4.4 4.3 5.8 3.9 5.6	5.0 5.2 3.5 5.6 4.7	6.1 4.0 6.0 4.7 3.2	5.8 4.0 4.7 5.2 4.3	3.2 3.1 2.7 3.0 2.6 5.3 4.1
C5: $\alpha < 1$ - Some Diversification 12.6%	2.1 2.3 15.2 2.2 2.2	3.3 3.5 10.4 3.3 3.5	3.8 4.8 6.8 3.9 4.7	5.2 5.3 2.2 6.4 4.9	3.1 2.5 2.6 4.4 2.4 5.6 3.4
C6: Early Revelation - Some Diversification 9.2%	7.7 2.9 1.9 7.7 3.8	4.8 1.8 11.2 4.6 1.6	4.7 4.2 5.5 5.7 3.9	3.3 7.7 2.5 1.9 8.6	1.5 1.2 0.7 5.9 3.5 6.3 4.9
C7: $\alpha < 1$ - Mild Diversification 16.1%	3.6 3.9 9.0 3.6 3.9	1.0 1.9 18.2 1.1 1.8	1.6 3.2 14.4 1.5 3.3	7.8 4.0 0.6 7.4 4.2	2.9 1.8 1.6 3.1 1.7 6.7 6.2
C8: Naive Diversification and Late Revelation 5.7%	0.0 0.0 24.0 0.0 0.0	0.6 2.4 18.9 0.7 1.4	4.9 4.8 4.6 4.9 4.8	5.4 5.6 2.0 5.6 5.4	2.5 0.8 2.4 0.6 5.4 5.2 7.1
C9: Late Revelation - mild Diversification 13.8%	0.0 0.0 24.0 0.0 0.0	1.4 1.5 18.2 1.4 1.5	0.0 0.0 24.0 0.0 0.0	9.1 1.7 1.5 8.3 3.4	2.0 2.5 2.0 2.0 3.7 5.8 6.0
C10: Early Revelation - mild Diversification 4.6%	9.3 0.5 2.2 11.9 0.1	8.4 3.9 1.0 8.5 2.2	10.3 0.0 3.2 10.5 0.0	3.9 7.1 1.2 4.8 7.0	2.6 1.7 1.1 7.0 0.7 8.8 2.1

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.

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 α 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 a period of time is minimized within an attack.[23] 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)).

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

[23]↑It is of course possible that subjects are playing out the attack process in their imagination, while reading the outcomes sequentially.

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 2.2 suggest that α may be picking up some measure of cognitive ability.[24] We include information self-reported by subjects in a post-experiment survey, such as field of study, a

[24]↑Choi et al. (2018) reports evidence suggesting a correlation between cognitive ability and probability weighting.

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 2.4 reports a series of censored tobit regressions, with the dependent variable for each network corresponding to the key summary statistics shown earlier in Figure 2.9: 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.

Table 2.4. Tobit Regression Analysis

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

The top row shows the effect that our measure of probability weighting (α, 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 α. 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, α 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 α on this allocation. Finally, Hypothesis 4 predicts that increasing α should increase the amount allocated to the top path in Network Green. While there is a statistically significant affect of α, 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, α 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). α 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 2.4 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.

2.5 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 (α) 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 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. Fréchette (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 research.

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 α (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 α 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 α is only picking up on cognitive ability. Future research could investigate why results from Network Green are anomalous, perhaps with an alternative elicitation of α 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.

3. BEHAVIORAL BANDITS: ANALYZING THE EXPLORATION VERSUS EXPLOITATION TRADE-OFF IN THE LAB

with Stanton Hudja

A version of this chapter has previously been published in the SSRN Electronic Journal. Citation: Hudja, S. and Woods, D. (2019). Behavioral Bandits: Analyzing the Exploration Versus Exploitation Trade-Off in the Lab. *SSRN Economic Journal*, http://dx.doi.org/10.2139/ssrn.3484498

3.1 Introduction

The dilemma of whether to explore an unfamiliar option or exploit a familiar option is pervasive in economics. For example, a CEO often decides between investing resources into a new product and an already established product. A farmer often decides between planting a new crop and an old crop. A consumer often chooses between eating at a new restaurant and a well-known restaurant. In these and many other examples, individuals must decide whether to forgo a predictable payoff to learn more about an unfamiliar, but potentially lucrative, option.

The single-agent exponential bandit model (i.e., the one-player case of the cooperative problem in Keller et al., 2005) provides a parsimonious model of this exploration versus exploitation trade-off. In economics, many models of exploration build on this rudimentary model. This model provides the foundation for models of dynamic public goods problems (Keller et al., 2005), innovation contests (Halac et al., 2017; Bimpikis et al., 2019), long-term contracts (Halac et al., 2016), moral hazard in teams (Bonatti & Hörner, 2011), and voting for reforms (Strulovici, 2010; Khromenkova, 2015). While many models build on the single-agent exponential bandit model, there is little empirical research on how well it describes individuals' resolution of the exploration versus exploitation trade-off.

We conjecture that exploration is influenced by behavioral factors that are generally unaccounted-for in the single-agent exponential bandit model: risk preferences, base rate neglect/conservatism, and non-linear probability weighting.[1] We propose these behavioral factors because they are ubiquitous in human decision making.[2] We address two questions that follow from this conjecture. First, how well does the single-agent exponential bandit model describe an individual's resolution of the exploration versus exploitation trade-off? Second, are any of our posited behavioral factors consistent with individual behavior? These questions are important from a theoretical perspective because they can pinpoint where inaccuracies in models of exploration may occur as well as allowing theorists to focus on the most empirically relevant behavioral factors. Additionally, they are important from a practical perspective because they can help identify which policies or contracts are likely to be effective in encouraging more efficient exploration.

We design a laboratory experiment to analyze the single-agent exponential bandit model and to quantify the impact of our posited behavioral factors. We use a laboratory experiment as it allows us to analyze a setting closely resembling the model environment. In our experiment, a subject continually chooses in near-continuous time between a risky and safe action. The safe action always pays a certain reward, while the risky action can be good or bad. A good risky action dominates the safe action and occasionally pays out a high reward. A bad risky action is dominated by the safe action and never pays out a reward. A subject is initially unsure of whether she has a good risky action and can only learn about the risky action by trying it out over time. If she receives a reward from the risky action, she knows that her risky action is good. If she continues to try out the risky action without ever receiving a reward, her belief that her risky action is good should continuously decrease.

A subject in this environment faces a trade-off. Experimentation with the risky action provides the immediate possibility of a high reward and provides valuable information about the likelihood of future rewards. However, experimentation comes at the cost of the safe action. The single-agent exponential bandit model predicts that a subject will follow a threshold strategy where she will choose the risky action for as long as her belief that the

[1] ↑An important recent extension to this model is Keller et al. (2019), who incorporate risk aversion.
[2] ↑For example, see Chapter 2 in Dhami (2016) for a textbook overview of the theoretical and empirical work on risk preferences and non-linear probability weighting and Chapter 19 for base rate neglect/conservatism.

risky action is good is sufficiently high. If her belief ever drops below this cutoff belief, the costs of experimentation outweigh the benefits and she will forever choose the safe action.

The single-agent exponential bandit model predicts that the length of time that a subject is willing to experiment will depend on various factors such as the discount factor, safe action, and prior belief. A subject is predicted to be willing to experiment longer as the discount factor increases because the continuation value of experimentation increases. A subject is predicted to be willing to experiment longer as the value of the safe action decreases because this decreases the opportunity cost of experimentation. Lastly, a subject's willingness to experiment is predicted to increase as the prior belief increases because this lengthens the time until the cutoff belief is reached.

The experiment consists of four treatments: the Baseline, "High Prior", "Low Safe Action", and "High Discount Factor" treatments. Each of the three non-Baseline treatments only differ from the Baseline treatment by one parameter. The High Prior, Low Safe Action, and High Discount Factor treatments isolate the impact of the prior, value of the safe action, and discount factor, respectively. These four treatments allow us to conduct a robust test of the single-agent exponential bandit model as we can test various comparative statics and point predictions.

The experimental data is used to test three hypotheses, with each subsequent hypothesis test being a more rigorous test of the model's predictions. The first hypothesis, which considers comparative statics, is that subjects become willing to experiment longer as the discount factor increases, the value of the safe action decreases, or the prior belief increases. The second hypothesis is that subjects increase their willingness to experiment by the predicted length when the discount factor increases, prior belief increases, or value of the safe action decreases. The third hypothesis is that subjects are willing to experiment for as long as predicted in each treatment. We find support for only the first hypothesis. Additionally, there is strong evidence of under-experimentation in three of our four treatments and some evidence of under-experimentation in the fourth.

The lack of support for the second and third hypotheses suggest that subjects have behavioral factors that influence their experimentation. We incorporate risk preferences, base rate neglect/conservatism, and non-linear probability weighting into a model of experimentation. We find that risk aversion, conservatism and inverse S-shaped probability weighting are statistically significant in our structural model. Our estimated structural model suggests that risk aversion is the predominant factor driving subjects' under-exploration in the experiment.

This research contributes to three strands of literature. The first strand is the theoretical literature on experimentation. Keller et al. (2005), building on Bolton and Harris (1999), analyze a game where all agents want to collect information on a risky action and can free-ride on other agents' costly experimentation. Strulovici (2010) and Khromenkova (2015) analyze a setting where potentially heterogeneous agents must collectively decide whether to experiment or not. As mentioned earlier, the theoretical literature on experimentation has also analyzed moral hazard in teams (Bonatti & Hörner, 2011), long-term contracting (Halac et al., 2016), and innovation contests (Halac et al., 2017; Bimpikis et al., 2019). This chapter contributes to this literature by providing empirical evidence that individuals do respond to factors that are common in exploration models such as the prior belief, cost of experimentation, and discount factor. However, our results also suggest that these models may be predicting too much experimentation and that risk aversion, conservatism, and non-linear probability weighting should be incorporated into models of exploration when possible.

The second strand is the literature on bandit experiments. Our research is related to the literature on bandit experiments where the risky action has a reward probability that can only take on one of two known values.[3] Hoelzemann and Klein (2020) analyze the game of strategic experimentation in Keller et al. (2005) in the lab and find that subjects appear to respond to strategic incentives by free-riding. Hudja (2019) analyzes Strulovici (2010) in the lab and finds that subjects appear to respond to payoff externalities imposed by other subjects. Unlike these papers, our focus is on individual behavior in an environment without the confounds of strategic incentives or externalities. Our research is closest to Banks et al. (1997), who fail to show that subjects respond to changes in the discount factor and

[3]↑There are other types of bandit experiments. For example, there have been experiments on bandits where the reward probability of the risky action is randomly drawn from a continuous distribution (Meyer and Shi, 1995; C. Anderson, 2001; C. Anderson, 2012). These types of bandit problems are not common in economics because they do not have closed form solutions. Banovetz and Oprea (2020) experimentally analyze a bandit problem where each arm reveals its type after one pull.

reward probabilities in a discrete time bandit where both good and bad risky actions can pay out rewards. Unlike Banks et al. (1997), our research suggests that subjects do respond to changes in environmental parameters like the discount factor.[4] A question that arises in this literature is the role of risk aversion in exploration. Both Banks et al. (1997) and Hudja (2019) fail to show that elicited risk aversion is a significant predictor of exploration. This is surprising as risk is inherent in exploration. A possible explanation is that these papers elicited risk aversion in a different environment than the experimental environment (Isaac and James, 2000; Charness et al., 2020). Our research contributes to this literature in part by finding a significant effect of risk aversion when estimating risk preferences within a bandit environment.

The third strand is the literature on continuous time experiments. Continuous time experiments are mostly continuous time versions of classic discrete time games (Oprea et al., 2011; Friedman and Oprea, 2012) and experiments featuring stochastic processes. Our experiment falls into the latter type of continuous time experiment. Most of the stochastic process experiments approximate either Brownian Motion (Oprea et al., 2009; S. T. Anderson et al., 2010; Oprea, 2014) or Poisson Processes (Hoelzemann and Klein, 2020; Hudja, 2019) in the lab. Our results suggest that subjects may deviate from theory that assumes linear probability weighting in the presence of Poisson Processes.

3.2 Theory

The theory motivating this experiment is borrowed from Strulovici (2010).[5] Time (t) is continuous and payoffs are discounted at a rate r. An agent continually decides between two actions. The first action is a safe action, which yields a flow payoff of s (> 0) per unit of time. The second action is a risky action, which can be either good or bad. A good risky action pays out rewards (magnitude h) at random times based on a Poisson process with

[4]↑We conduct power analyses that suggest that Banks et al. (1997) is under-powered and our experiment is well-powered. We conduct this analysis in Appendix C.4.

[5]↑The single-agent case of Strulovici (2010) restricts agents into choosing between the two actions in each period of time, they can not divide a resource between the actions as in Keller et al. (2005). This restriction does not change theoretical predictions and is convenient for an experimental implementation.

parameter λ. A bad risky action pays out nothing. The expected flow payoff of a good risky action is λh, which is greater than the flow payoff s.

The risky action has an initial probability p_0 of being good. An agent's belief (p) about the risky action evolves from the prior according to Bayes' rule. In the absence of a reward, an agent's belief is given by

$$p = \frac{p_0 \mathrm{e}^{-\lambda t}}{p_0 \mathrm{e}^{-\lambda t} + (1 - p_0)},$$

where t is the amount of time spent experimenting.[6] Note that in the absence of a reward, an agent's belief is decreasing in t. If and when a first reward arrives, an agent's belief jumps to one as she knows that she has a good risky action.

The optimal strategy depends on an agent's belief of the state of the risky action. An agent should implement the risky action if and only if her current belief that the risky action is good (p) is greater than or equal to a cutoff belief p_A. The cutoff belief is given by

$$p_A = \frac{s}{\lambda h + \frac{\lambda}{r}(\lambda h - s)}.$$

This cutoff belief solves the indifference equation $p\lambda h + p\lambda(\frac{\lambda h}{r} - \frac{s}{r}) = s$. The effect of the risky action on an unsure agent can be decomposed into two elements: (i) the expected payoff $p\lambda h$, and (ii) the jump in the value function (from $\frac{s}{r}$ to $\frac{\lambda h}{r}$ when she is indifferent) if a reward is received, which occurs at a rate λ with probability p. If the safe action is chosen, the payoff rate is s.

The cutoff belief implies that agents are non-myopic, i.e., that agents value the information generated by experimentation. A myopic agent would only experiment if the expected flow payoff of the risky action is greater than or equal to the flow payoff of the safe action. Thus, a myopic agent would experiment if and only if $p \geq \frac{s}{\lambda h}$, which is sub-optimal as this cutoff belief is greater than p_A.

The laboratory experiment focuses on how long subjects are willing to experiment, that is, how long they would choose the risky action in the absence of any rewards. The time that an agent is willing to experiment is found by solving for how long it would take for an

[6]↑Experimentation occurs when an agent chooses the risky action while she is still unsure of the state of the risky action.

agent's belief to reach her cutoff belief in the absence of any rewards from the risky action. This optimal experimentation time is given by

$$t^* = \frac{-ln\left(\frac{p_A(1-p_0)}{p_0(1-p_A)}\right)}{\lambda}.$$

This time is increasing in the prior belief and decreasing in the cutoff belief.

The experimental predictions are based on changes in the environmental parameters. A subject is predicted to experiment longer as the value (s) of the safe action decreases. As the value of the safe action decreases, the opportunity cost of experimentation decreases and agents become willing to experiment longer as they are willing to experiment at lower beliefs. A subject is also predicted to experiment longer as the discount factor (1 - r) increases. As the discount factor increases, agents place more value on the future rewards associated with a good risky action and agents become willing to experiment longer as they are willing to experiment at lower beliefs. Lastly, a subject is predicted to experiment longer as the prior (p_0) probability that the risky action is good increases. As the prior increases, it takes longer for agents to reach their cutoff belief in the absence of rewards and thus agents become willing to experiment longer.

3.2.1 Discrete Implementation

This research uses a discrete time approximation to test various predictions of the model. We utilize a discrete time approximation because it is not possible to implement a perfectly continuous time bandit problem in the laboratory as computers are limited by their internal clock. The approximation of this model is taken from Hudja (2019).

The approximation consists of dividing time into a number of ticks, each of length Δ seconds. In the approximation, only one decision can be made in a tick and only one payoff can be received in a tick. A bad risky action never returns h in a given tick, while a good risky action has a probability of $\lambda\Delta$ of returning h in a given tick. The safe action returns $s\Delta$, which is the payoff from exerting the safe action for Δ seconds in a continuous time bandit.

Table 3.1. Values of p_0, $s\Delta$, and δ for each Treatment

Treatment	Prediction	Prior	Safe Action	Discount Rate
Baseline:	130	$p_0 = 0.333$	$s\Delta = 0.5$	$\delta = 0.996$
High Prior:	199	$p_0 = 0.500$	$s\Delta = 0.5$	$\delta = 0.996$
Low Safe Action:	198	$p_0 = 0.333$	$s\Delta = 0.3$	$\delta = 0.996$
High Discount Factor:	199	$p_0 = 0.333$	$s\Delta = 0.5$	$\delta = 0.998\bar{3}$

Each treatment has a value of $\lambda\Delta = 0.01$ and $h = 155$.

The approximation also consists of replacing the infinite horizon of the continuous time problem with an indefinite horizon. In a given tick, there is a probability of $r\Delta$ that the tick will end the period. This results in a discount factor of $\delta = 1 - r\Delta$.

3.3 Experimental Design

The experiment is designed with two goals in mind. The first goal is to analyze how well the single-agent exponential bandit model describes individual behavior. The second goal is to create a dataset that allows us to econometrically test for risk preferences, base rate neglect/conservatism, and non-linear probability weighting.

3.3.1 Treatments and Parameters

The experiment consists of four treatments. These treatments are the Baseline treatment, the "High Prior" treatment, the "Low Safe Action" treatment, and the "High Discount Factor" treatment. Each treatment consists of a parameter set that has a unique combination of p_0, δ, and $s\Delta$. In each of these treatments, $\lambda\Delta$ is set to 0.01 and h is set to 155 experimental points.[7] The tick length Δ is set to 200 milliseconds in all treatments. Table 3.1 displays the parameters, and predictions, for each treatment in the experiment.

The experiment uses a within-subjects design where each session consists of subjects facing the Baseline treatment and one of the three other treatments. "High Prior" sessions isolate the effect of the prior on experimentation and consist of each subject facing the

[7]↑We fix $\lambda\Delta$ as it has a non-monotonic effect on experimentation, which can reduce power. An increase in $\lambda\Delta$ increases the myopic value of experimentation while increasing the rate of belief updating in the absence of a reward. We fix h as only the ratio of $\frac{h}{s\Delta}$ matters for non-behavioral predictions.

Table 3.2. Predictions for each Treatment in Discrete Time and Continuous Time

	Willingness to Experiment		**Cutoff Belief**	
Treatment	Discrete Time	Cont. Time	Discrete Time	Cont. Time
Baseline	130.0	130.2	0.112	0.112
High Prior	199.0	199.5	0.112	0.112
Low Safe Action	198.0	198.7	0.064	0.064
High Discount Factor	199.0	199.5	0.064	0.064

Baseline and High Prior treatments. "Low Safe Action" sessions isolate the effect of the safe action on experimentation and consist of each subject facing the Baseline and Low Safe Action treatments. "High Discount Factor" sessions isolate the effect of the discount factor on experimentation and consist of each subject facing the Baseline and High Discount Factor treatments. Within each session, subjects face twenty periods of the Baseline treatment and twenty periods of the session's other treatment. One half of the subjects in a session start out with the Baseline treatment and the other half of the subjects in a session start out with the session's other treatment.

Table 3.1 displays the predictions for each treatment. Appendix C.1 provides details for how these predictions were derived. In the Baseline treatment, subjects are predicted to be willing to experiment for 130 ticks.[8] In the High Prior treatment, subjects are predicted to be willing to experiment for 199 ticks. In the Low Safe Action treatment, subjects are predicted to be willing to experiment for 198 ticks. In the High Discount Factor treatment, subjects are predicted to be willing to experiment for 199 ticks. The parameters for the three non-Baseline treatments were chosen such that they would result in similar predictions, while satisfying the constraint of being simple for subjects to understand.

Table 3.2 compares the discrete time and continuous time predictions. The predictions for the discrete time approximation are close to the continuous time predictions. For each treatment, a subject's predicted willingness to experiment in the discrete time approximation is within one tick of their predicted willingness to experiment in continuous time. For each

[8]↑In this chapter, willingness to experiment refers to how long a subject is willing to experiment without ever obtaining a reward.

treatment, a subject's predicted cutoff belief in the discrete time approximation is within one one-hundredth of their predicted cutoff belief in continuous time.

3.3.2 Experiment

Instructions for the experiment were read aloud at the start of each session. The instructions were composed of a written component that outlined the experiment and a separate video that illustrated the experimental interface. After the instructions were read, subjects completed five comprehension questions that were each worth $1.00 when answered correctly. Upon completion of the comprehension questions, the session began.

The environment is described through an analogy of balls being drawn with replacement from a bag. Subjects can either draw a ball (the risky action) or not draw a ball (the safe action) in a given tick. There are two bags: (i) a "uniform" bag and (ii) a "mixed" bag. The uniform bag consists of 100 yellow balls. The mixed bag consists of 1 red ball and 99 yellow balls. In this analogy, the mixed bag is a good state, with a red ball being a reward and a yellow ball returning nothing. At the beginning of each period, one of the two bags is randomly selected for each subject. The mixed bag is selected with probability p_0 and the uniform bag is selected with probability $1 - p_0$. The bag stays the same throughout the period.

At the start of each period, subjects choose an initial action. Once a subject decides on an initial action, a five-second countdown begins. At the end of the five-second countdown, the first tick occurs. If a subject initially chooses to draw, she continually draws until, as an unsure decision maker, she decides to stop. Starting from the initial action, whenever a subject is unsure of the risky action and chooses to not draw, she is prevented from drawing for the rest of the period. Additionally, a subject is prevented from switching to the safe action once she obtains a reward.[9] Ticks continue until the random termination of the period.

Subjects receive feedback throughout the period through a graph displayed on their screen. A stationary red line is displayed at the current tick. The number of balls drawn is displayed to the right of the center of the red line. The current tick number is displayed to

[9]↑This prevents subjects from accidentally switching to the safe action after the initial reward reveals that the subject has a good risky action.

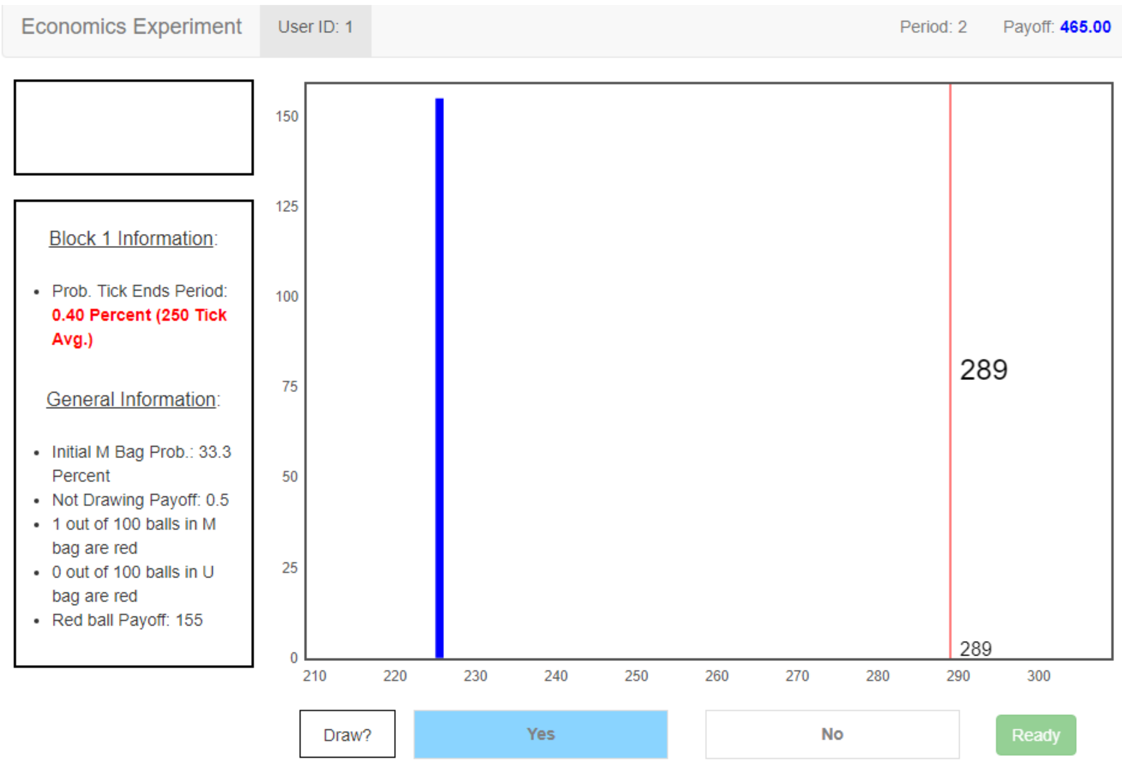


Figure 3.1. Example of the Experimental Interface

the right of the bottom of the red line. The payoff history for the last eighty ticks is shown to the left of the red line. At the beginning of each tick, subjects receive payoff information from the action preceding it. If the subject previously implemented the safe action, she sees a blue line of height $s\Delta$ drawn over the previous tick. If the subject previously implemented the risky action, she either sees no line (no reward occurred) or a blue line of height h (a reward occurred) drawn over the previous tick. Figure 3.1 displays an example of the screen from a High Discount Factor session.

At the end of the experiment, subjects completed a post-experimental survey. The post-experimental survey collected information on gender, race, country of origin, grade point average, year of schooling, and major.

3.3.3 Testable Hypotheses

The hypotheses are based on the theoretical predictions for the four treatments. The hypotheses focus on how long a subject is willing to experiment, that is, how long a subject is willing to implement the risky action without ever seeing a reward. Subjects are predicted to be willing to experiment for 130 ticks in the Baseline treatment, 199 ticks in the High Prior and High Discount Factor treatments, and 198 ticks in the Low Safe Action treatment.

