ESSAYS ON EXPERIMENTAL ECONOMICS AND INNOVATION

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Stanton Hudja

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 2020

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL

Dr. Tim Cason, Co-Chair

Department of Economics

Dr. Brian Roberson, Co-Chair

Department of Economics

Dr. David Gill

Department of Economics

Dr. Yaroslav Rosokha

Department of Economics

Approved by:

Dr. Brian Roberson

Director of Economics Doctoral Program

For Hope, Uziel, and Miriam.

ACKNOWLEDGMENTS

To my thesis committee Tim Cason (co-chair), Brian Roberson (co-chair), David Gill, and Yaroslav Rosokha, thank you for always being approachable, but also letting me make my own mistakes. I appreciate that I was able to struggle through my own research problems and approach all of you when I needed guidance. Your comments along the way have helped me become a better researcher and to see all of the hard work that is put into research.

Over my six years at Purdue University, I have been delighted to share an office with Mitch Johnston, Peter Wagner, Daniel Woods, Mary Kate Batistich, Chen Wei, and Junya Zhou. Thank you all for being great officemates and letting me obsess over small details. I would also like to thank Ben Raymond, Haiqing Zhao, and Prith Chaudhuri for always allowing me to use their office for econometrics related questions.

Finally, I would like to thank my friends and family for their ongoing support and encouragement. To Hope, Uziel, Georgina, Robin, Miriam, Jay and Laurie thank you for always being supportive throughout this process. To Monica, thank you for always keeping me grounded. To Joseph, thank you for always being a great friend.

TABLE OF CONTENTS

Page

Page

Page

Page

Page

LIST OF TABLES

Table Page

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

ABSTRACT

Hudja, Stanton Ph.D., Purdue University, May 2020. Essays on Experimental Economics and Innovation. Major Professor: Tim Cason and Brian Roberson.

My dissertation consists of four chapters. In the first chapter, I use a laboratory experiment to analyze how individuals resolve an exploration versus exploitation trade-off. The experiment implements a single-agent exponential bandit model. I find that, as predicted, subjects respond to changes in the prior belief, safe action, and discount factor. However, I commonly find that subjects give up on exploration earlier than predicted. I estimate a structural model that allows for risk aversion, base rate neglect/conservatism, and probability mis-weighting. I find support for risk aversion, conservatism, and probability mis-weighting as potential factors that influence subject behavior. Risk aversion appears to contribute to the finding that subjects explore less than predicted.

In the second chapter, I use a laboratory experiment to analyze how a group of voters experiment with a new reform. The experiment implements the continuous time Strulovici (2010) collective experimentation model. I analyze a subset of data where groups and single decision makers should eventually prefer to stop experimentation and abandon the reform. I find three results that are consistent with the modeled experimentation incentives. In this subset of data, groups stop experimentation earlier than single decision makers, wait longer to stop experimentation as the number of revealed winners increases, and stop experimentation earlier than the utilitarian optimum predicts. However, I also find that both groups and single decision makers stop experimentation earlier than predicted. Additional treatments show that this result is unlikely to be explained by standard explanations such as incorrect belief updating or risk aversion.

In the third chapter, I use a laboratory experiment to investigate the role of group size in an innovation contest. Subjects compete in a discrete time innovation contest, based on Halac et al. (2017), where subjects, at the start of each period, are informed of the aggregate number of innovation attempts. I compare two innovation contests, a two-person and four-person contest, that only differ by contest size and have the same probability of obtaining an innovation in equilibrium. The four-person contest results in more innovations and induces more aggregate innovation attempts than the two-person contest. However, there is some evidence that the two-person contest induces more innovation attempts from an individual than the four-person contest. Subjects' behavior is consistent with subjects placing more weight on their own failed innovation attempts, when updating their beliefs, than their competitors' failed innovation attempts.

In the fourth chapter, I investigate the role of performance feedback, in the form of a public leaderboard, in innovation competition that features sequential search activity and a range of possible innovation qualities. I find that in the subgame perfect equilibrium of contests with a fixed ending date (i.e., finite horizon), providing public performance feedback results in lower equilibrium effort and lower innovation quality. I conduct a controlled laboratory experiment to test the theoretical predictions and find that the experimental results largely support the theory. In addition, I investigate how individual characteristics affect competitive innovation activity. I find that risk aversion is a significant predictor of behavior both with and without leaderboard feedback and that the direction of this effect is consistent with the theoretical predictions.

INTRODUCTION

This thesis analyzes how individuals resolve the exploration versus exploitation trade-off. Throughout this thesis, I explore how individuals decide between exploiting a predictable payoff and exploring an unfamiliar, but potentially lucrative, option. I am concerned with three main questions. First, how well does economic theory capture the resolution of this trade-off? Second, how can we improve upon our current mechanisms for incentivizing exploration? Third, what behavioral factors influence individuals to explore unfamiliar options? While the papers in this thesis analyze different environments, they are connected by their attempt to address these questions.

In the first chapter, I analyze how individuals resolve a simple exploration versus exploitation trade-off in the laboratory. I implement the single-agent exponential bandit model in a laboratory experiment. In this study, I am directly addressing my first and third questions. I address the first question by analyzing how well the single-agent exponential bandit model describes individual behavior. I focus on how individuals respond to changes in the prior belief, safe action, and discount factor. I find that subjects respond in the correct direction to changes in these incentives. Subjects become more willing to experiment as their prior belief that experimentation is efficient increases, as the opportunity cost of experimentation decreases, and as they are induced to become more patient. However, I find that individuals under-experiment relative to the predictions of the single-agent exponential bandit model.

I both reconcile the discrepancy between theory and behavior and address the third question with a structural model. I incorporate risk aversion, base rate neglect/conservatism, and probability mis-weighting into the model of experimentation. Through maximum likelihood estimation, I find that each factor appears to contribute to experimentation. However, risk aversion appears to be the main factor that is contributing to the under-experimentation in the experiment. In fact, upon controlling for risk aversion, subjects appear to over-experiment.

In the second chapter, I analyze an environment where agents face an exploration versus exploitation trade-off that is complicated by the presence of other agents. I implement the Strulovici (2010) collective experimentation model in the laboratory. In this paper, individuals, with heterogeneous preferences for exploration, must collectively decide whether or not to experiment. I once again address the first and third questions. I address the first question by focusing on the difference between group experimentation and single-agent experimentation. Groups are predicted to be less willing to experiment than single-agents because of negative payoff exernalities that arise in the voting context. I find that subjects' behavior is consistent with this result. However, I find that both groups and single-agents experiment for less time than predicted.

I address both this under-experimentation and the third question by designing additional treatments. I explore three possible explanations for this under-experimentation: i) subjects may over-weight the information generated from experimentation relative to their prior, ii) subjects may overweight the random termination probability and iii) risk aversion. I design two new treatments. The first treatment removes the possibility of over-weighting the information generated from experimentation. The second treatment removes this same possibility and additionally removes the possibility that subjects over-weight the random termination probability. Additionally, I elicit risk aversion in each treatment. I find under-experimentation in each treatment and that there is a statistical difference between each of these treatments and the main experiment. I additionally find that risk aversion is not statistically significant in each treatment.

In my third chapter, I analyze the role of group size on exploration in innovation contests. I implement a discrete time version of a continuous time contest from Halac et al. (2017), which predicts invariance to contest size. In this paper, I address the first and second questions. I address these questions by implementing a two and four-person innovation contest, which have the same probability of resulting in an innovation in equilibrium, in the laboratory. I find that the four-person contest results in more innovations and induces more aggregate effort from contestants than the two-person contest. I attempt to address the result that four-person contests result in more innovations, and induce more aggregate effort, by developing a structural model based on differential weighting of experimentation. Under differential weighting of experimentation, subjects place more weight on the information generated by their own experimentation than information generated by other subjects' experimentation. Through Maximum Likelihood Estimation, the data appears to be consistent with differential weighting of experimentation.

In my fourth chapter, I analyze the role of leader-board feedback on exploration in research tournaments. In this paper, I analyze a research tournament with and without leader-board feedback in the laboratory. I address the first and second questions by analyzing the aggregate number of draws and the average quality of the winning innovation in each contest. I find that leader-board feedback actually reduces the number of aggregate draws in a research tournament and reduces the average quality of the winning innovation.

In addition to the first and second questions, this paper also addresses the third question. I analyze an individual's willingness to innovate in the contest with leader-board feedback, the contest without leader-board feedback, and an individual innovation task. The individual innovation task is designed to analyze an individual's willingness to innovate without competition. I analyze the effect of risk aversion, loss aversion, and the sunk cost fallacy on an individual's willingness to attempt an innovation. Additionally, I analyze the effect of grit, the big five characteristics, competitiveness, and achievement-striving on an individual's willingness to innovate. I find that risk aversion influences an individual's willingness to innovate in each environment. Risk aversion is the only factor that appears to influence behavior in all three environments.

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

with Daniel Woods

This paper uses a laboratory experiment to analyze how individuals resolve an exploration versus exploitation trade-off. The experiment implements a single-agent exponential bandit model. We find that, as predicted, subjects respond to changes in the prior belief, safe action, and discount factor. However, we commonly find that subjects give up on exploration earlier than predicted. We estimate a structural model that allows for risk aversion, base rate neglect/conservatism, and probability mis-weighting. We find support for risk aversion, conservatism, and probability mis-weighting as potential factors that influence subject behavior. Risk aversion appears to contribute to the finding that subjects explore less than predicted.

1.1 Introduction

The dilemma of whether to explore an uncertain option or exploit a familiar option is common in economics. For example, a CEO often chooses between investing resources into a new market and an established market. A farmer often chooses between planting a new crop and an old crop. A researcher often chooses between starting a new research agenda and continuing an old one. In these and many other examples, individuals must decide whether to forgo a predictable payoff to learn more about an uncertain, but potentially lucrative, option.

The single-agent exponential bandit model provides a simple model of this exploration versus exploitation trade-off. In economics, many models of exploration build on this rudimentary model. This model is a starting point 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 and Hörner, 2011), and voting for reforms (Strulovici, 2010; Khromenkova, 2015). While many models build on this bandit model, there is little empirical research on how well it describes individuals' resolution of the exploration versus exploitation trade-off.

In this paper, we aim to analyze how well this model describes individual behavior and to uncover behavioral factors that influence exploration. We address two main questions in this paper. First, how well does this model describe an individual's resolution of the exploration versus exploitation trade-off? Specifically, we focus on how well it describes individuals' response to changes in incentives and the decisions that they make. This question is important as inaccurate predictions that arise from this model are likely to arise in models that build on it. Second, which, if any, unaccounted-for behavioral factors are consistent with individual behavior? This question is important as models of exploration may not be including relevant factors.

We analyze the single-agent exponential bandit model in a laboratory experiment. We utilize a laboratory experiment as it allows us to analyze a setting closely resembling the model environment. In our experiment, an agent 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 bad risky action is dominated by the safe action and never pays out a reward. A good risky action dominates the safe action and occasionally pays out a (high) reward. An agent 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 continue to decrease.

A subject in this environment faces a trade-off. Experimentation with the risky action provides valuable information that the agent can use to update her belief that her risky action is good and provides the possibility of a high reward. 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 risky action is good is sufficiently high. If her belief ever drops below this cutoff belief, the cost of experimentation outweighs the immediate and long-term benefits of experimentation 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, value of the safe action, and prior belief. A subject is predicted to be willing to experiment longer as the discount factor increases because the option 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 is predicted to be willing to experiment longer 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. The High Prior, Low Safe Action, and High Discount Factor treatments only differ from the Baseline treatment by one parameter. The High Prior treatment has a higher prior than the Baseline treatment. The Low Safe Action treatment has a lower value of the safe action than the Baseline treatment. The High Discount Factor treatment induces a higher discount factor than the Baseline treatment.

The experimental data is used to test three hypotheses, with each subsequent hypothesis test being a less conservative test of the model's predictions. The first hypothesis, which considers only 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, we commonly find that subjects experiment less than predicted. There is strong evidence of under-experimentation in three of our four treatments and mixed evidence of under-experimentation in the other.

The experimental results suggest that subjects have unaccounted-for behavioral factors that influence their experimentation. The variation in the experimental treatments allows us to uncover these possible behavioral factors through Maximum Likelihood Estimation. We incorporate risk aversion, base rate neglect/conservatism, and probability mis-weighting into a model of experimentation. We find that subjects' behavior is consistent with risk aversion, conservatism, and probability mis-weighting. Risk aversion appears to be the main reason why subjects under-experiment.

This paper contributes to three strands of literature. The first 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), as well as Khromenkova (2015), analyzes a reverse setting where experimentation may influence individuals heterogeneously, but agents must collectively decide whether to experiment or not. The theoretical literature on experimentation has also analyzed moral hazard in teams (Bonatti and Hörner, 2011), long-term contracting (Halac et al., 2016), and innovation contests (Halac et al., 2017; Bimpikis et al., 2019). Our paper suggests that the comparative statics on the prior belief, cost of experimentation, and discount factor in these theoretical models should hold up in empirical studies. However, our paper also suggests that these models may be predicting too much experimentation and that risk aversion, conservatism, and probability mis-weighting should be considered when modeling experimentation environments.

The second is the literature on bandit experiments. Our paper relates to a type of bandit experiment where the risky action has a high reward probability that can only take on one of two known values. Papers that analyze this type of bandit problem are Banks et al. (1997), Hoelzemann and Klein (2018), and Hudja (2019). Our paper addresses two questions in this literature. The first question is whether subjects respond to incentives. Banks et al. (1997) fail to show that subjects respond to changes in the discount factor and high reward probabilities in a discrete time bandit where the bad risky action also pays out rewards. Hoelzemann and Klein (2018) analyze 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. Our paper is closest to Banks et al. (1997) and suggests that subjects do respond to changes in environmental parameters like the discount factor. Power analyses suggest that their paper is under-powered and our paper is well-powered. Our experiment, coupled with the power analyses, suggests that subjects respond to changes in environmental parameters. The second question is the role of risk aversion in experimentation. Banks et al. (1997) and Hudja (2019) fail to show that elicited risk aversion is correlated with subject behavior. Our paper suggests that risk aversion does play a role in experimentation and that it appears to contribute to under-experimentation. This difference may be due to the fact that we estimate risk aversion through maximum likelihood, while the previous papers do not.

The third is the literature on continuous time experiments. Continuous time experiments have mostly consisted of two types of experiments: (i) continuous time versions of classic discrete time games and (ii) 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; Anderson et al., 2010; Oprea, 2014) or Poisson Processes (Hoelzemann and Klein, 2018; Hudja, 2019) in the lab. Our experiment suggests that subjects may make mistakes when dealing with Poisson Processes as their behavior is consistent with probability mis-weighting.

1.2 Theory

The theory motivating this experiment is based on Keller et al. (2005).[1] The reader seeking greater detail than this section provides should consult Keller et al. (2005).

Time (t) is continuous and payoffs are discounted at a rate r. There is an individual who 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 bad risky action pays out nothing. A good risky action pays out a reward (magnitude h) at random times based on a Poisson process with parameter λ. 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 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

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

where t is the amount of time spent experimenting.[2] 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 and 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 (p) is greater than or equal to a cutoff belief p_A. The cutoff belief p_A is given by

[1]We test the predictions of the single-agent exponential bandit model in a slightly restricted environment. Subjects must choose between the safe and risky action in each period of time, they can not divide a resource between the two actions. This restriction has no bearing on theoretical predictions and is consistent with similar exponential bandit experiments such as Hoelzemann and Klein (2018). The theory that follows is for this restricted case.

[2]Experimentation occurs when an agent chooses the risky action while she is still uncertain of the state of the risky action (good or bad).

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

This cutoff belief solves the indifference equation $p\lambda h + \lambda p(\frac{\lambda h}{r} - \frac{s}{r}) = s$. The left-hand side of the indifference equation corresponds to the risky action, while the right-hand side corresponds to the safe action. 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.

One take-away from the cutoff belief is that agents are non-myopic. A myopic individual experiments if and only 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 experiments if and only if $p \geq \frac{s}{\lambda h}$. Agents are non-myopic as the predicted cutoff belief is below $\frac{s}{\lambda h}$. Myopic behavior is suboptimal as an agent does not value the information generated by experimentation (the future value of learning that the risky action is good).

The laboratory experiment focuses on how long individuals are willing to experiment, that is, how long they would choose the risky action in the absence of any rewards. The time that an individual is willing to experiment is found by solving for how long it would take for an individual's belief to reach her cutoff belief in the absence of any rewards from the risky action. This time is given by

$$\frac{-ln\left(\frac{p_A(1-p_0)}{p_0(1-p_A)}\right)}{\lambda},$$

where p_A is the cutoff belief and p_0 is the prior belief. 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. An agent is predicted to experiment longer as the value (s) of the safe action decreases. Intuitively, as the value of the safe action decreases, the opportunity cost of experimentation decreases and individuals become willing to experiment longer as they are willing to experiment at lower beliefs. An agent is also predicted to experiment longer as the discount factor (1 - r) increases. Intuitively, as the discount factor increases, individuals place more value on the future rewards associated with a good risky action and individuals become willing to experiment longer as they are willing to experiment at lower beliefs. Lastly, an agent 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 individuals to reach their cutoff belief in the absence of rewards and thus individuals become willing to experiment longer.

1.2.1 Discrete Implementation

This paper 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 continuous time bandit problem in the laboratory. The approximation of this model is based on 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.

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 δ, which is equal to $1 - r\Delta$.

Table 1.1.: Values of p_0, $s\Delta$, and δ for each treatment. Each treatment has a value of $\lambda\Delta = 0.01$ and $h = 155$. Myopic predictions are in parentheses.

Treatment	Prediction	Prior	Safe Action	Discount Rate
Baseline:	130 (5)	$p_0 = 0.333$	$s\Delta = 0.5$	$\delta = 0.996$
High Prior:	199 (74)	$p_0 = 0.500$	$s\Delta = 0.5$	$\delta = 0.996$
Low Safe Action:	198 (74)	$p_0 = 0.333$	$s\Delta = 0.3$	$\delta = 0.996$
High Discount Factor:	199 (5)	$p_0 = 0.333$	$s\Delta = 0.5$	$\delta = 0.998\bar{3}$

1.3 Experimental Design

The experiment is designed to analyze how well the single-agent exponential bandit model describes individual behavior. The primary goal of the experiment is to determine how subjects respond to changes in incentives. The experiment has two secondary goals. The first secondary goal is to determine whether subjects make optimal decisions. The second secondary goal is to create a dataset that allows us to econometrically test for unaccounted-for behavioral factors that may influence experimentation.

1.3.1 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.[3] The tick length Δ is set to 200 milliseconds in all treatments. Table 1.1 displays the treatments used in the experiment.

[3]We avoid analyzing these two variables. We choose to not analyze the effect of $\lambda\Delta$ on experimentation as the effect can be non-monotonic. 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 choose to not analyze the effect of h as only the ratio of $\frac{h}{s\Delta}$ matters for predictions.

Table 1.2.: 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

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 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 1.1 displays the predictions for each treatment. Section A.1 of the appendix provides details for how we derived these predictions. In the Baseline treatment, subjects are predicted to be willing to experiment for 130 ticks.[4] 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.

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

Table 1.2 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 treatment, an agent's predicted cutoff belief in the discrete time approximation is within one one-hundredth of their predicted cutoff belief in continuous time.

1.3.2 Experiment

Instructions for the experiment were read aloud at the start of the experiment. 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. Upon completion of the five comprehension questions, the session began.

The environment is described through an analogy of balls being drawn 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 a (ii) "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 drawn for each subject. The mixed bag is drawn with probability p_0 and the uniform bag is drawn with probability $1 - p_0$. The bag stays the same throughout the period.

At the start of each period, subjects have as much time as they would like to take 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 single decision maker, she decides to stop. Starting from the initial action, whenever a subject is

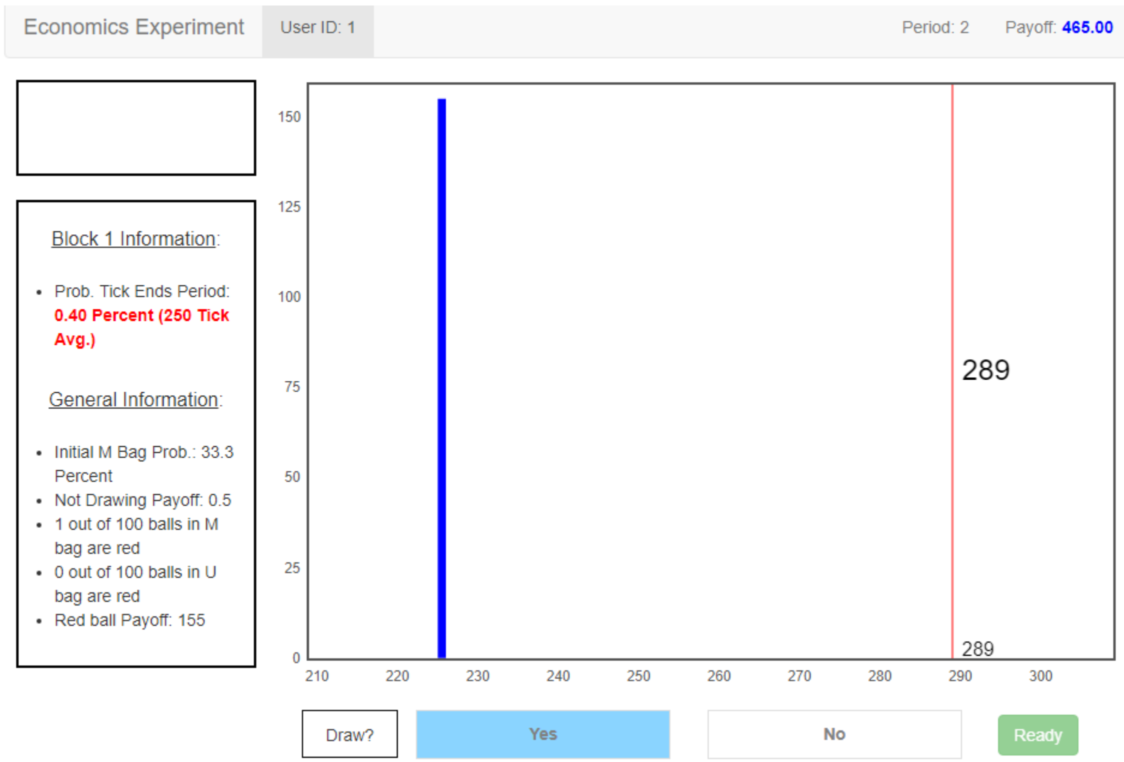


Figure 1.1.: Example of the experimental interface.

unsure of the risky action and chooses to not draw, she is prevented from drawing for the rest of the period.[5] Additionally, a subject is prevented from switching to the safe action once she obtains a reward.[6] Ticks continue until the random termination of the period.

Subjects receive feedback throughout the period through a graph displayed on their screen. A red line, that does not move, is drawn 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 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

[5]This decision should have no theoretical effect and gives the theory the best chance as it removes some possible decision error. This does leave the possibility that subjects trembling could lead to under-experimentation, since a tremble to stop is irreversible.

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

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 1.1 displays an example of the screen from a High Discount Factor session.

At the end of the experiment, subjects completed a post-experimental survey, which is displayed in section A.7 of the appendix. The post-experimental survey collected information on gender, race, country of origin, grade point average, year of schooling, and major. These demographics were collected to control for any heterogeneity in the treatments.[7]

1.3.3 Testable Hypotheses

The hypotheses are based on the theoretical predictions for the four treatments. The hypotheses focus on how long an individual is willing to experiment, that is, how long an individual 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 treatment, 198 ticks in the Low Safe Action treatment, and 199 ticks in the High Discount Factor treatment.

Eliciting how long an individual is willing to experiment is not always possible. 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, an individual 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 agent switches to

[7]We do not present the demographics in the results section because the demographics are generally non-informative and do not change any major results from the experiment. The role of gender may be interesting for future experiments to uncover as men appear to be more willing to experiment than women and the p-value associated with gender is slightly higher than 0.10. Risk aversion may play a small role in this difference. Using the structural estimation in section 1.5, and stratifying by gender, men have an estimated CRRA coefficient that is about 0.01 less than women's estimated CRRA coefficient. Men and women also have differences in base rate neglect/conservatism and probability mis-weighting that appear to contribute to this difference.

the safe action. In this case, it is never observed when an agent would stop experimenting.

This paper takes two approaches to mitigate these issues. The first approach is to analyze a subset of data where theory always predicts a switch to the safe action. 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.[8] The second approach is to use the Product Limit estimator to correct for the censoring that occurs. Details on the Product Limit estimator can be found in section A.2.1 of the appendix. These two approaches will be used in tandem to test the three hypotheses.

The hypotheses test how well theory describes subjects' willingness to experiment. Each subsequent hypothesis is a less conservative 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 leads us to 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.*

[8] Other cutoffs can be used. For example, the results are robust to using a cutoff of 250, that is, the period lasts at least 250 ticks and I use observations where the state is bad or the first reward occurs after 249 ticks.

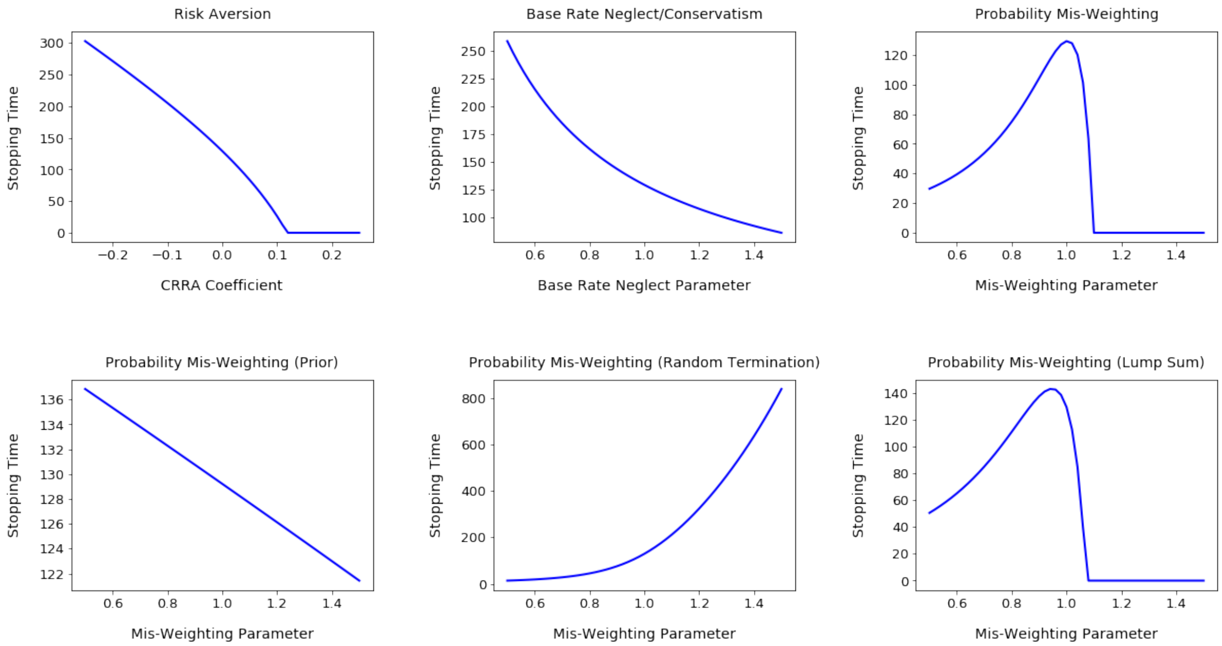


Figure 1.2.: First row of graphs display the effects of unilaterally increasing risk aversion, conservatism, and probability mis-weighting for the Baseline treatment. The second row of graphs display the effects of unilaterally increasing probability mis-weighting of the prior, random termination probability, and reward probability for the Baseline treatment.

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.

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

1.3.4 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 1.2. Subjects may exhibit risk aversion, base rate neglect/conservatism, and/or probability mis-weighting. This subsection focuses on how these behavioral factors may influence subject behavior.

Figure 1.2 displays how subject behavior in the Baseline treatment is predicted to change when risk aversion, base rate neglect, or probability mis-weighting is unilaterally varied.[9] These behavioral factors have similar effects on the other treatments. Risk aversion is modeled through CRRA utility. The first graph suggests that an individual becomes less willing to experiment as she becomes more risk averse. This is unsurprising as Keller et al. (2019) show that, in continuous time, experimentation is decreasing in risk aversion when $h > s$. Base rate neglect is modeled as each subject treating a yellow ball drawn as if it were ψ yellow balls drawn.[10] As ψ increases, subjects' beliefs decrease faster in the absence of any rewards. Unsurprisingly, the second graph suggests that subjects become less willing to experiment as ψ increases. Lastly, probability mis-weighting is modeled through a Prelec-I function.[11] Subjects may mis-weight the prior, reward probability, and random termination probability. The third graph shows a non-monotonic effect of probability mis-weighting on a subject's willingness to experiment.

The remaining graphs provide intuition for this non-monotonic effect. As the probability mis-weighting parameter (α) increases, the mis-weighted reward probability, random termination probability, and prior probability decrease for the Baseline treatment. The fourth graph suggests that as α increases, the mis-weighted prior de-

[9] Note that these are predictions for the discrete time approximation. More details on how these factors are modeled can be found in section 1.5.

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

[11] The Prelec-I function is given by $w(p) = e^{-(-ln(p))^{\alpha}}$. 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.

creases how long subjects are willing to experiment. Intuitively, subjects' perception of the prior is decreasing in α. The fifth graph suggests that as α increases, the mis-weighted random termination probability makes subjects willing to experiment longer. Intuitively, subjects' expectation of the period length is increasing in α. Lastly, the sixth graph suggests a non-monotonic effect of the mis-weighted reward probability on how long an individual is willing to experiment. Intuitively, beliefs decrease slower in the absence of rewards and the myopic value of experimentation decreases when α increases.

It is useful to reiterate the directional effects of these behavioral factors. First, risk aversion appears to decrease experimentation. Second, base rate neglect ($\psi > 1$) appears to decrease experimentation and conservatism ($\psi < 1$) appears to increase experimentation. Third, the effect of probability mis-weighting depends on the value of α.

1.3.5 Procedures

Experiments were run in the Vernon Smith Experimental Economics Laboratory at Purdue University. Experiments were run in August and September 2019. Experiments lasted for as little as 60 minutes and for as long as 95 minutes. 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. The standard deviation for earnings in the experiment was $6.54. Seventy-two subjects participated in the experiment, with twenty-four subjects in High Prior sessions, twenty-four subjects in Low Safe Action sessions, and twenty-four subjects in High Discount Factor sessions.

1.4 Results

This section takes two approaches towards testing the three hypotheses. The first approach is to use a subset of data where theory always predicts a switch to the safe

Table 1.3.: Summary statistics for the experiment. "HP Session" refers to the High Prior sessions, "LS" refers to the Low Safe Action sessions, and "HD Session" refers to the High Discount Factor sessions. The numbers without square brackets are the pooled averages from the Subset approach. The numbers inside of square brackets are the average of Product Limit estimated subject means from the Product Limit approach. "Difference" displays the difference between the summary statistics of the two treatments in each type of session.

Treatment	Prediction	HP Session	LS Session	HD 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	—	10.2 [15.9]	37.6 [31.9]	52.4 [70.1]

action. We will refer to this approach as the "Subset" approach. The second approach is to use the Product Limit estimator to correct for the censoring that occurs. We will refer to this approach as the "Product Limit" approach. We use results from both approaches as each approach has its own 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.[12] The Product Limit approach only uses subject means, but corrects for censored data. This section, including figures and graphs, will focus on the subset approach unless mentioned otherwise.

Section 1.4.1 addresses the first hypothesis. Section 1.4.2 addresses the second hypothesis. Section 1.4.3 addresses the third hypothesis.

1.4.1 Hypothesis 1

Hypothesis 1 states that subjects become willing to experiment longer when p_0 increases, δ increases, or $s\Delta$ decreases. Table 1.3 displays summary statistics from both the Subset approach and the Product Limit approach. The summary statistics

[12] Six percent of the data in the subset approach is censored.

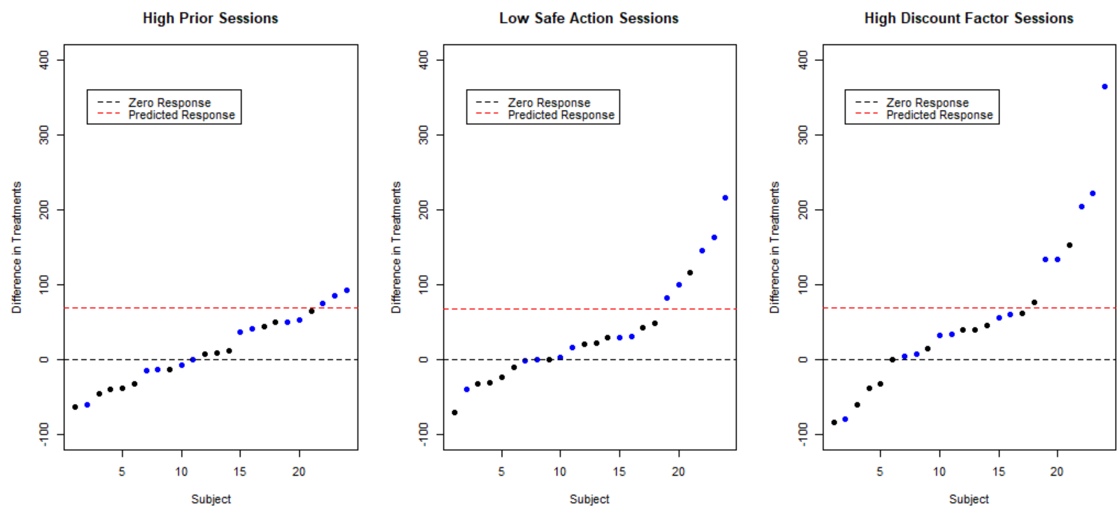


Figure 1.3.: Difference in subjects' average stopping time in the Baseline treatment and in the session's other treatment. The other treatment is the High Prior treatment in the first graph, the Low Safe Action treatment in the second graph, and the High Discount Factor in the third graph. The red line displays the predicted response to the treatment variable, while the black line displays no response to the treatment variable. Blue dots denote subjects who had the Baseline treatment first, while black dots denote subjects who had the Baseline Treatment second.

from both approaches suggest that subjects become willing to experiment longer when either p_0 or δ increases. Additionally, the summary statistics from both approaches suggest that subjects become willing to experiment longer when $s\Delta$ decreases.

Figure 1.3 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 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 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 in the High Discount Factor sessions have a higher mean stopping time in the High Discount Factor treatment than the Baseline treatment. While it appears that Hypothesis 1 holds, more formal analysis must be conducted.

Hypothesis 1 can be tested under both the Subset approach and the Product Limit approach. Random effect regressions, with subject level random effects, of the stopping time on the treatment can be run for the High Prior sessions, Low Safe Action sessions, and High Discount Factor sessions. The effect of increasing the prior is positive and significant at the five percent level (p-value=0.011). 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 backs up these results, except for the fact that the effect of increasing the prior is only significant at the ten percent level (p-value=.065).[13]

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

1.4.2 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. Table 1.3, however, 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 1.3 shows how subjects respond to a change in p_0, δ, and $s\Delta$. Figure 1.3 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 have a difference that is smaller than predicted in the High Prior sessions. Eighteen out of twenty-four subjects have a difference that is smaller than predicted in the Low Safe Action sessions. Seventeen out of twenty-four subjects have a difference that

[13]We ran bootstrapped regressions for each type of session. For example, for the change in the prior, we ran a bootstrapped regression of the difference in each High Prior session subject's Product Limit estimated mean stopping time for the Baseline treatment and the High Prior treatment. Bootstrapped regressions were run with 5000 bootstrap samples.

is less than predicted in the High Discount Factor sessions. While it appears that Hypothesis 2 does not hold, more formal analysis must still be conducted.

Hypothesis 2 can be tested under both the Subset approach and the Product Limit approach. As in Figure 1.3, a subject's response to the treatment variable can be calculated using the difference in a subject's mean stopping time in the Baseline treatment and their session's other treatment. For each variable, a bootstrapped regression, with 5000 bootstrap samples, can be run on the difference between each subject's response and their predicted response. 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=0.021). Subjects' response to an increase in the discount factor is insignificantly different than the predicted response at the ten percent level (p-value=0.609). The Product Limit approach is conducted similarly and backs up these results.

The High Prior Sessions can be used to test the assumption that cutoff beliefs are independent of the prior. We can test this assumption using the Subset approach.[14] In the High Prior sessions, the average cutoff belief 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 random effects regression with subject level random effects.

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.*

[14] Cutoff beliefs are not easily analyzable with survival analysis.

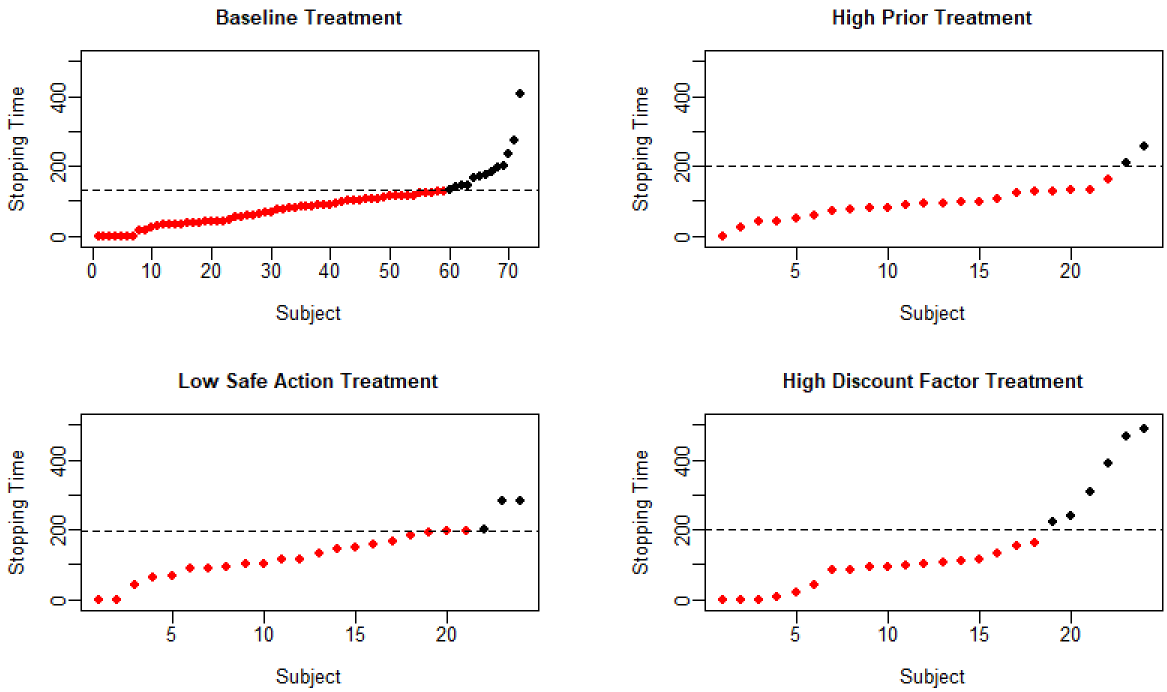


Figure 1.4.: Mean subject stopping times in each treatment. Red dots denote a mean stopping time lower than the prediction. Orange dots denote a mean stopping time equal to the prediction. Black dots denote a mean stopping time greater than the prediction.

1.4.3 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 section to see if subjects systematically under-experiment or over-experiment. Table 1.3 suggests that subjects are often willing to experiment for a shorter period of time than predicted by theory. Subjects in the High Prior Treatment appear to under-experiment by 100 ticks. Subjects in the Low Safe Action Treatment appear to under-experiment by at least 65 ticks. Subjects in the High Discount Factor treatment appear to under-experiment by at least 25 ticks.

Figure 1.4 compares, in each treatment, the stopping time of each subject to the predicted stopping time. Fifty-nine out of seventy-two subjects have an average stopping time below the prediction in the Baseline treatment. Twenty-two out of twenty-four subjects have an average stopping time below the prediction in the High Prior treatment. Twenty-one out of twenty-four subjects have an average stopping time below the prediction in the Low Safe Action Treatment. Eighteen out of twenty-four subjects have an average stopping time below the prediction in the High Discount Factor treatment, although some subjects have a much higher average stopping time than predicted. While it appears that subjects under-experiment, more formal analysis must be conducted.

Hypothesis tests, using both approaches, can be conducted to test whether subjects under-experiment. Random effects regressions, with subject level random effects, of the difference between each stopping time and its prediction can be run. 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=0.065). The Product Limit approach backs up 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=0.483).[15]

While subjects are often willing to experiment for a shorter period of time than predicted by theory, Table 1.3 suggests that stopping times are above the myopic prediction. A myopic subject is predicted to be willing to experiment for five ticks in the Baseline and High Discount Factor treatments and for seventy-four ticks in the High Prior and Low Safe Action treatments. Under both approaches, the mean stopping

[15]We ran bootstrapped regressions for each treatment. In each treatment, we regressed the difference of subjects' mean Product Limit Estimated stopping time and the predicted stopping time with 5000 bootstrap samples.

time in each treatment is above the myopic prediction. This result can be tested using both approaches. A random effects regression, with subject level random effects, of the difference between each stopping time and the myopic prediction can be 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 than myopia predicts at the five percent level in the High Prior Treatment (p-value=0.021). The Product Limit approach backs up 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=0.062).[16]

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.*

1.5 Estimating Behavioral Factors

The results section shows that subject behavior deviates from theory. In this section, we estimate a model in an attempt to better understand subject behavior. The goal of this section is to uncover behavioral factors that are consistent with subject behavior.

[16]We ran bootstrapped regressions for each treatment. In each treatment, we regressed the difference of subjects' mean Product Limit Estimated stopping time and the myopic stopping time with 5000 bootstrap samples.

1.5.1 Setup and Estimation

We focus on three possible deviations from theory: risk aversion, base rate neglect/conservatism, and probability mis-weighting. We include risk aversion in the model as risk is inherent in this problem and as other papers (Banks et al., 1997; Hudja, 2019) consider risk aversion in bandit experiments. We include probability mis-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 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 probability mis-weighting will be incorporated into belief updating. Base rate neglect/conservatism is modeled as a subject treating a tick of experimentation as if it is ψ ticks of experimentation.[17] Probability mis-weighting is modeled throughout this section through the Prelec-I function. Let $\tilde{p_0} = e^{-(-ln(p_0))^\alpha}$ be the mis-weighted value of p_0. Let $\tilde{\lambda}\Delta = e^{-(-ln(\lambda\Delta))^\alpha}$ be the mis-weighted value of $\lambda\Delta$. An agent's belief updating function, in the absence of a reward, can now be modeled as

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

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

Risk aversion and probability mis-weighting will be incorporated into the cutoff belief. Risk aversion is modeled using CRRA utility. Let $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$, with γ being the coefficient of risk aversion, be an agent's utility under risk aversion. Let $\tilde{\delta} = 1 - e^{-(-ln(1-\delta))^\alpha}$ be the mis-weighted value of the discount factor and once again let $\tilde{\lambda}\Delta = e^{-(-ln(\lambda\Delta))^\alpha}$ be the mis-weighted value of $\lambda\Delta$. A subject's cutoff belief can be found by value function iteration, where

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

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

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 a subject is willing to experiment can be found by finding where the belief updating function and cutoff belief intersect.