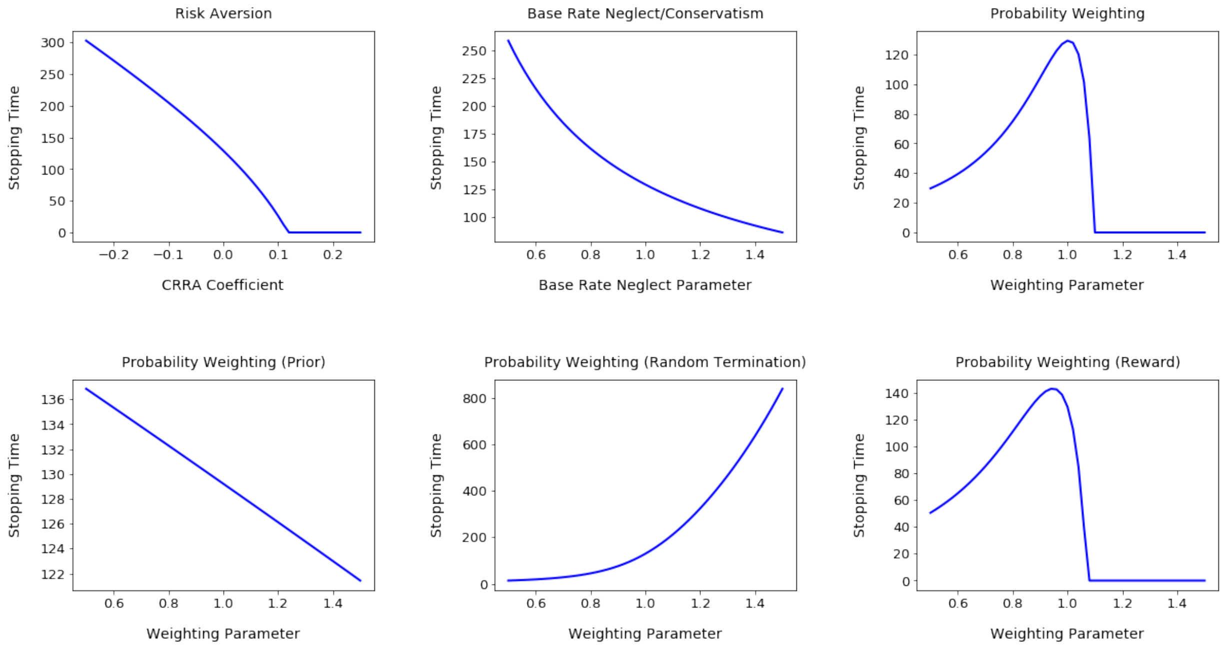
The hypotheses test how well theory describes subjects' willingness to experiment. Each subsequent hypothesis is a more rigorous test of theory. The first hypothesis focuses on how subjects respond to changes in the experimental parameters. Subjects should become willing to experiment longer when p_0 increases, δ increases, or $s\Delta$ decreases. This is our first hypothesis.

Hypothesis 1: *Subjects become willing to experiment longer when p_0 increases, δ increases, or $s\Delta$ decreases.*

The second hypothesis focuses on the magnitude of subjects' response to changes in the experimental parameters. For the given parameters, subjects should become willing to experiment for 69 more ticks when p_0 or δ increases and 68 more ticks when $s\Delta$ decreases. This leads us to our second hypothesis.

Hypothesis 2: *The length of time that subjects are willing to experiment increases by the correct magnitude when $s\Delta$, δ, or p_0 is changed.*

The third hypothesis focuses on the length of time that subjects are willing to experiment in each treatment. Subjects are predicted to be willing to experiment for 130 ticks in the Baseline treatment, 199 ticks in the High Prior treatment, 198 ticks in the Low Safe Action treatment, and 199 ticks in the High Discount Factor treatment. This leads us to our third hypothesis.



The first row of graphs display the effects of unilaterally increasing risk aversion, conservatism, and non-linear probability weighting for the Baseline treatment.
The second row of graphs display the effects of unilaterally increasing the probability weighting parameter with relation to only the prior, random termination probability, and reward probability for the Baseline treatment.

Figure 3.2. Impact of Behavioral Factors on Experimentation

Hypothesis 3: *The length of time that subjects are willing to experiment is as predicted in each treatment.*

3.3.4 Behavioral Factors

While subjects are predicted to behave according to the previous hypotheses, they may be influenced by behavioral factors that are unaccounted-for in Section 3.2. Subjects may exhibit risk preferences, base rate neglect/conservatism, and/or non-linear probability weighting. This subsection focuses on how these behavioral factors may influence subject behavior.

Figure 3.2 displays how subject behavior in the Baseline treatment is predicted to change when risk preferences, base rate neglect, or probability weighting is unilaterally varied.[10]

[10]↑Note that these are predictions for the discrete time approximation. More details on how these factors are modeled can be found in Section 3.5.

These behavioral factors have similar effects on the other treatments and these effects are shown in Appendix C.1. Risk preferences are modeled through CRRA utility ($u(x) = \frac{x^{1-\gamma}}{1-\gamma}$). The first graph suggests that a subject becomes less willing to experiment as she becomes more risk averse. This is consistent with Keller et al. (2019) as they show that experimentation is decreasing in risk aversion when $h > s$ in continuous time. Base rate neglect (conservatism) is modeled as each subject treating a tick of experimentation as if it were $\psi > 1$ (ψ<1) ticks.[11] Thus, belief updating in the presence of base rate neglect/conservatism is modeled as $\frac{p_0(1-\lambda\Delta)^{\psi\frac{t}{\Delta}}}{p_0(1-\lambda\Delta)^{\psi\frac{t}{\Delta}}+(1-p_0)}$, where $\frac{t}{\Delta}$ is the number of ticks spent experimenting in the absence of any rewards. As ψ increases, subjects' beliefs decrease faster in the absence of any rewards. Unsurprisingly, as a faster depreciation of beliefs should lead to less experimentation, the second graph suggests that subjects become less willing to experiment as ψ increases. Lastly, probability weighting is modeled through a Prelec-I function.[12] Subjects may non-linearly weight the prior, reward probability, and random termination probability. The third graph shows a non-monotonic effect of probability weighting on a subject's willingness to experiment.

The second row of graphs in Figure 3.2 provide intuition for the non-monotonic effect of probability weighting. As the probability weighting parameter (α) increases, the weighted reward probability, random termination probability, and prior probability decrease for the Baseline treatment. The fourth and the fifth graphs follow from the decreased prior and random termination probability, respectively. The sixth graph suggests a non-monotonic effect of the weighted reward probability on how long a subject is willing to experiment. The weighted reward probability results in two effects that move in opposite directions as α increases. As α increases, beliefs decrease slower in the absence of rewards, which increases a subject's willingness to experiment. However, the perceived expected flow payoff under a good risky action decreases as α increases.

[11]↑This approach is similar to the approach taken in Goeree et al. (2007) and Moreno and Rosokha (2016). Additionally, this behavioral factor is only relevant when an agent has never observed a reward.

[12]↑The Prelec-I function is given by $w(p) = \mathrm{e}^{-(-ln(p))^{\alpha}}$. When $\alpha < 1$ ($\alpha > 1$), agents over-weight (under-weight) low probability events and under-weight (over-weight) high probability events. When $\alpha = 1$, agents correctly weight probabilities.

3.3.5 Procedures

The experimental sessions were conducted in the Vernon Smith Experimental Economics Laboratory at Purdue University. Subjects were students at Purdue University who were recruited by email using ORSEE (Greiner, 2015) and who had not previously participated in a near-continuous time experiment. Sessions lasted between 60 and 95 minutes.[13] Subjects were paid for correct answers to the comprehension questions, three random periods in their first treatment, and three random periods in their second treatment. The average earnings for the experiment was $15.44. Seventy-two subjects participated in the experiment, with twenty-four subjects in each type of session.[14]

3.4 Results

The hypotheses focus on a subject's willingness to experiment, which is not always possible to observe. There are two cases where subjects do not reveal how long they are willing to experiment. The first case is where a subject receives a reward. In this case, a subject knows for certain that she has a good risky action and thus never switches to the safe action. The second case is where the period ends before an unsure subject switches to the safe action. In this case, it is never observed when a subject would stop experimenting.

We take two approaches to address these issues. The first approach is to analyze a subset of data where theory always predicts a switch to the safe action. We refer to this approach as the "Subset" approach. In this subset of data, the period lasts for at least 200 ticks and either (i) the state is bad or (ii) the first reward occurs after 199 ticks.[15] For hypothesis testing, we use the number of balls drawn in these observations.

The second approach is to use the Product Limit estimator to correct for the censoring that occurs.[16] We refer to this approach as the "Product Limit" approach. Details on the

[13] ↑The longer sessions are the High Discount Factor Sessions, which have longer periods on average.

[14] ↑The power analyses in Appendix C.4 suggest that our experiment is well-powered with this number of subjects.

[15] ↑Other cutoffs can be used. These results are robust for a cutoff of 250, that is, the period lasts at least 250 ticks and we use observations where the state is bad or the first reward occurs after 249 ticks.

[16] ↑The Product Limit estimator has commonly been used to deal with censoring in continuous time experiments (e.g. Oprea et al., 2009; S. T. Anderson et al., 2010; Calford and Oprea, 2017).

Table 3.3. Summary Statistics for the Experiment

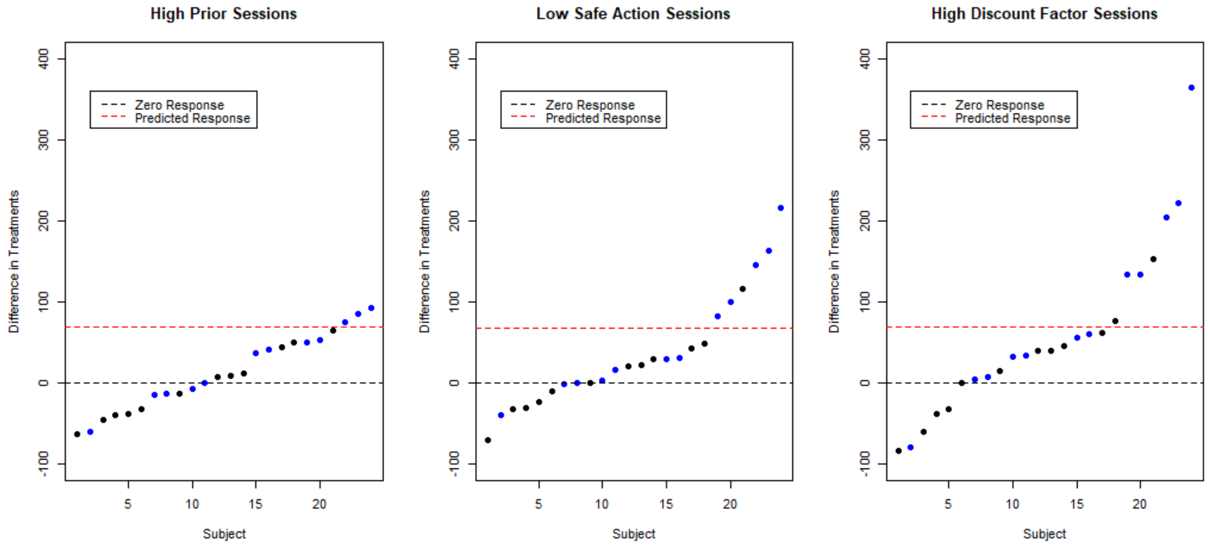
Treatment	Prediction	High Prior Session	Low Safe Session	High Discount Session
Baseline	130	86.1 [81.0]	93.4 [97.9]	90.7 [100.5]
High Prior	199	96.3 [96.9]	—	—
Low Safe Action	198	—	131.0 [129.8]	—
High Discount Factor	199	—	—	143.1 [170.6]
Difference in Means	—	10.2 [15.9]	37.6 [31.9]	52.4 [70.1]
Observations	—	263 [24]	298 [24]	396 [24]

Numbers without square brackets refer to the pooled averages from the Subset approach.
Numbers inside of square brackets refer to the average of Product Limit estimated subject means.

Product Limit estimator can be found in Appendix C.2. For hypothesis testing, we use the Product Limit estimator to create an appropriate mean for each subject and test these independent means using bootstrapped regressions. We bootstrap because it allows us to use the cardinal information contained in the data without making any assumptions about the distribution of the data (see page 52 of Moffatt (2016)). As our hypotheses require that we frequently test point predictions, it is especially important to use cardinal information. We use 9,999 bootstrap samples for all of our bootstrapping as suggested by MacKinnon (2002).

We use both approaches in tandem to test the three hypotheses as they both have strengths and weaknesses. The Subset approach only uses a subset of the data and includes some censored observations, but allows us to use common panel data econometric techniques.[17] The Product Limit approach only uses subject means, but corrects for censored data. Unless otherwise stated, we use the Subset approach for the following analysis.

Our experimental design allows us to test for order effects as we reversed the treatment order for half of our subjects. Our tests show that there is no statistically significant order effect in each treatment (shown in Appendix C.2). The elicited demographics can also be used as controls. For the following analysis, we ignore demographics for ease of exposition as our results are similar with or without demographic controls (also shown in Appendix C.2).



Blue dots denote subjects who had the Baseline treatment first, while black dots denote subjects who had the Baseline Treatment second. Different subjects are displayed for different sessions.

Figure 3.3. Difference in Average Stopping Time in the Baseline Treatment and the Session's Other Treatment

3.4.1 Hypothesis 1

Hypothesis 1 states that subjects become willing to experiment longer when p_0 increases, δ increases, or $s\Delta$ decreases. Table 3.3 displays summary statistics from both the Subset approach and the Product Limit approach. The summary statistics from both approaches suggest that subjects become willing to experiment longer when p_0 increases, δ increases, or $s\Delta$ decreases.

Figure 3.3 displays the difference in each subject's average stopping time in the Baseline treatment and their average stopping time in the session's other treatment. Fourteen out of twenty-four subjects (58.3%) in the High Prior sessions have a higher mean stopping time in the High Prior treatment than the Baseline treatment. Sixteen out of twenty-four subjects (66.7%) in the Low Safe Action sessions have a higher mean stopping time in the Low Safe Action treatment than the Baseline treatment. Eighteen out of twenty-four subjects (75.0%) in the High Discount Factor sessions have a higher mean stopping time in the High Discount

[17]↑Six percent of the data in the Subset approach is censored. The censoring for each treatment is shown in Appendix C.2.

Factor treatment than the Baseline treatment. While Table 3.3 and Figure 3.3 suggest that Hypothesis 1 holds, we will now conduct more formal analysis.

Hypothesis 1 is tested under both the Subset approach and the Product Limit approach. For each type of session, we run a regression, with subject level clustering and subject level random effects, of the stopping time on the treatment.[18] The effect of increasing the prior is positive and significant at the ten percent level (p-value=0.083).[19] Both the effect of increasing the discount factor and decreasing the safe action are positive and significant at the one percent level. The Product Limit approach is consistent with these results (the effect of increasing the prior has a p-value of 0.065).[20]

Result 1: *Subjects become willing to experiment longer when p_0 increases, δ increases, or $s\Delta$ decreases (evidence supporting Hypothesis 1).*

3.4.2 Hypothesis 2

Hypothesis 2 states that the length of time that subjects are willing to experiment increases by the correct magnitude when $s\Delta$, δ, or p_0 is changed. However, Table 3.3 suggests that Hypothesis 2 does not hold. The response to a change in p_0 appears to be less than twenty ticks. The response to a change in $s\Delta$ appears to be less than forty ticks.

Figure 3.3 shows how subjects respond to a change in p_0, δ, and $s\Delta$. Figure 3.3 plots the difference in each subject's mean stopping time for the baseline treatment and their session's other treatment. Twenty-one out of twenty-four subjects (87.5%) have a difference that is smaller than predicted in the High Prior sessions. Eighteen out of twenty-four subjects (75.0%) have a difference that is smaller than predicted in the Low Safe Action sessions. Seventeen out of twenty-four subjects (70.8%) have a difference that is less than predicted in the High Discount Factor sessions. While Table 3.3 and Figure 3.3 suggest that Hypothesis 2 does not hold, we will now conduct more formal analysis.

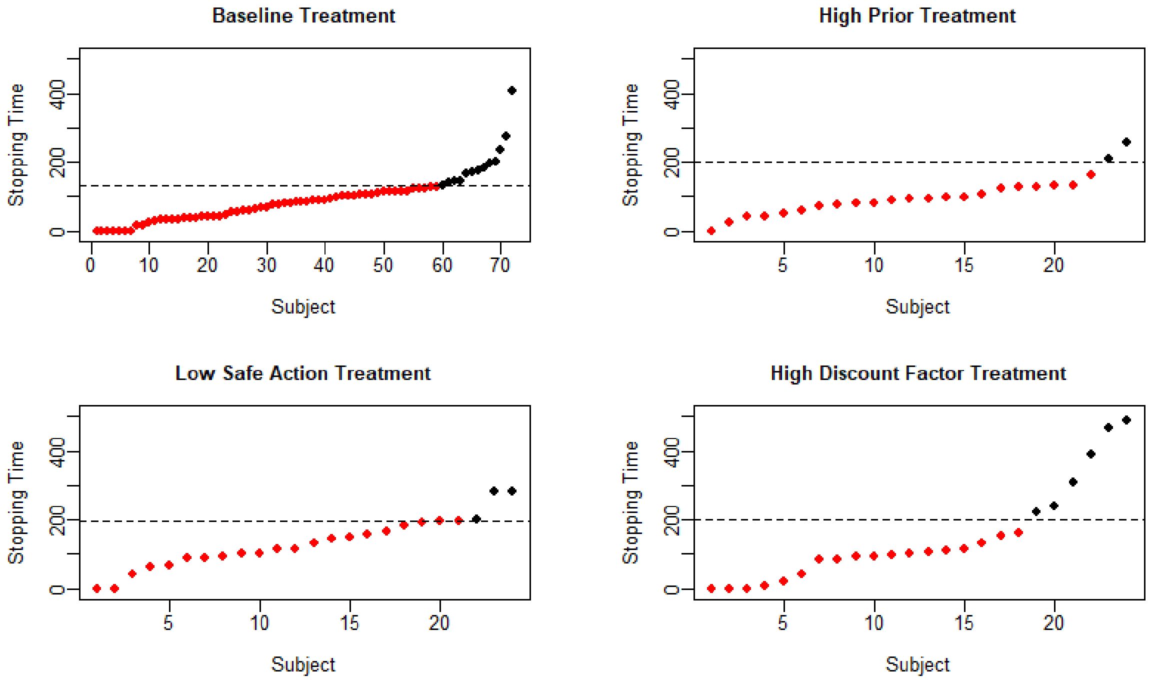
[18]↑Our results are robust to using fixed effects instead of random effects when applicable.
[19]↑The reported p-values are derived from conservative two-sided tests.
[20]↑We ran bootstrapped regressions for each type of session. For example, for the change in the prior, we regressed the difference in each High Prior session subject's Product Limit estimated mean stopping time for the Baseline treatment and the High Prior treatment on a constant.

Hypothesis 2 is tested under both the Subset approach and the Product Limit approach. For each variable, a bootstrapped regression is run of the difference between each subject's response to the treatment variable and their predicted response on a constant. A subject's response to the treatment variable is once again calculated using the difference in a subject's mean stopping time in the Baseline treatment and their session's other treatment. Subjects' response to an increase in the prior is significantly less than the predicted response at the one percent level. Subjects' response to an increase in the safe action is significantly less than the predicted response at the five percent level (p-value=0.022). Subjects' response to an increase in the discount factor is insignificantly different than the predicted response at the ten percent level (p-value=0.608). The Product Limit approach is conducted similarly and is consistent with these results.

The High Prior Sessions can be used to test the assumption that cutoff beliefs are independent of the prior. This test is important as this assumption is commonly made in models of exploration. This assumption will fail if subjects imperfectly Bayesian update and/or non-linearly weight the prior. We test this assumption using only the Subset approach.[21] In the High Prior sessions, the average cutoff belief (assuming no mis-perception of beliefs) is 0.19 in the Baseline treatment and 0.30 in the High Prior treatment. This difference is significant at the one percent level using a regression, that is clustered at the subject level and has subject level random effects, of the cutoff belief on the treatment. This result suggests that subjects are exhibiting base rate neglect/conservatism and/or non-linear probability weighting.

Result 2: *The length of time that subjects are willing to experiment increases by less than predicted when p_0 increases or when $s\Delta$ decreases (evidence against Hypothesis 2). Cutoff beliefs are not independent of the prior.*

[21]↑The Product Limit estimator can not easily be used to analyze cutoff beliefs as the Baseline and High Prior treatments have different prior beliefs and beliefs evolve non-linearly.



Red dots denote a mean stopping time lower than the prediction. Black dots denote a mean stopping time greater than or equal to the prediction.

Figure 3.4. Mean Subject Stopping Times in Each Treatment

3.4.3 Hypothesis 3

Hypothesis 3 states that the length of time that subjects are willing to experiment is as predicted in each treatment. It is clear that Hypothesis 3 does not hold as Hypothesis 2 does not hold. However, we use this subsection to see if subjects systematically under-experiment or over-experiment. Table 3.3 suggests that subjects are often willing to experiment for a shorter period of time than predicted by theory. Subjects in the Baseline, High Prior, Low Safe Action, and High Discount Factor treatments appear to under-experiment by at least 30, 100, 65, and 25 ticks, respectively.

Figure 3.4 compares, in each treatment, the stopping time of each subject to the predicted stopping time. Fifty-nine out of seventy-two subjects (81.9%) have an average stopping time below the prediction in the Baseline treatment. Twenty-two out of twenty-four subjects (91.7%) have an average stopping time below the prediction in the High Prior treatment. Twenty-one out of twenty-four subjects (87.5%) have an average stopping time below the prediction in the Low Safe Action Treatment. Eighteen out of twenty-four subjects (75.0%) have an average stopping time below the prediction in the High Discount Factor treatment, although a few subjects have a much higher average stopping time than predicted. While Table 3.3 and Figure 3.4 suggest that subjects under-experiment, we will now conduct more formal analysis.

We conduct statistical tests to investigate whether subjects under-experiment. We run regressions, with subject level clustering and subject level random effects, of the difference between each stopping time and its prediction on a constant. Subjects are overall willing to experiment for a shorter period of time than predicted by theory at the one percent level. Additionally, subjects are willing to experiment for a shorter period of time than predicted by theory at the one percent level in the Baseline, High Prior, and Low Safe Action treatments. Subjects are willing to experiment for a shorter period of time than predicted by theory at the ten percent level for the High Discount Factor treatment (p-value=0.068). The Product Limit approach is consistent with these results except for the High Discount Factor treatment, where subjects are willing to experiment for an insignificantly different period of time than predicted (p-value=0.484).[22]

While subjects are often willing to experiment for a shorter period of time than predicted by theory, they may be willing to experiment for a longer period of time than myopia predicts. A myopic subject is predicted to be willing to experiment for 5 ticks in the Baseline and High Discount Factor treatments and for 74 ticks in the High Prior and Low Safe Action treatments. Under both approaches, the mean stopping time in each treatment is above the myopic prediction. This result is tested using both approaches. A regression, that is clustered at the subject level and has subject level random effects, of the difference between each stopping time and the myopic prediction on a constant is run. Subjects are overall willing to experiment for a longer period of time than myopia predicts at the one percent level. Additionally, subjects are willing to experiment for a longer period of time than myopia predicts at the one percent level in the Baseline, Low Safe Action, and the High Discount Factor treatments. Subjects are willing to experiment for a longer period of time

[22]↑In each treatment, we ran a bootstrapped regression of the difference of subjects' Product Limit estimated mean stopping time and the predicted stopping time on a constant.

than myopia predicts at the five percent level in the High Prior Treatment (p-value=0.037). The Product Limit approach is consistent with these results except that subjects are only willing to experiment for a longer period of time than myopia predicts at the ten percent level in the High Prior Treatment (p-value=0.054).[23]

Result 3: *Subjects are overall willing to experiment for a shorter period of time than predicted by theory (evidence against Hypothesis 3). However, subjects are overall willing to experiment for a longer period of time than predicted by myopic behavior.*

3.5 Estimating Behavioral Factors

The previous section shows that subject behavior deviates from theory. In this section, we estimate a model to better understand subject behavior. The goal of this section is to determine whether subject behavior is consistent with risk preferences, base rate neglect/conservatism, and non-linear probability weighting.

3.5.1 Setup and Estimation

We focus on three possible deviations from theory: risk preferences, base rate neglect/conservatism, and non-linear probability weighting. We include risk preferences as risk is inherent in the risky action. We include non-linear probability weighting as subjects encounter various probabilities (reward probability, prior probability, random termination probability) in this experiment. Lastly, we include base rate neglect/conservatism as subjects may imperfectly Bayesian update and thus weight the information generated from experimentation too much or too little relative to the prior.

These deviations from theory are incorporated into both belief updating and the cutoff belief. Base rate neglect/conservatism and non-linear probability weighting are incorporated into belief updating. Base rate neglect/conservatism is modeled as a subject treating a tick

[23]↑In each treatment, we ran a bootstrapped regression of the difference of subjects' Product Limit estimated mean stopping time and the myopic stopping time on a constant.

of experimentation as if it is ψ ticks of experimentation.[24] Non-linear probability weighting is modeled using the Prelec-I function. Let $\tilde{p_0} = \mathrm{e}^{-(-ln(p_0))^{\alpha}}$ be the weighted value of p_0. Let $\widetilde{\lambda\Delta} = \mathrm{e}^{-(-ln(\lambda\Delta))^{\alpha}}$ be the weighted value of $\lambda\Delta$. A subject's belief updating function, in the absence of a reward, can now be modeled as

$$p = \frac{\tilde{p_0}(1-\widetilde{\lambda\Delta})^{\psi\frac{t}{\Delta}}}{\tilde{p_0}(1-\widetilde{\lambda\Delta})^{\psi\frac{t}{\Delta}} + (1-\tilde{p_0})}.$$

Notice that belief updating is in discrete time as we are focusing on the discrete approximation.

Risk preferences and non-linear probability weighting are incorporated into the cutoff belief. Risk preferences are modeled using CRRA utility. Let $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$ be a subject's utility function, with γ being the coefficient of risk aversion. Let $\tilde{\delta} = 1 - \mathrm{e}^{-(-ln(1-\delta))^{\alpha}}$ be the weighted value of the discount factor. A subject's cutoff belief is obtained through value function iteration, where

$$v(p) = max\left\{\frac{u(s)}{1-\tilde{\delta}}\ ,\ p\widetilde{\lambda\Delta} * (u(h) + \tilde{\delta} * \frac{\widetilde{\lambda\Delta}u(h)}{1-\tilde{\delta}}) + (1-p\widetilde{\lambda\Delta}) * \tilde{\delta} * v(p)\right\},$$

is the value function and p is the updated belief. The first part of the maximand is the value of stopping at this belief and the second part is the value of implementing the risky action at the current belief. The time that a subject is willing to experiment is found by finding where the belief updating function and cutoff belief intersect.