The parameters γ, ψ, and α can be 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

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

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

The model estimated in this section resembles a two-limit Tobit based on the aforementioned prediction. Let $pcensor_{i,t}$ 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 $0 < y_{i,t} < pcensor_{i,t}$ is equal to

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

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

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

The joint density for subject i can thus be written as

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

The log-likelihood can be written as $LogL = \sum_{i=1}^{n} lnL_i$. The model will use all of the data as this model can account for censoring. Maximum Likelihood estimation is used to estimate the parameters.

The maximized log-likelihood is equal to 8658.03.[18] 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 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).[19] These results suggest that risk aversion, conservatism, and probability mis-weighting influence subjects' experimentation decisions.

[18]The appendix, in section A.3, displays a model of experimentation that is based on the continuous time predictions. The results are similar to the following results.

[19]Standard errors back these results up except that α is only significant at the ten percent level. We utilize likelihood ratio tests since we calculate 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.

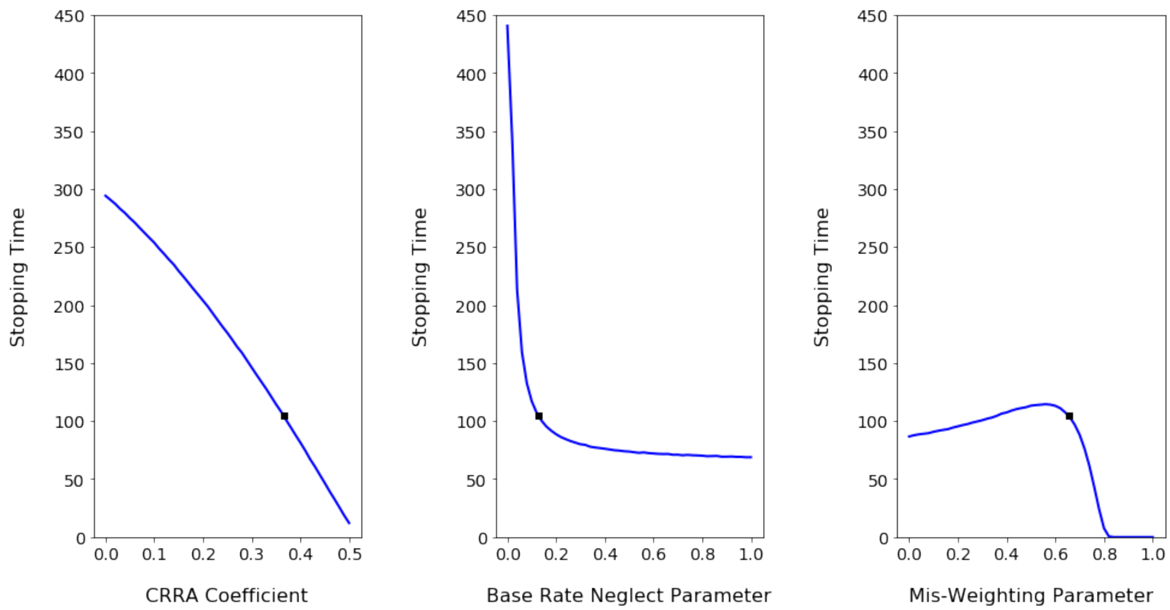


Figure 1.5.: Model predictions for the Baseline Treatment as each Behavioral factor is varied. These predictions are obtained through simulation and are for the subset approach. In each graph, one behavioral factor is varied, while the other two behavioral factors are held constant at the estimated levels. The black square denotes the prediction of the fully estimated model in section 1.5.1.

1.5.2 Effects

In this subsection, we explore the effects of each behavioral factor. Figure 1.5 displays the model predictions for the Baseline Treatment as each behavioral factor is varied. These effects are similar in other treatments. These predictions are mean predictions that are obtained by using the subset approach on one million period simulations.

The first graph 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 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

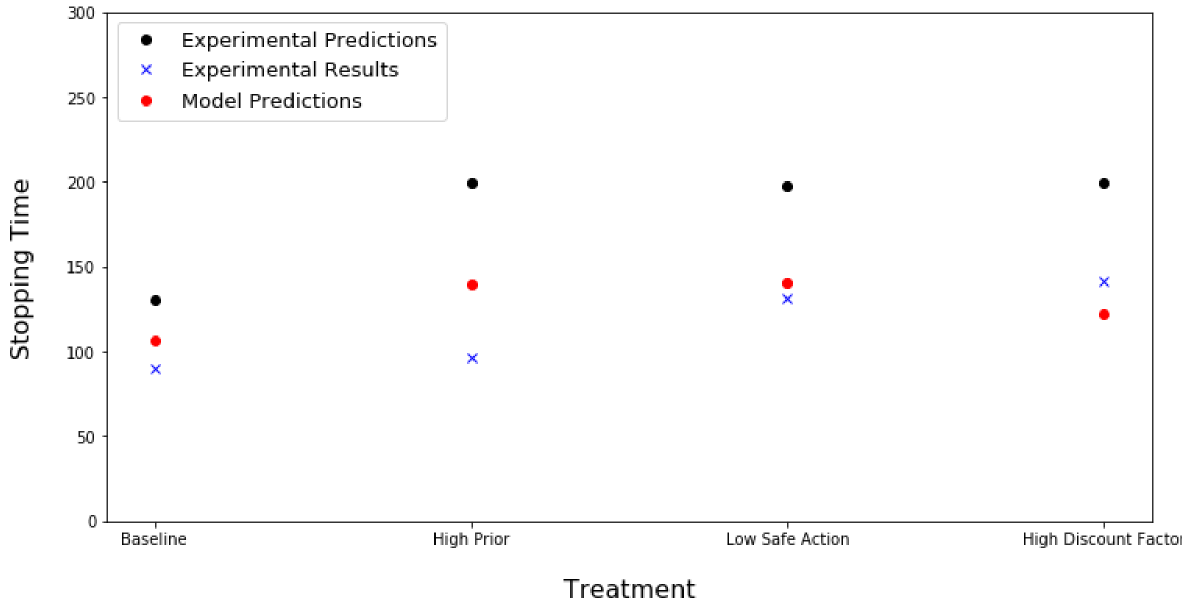


Figure 1.6.: Model predictions for each treatment using the subset approach. Model predictions are compared to the experimental predictions and to the average stopping times using the subset approach.

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 mis-weighting as α is varied (γ and ψ are held at their estimated values). The graph shows that turning the probability mis-weighting channel off ($\alpha = 1$) in the estimated model leads to less experimentation. This suggests that probability mis-weighting makes subjects willing to experiment longer.

1.5.3 Predictions

Figure 1.6 compares our model predictions to the experimental predictions. These predictions are mean predictions that are obtained by using the subset approach on one million period simulations. The 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 1.6 shows, the model predictions outperform the experimental predictions for each treatment in the experiment.

The model can also be compared to the experimental predictions of a previous study. Hudja (2019) reports an average stopping time for the last fifteen periods of its 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 same subset approach as in Hudja (2019), our model suggests an average stopping time of 128.3 ticks.

1.6 Power Analysis

The results section shows that subjects respond to changes in incentives. However, it is unclear whether this result holds in other bandit environments as an earlier study did not find a response to changes in incentives. Banks et al. (1997) analyze the role of incentives in a similar, yet different, bandit problem in the lab. In this section, we conduct a power analysis of both our paper and Banks et al. (1997) in an attempt to uncover whether subjects' response to incentives is unique to the exponential bandit environment.

The problem in Banks et al. (1997) has two differences from our bandit problem. First, they analyze a discrete time bandit problem. Second, the bad risky action in their problem can also pay out rewards. They analyze how subjects respond to a change in the discount factor and to a change in the arrival rates of the risky action. They find an insignificant effect of both changes at the ten percent level.

We conduct a power analysis of both our paper and their paper. Details on the power analyses can be found in section A.4 of the appendix. In the power analyses, we assume that subjects are noisy in that they deviate from the predicted cutoff belief.

We use two sources of noise for the power analysis.[20] Under both sets of noise, in our study, we find significant responses to changes in the discount factor, safe action, and prior belief 100 percent of the time at the five percent level. Under both sets of noise, in their paper, we find significant responses to a change in the discount factor less than 65 percent of the time and significant responses to a change in the risky action arrival rates less than 10 percent of the time at the 5 percent level. These percentages for Banks et al. (1997) fall below the recommended 80 percent for power (Moffatt, 2016).

1.7 Conclusion

This paper 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 probability mis-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 common under-experimentation found throughout this study suggests that individuals are sub-optimally experimenting. Lastly, our model suggests that subjects' behavior is consistent with risk aversion, conservatism and probability mis-weighting. Our model suggests that theorists should consider these behavioral factors when modeling experimentation environments.

[20] We use the High Prior Treatment in our experiment and the single-agent treatment in Hudja (2019) for measures of noise. We use this data as both datasets have a prior belief of 0.50, which is found in Banks et al. (1997). We want a similar spread in possible cutoff beliefs as in Banks et al. (1997).

There are many avenues for future research. First, other papers can focus on the effect of the arrival rate on experimentation. The arrival rate has a non-monotonic effect on the length of experimentation in that it can lead to an increase or decrease in experimentation depending on the parameters. It would be interesting to see if this non-monotonicity holds in the lab. Second, other papers can design treatments to isolate the effects of risk aversion, conservatism, and probability mis-weighting on experimentation. Our model is limited to what we modeled and thus may be picking up other factors instead of those three factors. Third, other experiments can place individuals into groups to see whether groups make better decisions in experimentation environments. Group decision making can be studied with and without communication. It would be interesting to observe discussion of experimentation strategies among group members.

2. VOTING FOR EXPERIMENTATION: A CONTINUOUS TIME ANALYSIS

This paper uses a laboratory experiment to analyze how a group of voters experiment with a new reform. The experiment implements the continuous time Strulovici (2010) collective experimentation model. I analyze a subset of data where groups and single decision makers should eventually prefer to stop experimentation and abandon the reform. I find three results that are consistent with the modeled experimentation incentives. In this subset of data, groups stop experimentation earlier than single decision makers, wait longer to stop experimentation as the number of revealed winners increases, and stop experimentation earlier than the utilitarian optimum predicts. However, I also find that both groups and single decision makers stop experimentation earlier than predicted. Additional treatments show that this result is unlikely to be explained by standard explanations such as incorrect belief updating or risk aversion.

2.1 Introduction

Societies often implement reforms that have uncertain and heterogeneous effects on the electorate. For example, tax reforms have effects that are hard to forecast and that vary for individuals in different states and tax brackets. Trade deals often have unintended consequences, which may determine the specific industries that gain or lose from a deal. Other examples of reforms that have both uncertain and heterogeneous effects are health care reforms, immigration reforms, and electoral reforms. When voters elect to implement one of these reforms, they are essentially electing to conduct a policy experiment where the reform must produce sufficient results or be repealed.

I analyze how a group of voters experiment with a new reform by testing the Strulovici (2010) model in a laboratory experiment. In the model, a group of voters continually decide by majority vote between a safe action that yields a constant identical payoff to everyone and a risky action that for each individual can be either good or bad. A bad type of risky action is dominated by the safe action and never returns a payoff. A good type of risky action dominates the safe action and occasionally pays out rewards. Voters are initially unsure of their type of risky action and learn about their type, which is independently distributed, through experimentation with the risky action. Voters who receive a reward from the risky action know their risky action is good; they are sure winners. Voters who have not yet received a reward from the risky action are unsure voters, who become more pessimistic about their type as experimentation goes on.

Each voter faces a collective experimentation problem because of the initial uncertainty of the risky action. Similar to an individual experimentation problem, an unsure voter faces an exploration versus exploitation trade-off: experimentation with the risky action comes at the cost of the safe action. However, an unsure voter must also consider future negative externalities that voters may impose upon one another. If the unsure voter has a bad type, her vote for the risky action can result in a majority of sure winners forming, trapping her into her (bad) risky action forever. If the unsure voter has a good type, her vote for the risky action can reveal that she is a winner, but this benefit may be short-lived as unsure voters may impose the safe action in the future.

Strulovici (2010) characterizes how unsure voters resolve this experimentation problem. He finds that the negative payoff externalites that can arise from voting weaken unsure voters' incentive to experiment with the risky action relative to a single decision maker, who only faces the exploration versus exploitation trade-off. The possibility of these externalities also reduces unsure voters' incentive to experiment relative to the utilitarian optimum. Finally, the possibility of these externalities increases unsure voters' incentive to experiment as the number of sure winners increases; the benefit of experimentation increases as a future sure winner is less likely to see the reform overturned.

I implement this model in a laboratory experiment to analyze how voters make collective decisions on a new reform in the face of these experimentation incentives. I employ a near continuous time approximation of the Strulovici (2010) model, which allows me to analyze an environment that has theoretical predictions close to the theoretical predictions of the continuous time environment. The experiment consists of two treatments. In the single-agent treatment, each subject faces an individual experimentation problem. In the majority-vote treatment, voters in a three person group face the collective experimentation problem.

The unique control of a laboratory experiment reveals how long groups, and single decision makers, are willing to try the reform. A laboratory experiment, unlike field data, allows for random assignment of groups to observations where a majority of sure winners can never form and for random assignment of single decision makers to observations where they never benefit from the reform. I analyze a subset of these observations where theory predicts that groups, and single decision makers, will stop trying the reform.

The results show that collective decisions on the reform are generally consistent with the modeled experimentation incentives. In particular, the results support three predictions for this subset of data: (i) groups stop trying the reform earlier than single decision makers, (ii) groups with zero winners stop trying the reform earlier than groups with one winner, and (iii) groups stop trying the reform earlier than the utilitarian optimum predicts. The first result is consistent with the weaker experimentation incentives under majority voting. The second result is consistent with the increasing experimentation incentives when a winner is observed. The third result is consistent with the socially inefficient experimentation incentives under majority voting. However, both groups and single decision makers stop trying the reform earlier than theory predicts.

This unexpected result is a product of subjects under-experimenting with the risky action. Two additional treatments are designed to rule out candidate explanations for under-experimentation. The first is similar to the single-agent treatment, but now I give unsure single decision makers the updated probability that their risky action is good. This treatment analyzes how belief updating affects experimentation. The second additional treatment is similar to the first additional treatment, but has a fixed period length. This treatment analyzes how discounting biases affect experimentation. These additional treatments, as well as elicited risk preferences, show that risk aversion, belief updating, and discounting biases are unlikely to be explanations for under-experimentation.

This paper contributes to three strands of literature. First is the literature on experimentation with multiple agents, which has mostly focused on strategic experimentation in which agents can free-ride on other agents' costly experimentation (Bolton and Harris, 1999). Keller et al. (2005) analyze strategic experimentation through an exponential bandit framework. Hoelzemann and Klein (2018) analyze strategic experimentation in the lab and find evidence of free-riding. Recent research on multi-agent experimentation has considered the case of collective experimentation in which an agent's experimentation may depend on other group members. Strulovici (2010) analyzes collective experimentation using exponential bandits. Khromenkova (2015) analyzes collective experimentation, but allows for voters of any type of risky action to learn its quality. Freer et al. (2018) build and experimentally test a collective experimentation model. They find that majority-voting performs better than their optimal rule in a three period signaling model. My paper is the first to experimentally analyze how collective experimentation differs from individual experimentation. I find that groups stop experimentation earlier than single-agents, as predicted by theory.

This paper also contributes to the literature on testing political economy models of reform. Previous studies have tested the Fernandez and Rodrick (1991) model, which shows that individual-specific uncertainty can prevent the adoption of efficiencyenhancing reforms. Cason and Mui (2003, 2005) report experimental results that supports the model. Paetzel et al. (2014) find that social preferences can mitigate this issue. My paper is the first to directly test the Strulovici (2010) model, which has new strategic issues that arise from voters learning over time and the initial uncertainty over the reform's efficiency.

Finally, this paper contributes to the literature on bandit experiments, which has mostly analyzed situations where the high payoff probability of an unknown risky action is randomly drawn from a continuous distribution. Meyer and Shi (1995) find under-experimentation in this type of problem and suggest shortened planning horizons as an explanation. Anderson (2001, 2012) finds under-experimentation in this type of problem and suggests ambiguity aversion as an explanation while providing evidence against hyperbolic discounting and risk aversion. In the current study, the high payoff probability of an unknown risky action can take on only one of two values. Banks et al. (1997) analyze a variant of this problem and find that strategies are unaffected by risk aversion. However, they do not analyze whether subjects under-experiment. I find that subjects under-experiment in this type of problem and that this under-experimentation is unlikely to be explained by incorrect belief updating, discounting biases, and risk aversion.

2.2 Theory

The experimental model is based on Strulovici (2010). This section closely follows section 2 of that article. The reader seeking greater detail (or proofs) than this section provides should consult Strulovici (2010).

Time (t) is continuous and payoffs of all individuals are discounted at a rate r. There is an odd number ($N \geq 1$) of individuals, with risk neutral preferences, who continually decide by majority vote between two actions. The two actions are the safe action, which yields a flow s per unit of time to all individuals and the risky action, which can be, for each player, either good or bad. The types of the risky action are independently distributed across the group.

If the risky action is bad for an individual, it always pays her nothing. If the risky action is good for an individual, it pays her rewards at random times, which depend on a Poisson process with parameter λ. The arrival of rewards is independent across individuals, and the magnitude of a reward is h. Thus, if the risky action is good for an individual, her expected payoff per unit of time from the risky action is $g = \lambda h$, which is greater than s.

The risky action has an initial probability p_0 of being good for each individual. This is common knowledge. All payoffs are publicly observed, so that everyone shares the same belief about any individual's type. Thus, the arrival of the first reward to an individual makes her publicly a sure winner. At any time, the group is divided into k sure winners, who always vote for the risky action, and $N - k$ unsure voters, who have the same probability p of having a good risky action. Unsure voters update their probability of having a good risky action according to Bayes' rule and become more pessimistic as experimentation goes on. An unsure voter learns about her type only from her own payoffs because types are independent.

When the group consists of one agent, this environment reduces to the environment of a single decision maker. The optimal strategy is for the risky action to be chosen if and only if the current belief (p) is greater than or equal to a cutoff belief p_A. The cutoff belief p_A is equal to

$$p_A = \frac{rs}{rg + \lambda(g - s)} \tag{2.1}$$

and solves the indifference equation $pg + \lambda p(\frac{g}{r} - \frac{s}{r}) = s$. The left-hand side of the indifference equation corresponds to the risky action, while the right-hand side corresponds to the safe action. The effect of the risky action on an unsure single decision maker can be decomposed into two elements: (i) the expected payoff pg, and (ii) the jump in the value function (from $\frac{s}{r}$ to $\frac{g}{r}$ as she is currently 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 myopic strategy is for the risky action to be chosen if and only if the current belief (p) is greater than or equal to a cutoff belief $p_M = \frac{s}{g}$. The myopic cutoff (p_M) is the probability below which the risky action yields a lower expected flow payoff than the safe action. The myopic strategy is suboptimal because a myopic decision maker does not value the information generated by experimentation. A myopic voter experiments less than a non-myopic voter (as $p_A < p_M$).

When a group decides by majority vote, collective decisions are determined by non-increasing cutoffs such that the risky action is played at time t if and only if $p_t > p(k_t)$, where k_t is the number of sure winners at that time. Starting with an appropriate initial belief p_0, the risky action is elected until unsure voters reach the threshold $p(0)$, at which point experimentation stops if no winner has yet been observed. If at least one winner has been observed experimentation continues until unsure voters reach another threshold $p(1) < p(0)$, and so forth. Experimentation means choosing the risky action when one's type is unknown. Only unsure voters experiment.

When there is a majority of unsure voters, decisions are dictated by their common interest unless and until they lose the majority. Cutoffs are therefore determined by unsure voters' preferences when they have the majority. These preferences are determined by the following Hamilton-Jacobi-Bellman (HJB) equation,

$$\begin{aligned} ru(k,p) =& max\{pg + \lambda p[w(k+1,p) - u(k,p)] \\ &+ \lambda p(N-k-1)[u(k+1,p) - u(k,p)] \\ &- \lambda p(1-p)\frac{du}{dp}(k,p), s\}, \end{aligned} \tag{2.2}$$

where $u(k,p)$ and $w(k,p)$ are unsure voters' and sure winners' respective value function when the state is (k,p). The first part of the maximand corresponds to the risky action, and the second corresponds to the safe action. The effect of the risky action on an unsure voter can be decomposed into four elements: (i) the expected payoff rate pg, (ii) the jump of the value function if a reward is received, which occurs at rate λ with probability p, (iii) the jump of the value function if another unsure voter receives a reward, which occurs at rate λ with probability $p(N-k-1)$ and (iv) the effect of Bayesian updating on the value function when no reward is observed. If the safe action is chosen, the payoff rate is s.

Since unsure voters have identical value functions, they unanimously decide to stop experimentation if p becomes too low, which occurs when the risky action part of (2) equals s. This determines the experimentation cutoffs in Theorem 1 (Strulovici, 2010).

Theorem 1 (Strulovici, 2010): *There exists a unique Markov equilibrium in undominated strategies. This equilibrium is characterized by cutoffs p(k) for* $k \in \{0, ..., N\}$*, such that R is chosen in state* (k, p) *if and only if* $p > p(k)$*. Furthermore, for all* $k \in \{0, ..., (N-1)/2\}$*,* $p_M > p(k) > p_A$*,* $p(k)$ *is decreasing in k for* $k \leq (N-1)/2$*, and* $p(k) = 0$ *for all* $k > (N-1)/2$*.*

Theorem 1 characterizes the cutoffs that determine behavior in equilibrium. The first property of the cutoffs is that unsure voters are not myopic ($p(k) < p_M$ for all $k \in \{0, ..., (N-1)/2\}$). Intuitively, all unsure voters prefer the risky action when the expected flow payoff is greater than the safe action. However, unsure voters consider the future rewards that come with being a winner. This property implies that groups are willing to implement the risky action past the myopic cutoff when unsure voters have the majority.

The second property of the cutoffs is that unsure voters are less willing to experiment than a single decision maker ($p(k) > p_A$ for all $k \leq (N-1)/2$). An unsure voter has less control over the action chosen as part of a group. In a group, a winner may have the safe action implemented. In a group, a loser's (bad type) experimentation can lead to a majority of sure winners forming. This lack of control decreases the incentive to experiment relative to a single decision maker. This property implies that groups, when unsure voters have the majority, are less willing to implement the risky action than an unsure single decision maker.

There are two more properties of theorem 1 worth discussing for the experiment: (i) when unsure voters have the majority, unsure voters become more willing to experiment as the number of sure winners increases ($p(k) \downarrow$ in k, for $k \leq (N-1)/2$) and (ii) once a majority of sure winners form, the risky action is implemented forever ($p(k) = 0$ for all $k > (N-1)/2$). The benefit of experimentation, and becoming a winner, increases as the number of winners increases. As the number of winners increases, the risky action is less likely to be overturned. This property implies that a group with one winner ($N \geq 3$) is willing to implement the risky action for longer than a group with zero winners. The last property implies that the risky action is implemented indefinitely in the experiment once a majority of sure winners forms.

2.2.1 Discrete Implementation

It is not possible to implement continuous time bandit problems in the laboratory, so I implement a close approximation. The approximation consists of two main parts. The first part is dividing time into a number of ticks, each of length Δ seconds. A group (or single-agent) can only employ one action in each tick. The second part is replacing the Poisson process of a good risky action with an appropriate Bernoulli process. Under a Bernoulli process, each subject can only receive one reward per tick.

The parameters of the approximation are based on the continuous time parameters. In the approximation, the probability of a subject receiving a reward from a good risky action, in a tick, is $\pi = \lambda\Delta$. This probability stems from the continuous time probability of obtaining at least one reward in Δ seconds $(1 - e^{-\lambda\Delta})$. The discount rate, in the approximation, for a tick is $\delta = 1 - r\Delta$, which is based on the continuous time discount rate for Δ seconds $(e^{-r\Delta})$.

The myopic benefit of the good risky action relative to the safe action must be the same in the approximation as it is in the continuous time model. In continuous time, a good risky action has an expected flow value of λh, while the safe action has a flow value of s. In the approximation, the expected payoff of a good risky action, in a tick, is $\pi h = \lambda\Delta h$. Thus, the payoff of the safe action (σ), in a tick, must be set to Δs.

Lastly, this model features an infinite horizon, which can not be implemented in the laboratory. However, the infinite horizon problem can be transformed into an equivalent indefinite horizon problem through the use of a random stopping rule. In the experiment, each tick has a probability of $q = r\Delta$ of ending the period. This probability induces a discount rate of $\delta = 1 - r\Delta$, which is the discount rate for the approximation.

2.3 Experimental Design and Testable Hypotheses

The treatments are designed to analyze group and single-agent experimentation. The primary goal of the experiment is to compare group experimentation to single-agent experimentation. The experiment has two secondary goals. The first secondary goal is to determine how the arrival of a sure winner affects group experimentation. The second secondary goal is to compare group experimentation to equilibrium and utilitarian predictions.

2.3.1 Treatments and Parameters

There are two treatments in the experiment. The single-agent treatment analyzes the single-agent setting (one subject in a group). The majority-vote treatment analyzes the majority-vote setting (three subjects in a group). The experiment uses a within-subjects design, where subjects start the experiment in one of the two treatments. Eighty-four subjects experience both treatments. Each treatment lasts for twenty-five periods.

Table 2.1.: Predicted cutoffs, in ticks, for the experiment and the continuous time model. In the majority-vote treatment, a switch to the safe action is predicted to occur if there are zero winners after tick 110 or one winner after tick 153. In the single-agent treatment, a switch to the safe action is predicted to occur if the single-agent is not a winner after tick 187. The 'Zero' column displays the cutoff for when there are zero winners, while the 'One' column displays the cutoff for when there is one winner.

	Majority-Vote		Single-Agent
	Zero	One	Single
Discrete (Ticks)	110.00	153.00	187.00
Continuous (Ticks)	110.77	153.55	187.18

The parameters for the experiment are chosen to simplify the environment. The group size is set at three and the prior probability of a good state, p_0, is set at 0.5. The tick length, Δ, is 200 milliseconds, which is consistent with other experiments in the continuous time literature (see Oprea et al., 2009; Anderson et al., 2010). The remaining parameters are $\pi = 0.01$, $\delta = 0.997$, $\sigma = \$0.01$ and $h = \$2.50$. The discount rate δ gives each period an expected period length of 333.33 ticks, which is greater than any of the predicted stopping times. These parameters approximate a continuous bandit that has parameters $\lambda = 0.05$, $r = 0.015$, $g = \$0.125$, and $s = \$0.05$.

The theoretical predictions for the experiment are shown in Table 2.1. Table 2.1 displays predictions for both the discrete time approximation and the continuous time model.[1] In the majority-vote treatment, groups are predicted to choose the risky action through tick 110. If there are zero winners after tick 110, the group switches to the safe action; otherwise, the group continues choosing the risky action through tick 153. If there is one winner after tick 153, the group switches to the safe action; otherwise, the group permanently chooses the risky action. In the single-agent treatment, subjects are predicted to choose the risky action through tick 187. If the

[1]The continuous time predictions are derived in section B.1 of the appendix and are based on equation (2) and other equations found in the proof of Theorem 1 in Strulovici (2010). The discrete time predictions are derived by value function iteration (details are in section B.2.1 of the appendix).

subject is not a winner after tick 187, she switches to the safe action; otherwise, she permanently chooses the risky action.

2.3.2 Beginning of a Treatment

Before the start of the experiment, instructions for the first treatment were passed out. Half of the sessions started off with twenty-five periods of the single-agent treatment, while the other half started off with twenty-five periods of the majority-vote treatment. After the instructions were read, subjects watched a video on the experimental interface, and then answered five comprehension questions that were each worth $1.00.

After the first treatment ended, subjects received instructions on the second treatment. After the instructions were read, subjects started the second treatment.

2.3.3 Single-Agent Treatment

The environment in the single-agent treatment is described through an analogy of balls being drawn from a bag. Subjects can either draw a ball (risky action) or not draw a ball (safe action) in a given tick. There are two bags: (i) a uniform bag, which consists of one-hundred yellow balls and (ii) a mixed bag, which consists of one red ball and ninety-nine yellow balls. In this analogy, the mixed bag is the 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, with equal probability, is randomly drawn for each subject. The bag stays the same throughout the period.

At the start of each period, subjects have as much time as they would like to choose an initial action. Once every subject chooses an initial action, a five second timer is displayed. When the timer hits zero, the first tick occurs. If a subject initially chooses to draw a ball, she continually draws until, as an unsure single 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

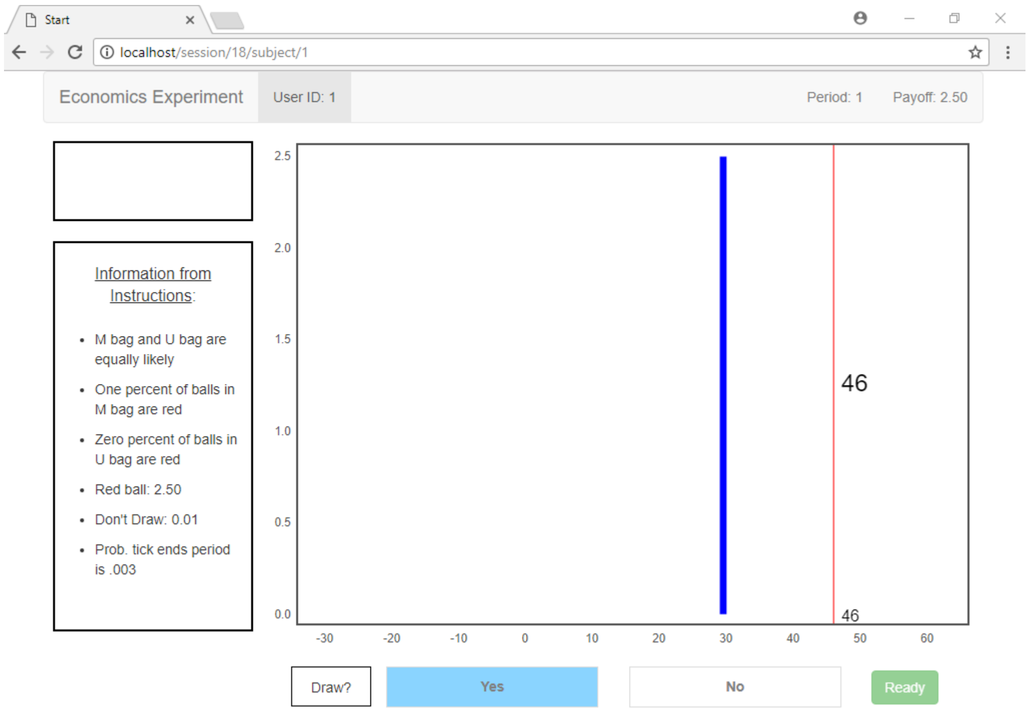


Figure 2.1.: Example of a subject's screen in the single-agent treatment. In this example, the (fictional) subject has drawn 46 balls in 46 ticks and obtained a reward.

of the period.[2] A subject must draw for the rest of the period once she receives a reward and becomes a winner.[3] Ticks continue until the random termination of the period.

Subjects receive feedback throughout the period through a graph displayed on their screen. A red line, that doesn't move, is drawn 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 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 safe action

[2]This forces unsure single-agents to decide between drawing in the next tick and never drawing for the rest of the period. This is the decision subjects make in theory. The advantage of this approach is that subjects cannot switch back to the risky action in periods that last longer than expected.
[3]Forcing a subject to automatically choose the risky action once she becomes a winner prevents her from later accidentally switching to the safe action and is consistent with theory.

previously occured, the subject sees a blue line of height σ drawn over the previous tick. If the risky action previously occured, the subject either sees no line (no reward occurred) or a blue line of height h (reward occurred) drawn over the previous tick. Figure 2.1 displays an example of the screen from the single-agent treatment.

At the end of the period, subjects receive summary information on the number of balls drawn, the number of red balls drawn, the payoff, and the period length (in ticks).

2.3.4 Majority-Vote Treatment

The environment of the majority-vote treatment differs slightly from the single-agent environment. The analogy of the single-agent treatment is expanded as it is now explained to subjects that they are a part of a fixed three-person group. Through an analogy of a coin being flipped three times, subjects are informed that the type of bag each group member is drawing from is independently determined at the start of each period.

At the start of the treatment, subjects are randomly matched into fixed groups of three. At the start of each period, subjects are given as much time as they would like to make an initial vote. Once everyone in the group has made an initial vote, a five second timer is displayed. When the timer hits zero, the first tick occurs. If the majority of initial votes are to draw a ball, each group member continually draws a ball from her own bag unless and until a majority composed of unsure voters votes to stop. Starting from the initial vote, whenever a majority composed of unsure voters votes to not draw, each group member cannot draw for the rest of the period.[4] A subject must vote 'yes' for the rest of the period once she receives a reward and becomes a winner. Thus, when a majority of voters receive rewards, each group

[4]This forces the group to stop drawing once a majority of unsure voters votes to not draw. Subjects are however allowed to change their vote as long as the group has never implemented the safe action. This approach was preferred over forcing a subject to always vote for the safe action once she initially votes for the safe action. The first voter to vote for the safe action may want to vote for the risky action later if experimentation continues and a sure winner is later revealed.

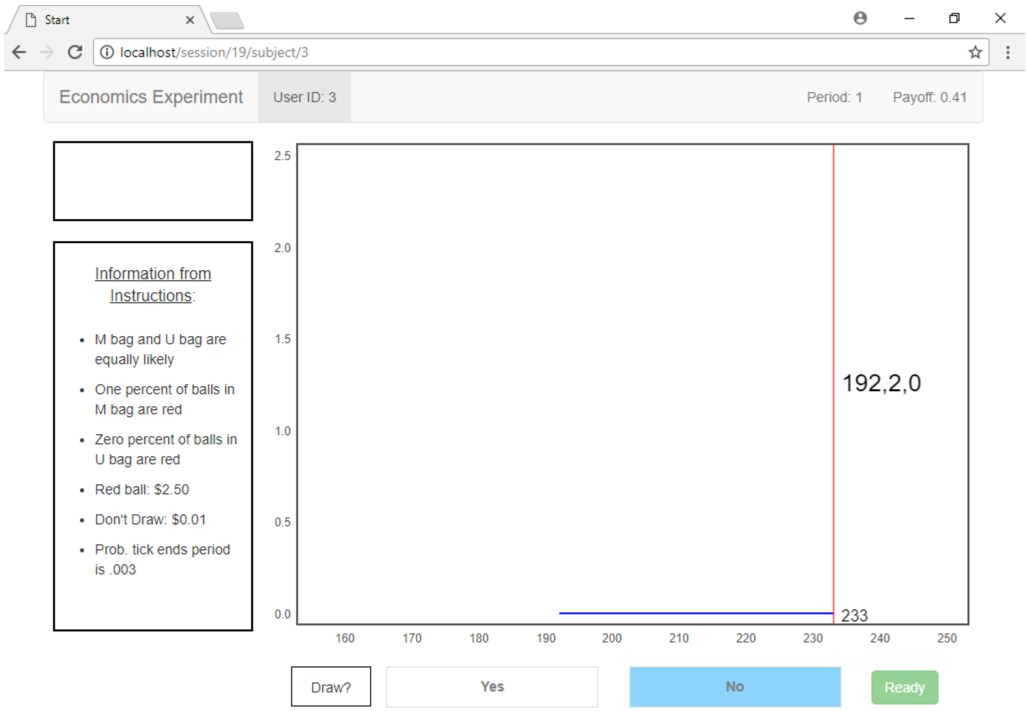


Figure 2.2.: Example of a subject's screen in the majority-vote treatment. In this example, the (fictional) subject has drawn 192 balls in 233 ticks and the group decided to stop drawing after the 192^{nd} tick. One subject in this group has obtained two red balls.

member draws a ball from their bag for the rest of the period. Ticks continue until the random termination of the period.

In each period, each subject receives feedback through a graph displayed on their screen. Figure 2.2 displays an example of the screen in the treatment. The graph is similar to the single-agent graph, but now there are three numbers to the right of the middle of the red line: (i) the number of balls drawn so far, (ii) the number of red balls drawn by subject B in the group, and (iii) the number of red balls drawn by subject C in the group. The number of red balls drawn by the other group members is provided to allow subjects to determine the number of sure winners in the group.

At the end of the period, subjects are informed of the number of balls they drew, the number of red balls they drew, their payoff and the period length. Subjects are also given information on the number of red balls each group member had drawn.

2.3.5 Testable Hypotheses

While theory focuses on how long voters are willing to experiment, testing these predictions is complicated by voters' heterogeneity. When voters are heterogeneous, the second voter to vote for the safe action stops experimentation for all unsure voters. For example, consider two group members who always immediately vote for the safe action. These voters prevent experimentation and prevent observation of how long the third voter is willing to experiment when there are zero winners.

Instead of analyzing how long unsure voters are willing to experiment, this paper analyzes how long groups are willing to choose the risky action when unsure voters have the majority. In other words, this paper analyzes how long groups (and single-agents) are willing to try the risky action. Trying the risky action refers to a group choosing the risky action when it does not yet have permanent support. Essentially, groups are trying the risky action until a majority of sure winners form, making the risky action irreversible, or until the group switches away from the risky action. Single-agents try the risky action when they experiment as they are the sole voter in the group.

Groups and single-agents are predicted to not reveal how long they are willing to try the risky action under two scenarios. The first scenario is when the risky action becomes permanent as there is never a switch to the safe action. This issue theoretically arises in the majority-vote treatment if a majority of sure winners is predicted to form and in the single-agent treatment if the single-agent is predicted to become a winner. The second scenario is that the period may end before there is a switch to the safe action. This issue theoretically arises if the period ends before a predicted switch to the safe action.

These two issues are mitigated by analyzing a subset of data where groups and single-agents should always reveal how long they are willing to try the reform. I refer to this dataset as the clean dataset, which consists of observations in the last fifteen periods of each treatment where two conditions are met. The first condition is that either a majority of group members have bad states or a single-agent has a bad state.[5] The second condition is that the period lasts at least 200 ticks.[6] The clean dataset provides an unbiased estimate of how long a group or single-agent is willing to try the reform. I conduct hypothesis tests on the clean dataset by analyzing when groups and single-agents switch to the safe action.

Theory predicts that a group is willing to try the risky action for 110 ticks when there are zero winners and for 153 ticks when there is one winner. Theory predicts that a single decision maker is willing to try the risky action for 187 ticks when she is unsure of the risky action. In the clean dataset, groups thus switch to the safe action earlier. This leads me to hypothesis 1.

HYPOTHESIS 1: *Groups switch to the safe action earlier than single decision makers.*

Theory predicts that a group with zero winners is willing to try the risky action for 110 ticks. Theory predicts that a group with one winner is willing to try the risky action for 153 ticks. In the clean dataset, a group thus switches to the safe action later upon observing a winner. This leads me to hypothesis 2.

[5] I remove the data where a majority of voters have good states rather than use survival analysis on those observations because survival analysis methods will be biased on those observations. Survival analysis requires that groups who are censored at a time t have the same prospect of survival as those who continue to be followed. This assumption does not theoretically hold when groups are censored when a majority of sure winners forms. The censored group almost always had one winner before being censored. Thus, if the censored group was not censored, it would have been theoretically willing to try the risky action for longer than the average groups that were not censored.

[6] In the appendix, I analyze a larger subset of data where this second condition is relaxed. I find very similar results. I relax this second condition by using survival analysis to analyze all the observations in the last fifteen periods of each treatment where either a majority of voters have bad states or single-agents have good states.

HYPOTHESIS 2: *Groups with one winner switch to the safe action later than groups with zero winners.*

Strulovici (2010) solves for the optimal experimentation cutoffs of a utilitarian social planner who wants to maximize the average expected payoff of the three voters. The utilitarian cutoffs, in ticks, are non-decreasing in the number of winners and occur later, for each number of winners ($k < \frac{N+1}{2}$), than the equilibrium cutoffs. There are two reasons why unsure voters are less willing to experiment than the utilitarian optimum requires. The first reason is that a utilitarian social planner does not face the control-sharing effects that unsure voters do. The second reason is that a utilitarian social planner takes into account sure winners' utility. Unsure voters do not consider sure winners' utility and thus sometimes impose the safe action when the risky action is more efficient.

Theory predicts that a utilitarian social planner that observes zero winners is willing to try the risky action for 132 ticks. Theory predicts that a utilitarian social planner that observes one winner is willing to try the risky action for 358 ticks.[7] In the clean dataset, groups thus switch to the safe action earlier than the utilitarian optimum predicts. This leads me to Hypothesis 3.

HYPOTHESIS 3: *Groups switch to the safe action earlier than the utilitarian optimum predicts.*

2.3.6 Procedures

Sessions were conducted at Purdue University in February 2018. Sessions lasted about 80 minutes. Subjects received payment for their answers to the comprehension

[7] These cutoffs are also solved by value function iteration (details are in section B.2.2 of the appendix).

Table 2.2.: Mean stopping times, in ticks, for the clean dataset. Overall equilibrium prediction for the majority-vote observations is an averaged prediction.

Stopping Time	Majority-Vote			Single-Agent		
	Equil.		Actual	Equil.		Actual
Overall	133.5	$\overset{**}{>}$	113.8	187.0	$\overset{***}{>}$	135.3
Winner Impossible	110.0	$<$	113.5	——		——
Winner Predicted	153.0	$\overset{***}{>}$	117.2	——		——
Winner Not Predicted	110.0	$>$	105.0	——		——

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

questions and received payment for five random periods in the single-agent treatment and five random periods in the majority-vote treatment.

2.4 Results

Section 2.4.1 focuses on the summary statistics of the clean dataset. The remaining three subsections report test of the three hypotheses.

2.4.1 General Results

Table 2.2 presents the mean stopping times (in ticks) derived from the clean dataset.[8] The mean stopping time in the majority-vote observations of the clean dataset is 113.8 ticks. This average is significantly less than the prediction of 133.5 ticks (p-value=0.026). Hypothesis tests in this subsection are conducted using bootstrapped regressions with 5000 samples clustered at the group level. The mean stopping time in the single-agent observations of the clean dataset is 135.3 ticks, which is also significantly less than predicted (p-value <0.001).[9]

[8]The clean dataset analyzes the last fifteen periods of each treatment. The first ten periods of each treatment can be found in the appendix.

[9]Subjects in each treatment appear to not be myopic as each of these cutoffs are significantly greater than the myopic cutoff of 41 ticks at the 1 percent level.

Table 2.2 also divides the majority-vote observations into three subsets based on the possibility of observing a winner.[10] The first subset, when a winner is impossible, is where all three group members have bad states. When a winner is impossible, the mean stopping time is 113.5 ticks, which is not significantly different from the predicted 110 ticks at the 10 percent level (p-value=0.650). The second subset, when a winner is predicted, is where one group member has a good state and the first reward arrives within 110 ticks of the risky action. When a winner is predicted, the mean stopping time is 117.2 ticks, which is significantly less than the predicted 153 ticks at the 1 percent level (p-value=0.008). Finally, the third subset, when a winner is not predicted but possible, is where one group member has a good state and the first reward arrives after 110 ticks of the risky action. Although a winner is not predicted, as groups should stop after 110 ticks, it is possible for a winner to be observed if groups implement the risky action longer than predicted. In this case, the mean stopping time is 105 ticks, which is not significantly different from the predicted 110 ticks (p-value=0.701). The early stopping time in the majority-vote observations is driven by observations where a winner is predicted.

Result 1: *Both groups and single-agents switch to the safe action earlier than predicted by theory.*

The stopping times are similar in each of the subsets of the majority-vote observations, which suggests that there may not be an effect of observing a winner. However, groups do not always observe a winner when a winner is predicted. In the majority-vote observations, the average stopping time is 141.7 ticks when a winner is predicted and observed. This average stopping time suggests that there is an effect of observing a winner. However, there may be self-selection that occurs as groups that

[10]The period lengths, states, and arrival of rewards, were all pre-generated in the experiment. This pre-generation allows for the majority-vote observations to be broken down into subsets. One set of period lengths were used throughout the experiment. States and the arrival of lump-sums were pre-generated for each subject in each period.

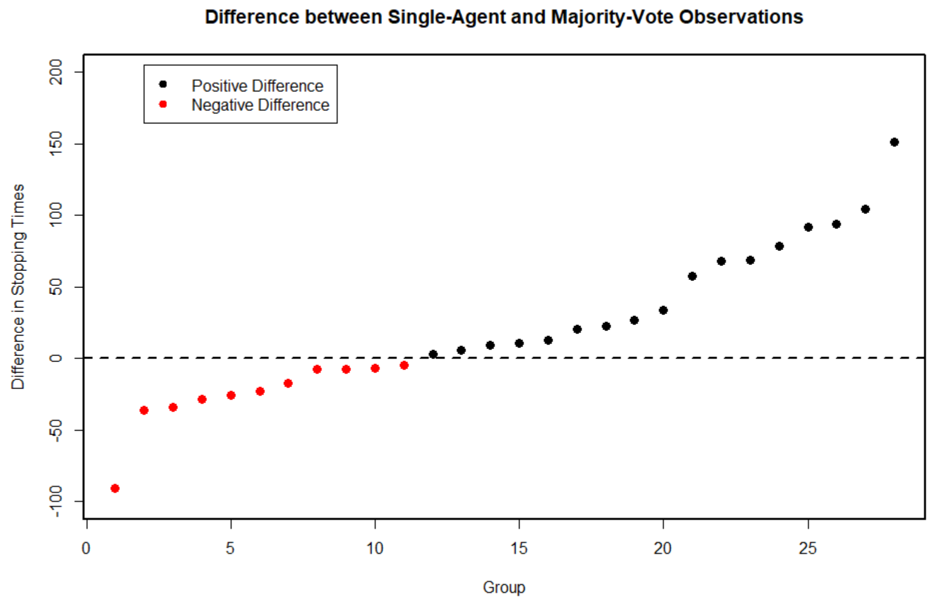


Figure 2.3.: Difference between the average group stopping time in the single-agent observations and majority-vote observations. A red dot indicates a longer mean stopping time in the majority-vote observations, while a black dot indicates a longer mean stopping time in the single-agent observations.

prefer to stop later are theoretically more likely to observe a winner. Section 2.4.3 addresses this issue.

2.4.2 Difference Between Treatments

Hypothesis 1 states that, in the clean dataset, groups switch to the safe action earlier than single-agents. Table 2.2 shows that the mean stopping time is greater in the single-agent observations than in the majority-vote observations (135.3 versus 113.8). Figure 2.3 gives a deeper analysis of the difference between the stopping times in the single-agent observations and the majority-vote observations. Figure 2.3 displays the difference between each group's average stopping time in the single-agent observations and the group's average stopping time in the majority-vote observations.[11] Most groups stop later in the single-agent observations.

These results can be formalized through a random effects regression with group level random effects. The regression includes a single indicator variable for whether the observation was in the single-agent treatment. The coefficient on this variable is positive (18.1) and significant at the five percent level (p-value=0.039).

Result 2: *Groups switch to the safe action earlier than single-agents (evidence supporting Hypothesis 1).*

2.4.3 Observing a Winner

The second hypothesis states that, in the clean dataset, groups with one winner switch to the safe action later than groups with zero winners. Table 2.2 suggests that groups stop later when they observe a winner. In the majority-vote observations, the mean stopping time is 117.2 ticks when theory predicts a winner versus 109.8 ticks when theory predicts a winner can not or will not be observed. This difference, in the majority-vote observations, increases when comparing groups that observe a winner to groups that can not (141.7 vs 113.5). However, this may be self-selection as groups that stop later may be more likely to observe a winner. While these results are consistent with the second hypothesis, an approach that controls for self-selection is needed.

The Cox regression can be used to answer the second hypothesis and control for self-selection. The Cox regression (Cox, 1972) estimates the impact of a covariate on the time to an event occurring. In this analysis, an event occurs when a switch to the safe action occurs. The Cox regression allows for the estimation of the impact of

[11]The mean used for the majority-vote observations is each group's mean stopping time in the majority-vote observations. The mean used for the single-agent observations is the mean stopping time for the pooled data of the three group members in the single-agent observations.

Table 2.3.: Results from a Cox regression that estimates the effect of observing a winner on a group's stopping time.

	Coeff.	Hazard Ratio	Z
Winner	-0.942	0.390	-4.96***
N		284	

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

time-dependent variables and can thus estimate the impact of observing a winner on the stopping time.

Table 2.3 displays the results of the Cox regression. The variable *Winner* is a time-dependent variable, which takes a value for each tick in a period. The variable *Winner* is equal to zero if no winners have been observed through the current tick and is equal to 1 if one winner has already been observed.[12] The Cox regression is clustered at the group level. The coefficient on *Winner* is -0.942 and the hazard ratio is 0.390 from the Cox regression. This result implies that groups without a winner are 2.57 times more likely, per unit of time, to switch to the safe action than groups with a winner. This result is significant at the 1 percent level. The Cox regression backs up the second hypothesis.

Result 3: *Groups with one winner switch to the safe action later than groups with zero winners (evidence supporting Hypothesis 2).*

Table 2.4.: Mean stopping times, in ticks, for each treatment. Stopping times use the last fifteen periods of each treatment. Overall utilitarian prediction for the majority-vote treatment averages the prediction for each observation in the clean dataset.

Stopping Time	Majority-Vote			Single-Agent		
	Util.		Actual	Util.		Actual
Overall	261.4	$\overset{***}{>}$	113.8	187.0	$\overset{***}{>}$	135.3
Winner Impossible	132.0	$\overset{**}{>}$	113.5	——		——
Winner Predicted	358.0	$\overset{***}{>}$	118.3	——		——
Winner Not Predicted	132.0	$\overset{***}{>}$	99.5	——		——

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

2.4.4 Utilitarian Cutoffs

The third hypothesis states that, in the clean dataset, groups switch to the safe action earlier than the utilitarian optimum predicts. The first result shows that groups stop too early relative to equilibrium, which implies that they also stop too early relative to the utilitarian optimum. However, it is interesting to see how groups stop relative to the utilitarian optimum in each subset of the majority-vote observations. Table 2.4 compares the mean stopping time in the majority-vote observations to the utilitarian optimum. The mean stopping time in the majority-vote observations is 113.8 ticks, which is less than half of the 261.4 ticks that the utilitarian optimum predicts. This result is significant at the 1 percent level. Hypothesis tests, for this subsection, are conducted using bootstrapped regressions with 5000 samples clustered at the group level.

The majority-vote observations can once again be broken down into subsets based on whether a winner is possible. However, the data in the last two subsets change due to the utilitarian predictions differing from equilibrium predictions. Under the

[12]For time-dependent variables in the Cox regression, it is crucial that a covariate can not reach forward in time (Therneau et al., 2018). This problem is avoided in my analysis as a group can only be recorded as observing a winner if it had in fact switched to the safe action after observing a winner.

utilitarian optimum, a winner is predicted when one group member has a good state and the first reward arrives within the first 132 ticks. A winner is not predicted, but possible, when one group member has a good state and the first reward arrives after the first 132 ticks. Groups stop significantly earlier than predicted by the utilitarian optimum in each subset.

Result 4: *Groups switch to the safe action earlier than the utilitarian optimum predicts (evidence supporting Hypothesis 3).*

2.5 Exploring Under-Experimentation

The first result shows that both groups and single decision makers stop trying the risky action earlier than predicted. This result is a product of under-experimentation from both voters and single decision makers. In this section, I analyze two additional treatments, using a between-subjects design, that each remove a potential bias from the individual experimentation problem. The individual experimentation environment is analyzed as it has fewer potential biases. There are thirty-two subjects in each additional treatment.

The first additional treatment is the belief treatment, which analyzes how belief updating affects experimentation with the risky action. The belief treatment is identical to the single-agent treatment except that subjects are now given the Bayesian update after every tick. The results from this treatment will show whether removing belief updating restores optimal experimentation. In both additional treatments, subjects play twenty-five periods and complete the Eckel and Grossman (2002) risk aversion task.

The second additional treatment is the no-discounting treatment, which analyzes how discounting biases affect experimentation with the risky action. One possible discounting bias is that subjects may overweight each tick's small probability of terminating the period. The no-discounting treatment is identical to the belief treatment

Table 2.5.: Mean stopping times, in ticks, for the single-agent treatments. Stopping times use the last fifteen periods of each treatment.

	Equil.		Actual
First Single-Agent Treatment	187.0	$\overset{***}{>}$	137.7
Belief Treatment	187.0	$\overset{***}{>}$	104.3
No-Discounting Treatment	187.0	$\overset{***}{>}$	142.2

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

except that the period length is now fixed to 333 ticks and the value of a reward is increased to 3.65 experimental dollars. Discounting biases are removed by removing the random termination of the period. The period length is set to 333 ticks as it is the tick length closest to the expected period length in the single-agent treatment. The value of a reward is increased to 3.65 experimental dollars to keep optimal experimentation the same as in the single-agent treatment. The results from this treatment can determine whether removing both belief updating and discounting restores optimal experimentation.

The results show that belief updating, discounting, and risk aversion are unlikely to explain why subjects under-experiment with the risky action. I focus on the last fifteen periods of each treatment and remove observations with good states or with a period length less than 200 ticks. Table 2.5 displays the average stopping time for the first single-agent treatment, the belief treatment, and the no-discounting treatment.[13] The average stopping time in the first single-agent treatment is 137.7, which is similar to the overall single-agent treatment stopping time. This stopping time is significantly less than the predicted 187 ticks at the 1 percent level using bootstrapped regressions clustered at the subject level.[14] Hypothesis tests in this section are conducted using bootstrapped regressions with 5000 bootstrap samples

[13] The "first single-agent treatment" is the single-agent data from the subjects who started off with twenty-five periods of the single-agent treatment

[14] I can cluster at the subject level because none of these subjects have yet interacted with their group members.

clustered at the subject level. The average stopping time in the belief treatment is 104.3 ticks, while the average stopping time in the no-discounting treatment is 142.2 ticks. Both of these averages are significantly less than the predicted 187 ticks at the 1 percent level. A random effects regression, with subject-level random effects, shows that these three treatments are each not significantly different from each other at the 10 percent level. Additionally, a random effects regression, with subject-level random effects, of stopping time on risk aversion is statistically insignificant at the 10 percent level in each treatment.

The lack of support for belief updating, discounting, and risk aversion suggests that heuristics may explain subjects' under-experimentation. One possible heuristic is that subjects prefer to stop experimentation around 100 ticks. However, only 2.4 percent of subjects have a Product Limit estimated mean stopping time, in bad states, between 95 and 105 in the last fifteen periods of the single-agent treatment. Another possibility is horizon truncation, which was suggested as a possible explanation for under-experimentation in bandit problems in Anderson (2001). Horizon truncation suggests that subjects approximate the solution to a dynamic programming problem by solving a short horizon version of the problem and then adding an adjustment factor for the omitted periods. If subjects were to under-adjust, this could explain under-experimentation.

2.6 Conclusion

This paper uses a laboratory experiment to analyze how a group of voters collectively experiment with a new, potentially heterogeneous, reform. The laboratory experiment implements the Strulovici (2010) collective experimentation model. I analyze a subset of data where groups and single decision makers should eventually abandon the reform. Groups' collective decisions support predictions of the model. However, both groups and single decision makers stop trying the reform earlier than predicted. I find that risk aversion, belief updating, and discounting biases are unlikely to explain this under-experimentation.

These results have consequences for policy experimentation outside of the laboratory. Collective decisions are consistent with experimentation incentives, which suggests that voting rules with stronger experimentation incentives should be implemented to increase social welfare. Strulovici (2010) describes two voting reforms with stronger experimentation incentives that should be considered. The first reform forces voters to agree upon experimentation cutoffs before implementation of the policy; social welfare theoretically increases as voters make decisions under fully aligned incentives. The second reform increases the number of votes required for the policy deterministically over time; this reduces the possibility of the safe action being chosen when the policy is socially efficient.

There are many avenues for future research. First, future studies can explore other possible explanations for why subjects under-experiment with the risky action. A few other possible explanations are cognitive biases and heuristics. Second, future studies can attempt to elicit experimentation cutoffs for unsure voters. I avoid the strategy method in this paper, but future papers can use my results to validate the elicitations from the strategy method. Finally, other voting rules, including the optimal voting rule of Strulovici (2010), can be tested in the laboratory. Other voting rules increase election of the risky action towards the utilitarian optimum.

3. IS EXPERIMENTATION INVARIANT TO GROUP SIZE? A LABORATORY ANALYSIS OF INNOVATION CONTESTS

This paper uses a laboratory experiment to investigate the role of group size in an innovation contest. Subjects compete in a discrete time innovation contest, based on Halac et al. (2017), where subjects, at the start of each period, are informed of the aggregate number of innovation attempts. I compare two innovation contests, a two-person and four-person contest, that only differ by contest size and have the same probability of obtaining an innovation in equilibrium. The four-person contest results in more innovations and induces more aggregate innovation attempts than the two-person contest. However, there is some evidence that the two-person contest induces more innovation attempts from an individual than the four-person contest. Subjects' behavior is consistent with subjects placing more weight on their own failed innovation attempts, when updating their beliefs, than their competitors' failed innovation attempts.

3.1 Introduction

Netflix launched an innovation contest, in 2006, to improve its current movie recommendation algorithm. The innovation contest offered $1 million to any team that could improve upon Netflix's algorithm by ten percent. Thousands of teams, from more than 100 countries, entered the contest and competed for the grand prize. The contest lasted for almost three years until one team won the $1 million prize in 2009.

Innovation contests, like the Netflix contest, often attract a large number of competitors. While it may appear that larger innovation contests are more likely to result in an innovation, this is not always predicted. Halac et al. (2017) analyze a contest that is similar to the Netflix contest in both prize-sharing scheme and disclosure policy. In this contest, they find that an increase in the number of contestants has no effect on the probability of obtaining an innovation. This result is surprising as empirical studies generally show that larger contests induce more aggregate effort (Sheremeta (2011)), which, in the Halac et al. (2017) environment, would result in larger contests being more likely to obtain an innovation.

In the Halac et al. (2017) environment, there is a principal who wants to obtain an innovation. The principal can implement various contests to try to obtain an innovation. One of these contests is the (continuous time) public winner-takes-all contest. In this contest, the principal rewards the full prize to the first agent that obtains an innovation and the principal discloses the arrival of an innovation to all agents as soon as it occurs. The possibility of an innovation depends on the state of nature. In a good state, each innovation attempt has the same probability of resulting in an innovation. In a bad state, an innovation attempt can never result in an innovation. The state of nature is initially unknown. In the case that no one has yet obtained an innovation, agents become more pessimistic about the state as their own innovation attempts fail and as they conjecture that other agents are trying and failing to produce an innovation.

The unique equilibrium of this contest is in stopping strategies. In equilibrium, an agent exerts effort until an innovation has been obtained or an agent's belief that an innovation is possible is low enough such that exerting effort decreases expected utility. In equilibrium, agents have correct beliefs about other agents' behavior and thus stop exerting effort when the number of aggregate failed innovation attempts is sufficiently high. As this critical level of failed innovation attempts is invariant to contest size, a change in contest size has no effect on the aggregate number of innovation attempts induced nor the probability of obtaining an innovation.

This paper analyzes a simple discrete time contest that has the same equilibrium in stopping strategies as the continuous time public winner-takes-all contest. This discrete time contest simplifies the environment of Halac et al. (2017) by informing agents, at the start of each period, of the aggregate number of innovation attempts. In this discrete time contest, agents no longer have to conjecture about other agents' past behavior. This contest further simplifies the environment by occurring in discrete time, which reduces the number of decisions that agents have to make. In order to keep incentives similar to the continuous time contest, agents are given the full prize in the case of a tie. This removes any strategic concern of tying, which drops out in continuous time. This discrete time contest is used in order to analyze whether this invariance can hold in a relatively simple laboratory environment.

The laboratory experiment consists of two treatments. In the first treatment, two subjects participate in the contest. In the second treatment, four subjects participate in the contest. While these two contests only differ by group size, they have the same probability of obtaining an innovation in equilibrium. A laboratory experiment is used as it allows for groups to be randomly assigned to contests with a good or bad state of nature. The random assignment of groups to contests with bad states allows for an unbiased estimate of each contest's ability to induce innovation attempts. In bad states, the contest can never end prematurely due to an innovation being obtained.

The data shows differences between the two-person and four-person contests. First, the four-person contest induces more aggregate innovation attempts than the two-person contest. This result is shown through the number of innovation attempts in bad states. The four-person contest has more innovation attempts in bad states than the two-person contest. Second, the four-person contest results in more innovations than the two-person contest. Third, as predicted by theory, there is some evidence that the two-person contest induces more innovation attempts from an individual than the four-person contest. This result is shown through bad states; a subject, on average, has more innovation attempts in bad states in the two-person contest. Thus, the first two results do not appear to be driven by subjects becoming more willing to attempt an innovation in larger contests.

These first two results are unexplained by theory. I propose differential weighting of experimentation as an explanation. This possible bias would imply that subjects place more weight on their own failed innovation attempts, when updating their beliefs, than their competitors' attempts. If this differential weighting of experimentation holds, subject behavior is less responsive to an increase in the number of competitors than predicted by theory. I develop and estimate a model of differential weighting of experimentation. Through Maximum Likelihood Estimation, I find that subject behavior is consistent with differential weighting of experimentation.

This paper contributes to three strands of literature. First is the literature on innovation contests. This literature has mostly focused on theoretically analyzing innovation contests. This paper implements a contest based on Halac et al. (2017), who theoretically model innovation contests with differing prize-sharing schemes and disclosure policies. Bimpikis et al. (2019) analyze prize-sharing schemes and disclosure policies for the first stage of a two stage contest. Other theory papers that analyze innovation contests are Choi (1991) and Chowdhury (2017). Recently, there has been experimental research conducted on innovation contests. Deck and Kimbrough (2017) experimentally analyze four types of innovation contests introduced in Halac et al. (2017). They find that contests where principals do not disclose successful innovation attempts outperform contests where principals disclose successful innovation attempts. My paper is the first to experimentally analyze how group size influences an innovation contest. I find that a larger contest results in more innovations than a smaller contest.

This paper also contributes to the literature on multi-agent bandit experiments. This literature has mostly focused on the single-agent bandit problem (Meyer and Shi (1995); Banks et al. (1997); Anderson (2001, 2012); Gans et al. (2007)). Recently, there have been experiments conducted on multi-agent bandit problems. Deck and Kimbrough (2017) compare different innovation contests in the laboratory. Hoelzemann and Klein (2018) analyze a game of strategic experimentation, where subjects are predicted to free-ride on other subjects' experimentation. They find evidence of free-riding in the lab. Hudja (2019) analyzes collective experimentation, where groups, under majority voting, are predicted to be less willing to try reforms than single decision makers. Hudja (2019) finds that the data is consistent with this result. My paper analyzes an environment where aggregate experimentation is predicted to be invariant to group size. I find that a change in group size has an effect on aggregate experimentation.

Finally, this paper contributes to the literature on the interaction between individual effort and contest size in contest experiments. Sheremeta (2011) finds that average effort is decreasing in the number of players in a Tullock contest. Morgan et al. (2012) also find this result, while Lim et al. (2014) find that average effort does not respond to the number of players. In all-pay auctions, Gneezy and Smorodinsky (2006) find that average effort decreases in the number of players, while Harbring and Irlenbusch (2003) find that average effort is weakly increasing in the number of players. In rank-order tournaments, Orrison et al. (2004) find that average effort does not change when the number of players increases, while List et al. (2010) find that average effort is decreasing in the number of players. My paper is the first to analyze how group size influences individual effort in an innovation contest environment. I find some evidence that a smaller contest induces more innovation attempts from an individual than a larger contest.

3.2 Theory

This section focuses on the contest implemented in the experiment. Section C.1 of the appendix displays details, and states the equilibrium, of the continuous time public winner-takes-all contest of Halac et al. (2017).

A principal wants to obtain an innovation. An innovation may or may not be possible depending on the state of nature. If the state is good, obtaining an innovation is possible; if the state is bad, obtaining an innovation is impossible. The principal has access to $N \geq 2$ agents who may work towards obtaining an innovation. Time is discrete and there are T periods in which an agent can work towards obtaining an innovation. In each period, an agent chooses whether to exert effort. Exerting effort costs an agent c. If an agent exerts effort in a good state, she obtains an innovation with probability λ (innovations are conditionally independent given the state). If an agent exerts effort in a bad state, she cannot obtain an innovation.

The principal implements an innovation contest that lasts T periods. In each period of the contest, an agent decides whether to exert effort. At the beginning of each period, the principal reveals whether an innovation has been obtained and the level of aggregate effort exerted in the contest.[1] The first agent to obtain an innovation receives a prize of $\overline{w}$. In the case of a tie, each agent who obtained an innovation receives $\overline{w}$.[2]

There is an initial probability p_0 that the state is good. Each agent updates their belief according to Bayes' rule. Thus, if an innovation has been obtained, the agent knows that the state is good. In the case that an innovation has not yet been obtained, the Bayesian update can be written as

$$\frac{p_0(1-\lambda)^{A^{t-1}}}{p_0(1-\lambda)^{A^{t-1}} + (1-p_0)}, \tag{3.1}$$

where A^{t-1} is the aggregate number of innovation attempts in the contest through the first $t-1$ periods.

An agent (denoted by i) chooses a strategy to maximize

$$\sum_{t=1}^{T} [(p_{t-1}\lambda\overline{w} - c)a_{i,t}](1 - p_0(1 - (1-\lambda)^{(A^{t-1})})), \tag{3.2}$$

where $a_{i,t}$ is the effort exerted by agent i in period t. The myopic expected payoff from exerting effort in a given period, when the prize has not yet been won, is denoted by $p_{t-1}\lambda\overline{w} - c$. An agent who exerts effort obtains the innovation, worth $\overline{w}$, with

[1]Theory is unchanged whether aggregate effort is public knowledge or hidden. Aggregate effort is provided as public knowledge in order to simplify the environment for subjects.
[2]A tie occurs when two or more agents obtain an innovation in the same period. Agents receive the same prize in the case of a tie in order to remove the strategic concerns of splitting the prize.

probability $p_{t-1}\lambda$. However, attempting an innovation costs c. The term $(1 - p_0(1 - (1-\lambda)^{(A_{t-1})}))$ is the probability that no agent has obtained an innovation up until period t.

It is immediate from (2) that there is always an equilibrium in stopping strategies. An equilibrium where each agent exerts effort until $p < \frac{c}{\lambda w}$ or until an innovation is obtained (or T is reached) always exists. There may be other equilibria depending on the parameters. However, it can be shown that any other equilibrium is outcome equivalent in the sense that the equilibrium induces the same level of aggregate effort and has the same probability of resulting in an innovation.[3] The previously mentioned equilibrium in stopping strategies will be referenced for predictions as it is intuitive and the experiment only tests predictions for the end of the contest.

3.3 Experimental Design

The treatments are designed to analyze the effect of group size in this contest. The experiment has two main goals. The first main goal of the experiment is to compare the aggregate effort induced by two contests that only differ by contest size and have the same probability of obtaining an innovation in equilibrium. The second main goal of the experiment is to create a dataset that I can use to test explanations for any possible differences between the two contests.

3.3.1 Treatments and Parameters

There are two treatments in this experiment. The first is the two-person treatment, where two subjects compete in an innovation contest. The second treatment

[3]Denote the level of aggregate effort in the previous equilibrium as SA. If an equilibrium predicts that $\sum_{i=1}^{N}\sum_{t=1}^{T} a_{i,t} > SA$ in a bad state, then there is at least one agent who would exert effort when $p < \frac{c}{\lambda w}$ in the absence of an innovation. This agent would be better off not exerting effort when $p < \frac{c}{\lambda w}$. If an equilibrium predicts that $\sum_{i=1}^{N}\sum_{t=1}^{T} a_{i,t} < SA$, in a bad state, then there is at least one agent who is exerting less effort, in the absence of an innovation, than they would if they were using the previous stopping strategy. This agent would be better off using the previous stopping strategy.

is the four-person treatment, where four subjects compete in an innovation contest. The experiment utilizes a between-subjects design. Thirty-two subjects participated in the first treatment and thirty-two subjects participated in the second treatment. Each treatment consists of thirty contests. There are two sessions per treatment. The experiments were implemented in z-Tree (Fischbacher, 2007).

The parameters for the experiment are chosen to simplify the environment. The initial probability that the state is good, p_0, is set at 0.75. The probability that effort results in an innovation in a good state, λ, is set at 0.10. The prize for an innovation, $\overline{w}$, is set at \$10.00. The cost of effort, c, is set at \$0.30. Finally, the length of the contest, T, is set to fifteen periods.

The predictions under these parameters are as follows. In each treatment, each subject is predicted to continue to exert effort until an innovation has been obtained or until nineteen units of effort have occurred. Thus, in the two-person treatment, each subject is predicted to exert up to ten units of effort in the absence of an innovation. In the four-person treatment, each subject is predicted to exert up to five units of effort in the absence of an innovation.

3.3.2 Experiment

Instructions for the experiment were passed out before the start of the experiment. After the instructions were read, subjects answered five comprehension questions that were each worth \$1.00. After the comprehension questions were answered, subjects started the experiment.

The environment is described to subjects through an analogy of balls being drawn from a bag. Subjects are informed that there are two bags: (i) a uniform bag, which consists of twenty blue balls and (ii) a mixed bag, which consists of two red balls and eighteen blue balls. In this analogy, the mixed bag is a good state, with a red ball being an innovation. Contests are described to subjects as cycles. In each contest, subjects have a 75 percent chance of drawing from the mixed bag and a 25 percent chance of drawing from the uniform bag. In each contest, subjects are drawing with replacement.

At the start of each contest, subjects are randomly matched. In each period of the contest, subjects have the choice to draw or not draw a ball. If a subject chooses to not draw a ball, she is paid a fixed amount. If a subject chooses to draw a ball, two outcomes can occur: (i) she draws a blue ball and is not paid anything or (ii) she draws a red ball, she wins the prize, and the contest ends.[4] In the case that the contest ends prematurely, each subject in the contest receives the opportunity cost for the remaining periods in the contest.

Subjects receive feedback throughout the experiment. At the beginning of each period, subjects are informed of the total number of balls drawn in the contest. Subjects receive a notification at the end of the period if the period has resulted in a red ball being drawn. This notification states that a red ball had been drawn and lets subjects know whether they drew a red ball. At the end of each contest, subjects are informed of the number of balls they had individually drawn, the aggregate number of balls drawn in the contest, the number of red balls drawn in the contest, and their individual payoff.

Upon completion of thirty contests, the Holt and Laury (2002) task was administered to elicit risk aversion. After the risk aversion task was completed, the Short Grit Scale (Duckworth and Quinn, 2009) was administered. The Short Grit Scale consists of eight questions that test for the psychological construct of grit.

[4]The contest ends once the prize is won to prevent subjects from making trivial decisions.

3.3.3 Pre-Generated Random Variables and Payment

The random variables in the experiment were pre-generated.[5] One sequence of contest states were drawn for the experiment. The sixth, ninth, tenth, sixteenth, twentieth, twenty-third, twenty-fourth, and twenty-ninth contests had bad states. The remaining contests had good states. For each contest with a good state, there was one sequence of innovations pre-generated. For example, the number of balls drawn in order to obtain an innovation in the seventh contest was the same in each grouping and each session.

3.3.4 Theoretical Predictions

Subjects are predicted to exert effort until an innovation has been obtained or until nineteen or more units of effort have been exerted. In the two-person contest, this results in subjects exerting effort until an innovation is obtained or ten periods have ended. In the four-person contest, this results in subjects exerting effort until an innovation is obtained or five periods have ended. In this subsection, I state the hypotheses that follow from these theoretical predictions.

The first hypothesis regards the level of aggregate effort that each contest induces. I refer to the level of aggregate effort that a contest induces as the level of aggregate effort that would occur if the contest ended without an innovation. Each contest theoretically induces the same level of aggregate effort as each contest theoretically induces twenty units of aggregate effort. This leads me to Hypothesis 1.

[5]In the instructions, subjects were not informed of this pre-generation, but they were informed of the distribution that resulted in these pre-generated random variables. I decided to pre-generate these random variables as pre-generation increases power for the hypothesis tests by decreasing noise. Additionally, I pre-generate so that any influence of the realization of random variables, on subject behavior, is held constant across sessions and treatments. Papers such as Engle-Warnick and Slonim (2006) and Dal Bó and Fréchette (2011) have shown that the realization of random variables can influence subject behavior. I decided to not inform subjects about the pre-generation because this information does not change beliefs or influence behavior and this information may lead to confusion from subjects.

Hypothesis 1: *Each treatment induces the same level of aggregate effort.*

The first hypothesis states that each treatment induces the same level of aggregate effort. The second hypothesis follows from the first hypothesis. As both treatments induce the same level of aggregate effort, each treatment is predicted to result in the same number of innovations in expectation. The probability of a contest resulting in an innovation is given by

$$p_0(1-(1-\lambda)^{A^*}),$$

where A^* is the level of aggregate effort induced by the contest. As each contest induces twenty units of aggregate effort, the probability of each contest resulting in an innovation is the same.

However, in the experiment, each treatment should have the same number of innovations regardless of the realizations of any random variables. This is due to the pre-generation of random variables. In each numbered contest, the level of aggregate effort needed to obtain an innovation is the same in each treatment. This leads me to hypothesis 2.

Hypothesis 2: *Each treatment results in the same number of innovations.*

The third hypothesis also follows from the first hypothesis. I refer to the level of individual effort that a contest induces as the average level of individual effort that would be observed if the contest ended without an innovation. As each contest induces twenty units of aggregate effort and there are less contestants in the two-person contest, the two-person contest is predicted to induce more effort from an individual than the four-person contest. This leads me to Hypothesis 3.

Hypothesis 3: *The two-person contest induces more effort from a subject than the four-person contest.*

In the results section, two sets of data will be used to test the three hypotheses. Hypotheses 1 and 3 will be tested using a subset of data where the contest can not end prematurely. I will analyze these hypotheses on contests with bad states. This allows for an unbiased estimate of the aggregate effort induced by a contest as the contest can not end prematurely due to an innovation being obtained. This also provides an unbiased estimate of each subject's willingness to exert effort in the contest. Hypothesis 2 will be tested using both good and bad states as I care about the overall innovation percentage.

3.3.5 Procedures

Sessions were run at Purdue University in April 2017. Subjects received payment from the comprehension questions, the contests, and the risk aversion task. Subjects were paid for each comprehension question, for two randomly chosen contests, and for one random decision in the risk aversion task.

3.4 Experimental Results

Section 3.4.1 displays summary statistics for the two treatments. Section 3.4.2 analyzes the three hypotheses. Section 3.4.3 analyzes the factors that influence individual behavior. Section 3.4.4 analyzes a belief-updating bias that may explain the data.

3.4.1 General Results

Table 3.1 displays the summary statistics for the two treatments. Table 3.1, and the results section, analyzes the last twenty contests.[6] Table 3.1 displays the mean

[6]Results for the first ten contests can be found in Section C.2 of the appendix. Figure 3 in Section C.2 of the appendix shows the mean level of aggregate effort in all of the bad states that subjects participated in. Behavior in the two-person treatment does not appear to stabilize until the tenth contest.

Table 3.1.: Mean statistics on the level of aggregate effort in bad states, the innovation percentage, and the level of individual effort in bad states. The equilibrium innovation percentage is found by averaging the predictions for each contest. The standard error of the mean is in parentheses.

Treatment:	Two-Person			Four-Person		
	Equil.		Actual	Equil.		Actual
Aggregate Effort in Bad States	20	$\overset{**}{>}$	15.02	20	$<$	21.15
	—		(0.66)	—		(0.90)
Innovation Percentage	0.70	$\overset{***}{>}$	0.61	0.70	$=$	0.70
	—		(0.03)	—		(0.04)
Individual Effort in Bad States	10	$\overset{**}{>}$	7.51	5	$<$	5.29
	—		(0.033)	—		(0.028)

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

level of aggregate effort in bad states, the innovation percentage, and the mean level of individual effort in bad states. The mean level of aggregate effort in the two-person contest in bad states is 15.02. This is significantly different from the predicted 20 at the five percent level (p-value=0.028) using a regression of the difference in each contest's realized and predicted level of aggregate effort on a constant.[7] Regressions in this subsection are all bootstrapped regressions with 5000 bootstrap samples clustered at the session level. The mean level of aggregate effort in the four-person contest in bad states is 21.15. This is not significantly different from the predicted 20 attempts at the ten percent level (p-value=0.498) using a regression of the difference in each contest's realized and predicted level of aggregate effort on a constant.

The innovation percentage is 61 percent in the two-person treatment. This is significantly less than the predicted 70 percent at the 1 percent level using a regression of the difference in each contest's realized innovation rate and predicted innovation

[7]Under-exertion of effort may seem surprising given that many contest experiments find over-exertion of effort. However, this contest utilizes a bandit framework and other papers that implement a bandit framework in the laboratory (Anderson (2001); Meyer and Shi (1995); Hudja (2019)) find under-experimentation.

rate on a constant.[8] For example, if a specific contest resulted in an innovation, this contest's difference is recorded as 0.30. If a specific contest did not result in an innovation, this contest's difference is recorded as -0.70. The innovation percentage is 70 percent in the four-person treatment. This is exactly at the predicted seventy percent and is thus not significantly different from the prediction.[9]

The mean level of individual effort in bad states is 7.51 in the two-person treatment. This is significantly different from the predicted 10 at the five percent level (p-value=0.028) using a regression of the difference in each individual's realized and predicted cumulative level of effort, in a given contest, on a constant. The mean level of individual effort in bad states is 5.29 in the four-person treatment. This is not significantly different from the predicted five at the ten percent level (p-value=0.498) using a regression of the difference in each individual's realized and predicted cumulative level of effort, in a given contest, on a constant.

Result 1: *Contests in the two-person treatment induce less aggregate effort, result in less innovations, and induce less individual effort than predicted. Contests in the four-person treatment do not significantly differ from the predicted level of induced aggregate effort, the predicted number of innovations, or the predicted level of induced individual effort.*

3.4.2 Hypotheses

Hypothesis 1 states that each treatment induces the same level of aggregate effort. Table 3.1, and the result that only the two-person contest induces less aggregate effort than predicted, suggest that the level of aggregate effort, in bad states, is greater in the four-person treatment than in the two-person treatment. Figure 3.1 addresses

[8]The predicted seventy percent stems from the pre-generated states and arrival of innovations.

[9]As shown in Figure 3.2, an innovation is not always obtained when it is predicted for the four person contest. There is one time when an innovation is obtained when it is not predicted and one time when an innovation is not obtained when it is predicted.

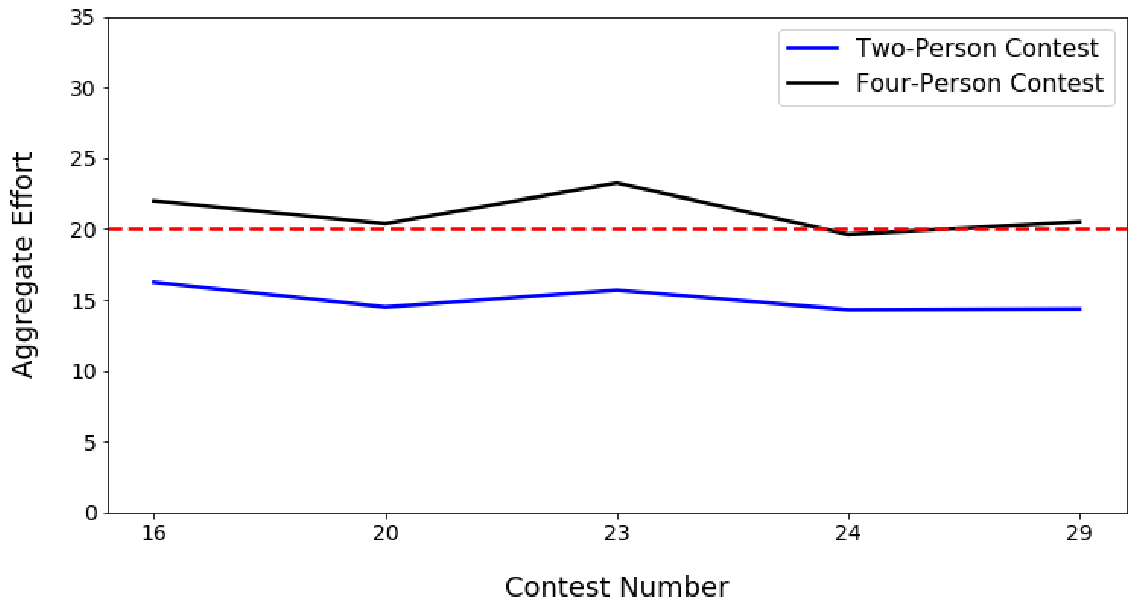


Figure 3.1.: Mean level of aggregate effort in bad states for each treatment. The red dotted line displays the equilibrium prediction for aggregate effort.

this possibility further by plotting the mean level of aggregate effort for each treatment in each contest with a bad state. The figure shows a clear separation between the two treatments. The four-person treatment has a higher mean level of aggregate effort in each contest with a bad state than the two-person treatment. This result can be formalized with a regression of the level of aggregate effort induced in each contest on the treatment, with 5000 bootstrap samples clustered at the session level. This regression shows that the four-person treatment has a significantly greater level of aggregate effort, in bad states, than the two-person treatment (p-value=.048).

Result 2: *Contests in the four-person treatment induce more aggregate effort than contests in the two-person treatment (evidence against Hypothesis 1).*

Hypothesis 2 states that each treatment results in the same number of innovations. Table 3.1, and the result that only the two-person contest results in less innovations

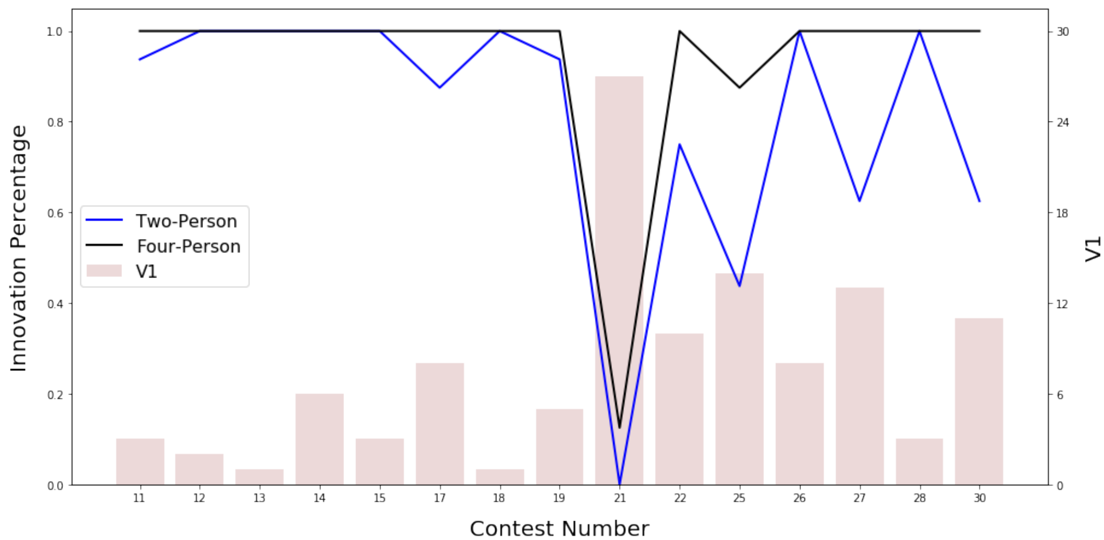


Figure 3.2.: Frequency of obtaining an innovation by treatment. The blue and black lines represent the two-person and four-person treatments, respectively. The bar graph displays the minimum number of balls drawn required to obtain an innovation, which is denoted by $V1$.

than predicted, suggest that the innovation percentage is greater in the four-person treatment than the two-person treatment. Additionally, the result that the four-person treatment induces more aggregate effort also suggests that the four-person treatment results in more innovations. Figure 3.2 addresses this possibility further by displaying the innovation percentage, for each treatment, in each period where an innovation is possible. Innovation percentage in bad states is not displayed as innovations are not possible in bad states. Figure 3.2 shows that the innovation percentage in the four-person treatment is greater than or equal to the innovation percentage in the two-person contest for each contest where an innovation is possible.[10] Hypothesis 2 can be more formally tested by running a regression of an indicator variable denot-

[10]While there appears to be a trend in the number of innovation attempts required to obtain an innovation, this trend does not appear to influence the results of this section. Figure 3.1 shows that the level of aggregate effort induced by each treatment is stable over time. This stability additionally implies that the level of individual effort induced in each treatment is stable over time and that a four-person contest should consistently be more likely to obtain an innovation than a two-person contest.

ing whether a contest resulted in an innovation on the treatment. This regression is bootstrapped regression with 5000 bootstrap samples clustered at the session level. This regression shows that the four-person contest is more likely to result in an innovation than the two-person contest (p-value<0.01).

Result 3: *Contests in the four-person treatment result in more innovations than contests in the two-person treatment (evidence against Hypothesis 2).*

Hypothesis 3 states that contests in the two-person treatment induce more effort from an individual than contests in the four-person treatment. While the four-person contest results in more innovations and induces more aggregate effort, the two-person contest may still induce more effort from an individual than the four-person contest. Table 3.1 suggests that contests in the two-person treatment actually do induce more individual effort than contests in the four-person treatment. This result can be formally tested using a regression of each subject's mean level of cumulative effort in bad state contests on the treatment, with 5000 bootstrapped samples clustered at the session level. This regression shows that individual effort is higher in bad states in the two-person contest than the four-person contest. This result is significant at the 10 percent level (p-value=.088).

Result 4: *There is some evidence that contests in the two-person treatment induce more effort from an individual than contests in the four-person treatment (some evidence supporting Hypothesis 3).*

Table 3.2.: Logistic regression of the choice to exert effort on multiple covariates. Standard errors in parentheses. Specifications (1) addresses the two-person treatment, while specifications (2) addresses the four-person treatment.

	Two Person Treat. Effort	Two Person Treat. Effort	Four Person Treat. Effort	Four Person Treat. Effort
Aggregate Effort	-0.197***	-0.234***	-0.269***	-0.270***
	(0.009)	(0.013)	(0.060)	(0.065)
Risk	——	-0.830***	——	-4.552***
	——	(0.049)	——	(0.719)
Grit	——	-0.650***	——	-0.183***
	——	(0.050)	——	(0.029)
Constant	0.481***	24.353***	3.399***	55.093***
	(0.070)	(1.844)	(0.299)	(8.080)
N	2,175	1,650	2,325	1,800

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

3.4.3 Individual Behavior

The previous subsection focused on testing the three hypotheses. This subsection moves away from the hypotheses by focusing on analyzing the factors that influence the decision to exert effort.