The parameters γ, ψ, and α are estimated through Maximum Likelihood Estimation. Let the prediction based solely on the parameter set and these parameters be denoted as $pred(set, \gamma, \psi, \alpha)$, where set is the parameter set a subject is facing. We assume that subjects make normally distributed errors around this prediction. A subject's willingness to experiment, in a given period, is thus given by

$$pred(set, \gamma, \psi, \alpha) + \epsilon_{\mathrm{i},t},$$

[24]↑As mentioned earlier, this is a similar approach as taken in Goeree et al. (2007) and Moreno and Rosokha (2016).

where $\epsilon_{\mathrm{i},t} \sim N(0, \sigma^2)$.

The model estimated in this section resembles a two-limit Tobit based on the aforementioned prediction. Let $pcensor_{\mathrm{i},t}$ denote the time of first possible censoring, which is either the time that the period ends or, if relevant, the minimum of the time that the period ends and the time of the first reward. The probability of a subject switching to the safe action before the first tick is equal to

$$\Phi\left(\frac{-pred(set, \gamma, \psi, \alpha)}{\sigma}\right).$$

The probability that a subject switches to the safe action at a time $0 < y_{\mathrm{i},t} < pcensor_{\mathrm{i},t}$ is equal to

$$\frac{1}{\sigma}\phi\left(\frac{y_{\mathrm{i},t} - pred(set, \gamma, \psi, \alpha)}{\sigma}\right).$$

The probability that a subject is censored (from above) in a period is equal to

$$1 - \Phi\left(\frac{pcensor_{\mathrm{i},t} - pred(set, \gamma, \psi, \alpha)}{\sigma}\right).$$

The joint density for subject i can thus be written as

$$\begin{aligned} L_{\mathrm{i}} = \prod_{t=1}^{T} & \left[\Phi\left(\frac{-pred(set, \gamma, \psi, \alpha)}{\sigma}\right)\right]^{I_{y_{\mathrm{i},t}=0}} \\ \times & \left[\frac{1}{\sigma}\phi\left(\frac{y_{\mathrm{i},t} - pred(set, \gamma, \psi, \alpha)}{\sigma}\right)\right]^{I_{0<y_{\mathrm{i},t}<pcensor_{\mathrm{i},t}}} \\ \times & \left[1 - \Phi\left(\frac{pcensor_{\mathrm{i},t} - pred(set, \gamma, \psi, \alpha)}{\sigma}\right)\right]^{I_{y_{\mathrm{i},t}=pcensor_{\mathrm{i},t}}}. \end{aligned}$$

The log-likelihood is written as $LogL = \sum_{\mathrm{i}=1}^{n} lnL_{\mathrm{i}}$. The model uses all of the data (2880 observations) as this model can account for censoring. We estimate the parameters through Maximum Likelihood Estimation.

The maximized log-likelihood is equal to 8658.03. Table 3.4 displays the estimated parameters from the estimation. The values of γ, ψ, and α are 0.36, 0.13, and 0.65, respectively. The value of σ is 166.89. The value of γ is significantly different than the restricted value of zero at the one percent level using a likelihood ratio test (restricted log-likelihood is equal to

Table 3.4. Estimates from the Structural Model

Parameter	Non-Behavioral Value	Estimated Value
γ	0.00	0.36
		[0.00]
ψ	1.00	0.13
		[0.02]
α	1.00	0.65
		[0.00]
Log-likelihood	—	8658.03

'Non-behavioral value' refers to the restriction based on the theory presented in Section 3.2. P-values from likelihood ratio tests are in square brackets.

8666.27). The value of ψ is significantly different than the restricted value of one at the five percent level using a likelihood ratio test (restricted log-likelihood is equal to 8660.73). The value of α is significantly different than the restricted value of one at the one percent level using a likelihood ratio test (restricted log-likelihood is equal to 8662.00).[25] These results suggest that risk aversion, conservatism, and non-linear probability weighting influence subjects' experimentation decisions.[26] In Appendix C.3, we show that these results are robust to various other specifications.[27]

3.5.2 Effects of Each Behavioral Factor

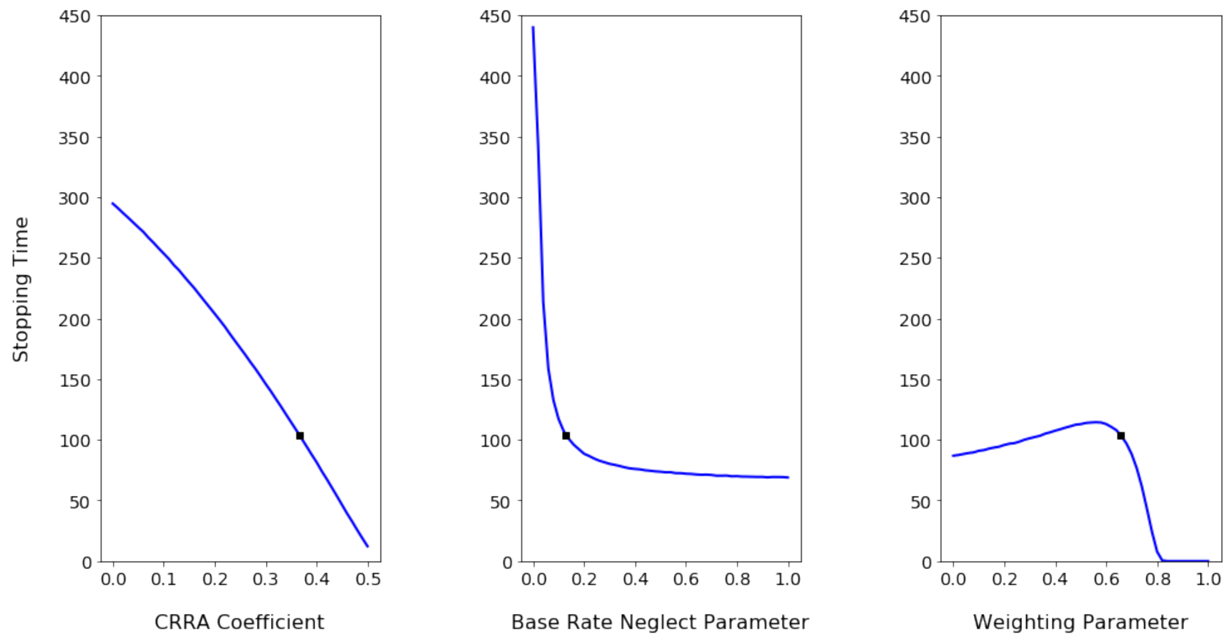
In this subsection, we explore the effects of each behavioral factor. Figure 3.5 displays the estimated model predictions for the Baseline Treatment as each behavioral factor is varied. These effects are similar in other treatments and are shown in Appendix C.3. These mean predictions are obtained by using the Subset approach on one million period simulations.

The first graph of Figure 3.5 displays the effect of risk aversion as the CRRA coefficient is varied (ψ and α are held at their estimated values). The graph shows that turning the

[25]↑Standard errors are consistent with these results except that α is only significant at the ten percent level. We prefer likelihood ratio tests as we obtain our standard errors through numerical differentiation. These standard errors are 0.15 for γ, 0.14 for ψ, 0.20 for α, and 3.69 for σ.

[26]↑While conservatism may seem counter-intuitive, we would expect it in this environment as there are large samples of new information (Griffin & Tversky, 1992).

[27]↑We estimate a model that allows for non-linear probability weighting to differ between the prior probability and the tick-based probabilities and a model that replaces the Prelec-I function with a Tversky and Kahneman (1992) probability weighting function.



Predictions are obtained through simulation and are for the Subset approach. In each graph, one behavioral factor is varied, while the other two behavioral factors are held constant at the estimated levels of the fully unrestricted structural model. The black square denotes the prediction of the fully estimated model in Section 3.5.1.

Figure 3.5. Estimated Model Predictions for the Baseline Treatment as each Behavioral Factor is Varied

risk aversion channel ($\gamma = 0$) off in the estimated model leads to more experimentation. This suggests that risk aversion is contributing to under-experimentation. The second graph displays the effect of conservatism as the base rate neglect parameter ψ is varied (γ and α are held constant at their estimated values). The graph shows that turning the conservatism channel off ($\psi = 1$) in the estimated model leads to less experimentation. This suggests that conservatism makes subjects willing to experiment longer. Lastly, the third graph displays the effect of probability weighting as α is varied (γ and ψ are held at their estimated values). The graph shows that turning the non-linear probability weighting channel off ($\alpha = 1$) in the estimated model leads to less experimentation. This suggests that non-linear probability weighting makes subjects willing to experiment longer. As risk aversion is the only factor that results in less exploration, it appears to be the predominant factor explaining under-exploration in the experiment.

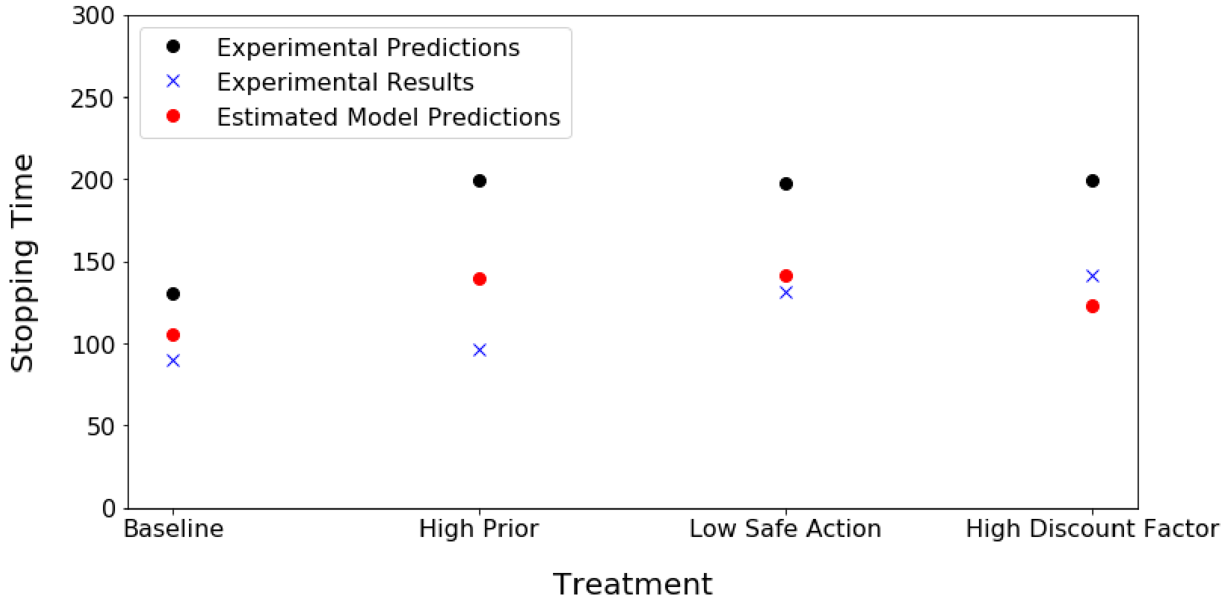


Figure 3.6. Estimated Model Predictions for each Treatment using the Subset Approach

3.5.3 Estimated Model Predictions

Figure 3.6 compares our estimated model predictions to the experimental predictions. We conduct this exercise in order to make sure that we are not fitting some treatments well at the expense of other treatments. These mean predictions are obtained by using the Subset approach on one million period simulations. The estimated model predicts an average stopping time of 105.88 ticks in the Baseline treatment, 139.10 ticks in the High Prior Treatment, 141.24 ticks in the Low Safe Action Treatment, and 122.77 ticks in the High Discount Factor treatment. As Figure 3.6 shows, the estimated model predictions outperform the experimental predictions for each treatment in the experiment.

The estimated model can also make out-of-sample predictions for a previous study. Hudja (2019) reports an average stopping time for the last fifteen periods in his sole single-agent exponential bandit treatment. The experimental prediction for this treatment is for subjects to stop after 187 ticks without observing a reward. The average stopping time that the paper reports is 135.3 ticks. Using the experimental parameters from Hudja (2019), our estimated model predicts an average stopping time of 128.3 ticks. Thus, the prediction from our estimated model is within ten ticks of the empirical average, while the experimental prediction is more than fifty ticks higher.

3.6 Conclusion

This research uses a laboratory experiment to analyze how individuals resolve an exploration versus exploitation trade-off. We analyze the predictions of the single-agent exponential bandit model in a laboratory experiment. We find that, as predicted, subjects respond to changes in the discount factor, safe action, and prior belief. However, we commonly find that subjects experiment less than predicted. Maximum Likelihood Estimation suggests that subjects' behavior is consistent with risk aversion, conservatism, and non-linear probability weighting.

These results have consequences for experimentation outside of the laboratory and for theory. The finding that subjects respond to changes in experimentation incentives suggests that institutions can increase experimentation by changing incentives. Specifically, it suggests that incentives like grant money can increase experimentation by reducing the cost of experimentation. The frequent under-experimentation found throughout this study suggests that individuals sub-optimally experiment. Lastly, our model suggests that subjects' behavior is consistent with risk aversion, conservatism and non-linear probability weighting. Our model suggests that theorists should consider these behavioral factors when modeling experimentation environments.

There are many avenues for future research. One possible avenue is that experiments could be conducted in more complicated exploration environments such as innovation contests and principal-agent problems. Our research suggests behavioral factors that can be used to motivate alternative hypotheses in these environments. Another avenue for future research is that experiments can help select possible interventions to mitigate under-experimentation. For example, we have suggested that risk aversion is a major contributor to under-experimentation. Since risk preferences can differ in the loss domain, a possible intervention would be to re-frame this problem as a loss. A future experiment could investigate whether this is effective in increasing experimentation.

REFERENCES

Abdallah, M., Naghizadeh, P., Cason, T., Bagchi, S., & Sundaram, S. (2019). Protecting Assets with Heterogeneous Valuations under Behavioral Probability Weighting, In *2019 ieee conference on decision and control (cdc)*.

Abdallah, M., Naghizadeh, P., Hota, A. R., Cason, T., Bagchi, S., & Sundaram, S. (2019). The Impacts of Behavioral Probability Weighting on Security Investments in Interdependent Systems, In *2019 american control conference (acc)*, Philadelphia.

Abeler, J., Falk, A., Goette, L., & Huffman, D. (2011). Reference points and effort provision. *American Economic Review*, *101*(2), 470–492. https://doi.org/10.1257/aer.101.2.470

Acemoglu, D., Malekian, A., & Ozdaglar, A. (2016). Network security and contagion. *Journal of Economic Theory*, *166*, 536–585. https://doi.org/10.1016/j.jet.2016.09.009

Acquisti, A., & Grossklags, J. (2007). What can behavioral economics teach us about privacy (A. Acquisti, S. Gritzalis, C. Lambrinoudakis, & S. di Vimercati, Eds.). In A. Acquisti, S. Gritzalis, C. Lambrinoudakis, & S. di Vimercati (Eds.), *Digital privacy: Theory, technologies and practices*. Auerbach Publications.

Alaba, F. A., Othman, M., Hashem, I. A. T., & Alotaibi, F. (2017). Internet of Things security: A survey. *Journal of Network and Computer Applications*, *88*, 10–28. https://doi.org/10.1016/J.JNCA.2017.04.002

An, B., Brown, M., Vorobeychik, Y., & Tambe, M. (2013). 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)*.

Anderson, C. (2001). *Behavioral Models of Strategies in Multi-Armed Bandit Problems* (Doctoral dissertation). California Institute of Technology.

Anderson, C. (2012). Ambiguity aversion in multi-armed bandit problems. *Theory and Decision*, *72*(1), 15–33. https://doi.org/10.1007/s11238-011-9259-2

Anderson, S. T., Friedman, D., & Oprea, R. (2010). Preemption games: Theory and experiment. *American Economic Review*, *100*(4), 1778–1803. https://doi.org/10.1257/aer.100.4.1778

Au, P. H., & Li, K. K. (2018). Bayesian Persuasion and Reciprocity: Theory and Experiment. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3191203

Banks, J., Olson, M., & Porter, D. (1997). An experimental analysis of the bandit problem. *Economic Theory*, *10*, 55–77.

Banovetz, J., & Oprea, R. (2020). *Complexity and Procedural Choice.*

Benartzi, S., & Thaler, R. H. (2001). Naive Diversification Strategies in Defined Contribution Savings Plans. *The American Economic Review*, *91*(1), 79–98. https://www.jstor.org/stable/2677899

Benhabib, J., Bisin, A., & Schotter, A. (2010). Present-bias, quasi-hyperbolic discounting, and fixed costs. *Games and Economic Behavior*, *69*(2), 205–223. https://doi.org/10.1016/j.geb.2009.11.003

Bier, V., Oliveros, S., & Samuelson, L. (2007). Choosing what to protect: Strategic defensive allocation against an unknown attacker. *Journal of Public Economic Theory*, *9*(4), 563–587.

Bimpikis, K., Ehsani, S., & Mostagir, M. (2019). Designing dynamic contests. *Operations Research*, *67*(2), 339–356. https://doi.org/10.1287/opre.2018.1823

Bleichrodt, H., & Pinto, J. L. (2000). A Parameter-Free Elicitation of the Probability Weighting Function in Medical Decision Analysis. *Management Science*, *46*(11), 1485–1496. https://doi.org/10.1287/mnsc.46.11.1485.12086

Bloch, F., Dutta, B., & Dziubinski, M. (2020). A game of hide and seek in networks. http://arxiv.org/abs/2001.03132

Boche, H., Naik, S., & Alpcan, T. (2011). Characterization of convex and concave resource allocation problems in interference coupled wireless systems. *IEEE Transactions on Signal Processing*, *59*(5), 2382–2394.

Bolton, P., & Harris, C. (1999). Strategic Experimentation. *Econometrica*, *67*(2), 349–374.

Bonatti, A., & Hörner, J. (2011). Collaborating. *American Economic Review*, *101*(2), 632–663. https://doi.org/10.1257/aer.101.2.632

Brandts, J., & Charness, G. (2000). Hot vs. Cold: Sequential Responses and Preference Stability in Experimental Games. *Experimental Economics*, *2*, 227–238.

Bruhin, A., Fehr-Duda, H., & Epper, T. (2010). Risk and Rationality: Uncovering Heterogeneity in Probability Distortion. *Econometrica*, *78*(4), 1375–1412. https://doi.org/10.3982/ECTA7139

Burdett, K., & Judd, K. (1983). Equilibrium Price Dispersion. *Econometrica*, *51*(4), 955–969.

Butera, L., Grossman, P. J., Houser, D., List, J. A., & Villeval, M. C. (2020). *A New Mechanism to Alleviate the Crises of Confidence in Science With An Application to the Public Goods Game*, GATE Lyon Saint-Etienne.

Calford, E., & Oprea, R. (2017). Continuity, Inertia, and Strategic Uncertainty: A Test of the Theory of Continuous Time Games. *Econometrica*. https://doi.org/10.3982/ECTA14346

Callen, M., Isaqzadeh, M., Long, J. D., & Sprenger, C. (2014). Violence and Risk Preference: Experimental Evidence from Afghanistan. *American Economic Review*, *104*(1), 123–148. https://doi.org/10.1257/aer.104.1.123

Camerer, C. F., Dreber, A., Forsell, E., Ho, T. H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., & Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, *351*(6280), 1433–1436. https://doi.org/10.1126/science.aaf0918

Camerer, C. F., Dreber, A., & Johannesson, M. (2019). Replication and other practices for improving scientific quality in experimental economics (A. Schram & A. Ule, Eds.). In A. Schram & A. Ule (Eds.), *Handbook of research methods and applications in experimental economics*. Edward Elgar Publishing.

Caplin, A., & Leahy, J. (2001). Psychological Expected Utility Theory and Anticipatory Feelings. *The Quarterly Journal of Economics*, *116*(1), 55–79. https://doi.org/10.1162/003355301556347

Caplin, A., & Leahy, J. (2004). The Supply of Information by a Concerned Expert. *The Economic Journal*, *114*(497), 487–505. https://doi.org/10.1111/j.0013-0133.2004.0228a.x

Capra, C. M. (2004). Mood-Driven Behavior in Strategic Interactions. *The American Economic Review: Papers and Proceedings*, *94*(2), 367–372.

Casari, M., & Cason, T. N. (2009). The strategy method lowers measured trustworthy behavior. *Economics Letters*, *103*(3), 157–159. https://doi.org/10.1016/j.econlet.2009.03.012

Cason, T. N., Friedman, D., & Hopkins, E. (2021). An Experimental Investigation of Price Dispersion and Cycles. *Journal of Political Economy*, *129*(3).

Chapman, J., Snowberg, E., Wang, S., & Camerer, C. (2018). *Loss Attitudes in the U.S. Population: Evidence from Dynamically Optimized Sequential Experimentation (DOSE)* (tech. rep.). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w25072

Charness, G., & Dufwenberg, M. (2006). Promises and Partnership. *Econometrica*, *74*(6), 1579–1601.

Charness, G., & Dufwenberg, M. (2011). Participation. *American Economic Review*, *101*(4), 1211–1237. https://doi.org/10.1257/aer.101.4.1211

Charness, G., Garcia, T., Offerman, T., & Villeval, M. C. (2020). Do measures of risk attitude in the laboratory predict behavior under risk in and outside of the laboratory? *Journal of Risk and Uncertainty*, *60*(2), 99–123. https://doi.org/10.1007/s11166-020-09325-6

Charness, G., & Rabin, M. (2002). Understanding Social Preferences with Simple Tests. *The Quarterly Journal of Economics*, *117*(3), 817–869. https://www.jstor.org/stable/4132490

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. https://doi.org/10.1016/J.JBEF.2015.12.001

Chen, R., & Chen, Y. (2011). The potential of social identity for equilibrium selection. *American Economic Review*, *101*(6), 2562–2589. https://doi.org/10.1257/aer.101.6.2562

Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, *99*(1), 431–457. https://doi.org/10.1257/aer.99.1.431

Choi, S., Kim, J., Lee, E., & Lee, J. (2018). *Probability Weighting and Cognitive Ability*.

Chowdhury, S. M. (2019). The Attack and Defense Mechanisms - Perspectives from Behavioral Economics and Game Theory. *Behavioral and Brain Sciences*, *42*, e121. https://doi.org/10.1017/S0140525X19000815

Chowdhury, S. M., Kovenock, D., Rojo Arjona, D., & Wilcox, N. T. (2016). *Focality and Asymmetry in Multi-battle Contests*. https://digitalcommons.chapman.edu/esi_working_papers/194/

Chowdhury, S. M., Kovenock, D., & Sheremeta, R. M. (2013). An experimental investigation of Colonel Blotto games. *Economic Theory*, *52*(3), 833–861. https://doi.org/10.1007/s00199-011-0670-2

Clark, D. J., & Konrad, K. A. (2007). Asymmetric Conflict: Weakest Link against Best Shot. *Journal of Conflict Resolution*, *51*(3), 457–469. https://doi.org/10.1177/0022002707300320

Cox, J. C., Smith, V. L., & Walker, J. M. (1988). Theory and individual behavior of first-price auctions. *Journal of Risk and Uncertainty*, *1*(1), 61–99. https://doi.org/10.1007/BF00055565

Cox, J. C., Smith, V. L., & Walker, J. M. (1992). Theory and Misbehavior of First-Price Auctions: Comment. *The American Economic Review*, *82*(5), 1392–1412.

Curley, S. P., Yates, J. F., & Abrams, R. A. (1986). Psychological Sources of Ambiguity Avoidance. *Organizational Behavior and Human Decision Processes*, *38*(2), 230–256.

Davis, D. D., & Holt, C. A. (1993). *Experimental Economics*. Princeton University Press.

De Clippel, G., Eliaz, K., & Knight, B. (2014). On the selection of arbitrators. *American Economic Review*, *104*(11), 3434–3458. https://doi.org/10.1257/aer.104.11.3434

Deck, C., & Sheremeta, R. M. (2012). Fight or Flight?: Defending against Sequential Attacks in the Game of Siege. *Journal of Conflict Resolution*, *56*(6), 1069–1088. https://doi.org/10.1177/0022002712438355

Dhami, S. (2016). *The Foundations of Behavioral Economic Analysis*. Oxford University Press.

Dighe, N. S., Zhuang, J., & Bier, V. M. (2009). Secrecy in defensive allocations as a strategy for achieving more cost-effective attacker detterrence. *International Journal of Performability Engineering*, *5*(1), 31–43.

Djawadi, B. M., Endres, A., Hoyer, B., & Recker, S. (2019). Network formation and disruption - An experiment are equilibrium networks too complex? *Journal of Economic Behavior and Organization*, *157*, 708–734. https://doi.org/10.1016/j.jebo.2018.11.004

Drouvelis, M., & Grosskopf, B. (2016). The effects of induced emotions on pro-social behaviour. *Journal of Public Economics*, *134*, 1–8. https://doi.org/10.1016/j.jpubeco.2015.12.012

Duvendack, M., Palmer-Jones, R., & Reed, W. R. (2017). What is meant by "Replication" and why does it encounter resistance in economics?, In *American economic review: Papers & proceedings*, American Economic Association. https://doi.org/10.1257/aer.p20171031

Dziubiński, M., & Goyal, S. (2013). Network design and defence. *Games and Economic Behavior*, *79*(1), 30–43. https://doi.org/10.1016/j.geb.2012.12.007

Dziubiński, M., & Goyal, S. (2017). How do you defend a network? *Theoretical Economics*, *12*(1), 331–376. https://doi.org/10.3982/te2088

Epper, T., & Fehr-Duda, H. (2018). *The Missing Link: Unifying Risk Taking and Time Discounting*, University of Zurich.

Fehr, E., Herz, H., & Wilkening, T. (2013). The lure of authority: Motivation and incentive effects of power. *American Economic Review*, *103*(4), 1325–1359. https://doi.org/10.1257/aer.103.4.1325

Fehr-Duda, H., de Gennaro, M., & Schubert, R. (2006). Gender, Financial Risk, and Probability Weights. *Theory and Decision*, *60*(2-3), 283–313. https://doi.org/10.1007/s11238-005-4590-0

Fehr-Duda, H., Epper, T., Bruhin, A., & Schubert, R. (2011). Risk and rationality: The effects of mood and decision rules on probability weighting. *Journal of Economic Behavior & Organization*, *78*(1-2), 14–24. https://doi.org/10.1016/J.JEBO.2010.12.004

Feng, S., Xiong, Z., Niyato, D., Wang, P., Wang, S. S., & Shen, X. S. (2020). Joint pricing and security investment in cloud security service market with user interdependency. *IEEE Transactions on Services Computing*.