Regression analysis can be used to understand the decision making process of subjects in the experiment. This subsection analyzes four different regressions. Each regression only includes bad state contests in the last twenty periods.[11] Additionally, each regression is clustered at the session level. Table 3.2 displays these four regressions.

The first two regressions focus on the decision to exert effort in the two-person treatment. The first regression is a logistic regression of an indicator variable denoting whether subject i exerted effort in period p of contest j on the current level of aggregate effort in the contest, contest dummy variables, and subject dummy variables.[12] The coefficient on aggregate effort is negative and significant at the one

[11]The regression only includes bad state contests as these are the only contests where subjects make an effort decision in each period. Thus, each period is equally represented for the period variable.

[12]Contest dummy variables are included instead of a contest number variable because not every contest, of the last twenty periods, is represented in this regression. However, the regression has

percent level. The second specification is similar to the first specification except that both risk aversion and grit are included in the regression. The coefficient on risk aversion is negative and significant at the one percent level and the coefficient on grit is negative and significant at the one percent level.[13] These regressions suggest that subjects in the two-person treatment are less likely to exert effort when the state is more likely to be bad and that a subject's decision to exert effort is decreasing in risk aversion and grit.

The last two regressions focus on the decision to exert effort in the four-person treatment. The third regression is similar to the first regression. The third regression is a logistic regression of an indicator variable denoting whether subject i exerted effort in period p of contest j on the current level of aggregate effort in the contest, contest dummy variables, and subject dummy variables. Similarly to the two-person treatment, the coefficient on aggregate effort is negative and significant at the one percent level. The fourth regression is similar to the third regression except that both risk aversion and grit are included in the regression. Similarly to the two-person treatment, the coefficient on risk aversion is negative and significant at the one percent level and the coefficient on grit is negative and significant at the one percent level. These regressions suggests that subjects in the four-person treatment are less likely to exert effort when the state is more likely to be bad and that a subject's decision to exert effort is decreasing in risk aversion and grit.

Result 6: *Subjects' willingness to exert effort is decreasing in aggregate effort, risk aversion, and grit.*

---

similar results with a contest number variable. Additionally, the results are similar if a period variable denoting the period in the contest is introduced.

[13]The coefficient on grit is counter-intuitive, but is consistent with a similar study. Hudja et al. (2019) analyze the decision to exert effort on various personal characteristics in an innovation contest where the principal rewards the best innovation at a pre-specified date. They find that the coefficient on grit is negative in all of the regressions and that it is significant in an individual innovation task.

3.4.4 Differential Weighting of Experimentation

This section has previously showed that the four-person contest induces more aggregate effort and results in more innovations than the two-person contest. This subsection explores differential weighting of experimentation as a possible explanation for these differences. Through Maximum Likelihood Estimation, I will analyze whether subject behavior is consistent with this bias.

Differential weighting of experimentation occurs if subjects place more weight on their own failed innovation attempts, when updating their beliefs, than their competitors' failed innovation attempts. It is similar to base rate neglect/conservatism, where subjects can place too much weight or too little weight on new information relative to their prior. However, under differential weighting of experimentation, subjects implicitly treat new information generated from their own innovation attempts differently than new information generated from their competitors' failed innovation attempts.

There are a few reasons to expect that individuals might implicitly differentially weight experimentation. First, the information obtained by experimentation is costly, while the information obtained by other individuals' experimentation is not. There is evidence that subjects overweight costly information relative to free information (Robalo and Sayag, 2018). Second, subjects receive monetary feedback, in the form of \$0.00 or \$10.00, when attempting an innovation and receive no monetary feedback while other agents attempt innovations. There is evidence that belief updating is larger under monetary feedback (Bennett et al., 2019). Lastly, the information from attempting an innovation may be more salient than the information from other agents' innovation attempts. This information may be more salient in part because it is costly and comes in the form of monetary feedback. There is evidence that more salient information receives considerably more weight than less salient information in belief updating (Camacho et al., 2011).

In this section, for ease of exposition, I estimate a risk-neutral model of experimentation. Section C.3 of the appendix estimates a model under risk aversion and finds similar results to this subsection. Subjects are modeled as using stopping strategies based on a cutoff belief.[14] Subjects are assumed to prefer to stop attempting an innovation once the myopic benefit of innovating is below the myopic cost of innovation. This cutoff belief is given by $\frac{c}{\lambda w}$. Notice that this is the same cutoff as the continuous time equilibrium in Halac et al. (2017). Subjects are allowed to have biases in their belief updating. Subjects are allowed to over-weight or under-weight their own failed innovation attempts. Let ψ_i denote the weight that subjects place on their own failed innovation attempts.[15] Additionally, subjects are allowed to over-weight or under-weight their competitors' failed innovation attempts. Let ψ_o denote the weight that subjects place on their competitors' failed innovation attempts. Belief updating, in the absence of an innovation, is now given by

$$\tilde{p} = \frac{p_0(1-\lambda)^{\psi_i D_{t-1} + \psi_o O_{t-1}}}{p_0(1-\lambda)^{\psi_i D_{t-1} + \psi_o O_{t-1}} + (1-p_0)},$$

where D_{t-1} is the number of failed innovation attempts that a subject has had up until period t and O_{t-1} is the number of failed innovation attempts that her competitors have had up until period t.

While subjects are assumed to use stopping strategies, subjects are allowed to make errors. I assume that subjects' errors become more frequent as their beliefs

[14] Subjects appear to use stopping strategies in the data. Overall, only 8.79 percent of effort decisions (in bad states in the last twenty periods) appear to be inconsistent with a stopping strategy. I consider an observation to be inconsistent with a stopping strategy if an individual has previously chosen to not exert effort in the contest but currently has decided to exert effort. Notice that this is a liberal notion of a violation as an individual who exerted effort in every period but the first is credited with fourteen deviations, but the first period may have actually been the deviation from the stopping strategy.

[15] Placing a weight on new information, in the belief updating process, is a common approach to modeling individuals placing too much weight (base rate neglect) or too little weight (conservatism) on new information relative to their prior. This approach to modeling base rate neglect/conservatism has been used in Goeree et al. (2007) and Moreno and Rosokha (2016).

get closer to the cutoff belief. The loglikelihood for this model can be written in the following way

$$LogL = \sum_{i=1}^{n} log \left[\prod_{contest=1}^{C} \prod_{period=1}^{P} \Phi \left(\tilde{p}\lambda\overline{w} - c \right)^{E} \left(1 - \Phi \left(\tilde{p}\lambda\overline{w} - c \right) \right)^{1-E} \right],$$

where E is an indicator variable for whether subject i exerted effort in the current period of the current contest. This loglikelihood is maximized over the bad state contests for the last twenty contests.

This loglikelihood is maximized at a value of 3053.35. The values of ψ_i and ψ_o that maximize the loglikelihood are 3.02 and 2.78, respectively. These parameters suggest that subjects place less weight on their competitors' failed innovation attempts than their own failed innovation attempts. The restriction that $\psi_i = \psi_o$ is rejected at the one-percent level using a likelihood ratio test (restricted loglikelihood is equal to 3066.33). Additionally, the value of ψ_i is significantly greater than one at the one percent level using a likelihood ratio test (restricted loglikelihood is equal to 3156.98). The value of ψ_o is significantly greater than one at the one percent level using a likelihood ratio test (restricted loglikelihood is equal to 3161.04). This model suggests that subject behavior is consistent with differential weighting of experimentation.

Differential weighting of experimentation is consistent with larger contests resulting in more innovations and inducing more aggregate effort. Assuming that subjects use the stopping strategy modeled in this section, subjects, in the absence of an innovation, are predicted to exert effort until

$$\frac{p_0(1-\lambda)^{3.02D_{t-1}+2.78O_{t-1}}}{p_0(1-\lambda)^{3.02D_{t-1}+2.78O_{t-1}} + (1-p_0)} < \frac{c}{\lambda\overline{w}}.$$

As the number of subjects increase in a contest, O_{t-1} becomes a larger share of aggregate effort and subjects are thus more willing to experiment at a given level of aggregate effort (in the absence of an innovation).

3.5 Conclusion

This paper uses a laboratory experiment to analyze the role of group size in an innovation contest. The innovation contest is based on the continuous time public winner-takes-all contest of Halac et al. (2017). The probability of obtaining an innovation, and the level of aggregate effort induced, is predicted to be invariant to group size in the continuous time public winner-takes-all contest. I analyze a simpler discrete time version of this contest in order to see if this invariance can be found in the laboratory. I compare two contests, a two-person and four-person contest, that only differ by contest size and have the same probability of obtaining an innovation in equilibrium.

The data shows that there are differences between the two contests. The four-person contest induces more aggregate effort and results in more innovations than the two-person contest. While the four-person contest induces more aggregate effort, the two-person contest induces more individual effort. I suggest that subjects may be placing more weight on their own failed innovation attempts, when belief updating, than their competitors' failed innovation attempts. I develop, and estimate, a model that allows subjects to place non-negative weight on both their own failed innovation attempts and their competitors' failed innovation attempts when forming beliefs. I find, through Maximum Likelihood Estimation, that subject behavior is consistent with subjects placing more weight on their own failed innovation attempts.

These results have consequences for both theory and contest design. The larger contest in this paper results in more innovations and induces more aggregate effort than the smaller contest. While this experiment focused on a contest with a very simple environment, this result has implications for innovation contests, like the continuous time public winner-takes-all contest, where there is no predicted effect of group size on obtaining an innovation. In these types of contests, contest designers should allow for unlimited entry into contests. Additionally, contest designers may consider spending advertising money in order to increase the number of competitors.

There are many avenues for future research. First, future studies can analyze innovation contests in an experimental environment that is closer to the theoretical environment. In this paper, I analyze a simpler environment in order to test whether the predicted invariance occurs. However, future studies can build off of this experiment and allow the prize to be split or not display the level of aggregate effort. Second, future studies can analyze how group size affects other types of innovation contests. For example, in Halac et al. (2017), the probability of an innovation being obtained in the continuous time hidden equal-sharing contest is non-monotonic in group size. A future study can explore whether this non-monotonicity holds in a laboratory experiment. Lastly, future studies can explore other reasons why a larger contest may result in more innovations. While I focus on a belief-based explanation, there may be other explanations.

4. PUBLIC LEADERBOARD FEEDBACK IN INNOVATION CONTESTS: A THEORETICAL AND EXPERIMENTAL INVESTIGATION

with Brian Roberson and Yaroslav Rosokha

We investigate the role of performance feedback, in the form of a public leaderboard, in innovation competition that features sequential search activity and a range of possible innovation qualities. We find that in the subgame perfect equilibrium of contests with a fixed ending date (i.e., finite horizon), providing public performance feedback results in lower equilibrium effort and lower innovation quality. We conduct a controlled laboratory experiment to test the theoretical predictions and find that the experimental results largely support the theory. In addition, we investigate how individual characteristics affect competitive innovation activity. We find that risk aversion is a significant predictor of behavior both with and without leaderboard feedback and that the direction of this effect is consistent with the theoretical predictions.

4.1 Introduction

Innovation contests play an increasingly important role in research and development applications ranging from algorithmic design problems, to graphic design and marketing, to scientific breakthroughs. For example, in 2009, Netflix ran a crowdsourcing contest, the Netflix Prize, with a $1 million reward and the objective to "substantially improve the accuracy of predictions about how much someone is going to enjoy a movie based on their movie preferences." One key feature of this contest was a real-time leaderboard that provided information regarding performance of the top submitted algorithms. Since then, leaderboards have become a common feature of crowd-sourcing contests (e.g., Kaggle.com, drivendata.org, challenge.gov). However, the extent to which leaderboards contribute to innovation quality and innovation effort is not well understood.

In this paper, we theoretically and experimentally examine sequential-search innovation competition with a public leaderboard and a fixed ending date (i.e. finite horizon) and compare it to innovation competition with private performance feedback. In each period of the innovation contest, participants have the opportunity to engage in a costly innovation search. The search yields –a priori uncertain– innovation quality. We refer to the maximum of the innovation qualities among all of the opportunities that she has developed in previous periods as the *score*. At the fixed end of the contest, the participant with the highest score wins a prize. In this context, our focus is on the effects of information disclosure in the form of a public leaderboard on effort provision and innovation quality. We provide new results on the characterization of the subgame perfect equilibrium for searched-based innovation competition with public-leaderboard feedback, and compare that to the case of private performance feedback without a leaderboard as characterized by Taylor (1995). We then use a controlled laboratory experiment to test the theoretical predictions on effort provision and innovation quality with and without public leaderboard feedback.

In our sequential-search environment, information disclosure in the form of public leaderboard feedback generates incentives that are reminiscent of the dollar auction and the penny auction.[1] The dollar auction is a dynamic ascending-price auction with public feedback of the highest standing bid (i.e. a leaderboard) and the following features: (i) the auction opens with a standing bid of zero, (ii) the standing bid may only be increased by a fixed bid increment (iii) bidding continues until no bidder is willing to increase the standing bid (by the fixed bid increment), (iv) the highest bidder wins the item up for auction, and (v) both the highest and the second highest bidders pay their bids. Escalation arises in this setting because the losing bidder would always be better off if she incrementally increased the standing bid and won

[1] See, for example, Hinnosaar (2016) on the penny auction and Shubik (1971) and O'Neill (1986) on the dollar auction.

the auction. Our sequential-search contest with public leaderboard feedback presents a similar opportunity for escalation. In particular, at each stage of the contest, each participant (i) has a sunk research cost, (ii) knows whether he or she is in the lead, and (iii) the trailing player can try to take the lead by expending an incremental search cost. We show that, in equilibrium, participants who trail in the competition provide more effort. However, we also show that, in equilibrium, both participants who are ahead and participants who are behind strategically reduce their effort as the leader's existing innovation quality increases.

The main takeaway from our theoretical analysis is that despite the potential for leaderboard feedback to escalate the competition, we find that the presence of a leaderboard generates both lower equilibrium expected effort and lower equilibrium expected innovation quality than would be achieved without the leaderboard. The results of our experiment largely confirm these theoretical predictions. In particular, the experiment consist of two main treatments of the competition with the leaderboard (*leaderboard feedback* treatment) and without the leaderboard (*private feedback* treatment). We find that the private-feedback treatment results in more effort and a higher quality of the winning innovation than the leaderboard-feedback treatment. We also experimentally confirm that current leaders tend to exert less effort than followers and that both leaders and followers become less willing to exert effort as the innovation quality increases.

Our paper contributes to several active streams of literature. First, we contribute to the literature on innovation competitions. The existing approaches include but are not limited to variations on the all-pay auctions (e.g., Che and Gale, 2003; Chawla et al., 2015), the exponential-bandit contests (e.g., Halac et al., 2017; Bimpikis et al., 2019), two-stage difference-form contests (e.g., Aoyagi, 2010; Klein and Schmutzler, 2017; Goltsman and Mukherjee, 2011; Gershkov and Perry, 2009; Yildirim, 2005), crowdsourcing contests (e.g,. Terwiesch and Xu, 2008; DiPalantino and Vojnovic, 2009; Erat and Krishnan, 2012; Ales et al., 2017), and dynamic contests (e.g., Lang et al., 2014; Seel and Strack, 2016). In terms of studies that focus on feedback in contests, our work is closely related to Mihm and Schlapp (2018) who examine a two-period contest with leaderboard feedback, private feedback, and no feedback. The authors show that the level of uncertainty may interact with the designer's objective (i.e., average effort or best performance) and lead to feedback being optimal for some combination(s) of uncertainty and objective. Regarding models on search-based innovation competitions, our work is most closely related to Taylor (1995), Fullerton and McAfee (1999), and Baye and Hoppe (2003). In particular, although the existing literature on search-based innovation competition has considered the case of private feedback, our study is the first (to our knowledge) to provide equilibrium predictions for dynamic contests with the leaderboard feedback in a finite-horizon setting.

Second, we contribute to the experimental literature on feedback in contests. Relevant recent experimental work shows that feedback may not always be desired include Kuhnen and Tymula (2012), Ludwig and Lünser (2012), and Deck and Kimbrough (2017). Deck and Kimbrough (2017) experimentally confirm that in Halac et al. (2017) setting, withholding information leads to better innovation outcomes. This result arises from the fact that the information that your opponents have not procured the zero-one innovation lowers your own belief about the probability that innovation is possible. That is, information may be discouraging and, thus, hiding information may be valuable. In the dynamic effort provision setting in which there is range of possible outcomes Kuhnen and Tymula (2012) and Ludwig and Lünser (2012) find that feedback influences the dynamics of effort provision but not total effort. Our experimental results are consistent with some of the findings on the dynamics of effort provision observed in these papers. In particular, we find that leaders tend to reduce their effort, whereas followers tend to increase their effort. It is important to note, however, that these findings are not generalizable to all contest settings. In fact, in a recent survey, Dechenaux et al. (2015) highlight that in some cases feedback may result in the trailing player dropping out (e.g., Fershtman and Gneezy, 2011).

Finally, our work is related to the literature on factors that motivate individuals to innovate. In particular, on the experimental side, recent studies have examined the role of incentives (Ederer and Manso, 2013), preferences (Herz et al., 2014; Rosokha and Younge, 2017), and biases (Herz et al., 2014). On the empirical side, two recent surveys by Astebro et al. (2014) and Koudstaal et al. (2015) highlight that entrepreneurs are typically less risk and loss averse. In the current paper, we consider the extent to which risk aversion, loss aversion, and sunk-cost fallacy play a role in a search-based innovation competition.[2] Specifically, as part of our experiment, we elicited those three measures with incentivized multiple-price list tasks. In addition, we asked subjects to complete several unincentivized personality questionnaires. We find that risk aversion is a significant predictor of the number of costly innovation actions in the contest, with more risk-averse subjects taking fewer actions. However, we did not find that loss aversion, sunk-cost fallacy, or unincentivized measures of personality were predictive of subjects' behavior in the contest.

The rest of the paper is organized as follows: in section 4.2, we present the theoretical model. In section 4.3, we provide details of the experimental design. In section 4.4, we develop predictions for our environment and organize them into four hypotheses. In section 4.5, we present main results of the experiment. Finally, in section 4.6, we conclude.

4.2 Theory

Consider a two-player T-period dynamic innovation contest, along the lines of Taylor (1995). In this model, innovation activity takes the form of a search process with perfect recall. In each period $t \in \{1, \dots, T\}$, each player $i \in \{1, 2\}$ has the opportunity to exert effort at a cost of $c > 0$. If player i exerts effort, she obtains an innovation, with quality level $s_{i,t}$, a random variable that is distributed according to

[2] We focus on risk aversion and loss aversion as characteristics that have been documented to matter in the lab (e.g., Herz et al., 2014; Rosokha and Younge, 2017) and field (Astebro et al., 2014; Koudstaal et al., 2015) settings. In addition, we consider the sunk-cost fallacy because it has been shown to affect behavior in a related setting of penny auctions (Augenblick, 2015). Penny auctions are auctions in which agents pay to bid and the value of the item decreases after each bid. Augenblick (2015) shows theoretically how the sunk cost fallacy can lead to auctioneers making profit and finds empirical support for the sunk-cost fallacy in online penny auction data. Although our environment shares elements similar to the penny auction, we do not find evidence of the sunk-cost fallacy.

F, where F has a continuous and strictly positive density everywhere on its support, which is assumed to be a convex subset of R_+ with a lower bound of 0.[3] In the event that player i does not exert effort in period t, let $s_{i,t} = 0$. Player i's innovation "score" at the end of period t is denoted by $\bar{s}_{i,t} \equiv \max\{s_{i,1}, \ldots, s_{i,t}\}$. After T periods, the contest ends and the player with the higher innovation score at the end of period T, that is, the player i with $\bar{s}_{i,T} = \max\{\bar{s}_{1,T}, \bar{s}_{2,T}\}$, is awarded a prize with value $v \geq 2c$.[4] In the case of a tie, the winner is randomly chosen.

We examine two levels of feedback in the dynamic-innovation contest: (i) private feedback and (ii) leaderboard feedback. With the private-feedback innovation contest, at the beginning of each period t, each player i knows her current score ($\bar{s}_{i,t-1}$) and at the end of period t, player i observes her period t innovation quality $s_{i,t}$. With the leaderboard-feedback innovation contest, at the beginning of each period t, each player i knows, in addition to her own private feedback, the current max score,[5] $\max\{\bar{s}_{1,t-1}, \bar{s}_{2,t-1}\}$. In the following subsection, we characterize the subgame perfect equilibrium for the public-feedback innovation contest.

Throughout the rest of the paper, we use the convention, due to Taylor (1995), of referring to each draw of an innovation quality $s_{i,t}$ as a new innovation. Note, however, that an equivalent interpretation is that player i is working on one specific innovation and that each draw of an innovation quality $s_{i,t}$ is in regards to searching over quality improvements to that particular innovation. Depending on the application, this second interpretation may be more natural.

[3] In the experiment, we assume that innovations are exponentially distributed ($F(x;\lambda) = 1 - e^{-\lambda x}$ and $f(x;\lambda) = \lambda e^{-\lambda x}$, where $\lambda > 0$ is the rate parameter).

[4] Our analysis can be extended to the case of $v \in [c, 2c)$, but for the sake of brevity, we focus here on the case in which $v \geq 2c$.

[5] Note that the game in which, at the beginning of each period, each player observes both of the players' current scores is theoretically equivalent to the game in which, at the beginning of each period, each player observes her own score and the maximum of the players' scores.

4.2.1 Subgame Perfect Equilibrium in Innovation Contests

Private Feedback

The subgame perfect equilibrium for the private-feedback innovation contest is characterized by Taylor (1995). In particular, Proposition 2 of that paper establishes that the unique subgame perfect equilibrium takes the form of a stopping rule in which each player i continues to exert effort until her max score hits a threshold – denoted by ξ_i – and she stops exerting effort.

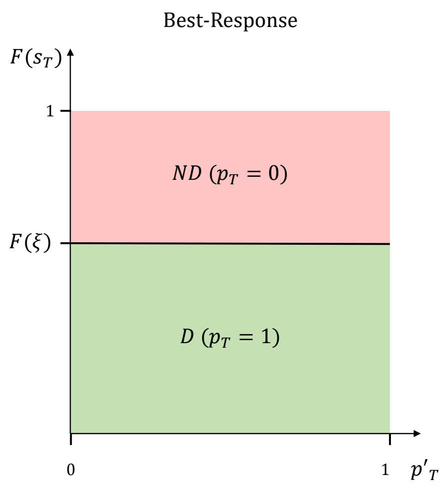


Figure 4.1.: Period T local best response for Private Feedback. s_T is own score in period T. $F(.)$ is the distribution of innovation quality. p'_T is the probability that the other player draws in period T. $ND(p_T = 0)$ is the decision not to draw. $D(p_T = 1)$ is the decision to draw. ξ is threshold determined by equation (4.1).

The equilibrium value of the threshold ξ_i is determined by the equation

$$v\int_{\xi_i}^{\infty}(1 - F^T(\xi_i))\frac{F(x) - F(\xi_i)}{1 - F(\xi_i)}dF(x) - c = 0. \tag{4.1}$$

For example, in our experiment, we assume that when a player exerts effort in a given period the quality of the innovation in that period is a random variable that is distributed according to $F(x;\lambda) = 1 - e^{-\lambda x}$ with $\lambda = 0.125$, which implies that for $T = 10$, the unique subgame perfect equilibrium stopping rule has a threshold of $\xi = 12.16$.

Leaderboard Feedback

In Appendix D.1, we characterize the SPNE in the leaderboard-feedback innovation contest for the case of a general utility function that may allow for risk aversion, loss aversion, and sunk-cost fallacy considerations to be modeled. For simplicity, we focus, here, on the case of risk neutral players. Let f_t (l_t) denote the follower (leader) in an arbitrary period t. We begin by characterizing the final-stage local equilibrium strategies and corresponding equilibrium expected payoffs, and then make our way back through the game tree. In the final period T, if the max score at the beginning of period T is s_T, then we have the following matrix game:

Table 4.1.: The local subgame of period T.

		f_T (follower) D	f_T (follower) ND
l_T (leader)	D	$\frac{v(1+F(s_T)^2)}{2} - c$, $\frac{v(1-F(s_T)^2)}{2} - c$	$v - c$, 0
	ND	$vF(s_T)$, $v(1 - F(s_T)) - c$	v, 0

From Table 4.1, we see that the period T follower's (f_T's) final-stage local expected payoff from choosing to draw (D) when the period T leader (l_T) chooses not to draw (ND) is $v(1 - F(s_T)) - c$. Similarly, f_T's expected payoff from choosing D when l_T chooses D is $vF(s_T)(1 - F(s_T)) + \frac{v(1-F(s_T))^2}{2} - c = \frac{v(1-F(s_T)^2)}{2} - c$. Regardless of l_T's period T action, the payoff to f_T from choosing ND in period T is 0. The expected payoffs for the period T leader (l_T) follow along similar lines.

To calculate the final-stage local equilibrium, let p_{l_T} (p_{f_T}) denote the probability that the period T leader l_T (period T follower f_T) draws in period T. Figure 4.2

presents the players' best-response correspondences as a function of the leader's max score at the beginning of period-T, s_T, and of the probability that the opponent draws in period T and receives a stochastic period-T innovation quality distributed according to $F(\cdot)$.

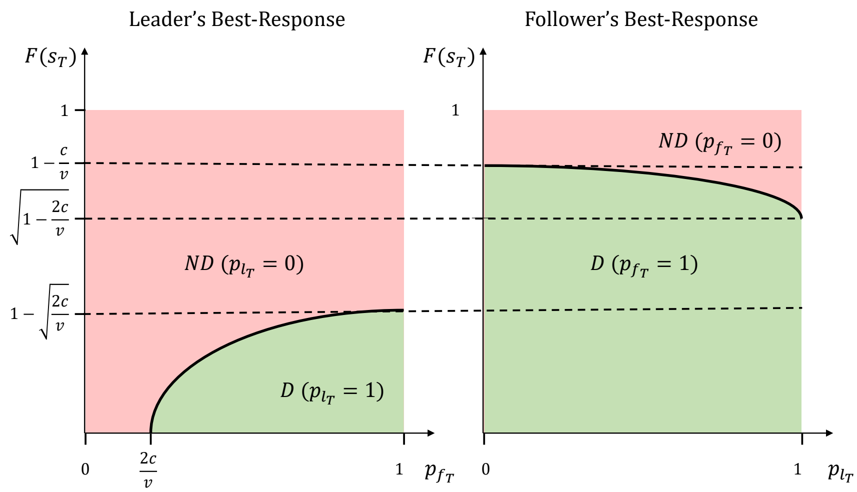


Figure 4.2.: Period T local best responses for Leaderboard Feedback. s_T is the score in period T. $F(.)$ is the distribution of innovation quality. p_{f_T} is the probability that follower draws in period T. p_{l_T} is the probability that the leader draws in period T. $ND(p_{i_T}=0)$ is the decision not to draw by player $i \in \{leader, follower\}$. $D(p_{i_T}=1)$ is the decision to draw by player $i \in \{leader, follower\}$.

Proposition 1 characterizes the final-stage local equilibrium strategies and expected payoffs that follow directly from the best-response correspondences given in Figure 4.2. In particular, if $1-\sqrt{\frac{2c}{v}} \geq F(s_T)$ and $p_{f_t}=1$ then we see from the Leader's Best-Response panel of Figure 4.2 that the leader's best response is $D(p_{l_T}=1)$. Similarly, if $1-\sqrt{\frac{2c}{v}} \geq F(s_T)$ then we see from the Follower's Best-Response panel of Figure 4.2 that for any value of $p_{f_t} \in [0,1]$ the follower's best response is $D(p_{f_T}=1)$. The remaining cases of values of $F(s_T)$ follow along similar lines.

Proposition 4.2.1 *The final-stage local equilibrium strategies are characterized as follows:*

$$\begin{cases} \text{Both draw} & \text{if } 1-\sqrt{\frac{2c}{v}} \geq F(s_T) \\ \text{only follower draws} & \text{if } 1-\frac{c}{v} \geq F(s_T) > 1-\sqrt{\frac{2c}{v}} \\ \text{neither draws} & \text{if } F(s_T) > 1-\frac{c}{v} \end{cases}.$$

The corresponding final-stage local equilibrium expected payoffs for the leader and follower are given in Figure 4.2.

To calculate the subgame perfect equilibrium strategies, we may take the Proposition 1 final-stage local expected payoffs and work back through the game tree to stage $T-1$. The only (computational) issue in continuing the backward induction process all the way to the root of the game in stage 1 is the calculation of the expected continuation payoffs in the period t local subgame. We provide details on these calculations in Appendix D.1.

4.3 Experimental Design

In this section, we describe the experimental design and provide predictions for our experiment using the theory developed above. In particular, the primary goal of the experiment is to address the role of feedback in sequential-search innovation competition. To this end, the main part of our experiment consists of two within-subject treatments: (i) a public feedback treatment and (ii) a leaderboard feedback treatment. In addition to the primary goal, our aim is to better understand factors that may influence individuals to innovate. To this end, our design includes an individual search task that removes the strategic aspect present in the two competitions and the elicitation of individual (e.g., risk aversion) and personality (e.g., grit) characteristics that may be important in an innovation setting. Next, we elaborate on details of the design and our implementation of the experiment.

4.3.1 Private-Feedback and Leaderboard-Feedback Contests

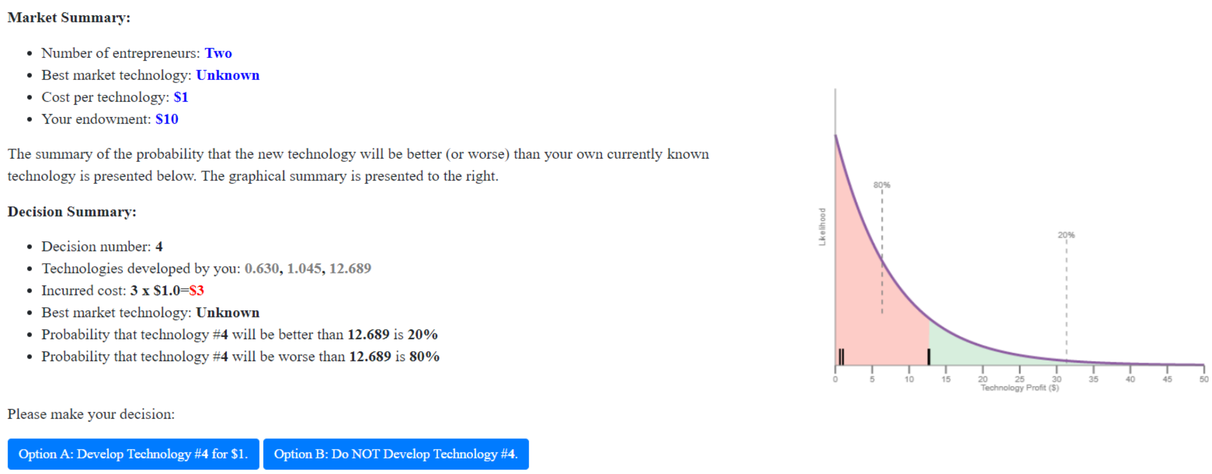
At the beginning of the experiment, each subject individually reads instructions that are displayed on their computer screen. In particular, we implemented a within-subject design, whereby each subject starts the experiment with either eight private-feedback contests or eight leaderboard-feedback contests and then switches to the other feedback type for contests 9 through 16. Thus, before contests 1 and 9, subjects are provided with detailed instructions and practice tasks that explain the setting of the upcoming eight contests. During the practice tasks, subjects were matched with a computer that made decisions randomly, and subjects were informed about the random behavior of the opponent in the practice task. A copy of the instructions used in the experiment and the practice tasks is provided in Appendix D.3.

Each contest consists of two subjects matched for 10 periods of decision-making. Prior to the first period, each subject is given an endowment of \$10.00. Within each period, subjects have the opportunity to pay a cost $c = \$1.00$ to draw an innovation quality from an exponential distribution with parameter $\lambda = 0.125$. At the end of 10 periods, the contest ends and the subject with the highest-quality innovation (the highest score) wins the prize of $v = \$10.00$. Each subject keeps any money left over from her endowment. These parameters were chosen to simplify the environment and were the same for the private and leaderboard treatments as well as for the individual search task described in section 4.3.2.

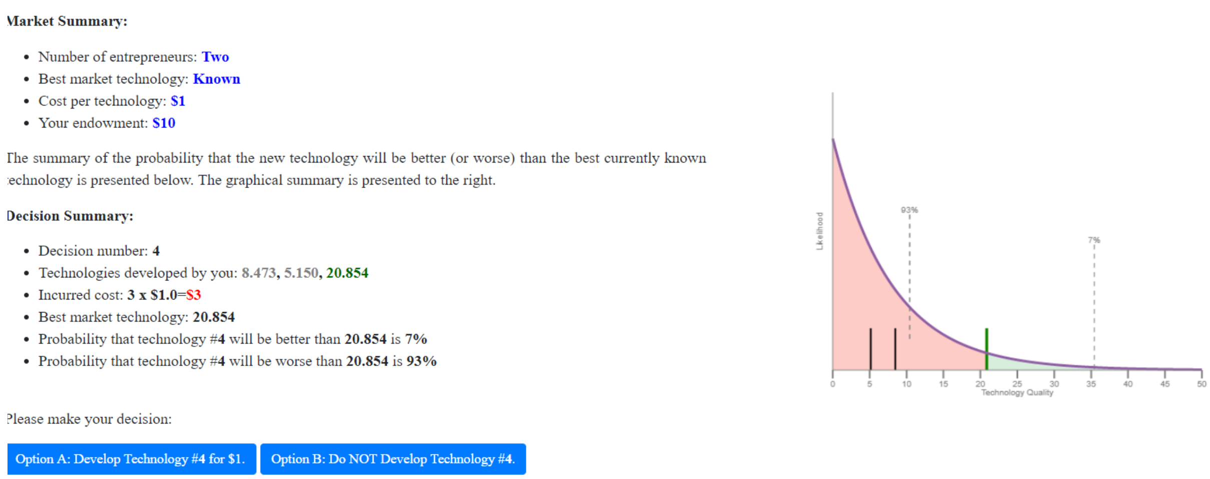
The first treatment is a two-player pravate-feedback contest in which each subject only receives feedback on their own innovations. Specifically, in each period, subjects decide whether to innovate. Although subjects know the quality of their own innovation, they do not know whether they are winning or losing until all decision periods are over. That is, the winning innovation is revealed only at the end of the contest. A screenshot of the private-feedback treatment is presented in Figure 4.3(a). In particular, during each period, each subject has access to the number of times she has drawn, the quality of each of the past innovations she has drawn, and her current innovation score (her innovation with the highest quality). To simplify decision-making, subjects are told the probability that an additional draw will result in a higher individual innovation score. At the end of the contest, subjects are informed of the winner of the contest and the amount of money they have earned for the contest.

The second treatment is a two-player leaderboard-feedback contest in which each subject receives feedback on hew own innovation as well as the innovation that is currently leading the contest. Specifically, similar to the private-feedback contest, in each period of the leaderboard-feedback contest, subjects decide whether to innovate; however, the contest's best innovation is now revealed at the start of each period. Thus, each participant knows whether she is a leader or a follower. A screenshot of the leaderboard-feedback treatment is presented in Figure 4.3(b). Although most aspects of the leaderboard-feedback treatment are the same as in the private-feedback treatment, subjects receive additional feedback regarding the current highest score in the contest. That is, subjects always know whether they are currently winning or losing the contest and the probability that their next draw will result in their score being higher than the current maximum score.[6]

[6]Subjects are no longer shown the probability that an additional draw will result in a higher individual innovation score.



(a) Private Feedback



(b) Leaderboard Feedback

Figure 4.3.: Screenshots of the Experimental Interface

4.3.2 Individual Tasks and Questionnaires

After completing both treatments, subjects were presented with several individual tasks. In particular, subjects completed three elicitation tasks: (i) a risk-aversion task, (ii) a loss-aversion task, and (iii) a sunk-cost-fallacy task. In each of these three tasks, subjects chose one of two options for each of the 20 decisions. The decisions were organized into a multiple price list as is common in the literature (e.g., Holt and Laury, 2002; Rubin et al., 2018). In particular, the first task was the risk-aversion task. In this task each participant chose between a risky option (50% chance of \$10.00 and a 50% percent chance of \$0.00) and a safe option that was varied across decisions (started at \$0.50 and increased by \$0.50 in each subsequent decision). The second task was the loss-aversion task. In this task each participant chose between a safe option of \$0.00 and a risky option had a 50% chance at \$0.00 and a 50% chance of a loss (varied from $-\$0.50$ to $-\$10.00$ in increments of \$0.50). The third elicitation task was the sunk-cost-fallacy task. In this task, subjects were given an endowment of \$15.00 and were required to pay \$5.00 to initiate a project. Each subject then decided whether to complete the project at various completion costs. Completing the project was always worth \$7.50; however, the cost varied between decisions. The completion cost started at \$0.50 and increased by \$0.50 in each subsequent decision. The sunk-cost fallacy occurs if the subject completes the project at a cost greater than \$7.50. Screenshots of the three individual elicitation tasks are presented in Figures D.5-D.7 in the Appendix.

In addition to the above elicitation tasks, each subject participated in eight individual search tasks. The individual search tasks were similar to the two contests except that the human opponent was replaced with an existing innovation of a known quality. In particular, the existing innovation took on five values: 15.177, 16.832, 18.421, 20.205, and 23.966.[7] Each subject saw all five values and the values 15.177, 18.421, and 23.966 were repeated twice. The five values were displayed in random order. If the subject ends the period with an innovation of greater quality than the existing innovation, she won \$10.00. Thus, these tasks allow us to analyze individual behavior in a similar environment but without competition against another human subject. A screenshot of the individual search task is presented in Figure D.8 in the Appendix.

[7]These values correspond to 85, 88, 90, 92, and 95 percentiles of the exponential distribution, respectively. In particular, the risk-neutral agent would be indifferent between drawing and not drawing if the existing innovation was 18.421.

The experiment concluded with three unincentivized personality questionnaires. In particular, the first questionnaire measured the psychological construct of grit through the 12-item Grit Scale (Duckworth et al., 2007). The second questionnaire measured the big five characteristics (agreeableness, extraversion, neuroticism, openness, and conscientiousness) through the 44-item big-five inventory (John and Srivastava, 1999). The third questionnaire measured achievement-striving and competitiveness through the 10- and 6-item scales obtained from the International Personality Item Pool.[8]

4.3.3 Experimental Administration

All parts of the experiment, including instructions, innovation contests, individual elicitation tasks, and personality questionnaires, were implemented in oTree (Chen et al., 2016). In total, subjects participated in 27 compensation-relevant tasks. Specifically, the compensation-relevant tasks included the eight private-feedback contests, the eight leaderboard-feedback contests, the risk-aversion elicitation task, the loss-aversion elicitation task, the sunk-cost-elicitation task, and the eight individual search tasks. At the end of the experiment, two of these 27 tasks were chosen at random by the computer for payment.

In total, 96 students were recruited on the campus of Purdue University using ORSEE software (Greiner, 2015). Participants were split into 12 sessions, with eight participants per session. As mentioned above, to ensure that the order of treatments did not affect the main results, half of the sessions started out with eight private-feedback contests, while the other half of the sessions started out with eight leaderboard-feedback contests. The experimental lasted under 60 minutes, with average earnings of $19.91.

[8] https://ipip.ori.org/

4.4 Predictions

In this section, we present predictions for the experiment that were obtained by computationally solving for the sequential equilibrium described in section 4.2. In particular, using the model, 1 million contests were simulated and the resulting predictions were organized into four hypotheses: the first hypothesis pertains to the comparison of the private- and leaderboard-feedback contests; the second hypothesis pertains to the comparison of leader and follower behavior; the third hypothesis pertains to the dynamics of the draws in the two contests; and the fourth hypothesis pertains to the role of individual characteristics such as risk aversion, loss aversion, and the sunk-cost fallacy.[9]

Table 4.2.: Displays the summary of predictions. Aggregate draws refers to the predicted number of draws that occurs in a contest in each treatment. Winning innovation refers to the predicted quality of the winning innovation in each treatment. Known score refers to the individual score in the private-feedback treatment and the maximum score in the leaderboard-feedback treatment. The third row displays the draw rate of the leader and the follower in periods 2, 6, and 10 of the experiment. The fourth row displays the draw rate in periods 2, 6, and 10 of the experiment for known scores in the 20th-80th percentiles for that period. The fifth row displays the difference in draw rates for known scores in the lower half and the upper half of the known score distribution for periods 2, 6, and 10.

	Private Feedback	Leaderboard Feedback
Winning Innovation	23.42	21.84
Aggregate Draws	8.36	6.34
Proportion of Draws		
Leader		
Known Score 0–15	0.67/0.30/0.03	0.59/0.04/0.00
Known Score 15–25	0.11/0.02/0.00	0.00/0.00/0.00
Follower		
Known Score 0–15	0.90/0.62/0.37	0.59/0.55/1.00
Known Score 15–25	0.58/0.19/0.08	0.14/0.38/0.32

The top part of Table 4.2 shows that a contest with private feedback is predicted to induce more draws (8.36) and result in a greater winning innovation score (23.42)

[9]One million contests were simulated for each value of each bias parameter.

than a contest with leaderboard feedback (6.34 draws; winning innovation of 21.84). We summarize this prediction with Hypothesis 1.

Hypothesis 1 *The private-feedback contest leads to more draws and a higher winning innovation than the leaderboard-feedback contest.*

The bottom part of Table 4.2 presents the proportion of draws broken down by the period of the contest (presented as a triple of 2nd/6th/10th period), the current score (grouped into ranges 0–15 and 15–25), and whether the player was a leader or a follower.[10] By comparing the proportion of draws between leaders and followers, the follower is clearly predicted to be at least as likely to draw as the leader across most of the ranges of innovation scores and periods.[11] We summarize this prediction with Hypothesis 2.

Hypothesis 2 *Followers draw more frequently than leaders.*

The bottom part of Table 4.2 also provides an insight regarding the dynamics of decision-making. In the private-feedback treatment, as the individual innovation score increases, each player becomes less willing to draw. This can be seen by comparing the proportion of draws between relatively low individual scores (0–15) and relatively high individual scores (15–25) for both leaders and followers. Additionally, in the leaderboard-feedback treatment, as the maximum score increases, each player becomes less willing to draw. This can be seen by comparing the proportion of draws between relatively low maximum scores (0–15) and relatively high maximum scores (15–25) for both leaders and followers. We summarize this prediction with Hypothesis 3.

Hypothesis 3 *Players become less willing to draw as their individual score increases in the private-feedback treatment and as the maximum score increases in the leaderboard-feedback treatment.*

[10] Figures D.10 and D.11 in Appendix D.4 present further evidence on the proportion of draws obtained from our computational model using simulations.

[11] Overall, leaders draw 8.73% of the time in the simulated contests and followers draw 39.20% of the time in the simulated contests.

Lastly, we incorporate three behavioral characteristics: risk aversion, loss aversion, and the sunk-cost fallacy.[12] The three panels of Figure 4.4 present the comparative statics as we vary these characteristics one at a time. For example, to vary risk aversion, we model both players as having a CRRA utility function with parameter γ, and we vary this parameter across a range of values typically observed in the experimental literature.

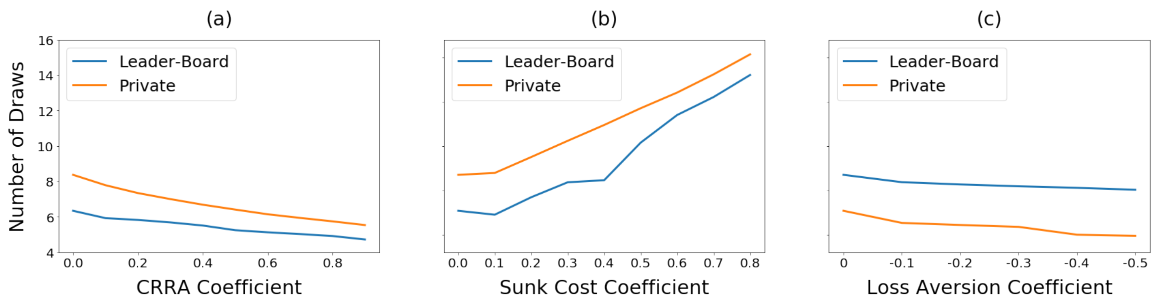


Figure 4.4.: Displays the decision to draw and the comparative statics. This figure displays equilibrium predictions under different levels of (a) risk aversion, (b) sunk cost fallacy, and (c) loss aversion. The orange line is the private-feedback treatment, while the blue line is the leaderboard-feedback treatment.

Figure 4.4 shows that as risk aversion and loss aversion increase, the number of total draws made in the contest decreases. The sunk-cost fallacy, however, has an opposite effect. In particular, as the sunk-cost fallacy increases, we observe more total draws. We summarize these predictions with Hypothesis 4.

Hypothesis 4 *The number of draws increases with (a) a decrease in risk aversion, (b) a decrease in loss aversion, (c) an increase in sunk-cost fallacy.*

[12]Specifications of the three utility functions as well as the general procedure for obtaining predictions are provided in Appendix D.2.

4.5 Results

In this section, we present the results of our experiment. In particular, first, in section 4.5.1 we compare the outcomes of the private and leaderboard treatments. Next, in section 4.5.2, we test for differences in behavior between the leader and the follower. Then, in section 4.5.3 we consider the dynamics observed in the experimental data. Finally, in section 4.5.4, we discuss the role of individual characteristics in determining innovation-contest outcomes.

4.5.1 Private vs Leaderboard Contests

The columns of Table 4.3 display the summary statistics from the two treatments. In particular, the table is divided down into two parts. In the top part, we present the aggregate results on the final innovation quality and the total number of draws that we observed in each of the treatments, on average. In the bottom part, we present the results on the proportion of draws conditional on the period in the game (periods 2, 6, and 10 are separated by "/"), current score (we group scores into two ranges 0–15 and 15–25), and whether the decision-maker was a leader or a follower.[13]

[13] Recall that while the role of leader/follower is known to the decision-making in the leaderboard treatment, it is not known to the decision-makers in the private feedback.

Table 4.3.: Displays the results of the contests. Aggregate draws refers to the predicted number of draws that occurs in a contest in each treatment. Winning innovation refers to the predicted quality of the winning innovation in each treatment. The third row displays the draw rate of the leader and the follower in periods 2, 6, and 10 of the experiment. The fourth row displays the draw rate in periods 2, 6, and 10 of the experiment for scores that range in the 20th-80th percentile for that period. The fifth row displays the difference in draw rates for scores in the lower half and the upper half of the score distribution for periods 2, 6, and 10.

	Private Feedback	Leaderboard Feedback
Winning Innovation	22.87	21.47
Aggregate Draws	8.50	7.54
Proportion of Draws		
Leader		
Known Score 0–15	0.59/0.60/0.33	0.37/0.36/0.20
Known Score 15–25	0.16/0.16/0.11	0.08/0.08/0.07
Follower		
Known Score 0–15	0.61/0.64/0.40	0.60/0.59/0.63
Known Score 15–25	0.45/0.41/0.38	0.49/0.50/0.49

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

The top part of Table 4.3 shows the average number of contest draws and the average value of the winning innovation in each treatment. In particular, in the private-feedback treatment, the average number of draws (8.50) and the average value of the winning innovation (22.87) are not significantly different from the theoretically predicted values (8.36 draws, p-value 0.67; score of 23.42, p-value 0.36).[14] In terms of the leaderboard feedback, we also find no difference in the value of the winning innovation between theory and the experiment (21.84 vs. 21.47, p-value 0.42). However, we do find a difference between theory and the experiment in terms of the number of draws for the leaderboard-feedback treatment (6.34 vs. 7.54, p-value 0.000).[15]

The main focus of the aggregate results is on the comparison between private and leaderboard feedback (i.e., Hypothesis 1). Table 4.3 shows that in our experiment, the number of draws in the private-feedback contest (8.50) is greater than in the

[14]Hypothesis tests in this subsection are conducted using bootstrapped regressions, with 5,000 bootstrap samples, on the session-level averages.

[15]This difference in significance between draws and winning scores for the leader-board feedback treatment may be a product of the realization of the random draws.

leaderboard-feedback contest (7.54). We test whether this difference is significant using a random-effects regression with session-level effects. We find that this difference is significant (p-value=0.000). Similarly, Table 4.3 shows that the winning technology is greater in a private-feedback contest (22.87) than a leaderboard-feedback contest (21.47). Again, using a random-effects regression with session-level effects, we find that this difference is significant (p-value=0.029). We summarize these tests with Result 1.

Result 1 *A private-feedback contest results in more draws and a greater winning innovation value than a leaderboard-feedback contest (evidence supporting Hypothesis 1).*

4.5.2 Leaders vs. Followers

The bottom part of Table 4.3 shows that the proportion of time that a follower draws is greater than the proportion of time that a leader draws. While the difference is observed in both the private and leaderboard treatments, the difference is much larger in the latter. Figure 4.5 presents further evidence regarding this comparison. Formally, each panel of the figure shows a panel data logistic regression of the decision to draw on the maximum score. The bottom row of the figure presents the comparison of of the leader's decision (blue) and the follower's decision (red). The figure clearly shows that in almost every combination of period and maximum score, followers are more likely to draw than leaders. Thus, Figure 4.5 suggests that Hypothesis 2 holds.[16]

[16]Figure B5 provide similar figures for the remaining periods.

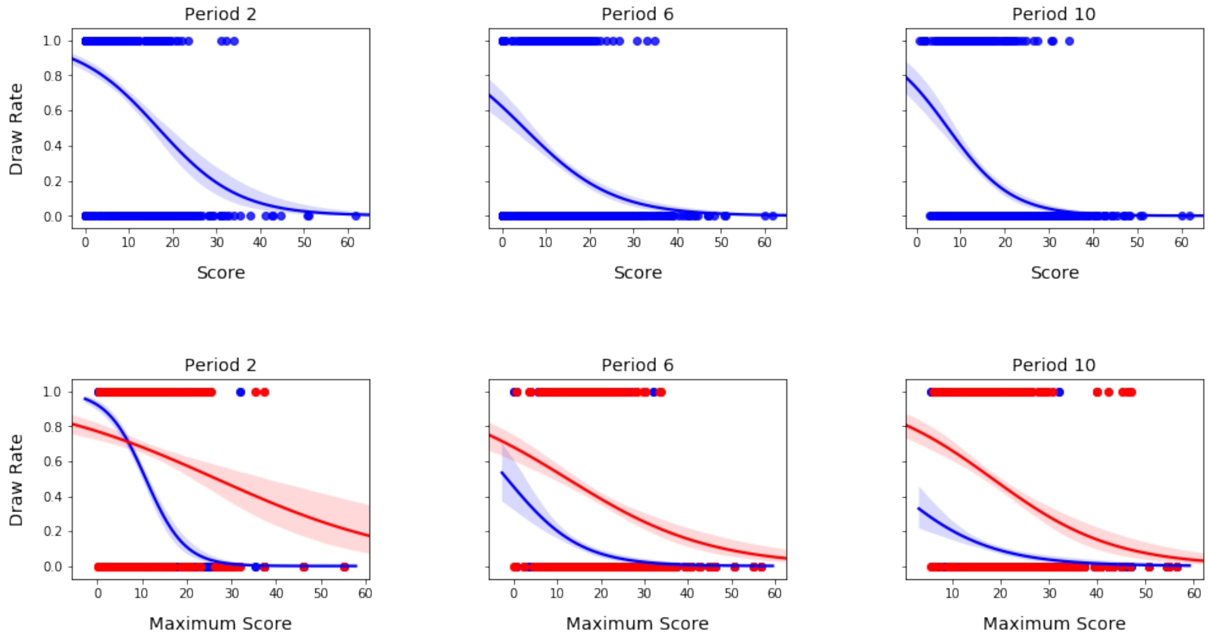


Figure 4.5.: Displays the decision to draw in the Leaderboard-Feedback Treatment. This figure displays two sets of graphs. The first set of graphs display logistic regressions of the decision to draw in the pravate-feedback treatment for periods 2, 6, and 10. The second set of graphs display logistic regressions of the leader's decision (blue) to draw and the follower's decision (red) to draw in the leaderboard-feedback treatment for periods 2, 6, and 10.

To formally test the difference between leader and follower behavior, we use a panel data logistic regression. In particular, we regress the decision to draw on an indicator variable for whether the subject was a leader, while accounting for subject-level random effects and clustering standard errors at the session level.[17] The coefficient on the leader variable is negative and significant at the 1% level. We summarize these observations with Result 2.

Result 2 *Leaders draw less frequently than followers in the leaderboard-feedback treatment (evidence supporting Hypothesis 2).*

[17]Note that the regression is run on the observations where the score is greater than zero (and thus there is a leader and a follower).

4.5.3 Dynamics of Decision Making

Figure 4.5 suggests that subjects are less willing to draw as the individual score increases in the private-feedback treatment and as the maximum score increases in the leaderboard-feedback treatment. To formally test Hypothesis 3, we run panel data logisitic regressions, with subject-level random effects and session-level clustered standard errors, of the decision to draw on the individual score. We run these regressions for the last nine periods of the private-feedback treatment. We find that in each of the regressions, the coefficient on the individual score is negative and significant at the 1% level. Additionally, we run similar regressions for the leaderboard-feedback treatment with the difference being that the decision to draw is regressed on the maximum score. Again, for each of the regressions, the coefficient on the maximum score is negative and significant at the 1% level. We summarize these results with Result 3.

Result 3 *Subjects are less willing to draw as their individual score increases in the private-feedback treatment and as the maximum score increases in the leaderboard-feedback treatment (evidence supporting Hypothesis 3).*

4.5.4 Role of Individual Characteristics

In our experiment, subjects completed various elicitation tasks. We used these tasks to shed light on factors that may influence subjects' decision to draw. Table 4.4 displays three sets of regressions that analyze the decision to draw on the elicited characteristics.[18] In particular, the regressions are carried out using a panel data logistic regression with subject-level random effects, and standard errors are obtained by clustering at the session level.

Table 4.4 shows that the regression analyses yield results consistent with our prior analysis in terms of the role of the treatments and leader/follower behavior. In terms

[18] We relegate regressions on the individual search task to the appendix as the results are similar to the results found in Table 4.4.

of elicited individual characteristics, we find that risk aversion has a significantly negative effect across a number of specifications. At the same time, we find that our measures of loss aversion and sunk-cost fallacy are not significant in any of the specifications. We summarize these results with Result 4.

Result 4 *Risk aversion leads to a lower likelihood of drawing an innovation (evidence supporting Hypothesis 4a).*

Recall that in addition to the incentivized elicitation of risk aversion, loss aversion, and the sunk-cost fallacy, we conducted a number of non-incentivized personality questionnaires that addressed personality characteristics. In particular, in addition to a broad questionnaire (i.e., Big 5), we selected a few characteristics as potentially important to behavior in an innovation-contest setting (i.e., Grit and Competitiveness). Table 4.4 shows that virtually no personality characteristics are significant in explaining drawing behavior for any of the regression specifications.[19]

[19]Table D.1 in the Appendix provides an alternative specification of this regression in which we first carry out factor analysis to identify orthogonal factors present in the questionnaire. The regression results stay largely the same. Additionally, the reader may be concerned that there is endogeneity of the individual score in the private feedback treatment and endogeneity of the maximimum score in the leader-board feedback treatment. We instrument for the former variable by creating a variable for each individual's score immediately after the first period. We instrument for the latter variable by creating a variable for the maximum score immediately after the first period. In a pooled regression, with these two instruments, we find similar results as before. Risk aversion is the only significant behavioral factor.

Table 4.4.: Displays the results of the regressions. The regression pools the data from the individual search tasks, the private-feedback treatment, and the leaderboard-feedback treatment.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. Var.:	Pooled		Private			Leaderboard	
Draw Decision		All	Leader	Follower	All	Leader	Follower
L-Board	-0.70***	—	—	—	—	—	—
	(0.20)	—	—	—	—	—	—
Priv. x Score	-0.17***	-0.21***	-0.25***	-0.18***	—	—	—
	(0.01)	(0.02)	(0.04)	(0.02)	—	—	—
L-Board x MaxScore	-0.11***	—	—	—	-0.11***	-0.23***	-0.11***
	(0.01)	—	—	—	(0.01)	(0.02)	(0.01)
Period	-0.12***	-0.13***	-0.19***	-0.11***	-0.10***	-0.24***	-0.03
	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.05)
Risk Aversion	-1.13**	-1.41**	-1.50	-1.20**	-1.05**	-1.01	-0.31
	(0.50)	(0.72)	(1.32)	(0.56)	(0.46)	(0.87)	(1.15)
Loss Aversion	-0.22	-0.10	1.15	-0.83	-0.30	-1.12	-0.43
	(0.65)	(0.83)	(1.01)	(0.70)	(0.63)	(1.09)	(0.89)
Sunk Cost Fallacy	0.06	0.14	-1.07	0.25	-0.12	-0.55	0.02
	(0.61)	(0.94)	(0.87)	(0.96)	(0.45)	(0.87)	(0.94)
Grit	-0.15	-0.28	-0.53	-0.05	-0.02	-0.17	-0.20
	(0.24)	(0.39)	(0.47)	(0.33)	(0.16)	(0.40)	(0.47)
Competitiveness	-0.18	0.12	0.00	0.42	-0.43	-0.07	-0.27
	(0.31)	(0.43)	(0.47)	(0.42)	(0.28)	(0.38)	(0.36)
Achievement Striving	0.38	0.18	0.06	-0.18	0.57	0.66	0.08
	(0.39)	(0.54)	(0.68)	(0.49)	(0.36)	(0.46)	(0.70)
Extraversion	0.04	-0.03	0.09	-0.07	0.09	-0.21	0.13
	(0.10)	(0.13)	(0.17)	(0.11)	(0.11)	(0.23)	(0.13)
Agreeableness	0.19	0.09	0.03	0.00	0.27	0.20	0.26
	(0.22)	(0.28)	(0.35)	(0.29)	(0.22)	(0.32)	(0.33)
Neuroticism	0.06	0.07	-0.14	0.12	0.04	0.13	-0.08
	(0.13)	(0.17)	(0.22)	(0.15)	(0.13)	(0.22)	0.26
Openness	-0.18	-0.18	-0.27	-0.25	-0.23	-0.46	-0.23
	(0.17)	(0.26)	(0.33)	(0.25)	(0.15)	(0.33)	(0.26)
Conscientiousness	0.04	0.30	0.43	0.10	-0.24	-0.34	-0.41
	(0.28)	(0.49)	(0.46)	(0.45)	(0.17)	(0.60)	(0.34)
Constant	0.83	0.91	5.38**	1.03	0.58	2.62**	4.24*
	(1.44)	(2.00)	(2.46)	(1.95)	(1.26)	(1.13)	(2.46)
Observations	15,360	7,680	3,451	3,451	7,680	3,411	3,411

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

4.6 Conclusion

In this paper, we investigate the role of leaderboard feedback in sequential-search innovation competition. In particular, our contribution is threefold. First, we contribute to the existing theoretical literature by developing a model of dynamic scoring contests with a finite horizon and perfect recall. Our work is the first (to our knowledge) to formally provide an equilibrium prediction for the environment with leaderboard feedback. Specifically, we show that leaderboard feedback may result in lower effort as captured by the number of costly innovation decisions, which in turn yields worse innovation quality of the innovation competition than providing private feedback.

Second, we contribute to the experimental literature that investigates contest and innovation competitions. Our experiment yields several results that support theory. Specifically, we find that for a two-player finite-horizon contest, leaderboard feedback yields less effort and lower innovation quality than private feedback. We also find that the internal dynamics present in the data are consistent with the model. In particular, when feedback is provided, leaders of the contest reduce their effort, whereas followers do not. In addition, as the quality of innovation increases, agents become less likely to invest resources to generate a new innovation.

Finally, our work also contributes to a stream of literature that studies the role of individual characteristics in determining an individual's propensity to innovate. In particular, we elicit three individual characteristics that have been shown to be important in the innovation and contest setting: risk aversion, loss aversion, and the sunk-cost fallacy. We find that among these individual characteristics, risk aversion stands out as being an important driver of behavior in our experiment. At the same time, loss aversion and sunk cost fallacy are not significant in explaining the data. In addition, we find no evidence that personality characteristics are predictive of behavior in the dynamic contests studied in this paper.

Our work has several shortcomings that open interesting avenues for future research. First, our theoretical model and laboratory experiment investigate a finite-horizon innovation competition. Comparing it to the an infinite-horizon setting would be interesting. Second, we considered a two-player contest, the extent to which these results translate to a setting with more than two players is not known. Finally, subjects in our experiment participated in the contest (although they had an option not to draw). Investigating the extent to which our results hold if subjects could select to withdraw from the contests entirely would be interesting.

BIBLIOGRAPHY

Ales, L., Cho, S.-H., and Körpeoğlu, E. (2017). Optimal award scheme in innovation tournaments. *Operations Research*, 65:693–702.

Anderson, C. (2001). Behavioral models of strategies in multi-armed bandit problems. *PhD Thesis*.

Anderson, C. (2012). Ambiguity aversion in multi-armed bandit problems. *Theory and Decision*, 72:15–33.

Anderson, S., Friedman, D., and Oprea, R. (2010). Preemption games: Theory and experiment. *American Economic Review*, 100:1778–1803.

Aoyagi, M. (2010). Information feedback in a dynamic tournament. *Games and Economic Behavior*, 70:242–260.

Astebro, T., Herz, H., Nanda, R., and Weber, R. A. (2014). Seeking the roots of entrepreneurship: Insights from behavioral economics. *Journal of Economic Perspectives*, 28:49–70.

Augenblick, N. (2015). The sunk-cost fallacy in penny auctions. *The Review of Economic Studies*, 83:58–86.

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

Baye, M. R. and Hoppe, H. C. (2003). The strategic equivalence of rent-seeking, innovation, and patent-race games. *Games and economic behavior*, 44:217–226.

Bennett, D., Sasmita, K., Maloney, R., Murawsk, C., and Bode, S. (2019). Monetary feedback modulates performance and electrophysiological indices of belief updating in reward learning. *Psychophysiology*, 56.

Bimpikis, K., Ehsani, S., and Mostagir, M. (2019). Designing dynamic contests. *Operations Research*, 67:339–356.

Bolton, P. and Harris, C. (1999). Strategic experimentation. *Econometrica*, 67:349–374.

Bonatti, A. and Hörner, J. (2011). Collaborating. *American Economic Review*, 101:632–663.

Camacho, N., Donkers, B., and Stremersch, S. (2011). Predictably non-bayesian: Quantifying salience effects in physician learning about drug quality. *Marketing Science*, 30:305–320.

Cason, T. and Mui, V. (2003). Testing political economy models of reform in the laboratory. *American Economic Review Papers and Proceedings*, 93:208–212.

Cason, T. and Mui, V. (2005). Uncertainty and resistance to reform in laboratory participation games. *European Journal of Political Economy*, 21:708–737.

Chawla, S., Hartline, J. D., and Sivan, B. (2015). Optimal crowdsourcing contests. *Games and Economic Behavior*.

Che, Y.-K. and Gale, I. (2003). Optimal design of research contests. *American Economic Review*, 93:646–671.

Chen, D. L., Schonger, M., and Wickens, C. (2016). otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Choi, J. (1991). Dynamic r&d competition under 'hazard rate' uncertainty. *RAND Journal of Economics*, 22:596–610.

Chowdhury, C. (2017). The all-pay auction with nonmonotonic payoff. *Economic Journal*, 84:375–390.

Dechenaux, E., Kovenock, D., and Sheremeta, R. M. (2015). A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics*, 18:609–669.

Deck, C. and Kimbrough, E. O. (2017). Experimenting with contests for experimentation. *Southern Economic Journal*, 84:391–406.

DiPalantino, D. and Vojnovic, M. (2009). Crowdsourcing and all-pay auctions. In *Proceedings of the 10th ACM conference on Electronic commerce*, pages 119–128. ACM.

Duckworth, A. and Quinn, P. (2009). Development and validation of the short grit scale (grits). *Journal of Personality Assessment*, 91:166–174.

Duckworth, A. L., Peterson, C., Matthews, M. D., and Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of personality and social psychology*, 92:1087.

Eckel, C. and Grossman, P. (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior*, 23:281—-295.

Ederer, F. and Manso, G. (2013). Is pay for performance detrimental to innovation? *Management Science*, 59:1496–1513.

Erat, S. and Krishnan, V. (2012). Managing delegated search over design spaces. *Management Science*, 58:606–623.

Fernandez, R. and Rodrick, D. (1991). Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *American Economic Review*, pages 1146–1155.

Fershtman, C. and Gneezy, U. (2011). The tradeoff between performance and quitting in high power tournaments. *Journal of the European Economic Association*, 9:318–336.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10:171–178.

Freer, M., Martinelli, C., and Wang, S. (2018). Collective experimentation: A laboratory study. *Working Paper*.

Fullerton, R. L. and McAfee, R. P. (1999). Auctionin entry into tournaments. *Journal of Political Economy*, 107:573–605.

Gans, N., Knox, G., and Croson, R. (2007). Simple models of discrete choice and their performance in bandit experiments. *Operations Management*, 9:383–408.

Gershkov, A. and Perry, M. (2009). Tournaments with midterm reviews. *Games and Economic Behavior*, 66:162–190.

Gneezy, U. and Smorodinsky, R. (2006). All-pay auctions – an experimental study. *Journal of Economic Behavior and Organization*, 61:255–275.

Goeree, J., Palfrey, T., Rogers, B., and McKelvey, R. (2007). Self-correcting information cascades. *The Review of Economic Studies*, 74:733–762.

Goltsman, M. and Mukherjee, A. (2011). Interim performance feedback in multistage tournaments: The optimality of partial disclosure. *Journal of Labor Economics*, 29:229–265.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with orsee. *Journal of the Economic Science Association*, 1:114–125.

Halac, M., Kartik, N., and Liu, Q. (2016). Optimal contracts for experimentation. *Review of Economic Studies*, 83:1040–1091.

Halac, M., Kartik, N., and Liu, Q. (2017). Contests for experimentation. *Journal of Political Economy*, 125:1523–1569.

Harbring, C. and Irlenbusch, B. (2003). An experimental study on tournament design. *Labour Economics*, 10:443–464.

Herz, H., Schunk, D., and Zehnder, C. (2014). How do judgmental overconfidence and overoptimism shape innovative activity? *Games and Economic Behavior*, 83:1–23.

Hinnosaar, T. (2016). Penny auctions. *International Journal of Industrial Organization*, 48:59–87.

Hoelzemann, J. and Klein, N. (2018). Bandits in the lab. *Working Paper*.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92:1644–1655.

Hudja, S. (2019). Voting for experimentation: A continuous time analysis. *Working Paper*.

John, O. P. and Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, 2:102–138.

Keller, G., Novák, V., and Willems, T. (2019). A note on optimal experimentation under risk aversion. *Journal of Economic Theory*, 179:476–487.

Keller, G., Rady, S., and Cripps, M. (2005). Strategic experimentation with exponential bandits. *Econometrica*, 73:39–69.

Khromenkova, D. (2015). Collective experimentation with breakdowns and breakthroughs. *Working Paper*.

Klein, A. H. and Schmutzler, A. (2017). Optimal effort incentives in dynamic tournaments. *Games and Economic Behavior*, 103:199–224.

Koudstaal, M., Sloof, R., and Van Praag, M. (2015). Risk, uncertainty, and entrepreneurship: Evidence from a lab-in-the-field experiment. *Management Science*, 62:2897–2915.

Kuhnen, C. M. and Tymula, A. (2012). Feedback, self-esteem, and performance in organizations. *Management Science*, 58:94–113.

Lang, M., Seel, C., and Strack, P. (2014). Deadlines in stochastic contests. *Journal of Mathematical Economics*, 52:134–142.

Lim, W., Matros, A., and Turocy, T. (2014). Bounded rationality and group size in tullock contests: Experimental evidence. *Journal of Economic Behavior and Organization*, 99:155–167.

List, J., van Soest, D., Stoop, J., and Zhou, H. (2010). On the role of group size in tournaments: Theory and evidence from lab and field experiments. *Working Paper*.

Ludwig, S. and Lünser, G. K. (2012). Observing your competitor–the role of effort information in two-stage tournaments. *Journal of Economic Psychology*, 33:166–182.

Meyer, R. and Shi, Y. (1995). Sequential choice under ambiguity: Intuitive solutions to the armed-bandit problem. *Management Science*, 41:817–834.

Mihm, J. and Schlapp, J. (2018). Sourcing innovation: On feedback in contests. *Management Science*, 65:559–576.

Moffatt, P. G. (2016). *Experimetrics*. Palgrave.

Moreno, O. and Rosokha, Y. (2016). Learning under compound risk vs. learning under ambiguity - an experiment. *Journal of Risk and Uncertainty*, 53:137–162.

Morgan, J., Orzen, H., and Sefton, M. (2012). Endogenous entry in contests. *Economic Theory*, 51:435–463.

O'Neill, B. (1986). International escalation and the dollar auction. *Journal of Conflict Resolution*, 30:33–50.

Oprea, R. (2014). Survival versus profit maximization in a dynamic stochastic experiment. *Econometrica*, 82:2225–2255.

Oprea, R., Friedman, D., and Anderson, S. (2009). Learning to wait: A laboratory investigation. *The Review of Economic Studies*, 76:1103–1124.

Orrison, A., A., S., and K., W. (2004). Multiperson tournaments: An experimental examination. *Management Science*, 50:268–279.

Paetzel, F., Sausgruber, R., and Traub, S. (2014). Social preferences and voting on reform: An experimental study. *European Economic Review*, 70:36–55.

Robalo, P. and Sayag, R. (2018). Paying is believing: The effect of costly information on bayesian updating. *Journal of Economic Behavior and Organization*, 156:114–125.

Rosokha, Y. and Younge, K. (2017). Motivating innovation: The effect of loss aversion on the willingness to persist. *Review of Economics and Statistics*, pages 1–45.

Rubin, J., Samek, A., and Sheremeta, R. M. (2018). Loss aversion and the quantity–quality tradeoff. *Experimental Economics*, 21:292–315.

Seel, C. and Strack, P. (2016). Continuous time contests with private information. *Mathematics of Operations Research*, 41:1093–1107.

Sheremeta, R. (2011). Contest design: An experimental investigation. *Economic Inquiry*, pages 573–590.

Shubik, M. (1971). The dollar auction game: A paradox in noncooperative behavior and escalation. *Journal of conflict Resolution*, 15:109–111.

Strulovici, B. (2010). Learning while voting: Determinants of collective experimentation. *Econometrica*, 78:933–971.

Taylor, C. R. (1995). Digging for golden carrots: An analysis of research tournaments. *The American Economic Review*, pages 872–890.

Terwiesch, C. and Xu, Y. (2008). Innovation contests, open innovation, and multiagent problem solving. *Management science*, 54:1529–1543.

Therneau, T., Crowson, C., and Atkinson, E. (2018). Using time dependent covariates and time dependent coefficients in the cox model. https://cran.r-project.org/web/packages/survival/vignettes/timedep.pdf.

Yildirim, H. (2005). Contests with multiple rounds. *Games and Economic Behavior*, 51:213–227.

A. APPENDIX FOR: BEHAVIORAL BANDITS: ANALYZING THE EXPLORATION VERSUS EXPLOITATION TRADE-OFF IN THE LAB

A.1 Theory Appendix

This section of the appendix focuses on how predictions are obtained. The first subsection shows how the continuous time predictions are obtained. The second subsection shows the discrete time approximation predictions are obtained. The third subsection displays figures that are related to theory such as payoff hills and behavioral predictions for the non-Baseline treatments.

A.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

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

A.1.2 Discrete Time Predictions

The discrete time predictions are found by using 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. This value function iteration process is similarly used for sections 1.3.4 and 1.5.

A.1.3 Additional Figures

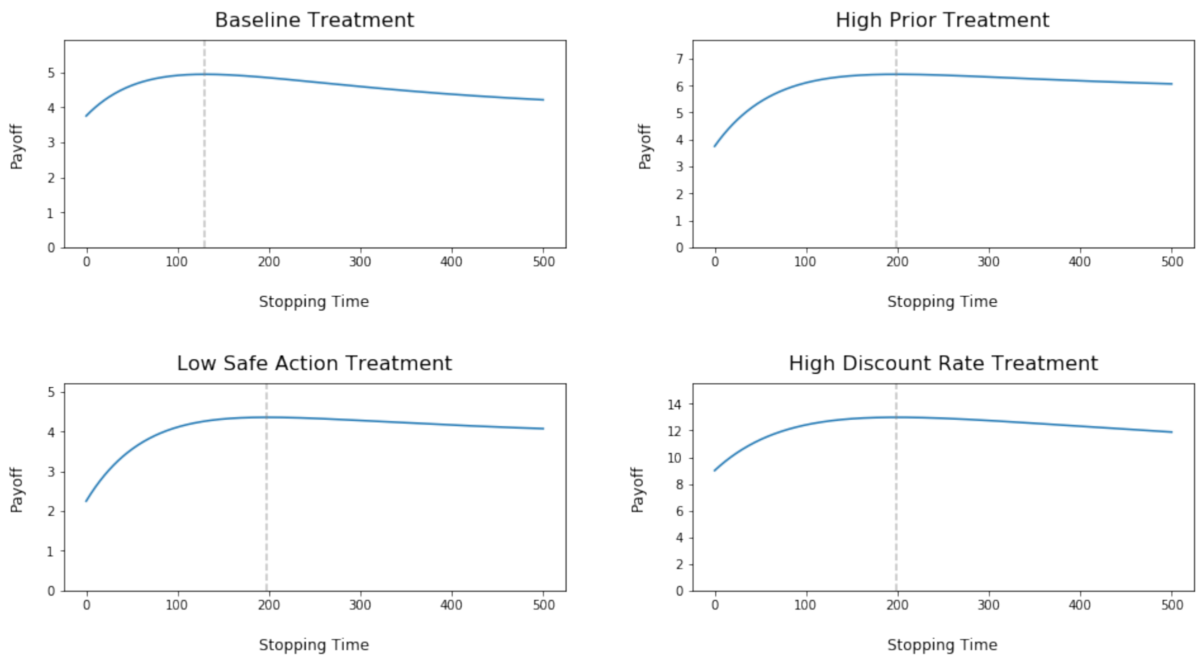


Figure A.1.: Payoff hills for each treatment of the experiment. Payoffs are shown for stopping times between 0 and 500. A grey dotted line is drawn at the predicted stopping time.

Figure A.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.

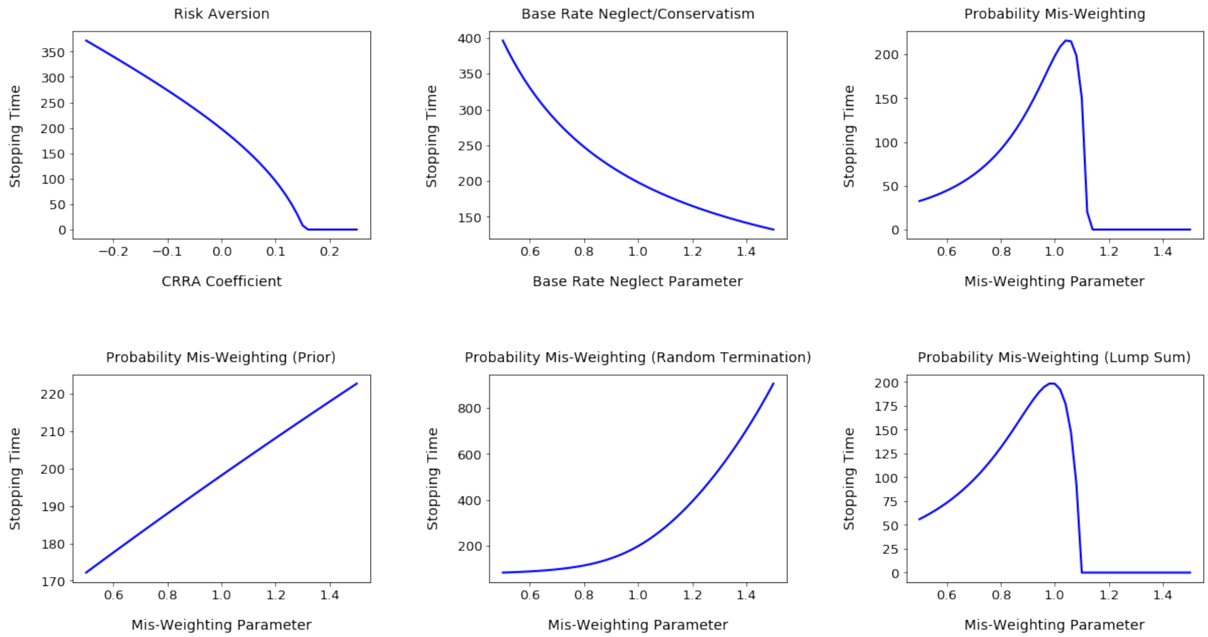


Figure A.2.: Predictions for the High Prior Treatment as different behavioral factors are unilaterally varied.

Figure A.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 1.2 except that the length of time that an individual is willing to experiment is now increasing in the mis-weighted prior as α increases. This occurs because the Prelec-I function is centered around 0.368 and thus the mis-weighted prior gets further away from 0.368 and closer towards it's true value as α increases from 0 to 1. As α increases from 1, the mis-weighted prior continues to increase. As the mis-weighted prior is increasing in α, subjects unilaterally become willing to experiment longer.

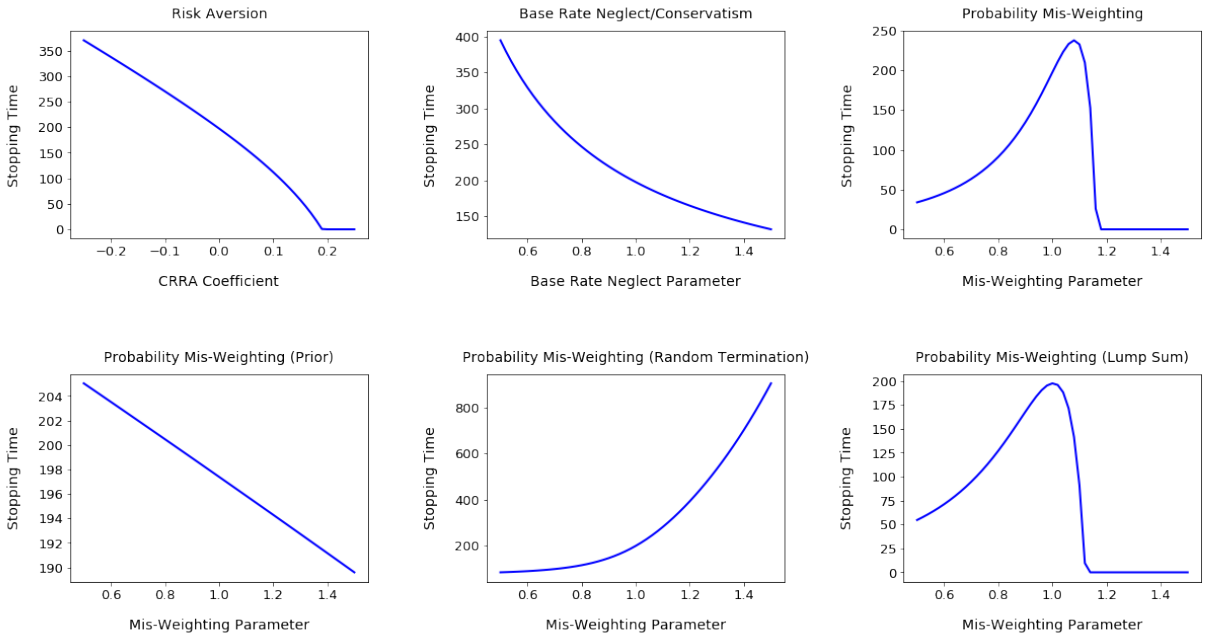


Figure A.3.: Predictions for the Low Safe Action Treatment as different behavioral factors are unilaterally varied.

Figure A.3 displays the predictions for the Low Safe Action treatment as various behavioral factors are varied. These responses to the behavioral factors are similar to the responses in Figure 1.2.

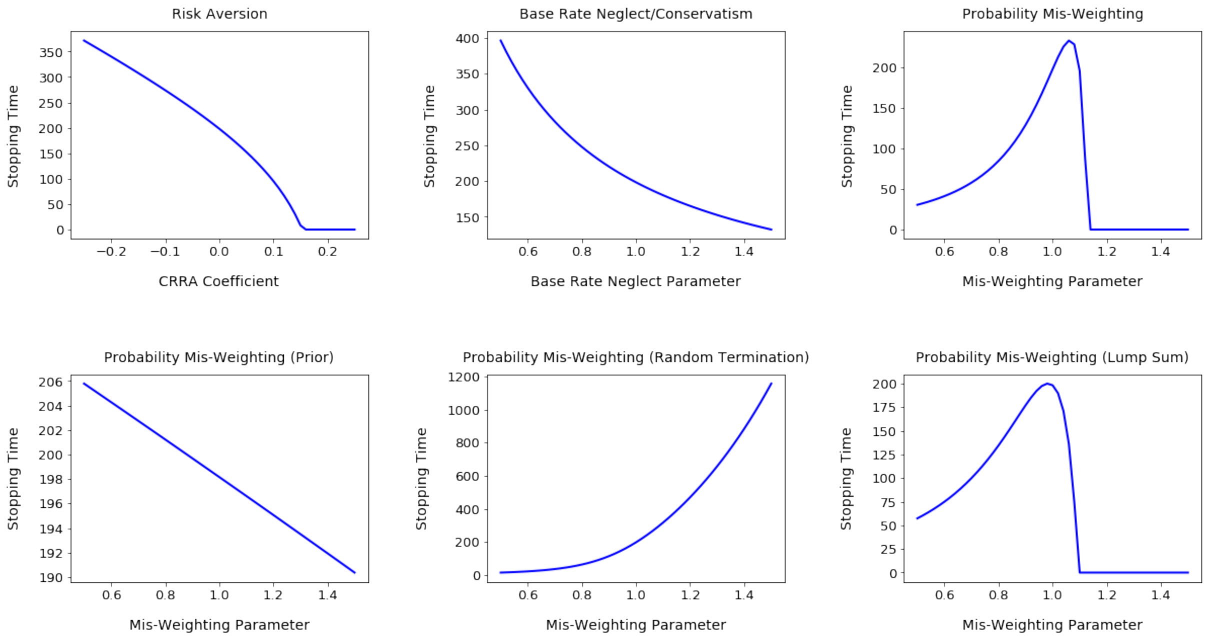


Figure A.4.: Predictions for the High Discount Factor Treatment as different behavioral factors are unilaterally varied.

Figure A.4 displays the predictions for the High Discount Factor treatment as various behavioral factors are varied. These responses to the behavioral factors are similar to the responses in Figure 1.2.

A.2 Empirical Appendix

This section of the appendix complements the results section of the paper. The first subsection focuses on details about the Product Limit Estimator. The second subsection displays figures that are related to the results section.

A.2.1 Product Limit Estimator

This subsection displays how the Product Limit estimator was used to conduct hypothesis testing. The first part of this subsection explains the Product Limit estimator. This part is taken from the explanation given in Hudja (2019). The second part of this subsection explains how the Product Limit estimator was used for hypothesis testing.

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. We then use a bootstrapped regression with 5000 bootstrap samples to test whether the appropriate mean (or the difference in appropriate means) is significantly different from the predicted value.

A.2.2 Additional Figures

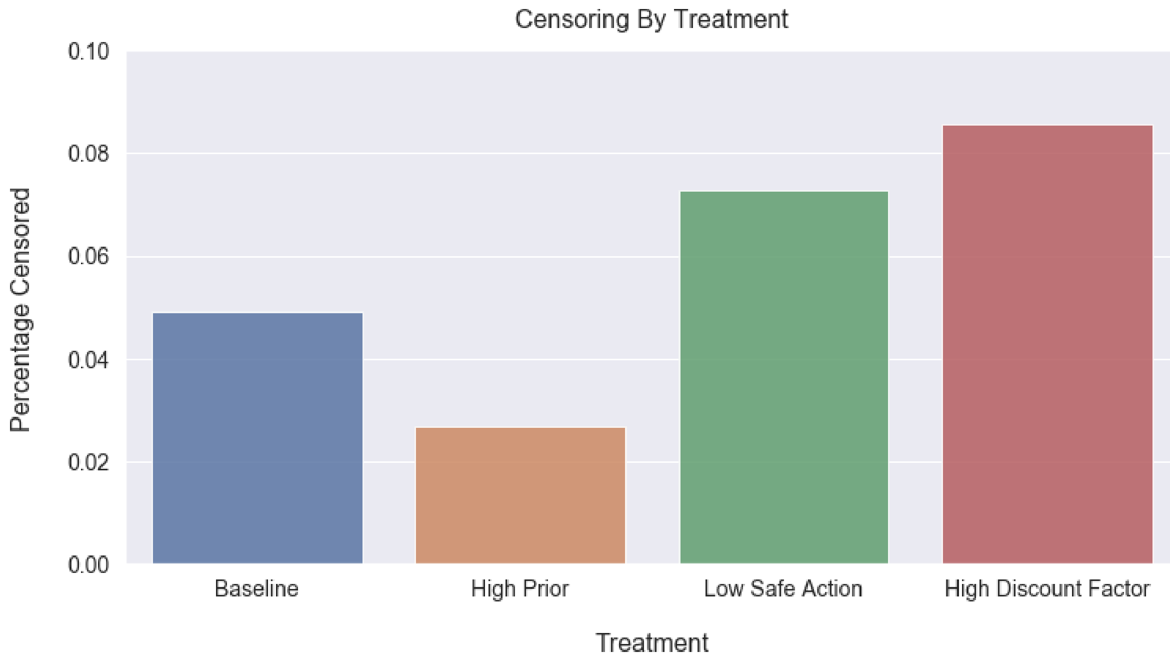


Figure A.5.: Rate of censoring in each treatment under the subset approach.

Figure A.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. This figure may explain why there are some discrepancies between the subset approach and the Product Limit approach for the High Discount Factor treatment as the most censoring occurs in the High Discount Factor Treatment.