Fielder, A., Panaousis, E., Malacaria, P., Hankin, C., & Smeraldi, F. (2014). Game theory meets information security management, In *International information security conference (ifip)*.

Fréchette, G. R. (2015). Experiments: Professionals Versus Students (G. Fréchette & A. Schotter, Eds.). In G. Fréchette & A. Schotter (Eds.), *Handbook of experimental economic methodology*. Oxford University Press.

Fréchette, G. R., Lizzeri, A., & Perego, J. (2018). *Rules and Commitment in Communication*.

Frey, B. J., & Dueck, D. (2007). Clustering by passing messages between data points. *Science*, *315*. https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.121.3145

Friedman, D. (1992). Theory and Misbehavior of First-Price Auctions: Comment. *American Economic Review*, *82*(5), 1374–1378. https://doi.org/10.2307/2117485

Friedman, D., & Oprea, R. (2012). A Continuous Dilemma. *American Economic Review*, *102*(1), 337–363. https://doi.org/10.1257/aer.102.1.337

Friedman, D., & Sunder, S. (1994). *Experimental Methods: A Primer for Economists*. Cambridge University Press.

Gartner. (2018). Gartner Forecasts Worldwide Information Security Spending to Exceed $124 Billion in 2019. https://www.gartner.com/en/newsroom/press-releases/2018-08-15-gartner-forecasts-worldwide-information-security-spending-to-exceed-124-billion-in-2019

Gelman, A., & Carlin, J. (2014). Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspectives on Psychological Science*, *9*(6), 641–651. https://doi.org/10.1177/1745691614551642

Goeree, J. K., & Holt, C. A. (2005). An experimental study of costly coordination. *Games and Economic Behavior*, *51*(2 SPEC. ISS.), 349–364. https://doi.org/10.1016/j.geb.2004.08.006

Goeree, J. K., Holt, C. A., & Palfrey, T. R. (2003). Risk averse behavior in generalized matching pennies games. *Games and Economic Behavior*, *45*(1), 97–113. https://doi.org/10.1016/S0899-8256(03)00052-6

Goeree, J. K., Palfrey, T. R., Rogers, B. W., & Mckelvey, R. D. (2007). Self-Correcting Information Cascades. *The Review of Economic Studies*, *74*(3), 733–762.

Goyal, S., & Vigier, A. (2014). Attack, Defence, and Contagion in Networks. *The Review of Economic Studies*, *81*(4), 1518–1542. https://doi.org/10.1093/restud/rdu013

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with ORSEE. *Journal of the Economic Science Association*, *1*(1), 114–125. https://doi.org/10.1007/s40881-015-0004-4

Griffin, D., & Tversky, A. (1992). The Weighing of Evidence and the Determinants of Confidence. *Cognitive Psychology*, *24*(3), 411–435.

Guan, P., He, M., Zhuang, J., & Hora, S. C. (2017). Modeling a multitarget attacker–defender game with budget constraints. *Decision Analysis*, *14*(2), 87–107.

Halac, M., Kartik, N., & Liu, Q. (2016). Optimal contracts for experimentation. *Review of Economic Studies*, *83*(3), 1040–1091. https://doi.org/10.1093/restud/rdw013

Halac, M., Kartik, N., & Liu, Q. (2017). Contests for Experimentation. *Journal of Political Economy*, *125*(5), 1523–1569.

Harrison, G. W. (1989). Theory and Misbehavior of First-Price Auctions. *The American Economic Review*, *79*(4), 749–762. https://www.jstor.org/stable/1827930

Harrison, G. W. (1992). Theory and Misbehavior of First-Price Auctions: Reply. *The American Economic Review*, *82*(5), 1426–1443.

Hoelzemann, J., & Klein, N. (2020). *Bandits in the Lab*, Université de Montréal.

Holt, C. A., & Laury, S. K. (2002). Risk Aversion and Incentive Effects. *American Economic Review*, *92*(5), 1644–1655. https://doi.org/10.1257/000282802762024700

Homer, J., Zhang, S., Ou, X., Schmidt, D., Du, Y., Rajagopalan, S. R., & Singhal, A. (2013). Aggregating vulnerability metrics in enterprise networks using attack graphs. *Journal of Computer Security*, *21*(4), 561–597. https://doi.org/10.3233/JCS-130475

Hota, A. R., Clements, A. A., Bagchi, S., & Sundaram, S. (2018). A Game-Theoretic Framework for Securing Interdependent Assets in Networks (S. Rass & S. Schauer, Eds.). In S. Rass & S. Schauer (Eds.), *Game theory for security and risk management: From theory to practice*. Springer. https://doi.org/10.1007/978-3-319-75268-6{_}7

Hota, A. R., Clements, A. A., Sundaram, S., & Bagchi, S. (2016). Optimal and Game-Theoretic Deployment of Security Investments in Interdependent Assets. Springer. https://doi.org/10.1007/978-3-319-47413-7{_}6

Hoyer, B., & Rosenkranz, S. (2018). Determinants of Equilibrium Selection in Network Formation: An Experiment. *Games*, *9*(4), 89. https://doi.org/10.3390/g9040089

Huck, S., Seltzer, A. J., & Wallace, B. (2011). Deferred compensation in multiperiod labor contracts: An experimental test of Lazear's model. *American Economic Review*, *101*(2), 819–843. https://doi.org/10.1257/aer.101.2.819

Hudja, S. (2019). *Voting for Experimentation: A Continuous Time Analysis.*

Humayed, A., Lin, J., Li, F., & Luo, B. (2017). Cyber-Physical Systems Security—A Survey. *IEEE Internet of Things Journal*, *4*(6), 1802–1831. https://doi.org/10.1109/JIOT.2017.2703172

Ifcher, J., & Zarghamee, H. (2011). Happiness and time preference: The effect of positive affect in a random-assignment experiment. *American Economic Review*, *101*(7), 3109–3129. https://doi.org/10.1257/aer.101.7.3109

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), 2–8. https://doi.org/10.1371/journal.pmed.0020124

Ioannidis, J. P., Stanley, T. D., & Doucouliagos, H. (2017). The Power of Bias in Economics Research. *Economic Journal*, *127*(605), F236–F265. https://doi.org/10.1111/ecoj.12461

Isaac, R. M., & James, D. (2000). Just Who Are You Calling Risk Averse? *Journal of Risk and Uncertainty*, *20*(2), 177–187. https://doi.org/10.1023/A:1007881008177

Jauhar, S., Chen, B., Temple, W. G., Dong, X., Kalbarczyk, Z., Sanders, W. H., & Nicol, D. M. (2015). Model-Based Cybersecurity Assessment with NESCOR Smart Grid Failure Scenarios, In *2015 ieee 21st pacific rim international symposium on dependable computing (prdc)*, IEEE. https://doi.org/10.1109/PRDC.2015.37

Kagel, J. H., & Roth, A. E. (1992). Theory and Misbehavior of First-Price Auctions: Comment. *The American Economic Review*, *82*(5), 1379–1391.

Kamenica, E. (2019). Bayesian Persuasion and Information Design. *Annual Review of Economics*, *11*(1), 080218–025739. https://doi.org/10.1146/annurev-economics-080218-025739

Kamenica, E., & Gentzkow, M. (2011). Bayesian Persuasion. *American Economic Review*, *101*(6), 2590–2615. https://doi.org/10.1257/aer.101.6.2590

Keller, G., Novák, V., & Willems, T. (2019). A note on optimal experimentation under risk aversion. *Journal of Economic Theory*, *179*, 476–487. https://doi.org/10.1016/j.jet.2018.11.006

Keller, G., Rady, S., & Cripps, M. (2005). Strategic Experimentation with Exponential Bandits. *Econometrica*, *73*(1), 39–68.

Khromenkova, D. (2015). *Collective Experimentation with Breakdowns and Breakthroughs*, University of Mannheim.

Kirchsteiger, G., Rigotti, L., & Rustichini, A. (2006). Your morals might be your moods. *Journal of Economic Behavior and Organization*, *59*(2), 155–172. https://doi.org/10.1016/j.jebo.2004.07.004

Kosfeld, M. (2004). Economic Networks in the Laboratory: A survey. *Review of Network Economics*, *3*(1).

Kőszegi, B., & Rabin, M. (2006). A Model of Reference-Dependent Preferences. *The Quarterly Journal of Economics*, *121*(4), 1133–1165. https://doi.org/10.1093/qje/121.4.1133

Kovenock, D., & Roberson, B. (2018). The Optimal Defense of Networks of Targets. *Economic Inquiry*, *56*(4), 2195–2211. https://doi.org/10.1111/ecin.12565

Kovenock, D., Roberson, B., & Sheremeta, R. M. (2019). The attack and defense of weakest-link networks. *Public Choice*, *179*(3-4), 175–194. https://doi.org/10.1007/s11127-018-0618-1

Lazear, E. P. (1979). Why Is There Mandatory Retirement? *Journal of Political Economy*, *87*(6), 1261–1284. https://www.jstor.org/stable/1833332

Lee, E. (2015). The Past, Present and Future of Cyber-Physical Systems: A Focus on Models. *Sensors*, *15*(3), 4837–4869. https://doi.org/10.3390/s150304837

Leibowitz, H., Piotrowska, A. M., Danezis, G., & Herzberg, A. (2019). No Right to Remain Silent: Isolating Malicious Mixes, In *28th usenix security symposium (usenix security 19)*, Santa Clara, CA, USENIX Association. https://www.usenix.org/conference/usenixsecurity19/presentation/leibowitz

List, J. A., Sadoff, S., & Wagner, M. (2011). So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design. *Experimental Economics*, *14*(4), 439–457. https://doi.org/10.1007/s10683-011-9275-7

Loewenstein, G. (1987). Anticipation and the Valuation of Delayed Consumption. *The Economic Journal*, *97*(387), 666. https://doi.org/10.2307/2232929

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, *151*, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

MacKinnon, J. G. (2002). Bootstrap Inference in Econometrics. *The Canadian Journal of Economics / Revue canadienne d'Economique*, *35*(4), 615–645.

Maniadis, Z., Tufano, F., & List, J. A. (2014). One swallow doesn't make a summer: New evidence on anchoring effects. *American Economic Review*, *104*(1), 277–290. https://doi.org/10.1257/aer.104.1.277

Maniadis, Z., Tufano, F., & List, J. A. (2015). How to make experimental economics research more reproducible: Lessons from other disciplines and a new proposal. *Research in Experimental Economics*, *18*, 215–230. https://doi.org/10.1108/S0193-230620150000018008

McBride, M., & Hewitt, D. (2013). The enemy you can't see: An investigation of the disruption of dark networks. *Journal of Economic Behavior & Organization*, *93*, 32–50. https://doi.org/10.1016/j.jebo.2013.07.004

McFadden, D. L. (1976). Quantal Choice Analysis: A Survey. *Annals of Economic and Social Measurement*, *5*(4).

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior*, *10*(1), 6–38. https://doi.org/10.1006/GAME.1995.1023

Merlo, A., & Schotter, A. (1992). Theory and Misbehavior of First-Price Auctions: Comment. *The American Economic Review*, *82*(5), 1413–1425.

Mersinas, K., Hartig, B., Martin, K. M., & Seltzer, A. (2016). Are information security professionals expected value maximizers?: An experiment and survey based test. *Journal of Cybersecurity*, *2*(1), 57–70. https://doi.org/10.1093/cybsec/tyw009

Meyer, R. J., & Shi, Y. (1995). Sequential Choice under Ambiguity: Intuitive Solutions to the Armed-Bandit Problem. *Management Science*, *41*(5), 817–834. https://www.jstor.org/stable/2633100

Modelo-Howard, G., Bagchi, S., & Lebanon, G. (2008). 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)*.

Moffatt, P. G. (2016). *Experimetrics : Econometrics for Experimental Economics*. Palgrave.

Moreno, O. M., & Rosokha, Y. (2016). Learning under compound risk vs. learning under ambiguity – an experiment. *Journal of Risk and Uncertainty*, *53*(2-3), 137–162. https://doi.org/10.1007/s11166-017-9250-6

Nguyen, K. C., Alpcan, T., & Basar, T. (2010). Stochastic Games for Security in Networks with Interdependent Nodes. http://arxiv.org/abs/1003.2440

Nguyen, Q. (2017). *Bayesian persuasion: Evidence from the laboratory*, Utah State University.

Nikoofal, M. E., & Zhuang, J. (2012). Robust allocation of a defensive budget considering an attacker's private information. *Risk Analysis: An International Journal*, *32*(5), 930–943.

Nithyanand, R., Starov, O., Zair, A., Gill, P., & Schapira, M. (2016). Measuring and mitigating AS-level adversaries against Tor, In *Network & distributed system security symposium (ndss)*.

Oprea, R. (2014). Survival versus Profit Maximization in a Dynamic Stochastic Experiment. *Econometrica*, *82*(6), 2225–2255. https://doi.org/10.3982/ecta10751

Oprea, R., Friedman, D., & Anderson, S. T. (2009). Learning to Wait: A Laboratory Investigation. *Review of Economic Studies*, *76*(3), 1103–1124. https://academic.oup.com/restud/article/76/3/1103/1591608

Oprea, R., Henwood, K., & Friedman, D. (2011). Separating the Hawks from the Doves: Evidence from continuous time laboratory games. *Journal of Economic Theory*, *146*(6), 2206–2225. https://doi.org/10.1016/J.JET.2011.10.014

Pal, R., & Golubchik, L. (2010). Analyzing self-defense investments in internet security under cyber-insurance coverage, In *2010 ieee 30th international conference on distributed computing systems*, IEEE.

Paté-Cornell, M. E., Kuypers, M., Smith, M., & Keller, P. (2018). Cyber Risk Management for Critical Infrastructure: A Risk Analysis Model and Three Case Studies. *Risk Analysis*, *38*(2), 226–241. https://doi.org/10.1111/risa.12844

Prelec, D. (1998). The Probability Weighting Function. *Econometrica*, *66*(3), 497. https://doi.org/10.2307/2998573

Pronin, E., & Ross, L. (2006). Temporal differences in trait self-ascription: When the self is seen as an other. *Journal of Personality and Social Psychology*, *90*(2). https://doi.org/10.1037/0022-3514.90.2.197

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, *3*(4), 323–343. https://doi.org/10.1016/0167-2681(82)90008-7

Rutström, E. E., & Wilcox, N. T. (2009). Stated beliefs versus inferred beliefs: A methodological inquiry and experimental test. *Games and Economic Behavior*, *67*(2), 616–632. https://doi.org/10.1016/j.geb.2009.04.001

Sheremeta, R. M. (2019). The attack and defense games. *Behavioral and Brain Sciences*, *42*, e140. https://doi.org/10.1017/S0140525X19000931

Sheyner, O., & Wing, J. (2003). Tools for generating and analyzing attack graphs, In *International symposium on formal methods for components and objects (fmco)*, Springer. https://doi.org/10.1007/978-3-540-30101-1{_}17

StataCorp. (2013). *Stata 13 Power and Sample Size Reference Manual*. College Station, TX, Stata Press.

Strulovici, B. (2010). Learning While Voting: Determinants of Collective Experimentation. *Econometrica*, *78*(3), 933–971. https://doi.org/10.3982/ecta8011

Sun, X., Shen, C., Chang, T.-H., & Zhong, Z. (2018). Joint resource allocation and trajectory design for UAV-aided wireless physical layer security, In *2018 ieee globecom workshops (gc wkshps)*, IEEE.

Tanaka, T., Camerer, C. F., & Nguyen, Q. (2010). Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam. *American Economic Review*, *100*(1), 557–571. https://doi.org/10.1257/aer.100.1.557

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. https://doi.org/10.1007/BF00122574

Wu, D., Xiao, H., & Peng, R. (2018). Object defense with preventive strike and false targets. *Reliability Engineering & System Safety*, *169*, 76–80.

Wu, W., & Ye, B. (2019). *Competition in Persuasion: An Experiment* (tech. rep.).

Xie, P., Li, J. H., Ou, X., Liu, P., & Levy, R. (2010). Using Bayesian networks for cyber security analysis, In *Proceedings of the international conference on dependable systems and networks (dns)*. https://doi.org/10.1109/DSN.2010.5544924

Yang, R., Kiekintveld, C., Ordonez, F., Tambe, M., & John, R. (2011). Improving resource allocation strategy against human adversaries in security games, In *Twenty-second international joint conference on artificial intelligence (ijcai)*.

Zhang, L., & Ortmann, A. (2013). *Exploring the meaning of significance in experimental economics*.

Ziliak, S. T., & McCloskey, D. N. (2004). Size matters: The standard error of regressions in the American Economic Review. *Journal of Socio-Economics*, *33*(5), 527–546. https://doi.org/10.1016/j.socec.2004.09.024

A. CHAPTER 1 APPENDIX

A.1 Example Stata and Python Code for a Normal DGP Simulation based Power Analysis

A.1.1 Stata Code

```
clear all

// num_sims = number of generated data-sets (the bigger the better)
scalar num_sims = 1000
// n - number of observations per treatment_vector
scalar num_obs = 63

// DGP params:
scalar intercept = .5
scalar tau = .25
// subscript 0, T=0, similarly 1, T=1
scalar std0 = .5
scalar std1 = .5

Python:

# importing packages
import numpy as np
import pandas as pd
# importing stata interface package,
# so that both stata and Python have access to variable 'num_sims'
from sfi import Scalar

# n - short for num_obs - reading in variable from stata
n = int( Scalar.getValue('num_obs') )
```

```
# reading in variables from stata
num_sims = int( Scalar.getValue('num_sims') )
intercept = float( Scalar.getValue('intercept') )
tau = float( Scalar.getValue('tau') )
std0 = float( Scalar.getValue('std0') )
std1 = float( Scalar.getValue('std1') )

treatment_vector = np.repeat([0,1],n*num_sims)

# set seed for consistent results, comment out or change seed otherwise
np.random.seed(1)

# drawing noise terms
noise0 = np.random.normal(0,std0,n*num_sims)
noise1 = np.random.normal(0,std1,n*num_sims)
# putting the noise together in a manner consistent with treatment vector
noise_vector = np.append(noise0,noise1)

# generating the outcome according to the specified DGP
outcome_vector = intercept + tau*treatment_vector + noise_vector

# the data-set is put into 'power blocks',
# so one data-set is where power_block==0,
# another where power_block==1, etc.
power_block_vector = np.tile(np.repeat(np.arange(num_sims),n),2)

# putting it into a format that stata will like
matrix = np.stack([treatment_vector,outcome_vector,power_block_vector]

,axis=-1)
```

```
df = pd.DataFrame(data=matrix, columns=["treatment", "outcome",

"power_block"])
df.to_stata('pa_sims.dta')

end

use "pa_sims.dta"

// so we can count the number of tests rejected
gen rejected = 0

// a is desired alpha level
local a = .05

// needed for loop
local ns = num_sims-1

forval i = 0 / `ns' {
// perform the t-test
quietly ttest outcome if power_block == `i', by(treatment)
// save the p_value
scalar p_val = r(p)
// record if the null was rejected
quietly replace rejected = 1 if p_val <= `a' & power_block == `i'
}


quietly summ rejected
// the mean of the rejected variable is the power
```

```
display as text "Power: " as result r(mean)

// // results from non-simulation ttest for comparison
scalar mean1 = intercept + tau
sampsi `=scalar(intercept)' `=scalar(mean1)', ///
 sd1(`=scalar(std0)') sd2(`=scalar(std1)') p(.8)
```

A.1.2 Discussion

Notice the lack of loops in the Python code that generates the data-sets, vectorizing these improves speed performance substantially. In practice, one should avoid loops as often as practically possible, unless there are memory issues, or if there are parallellization techniques (think multi-core processors) that can be easily implemented. Unfortunately Stata offers little opportunities for vectorization, and thus must loop over each power block to conduct the statistical test. This loop process is very parallelizable as the results of one loop do not affect the others so they could be conducted simultaneously. Stata does offer easy access to parallelization techniques through its (non-standard, costly) MP version, but does not currently support parallelization over loops. There are also community created packages for parallel computing in Stata, however, these are (understandably) much more difficult to implement for a regular user. One last thing to mention is that the Python section of the code has exported the generated data-set to a Stata dta file which Stata then reads back in, which is not strictly necessary if the package "`from sfi import Data`" is utilized. However, for more complicated examples Stata does not seem to be very stable when running Python code, and thus it is recommended to run the Python code separately, export the data in a Stata readable file, and then run the do file that uses that data.

It is possible to conduct this particular example in Python, and as it is vectorized it is several orders of magnitude faster than the given Stata code. This code is provided in the section below. However, for more complicated statistical tests, Python does not have readily available and user friendly packages for many tests Stata can conduct (or if they do, they are not vectorizable), thus requiring the use of looping in Stata.

A.1.3 Python Code

```
import numpy as np


mean1 = .5
std1 = .5
mean2 = .75
std2 = std1


n=63
num_sims = 1000
num_reject = 0
alpha = .05

np.random.seed(1)

d1 = np.random.normal(mean1,std1,(n,num_sims))
d2 = np.random.normal(mean2,std2,(n,num_sims))

p_vals = sps.ttest_ind(d1,d2,axis=0)[1]
rejected = np.logical_not(p_vals>=alpha)

print('power:',np.mean(rejected))
```

A.1.4 Python Code for QR Simulations of Bayesian Persuasion

```
import scipy.stats as sps
import numpy as np


# functions for bayesian persuasion:
```

```
# returns the probability that a red message would be generated
def prob_succeed(prior,x):
    return prior*1.0 + (1.0-prior)*(1-x)

# can be used to calculate the optimal x
# to get receivers to a certain posterior
def optimal_x(prior, posterior_improvement):
    # note this function takes in the prior as a
    # percentage (e.g. 35% instead of .35)
    # for backwards compatibility reasons
    # if you'd like to change that, remove the '/100' from below
    p = prior/100
    q = (prior + posterior_improvement)/100
    out = (p/q-p)/(1-p) - 1
    return out*-1

# prob succeed but with x instead of change in prior
# prior and x are between 0 and 1 (i.e. probabilities, not %ages)
def prob_succeedX(prior,x):
    return prior*1.0 + (1.0-prior)*(1-x)

# for a given x, calculates the new prior
# conditional on a red message

def update_prior(prior,x):
    return prior/(prior+(1-x)*(1-prior))

# calculates the expected profit for a receiver,
# for each possible posterior, what is the expected profit of choosing
# red or blue, given the payoffs
```

```
# returns a 2*(len(posteriors)) matrix,
# first row is ep_red,
# second_row is ep_blue
def receiver_ep(posteriors,pay_match,pay_mismatch):
    red_vec = posteriors*pay_match+(1-posteriors)*pay_mismatch
    blue_vec = posteriors*pay_mismatch + (1-posteriors)*pay_match
    return np.vstack((red_vec,blue_vec))

# calculates the sender's ep for each possible x decision
# he is allowed to make,
# given the qre probability matrix of the reciever
# for each induced posterior
# (should map 1to1 to x decisions) (receiver_probs)
def sender_ep(possible_x_decisions,receiver_probs,pay_red,pay_blue,

receiver_prior):
    prob_red = receiver_probs[0,:]
    prob_succs = prob_succeedX(receiver_prior,possible_x_decisions)
    return (prob_succs*(prob_red*pay_red+(1-prob_red)*pay_blue)
    + (1-prob_succs)*pay_blue)

# given a matrix of expected profits, where each row
# represents a decision,
# and each column represents a state,
# will return a probability matrix with the
# probability of each decision given the state
# according by the logit QRE function with the parameter lamb
# (lamb=lambda, but that word is reserved in python)
# (also seems to work just fine for a vector)
def epmat2qre(ep_mat,lamb):
```

```
    ep_lamb = ep_mat*lamb
    mx = np.max(ep_lamb,axis=0)
    exp = np.exp(ep_lamb-mx)
    return exp/np.sum(exp,axis=0)

# assume senders can set x=0,.01,.02, ... , .99, 1.
# i.e., integer percentages
all_actions = np.array([.01*i for i in range(101)])
# ... meaning there is a finite number of posteriors they can induce
# this function calculates the induced posterior
# for each x they could choose
# priors and x's should be between 0 and 1
# possible x's should be a list
def possible_posteriors(prior,possible_xs):
    return update_prior(prior,np.array(possible_xs))

# calculates the qre equilibrium via backwards induction
# takes in the receiver's expected payoff matrix -
# calculate using receiver_ep function above
def qre_equilibrium(pay_red,pay_blue,sender_lamb,receiver_lamb,
receiver_prior,receiver_ep):
    receiver_probs = epmat2qre(receiver_ep,receiver_lamb)
    sep = sender_ep(np.array(all_actions),receiver_probs,
    pay_red,pay_blue,receiver_prior)
    sqre = epmat2qre(sep,sender_lamb)
#     weighted_x = np.sum(np.array(possible_xs)*sqre)
    return sqre



# setting the bayesian persuasion parameters
```

```
pay_high = 2
pay_low = 1
prior_1 = 1/3
prior_2 = 1/5
pay_match = 2
pay_mismatch = 1

# setting the qre parameters
lambs = 11.26
lambr = lambs

# this particular function provides similar functionality to
# np.random.choice
# but instead takes in an entire matrix (to be applied row-wise)
# IN: choices N*M matrix, probs N*M matrix, OUT: choices N*1 vector
def choice_vec(choices,probs):
    cumprobs = probs.cumsum(axis=1)
    draws = np.random.random(probs.shape[0])
    oneminus = cumprobs-draws.reshape(draws.shape[0],1)
    cols = np.argmax(oneminus>0,axis=1)
    rows = np.arange(cumprobs.shape[0])
    return choices[rows,cols]

# first calculate receiver's expected profit
full_rep1 = receiver_ep(np.array(possible_posteriors(prior_1,
all_actions)),pay_match,pay_mismatch)
full_rep2 = receiver_ep(np.array(possible_posteriors(prior_2,
all_actions)),pay_match,pay_mismatch)

# now calculate sender's noisy BR to the receiver actions
```

```
dist1 = qre_equilibrium(pay_high,pay_low,lambs,lambr,prior_1,full_rep1)
dist2 = qre_equilibrium(pay_high,pay_low,lambs,lambr,prior_2,full_rep2)

num_sims = 1000
n = 63

# set seed for consistent results
np.random.seed(1)
actions1 = choice_vec(np.tile(all_actions,(num_sims*n,1)),
np.tile(dist1,(num_sims*n,1)))
np.random.seed(1000000)
actions2 = choice_vec(np.tile(all_actions,(num_sims*n,1)),
np.tile(dist2,(num_sims*n,1)))

# actions are currently a vector, which is handy for stata
# however, for vectorization, putting it as a matrix
actions1 = actions1.reshape(n,num_sims)
actions2 = actions2.reshape(n,num_sims)

alpha = .05

# conducting the t-tests and counting the number of rejections
p_vals = sps.ttest_ind(actions1,actions2,axis=0)[1]
rejected = np.logical_not(p_vals>=alpha)

print('power:',np.mean(rejected))
```

A.2 D-optimal Example With Treatment-specific Standard Deviation

Consider the following (extreme) illustrative example. Assume an artificial DGP of the form $y = b + cx_i + (x_i + 1)^{2.5}\epsilon_i$, $\epsilon_i \sim N(0, \sigma^2)$, and where the possible treatment levels are $x_i \in [0, 1]$. From the DGP it can be seen that a higher x_i results in noisier data. To ease notation, consider a case where 2 subjects will be assigned to each treatment (the general pattern holds for any distribution of n over treatments), and denote the two treatment intensities to be chosen as w and x, with $x > w$. The Log Likelihood would be as follows:

$LL(x, w) = 4k - \frac{2}{(w+1)^{5.0}}\left(-b - cw + y\right)^2 - \frac{2}{(x+1)^{5.0}}\left(-b - cx + y\right)^2$

Differentiating with respect to b and c (the parameters we seek to accurately estimate) and forming the information matrix yields:

$det(I) = -\left(-\frac{4w}{(w+1)^{5.0}} - \frac{4x}{(x+1)^{5.0}}\right)^2 + \left(-\frac{4w^2}{(w+1)^{5.0}} - \frac{4x^2}{(x+1)^{5.0}}\right)\left(-\frac{4}{(w+1)^{5.0}} - \frac{4}{(x+1)^{5.0}}\right)$

Utilizing a grid search where $w, x \in [1\mathrm{e} - 10, 1]$, the determinant is maximized where $w = 1\mathrm{e} - 10$ and $x = 0.668$. The smallest possible value is selected for w, as this is where noise is minimized. As for x there exists a trade-off, increasing it increases separation between the treatments, but also increases the noise of the observations in the x treatment, meaning it is not optimal to increase x all the way to 1.