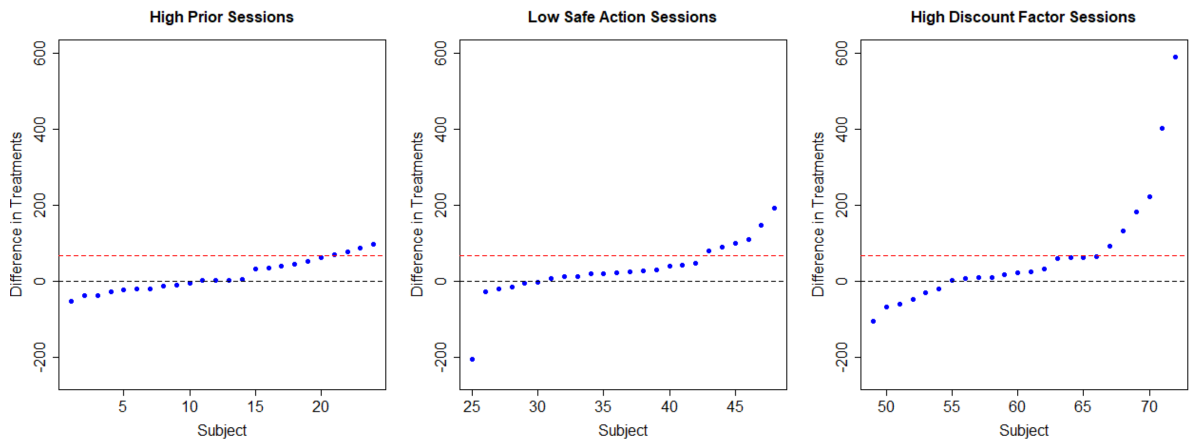


Figure A.6.: Difference in subjects' average Product Limit estimated stopping time in the Baseline treatment and in the session's other treatment. The other treatment is the High Prior treatment in the first graph, the Low Safe Action treatment in the second graph, and the High Discount Factor in the third graph. The red line displays the predicted response to the treatment variable, while the black line displays no response to the treatment variable.

Figure A.6 displays the Product Limit estimated version of Figure 1.3. For each subject, a mean Product Limit estimated stopping time for the Baseline treatment and for their session's other treatment was calculated. Figure A.6 plots the difference in these two means. These graphs are similar to the graphs in Figure 1.3 as a majority of subjects in each session appear to respond in the correct direction to a change in the treatment variable. The graph for the High Discount Factor sessions illustrates why the Product Limit approach fails to reject that subjects properly respond to an increase in the discount factor. There are a few subjects that appear to over-respond by a large factor to the change in the discount factor, which mitigates the effect of the subjects who appear to under-respond to the change in the discount factor.

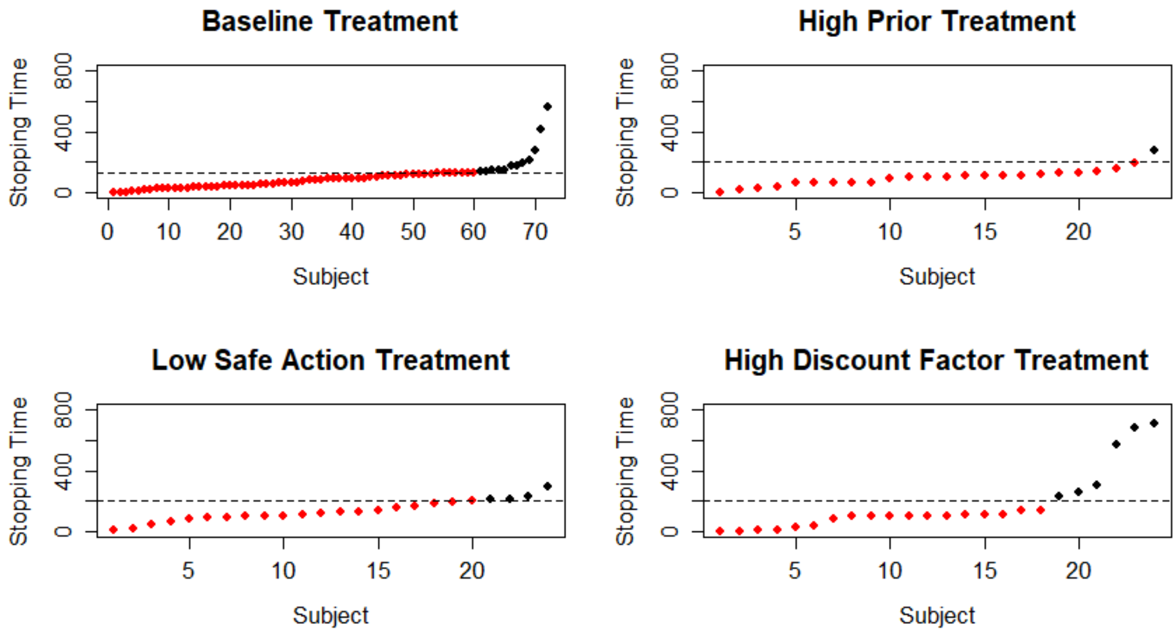


Figure A.7.: Mean Product Limit estimated stopping times in each treatment. Red dots denote a mean stopping time lower than the prediction. Orange dots denote a mean stopping time equal to the prediction. Black dots denote a mean stopping time greater than the prediction.

Figure A.7 displays the mean Product Limit estimated stopping time for each subject in each treatment. These graphs shed light on why we fail to reject that subjects experiment for the correct length of time in the High Discount Factor Treatment. As shown in the fourth graph, there are a few subjects that appear to over-experiment by a large margin, which mitigates the effect of the subjects who appear to under-experiment in the High Discount Factor treatment.

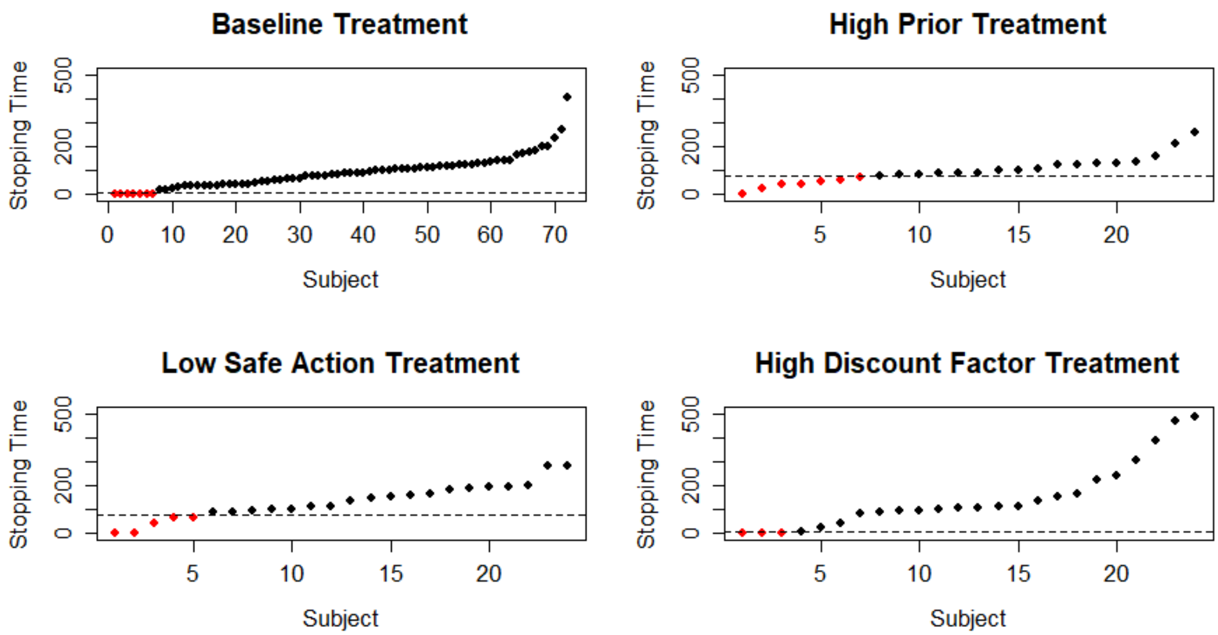


Figure A.8.: Mean stopping times in each treatment. Red dots denote a mean stopping time lower than the myopic prediction. Orange dots denote a mean stopping time equal to the myopic prediction. Black dots denote a mean stopping time greater than the myopic prediction.

Figure A.8 compares mean stopping times for each subject in each treatment to the myopic predictions. This figure uses the subset approach. The myopic prediction is equal to 5 for the Baseline and High Discount Factor treatments and is equal to 74 for the Low Safe Action and High Prior Treatments. In each figure, it is clear that a majority of subjects in each treatment experiment longer than the myopic prediction.

Figures A.9 compares mean stopping times for each subject in each treatment to the myopic predictions. This figure uses the Product Limit approach. The myopic prediction is equal to 5 for the Baseline and High Discount Factor treatments and is equal to 74 for the Low Safe Action and High Prior Treatments. In each figure, it is clear that a majority of subjects in each treatment experiment longer than the myopic prediction.

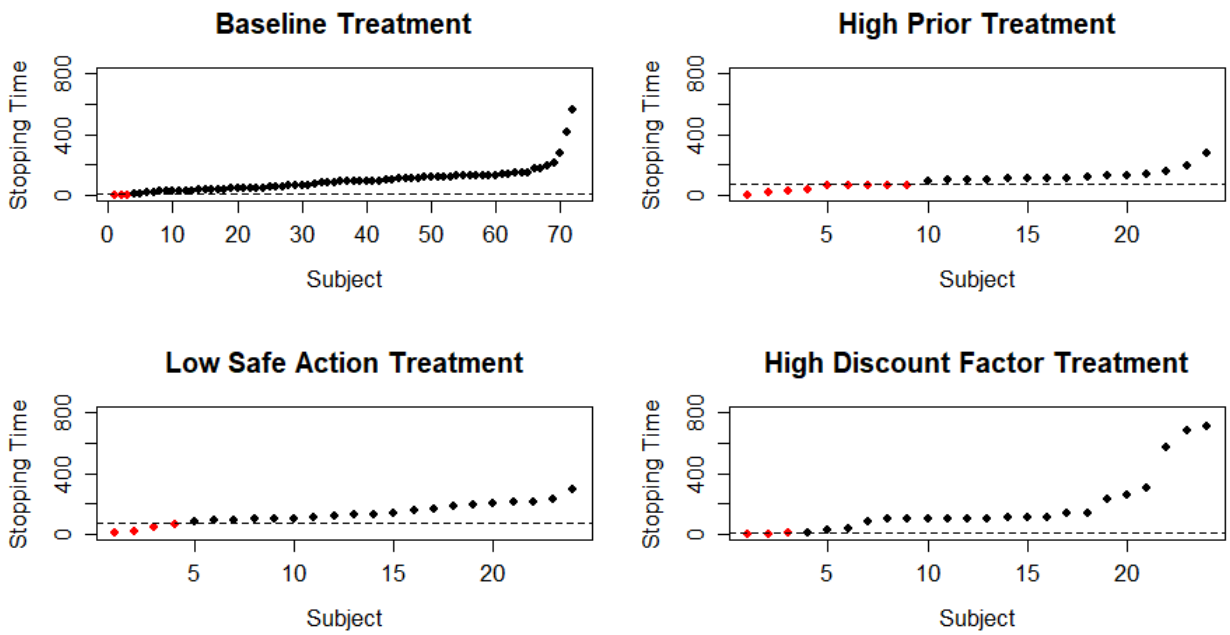


Figure A.9.: Mean Product Limit estimated stopping times in each treatment. Red dots denote a mean stopping time lower than the myopic prediction. Orange dots denote a mean stopping time equal to the myopic prediction. Black dots denote a mean stopping time greater than the myopic prediction.

Figure A.10 displays the predicted losses for each treatment. These are obtained through calculating each subject's average stopping time in a specific treatment using the subset approach. The predicted loss is obtained by differencing the amount of money expected from that stopping time and the amount of money expected from the optimal stopping time. The predicted losses are relatively small, but this is a feature of the exponential bandit problem as subjects are guaranteed a decent amount of money if they never experiment. However, in each treatment, there is a robust number of subjects who lose at least 50 cents.

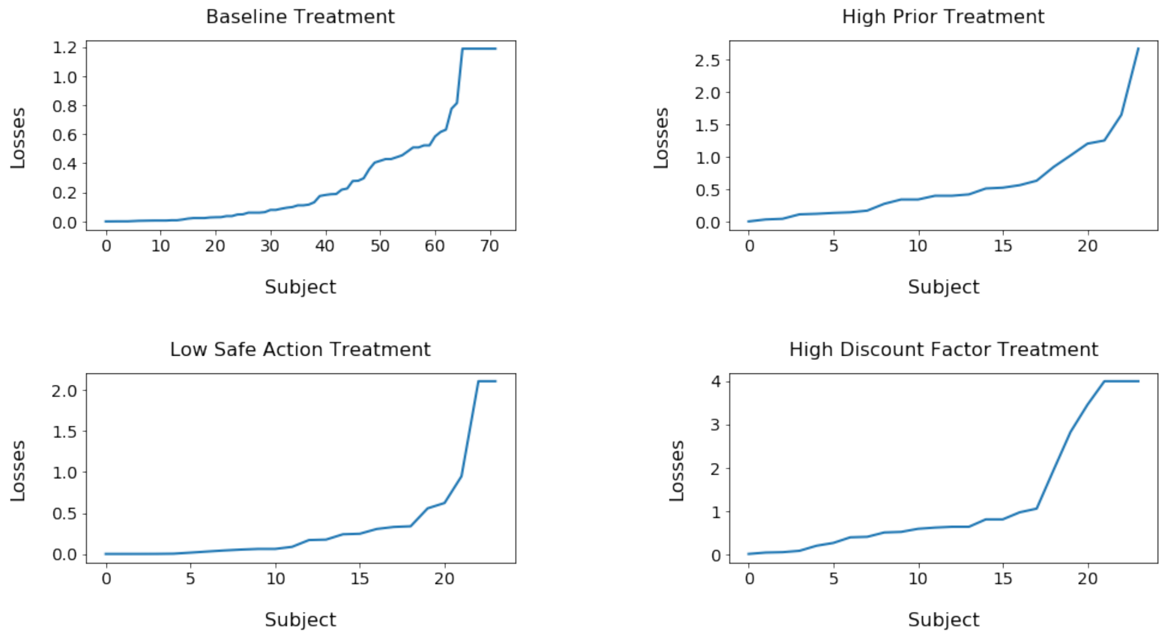


Figure A.10.: CDFs of predicted losses for each treatment.

A.3 Model Appendix

This section of the appendix addresses the model section of the paper. The first subsection focuses on a continuous time model that can be estimated which backs up the results of the discrete time approximation model. The second subsection displays figures that complement the analysis of the model section.

A.3.1 Continuous Time Model

In this subsection, we display the details for a continuous time version of the model that is estimated in section 1.5. This model is based on the model for experimentation under risk aversion in continuous time (Keller et al., 2019) and uses continuous time predictions to uncover behavioral factors.

In this section, we focus on the same behavioral factors as before: risk aversion, base rate neglect/conservatism, and probability mis-weighting. However, we assume that these behavioral factors influence the continuous time environment. Base rate neglect/conservatism is modeled similarly as before. Subjects are assumed to treat a second of experimentation as if it where ψ seconds of experimentation in the absence of a reward. Probabiliy mis-weighting of the prior is treated as it was before. The mis-weighted value of the prior is equal to $e^{-(-ln(p_0))^\alpha}$. Subjects are still assumed to mis-weight the probability that a good risky action results in a reward in a tick. This probability is still equal to $e^{-(-ln(\lambda\Delta))^\alpha}$. However, this results in the discrete time approximation approximating an arrival rate of $\tilde{\lambda} = e^{-(-ln(\lambda\Delta))^\alpha} \times \frac{1}{\Delta}$. Thus, we use this value of $\tilde{\lambda}$ for continuous time predictions. In this case, belief updating can be written, in the absence of any rewards, as

$$\frac{\tilde{p_0}e^{-\psi\tilde{\lambda}t}}{\tilde{p_0}e^{-\psi\tilde{\lambda}t} + (1 - \tilde{p_0})}.$$

In this formulation of the problem, we have an explicit form of the cutoff belief. Once again risk aversion is modeled through CRRA utility as $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$. The mis-weighting of the arrival rate is modeled in the same way as it is in the previous paragraph. Subjects mis-weight the discount rate as $e^{-(-ln(1-r\Delta))^\alpha}$, which results in the induced discount rate as approximating a discount rate of $\tilde{r} = e^{-(-ln(1-r\Delta))^\alpha} \times \frac{1}{\Delta}$. The continuous time cutoff belief, with risk aversion, now has a closed form solution, from Keller et al. (2019), which can be written as

$$\frac{u(s)}{\tilde{\lambda}u(h) + \frac{\tilde{\lambda}}{\tilde{r}}(\tilde{\lambda}u(h) - u(s))}.$$

The log-likelihood can be developed in the following way as in the discrete time model. The maximized value of the log-likelihood is equal to 8659.16. The estimated value of γ is 0.42, the estimated value of ψ is 0.20, and the estimated value of α is 0.74. The estimated value of σ is 166.78. The value of γ is significantly different from zero at the one percent level using a likelihood ratio test (restricted log-likelihood is equal to 8662.25). The value of ψ is significantly different from one at the ten percent level using a likelihood ratio test (restricted log-likelihood is equal to 8660.97). The value of α is significantly different from one at the five percent level using a likelihood ratio test (restricted log-likelihood is equal to 8662.05).

Standard errors are calculated through numerical differentiation. The standard error for γ is 0.16, the standard error for ψ is 0.17, the standard error for α is 0.13, and the standard error for σ is 3.69. Using these standard errors, γ is significantly different from zero at the one percent level, ψ is significantly different from one at the one percent level, and α is significantly different from one at the ten percent level (t-statistic=1.95). These results are similar to the discrete time approximation model as subjects appear to display risk aversion, conservatism, and probability mis-weighting.

A.3.2 Additional Figures

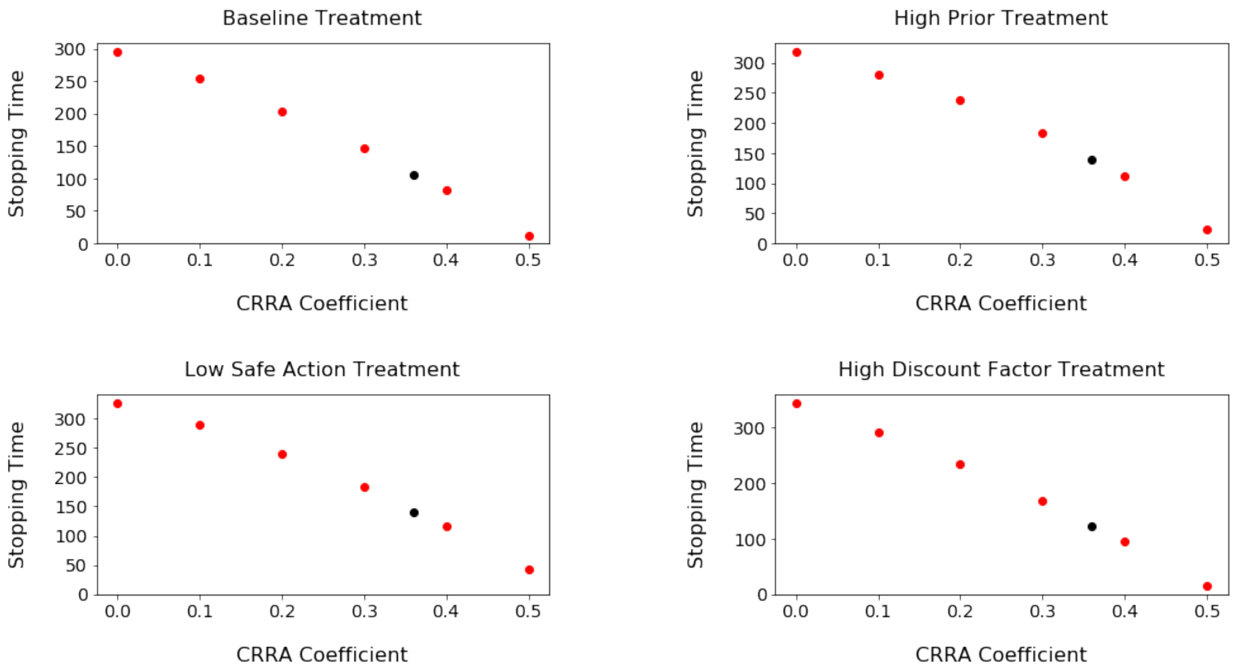


Figure A.11.: The effect of unilaterally changing the CRRA coefficient on subset approach predictions from the model. The black dot denotes the subset approach prediction from the model using the model's actual estimated CRRA coefficient.

Figure A.11 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.

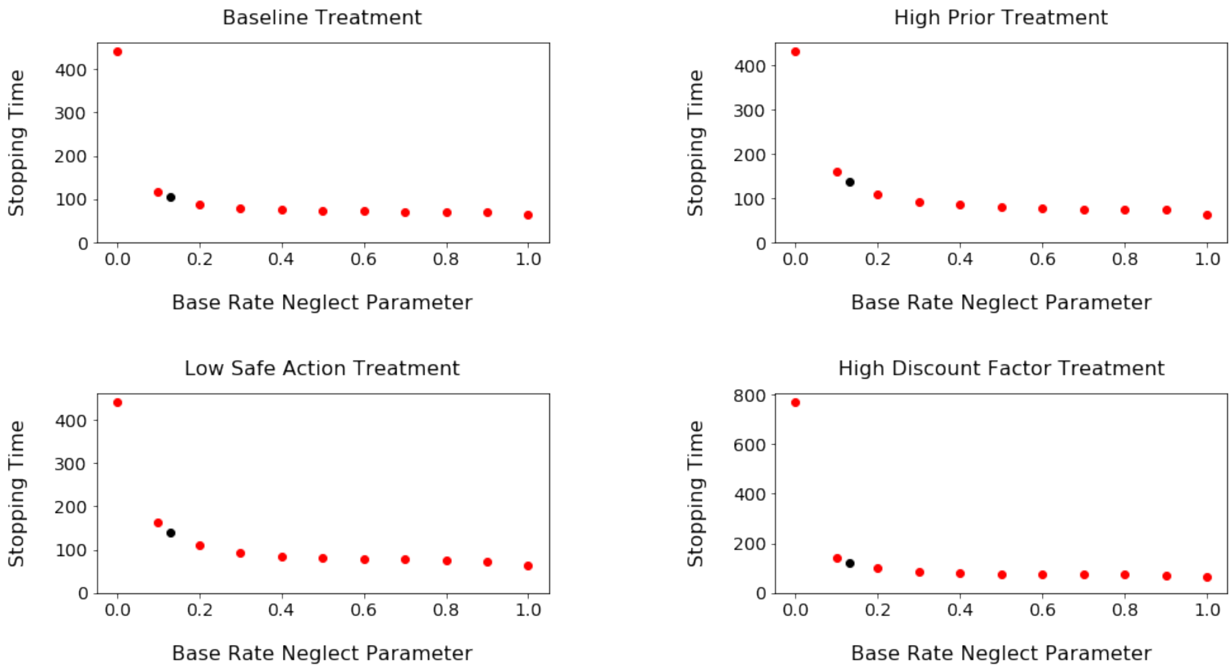


Figure A.12.: The effect of unilaterally changing the base rate neglect parameter on subset approach predictions from the model. The black dot denotes the subset approach prediction from the model using the model's actual estimated base rate neglect parameter.

Figure A.12 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.

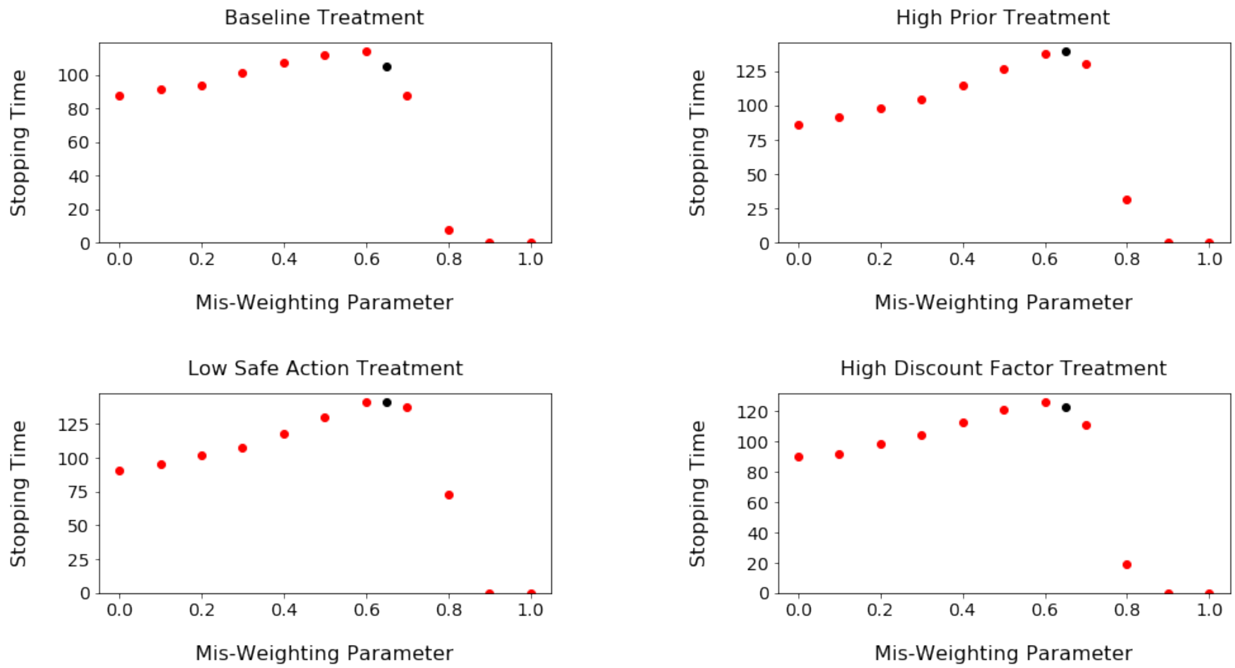


Figure A.13.: The effect of unilaterally changing the probability mis-weighting parameter on subset approach predictions from the model. The black dot denotes the subset approach prediction from the model using the model's actual estimated probability mis-weighting parameter.

Figure A.13 displays the prediction of each treatment, using the subset approach, when varying the probability mis-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 probability mis-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.

A.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 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.

[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 $\frac{p_0(g)^{cutpoint}(1-g)^{-cutpoint}}{p_0(g)^{cutpoint}(1-g)^{-cutpoint}+(1-p_0)}$.

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 experienced, 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 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 the 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 with a 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 with a 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 64 percent of the time at the five percent level. Using the Hudja (2019) noise, we find that subjects respond to δ 60 percent of the time at the five percent level. Using the

High Prior Treatment noise, we find that subjects respond to g 7 percent of the time at the five percent level. Using the Hudja (2019) noise, we find that subjects respond to δ 5 percent of the time at the five percent level. Both of these are lower than the recommended 80 percent (Moffatt, 2016).

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.

A.5 Instructions

In this section of the appendix, we display the instructions for a High Discount Rate session. The instructions for the High Prior and Low Safe Action sessions are similar to these instructions except that the prior and value of the safe action are respectively changing between blocks.

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. 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 0.50 points in the current tick and 0.50 points in each of the remaining ticks. For 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 can change.** This probability can only change between blocks. Within each block, this probability stays the same.

At the start of a block, the probability that the period will end in the current tick will be displayed. In order to start the block, you are required to correctly answer a question on this probability.

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, 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 for each tick that you do not draw a ball. 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.

A.6 Video Transcript

In this section, we display the transcript of the interface video that we presented to subjects at the start of the High Discount Rate session. We display the transcript in an outline format. This outline format is read from when recording the video. The transcript for the High Prior session and the Low Safe Action session is similar to this transcript except details on the block information and parameter values are changed.

Transcript

This video is designed to show you how to use the experimental interface.

- The examples in this video are made up to illustrate the interface
- There are five components to the interface:
  - The first component is the timer
    - The timer is in the upper left corner
    - Once you have made an initial decision, the timer counts down from 5 seconds to 0 seconds
    - When the timer hits 0 seconds the first tick starts
  - The second component is the table of information
    - The table of information provides Block information and General information
      - The block information tells you the probability that the current tick ends the period
      - The probability that the current tick ends the period may change between blocks, but will not change within blocks
      - In the first period of a new block, there is a question that must be answered to start the period

      - The general information gives you regular information from the instructions
    - After the last tick has occurred, the table of information gives summary information on the current period
  - The third component is the graph
    - The graph displays information relevant to the current tick
    - We will come back to the graph once the first tick starts
  - The fourth component is the decision buttons
    - The decision buttons allow you to draw or not draw a ball
    - The buttons are the "yes" button and the "no" button
    - To draw, click on the "yes" button or use the left-arrow key
    - To not draw, click on the "no" button or use the right-arrow key
    - To finalize the initial decision, click the ready button
  - The fifth component is the payoff, which displays your current payoff in the period in points

- For this hypothetical example, I will click the "yes" button and start off drawing balls
- I will click the ready button to start the timer
- Upon clicking the ready button, the five second timer will start
- Once the timer hits zero, the ticks start and the graph is displayed
- I will now click the ready button
- The graph has a few features worth mentioning
  - A red line that doesn't move is drawn at the current tick
  - To the right of the middle of the red line is the number of balls drawn so far

  - To the right of the bottom of the red line is the current tick number
  - To the left of the red line is the payoff history in each of the last eighty ticks
  - Notice that we have not obtained any payoff yet so we are drawing yellow balls
  - Since we have not drawn a red ball yet, we can switch to not drawing balls
  - I will now switch to not drawing balls by using the right-arrow key
  - We now have a payoff in each tick of 0.50 points, as shown by the blue line
  - The number of balls drawn, so far, are now constant
  - Notice that we can not change our action to drawing a ball
  - Now we wait until the random termination of the period
- At the end of the period, the table of information provides us with information from the current period
  - It shows us the payoff we receive, the number of balls drawn, how many ticks had occurred, and the number of red balls drawn
  - To continue to the next period, we click the continue button
- We will go over one more example before returning to the instructions
- Once again, we will choose an initial action, for this example I will choose to start off drawing balls
  - This is done just to show you more features of the interface
- I will click on the ready button to initialize the timer
- The five second-timer starts and when the timer hits zero, the first tick occurs
- In this example, I have received a red ball

- The line of height 155 shows that I have received a red ball
- Notice that now I can not change my decision
- The “no” button is deactivated and clicking on it doesn’t change the decision

• Notice that we have obtained more than one red ball, you are not limited to one red ball in a period

• At the end of the period, we read the table of information and then move on to the next period by clicking the continue button

A.7 Post-Experimental Questionnaire

In this section, we display the questionnaire that subjects completed at the end of the experiment. This questionnaire was used to collect demographic information on the subjects.

Questionnaire

Question 1: What is your gender? Please write on the line below.

____

Question 2: What country were you born in? Please write on the line below.

____

Question 3: What is the main field of study for your undergraduate degree? Please circle one option.

- Management/Business
- Economics
- Humanities
- Liberal Arts
- Education
- Engineering
- Science
- Social Sciences
- Agriculture
- Pharmacy
- Nursing
- Other

Question 4: What do you consider your primary racial identity? Please circle one option.

- Asian
- Black
- Caucasian
- Hispanic
- Other

Question 5: What is your GPA? Please circle one option.

- 3.5-4.0
- 3.0-3.5
- 2.5-3.0
- 2.0-2.5
- Below 2.0

Question 6: Are you an undergraduate student (which year) or a graduate student? Please circle one option.

- First Year
- Second Year
- Third Year
- Fourth Year or Above
- Graduate Student

B. APPENDIX FOR: VOTING FOR EXPERIMENTATION: A CONTINUOUS TIME ANALYSIS

B.1 Theoretical Appendix

The theoretical appendix shows how the equilibrium predictions from Strulovici (2010) are calculated. The predictions follow from the proof of Theorem 1 in the appendix of Strulovici (2010).

Using the smooth-pasting condition of the Hamilton-Jacobi-Bellman equation (2), the indifference threshold $p(1)$ must solve

$$pg + \lambda p \left(\frac{g}{r} - \frac{s}{r}\right) + \lambda p \left(\frac{pg}{r} - \frac{s}{r}\right) = s. \tag{B.1}$$

Notice that $w(2,p) = \frac{g}{r}$ and that $u(2,p) = \frac{pg}{r}$. When there are two winners, sure winners have the majority and the risky action is chosen forever. The smooth-pasting property implies that the derivative of the value function drops out when solving for $p(1)$. The value matching property implies that the value function is equal to $\frac{s}{r}$. The left-hand side of (3) is increasing in p, equal to 0 if $p = 0$ and higher than s if $p = 1$. Therefore, the equation has a unique root, which can be expressed as

$$p(1) = \frac{rs}{rg + \lambda(g - s) + \lambda(p(1)g - s)}. \tag{B.2}$$

When $p > p(1)$ all unsure voters vote for R, when $p \leq p(1)$ all unsure voters vote for S. Once all unsure voters vote for S, no learning occurs and S is played forever. The cutoff $p(1)$ determines the value functions $w(1,p)$ and $u(1,p)$ of sure and unsure voters, which are computable in closed form:

$$w(1,p) = \frac{g}{r} - \frac{g-s}{r}\left(\frac{1-p}{1-p(1)}\right)^2 \left(\frac{\Omega(p)}{\Omega(p(1))}\right)^{\frac{r}{\lambda}} \tag{B.3}$$

and

$$u(1,p) = \frac{pg}{r} - \frac{s-p(1)g}{r}\left(\frac{1-p}{1-p(1)}\right)^2 \left(\frac{\Omega(p)}{\Omega(p(1))}\right)^{\frac{r}{\lambda}} \tag{B.4}$$

for $p \geq p(1)$, where $\Omega(p) = \frac{(1-p)}{p}$. These functions are increasing in p. Using these new functions I can solve for $p(0)$, which must solve

$$pg + p\lambda(w(1,p) - \frac{s}{r}) + 2p\lambda(u(1,p) - \frac{s}{r}) = s. \tag{B.5}$$

There is a unique root for $p(0)$ as the left-hand side is increasing in p, equal to 0 for $p = 0$ and above s for $p = 1$.

B.2 Value Function Iteration Appendix

The value function iteration appendix shows how the discrete time equilibrium and discrete time utilitarian optimum predictions were obtained. The first subsection discusses the equilibrium predictions, while the second subsection discusses the utilitarian predictions.

B.2.1 Discrete Time Equilibrium Predictions

The single-agent predictions are found by using 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 1001 values of p from .5 to $\frac{.5*(1-\lambda\Delta)^{1000}}{.5*(1-\lambda\Delta)^{1000}+.5}$. The initial guess for $u(p)$ is $\frac{s}{r\Delta}$ at each value of p. This initial guess for $u(p)$, for each of the 1001 values of p, is chosen for all of the value function iterations done in this subsection.

The one-winner predictions are found by using value function iteration on the following equation:

$$\begin{aligned} u(p) =& max\{\frac{s}{r\Delta}, p * \lambda\Delta * (h + (1 - r\Delta) * \frac{\lambda\Delta * h}{r\Delta}) \\ &+ (p * \lambda\Delta * (1 - p * \lambda\Delta)) * (1 - r\Delta) * \frac{p' * \lambda\Delta * h}{r\Delta} \\ &+ ((1 - p\lambda\Delta)^2) * (1 - r\Delta) * u(p')\}. \end{aligned}$$

In order for the zero-winner predictions to be calculated, the value of becoming a winner and the value of being an unsure voter with one winner must be calculated. In order to calculate these values, I must use the optimal stopping time from when there is one winner, which is at 153 ticks. Let $draws(p)$ $(= ln\left(\frac{p(1-p_0)}{(1-p)p_0}\right)/(ln(1-\lambda\Delta)))$ be the number of draws that have already occurred. Notice that there is a unique number of draws that have occurred at each p since our prior is .5. If $draws(p) < 153$, the value of becoming a winner equals:

$$\begin{aligned} Winner(p) =& \lambda\Delta * h * \frac{(1 - (1 - r\Delta)^{153-draws(p)})}{r\Delta} \\ &+ ((1 - p + p * (1 - \lambda\Delta)^{153-draws(p)})^2) * ((1 - r\Delta)^{153-draws(p)}) * \frac{s}{r\Delta} \\ &+ (1 - ((1 - p + p * (1 - \lambda\Delta)^{153-draws(p)})^2)) * ((1 - r\Delta)^{153-draws(p)}) * \frac{\lambda\Delta * h}{r\Delta}. \end{aligned}$$

If the value of $draws(p) \geq 153$, the value of becoming a winner equals:

$$Winner(p) = \frac{s}{r\Delta}$$

as the safe action will be imposed.

The value of being an unsure voter when there is one winner and when $draws(p) <$ 153 equals:

$$\begin{aligned}
Unsure(p) =& p * \lambda\Delta * h * \frac{(1-(1-r\Delta)^{153-draws(p)})}{r\Delta} \\
&+ ((1-p+p*(1-\lambda\Delta)^{153-draws(p)})^2) * ((1-r\Delta)^{153-draws(p)}) * \frac{s}{r\Delta} \\
&+ (1-(1-p+p*(1-\lambda\Delta)^{153-draws(p)})) * ((1-r\Delta)^{153-draws(p)}) * \frac{\lambda\Delta * h}{r\Delta} \\
&+ (1-p+p*(1-\lambda\Delta)^{153-draws(p)}) * (1-(1-p+p*(1-\lambda\Delta)^{153-draws(p)})) \\
&* ((1-r\Delta)^{153-draws(p)}) * \frac{p_{153} * \lambda\Delta * h}{r\Delta},
\end{aligned}$$

where $p_{153} = \frac{p*(1-\lambda\Delta)^{153-draws(p)}}{p*(1-\lambda\Delta)^{153-draws(p)}+(1-p)}$. If the value of $draws(p) \geq 153$, the value of becoming an unsure voter with one winner is:

$$Unsure(p) = \frac{s}{r\Delta}.$$

The zero-winner predictions are found by using value function iteration on the following equation:

$$\begin{aligned}
u(p) =& max\{\frac{s}{r\Delta}, p * \lambda\Delta * ((1-p*\lambda\Delta)^2) * (h + (1-r\Delta) * winner(p')) \\
&+ p * \lambda\Delta * (1-(1-p*\lambda\Delta)^2) * (h + (1-r\Delta) * \frac{\lambda\Delta * h}{r\Delta}) \\
&+ 2 * ((1-p\lambda\Delta)^2) * (p * \lambda\Delta) * (1-r\Delta) * unsure(p') \\
&+ (1-p\lambda\Delta) * ((p\lambda\Delta)^2) * (1-r\Delta) * p' * \frac{\lambda\Delta * h}{r\Delta} \\
&+ ((1-p\lambda\Delta)^3) * (1-r\Delta) * u(p')\}.
\end{aligned}$$

B.2.2 Discrete Time Utilitarian Predictions

The utilitarian prediction for when there is more than one winner is to always choose the risky action. This occurs because the risky action provides a higher group payoff than the safe action when there are two or more winners. The one-winner predictions are found using value function iteration on the following equation:

$$\begin{aligned}
u(p) =& max\{\frac{3s}{r\Delta}, \lambda\Delta * h + 2 * p * \lambda\Delta * h \\
&+ ((p * \lambda\Delta)^2) * ((1 - r\Delta)) * \frac{3 * h * \lambda\Delta}{r\Delta} \\
&+ 2 * (p * \lambda\Delta) * (1 - p * \lambda\Delta) * (1 - r\Delta) * \frac{2 * \lambda\Delta * h + p' * \lambda\Delta * h}{r\Delta} \\
&+ ((1 - p * \lambda\Delta)^2) * (1 - r\Delta) * u(p')\}.
\end{aligned}$$

Value function iteration now preceeds in the same way as in the equilibrium predictions but the intial guess at each p is now $\frac{3}{r\Delta}$.

In order to calculate the one-winner predictions, the utilitarian value when there is one winner must be calculated. In order to calculate this value, the optimal utilitarian stopping time when there is one winner, which is 358 ticks, must be used. Once again, $draws(p)$ is the number of draws that have occured. There is a unique value of $draws(p)$ at each p since we have a prior of .5. The utilitarian value when there is one winner, if $draws(p) < 358$ is given by the following equation:

$$\begin{aligned}
Winner(p) =& (\lambda\Delta * h + 2 * p * \lambda\Delta * h) * \frac{1 - (1 - r\Delta)^{358 - draws(p)}}{r\Delta} \\
&+ ((1 - p + p * (1 - \Delta)^{358 - draws(p)})^2) * ((1 - r\Delta)^{358 - draws(p)}) * \frac{3 * s}{r\Delta} \\
&+ ((1 - (1 - p + p * (1 - \Delta)^{358 - draws(p)}))^2) * ((1 - r\Delta)^{358 - draws(p)}) * \frac{3 * \lambda\Delta * h}{r\Delta} \\
&+ 2 * (1 - (1 - p + p * (1 - \Delta)^{358 - draws(p)})) * (1 - p + p * (1 - \Delta)^{358 - draws(p)}) \\
&* ((1 - r\Delta)^{358 - draws(p)}) * \frac{2 * \lambda\Delta * h + p_{358} * \lambda\Delta * h}{r\Delta},
\end{aligned}$$

where $p_{358} = \frac{p*(1-\lambda\Delta)^{358-draws(p)}}{p*(1-\lambda\Delta)^{358-draws(p)}+(1-p)}$. If $draws(p) \geq 358$, the value of one winner is given by the following equation:

$$Winner(p) = \frac{3 * s}{r\Delta}.$$

The zero-winner predictions are found using value function iteration on the following equation:

$$
\begin{aligned}
u(p) = & max\{\frac{3s}{r\Delta}, 3 * p * \lambda\Delta * h + ((p * \lambda\Delta)^3) * ((1 - r\Delta) * \frac{\lambda\Delta * h * 3}{r\Delta} \\
& + 3 * ((p * \lambda\Delta)^2) * (1 - p * \lambda\Delta) * (1 - r\Delta) * \frac{2 * \lambda\Delta * h + p' * \lambda\Delta * h}{r\Delta} \\
& + 3 * ((1 - p * \lambda\Delta)^2) * p * \lambda\Delta * (1 - r\Delta) * winner(p') \\
& + ((1 - p * \lambda\Delta)^3) * (1 - r\Delta) * u(p')\}
\end{aligned}
$$

B.3 Data Appendix

The data appendix is broken down into two subsections. The first subsection analyzes the same hypotheses as the results section, but tests these hypotheses on a larger dataset using survival analysis. The second subsection in the appendix analyzes the first ten periods of each treatment in the experiment.

B.3.1 Larger Dataset

This subsection analyzes the same hypotheses as the results section, but on a larger subset of data than was used in the results section. This section conducts hypotheses on observations, in the last fifteen periods of each treatment, where a majority of group members have bad states in the majority-vote treatment and where single-agents have bad states in the single-agent treatment. I refer to these observations as the larger dataset. There are three hypotheses for the larger dataset: (i) groups stop earlier than single-agents, (ii) groups with one winner stop later than groups with zero winners, and (iii) groups stop earlier than the utilitarian optimum predicts. These hypotheses will be conducted using survival analysis which corrects for the censoring that occurs due to the random termination of the period.

This subsection, unless noted otherwise, conducts tests on groups' Product Limit estimated mean stopping times. This analysis gets rid of the dependence that occurs when using survival analysis on multiple observations from the same group. The average of the groups' Product Limit estimated mean stopping times for when a winner is impossible is 108.8. This result is not statistically significant from the predicted 110 ticks at the 10 percent level (p-value=.835). Hypothesis tests for this paragraph are conducted using bootstrapped regressions, on the group Product Limit estimated mean stopping times, with 5000 bootstrap samples. The average of the groups' Product Limit estimated mean stopping times for when a winner is predicted is 111.8. This result is statistically significant from the predicted 153 ticks at the 1 percent level. Lastly, the average of the groups' Product Limit estimated mean stopping times for when a winner is not predicted is 100.3. This result is not statistically significant from the predicted 110 at the 10 percent level (p-value=.278). These results back up the results from the summary statistics of the clean dataset.