A.3 Differential Treatment Variance in Bayesian Persuasion

This section explores the impact of each treatment having different subject variability in the Bayesian Persuasion environment specified in Section 1.4. If the standard deviation differs substantially between treatments, then for optimal power more subjects should be allocated to the more noisy treatment, in proportion with the relative variances (List et al., 2011). Table A.1 presents the required sample size to hit a power of 80% for a selection of larger and smaller variances across each treatment. As evident in Table A.1, a higher variance requires more subjects, and assigning subjects efficiently by allowing the sample size to differ across treatments results in a lower total number of required observations. There are two questions that would need to be answered ex-ante in order to allocate different numbers of subjects to each treatment. Firstly, what treatment would have more variation in subject behavior? Secondly, what would be the ratio of the treatment standard deviations? Ex-post this should be less of an issue, as there should be reasonably reliable estimates of the variation in subject behavior by treatment, but this insight would only really be of use to replication studies. Ex-ante, the QR framework can provide estimates of the anticipated standard deviation of each treatment. If the estimated standard deviations differ substantially, then consideration of non-equal subject allocations may be prudent.

Table A.1. Effects of Unequal Variance

	$\sigma_1 = .25$	$\sigma_1 = .5$	$\sigma_1 = .75$
$\sigma_0 = .25$	16,16	40,40 24,48	79,79 32,96
$\sigma_0 = .5$	40,40 48,24	63,63	103,103 79,119
$\sigma_0 = .75$	79,79 96,32	103,103 119,79	142,142

A.4 Ex-ante Power Analysis of Omitted Papers

A.4.1 De Clippel et al. (2014)

De Clippel et al. (2014) experimentally analyze two rules for the selection of different outcomes (motivated in the paper as arbitrators), the short-list (SL) and veto-rank (VR) methods. In the SL method, one player is randomly selected to create a list of $n < N$ (here $n = 3$, $N = 5$) options, of which the other player chooses the final outcome from. In the VR method, both players remove $m < N$ options from contention (here $m = 2$), and then rank the remaining options. The selected option is the option with the lowest sum of ranks that has not been removed/vetoed.

Ex-ante Power Analysis

The particular treatment effect is that the SL method induces higher average payoffs than the VR method for the preference orderings [a,b,c,d,e] and [b,a,c,d,e] for players 1 and 2 respectively (denoted $Pf2$ in the paper). Unfortunately, this experiment was relatively novel, and there does not appear to be any closely related experiments from which to obtain estimates of λ. For the purposes of this exercise, I will consider the data from the other three preference orderings that were also reported in De Clippel et al. (2014), but that are not involved with the given hypothesis. This generates a 'past experiment' of sorts, and somewhat resembles the problem an experimenter might face.

I fit the QRE λ separately on the VR and SL environments, as despite their similarities as both being methods to determine an outcome for the group, the strategic environments differ substantially. The SL environment is sequential, with the second mover only needing to choose which of the 3 options to take, and the first mover only needing to choose out of 10 possible short-lists. Whereas the VR environment is simultaneous, with each player needing to choose from 60 possible veto rank combinations. The VR environment has multiple Nash equilibria, so I consider only the logit path in this game.[1] Fitting the QRE model to all of the preference orderings except for the one of interest yields $\lambda_{VR} = 10.00$ and $\lambda_{SL} = 2.49$. For the preference ordering of interest, 500 total observations per treatment are required for a power of 80%. This is obtained by having 50 pairs of subjects observed over the 10 rounds in the experiment (so 200 subjects in total).[2]

Optimal Experimental Design

There are a few different factors that could be varied here, such as the number of options, the size of the short-list, the number of vetos, the payoff for each action, and the preference orderings. I focus on the latter only, as fitting a QRE to the VR condition takes a substantial amount of computing time due to the logit path restriction, making considering all possible dimensions intractable. I consider 10 randomly sampled preference orderings that were not considered in the original experiment. Like the original environment, I hold player 1's preference ordering to be $[a, b, c, d, \mathrm{e}]$, while player 2's preference ordering changes. For comparison, I also include the four original preference orderings, denoted $Pf1 - 4$, and denote the newly considered preference orders as $PfX1 - 10$. The resulting analysis is

[1]↑The logit path is the unique path along λ that connects the uniform random play when $\lambda = 0$ to one particular Nash equilibrium when $\lambda \to \infty$. The Nash equilibrium that is converged to is called the logit solution, and is a refinement of the Nash equilibrium.

[2]↑It should be noted that as there is random re-matching within the session, then the t-test really should be conducted on the average at the session level. However, I report this approach as this is how the statistical test was conducted in the original paper.

presented in Table A.2, which reveals a clear separation between low to moderately powered preference orderings, and very high powered preference orderings with a simulation power at or near 100%. I focus on the high powered preference sets, and seek to differentiate between them by setting $N = 50$, $\alpha = 0.01$, and by considering a closed form solution of power that is not subject to random noise. This is presented in Table A.3, and yields a 'winner' of $PfX10-[d, a, \mathrm{e}, c, b]$. It should be noted that here is another example of not just maximizing the treatment effect size. $PfX10$ gains its edge over $PfX7$ despite having a lower τ, due to the decrease in standard deviation in the SL treatment. Using $PfX10$ instead of $Pf2$ would result in a required number of observations of $N = 22$ per treatment, which I round up to the minimum of one session for each treatment with a session size of 20, so $N = 100$ observations per treatment, or 40 subjects in total.

A.4.2 Huck et al. (2011)

Huck et al. (2011) consider an experimental test of the Lazear model of deferred compensation (Lazear, 1979). In their experiment, the firm initially sets the wage profile W_1, W_2, and W_3, which is the worker's wage in periods $t = 1$, $t = 2$, and $t = 3$ respectively. Whether the firm can commit to this wage profile is the treatment of interest. The worker makes effort decisions at $t = 1.5$ and $t = 2.5$, and can choose $E_t \in \{L, M, H\}$. If the worker chooses L, he is fired with probability p, and does not receive the wages that would follow (i.e. either W_2 and W_3 if fired at $t = 1.5$, and W_3 if fired at $t = 2.5$). Both the firm's payoff and the worker's cost is increasing in the chosen effort level. If the worker is fired at $t = 1.5$, then the firm receives the benefit of L effort at $t = 2.5$, at no cost to either worker or firm. The particular parameters chosen for the experiment were $p = 0.5$, costs of effort $C_L = 0$, $C_M = 20$, $C_H = 40$, firm revenues $Z_L = 50$, $Z_M = 100$, $Z_H = 140$, and wage offers could be $W_t \in \{0, 1, 2, ..., 119, 120\}$. The equilibrium is solved via backwards induction, and under commitment the firm's equilibrium behavior is to set a deferred compensation contract of $W_1 = 0$, $W_2 \in [0, 20]$ and $W_3 = 60 - W_2$, which induces the worker to choose M at both effort decisions. Without commitment, the equilibrium is that the firm offers no wages and the worker always chooses L effort. The hypothesis in question is whether the deferred compensation portion of the wage profile $W_2 + W_3$ differs between the commitment and no commitment cases.

Table A.2. De Clippel et al. (2014) - Power of Preference Orderings at $N = 300$

Player 2 Preference Ordering	μ_{SL},μ_{VR} τ σ_{SL},σ_{VR}	Power with $\alpha = .05$
$Pf1$ - [e, d, c, b, a]	1.0,1.0 0.0 0.0,0.0	0.0%
$Pf2$ - [b, a, c, d, e]	1.46,1.54 0.08 0.51,0.41	0.582%
$Pf3$ - [c, b, a, d, e]	1.33,1.37 0.04 0.43,0.34	29.6%
$Pf4$ - [e, c, a, b, d]	1.16,1.29 0.13 0.32,0.31	99.8%
$PfX1$ - [e, b, a, c, d]	1.23,1.34 0.11 0.38,0.32	96.4%
$PfX2$ - [e, d, c, a, b]	1.02,1.03 0.01 0.11,0.11	20.1%
$PfX3$ - [b, c, d, e, a]	1.34,1.59 0.25 0.48,0.31	100%
$PfX4$ - [d, c, e, b, a]	1.09,1.17 0.08 0.27,0.17	99.9%
$PfX5$ - [c, e, b, a, d]	1.2,1.22 0.02 0.36,0.28	11.6%
$PfX6$ - [a, c, b, d, e]	1.53,1.86 0.33 0.57,0.31	100%
$PfX7$ - [a, d, e, b, c]	1.49,1.87 0.38 0.61,0.31	100%
$PfX8$ - [c, b, a, e, d]	1.32,1.41 0.09 0.44,0.29	81.3%
$PfX9$ - [b, e, c, a, d]	1.32,1.58 0.25 0.48,0.32	100%
$PfX10$ - [d, a, e, c, b]	1.26,1.6 0.34 0.47,0.3	100%

Table A.3. De Clippel et al. (2014) - Power of Selected Preference Orderings at $N = 50$

Player 2 Preference Ordering	μ_{SL},μ_{VR} τ σ_{SL}, σ_{VR}	Power with $\alpha = .01$
$Pf4$ - [e, c, a, b, d]	1.13,1.29 0.15 0.34,0.31	42.4%
$PfX1$ - [e, b, a, c ,d]	1.23,1.34 0.11 0.38,0.32	15.9%
$PfX3$ - [b, c, d, e, a]	1.33,1.59 0.25 0.48,0.31	71.0%
$PfX4$ - [d, c, e, b, a]	1.09,1.17 0.09 0.27,0.17	25.8%
$PfX6$ - [a, c, b, d, e]	1.53,1.86 0.33 0.57,0.31	86.2%
$PfX7$ - [a, d, e, b, c]	1.49,1.87 0.38 0.61,0.3	91.2%
$PfX9$ - [b, e, c, a, d]	1.32,1.58 0.25 0.47,0.32	70.2%
$PfX10$ - [d, a, e, c, b]	1.27,1.6 0.34 0.47,0.29	95.5%

Ex-ante Power Analysis

It is relatively straightforward to fit a QRE to either of these sequential environments, so all that remains is to specify a reasonable λ. However, unfortunately there does not appear to be any tangentially related experimental environments with which to obtain this from. Instead, for the purposes of this exercise, I conduct a thought experiment where I only observe data from one of the treatments, and then use that to fit the λ. This is another way to try and make an analogous situation to what an experimenter might face, using a previous experimental setup as a baseline, and proposing to test a new environment. I run this exercise both ways, i.e. I assume I only have the commitment data to fit a λ to then generate predictions for the no commitment experiment, and vice versa. From either approach I obtain similar estimates of λ, $\lambda_{FCT} = 0.0414$ and $\lambda_{NCT} = 0.0552$. This yields predictions as presented in Table A.4. The QRE simulation has performed quite well in terms of the treatment effect size, but less so in terms of point predictions. It has predicted the differential standard deviation well, but not so much from the FCT treatment. The statistical test that was conducted was a Mann-Whitney test on the average deferred wages by session, which consisted of groups of 10 (5 firms and 5 workers), who interacted over 20 rounds. The averaging by session reduces the noise substantially, and I find that in conjunction with to the large treatment effect size, the test is very strongly powered even for only a small number of independent observations.

Table A.4. Huck et al. (2011) QRE Power Analysis

Treatment	λ	μ_{NCT}	σ_{NCT}	μ_{FCT}	σ_{FCT}	τ	Sessions per Treatment
FCT	0.0414	45.75	2.93	73.29	3.66	27.54	3
NCT	0.0552	34.95	2.34	68.63	3.20	33.68	3
Actual	0.0480*	21.31	3.65	54.74	12.61	33.43	6**

*QRE fit on both data-sets.

**Number of sessions used in Huck et al. (2011).

Optimal Experimental Design

As the test is already strongly powered even for very low number of observations, via simulation there are issues with 'perfectly powered' tests with a power of 100%. It is not possible to improve from a perfectly powered test, therefore I cannot investigate changes in power due to changes in the experimental design. Furthermore, as the statistical test is so strongly powered, it is of little practical importance to improve power. Therefore, for this particular paper, I do not conduct any thought exercise with regards to changes in the experimental design.

B. CHAPTER 2 APPENDIX

B.1 Discussion of Numerical Solutions and Unique Solutions

B.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 α as a, his objective function with the budget constraint substituted for x is the following:

$$\left(1 - \mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}}\right)\left(\left(\left(1 - \mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}}\right)\mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{y}{z}}\right)\right)^{a}} + \mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}}\right) + \mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}}\right) \tag{B.1}$$

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

$$
\begin{aligned}
&-\frac{a\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}\mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}}\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}}{4z\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)}\Bigg(\Bigg(1 \\
&-\mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}}\Bigg)\mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{y}{z}}\right)\right)^{a}}+\mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}}\Bigg) \\
&+\frac{a\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}\mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}}\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}}{4z\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)} \\
&+\left(1-\mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}}\right)\Bigg(-\frac{a\left(-\log\left(1-\mathrm{e}^{-\frac{y}{z}}\right)\right)^{a}\mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{y}{z}}\right)\right)^{a}}\mathrm{e}^{-\frac{y}{z}}}{z\left(1-\mathrm{e}^{-\frac{y}{z}}\right)\log\left(1-\mathrm{e}^{-\frac{y}{z}}\right)}\Bigg(1 \\
&-\mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}}\Bigg) \\
&+\frac{a\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}\mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}}\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}}{4z\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)} \\
&-\frac{a\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}\mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{y}{z}}\right)\right)^{a}}\mathrm{e}^{-\left(-\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\right)^{a}}\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}}{4z\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)\log\left(1-\mathrm{e}^{-\frac{1}{z}\left(\frac{B}{4}-\frac{y}{4}\right)}\right)}\Bigg)=0
\end{aligned}
\tag{B.2}
$$

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 B.1 presents such an analysis graphically for selected combinations of B, z, and a.

The top left graph in Figure B.1 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 $log(1-\mathrm{e}^{\frac{B-y}{4z}})$. 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 α 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 ($B = 24$, $z = 18.2$, $z = 31.1$).

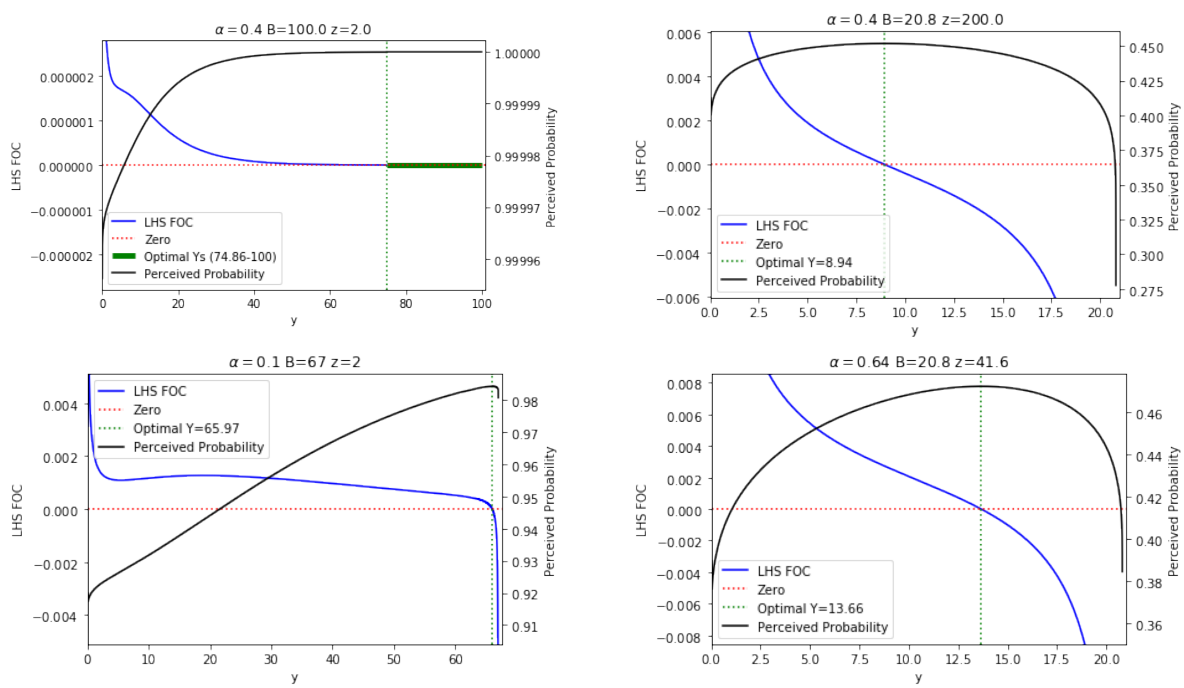


Figure B.1. Selected Numerical Solutions for Common Edge and $p(x_i) = 1 - e^{\frac{-x_i}{z}}$

Several general patterns are evident from the graphs in Figure B.1. 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 α (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.

B.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:

$$\begin{aligned}
&e^{-(-b\log(y)+b\log(z))^a} + 2e^{-\left(-\log\left(\left(\frac{B}{4z}-\frac{y}{4z}\right)^b\right)\right)^a} \\
&- 2e^{-\left(-\log\left(\left(\frac{B}{4z}-\frac{y}{4z}\right)^b\right)\right)^a} e^{-(-b\log(y)+b\log(z))^a} \\
&- e^{-2\left(-\log\left(\left(\frac{B}{4z}-\frac{y}{4z}\right)^b\right)\right)^a} + e^{-2\left(-\log\left(\left(\frac{B}{4z}-\frac{y}{4z}\right)^b\right)\right)^a} e^{-(-b\log(y)+b\log(z))^a}
\end{aligned} \tag{B.3}$$

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

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 B.2 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

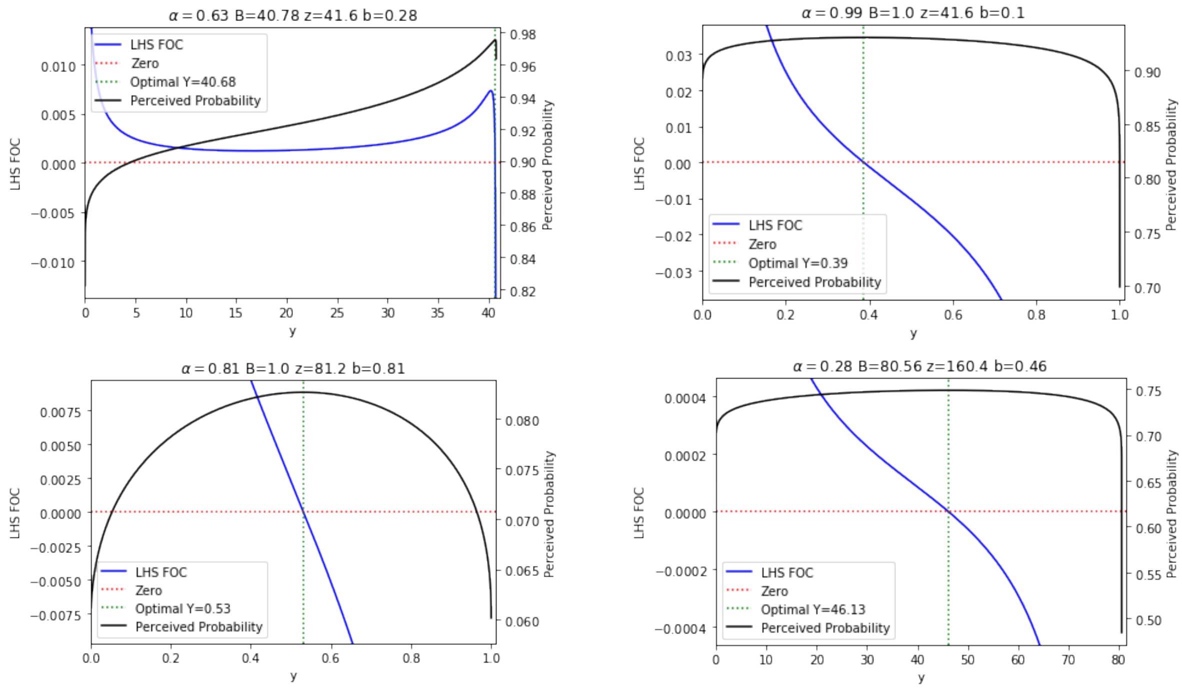


Figure B.2. Selected Numerical Solutions for Common Edge and $p(x_{\mathrm{i}}) = (\frac{x_{\mathrm{i}}}{z})^b$

where the first order condition is not negative as y tends to B. These are characterized by high values of αs, Bs, zs, and bs, 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.

B.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.

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

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, $\mathcal{L}$, α, 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 B.3, B.4, and B.5, which represent the three types of patterns that generally emerge. The top graph in each figure is a heatmap[1] of the expression representing both first order conditions, while the four following graphs are each first order

[1]↑Note, the heatmap is truncated for high values of F so we can more easily focus on values where it is close to zero.

condition holding one parameter fixed at the optimum value. The latter is to support the heatmap analysis, for example, the heatmap in Figure B.4 looks rather flat along the $\mathcal{L}$ dimension, but the 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 singled 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.

B.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. 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 $1 - \mathrm{e}^{\frac{-x}{z}} + \mathrm{e}^{\frac{-x}{z}}(1 - \mathrm{e}^{\frac{-x}{z}}) = 1 - \mathrm{e}^{\frac{-x-y}{z}}$. 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 $p(x_{\mathrm{i}}) = 1 - \mathrm{e}^{\frac{-x_{\mathrm{i}}}{z}}$, 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 in 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.