The first hypothesis states that, in the larger dataset, groups stop earlier than single-agents. I create a Product Limit estimated mean for each group in the majority-vote observations of the larger dataset and the single-agent observations of the larger dataset. I create a Product Limit estimated mean in the single-agent observations by pooling the three group members data and taking the Product Limit estimated average. I test the hypothesis by taking the difference between the two averages for each group and running a bootstrapped regression with 5000 bootstrap samples. The average group stops 36.0 ticks later in the single-agent observations than in the majority-vote observations. This result is significant at the 1 percent level (p-value=0.006). This gives evidence for the first hypothesis on the larger dataset and is a similar result to the first hypothesis on the clean dataset.

The second hypothesis states that, in the larger dataset, groups with one winner stop earlier than groups with zero winners. A Cox regression on this subset of data has a coefficient of -.9861, and a hazard ratio of .373, which is significant at the 1 percent level. This hazard ratio suggests that groups with zero winners are 2.68 times more likely, per unit of time, to switch to the safe action. This gives evidence for the second hypothesis on the larger dataset and is a similar result to the second hypothesis on the clean dataset.

The third hypothesis states that, in the larger dataset, groups stop earlier than the utilitarian optimum predicts. When a winner is impossible, according to the utilitarian optimum, in the larger dataset, the average of the groups' average stopping times is 108.8. This result is significantly less than the predicted 132 at the 1 percent level. Hypothesis tests are conducted by bootstrapped regressions with 5000 boostrap samples. When a winner is predicted, according to the utilitarian optimum, in the larger dataset, the average of the groups' average stopping times is 114.2. This result is significantly less than the predicted 358 at the 1 percent level. When a winner is

Table B.1.: Mean stopping time (in ticks) of majority-vote observations and single-agent observations. The equilibrium prediction of overall implementation in the majority-vote treatment was calculated from the states and rewards drawn for the experiment.

Stopping Time	Majority-Vote			Single-Agent		
	Equil.		Actual	Equil.		Actual
Overall	125.9	$\overset{*}{<}$	173.5	187.0	$\overset{***}{>}$	144.4
Winner Impossible	110.0	$<$	111.2	—		—
Winner Predicted	153.0	$\overset{**}{<}$	240.8	—		—
Winner Not Predicted	110.0	$\overset{*}{<}$	164.0	—		—

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

impossible, according to the utilitarian optimum, in the larger dataset, the average of the groups' average stopping times is 92.1. This result is significantly less than the predicted 132 at the 1 percent level. This gives evidence for the third hypothesis on the larger dataset and is a similar result to the third hypothesis on the clean dataset.

B.3.2 First Ten Periods of Each Treatment

Throughout the paper, I analyzed observations from the last fifteen periods of each treatment. I ignored the first ten periods of each treatment because of learning. Subjects may be learning about their own strategies, the Bernoulli process, and the random stopping rule. This subsection analyzes the experimental data in the first ten periods of each treatment. The data observed is the data from the first ten periods of each treatment, where the period lasted at least 200 ticks, and where a majority of group members had bad states and single-agents had bad states.

Table B.1 displays the mean stopping time in the first ten periods of each treatment. The data is analyzed in the same way as in the results section. Overall, groups stop later than predicted. This result is significant at the ten percent level (p-value=0.07). Hypothesis tests in this subsection are conducted with a bootstrapped

Table B.2.: Displays the results from the cox regression that estimates the effect of a time-dependent covariate, *Winner*, on implementation time of the risky action in the first ten periods of each treatment.

	Coeff.	Hazard Ratio	Z
Winner	-1.458	0.233	-5.43***
N		192	

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

regression with 5000 bootstrap samples clustered at the group level. When a winner is impossible, there is no statistical difference between stopping time and the equilibrium prediction. When a winner is predicted, groups stop later than the predicted 153 ticks. This result is significant at the five percent level (p-value=0.043). When a winner is not predicted, but possible, the stopping time is later than the predicted 110 ticks. This result is significant at the 10 percent level (p-value=0.077).

Single-agents stop earlier than predicted. Subjects stop earlier than the equilibrium prediction of 187. This result is significant at the one percent level. The results in Table B.1 suggest that groups and single-agents both decrease their stopping time as the experiment goes on. However, groups start off stopping later than predicted, while single-agents start off stopping earlier than predicted. These results also suggest that groups initially stop later than single-agents do. It is apparent from these results that hypothesis 1 does not hold in the first ten periods of each treatment. It is also apparent from these results that hypothesis 3 does hold.

Finally, Table B.2 estimates the impact of a winner on the stopping time in the first ten periods of the majority-vote treatment. The coefficient on *Winner* is negative and the hazard ratio is less than one. Thus, observing a winner makes groups less likely to witch to the safe action in the next tick. The hazard ratio decreases by 77 percent upon observing a winner. This result is significant at the 1 percent level and consistent with Table 6, which shows that stopping times are greater in the observations where a winner is possible than in the observations where a winner is impossible.

B.4 Instructions

In this section, I display the instructions for the majority-vote treatment and the belief treatment, which is an additional treatment to analyze under-experimentation. The instructions for the single-agent treatment are similar to the instructions for the majority-vote treatment except that there is no mentioning of a group. The instructions for the no-discounting treatment are similar to the belief treatment except that the period length is fixed for 333 ticks and the value of a reward increases to 3.65 experimental dollars. The part 2 that is mentioned in the belief treatment is the risk aversion elicitation task.

B.4.1 Instructions for the Majority-Vote Treatment

Part 1 - 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 will have the opportunity to earn $5.00 based on how you answer five comprehension questions about these instructions. For each correct answer to a question you will earn $1.00. You will be able to refer to these written instructions as you answer the questions.

These instructions are for the first part of a two-part experiment. The choices made in this part of the experiment will **in no way affect** your earnings in the second part of the experiment. From this point forward, all units of account will be in experimental dollars. At the end of the experiment, experimental dollars will be converted to U.S. dollars at the rate of 2 U.S. dollars for every 5 experimental dollars (i.e. .05

experimental dollars are worth \$0.02).

Background

In this experiment, imagine that there are two types of bags. The first type is a 'mixed' bag (denoted by the letter M), which contains **1 red ball and 99 yellow balls**. 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 is a 'uniform' bag (denoted by the letter U), which contains **0 red balls and 100 yellow balls**. 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 do not change.

At the beginning of this part of the experiment, you will be randomly grouped with two other people. You will be grouped with these **two other individuals for all periods** in this part of the experiment. Each member of the group has their own M bag and U bag.

You will be participating in 25 periods in this part of the experiment. At the start of a new period, each group member has one of their bags randomly and independently selected for them. The type of bag selected for you may differ from the type of bag selected for another group member. At the start of a new period, there is a **50 percent** chance that the bag you are drawing from is your M bag and a **50 percent** chance that the bag you are drawing from is your U bag. Your M bag is therefore just as likely as your U bag to be selected.

The random selection of bags is independently determined before each period, as if by flipping a coin three times. If the coin on the first flip lands on "heads" the first member of the group has their M bag selected, but if the coin lands on "tails" the first member of the group has their U bag selected. The outcome of the second flip determines the bag for the second member of the group. The outcome of the third flip determines the bag for the third member of the group. Notice that the type of bag selected for another member of your group **does not affect the type of bag selected for you**.

The bag, for each person, 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**.

Period

Each period consists of many 'ticks'. In each tick, each group member may draw a ball from their own bag. Each tick lasts for a fifth of a second (i.e. five ticks per second), and ticks continually occur until the period ends. The period ends with a small probability each tick (3 out of 1000). Imagine 1000 tickets, numbered 1 through 1000 are placed in a box. It is as if, after every tick, a ticket is randomly drawn. If the number on the ticket is 1, 2, or 3, the period ends. If the number on the ticket is 4 through 1000, the period continues and the ticket is placed back in the box, so the contents of the box do not change. Under this probability, the average period length is 333.33 ticks or 66.66 seconds. Many periods will last less than the average, and a few will last much longer.

As mentioned above, in each tick, each group member may draw a ball from their own bag. Each group member will be asked an initial question of whether or not they would like to vote to draw a ball. In general, the vote you have previously cast will be continually cast until you decide to change it. If a majority (two or more) of members in your group initially vote to draw a ball, each member will continually draw a ball from their own bag in each tick until a majority vote to stop drawing a ball. At any tick, if a majority of group members vote to draw a ball, everyone in the group draws a ball from their bag. Once a majority of group members have voted to stop drawing a ball, each member can **no longer draw a ball** from their own bag for the rest of the ticks in the period. If a majority vote to initially not draw a ball, each member **cannot draw a ball** for any of the ticks in the period.

If you draw a ball, you can either draw a red ball or a yellow ball. If you draw a red ball, you obtain **2.50 experimental dollars**. Your group members will be informed that you drew a red ball. Once you have obtained a red ball **you automatically vote to draw a ball** for the rest of the ticks in the period (regardless of your previous vote). Notice that if two group members draw a red ball, you will draw a ball for the rest of the ticks in the period. You **can obtain** more than one red ball in a period. For example, if you draw a red ball, you can obtain a payoff of 2.50, 5.00, 7.50 experimental dollars, etc. from the red balls you draw in that period. If you draw a yellow ball, you will not be compensated for it. Notice that if you have drawn yellow balls in all ticks so far, you can vote to stop drawing a ball. If a majority of group members vote to stop drawing a ball, you obtain **0.01 experimental dollars** in the current tick and **0.01 experimental dollars** in each of the remaining ticks.

Five periods from this part will be randomly chosen for payment. The periods chosen for you may differ from the periods chosen for others.

Interface

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

Recap

The following is a recap of important parts of the instructions

- 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

- For each period, it is equally likely that you are drawing from your M bag or your U bag. The type of bag selected for one of your group members does not affect the type of bag selected for you

- In each tick, if a majority of group members vote to draw a ball, all group members draw a ball from their own bag. If a majority of group members vote to not draw a ball, all group members do not draw a ball for the rest of the ticks in the period

- If you draw a red ball, you obtain 2.50 experimental dollars and automatically vote to draw a ball in the remainder of the ticks in the period. You can obtain more than one red ball in a period

- If you draw a yellow ball, you are not compensated

- If a majority of group members vote to stop drawing a ball, you obtain .01 experimental dollars in the current tick and each of the remaining ticks (each group member can no longer draw a ball from their bag for the rest of the ticks in the period)

B.4.2 Instructions for the Belief Treatment

Part 1 - 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 will have the opportunity to earn \$5.00 based on how you answer five comprehension questions about these instructions. For each correct answer to a question you will earn \$1.00. You will be able to refer to these written instructions as you answer the questions.

These instructions are for the first part of a two-part experiment. The choices made in this part of the experiment will **in no way affect** your earnings in the second part of the experiment. From this point forward, all units of account will be in experimental dollars (unless noted otherwise). At the end of the experiment, experimental dollars will be converted to U.S. dollars at the rate of 2 U.S. dollars for every 5 experimental dollars (i.e. .05 experimental dollars are worth \$0.02).

Background

In this experiment, imagine that there are two types of bags. The first type is a 'mixed' bag (denoted by the letter M), which contains **1 red ball and 99 yellow balls**. 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 is a 'uniform' bag (denoted by the letter U), which contains **0 red balls and 100 yellow balls**. 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 do not change.

You will be participating in 25 periods in this part of the experiment. At the start of a new period, there is a **50 percent** chance that the bag you are drawing from is an M bag and a **50 percent** chance that the bag you are drawing from is a U bag. An M bag is therefore just as likely as a U bag to be chosen. This random choice of the bag is independently determined before each period, as if by flipping a coin. If the coin lands on "heads" an M bag is used, but if the coin lands on "tails" a U bag is used.

The 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**.

Period

Each period consists of many 'ticks'. In each tick, a ball may be drawn from the bag. Each tick lasts for a fifth of a second (i.e. five ticks per second), and ticks continually occur until the period ends. The period ends with a small probability each tick (3 out of 1000). Imagine 1000 tickets, numbered 1 through 1000 are placed in a box. It is as if, after every tick, a ticket is randomly drawn. If the number on the ticket is 1, 2, or 3, the period ends. If the number on the ticket is 4 through 1000, the period continues and the ticket is placed back in the box, so the contents of the box do not change. Under this probability, the average period length is 333.33 ticks or 66.66 seconds. Many periods will last less than the average, and a few will last much longer.

As mentioned above, in each tick, a ball may be drawn from the bag. You will be asked an initial question of whether or not you would like to draw a ball. If you initially choose to draw a ball, you will continually draw a ball in each tick until you decide to stop drawing a ball. Once you have chosen to stop drawing a ball, you can **no longer draw a ball** for the rest of the ticks in the period. 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, you can either draw a red ball or a yellow ball. If you draw a red ball, you obtain **2.50 experimental dollars**. Once you have obtained a red ball **you automatically draw a ball** for the rest of the ticks in the period. You **can obtain** more than one red ball in a period. For example, if you draw a red ball, you can obtain a payoff of 2.50, 5.00, 7.50 experimental dollars, etc. based on how many red balls you draw in that period. If you draw a yellow ball, you will not be compensated for it. Notice that 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 obtain **.01 experimental dollars** in the current tick and **.01 experimental dollars** in each of the remaining ticks.

Five periods from this part will be randomly chosen for payment. The periods chosen for you may differ from the periods chosen for others.

Conditional Probability

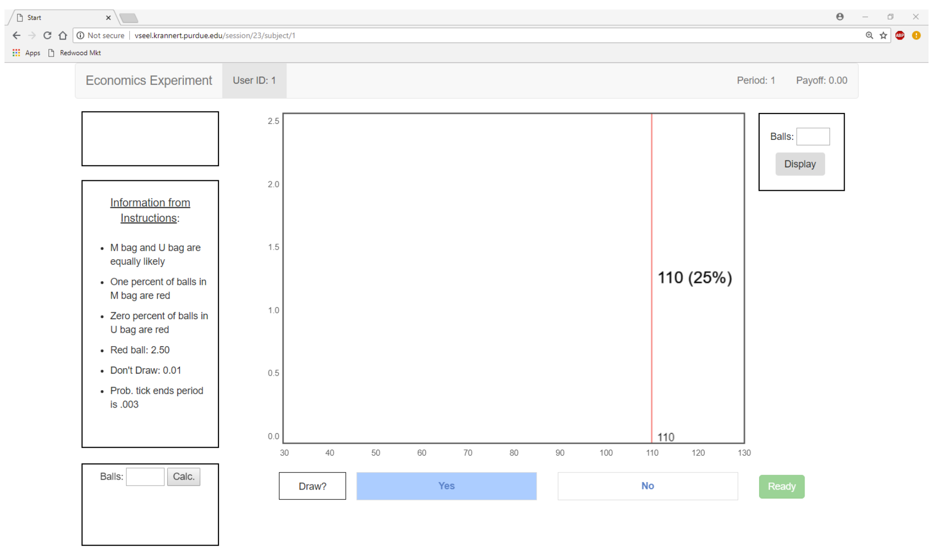
The interface displays a conditional probability when you have only drawn yellow balls. The conditional probability is the probability that you have the M bag based on the balls that you have already observed. While your bag is initially chosen by a coin flip, the balls you draw provide information about how likely you are to have the M bag. For example, if you were to observe 1,000 yellow balls and 0 red balls, you know that it is unlikely that you have the M bag. The conditional probability tells you how likely it is.

To illustrate how the conditional probability is calculated, consider a (hypothetical) subject who has drawn 110 yellow balls and 0 red balls. In this case, the conditional probability is equal to

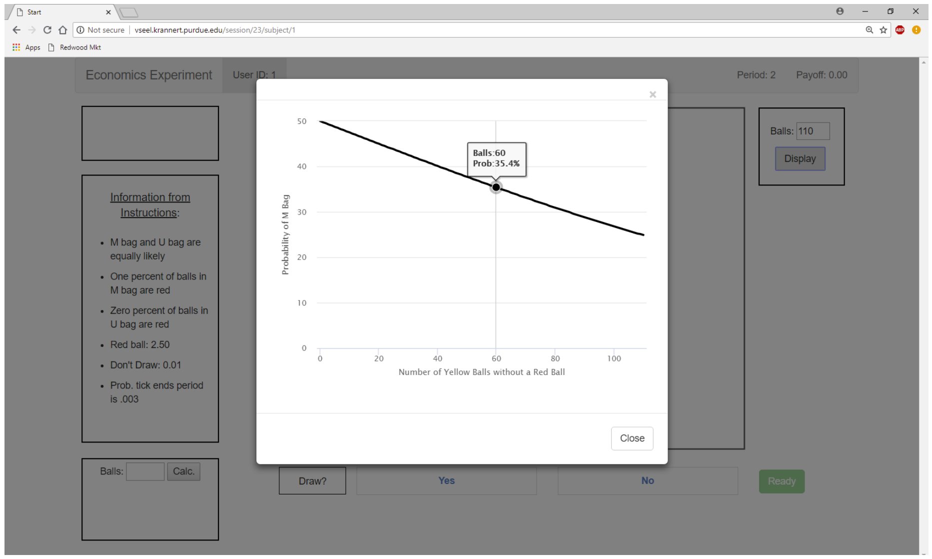
$$\frac{\text{P(You have an M bag and you draw 110 Yellow Balls and 0 Red Balls)}}{\text{P(You Draw 110 Yellow Balls and 0 Red balls)}}$$

This probability calculates how likely it is that the M bag is responsible for you observing 110 yellow balls and 0 red balls. This probability is equal to (approximately) .25 (= 25%), which implies that if you were to happen to observe a large number of periods with 110 yellow balls and 0 red balls, about 25 percent of these periods should have M bags.

The conditional probability can change as you draw more balls. The conditional probability will change when the balls you draw provide more information about which bag you are drawing from. As you draw, the conditional probability will continue to decrease if you have not yet drawn a red ball. If you have not yet drawn a red ball, each additional yellow ball is more evidence that is consistent with you having the U bag (the U bag has more yellow balls). You should only consider the current conditional probability and treat this probability as the probability that you have the M bag.



The experimental interface has two built-in features to display the conditional probability. The first feature is the calculator in the lower-left box. As an example, if you enter the number 110 into the form and click on the button, the box will display the conditional probability when you draw 110 yellow balls and 0 red balls. The second feature is in the upper-right box. As an example, if you enter the number 110 into the form and click on the "Display" button, a graph will display how the conditional probability evolves as you draw 110 yellow balls without a red ball. Both features display probabilities to the nearest tenth of a percentage point. The following image displays the graph.



Interface

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

Recap

The following is a recap of important parts of the instructions

- 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 beginning of each period, it is equally likely that an M bag or a U bag is chosen
- The conditional probability states the probability that you are drawing from the M bag based on the balls you have already drawn
- If you draw a red ball, you obtain 2.50 experimental dollars and automatically draw a ball in the remainder of the ticks in the period. You can obtain more than one red ball in a period

- If you draw a yellow ball, you are not compensated
- If you do not draw a ball, you obtain 0.01 experimental dollars in the current tick and each of the remaining ticks (you can no longer draw a ball for the rest of the ticks in the period)

Calculating the Conditional Probability

This page is provided to specifically show how the conditional probability is calculated. This page is provided for individuals who want to learn more about how the conditional probability is calculated. Reading this page is optional and will not affect your earnings as the computer automatically calculates the conditional probability in the case where you have only drawn yellow balls.

The conditional probability if you have drawn X yellow balls are drawn and 0 red balls is given by:

$$\frac{\text{P(You have an M bag and you draw X Yellow Balls and 0 Red Balls)}}{\text{P(You Draw X Yellow Balls and 0 Red Balls)}} = \frac{.5 \times .99^X}{.5 \times .99^X + .5}.$$

The probability that you have an M bag and that you have drawn X Yellow Balls and 0 Red Balls is given by $.5*.99^X$ because you have an initial probability of .5 of having an M bag and have a probability of $.99^X$ that the first X balls from an M bag are yellow. The denominator includes the probability that you have a U bag and that the first X balls from the U bag are yellow. This probability is equal to .5 because the first X balls from the U bag will always be yellow.

C. APPENDIX FOR: IS EXPERIMENTATION INVARIANT TO GROUP SIZE? A LABORATORY ANALYSIS OF INNOVATION CONTESTS

C.1 Theory Appendix

The theory appendix explains the details of the continuous time public winner-takes-all contest. This section is based heavily off of Halac et al. (2017).

In the continuous time public winner-takes-all contest, the full prize is awarded to the first agent who succeeds. Let $A_{i,z}$ denote (i's conjecture of) the aggregate effort exerted by i's opponents at time z as long as no agent has obtained a success by z. An agent i's problem can be written as

$$\overset{max}{(a_{i,t})_{t\in[0,T]}} \int_0^T [(p_{i,t}\lambda\overline{w} - c)a_{i,t}]e^{-\int_0^t p_{i,z}\lambda(a_{i,z}+A_{-i,z})dz}dt,$$

where $p_{i,t}$ is i's belief that the state is good at time t (as long as no success has been obtained), given by

$$p_{i,t} = \frac{p_0 e^{-\int_0^T \lambda(a_{i,z}+A_{-i,z})dz}}{p_0 e^{-\int_0^T \lambda(a_{i,z}+A_{-i,z})dz} + (1-p_0)}.$$

Note that $e^{-\int_0^t p_{i,z}\lambda(a_{i,z}+A_{-i,z})dz}$ is the agent's belief that no success will be obtained by time t.

Since the agent's belief $p_{i,t}$ is decreasing over time, the unique solution to this problem is $a_{i,t} = 1$ if $p_{i,t} \geq \frac{c}{\lambda\overline{w}}$ and $a_{i,t} = 0$ otherwise. It follows that in a continuous time public winner-takes-all contest with deadline T, there is a unique equilibrium. All agents exert effort until either a success is obtained or the public belief reaches $\frac{c}{\lambda\overline{w}}$ (or the contest ends). The probability of obtaining an innovation is independent of the number of agents as multiple agents succeeding at the same instant is second order.

Table C.1.: Mean statistics on the innovation percentage, the level of aggregate effort in bad states, and the level of individual effort in bad states. The equilibrium innovation percentage is found by averaging the predictions for each contest.

Treatment:	Two-Person			Four-Person		
	Equil.		Actual	Equil.		Actual
Innovation Percentage	0.70	$\overset{**}{>}$	0.66	0.70	$=$	0.70
Aggregate Effort in Bad States	20	$>$	18.25	20	$\overset{***}{<}$	25.42
Individual Effort in Bad States	10	$>$	9.13	5	$\overset{***}{<}$	6.35

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

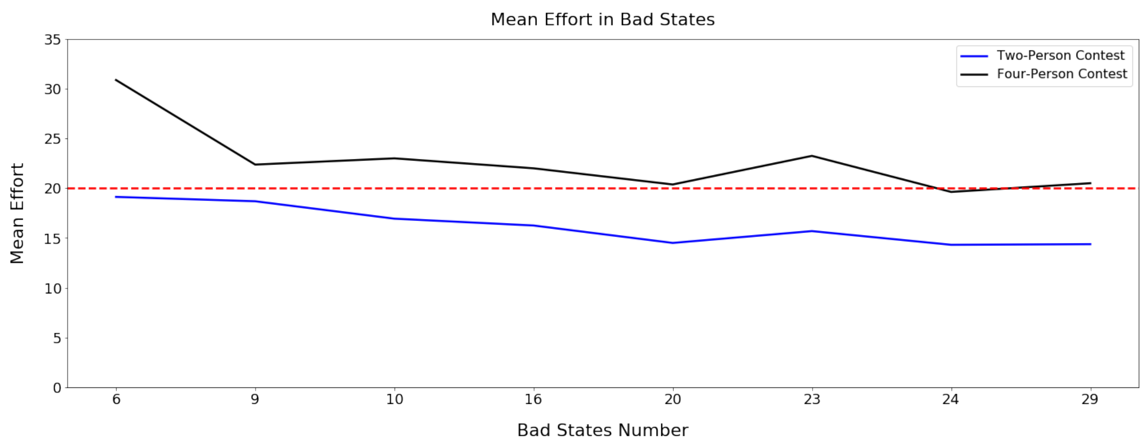


Figure C.1.: Mean level of aggregate effort in bad states for each treatment. The red dotted line displays the equilibrium prediction for aggregate effort.

C.2 First Ten Periods Appendix

Table C.1 displays the summary statistics for the first ten periods. In the two-person treatment, the innovation percentage is significantly less than the predicted 70 percent at the five percent level (p-value=.034). In the two-person treatment, the aggregate effort in bad states is not significantly different than the predicted 20 at the ten percent level. Both of these hypotheses are conducted with bootstrapped regressions, clustered at the session level, with 5000 bootstrap samples. The individual effort in bad states is not significantly different than the predicted 10 at the ten percent level. This hypothesis is conducted with a bootstrapped regression on subject means, with 5000 bootstrap samples, clustered at the session level.

In the four-person treatment, the innovation percentage is equal to the predicted 70 percent. In the four-person treatment, the aggregate effort in bad states is significantly greater than the predicted 20 at the one percent level. This hypothesis is conducted with bootstrapped regressions, clustered at the session level, with 5000 bootstrap samples. The individual effort in bad states is significantly greater than the predicted five at the one percent level. This hypothesis is conducted with a bootstrapped regression on subject means, with 5000 bootstrap samples, clustered at the session level.

The first hypothesis can once again be tested using a regression of the innovation percentage on the treatment. The difference of 3.75 percent is significant at the ten percent level using a bootstrapped regression with 5000 bootstrap samples, clustered at the session level (p-value=0.052). This provides some evidence for Hypothesis 1. The second hypothesis can be tested using a bootstrapped regression of the level of aggregate effort on the treatment. The difference of 7.17 is significant at the 1 percent level using a bootstrapped regression with 5000 bootstrap samples, clustered at the session level. This supports Hypothesis 2. The third hypothesis can once again be tested using a regression of the number of draws, in bad states, on the treatment. The difference of -2.38 is significant at the 1 percent level when using a bootstrapped regression of the subject means with 5000 bootstrap samples, clustered at the session level. This supports Hypothesis 3. These results back up the results from the last twenty contests.

C.3 Model With Risk Aversion

In this section, I will develop and estimate a model of differential weighting of experimentation under risk aversion. The model in this section only differs from subsection 3.4.4 in that subjects are risk averse. Subjects are assumed to have CRRA utility, that is, $u(x) = \frac{x^{1-r}}{1-r}$, where r is the level of risk aversion. The optimal stopping strategy under risk aversion is for an individual to exert effort as long as a prize has not been obtained and that the belief is greater than or equal to $\frac{u(c)}{\lambda u(\overline{w})}$.[1] The rest of the model follows from subsection 3.4.4. Belief updating, in the absence of an innovation, is still given by

$$\tilde{p} = \frac{p_0(1-\lambda)^{\psi_i D_{t-1}+\psi_o O_{t-1}}}{p_0(1-\lambda)^{\psi_i D_{t-1}+\psi_o O_{t-1}} + (1-p_0)}.$$

The loglikelihood for this model can be written by

$$LogL = \sum_{i=1}^{n} log\left[\prod_{contest=1}^{C}\prod_{period=1}^{P}\Phi\left(\tilde{p}\lambda u(\overline{w}) - u(c)\right)^{E}\left(1-\Phi\left(\tilde{p}\lambda u(\overline{w}) - u(c)\right)\right)^{1-E}\right].$$

This loglikelihood is maximized at a value of 3046.68. The estimated values of ψ_i and ψ_o are 2.67 and 2.42, respectively. These parameters suggest that subjects place less weight on their competitors' failed innovation attempts than their own failed innovation attempts. The estimated value of r is 0.09. This parameter suggests that

[1]This can be shown through the following argument. In period T, the value at belief p is given by $V(p,T) = max\{u(c), p\lambda u(\overline{w})\}$, where the first term in the maximand is the utility of not exerting effort and the second term in the maximand is the expected utility of exerting effort. Thus, when $p < \frac{u(c)}{\lambda u(\overline{w})}$, an individual does not exert effort and $V(p,T) = u(c)$. When $p > \frac{u(c)}{\lambda u(\overline{w})}$, an individual exerts effort and $V(P,T) > u(c)$. When $p = \frac{u(c)}{\lambda u(\overline{w})}$, an individual is indifferent between exerting effort and not exerting effort and $V(P,T) = u(c)$. Assume that an individual does not exert effort in period t and $V(p,t) = (T-t+1) * u(c)$ if $p < \frac{u(c)}{\lambda u(\overline{w})}$. Assume that an individual exerts effort in period t and $V(P,t) > (T-t+1) * u(c)$ if $p > \frac{u(c)}{\lambda u(\overline{w})}$. Assume that an individual exerts effort in period t and $V(P,t) = (T-t+1) * u(c)$ if $p = \frac{u(c)}{\lambda u(\overline{w})}$. In period t-1, the value at belief p is given by $V(p,t-1) = max\{(T-t+2) * u(c), p\lambda(u(\overline{w}) + (T-t+1) * u(c)) + (1-p\lambda) * V(p,t)\}$, which can be rewritten as $V(p,t-1) = max\{u(c), p\lambda(u(\overline{w})) + (1-p\lambda) * (V(p,t) - (T-t+1) * u(c))\}$. If $p > \frac{u(c)}{\lambda u(\overline{w})}$, an individual is better off exerting effort. If $p < \frac{u(c)}{\lambda u(\overline{w})}$, an individual is better off not exerting effort. If $p = \frac{u(c)}{\lambda u(\overline{w})}$, an individual is indifferent between exerting and not exerting effort.

subjects are risk averse, which is consistent with subsection 3.4.3. The restriction that $\psi_i = \psi_o$ is rejected at the one-percent level using a likelihood ratio test (restricted loglikelihood is equal to 3062.39). Additionally, the value of ψ_i is significantly greater than one at the one percent level using a likelihood ratio test (restricted loglikelihood is equal to 3153.10). The value of ψ_o is significantly greater than one at the one percent level using a likelihood ratio test (restricted loglikelihood is equal to 3112.18). The value of r is significantly greater than zero at the one percent level using a likelihood ratio test (restricted loglikelihood is equal to 3053.35). This model suggests that subject behavior is consistent with differential weighting of experimentation.

Differential weighting of experimentation is consistent with larger contests resulting in more innovations and inducing more aggregate effort. Assuming that subjects use the stopping strategy modeled in this section, subjects, in the absence of an innovation, are predicted to exert effort until

$$\frac{p_0(1-\lambda)^{2.67D_{t-1}+2.42O_{t-1}}}{p_0(1-\lambda)^{2.67D_{t-1}+2.42O_{t-1}}+(1-p_0)} < \frac{u(c)}{\lambda u(\overline{w})}.$$

As the number of subjects increase in a contest, O_{t-1} becomes a larger share of aggregate effort and subjects are thus more willing to experiment at a given level of aggregate effort (in the absence of an innovation).

C.4 Instructions

In this section, I will display the instructions for the four-person treatment. The instructions for the two-person treatment are similar to these instructions except that subjects are in four-person contests. The "part 2" referenced in this set of instructions is the risk elicitation task. Upon completion of the instructions, subjects went on to answer five comprehension questions.

Instructions

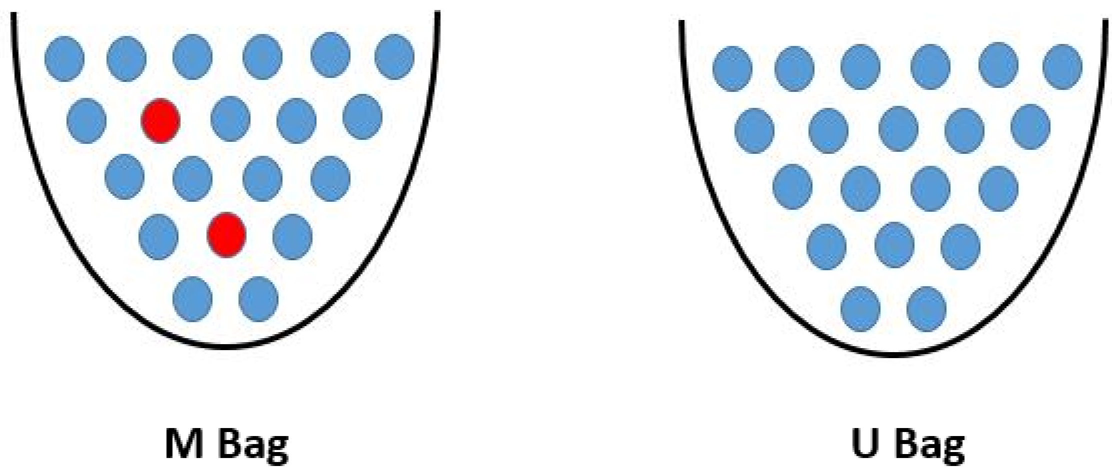
This experiment is a study of group and individual decision making. The amount of money you earn depends partly on the decisions that you make and thus you should read these instructions carefully. The money you earn will be paid privately to you, in cash, at the end of the experiment.

The experiment is divided into two parts. These are the instructions for Part 1. You will receive further instructions for Part 2 once Part 1 is completed. Please note that your decisions in Part 1 will in **no way** affect your earnings (or the earnings of others) in Part 2.

Background

At the start of the experiment, you will have the opportunity to earn $5.00 based on how you answer five comprehension questions about these instructions. A correct answer to a question is worth $1.00. You will be able to refer to these written instructions as you answer the questions.

In this experiment, imagine that there are two bags of balls. The first bag of balls is a 'mixed' bag (denoted by the letter M), which contains **2 red balls and 18 blue balls**. If you draw a ball from the M bag, there is a **10 percent** chance that you will draw a red ball and a **90 percent** chance that you will draw a blue ball. The second bag of balls is a 'uniform' bag (denoted by the letter U), which contains **0 red balls and 20 blue balls**. If you draw a ball from the 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 blue ball. The following images are a graphical representation of the two bags.



You will be participating in 30 repetitions (called 'cycles') of a game. Each cycle will consist of 15 periods. At the start of a new cycle, there is a **75 percent** chance that the bag you are drawing from is the M bag and a **25 percent** chance that the bag you are drawing from is the U bag. The M bag is therefore three times more likely than the U bag. This random choice of the bag is independently determined before each cycle, as if by throwing a 4-sided die. If a 1, 2, or 3 came up on the die then the M bag is used, but if a 4 came up on the die then the U bag is used.

The bag does not change within a cycle. Balls you draw in periods 1 through 15, within the same cycle, 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.

At the start of each period in a given cycle, you will have the opportunity to either draw a ball from a bag, or collect **$0.30**. If you choose to draw a ball from a bag, two things can occur: you can either draw a **red ball**, or you can draw a **blue**

ball. If you draw a red ball, **the cycle is over** and you will be compensated for the red ball. If you draw a blue ball, the cycle may or may not continue and you will **not** be compensated for the blue ball. The ball you draw will be put back into the bag.

At the start of each cycle, you will be randomly paired with three subjects. You will be paired with those subjects for each period of the cycle.

You and the subjects you are paired with will be drawing from four separate but identical bags. Thus, you and the subjects you are paired with will have **the same probability of drawing a red ball** in each period, because either you will all be drawing from M bags or will all be drawing from U bags. Any ball that is drawn from any bag will be put back in the bag at the end of the period. Thus, the composition of the bag you are drawing from remains the same in each period of the cycle. For example, if you were drawing from the M bag, you would have the same probability of drawing a red ball in each period of the cycle.

Compensation

At the start of each period in the cycle, you have the choice to draw a ball from the bag or to collect **$0.30**. If you draw a red ball from the bag you will be compensated for it. **A red ball is worth $10.00, regardless of what the subjects you are paired with do**. If you draw a red ball, you will receive $10.00 in that period. If you draw a blue ball, regardless of what the subjects you are paired with do, you will receive **$0.00** for the period.

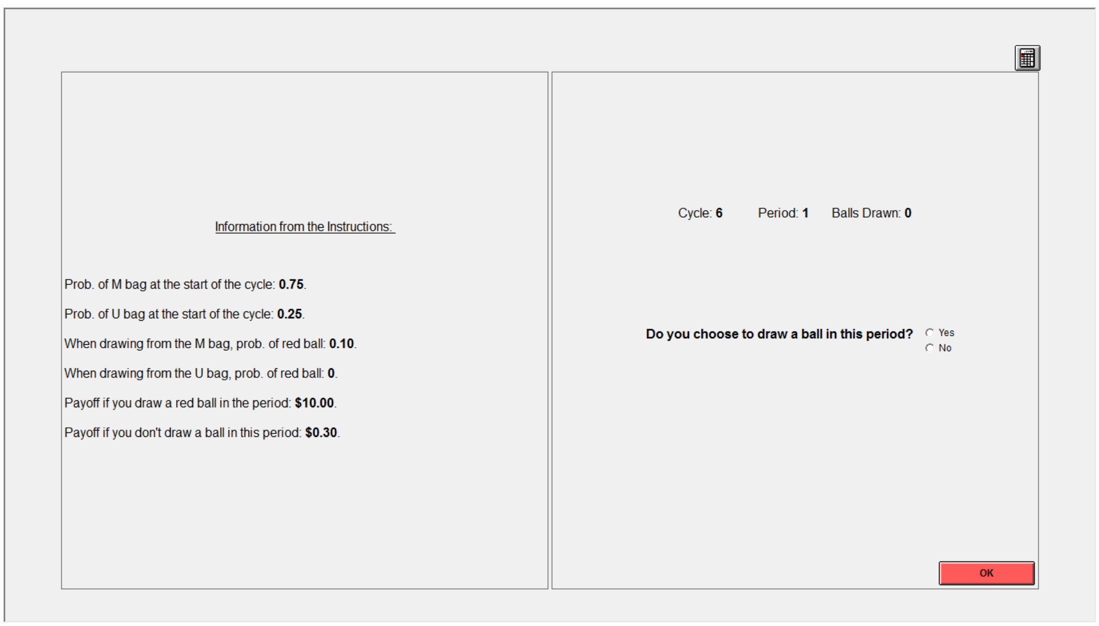
If you or any of the subjects you are paired with draws a red ball, the cycle will end. You **will** be compensated for the rest of the periods in the cycle. For example, if a red ball is drawn with eleven periods remaining in the cycle, your profit from the cycle will be the sum of the payoffs from the previous and current periods in the cycle and

$0.30 for each of the remaining eleven periods. To reiterate, you will receive the payoffs from the previous and current periods in the cycle and $3.30 ($0.30 for each of the remaining eleven periods).

Two of the thirty cycles will be randomly chosen for payment. The cycles chosen for you may be different than the ones drawn for other participants.

Example

Below is an example of what a period will look like.



In each period, you will get a summary of important information from the instructions on the left side of the screen. In the top right corner of the screen is a calculator button. If you would like to make calculations before you decide to draw a ball, you can click on the button and a calculator will show up. The right side of the screen also has information with regards to the current cycle, the current period, and the amount of balls drawn in the current cycle by **you and the subjects you are paired with**. In the center of the right side of the screen is a prompt asking you if you choose to draw a ball in the current period. Once you have made your decision as to whether or not you will draw a ball in the current period, press the "OK" button to move on. You will have as much time as you like to make your decision.

At the end of the cycle, you will get information about the cycle that just ended. You will be told whether or not a red ball has been drawn, how many people drew a red ball and how many balls were drawn in the cycle. You will receive a breakdown of your payoff from the current cycle in terms of the periods you didn't draw a ball and the periods that you did draw a ball.

Recap

The following is a recap of important parts of the instructions

- There are two bags: a 'mixed' (M) bag with 2 red balls and 18 blue balls and a 'uniform' (U) bag with 0 red balls and 20 blue balls
- For each cycle, there is a 75 percent chance you and the subjects that you are paired with are drawing from M bags and a 25 percent chance that you and the subjects that you are paired with are drawing from U bags
- In each period, you can either draw a ball or collect $0.30
- If you draw a red ball you earn $10.00 for the current period
- The cycle ends as soon as someone draws a red ball or at the end of 15 periods
- In the case that the cycle ends before 15 periods are completed, you will be compensated for the remaining periods in the cycle

D. APPENDIX FOR: PUBLIC LEADERBOARD FEEDBACK IN INNOVATION CONTESTS: A THEORETICAL AND EXPERIMENTAL INVESTIGATION

D.1 SPNE for Finite-horizon Leaderboard-Feedback Innovation Contest

In this appendix, we describe the process for characterizing the subgame perfect Nash equilibria of the finite horizon leaderboard-feedback innovation contest. Recall that f_t (l_t) denotes the follower (leader) in an arbitrary period t. We begin by characterizing the final-stage local equilibrium strategies and corresponding equilibrium expected payoffs, and then make our way back through the game tree. We assume that: (i) utility is time separable and (ii) the utility $u(\cdot)$ in each period displays constant absolute risk aversion (CARA), where for convenience we set $u(x) = \left(1 - e^{-xR}\right)R$ for $R > 0$ and $u(x) = x$ for $R = 0$. Although our focus in this appendix is on a utility function that displays risk aversion, it is straightforward to extend the analysis below to allow for loss aversion and sunk-cost fallacy considerations.

Period T

Let p_{l_T} [p_{f_T}] denote the probability that the period T leader l_T [period T follower f_T] draws in period T, and let $\pi_{f_T}(D, p_{l_T}|s_T)$ denote the the payoff to the period T follower f_T from drawing in period T given p_{l_T} and the score s_T. In the final period T, if the max score at the beginning of period T is s_T, then the benefit to the period

T follower from drawing (i.e. $p_{f_T} = 1$) when the period T leader does not draw (i.e. $p_{l_T} = 0$) is

$$\pi_{f_T}(D, p_{l_T} = 0|s_T) = (1 - F(s_T))u(v - c) + F(s_T)u(-c). \tag{D.1}$$

Next, the benefit to the period T follower from drawing when the period T leader does draw is

$$\pi_{f_T}(D, p_{l_T} = 1|s_T) = \left[\frac{1 - [F(s_T)]^2}{2}\right] u(v - c) + \left[\frac{1 + [F(s_T)]^2}{2}\right] u(-c). \tag{D.2}$$

Thus, at the beginning of period T and given any $p_{l_T} \in [0, 1]$, we have that

$$\pi_{f_T}(D, p_{l_T}|s_T) = (1 - p_{l_T})\pi_{f_T}(D, p_{l_T} = 0|s_T) + p_{l_T}\pi_{f_T}(D, p_{l_T} = 1|s_T). \tag{D.3}$$

For all $p_{l_T} \in [0, 1]$, the payoff to the period T follower from not drawing in period T, denoted $\pi_{f_T}(ND, p_{l_T}|s_T)$, is 0.