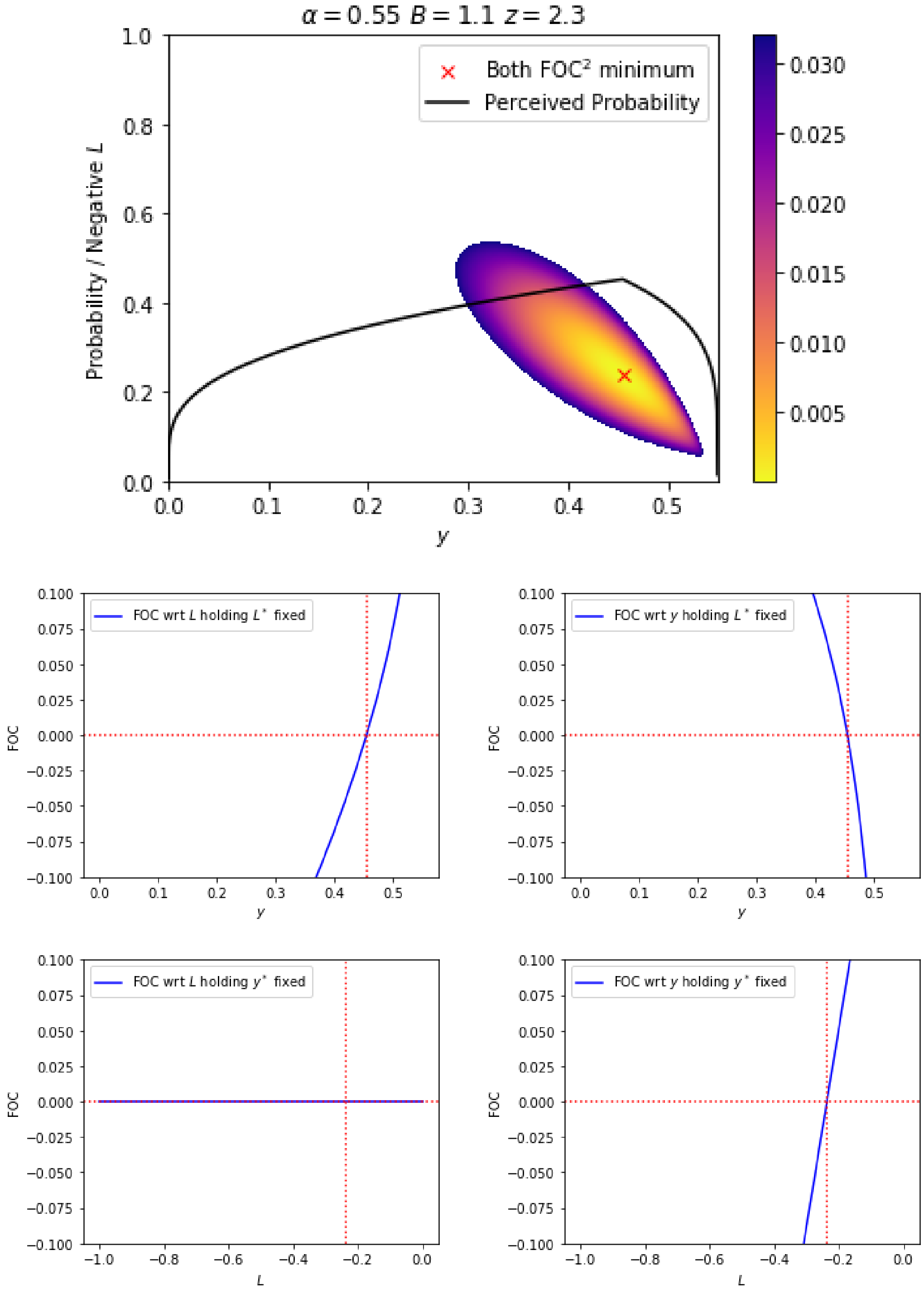


Figure B.3. Network Green - $\alpha = .55$, $B = 1.1$, $z = 2.3$

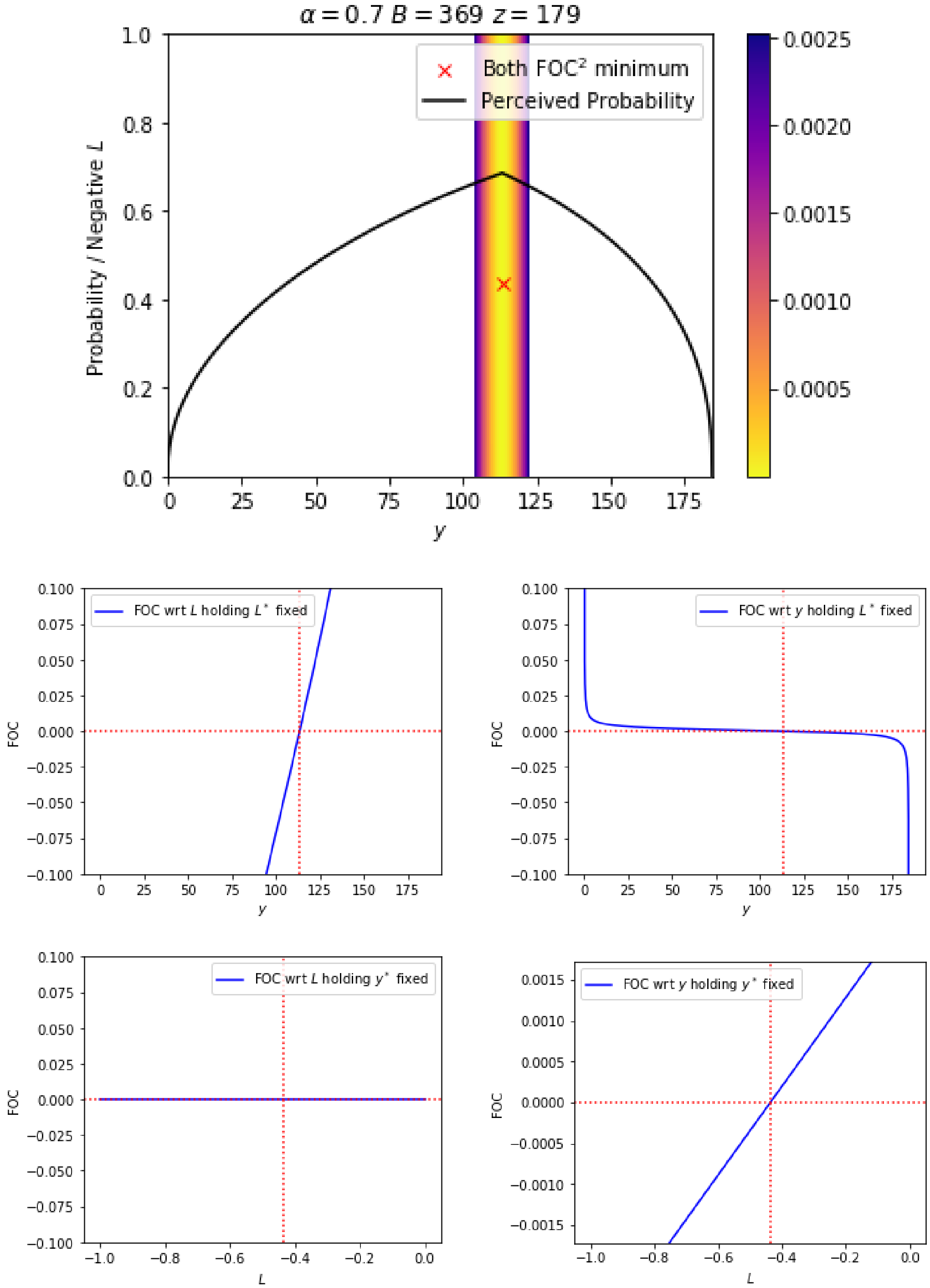


Figure B.4. Network Green - $\alpha = .7$, $B = 369$, $z = 179$

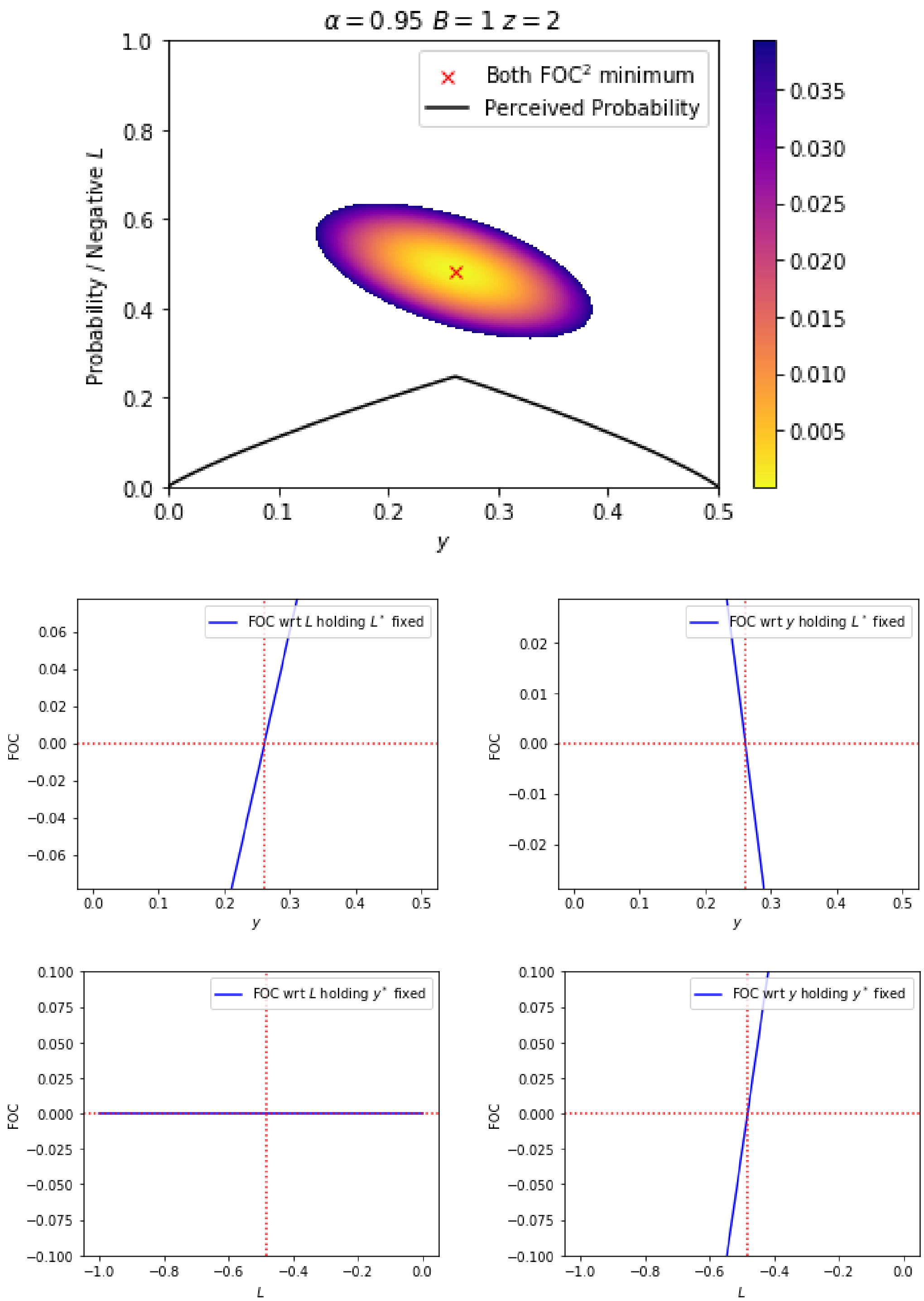


Figure B.5. Network Green - $\alpha = .95$, $B = 1$, $z = 2$

B.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.

B.2 DOSE Procedures

B.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 α and λ. 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 λ, 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).

B.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 λ 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 λ, where higher λ means subjects are getting closer to perfectly best responding. We start with a low λ, go through the logit responses of the aforementioned α and σ 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 λ is increased slightly. This process continues until the 80% threshold is reached. The ending λ_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 α, σ and λ. 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).

B.3 Hypothesis 2 Robustness Tests

Table B.1. One-sided p-values from Wilcoxon Signed Rank Tests of Hypothesis 2 for Different α Thresholds

	.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

B.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 B.6 and B.7 present the CDFs of the estimated σ and α from this measure.

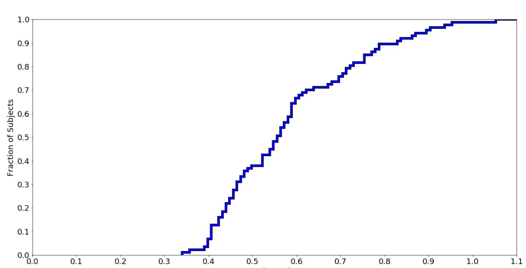


Figure B.6. CDF of Elicited σ

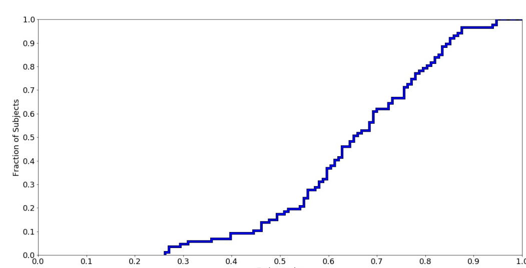


Figure B.7. CDF of Elicited α

Table B.2 presents the Spearman's ρ 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 α.

Table B.2. Full Spearman's ρ Table.

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

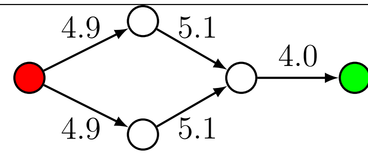
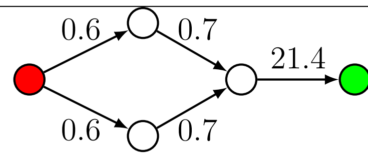
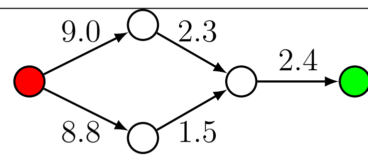
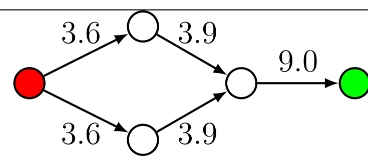
Bold if $p < 0.05$, italics if $0.05 < p < 0.1$.

AT=Network Attack Task, LT= Binary Lottery Task

B.5 Individual Network Cluster Analysis

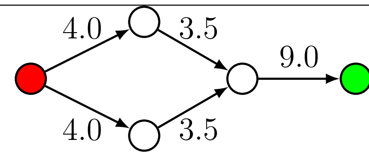
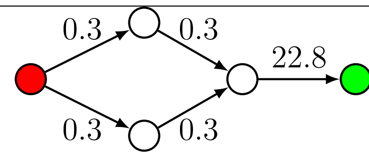
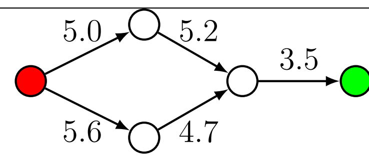
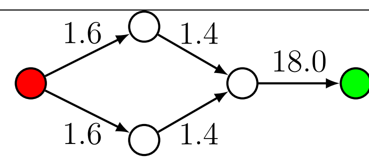
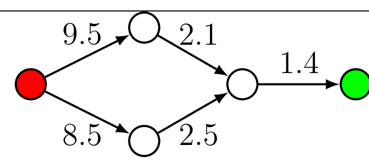
B.5.1 Network Red

Table B.3. Network Red Cluster Analysis

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%	

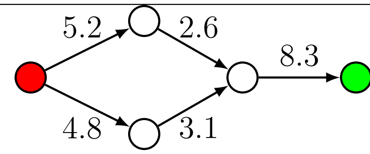
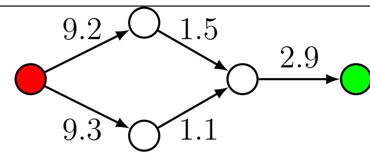
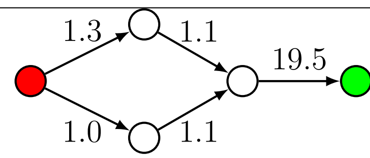
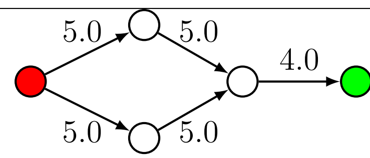
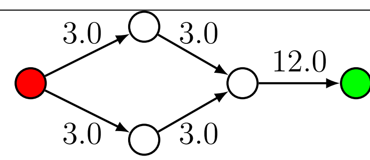
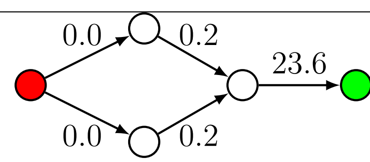
B.5.2 Network Orange

Table B.4. Network Orange Cluster Analysis

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%	

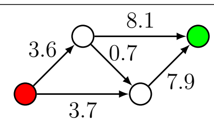
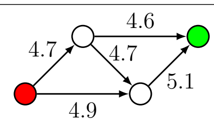
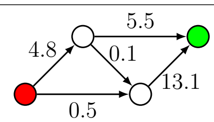
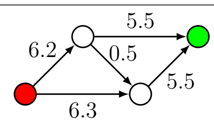
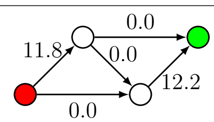
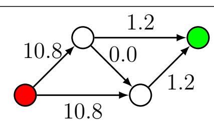
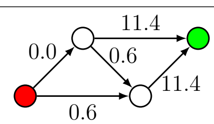
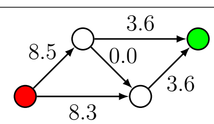
B.5.3 Network Yellow

Table B.5. Network Yellow Cluster Analysis

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%	

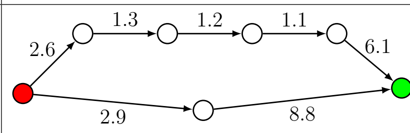
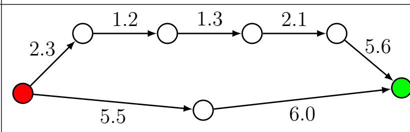
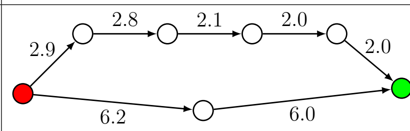
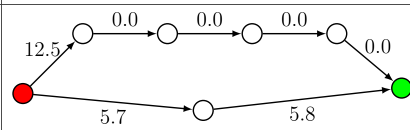
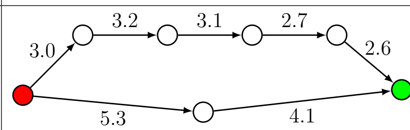
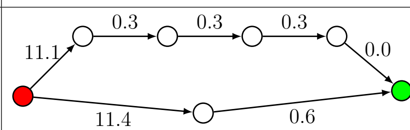
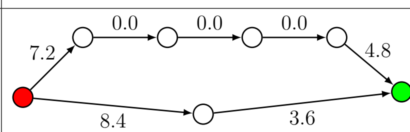
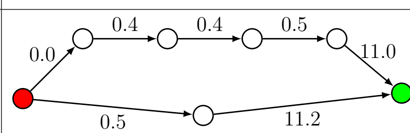
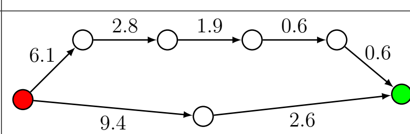
B.5.4 Network Blue

Table B.6. Network Blue Cluster Analysis

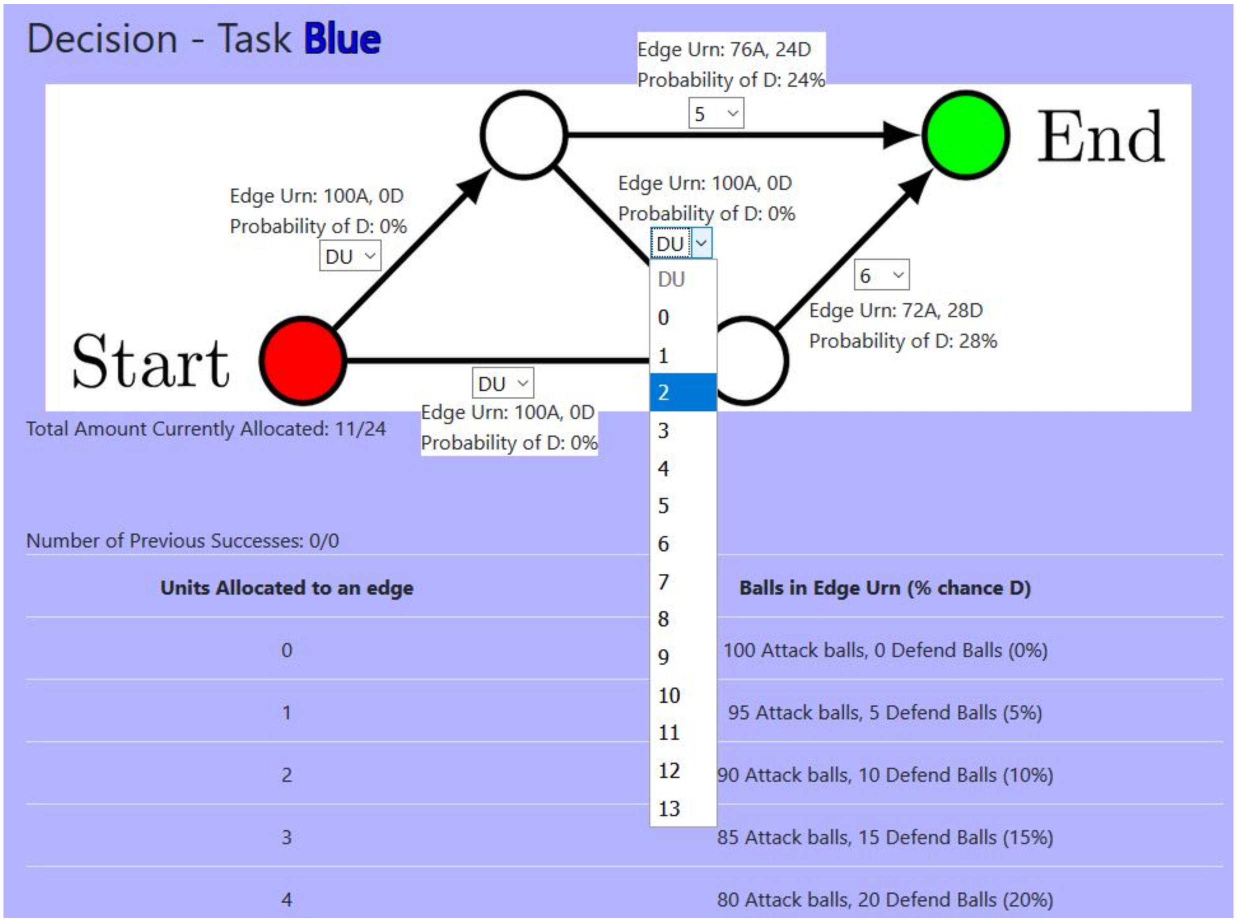
<table>
<tr><th>Cluster Name
% of subjects in Cluster</th><th>Average Edge Allocation</th></tr>
<tr><td>Late Revelation - Mild Diversification
24.1%</td><td>
</td></tr>
<tr><td>Naive Diversification
12.6%</td><td>
</td></tr>
<tr><td>‘False Common Edge’ Heuristic
3.4%</td><td>
</td></tr>
<tr><td>Mild Diversification and Mild Early Revelation
18.4%</td><td>
</td></tr>
<tr><td>Near Optimal
2.3%</td><td>
</td></tr>
<tr><td>Near Optimal - Early Revelation
9.2%</td><td>
</td></tr>
<tr><td>Late Revelation - Very Mild Diversification
20.7%</td><td>
</td></tr>
<tr><td>Near Optimal - Some Early Revelation
9.2%</td><td>
</td></tr>
</table>

B.5.5 Network Green

Table B.7. Network Green Cluster Analysis

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%	

B.6 Experiment Interface



B.7 Experiment Instructions

B.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.

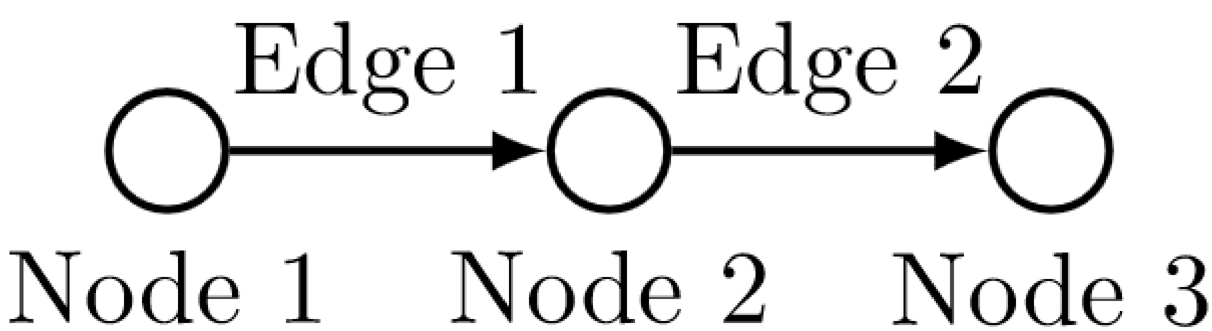
In this experiment, you will participate in Task 1, and then 6 additional colored tasks. The Task 1 instructions are given on a separate piece of paper. For the additional colored tasks, you have been given instructions printed on different colored paper. The color of the paper coincides to the name 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.

B.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.

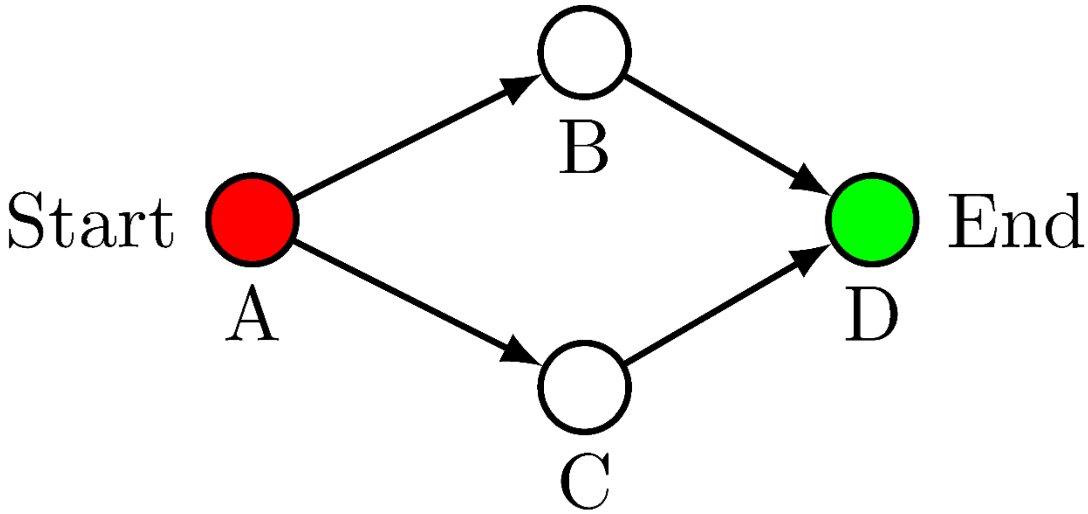
The ball for your chosen Option is not drawn until the end of the 20 decision periods. At the end of Task 1 you will be shown the randomly selected period, the Options that were available in that period, the Option you chose, the ball that was drawn from the chosen Option, and your final payoff for Task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

B.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.

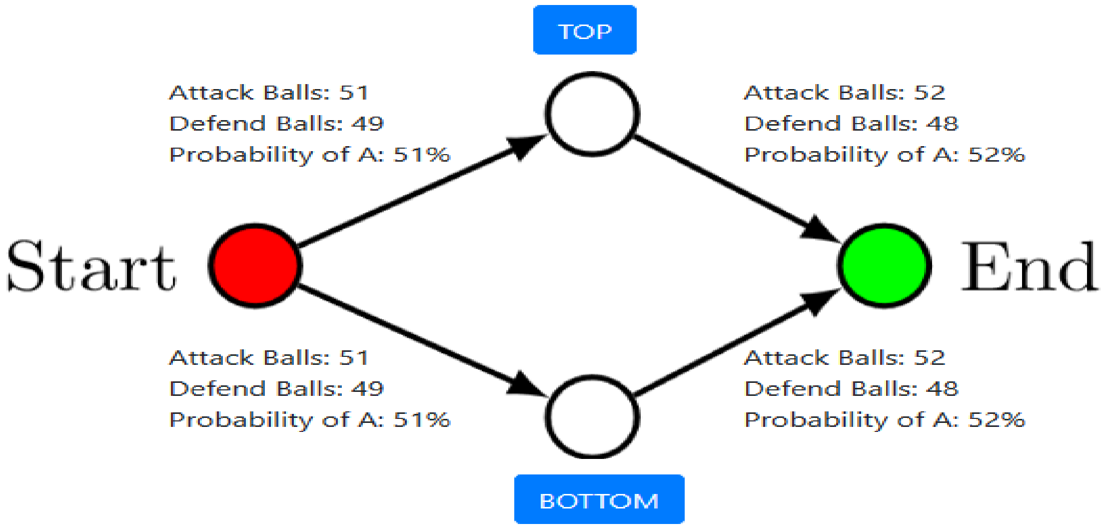
There are two roles in this task, Attacker and Defender. **You will be playing in the role of the Attacker against a computerized Defender. As an Attacker, your objective is to capture the node labelled End in the figure below.**



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

B.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.

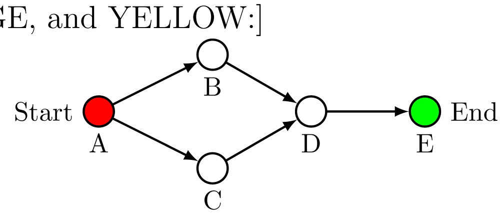
[ALL TASKS:]

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.

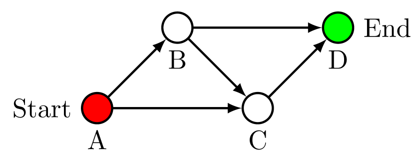
There are two roles in this task, Attacker and Defender. **You will be playing in the role of the Defender against a computerized Attacker. As a Defender, your objective is to prevent the Attacker from capturing the node labelled End in the figure below.**

[TASKS RED, ORANGE, and YELLOW:]



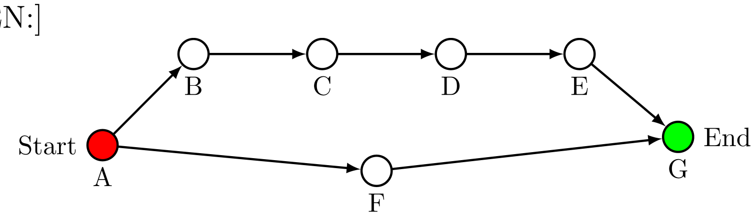
The Attacker starts at Node A, labelled Start, and can attempt to capture any other connected Node via the Edges in the direction of the arrows. For example, from Node A the attacker can attempt to capture Node B or Node C. If the Attacker captures the Node, it moves to that Node, and can then attempt 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[2].

[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 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:]

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

[2] ↑This should read 'Node C' instead of 'Node D'. We present the instructions in their original form with typos included for replication purposes.

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%)

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:]

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	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 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**

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%)

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.

C. CHAPTER 3 APPENDIX

C.1 Theory Appendix

This section focuses on how predictions are obtained. The first subsection shows how the continuous time predictions are obtained. The second subsection shows how the discrete time approximation predictions are obtained. The third subsection displays the payoff hills and the behavioral predictions for the non-Baseline treatments.

C.1.1 Continuous Time Predictions

The continuous time predictions are taken from Strulovici (2010). The Hamilton-Jacobi-Bellman equation for this problem can be written as

$$ru(p) = max\left\{p\lambda h + \lambda p\left(\frac{\lambda h}{r} - u(p)\right) - \lambda p(1-p)\frac{du}{dp}u(p), s\right\}.$$

Through the smooth-pasting condition, and value matching conditions, this equation can be rewritten as

$$s = p\lambda h + \lambda p\left(\frac{\lambda h}{r} - \frac{s}{r}\right).$$

This leads to a cutoff belief of

$$p_A = \frac{s}{\lambda h + \frac{\lambda}{r}(\lambda h - s)}.$$

C.1.2 Discrete Time Predictions

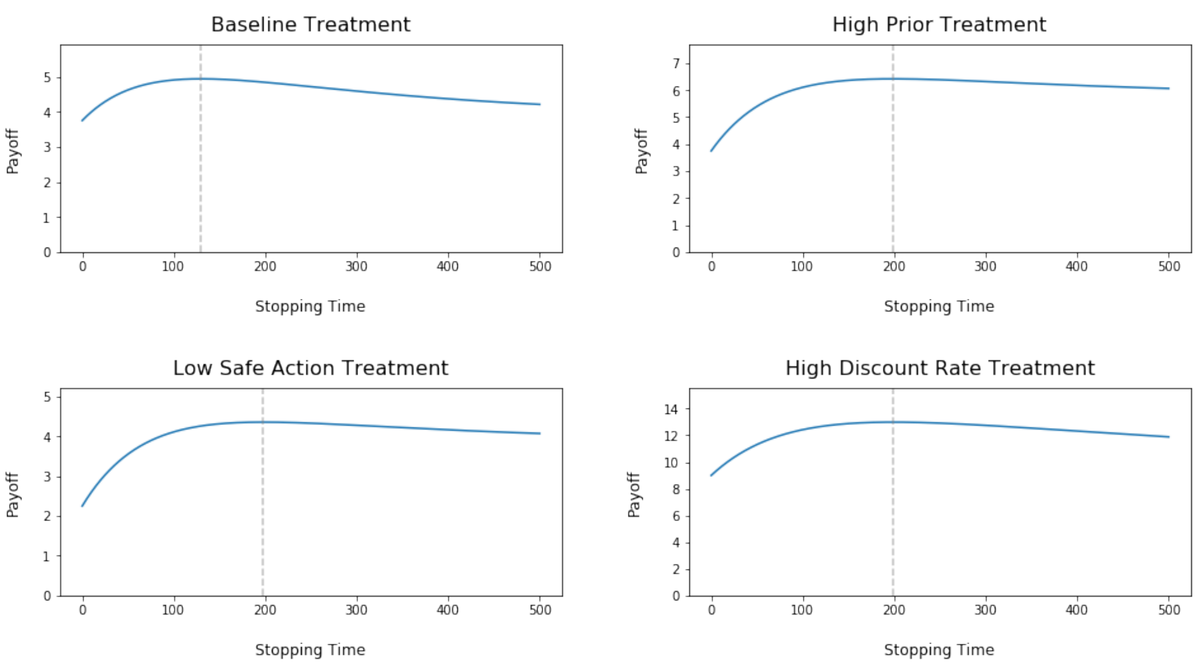
The discrete time predictions are found by first solving for the cutoff belief and then finding the minimum number of ticks, in the absence of a reward, that results in a belief smaller than the cutoff belief. The cutoff belief can be found by value function iteration on the following equation:

$$u(p) = max\{\frac{s}{r\Delta}, p * \lambda\Delta * (h + (1 - r\Delta) * \frac{\lambda\Delta * h}{r\Delta}) + (1 - p * \lambda\Delta) * (1 - r\Delta) * u(p)\},$$

where $p = \frac{p(1-\lambda\Delta)}{p(1-\lambda\Delta)+(1-p)}$. The value function iteration consists of 10001 values of p starting at 0 and increasing by increments of 0.0001 until 1 is reached. The initial guess for $u(p)$ is $\frac{s}{r\Delta}$ at each value of p. The value of $u(p)$ is obtained through interpolation. The predictions are then found by solving for the smallest number of ticks such that $\frac{p_0(1-\lambda\Delta)^{ticks}}{p_0(1-\lambda\Delta)^{ticks}+(1-p_0)}$ is smaller than the cutoff belief. This process is used similarly for Sections 3.3.4 and 3.5.

C.1.3 Payoff Hills and Behavioral Predictions

Figure C.1 displays the payoff hills for each treatment. These payoff hills are based on the discrete time approximation. Each payoff hill has a similar structure where payoffs are increasing at a decreasing rate as the stopping time approaches the optimal stopping time from zero. Payoffs are decreasing at a relatively flat rate as the stopping time increases from the optimal stopping time.



(The gray dotted line is the predicted stopping time.)

Figure C.1. Payoff Hills for each Treatment of the Experiment

Section 3.3.4 displays the effects of unilaterally increasing risk aversion, conservatism, and non-linear probability weighting for the Baseline treatment. These effects can be shown for the three non-baseline treatments.

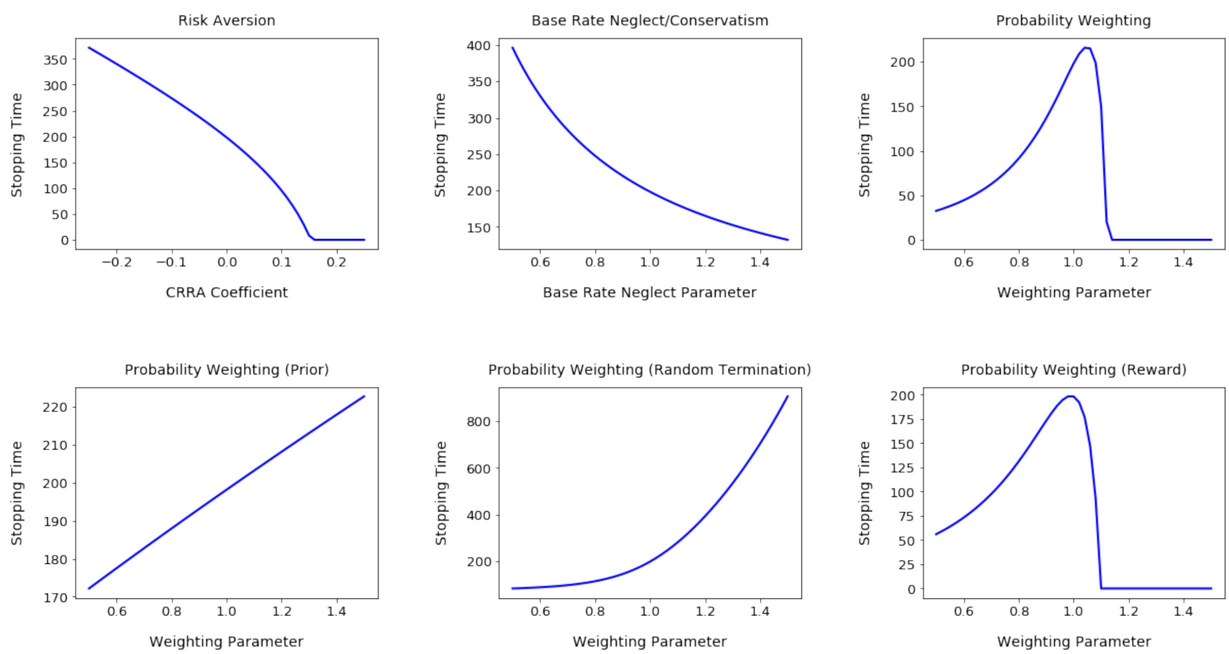


Figure C.2. Predictions for the High Prior Treatment as Different Behavioral Factors are Unilaterally Varied

Figure C.2 displays the predictions for the High Prior Treatment as various behavioral factors are uniformly varied. These responses to the behavioral factors are similar to the responses in Figure 3.2 except that the length of time that an individual is willing to experiment is now increasing in the weighted prior as α increases. This occurs because the Prelec-I function is centered around 0.368 and thus the weighting of the prior gets further away from 0.368 and closer towards its true value as α increases from 0 to 1. As α increases from 1, the weighting of the prior continues to increase. As the weighted prior is increasing in α, subjects unilaterally become willing to experiment longer.

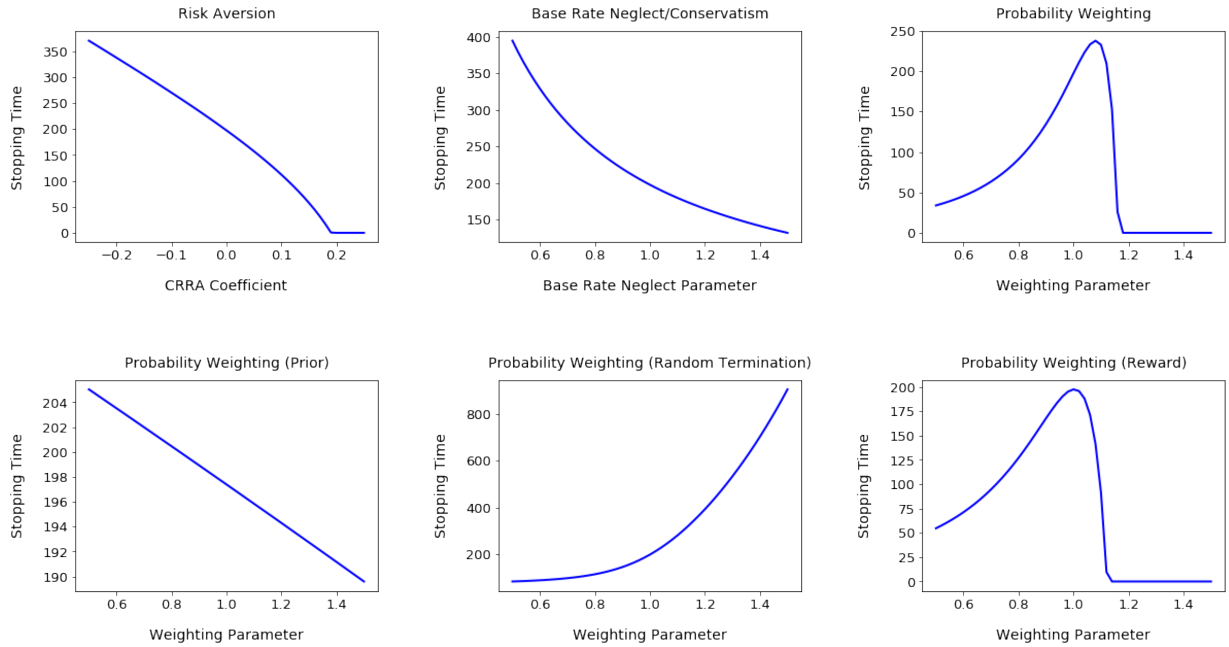


Figure C.3. Predictions for the Low Safe Action Treatment as Different Behavioral Factors are Unilaterally Varied

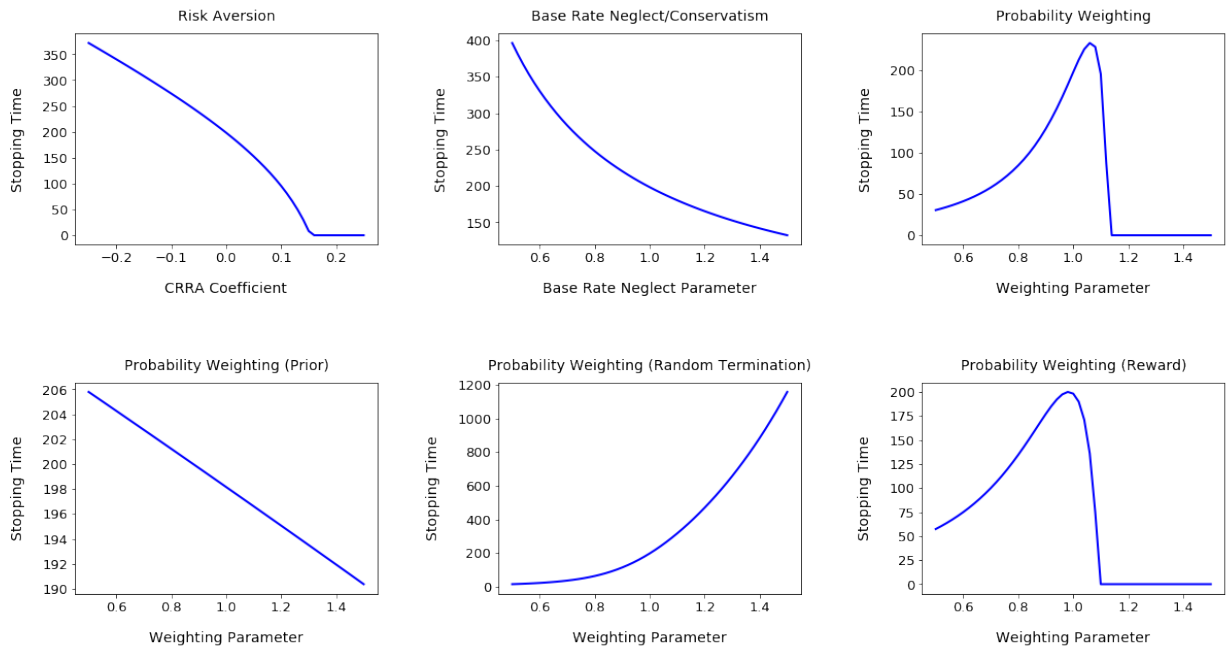


Figure C.4. Predictions for the High Discount Factor Treatment as Different Behavioral Factors are Unilaterally Varied

Figure C.3 displays the predictions for the Low Safe Action treatment as various behavioral factors are varied. Figure C.4 displays the predictions for the High Discount Factor treatment as various behavioral factors are varied. The responses to the behavioral factors, in these two figures, are similar to the responses in Figure 3.2.

C.2 Empirical Appendix

This section complements the results section of Chapter 3. The first subsection focuses on details about the Product Limit estimator. The second subsection displays tests for order effects and has regressions that control for demographics. The third subsection shows the censoring rates in the data used in the Subset approach.

C.2.1 Product Limit Estimator

This subsection displays how the Product Limit estimator was used to conduct hypothesis tests. The Product Limit estimator uses the information contained in the censored observations to correct for censoring bias. The Product Limit estimator uses the implementation time of the risky action. Let t_i denote the observed implementation time in cases of stopping or censoring. Note that censoring occurs when (i) the period ends before an unsure agent switches to the safe action or (ii) an agent obtains a reward. Each t_i is ordered from smallest to largest. For each t_i, let d_i denote the number of events (stops) at t_i, and let n_i denote the number at risk right before t_i (stopped or censored at or after). The Product Limit estimator of the CDF is thus $F(t) = 1 - \prod_{t_i \leq t} \frac{n_i - d_i}{n_i}$. We calculate the full CDF over observed implementation times and use it to calculate mean stopping times.

We conduct analysis using the Product Limit estimator in the following way. We use the Product Limit estimator to create a mean stopping time for each subject for each treatment. The mean stopping time is found by finding the area under the survival curve over the range from 0 to the maximum observed implementation time. We then use a bootstrapped regression to test whether the appropriate mean (or the difference in appropriate means) is significantly different from the predicted value. This regression is a regression of the appropriate variable on a constant.

One important assumption in Product Limit estimation is that at any time observations that are censored have the same survival prospects as those that were not censored. We satisfy this assumption because we create Product Limit estimated means for each subject (for each treatment). If a subject who was censored at a time j was not censored, we would expect her to behave similarly to other times that she wasn't censored at time j.

C.2.2 Order Effects and Demographics

Table C.1. Tests of Order Effects in each Treatment

	Baseline	High Prior	Low Safe Action	High Discount Factor
	Stopping Time	Stopping Time	Stopping Time	Stopping Time
Order	12.95	-6.448	-4.591	29.07
	(10.12)	(22.96)	(31.05)	(56.87)
Constant	84.23***	101.4***	133.3***	132.3***
	(9.361)	(20.74)	(21.79)	(45.61)
Observations	449	112	151	246

Random effects regressions of stopping time on an order dummy variable.
Clustered at the subject level with subject level random effects.
Order dummy variable indicates whether or not a subject did not start with the Baseline treatment.
Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C.1 tests for order effects in each treatment. In each of the regressions, the order dummy variable is not statistically significant at the ten percent level.

Table C.2. Demographic Robustness Checks

	High Prior Sess.	Low Safe Action Sess.	High Discount Factor Sess.
	Stopping Time	Stopping Time	Stopping Time
Non-Basline Treat.	15.79*	37.02**	58.57***
	(9.51)	(14.76)	(21.77)
Male	25.09	19.04	100.01**
	(35.03)	(36.53)	(46.61)
Caucasian	-16.31	71.37	-89.87
	(33.69)	(55.24)	(142.77)
Hispanic	-127.61*	0.00	144.16
	(66.14)	(.)	(96.76)
Second Year	-80.70**	8.47	-3.97
	(37.91)	(38.71)	(23.88)
Third Year	-4.02	-35.60	-61.13
	(37.59)	(66.37)	(42.82)
Graduate Student	-68.01	167.09***	-103.28***
	(66.19)	(49.37)	(29.91)
Economics	138.43	-36.33	0.00
	(106.82)	(70.67)	(.)
Engineering	109.18**	14.42	48.42
	(45.41)	(55.47)	(34.53)
Management/Business	-58.11	52.57*	83.33
	(36.81)	(31.77)	(73.07)
Science	-50.44	63.18	68.27
	(36.61)	(55.95)	(51.80)
India	0.00	-45.74	43.23**
	(.)	(75.27)	(20.90)
United States of America	-15.85	-116.88	166.15
	(35.89)	(75.48)	(111.18)
Constant	137.55***	96.83	-108.93*
	(28.54)	(62.83)	(58.88)
Other Demographics	YES	YES	YES
Observations	263	298	397

Clustering and random effects at the subject level.
Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C.2 displays the tests of Hypothesis 1 when we control for demographics. The regressions show that the results are similar to the results in the main body of the text when we control for demographics. A likelihood ratio test suggests that when including demographics, we should not include random effects for the High Discount Factor session. However, the results are similar with the coefficient on the High Discount Factor Treatment now significant at the five percent level (p-value=0.016).

C.2.3 Censoring

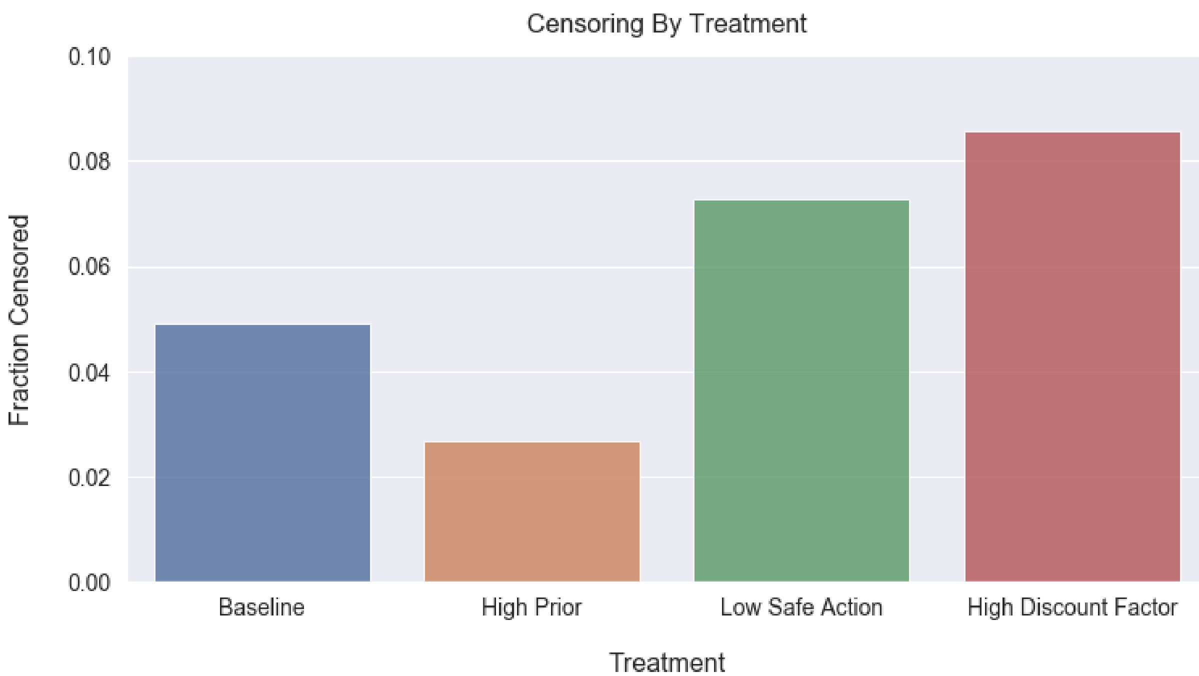


Figure C.5. Rate of Censoring in each Treatment under the Subset Approach

The Subset approach utilizes a subset of data where censoring should be mitigated. However, there is still some censoring in this dataset. Figure C.5 displays the rate of censoring in each treatment under the Subset approach. Censoring is lowest in the High Prior treatment, followed by the Baseline treatment, Low Safe Action treatment, and High Discount Factor treatment.

C.3 Model Appendix

Table C.3. Estimates from Robustness Checks on the Structural Model

Parameter	Non-Behavioral Value	Section 3.5 Model	Sep. Prob. Weights	TK Weighting
γ	0.00	0.36	0.36	0.38
		[0.00]	[0.00]	[0.02]
ψ	1.00	0.13	0.13	0.10
		[0.02]	[0.01]	[0.04]
α_1	1.00	0.65	0.54	0.53
		[0.00]	[0.00]	[0.02]
α_2	1.00	—	0.66	—
		—	[0.04]	—
Log-likelihood	—	8658.03	8657.63	8659.32

Non-behavioral value refers to the restriction based on the theory presented in Section 3.2.
"Section 3.5 Model" refers to the structural model estimated in Section 3.5.
'Sep. Prob. Weights" refer to the model that allows subjects to weight the prior separately than the other probabilities.
"TK Weighting" refers to the model that replaces the Prelec-I function with a Tversky and Kahneman (1992) probability weighting function.
P-values from likelihood ratio tests are in square brackets.

This section addresses the model section of Chapter 3. The first subsection displays robustness checks on the model shown in Section 3.5. The second subsection displays the effect of each behavioral factor in each treatment.

C.3.1 Robustness Checks

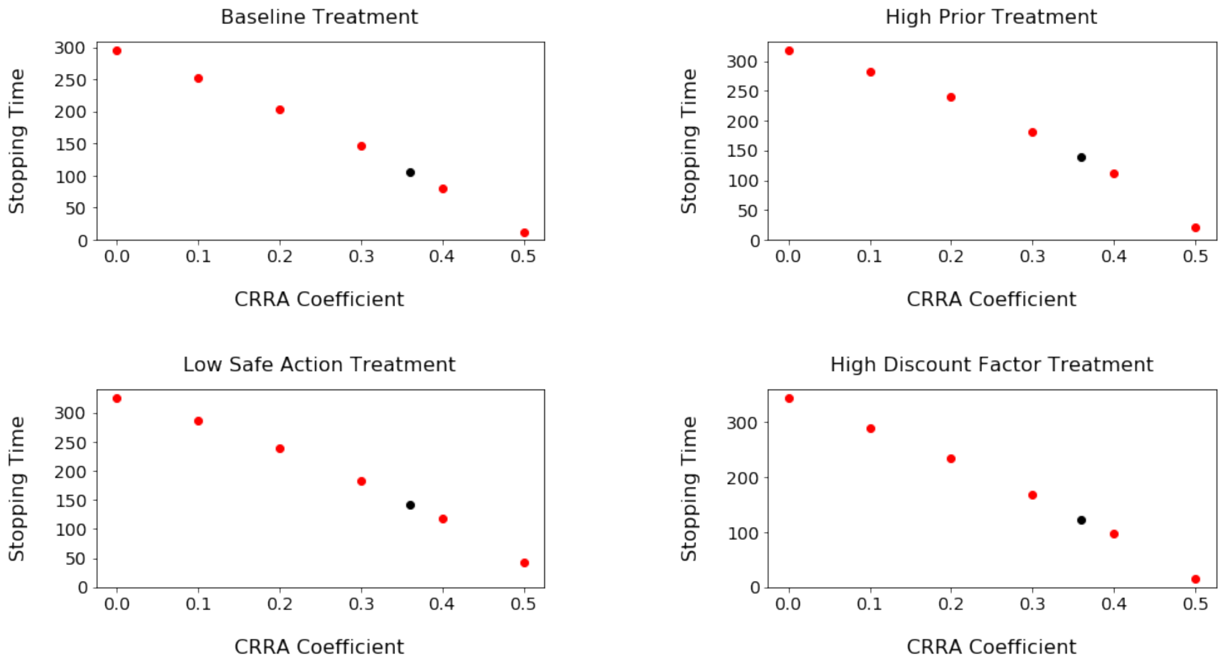
In this subsection, we focus on two main robustness checks to the model. First, we allow for the possibility that subjects weight the prior probability differently from the other probabilities. Second, we replace the Prelec-I probability weighting function with the Tversky and Kahneman (1992) probability weighting function.

In the first robustness check, we allow for the possibility that subjects weight the prior differently than the other probabilities. One reason why this may occur is that the tick-based probabilities (reward probability and random termination probability) are much smaller than the prior probability. Another reason why this may occur is that subjects continually experience these tick-based probabilities in a period, which they do not for the prior probability. This robustness check results in a model similar to the model in Section 3.5, except that now non-linear probability weighting is allowed to vary between the prior probability (α_1) and the tick-based probabilities (α_2). Note that the tick based-probabilities share the same probability weighting parameter. As shown in Table C.3, this model results in similar estimates of γ and ψ as our model in Section 3.5. Additionally, we fail to reject the hypothesis at the ten percent level, using a likelihood ratio test, that subjects have an equal probability weighting parameter for the prior probability and the tick-based probabilities.

In the second robustness check, we allow for a different probability weighting function. We replace the Prelec-I function with the Tversky and Kahneman (1992) probability weighting function. We do this to make sure that our results are not subject to the particular weighting function that we use. This weighting function is given by $w(p) = \frac{p^\alpha}{(p^\alpha+(1-p)^\alpha)^{\frac{1}{\alpha}}}$. The model estimated in this robustness check is identical to the model in Section 3.5 except for this different probability weighting function. As shown in Table C.3, the model results in similar estimates of γ and ψ as in the model in Section 3.5. Additionally, our model in Section 3.5 is preferred to this model (as determined by both AIC and BIC).

C.3.2 Effects for Each Treatment

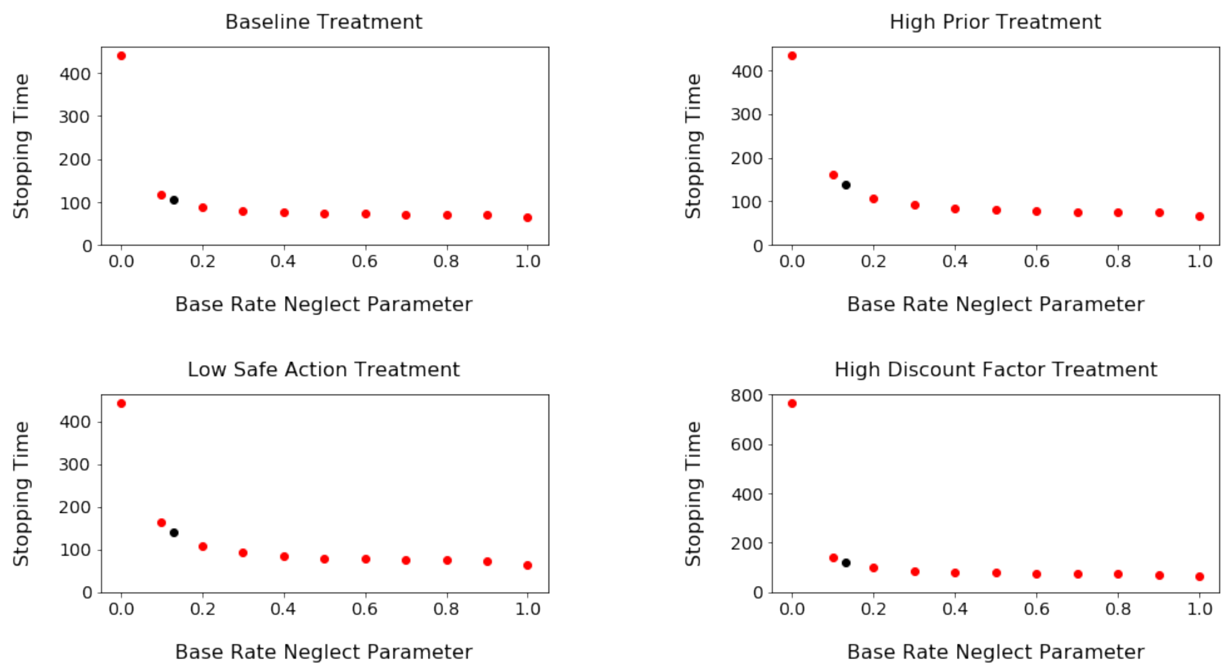
This subsection displays the effect of each behavioral parameter in each treatment. First, the effect of risk aversion is displayed in each treatment. Second, the effect of conservatism is displayed in each treatment. Third, the effect of non-linear probability weighting is displayed in each treatment.



The black dot denotes the Subset approach prediction from the model using the model's actual estimated CRRA coefficient.

Figure C.6. The Effect of Unilaterally Changing the CRRA Coefficient on Subset Approach Predictions from the Model

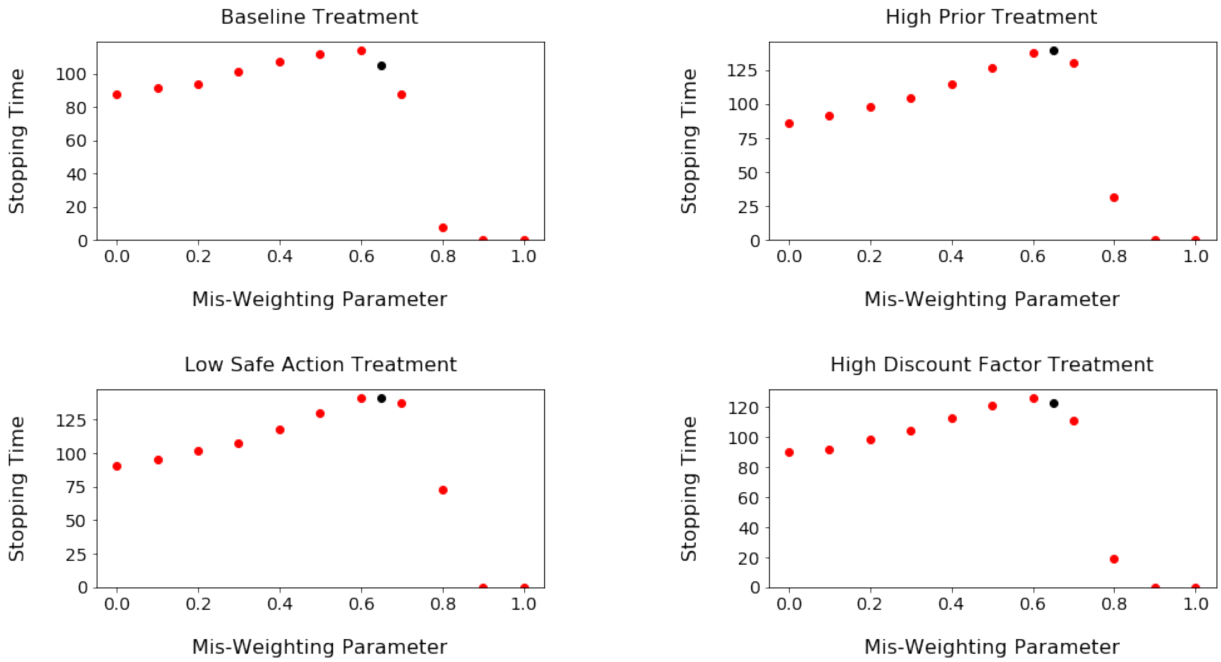
Figure C.6 displays the predictions for each treatment, using the Subset approach, when varying the CRRA coefficient. Essentially, these are the predictions using the Subset approach of varying γ while ψ, α, and σ are held constant. These graphs show that risk aversion is contributing to subjects' under-experimentation as risk neutral subjects, controlling for the other behavioral factors, would experiment longer in each treatment.



The black dot denotes the Subset approach prediction from the model using the model's actual estimated base rate neglect parameter.

Figure C.7. The Effect of Unilaterally Changing the Base Rate Neglect Parameter on Subset Approach Predictions from the Model

Figure C.7 displays the prediction of each treatment, using the Subset approach, when varying the base rate neglect parameter. Essentially, these are the predictions using the Subset approach of varying ψ while γ, α, and σ are held constant. These graphs show that conservatism is unilaterally making subjects want to experiment longer as an individual with $\psi = 1$, would experiment for a shorter period of time in each treatment.



The black dot denotes the Subset approach prediction from the model using the model's actual estimated probability weighting parameter.

Figure C.8. The Effect of Unilaterally Changing the Probability Weighting Parameter on Subset Approach Predictions from the Model

Figure C.8 displays the prediction of each treatment, using the Subset approach, when varying the non-linear probability weighting parameter. Essentially, these are the predictions using the Subset approach of varying α while γ, ψ, and σ are held constant. These graphs show that non-linear probability weighting in this environment is unilaterally making subjects want to experiment longer as an individual with $\alpha = 1$, would experiment for a shorter period of time in each treatment.

C.4 Power Analysis Appendix

In this section we describe our power analysis of Banks et al. (1997) and of our own paper. This section starts by describing the bandit problem in Banks et al. (1997) and continues on to describe their experimental design, the specifications used for the power analysis, and the results of the power analysis.

In their bandit problem, there is a safe action that pays out 50 tokens (five tokens is equivalent to one cent). The risky action has a 50 percent chance of being good. A good risky action has a probability of g of returning 100 tokens and a probability of $1-g$ of returning 0 tokens. A bad risky action has a probability of $1-g$ of returning 100 tokens and a probability of g of returning 0 tokens. In the problem, subjects discount future payoffs through a discount factor of δ.

There are four treatments in their experiment. The experiment uses a 2x2 factorial design where δ and g are varied. The predictions for each treatment are based on cutpoints, which is the difference between the number of high payoffs and low payoffs from the risky action that makes an agent indifferent between the risky and safe action. For example, if an agent is indifferent between the risky action and safe action when she has observed five low payoffs and three high payoffs from the risky action, her cutpoint is negative two. Due to the setup of the problem, each cutpoint results in a unique belief.[1] They analyze cutpoints by increments of 0.5.

The predictions for their four treatments is as follows. The first treatment, where $\delta = 0.8$ and $g = 0.7$, has a cutpoint of -0.5 and a cutoff belief of 0.32. The second treatment, where $\delta = 0.8$ and $g = 0.9$ has a cutpoint of -0.5 and a cutoff belief of 0.20. The third treatment, where $\delta = 0.9$ and $g = 0.7$, has a cutpoint of -1.0 and a cutoff belief of 0.23. The fourth treatment, where $\delta = 0.9$ and $g = 0.9$ has a cutpoint of -0.5 and a cutoff belief of 0.12.

The experiment consists of subjects repeatedly facing this bandit problem. The discount factor is induced through a random termination of the period. In each treatment, each subject repeatedly faces this bandit problem until either a maximum amount of time (60 minutes) occurs or a maximum number of periods (5) occurs.

They analyze the data by developing a best cutpoint for each subject. The best cutpoint is the cutpoint (c) that results in the smallest number of observed deviations from a subject's pooled data. They calculate a best cutpoint for each subject and run a regression of the cutpoint on a constant, δ, γ, $\delta \times \gamma$, a dummy variable for whether a subject is expe-

[1]↑This occurs because the probability of a high payoff in a good risky action is the complement of the probability of a high payoff in a bad risky action. In this environment, the Bayesian update can be written as $\frac{p_0(g)^{cutpoint}(1-g)^{-cutpoint}}{p_0(g)^{cutpoint}(1-g)^{-cutpoint}+(1-p_0)}$.

rienced, and the subject's elicited risk preference. The coefficients on δ (p-value=0.175), γ (p-value=0.616), and $\delta \times \gamma$ (p-value=0.148) are all insignificant at the ten percent level.

We conduct a power analysis of their paper in the following way. We assume that, in each treatment, subjects' preferred cutoff beliefs are normally distributed around the predicted cutoff belief. We focus on cutoff beliefs as it is a simple way to compare the power analysis of their paper to our paper as we have much higher predicted cutpoints. Additionally, we assume that, in each period, subjects' cutoff beliefs are normally distributed around their predicted cutoff belief. We are essentially assuming that subjects make mistakes in each period.

We obtain estimates for the between variation and within variation of cutoff beliefs from previous data. We utilize both the High Prior Treatment in this experiment and the single-agent treatment of Hudja (2019). We utilize these treatments as the prior belief is 0.5 in each treatment which is the same prior belief as in Banks et al. (1997). The initial prior belief is important as it gives us a range of possible cutoff beliefs that is similar to the range of possible cutoff beliefs in Banks et al. (1997). We obtain the between variation by taking the standard deviation of subjects' mean cutoff beliefs in each of our datasets. We obtain the within variation by taking the mean of each subject's cutoff belief standard deviation. We use a Subset approach in each dataset to get these variations.

We run 100 simulations of their experiment, for each source of noise, based on their design. There are twenty subjects in the first treatment, nineteen subjects in the second treatment, eighteen subjects in the third treatment, and nineteen subjects in the fourth treatment. In each simulation, subjects are given a cutoff belief that is randomly drawn from a normal distribution with the predicted cutoff belief as the mean and the between standard deviation taken from one of the two noise sources. We have each simulated subject play five periods of each bandit. In each period, each subject's cutoff belief is normally distributed with their preferred cutoff belief as the mean and the within standard deviation taken from one of the two noise sources.

Once a simulation is run, we calculate the best cutpoint for each subject in the following way. We count the number of observed deviations for cutpoints between -5 and 5 in increments of 1 for each subject. For example, if a subject draws when the difference between the high payoffs and low payoffs is equal to 1, then all cutpoints above 1 receive an observed deviation. We then treat the cutpoint that has the smallest number of deviations as the best cutpoint. In the case that multiple cutpoints have the smallest number of deviations, the midpoint of this best cutpoint interval is taken. This is similar to how best cutpoints are calculated in Banks et al. (1997).

For each simulation, and each set of noise, we run a regression of the best cutpoint on a dummy variable for a high value of δ, a dummy variable for a high value of g, and the interaction of the previous two dummy variables. This is the same regression that is run in Banks et al. (1997) except that risk aversion and experience is removed from the regression. We avoid risk and experience in order to simplify the analysis and this is what we would have done if we were doing an ex-ante power analysis.

We conduct this process for each level of noise. We measure power for a specific treatment variable by counting the number of times that there is neither a response to the relevant dummy variable nor the interaction that includes that relative dummy variable. Using the High Prior Treatment noise, we find that subjects respond to δ 74 percent of the time at the five percent level. Using the Hudja (2019) noise, we find that subjects respond to δ 61 percent of the time at the five percent level. Using the High Prior Treatment noise, we find that subjects respond to g 10 percent of the time at the five percent level. Using the Hudja (2019) noise, we find that subjects respond to g 7 percent of the time at the five percent level. Both of these are lower than the generally accepted 80 percent.

We conduct a power analysis of our paper by simulating the experiment 100 times for each level of noise. We use the exact setup of the experiment, but once again assume that subjects' preferred cutoff beliefs are normally distributed around the predicted cutoff belief (using between standard deviation from one of our two noise sources) and that subjects' cutoff beliefs in each period are normally distributed around their preferred cutoff belief (using within standard deviation from one of our two noise sources). We analyze responses to the treatment variables in the same way that we do in the results section. For each level of noise, we find that subjects always respond to a change in p_0, $s\Delta$, and δ.

C.5 Instructions

In this section, we present the text in the instructions that is common to all sessions. We then annotate this with the text unique to each session. For the High Discount Factor sessions we use italics *like this*, for the High Prior sessions we use square parentheses [like this], and for the Low Safe Action session we use curly brackets {like this}.

Instructions

This experiment is a study of economic decision making. The amount of money that you earn depends partly on the decisions that you make and thus you should read these instructions carefully. The money that you earn will be paid privately to you, in cash, at the end of the experiment.

At the start of the experiment, you can earn \$5.00 by answering five comprehension questions about these instructions. For each correct answer to a question you will earn \$1.00. You can refer to these written instructions as you answer the questions.

From this point forward, all units of account will be in experimental points. At the end of the experiment, points will be converted to U.S. dollars at the rate of 1 U.S. dollar for every 100 points (i.e. 1 point is worth \$0.01).

Overview: Bags

In each period of this experiment, you will make decisions on whether to draw from a bag. In this experiment, imagine that there are two types of bags. The first type of bag is a 'mixed' bag (denoted by the letter M). **An M bag contains 1 red ball and 99 yellow balls.** Thus, if you draw a ball from an M bag, there is a 1 percent chance that you will draw a red ball and a 99 percent chance that you will draw a yellow ball. The second type of bag is a 'uniform' bag (denoted by the letter U). **A U bag contains 0 red balls and 100 yellow balls.** Thus, if you draw a ball from a U bag, there is a 0 percent chance that you will draw a red ball and a 100 percent chance that you will draw a yellow ball. After every draw the drawn ball is replaced back into the bag, so the bag contents and the chances of drawing the balls of each color from your current bag never change.

At the start of a new period, an M bag or a U bag will be randomly chosen. [The probability that the U bag will be chosen is listed on the experimental interface to the right of "Initial M Bag Prob". For example, imagine that at] *At* {At} the start of a new period, there is a 33.3 percent chance that the M bag is chosen and a 66.7 percent chance that the U bag is chosen. It is as if a six-sided die is rolled in the beginning of the period. If the die lands on 1, or 2 an M bag is used. If the die lands on 3, 4, 5 or 6 a U bag is used.

The type of bag does not change within a period. Balls you draw within the same period will all be drawn from the same bag, and the drawn ball will be put back into the bag after each drawing, **so the contents of the bag do not change within a period.**

Drawing

Each period consists of many 'ticks'. Each tick lasts for a fifth of a second (i.e. five ticks per second), and ticks continually occur until the period ends. How a period ends will be discussed later. In each tick, you may draw a ball from the bag. You will be asked whether or not you would like to initially draw a ball. If you initially choose to draw a ball, you will keep drawing a ball in each tick until you decide to stop drawing a ball. **If you ever choose to stop drawing a ball, you can no longer draw a ball for the rest of the ticks in the period.** This also means that if you decide to initially not draw a ball, you cannot draw a ball for any of the ticks in the period.

If you draw a ball, this ball will either be red or yellow. If you draw a red ball, you earn **155 points** in the current tick. **Once you have drawn one red ball you automatically draw** a ball for the rest of the ticks in the period. Notice that you can draw multiple red balls in a period. For example, if you draw a red ball, you can earn a period payoff of 155, 310, 465 points, etc., based on how many red balls you draw in that period. If you draw a yellow ball you will earn 0 points in the current tick. If you have drawn yellow balls in all ticks so far, you can choose to stop drawing a ball. If you choose to stop drawing a ball you will receive {a payoff in each of the remaining ticks. This payoff is listed on the experimental interface to the right of "Not Drawing Payoff". For example, imagine that this payoff is 0.50 points. If you chose to stop drawing a ball you will receive} 0.50 points in the current tick and 0.50 points in each of the remaining ticks. *For* [For] {In this} example, if you do not draw for 100 ticks, you receive 50 experimental points for those 100 ticks.

How a Period Ends

There is a probability that the period will end in the current tick. The probability that the period will end in the current tick is [listed on the experimental interface to the right of "Prob. Tick Ends Period". For example, imagine that the probability that the current tick ends the period] is 0.4 percent. It is as if 3000 tickets, numbered 1 through 3000 are placed in a box and a ticket is randomly drawn after every tick. If the number on the ticket is 1 through 12, the period ends. If the number on the ticket is 13 through 3000, the period continues and the ticket is placed back in the box, so the contents of the box never change. Under this [example] probability, the average period length is 250 ticks (i.e. 50 seconds).

Blocks

You will be participating in two twenty period blocks. **Throughout the experiment, only the probability that the *period will end in the current tick* [M bag is chosen at the start of the period] {payoff for not drawing} can change.** This *probability* [probability] {payoff} can only change between blocks. Within each block, this *probability* [probability] {payoff} stays the same.

At the start of a block, the *probability that the period will end in the current tick* [M bag is chosen at the start of the period] {payoff for not drawing} will be displayed. In order to start the block, you are required to correctly answer a question on this *probability* [probability] {payoff}.

You will be paid for three random periods in each block (and for your answers to the comprehension questions).

Interface

To learn about the interface, please watch the video being shown on the projector.

Summary

- There are two types of bags: the M bag with 1 red ball and 99 yellow balls and the U bag with 0 red balls and 100 yellow balls
- At the start of each period, [the M bag is randomly chosen based on the probability listed to the right of "Initial Prob. of M bag" on the experimental interface. This probability can only change between blocks. Within each block, this probability stays the same.] *there is a 33.3 percent chance that the M bag is randomly chosen for the period and a 66.7 percent chance that the U bag is randomly chosen for the period* {there is a 33.3 percent chance that the M bag is randomly chosen for the period and a 66.7 percent chance that the U bag is randomly chosen for the period}
- You earn 155 points for each red ball that you draw. If you ever draw a red ball in a period, you automatically draw a ball in the remainder of the ticks in the period. You can draw more than one red ball in a period
- You earn 0 points for each yellow ball that you draw
- You earn *0.5 points* [0.5 points] {the payoff listed to the right of "Not Drawing Payoff" on the experimental interface} for each tick that you do not draw a ball. {This payoff can only change between blocks. Within each block, this payoff stays the same.} Once you decide to not draw a ball, you can no longer draw a ball for the rest of the ticks in the period

- *The probability that the current tick will end the period is listed to the right of "Prob. Tick Ends Period" on the experimental interface. This probability can only change between blocks. Within each block, this probability stays the same.* [There is a 0.4 percent chance that the period will end in the current tick.] {There is a 0.4 percent chance that the period will end in the current tick.}

VITA

EDUCATION

- Currently Pursuing: *Doctorate of Philosophy*, Economics, Purdue University, 2015-2021.
- *Master of Commerce with Distinction, Economics*, University of Canterbury, 2012-2013.
- *B.Sc. (Hons.) (First Class), Economics,* University of Canterbury, 2006-2010.

RESEARCH INTERESTS

- Behavioral, Experimental, and Computational Economics, Industrial Organization.

WORKING PAPERS

- "Improving Ex-ante Power Analysis With Quantal Response Simulations"
- "Behavioral Bandits: Analyzing the Exploration Versus Exploitation Trade-off in the Lab" (with Stanton Hudja).

PUBLICATIONS

- "Network Defense and Behavioral Biases: An Experimental Study" (with Mustafa Abdallah, Saurabh Bagchi, Shreyas Sundaram, and Timothy Cason). ***Experimental Economics*** (forthcoming).
- "Nice to You, Even Nicer to Me: Does Self-Serving Generosity Diminish Reciprocal Behavior?" (with Maroš Servátka), ***Experimental Economics*** (2019), 22, 506-529.
- "Price-Setting and Attainment of Equilibrium: Posted Offers Versus An Administered Price" (with Sean Collins, Duncan James, and Maroš Servátka), ***Games and Economic Behavior*** (2017), 106, 277-293.

- "Testing Psychological Forward Induction and the Updating of Beliefs in the Lost Wallet Game" (with Maroš Servátka), ***Journal of Economic Psychology*** (2016), 56, 116-125.

RESEARCH GRANTS

- IFREE Small Grants Program, The International Foundation for Research in Experimental Economics, 2019.
- Doctoral Research Funds Grant, Krannert School of Management, 2019.

TEACHING

- Instructor, Purdue University: Game Theory (2017).
- Teaching Assistant, Purdue University: Principles of Microeconomics (2015), Intermediate Macroeconomics (2017), Game Theory (2018).
- Teaching Assistant, University of Canterbury: Introduction to Microeconomics (2010), Introduction to Macroeconomics (2010), Game Theory (2012, 2013).

HONORS AND AWARDS

- Purdue University: Ross Fellowship, 2015, PGSG Professional Grant, 2017, Krannert Certificate for Distinguished Teaching, 2017, Bilsland Dissertation Fellowship, 2018.
- University of Canterbury: University Prizes, 2014, UC Master's Scholarship, 2011, Summer Research Scholarship, 2009 and 2011, Economics Department Book Prize, 2008, Madam Tiong Guok Hua Memorial Prize in Economics, 2007.
- New Zealand Association of Economists: Graduate Study Award, 2013. New Zealand Qualifications Authority: NCEA Scholarship in Economics and Statistics, 2006.

CONFERENCES

- Presented *Behavioral Bandits: Analyzing the Exploration Versus Exploitation Trade-off in the Lab* at the ESA Job-market Candidates Seminar Series 2020.

- Presented *Using QRE Simulations for Power Analysis* at the ESA Global Meeting 2020.
- Presented *Competition in Information: An Experimental Test of Bayesian Persuasion in Search Markets* at the Southern Economic Association Annual Meetings 2018, the Australia New Zealand Workshop in Experimental Economics 2018, and the ESA World Meeting 2019.
- Presented *Does Self-serving Generosity Diminish Reciprocal Behaviour?* at the Bratislava Economic Meeting 2014, the Asia-Pacific Meeting of the Economic Science Association 2014, and the 54th New Zealand Association of Economists Annual Conference 2013.
- Presented *Testing Psychological Forward Induction in the Lost Wallet Game* at the 53rd New Zealand Association of Economists Annual Conference 2012.