For the characterization of when player f_T is indifferent between drawing and not drawing as a function of the beginning of period T leader score s_T and the leader's final-stage-local strategy p_{l_T}, it will be convenient to refer to the change in player f_T's payoff in moving from drawing to not drawing given that either $p_{l_T} = 0$ or $p_{l_T} = 1$, which we denote by $\Delta\pi_{f_T}(p_{l_T} = 0|s_T)$ and $\Delta\pi_{f_T}(p_{l_T} = 1|s_T)$ respectively, where

$$\Delta\pi_{f_T}(p_{l_T} = 0|s_T) = \pi_{f_T}(ND, p_{l_T} = 0|s_T) - \pi_{f_T}(D, p_{l_T} = 0|s_T) \tag{D.4}$$

and

$$\Delta\pi_{f_T}(p_{l_T} = 1|s_T) = \pi_{f_T}(ND, p_{l_T} = 1|s_T) - \pi_{f_T}(D, p_{l_T} = 1|s_T) \tag{D.5}$$

If

$$\frac{\pi_{f_T}(D, p_{l_T} = 0|s_T)}{\pi_{f_T}(D, p_{l_T} = 0|s_T) - \pi_{f_T}(D, p_{l_T} = 1|s_T)} \in [0, 1]$$

then for

$$p_{l_T}^{indiff} = \frac{\Delta\pi_{f_T}(p_{l_T}=0|s_T)}{\Delta\pi_{f_T}(p_{l_T}=0|s_T) - \Delta\pi_{f_T}(p_{l_T}=1|s_T)} = \frac{(1-F(s_T))u(v-c) + F(s_T)u(-c)}{(u(v-c)-u(-c))\,\frac{1}{2}\,(1-[F(s_T)]^2)} \tag{D.6}$$

it follows from equation (D.3) that

$$\pi_{f_T}(D, p_{l_T}^{indiff}|s_T) = \pi_{f_T}(ND, p_{l_T}^{indiff}|s_T) = 0$$

and the period T follower is indifferent between drawing and not drawing. Because $\Delta\pi_{f_T}(p_{l_T} = 0|s_T) \leq \Delta\pi_{f_T}(p_{l_T} = 1|s_T)$, it follows that if $\Delta\pi_{f_T}(p_{l_T} = 0|s_T) = -\pi_{f_T}(D, p_{l_T} = 0|s_T) > 0$, then player f_T would have incentive to not draw for all $p_{l_T} \in [0,1]$. Similarly, if $\Delta\pi_{f_T}(p_{l_T} = 1|s_T) = -\pi_{f_T}(D, p_{l_T} = 1|s_T) < 0$, then player f_T would have incentive to draw for all $p_{l_T} \in [0,1]$. Thus, it follows that for the term $p_{l_T}^{indiff}$ defined by equation (D.6) to take values in the interval $[0,1]$, it must be the case that $\Delta\pi_{f_T}(p_{l_T} = 0|s_T) = -\pi_{f_T}(D, p_{l_T} = 0|s_T) \leq 0$ and $\Delta\pi_{f_T}(p_{l_T} = 1|s_T) = -\pi_{f_T}(D, p_{l_T} = 1|s_T) \geq 0$, or equivalently, $F(s_T) \in \left[\sqrt{\frac{u(v-c)+u(-c)}{u(v-c)-u(-c)}}, \frac{u(v-c)}{u(v-c)-u(-c)}\right]$.[1]

For the purpose of stating player f_T's final-stage-local best-response correspondence as a function of $(p_{l_T}, s_T) \in [0,1] \times \text{supp}(F)$, let

$$\Sigma_{f_T}^{indiff} = \left\{ s_T \middle| \Delta\pi_{f_T}(p_{l_T} = 0|s_T) \leq 0 \text{ and } \Delta\pi_{f_T}(p_{l_T} = 1|s_T) \geq 0 \right\}$$

denote the set of period T beginning scores s_T such that $p_{l_T}^{indiff} \in [0,1]$. Similarly, let

$$\Sigma_{f_T}^{1} = \left\{ s_T \middle| \Delta\pi_{f_T}(p_{l_T} = 1|s_T) < 0 \right\}$$

and let

$$\Sigma_{f_T}^{0} = \left\{ s_T \middle| \Delta\pi_{f_T}(p_{l_T} = 0|s_T) > 0 \right\}$$

[1]Note that in the case of risk neutrality, the equation (D.6) expression for $p_{l_T}^{indiff}$ becomes $p_{l_T}^{indiff} = \frac{v(1-F(s_T))-c}{\frac{v}{2}(1-F(s_T))^2}$ which takes values in $[0,1]$ when $F(s_T) \in \left[\sqrt{1-\frac{2c}{v}}, 1-\frac{c}{v}\right]$.

and note that $\Sigma_{f_T}^{indiff}$, $\Sigma_{f_T}^{1}$, and $\Sigma_{f_T}^{0}$ form a partition of supp(F). Player f_T's final-stage-local best-response correspondence is given by:

$$BR_{f_T}(p_{l_T}|s_T) = \begin{cases} p_{f_T} = 1 & \text{if } s_T \in \Sigma_{f_T}^{1} \\ & \text{or } s_T \in \Sigma_{f_T}^{indiff} \text{ and } p_{l_T} < p_{l_T}^{indiff} \\ & \\ p_{f_T} \in [0,1] & \text{if } s_T \in \Sigma_{f_T}^{indiff} \text{ and } p_{l_T} = p_{l_T}^{indiff} \\ & \\ p_{f_T} = 0 & \text{if } s_T \in \Sigma_{f_T}^{0} \\ & \text{or } s_T \in \Sigma_{f_T}^{indiff} \text{ and } p_{l_T} > p_{l_T}^{indiff} \end{cases} \tag{D.7}$$

Moving on to the period T leader's problem, the payoff to the period T leader from not drawing when the period T follower draws is

$$\pi_{l_T}(ND, p_{f_T} = 1|s_T) = F(s_T)u(v)$$

verses a payoff of

$$\pi_{l_T}(D, p_{f_T} = 1|s_T) = \left[\frac{1 + [F(s_T)]^2}{2}\right] u(v-c) + \left[\frac{1 - [F(s_T)]^2}{2}\right] u(-c).$$

when both the period T and the period T follower draw. Similarly, the payoff to the period T leader from not drawing when the period T follower does not draw is

$$\pi_{l_T}(ND, p_{f_T} = 0|s_T) = u(v)$$

verses a payoff of

$$\pi_{l_T}(D, p_{f_T} = 0|s_T) = u(v-c)$$

from drawing. Thus, the payoff to the period T leader from drawing in period T given any $p_{f_T} \in [0,1]$, denoted $\pi_{l_T}(D, p_{f_T}|s_T)$ is

$$\pi_{l_T}(D, p_{f_T}|s_T) = (1 - p_{f_T})\pi_{l_T}(D, p_{f_T} = 0|s_T) + p_{f_T}\pi_{l_T}(D, p_{f_T} = 1|s_T) \tag{D.8}$$

and the payoff to the period T leader from not drawing in period T, denoted $\pi_{l_T}(ND, p_{f_T}|s_T)$ is

$$\pi_{l_T}(ND, p_{f_T}|s_T) = (1 - p_{f_T})\pi_{l_T}(ND, p_{f_T} = 0|s_T) + p_{f_T}\pi_{l_T}(ND, p_{f_T} = 1|s_T). \tag{D.9}$$

To define $p_{f_T}^{indiff}$, we use the expressions $\Delta\pi_{l_T}(p_{f_T} = 0|s_T)$ and $\Delta\pi_{l_T}(p_{f_T} = 1|s_T)$ where

$$\Delta\pi_{l_T}(p_{f_T} = 0|s_T) = \pi_{l_T}(ND, p_{f_T} = 0|s_T) - \pi_{l_T}(D, p_{f_T} = 0|s_T) \tag{D.10}$$

and

$$\Delta\pi_{l_T}(p_{f_T} = 1|s_T) = \pi_{l_T}(ND, p_{f_T} = 1|s_T) - \pi_{l_T}(D, p_{f_T} = 1|s_T). \tag{D.11}$$

It follows from equations (D.8) and (D.9), that if

$$\frac{\pi_{l_T}(ND, p_{f_T} = 0|s_T) - \pi_{l_T}(D, p_{f_T} = 0|s_T)}{[\pi_{l_T}(ND, p_{f_T} = 0|s_T) - \pi_{l_T}(D, p_{f_T} = 0|s_T)] - [\pi_{l_T}(ND, p_{f_T} = 1|s_T) - \pi_{l_T}(D, p_{f_T} = 1|s_T)]} \in [0,1]$$

then for

$$\begin{aligned} p_{f_T}^{indiff} &= \frac{\Delta\pi_{l_T}(p_{f_T} = 0|s_T)}{\Delta\pi_{l_T}(p_{f_T} = 0|s_T) - \Delta\pi_{l_T}(p_{f_T} = 1|s_T)} \\ &= \frac{u(v) - u(v-c)}{(1 - F(s_T))u(v) - (u(v-c) - u(-c))\,\frac{1}{2}\,(1 - [F(s_T)]^2)} \end{aligned} \tag{D.12}$$

it follows from equations (D.8) and (D.9) that

$$\pi_{l_T}(D, p_{f_T}^{indiff}|s_T) = \pi_{l_T}(ND, p_{f_T}^{indiff}|s_T) = 0$$

and the period T leader is indifferent between drawing and not drawing.

Next, because $\Delta\pi_{l_T}(p_{f_T} = 0|s_T) \geq \max\{0, \Delta\pi_{l_T}(p_{f_T} = 1|s_T)\}$, it follows that if $\Delta\pi_{l_T}(p_{f_T} = 1|s_T) > 0$ then for all $p_{f,T} \in [0,1]$ player l_T would have incentive to not draw. For the term p_{f_T} defined by equation (D.12) to take values in the interval $(0,1)$, it must be the case that $\Delta\pi_{l_T}(p_{f_T} = 1|s_T) \leq 0$.

In a manner similar to that used above for player f_T's final-stage-local best-response correspondence, we let

$$\Sigma_{l_T}^{indiff} = \left\{ s_T \middle| \Delta\pi_{l_T}(p_{f_T} = 1|s_T) \leq 0 \right\}$$

denote the set of period T beginning scores s_T such that $p_{l_T}^{indiff} \in [0,1]$. Similarly, let

$$\Sigma_{l_T}^{0} = \left\{ s_T \middle| \Delta\pi_{l_T}(p_{f_T} = 1|s_T) > 0 \right\}$$

and note that $\Sigma_{f_T}^{indiff}$ and $\Sigma_{f_T}^{0}$ form a partition of $\text{supp}(F)$. Then, the period T leader's final-stage local best-response correspondence as a function of $(p_{f_T}, s_T) \in [0,1] \times \text{supp}(F)$ may be written as,

$$BR_{l_T}(p_{f_T}|s_T) = \begin{cases} p_{l_T} = 1 & \text{if } s_T \in \Sigma_{l_T}^{indiff} \text{ and } p_{f_T} > p_{f_T}^{indiff} \\ \\ p_{l_T} \in [0,1] & \text{if } s_T \in \Sigma_{l_T}^{indiff} \text{ and } p_{f_T} = p_{f_T}^{indiff} \\ \\ p_{l_T} = 0 & \text{if } s_T \in \Sigma_{l_T}^{0} \\ & \text{or } s_T \in \Sigma_{l_T}^{indiff} \text{ and } p_{f_T} < p_{f_T}^{indiff} \end{cases} \tag{D.13}$$

Combining the period T follower's final-stage-local best-response correspondence from equation (D.7) with the period T leader's final-stage-local best-response correspondence from equation (D.13), we can now solve for the subgame perfect final-stage-local equilibrium strategies.

First note that because $\Delta\pi_{f_T}(p_{l_T} = 1|s_T) \geq 0$ implies that $\Delta\pi_{l_T}(p_{f_T} = 1|s_T) \geq 0$, it follows that $\Sigma_{l_T}^{indiff} \cap \Sigma_{f_T}^{indiff} = \emptyset$ and thus, there exists no non-degenerate final-stage-local equilibrium. Futhermore, note that $\Sigma_{l_T}^{indiff} \subset \Sigma_{f_T}^{1}$ and that $\Sigma_{f_T}^{indiff} \subset \Sigma_{l_T}^{0}$. For final-stage-local pure-strategy equilibria, we have the following:

$$\begin{cases} \text{Both draw} & \text{if } s_T \in \Sigma_{l_T}^{indiff} \subset \Sigma_{f_T}^{1} \\ \text{only follower draws} & \text{if } s_T \in \Sigma_{l_T}^{0} \cap \Sigma_{f_T}^{1} \\ \text{neither draws} & \text{if } s_T \in \Sigma_{l_T}^{0} \cap \left(\Sigma_{f_T}^{0} \cup \Sigma_{f_T}^{indiff}\right) \end{cases}$$

Note that there exists an $\overline{s}_{B,T} \in [0,1]$ such that the set $\Sigma_{l_T}^{indiff} \subset \Sigma_{f_T}^{1}$ is equivalent to $[0, \overline{s}_{B,T}]$. Similarly, there exists a $\underline{s}_{N,T} \in [0,1]$ such that the set $\Sigma_{l_T}^{0} \cap \left(\Sigma_{f_T}^{0} \cup \Sigma_{f_T}^{indiff}\right)$ is equivalent to $[\underline{s}_{N,T}, 1]$. The remaining set $\Sigma_{l_T}^{0} \cap \Sigma_{f_T}^{1}$ is equivalent to $[\overline{s}_{B,T}, \underline{s}_{N,T}]$. At the points where there exist multiple equilibria (i.e. $\overline{s}_{B,T}$ and $\underline{s}_{N,T}$) we will make the simplifying assumption that the player that is indifferent between drawing and not drawing chooses to draw. That is, at $s_T = \overline{s}_{B,T}$ we focus on the final-stage-local equilibrium in which both player's draw and at $s_T = \underline{s}_{N,T}$ we focus on the final-stage-local equilibrium in which player f_T draws. Given $\overline{s}_{B,T}$ and $\underline{s}_{N,T}$, the final-stage-local equilibria may be characterized as:

$$\begin{cases} \text{Both draw} & \text{if } s_T \in [0, \overline{s}_{B,T}] \\ \text{only follower draws} & \text{if } s_T \in (\overline{s}_{B,T}, \underline{s}_{N,T}] \\ \text{neither draws} & \text{if } s_T \in (\underline{s}_{N,T}, 1] \end{cases}$$

The corresponding subgame perfect final-stage local equilibrium expected payoffs for the leader and follower, respectively, are

$$\begin{cases} \pi_{l_T}(D, p_{f_T} = 1|s_T) \;\&\; \pi_{f_T}(D, p_{l_T} = 1|s_T) & \text{if } s_T \in [0, \overline{s}_{B,T}] \\ \pi_{l_T}(ND, p_{f_T} = 1|s_T) \;\&\; \pi_{f_T}(D, p_{l_T} = 0|s_T) & \text{if } s_T \in (\overline{s}_{B,T}, \underline{s}_{N,T}] \\ \pi_{l_T}(ND, p_{f_T} = 0|s_T) \;\&\; \pi_{f_T}(ND, p_{l_T} = 0|s_T) & \text{if } s_T \in (\underline{s}_{N,T}, 1] \end{cases}$$

Periods 1 to $T-1$

In moving from period T to any period $t \in \{1, \ldots, T-1\}$, the procedure for calculating the subgame perfect period-t-local equilibrium strategies and payoffs follows along the exact same lines as in period T given the changes to the expressions $\pi_{f_t}(p_{f_t}, p_{l_t}|s_t)$ and $\pi_{l_t}(p_{l_t}, p_{f_t}|s_t)$ respectively. In particular, for each period $t \in \{1, \ldots, T-1\}$ we take the period $t+1$ continuation payoffs as given and then calculate $\pi_{f_t}(p_{f_t}, p_{l_t}|s_t)$ and $\pi_{l_t}(p_{l_t}, p_{f_t}|s_t)$. Note that in the case of $t \in \{1, \ldots, T-1\}$, there are twelve possible transitions to consider:

Outcome	in $t+1$			s_{t+1} is such that:
	State	Leader $[l_{t+1}]$	Draws $\|s_{t+1}$	
O_1	$s_{t+1} = s_t$	l_t	Neither	$BR_{l_{t+1}}(ND\|s_{t+1}) = ND$ & $BR_{f_{t+1}}(ND\|s_{t+1}) = ND$
O_2	$s_{t+1} = s_t$	l_t	f_{t+1}	$BR_{l_{t+1}}(D\|s_{t+1}) = ND$ & $BR_{f_{t+1}}(ND\|s_{t+1}) = D$
O_3	$s_{t+1} = s_t$	l_t	l_{t+1}	$BR_{l_{t+1}}(ND\|s_{t+1}) = D$ & $BR_{f_{t+1}}(D\|s_{t+1}) = ND$
O_4	$s_{t+1} = s_t$	l_t	Both	$BR_{l_{t+1}}(D\|s_{t+1}) = D$ & $BR_{f_{t+1}}(D\|s_{t+1}) = D$
O_5	$s_{t+1} > s_t$	l_t	Neither	$BR_{l_{t+1}}(ND\|s_{t+1}) = ND$ & $BR_{f_{t+1}}(ND\|s_{t+1}) = ND$
O_6	$s_{t+1} > s_t$	l_t	f_{t+1}	$BR_{l_{t+1}}(D\|s_{t+1}) = ND$ & $BR_{f_{t+1}}(ND\|s_{t+1}) = D$
O_7	$s_{t+1} > s_t$	l_t	l_{t+1}	$BR_{l_{t+1}}(ND\|s_{t+1}) = D$ & $BR_{f_{t+1}}(D\|s_{t+1}) = ND$
O_8	$s_{t+1} > s_t$	l_t	Both	$BR_{l_{t+1}}(D\|s_{t+1}) = D$ & $BR_{f_{t+1}}(D\|s_{t+1}) = D$
O_9	$s_{t+1} > s_t$	f_t	Neither	$BR_{l_{t+1}}(ND\|s_{t+1}) = ND$ & $BR_{f_{t+1}}(ND\|s_{t+1}) = ND$
O_{10}	$s_{t+1} > s_t$	f_t	f_{t+1}	$BR_{l_{t+1}}(D\|s_{t+1}) = ND$ & $BR_{f_{t+1}}(ND\|s_{t+1}) = D$
O_{11}	$s_{t+1} > s_t$	f_t	l_{t+1}	$BR_{l_{t+1}}(ND\|s_{t+1}) = D$ & $BR_{f_{t+1}}(D\|s_{t+1}) = ND$
O_{12}	$s_{t+1} > s_t$	f_t	Both	$BR_{l_{t+1}}(D\|s_{t+1}) = D$ & $BR_{f_{t+1}}(D\|s_{t+1}) = D$

Note that although O_3, O_7 and O_{11} do not arise in equilibrium [i.e. there exists no t with a period-t-local equilibrium in which only the leader draws], we include that here as a possibility. Also observe that in states O_5-O_8 it must be the case that l_t draws and in states O_9-O_{12} it must be the case that f_t draws.

For the period-t follower we have:

$$\begin{aligned}
\pi_{f_t}(D, p_{l_t} = 0|s_t) = \mathrm{Prob}(O_1|s_t, D, p_{l_t} = 0)E\left(\pi_{f_{t+1}}(ND, p_{l_{t+1}} = 0|s_{t+1})|O_1\right)\\
\mathrm{Prob}(O_2|s_t, D, p_{l_t} = 0)E\left(\pi_{f_{t+1}}(D, p_{l_{t+1}} = 0|s_{t+1})|O_2\right)\\
+\mathrm{Prob}(O_3|s_t, D, p_{l_t} = 0)E\left(\pi_{f_{t+1}}(ND, p_{l_{t+1}} = 1|s_{t+1})|O_3\right)\\
+\mathrm{Prob}(O_4|s_t, D, p_{l_t} = 0)E\left(\pi_{f_{t+1}}(D, p_{l_{t+1}} = 1|s_{t+1})|O_4\right)\\
+\mathrm{Prob}(O_9|s_t, D, p_{l_t} = 0)E\left(\pi_{l_{t+1}}(ND, p_{f_{t+1}} = 0|s_{t+1})|O_5\right)\\
\mathrm{Prob}(O_{10}|s_t, D, p_{l_t} = 0)E\left(\pi_{l_{t+1}}(ND, p_{f_{t+1}} = 1|s_{t+1})|O_6\right)\\
+\mathrm{Prob}(O_{11}|s_t, D, p_{l_t} = 0)E\left(\pi_{l_{t+1}}(D, p_{f_{t+1}} = 0|s_{t+1})|O_7\right)\\
+\mathrm{Prob}(O_{12}|s_t, D, p_{l_t} = 0)E\left(\pi_{l_{t+1}}(D, p_{l_{t+1}} = 1|s_{t+1})|O_8\right)
\end{aligned} \tag{D.14}$$

$$
\begin{aligned}
\pi_{f_t}(D, p_{l_t} = 1|s_t) = {} & \mathrm{Prob}(O_1|s_t, D, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(ND, p_{l_{t+1}} = 0|s_{t+1})|O_1\right) \\
& \mathrm{Prob}(O_2|s_t, D, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(D, p_{l_{t+1}} = 0|s_{t+1})|O_2\right) \\
& + \mathrm{Prob}(O_3|s_t, D, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(ND, p_{l_{t+1}} = 1|s_{t+1})|O_3\right) \\
& + \mathrm{Prob}(O_4|s_t, D, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(D, p_{l_{t+1}} = 1|s_{t+1})|O_4\right) \\
& + \mathrm{Prob}(O_5|s_t, D, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(ND, p_{l_{t+1}} = 0|s_{t+1})|O_5\right) \\
& \mathrm{Prob}(O_6|s_t, D, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(D, p_{l_{t+1}} = 0|s_{t+1})|O_6\right) \\
& + \mathrm{Prob}(O_7|s_t, D, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(ND, p_{l_{t+1}} = 1|s_{t+1})|O_7\right) \\
& + \mathrm{Prob}(O_8|s_t, D, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(D, p_{l_{t+1}} = 1|s_{t+1})|O_8\right) \\
& + \mathrm{Prob}(O_9|s_t, D, p_{l_t} = 1)E\left(\pi_{l_{t+1}}(ND, p_{f_{t+1}} = 0|s_{t+1})|O_5\right) \\
& \mathrm{Prob}(O_{10}|s_t, D, p_{l_t} = 1)E\left(\pi_{l_{t+1}}(ND, p_{f_{t+1}} = 1|s_{t+1})|O_6\right) \\
& + \mathrm{Prob}(O_{11}|s_t, D, p_{l_t} = 1)E\left(\pi_{l_{t+1}}(D, p_{f_{t+1}} = 0|s_{t+1})|O_7\right) \\
& + \mathrm{Prob}(O_{12}|s_t, D, p_{l_t} = 1)E\left(\pi_{l_{t+1}}(D, p_{l_{t+1}} = 1|s_{t+1})|O_8\right)
\end{aligned}
\tag{D.15}
$$

$$
\begin{aligned}
\pi_{f_t}(ND, p_{l_t} = 0|s_t) = {} & \mathrm{Prob}(O_1|s_t, ND, p_{l_t} = 0)E\left(\pi_{f_{t+1}}(ND, p_{l_{t+1}} = 0|s_{t+1})|O_1\right) \\
& \mathrm{Prob}(O_2|s_t, ND, p_{l_t} = 0)E\left(\pi_{f_{t+1}}(D, p_{l_{t+1}} = 0|s_{t+1})|O_2\right) \\
& + \mathrm{Prob}(O_3|s_t, ND, p_{l_t} = 0)E\left(\pi_{f_{t+1}}(ND, p_{l_{t+1}} = 1|s_{t+1})|O_3\right) \\
& + \mathrm{Prob}(O_4|s_t, ND, p_{l_t} = 0)E\left(\pi_{f_{t+1}}(D, p_{l_{t+1}} = 1|s_{t+1})|O_4\right)
\end{aligned}
\tag{D.16}
$$

$$
\begin{aligned}
\pi_{f_t}(ND, p_{l_t} = 1|s_t) = \text{Prob}(O_1|s_t, ND, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(ND, p_{l_{t+1}} = 0|s_{t+1})|O_1\right) \\
\text{Prob}(O_2|s_t, ND, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(D, p_{l_{t+1}} = 0|s_{t+1})|O_2\right) \\
+ \text{Prob}(O_3|s_t, ND, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(ND, p_{l_{t+1}} = 1|s_{t+1})|O_3\right) \\
+ \text{Prob}(O_4|s_t, ND, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(D, p_{l_{t+1}} = 1|s_{t+1})|O_4\right) \\
+ \text{Prob}(O_5|s_t, ND, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(ND, p_{l_{t+1}} = 0|s_{t+1})|O_5\right) \\
\text{Prob}(O_6|s_t, ND, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(D, p_{l_{t+1}} = 0|s_{t+1})|O_6\right) \\
+ \text{Prob}(O_7|s_t, ND, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(ND, p_{l_{t+1}} = 1|s_{t+1})|O_7\right) \\
+ \text{Prob}(O_8|s_t, ND, p_{l_t} = 1)E\left(\pi_{f_{t+1}}(D, p_{l_{t+1}} = 1|s_{t+1})|O_8\right)
\end{aligned} \tag{D.17}
$$

Given the expressions in equations (D.14)-(D.17) for the period-t follower and the corresponding calculations for the period-t leader, the period-t-local equilibrium can be calculated by: (i) forming the period-t version of the 'Δ' expressions in equations (D.4), (D.5), (D.10), and (D.11), (ii) using the period-t version of the 'Δ' expressions to form the period t indifference conditions (D.6) and (D.12) and construct each player's period-t-local best-response correspondences as in equations (D.13) and (D.7), and (iii), using the player's period-t-local best-response correspondences characterize the period-t-local equilibrium.

As an example, consider the case of $t = T - 1$. Recall the characterization of the final-stage-local pure-strategy equilibrium:

$$
\begin{cases}
\text{Both draw} & \text{if } s_T \in [0, \overline{s}_{B,T}] \\
\text{only follower draws} & \text{if } s_T \in (\overline{s}_{B,T}, \underline{s}_{N,T}] \\
\text{neither draws} & \text{if } s_T \in (\underline{s}_{N,T}, 1]
\end{cases}
$$

Note that in period $T-1$, we know that there exists no period T equilibrium in which only l_T draws. Thus, there is no possible transition from state $T-1$ to state T in the form of outcomes O_3, O_7, and O_{11}.

If the max score at the beginning of period $T-1$ is s_{T-1}, then the probabilities $\text{Prob}(O_j|\cdot)$, for $j = 1, \ldots, 12$ in equation (D.14) are given by:

$$\text{Prob}(O_1|s_{T-1}, D, p_{l_{T-1}} = 0) = \begin{cases} F(s_{T-1}) & \text{if } s_{T-1} \in (\underline{s}_{N,T}, 1] \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Prob}(O_2|s_{T-1}, D, p_{l_{T-1}} = 0) = \begin{cases} F(s_{T-1}) & \text{if } s_{T-1} \in (\overline{s}_{B,T}, \underline{s}_{N,T}] \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Prob}(O_3|s_{T-1}, D, p_{l_{T-1}} = 0) = 0$$

$$\text{Prob}(O_4|s_{T-1}, D, p_{l_{T-1}} = 0) = \begin{cases} F(s_{T-1}) & \text{if } s_{T-1} \in [0, \overline{s}_{B,T}] \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Prob}(O_9|s_{T-1}, D, p_{l_{T-1}} = 0) = \begin{cases} 1 - F(\underline{s}_{N,T}) & \text{if } s_{T-1} \in [0, \underline{s}_{N,T}] \\ 1 - F(s_{T-1}) & \text{if } s_{T-1} \in (\underline{s}_{N,T}, 1] \end{cases}$$

$$\text{Prob}(O_{10}|s_{T-1}, D, p_{l_{T-1}} = 0) = \begin{cases} F(\underline{s}_{N,T}) - F(\overline{s}_{B,T}) & \text{if } s_{T-1} \in [0, \overline{s}_{B,T}] \\ F(\underline{s}_{N,T}) - F(s_{T-1}) & \text{if } s_{T-1} \in (\overline{s}_{B,T}, \underline{s}_{N,T}] \\ 0 & \text{if } s_{T-1} \in (\underline{s}_{N,T}, 1] \end{cases}$$

$$\text{Prob}(O_{11}|s_{T-1}, D, p_{l_{T-1}} = 0) = 0$$

$$\text{Prob}(O_{12}|s_{T-1}, D, p_{l_{T-1}} = 0) = \begin{cases} F(\overline{s}_{B,T}) - F(s_{T-1}) & \text{if } s_{T-1} \in [0, \overline{s}_{B,T}] \\ 0 & \text{if } s_{T-1} \in (\overline{s}_{B,T}, 1] \end{cases}$$

The corresponding probabilities for equations (D.15)-(D.17) follow directly. This completes the description of the process for characterizing the subgame perfect Nash equilibria of the finite horizon leaderboard-feedback innovation contest.

D.2 Incorporating Behavioral Characteristics

We obtain predictions for risk aversion, loss aversion, and the sunk cost fallacy using the following procedure:

- First, for a maximum score in the leader-board feedback treatment and an individual score in the private feedback treatment, we calculate the expected utility from drawing or not drawing in the last period. At this stage, we incorporate the relevant behavioral characteristic (risk aversion, loss aversion, sunk cost fallacy) into that calculation and repeat this process for various scores in each treatment.

- We then calculate the expected utility, and the optimal decisions, in the penultimate period for the same scores. We calculate the expected utility of drawing and not drawing in the penultimate period through backward induction as we have solved for the last period.

- We continue this process using backward induction. Once we have solved for the optimal decisions for each score and period, we use simulations to make contest predictions.

We use the following specifications:

- **Risk aversion** is modeled using CRRA utility, that is, $u(x) = \frac{x^{1-r}}{1-r}$.

- **Loss aversion** is modeled as an individual being reference dependent around losses. Let TC be the total cost an agent has spent in the contest and E be the agent's endowment. When an individual loses the contest, her utility is given by $E - \lambda * TC$, where $\lambda > 1$. Note that an individual can never lose money when she wins the prize in our experiment. When an individual wins the contest, her utility is given by $E + V - TC$, where V is the prize value.

- The **sunk cost fallacy** is modeled as an individual having a preference for drawing when she has accumulated sunk costs in the contest. An individual's expected utility in the last period from drawing is given by $E - TC + \alpha * TC + p(V) * V$, where $\alpha > 0$ and $p(V)$ is the probability that she wins the contest.

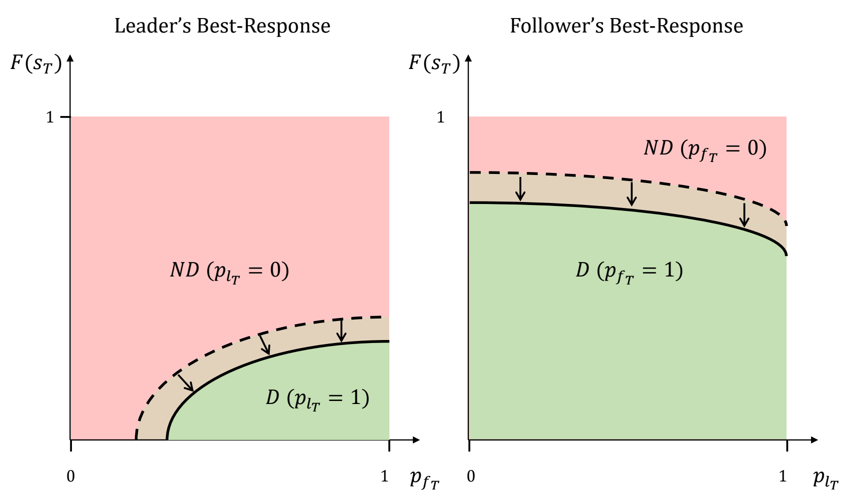


Figure D.1.: The effect of risk aversion on period T local best responses for Leaderboard Feedback.

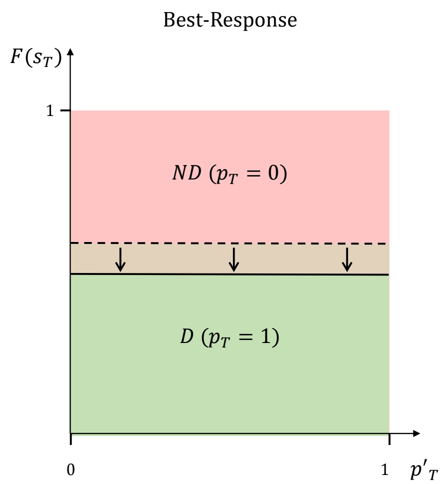


Figure D.2.: The effect of risk aversion on period T local best responses for Private Feedback.

D.3 Experimental Instructions

D.3.1 Introduction

Welcome and thank you for participating! Today's experiment will last about 60 minutes. Everyone will earn at least $5. If you follow the instructions carefully, you might earn even more money. This money will be paid at the end of the experiment in private and in cash.

It is important that during the experiment you remain silent. If you have a question or need assistance of any kind, please raise your hand, but do not speak - and an experiment administrator will come to you, and you may then whisper your question. In addition, please turn off your cell phones and put them away during the experiment. Anybody that violates these rules will be asked to leave.

In this experiment you will face 27 tasks in which you will take the role of an entrepreneur. Prior to each task, you will be provided with the information regarding the task. At the end of the experiment, two of the tasks will be chosen randomly to determine your actual money earnings. Thus, your decisions in one task will not affect your earnings in any other task. In addition, at the end of the 27 tasks, you will be asked to fill out several questionnaires.

Next, you will be provided detailed information pertaining to Task #1-8 of the experiment. Before starting with the actual tasks, you will face one practice task. Your compensation for the experiment will not depend on the practice task

D.3.2 Tasks #1–8: Description

In Tasks #1–8 of the experiment, you will be given an endowment of $10 and choose whether to develop up to 10 technologies at a cost of $1 per technology. The quality of each technology is uncertain and will be determined randomly using the probability distribution to the right. However, only the best technology can be brought to the market and yield revenue.

The decisions whether to develop a technology will be made sequentially. In particular, you will first decide whether to develop technology #1. If you decide to do so, you will incur a cost of \$1 and observe the quality of technology #1. Next, you will decide whether to develop technology #2. If so, you will incur a cost of \$1. And so on. Each new technology will be obtained using an independent draw from the distribution to the right. That is, quality of technology #2 does not depend on technology #1, quality of technology #3 does not depend on technology #2, etc. At each decision, you will be provided with the summary information in the graphical and text forms.

For example, suppose you have developed 4 technologies. Each of them will be marked on the graph with a line. At the time of each decision, you will be provided with the probability that a new technology will be better (or worse) than the best known technology. For example, suppose you are deciding whether to develop technology #5, then the probability that technology #5 will be better than the best known technology is shaded in green, and is equal to 36%. The probability that technology #5 will be worse than the best known technology is shaded in red, and is equal to 64%.

For each task, you will be randomly matched with another participant in this room. Each of you will simultaneously and independently decide whether to develop up to 10 technologies (one technology at a time). At the time of each decision you will not know the technology that has the best quality among all of the technologies developed so far (either by you or by the participant that you are matched with). After all of the decisions have been made, the best technology developed in during the task (either by you or by the participant that you are matched with) will be revealed. The best technology will be adopted by the market and yield \$10 revenue.

At this time you can get some experience of drawing from the distribution. You can click 'Draw' to draw a random number from the distribution. You can also click 'Reset' to clear all the draws. Reminder, each draw is independent from all other draws. Note, that although the diagram shows domain to be [0,50], the domain is unbounded and there is a small chance (less than a quarter of one percent) that a draw from the distribution will exceed 50. When you are done drawing random numbers from the distribution, please click 'Continue to Practice Task'.

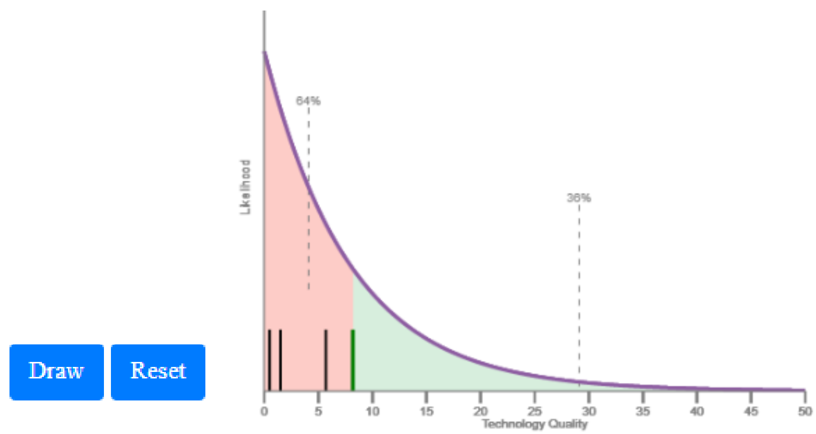


Figure D.3.: Screenshots of the distribution presented in the instructions.

D.3.3 Tasks #1–8: Practice Task

For each of the Tasks #1-8, you will be randomly matched with another participant in this room. That is, there will be new random rematching at the beginning of each task, but the matching will stay fixed within a task. Each participant will be given an endowment of $10 and able to develop up to 10 technologies at the cost of $1 per technology. At the time of each decision you **will not** know the technology that has the best quality among all of the technologies developed so far (either by you or by the participant that you are matched with). Note that only the best technology among the two of you can be brought to the market and yield revenue. **The best technology will generate a revenue of $10.**

For the practice task, you will make a sequence of decisions in this setting, however, unlike the actual tasks, for the practice tasks you will be matched with a computer that chooses randomly.

Market Summary:

- Number of entrepreneurs: **Two**
- Best market technology: **Unknown**
- Cost per technology: **$1**
- Your endowment: **$10**

You will make a sequence of 10 decisions. Each decision is a choice between two options:

- Option A: develop another technology at a cost of $1
- Option B: do NOT develop another technology

The summary of the most current information is presented below.

Decision Summary:

- Decision number: **6**
- Technologies developed by you: **9.690**, **7.572**, **4.079**
- Incurred cost: **3 x $1=$3**
- Best market technology: **Unknown**
- Probability that technology #**6** will be better than **9.690** is **30%**
- Probability that technology #**6** will be better than **9.690** is **70%**

Please make your decisions:

Option A: Develop Technology #6 for $1. | Option B: Do NOT Develop Technology #6.

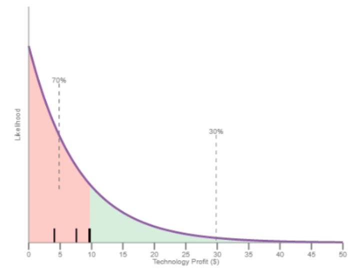


Figure D.4.: Screenshots of the practice task.

D.4 Additional Tables and Figures

Please make a choice for **each of the 20 decisions** in this task.

Decision	Option A	Option B	Your Choice
#1	**$10** with 50% chance; **$0** with 50% chance	$0.5	A
#2	**$10** with 50% chance; **$0** with 50% chance	$1	A
#3	**$10** with 50% chance; **$0** with 50% chance	$1.5	A
#4	**$10** with 50% chance; **$0** with 50% chance	$2	A
#5	**$10** with 50% chance; **$0** with 50% chance	$2.5	A
#6	**$10** with 50% chance; **$0** with 50% chance	$3	A
#7	**$10** with 50% chance; **$0** with 50% chance	$3.5	A
#8	**$10** with 50% chance; **$0** with 50% chance	$4	A
#9	**$10** with 50% chance; **$0** with 50% chance	$4.5	A
#10	$10 with 50% chance; $0 with 50% chance	**$5**	B
#11	$10 with 50% chance; $0 with 50% chance	**$5.5**	B
#12	$10 with 50% chance; $0 with 50% chance	**$6**	B
#13	$10 with 50% chance; $0 with 50% chance	**$6.5**	B
#14	$10 with 50% chance; $0 with 50% chance	**$7**	B
#15	$10 with 50% chance; $0 with 50% chance	**$7.5**	B
#16	$10 with 50% chance; $0 with 50% chance	**$8**	B
#17	$10 with 50% chance; $0 with 50% chance	**$8.5**	B
#18	$10 with 50% chance; $0 with 50% chance	**$9**	B
#19	$10 with 50% chance; $0 with 50% chance	**$9.5**	B
#20	$10 with 50% chance; $0 with 50% chance	**$10**	B

Submit Decisions

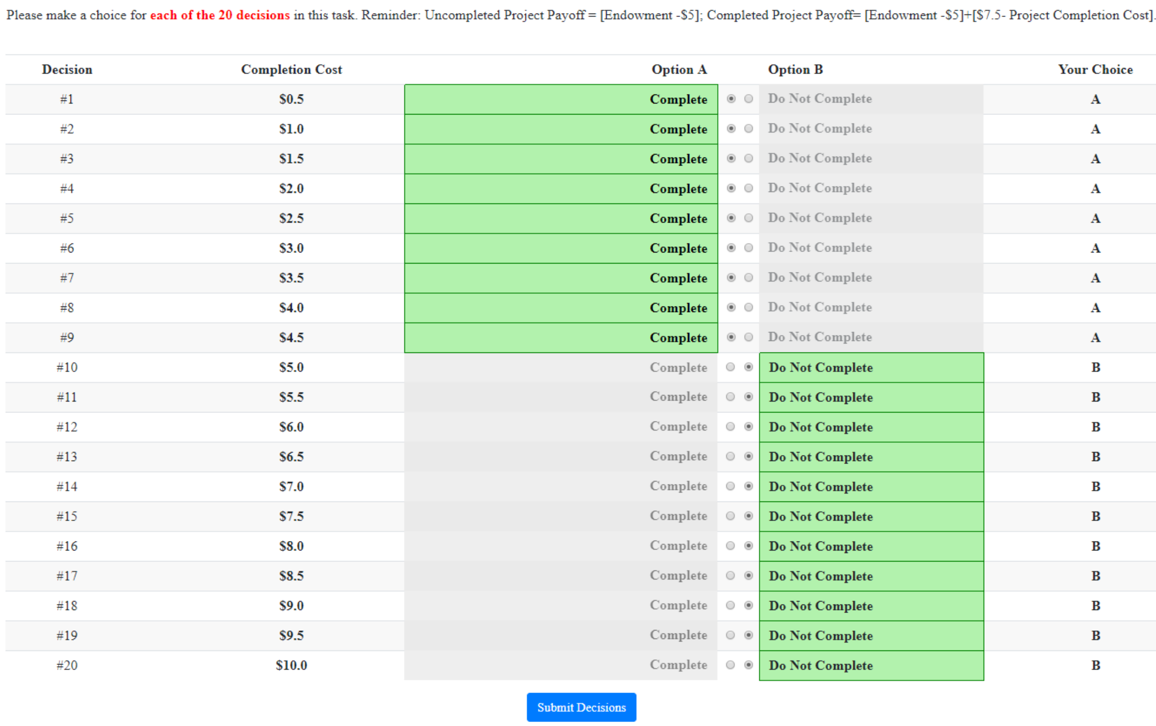
Figure D.5.: Screenshots of the risk aversion elicitation task.

Please make a choice for **each of the 20 decisions** in this task.

Decision	Option A	Option B	Your Choice
#1	-$0.5 with 50% chance; **$5.00** with 50% chance	$0.00	**A**
#2	-$1 with 50% chance; **$5.00** with 50% chance	$0.00	**A**
#3	-$1.5 with 50% chance; **$5.00** with 50% chance	$0.00	**A**
#4	-$2 with 50% chance; **$5.00** with 50% chance	$0.00	**A**
#5	-$2.5 with 50% chance; **$5.00** with 50% chance	$0.00	**A**
#6	-$3 with 50% chance; **$5.00** with 50% chance	$0.00	**A**
#7	-$3.5 with 50% chance; **$5.00** with 50% chance	$0.00	**A**
#8	-$4 with 50% chance; **$5.00** with 50% chance	$0.00	**A**
#9	-$4.5 with 50% chance; **$5.00** with 50% chance	$0.00	**A**
#10	-$5 with 50% chance; $5.00 with 50% chance	**$0.00**	**B**
#11	-$5.5 with 50% chance; $5.00 with 50% chance	**$0.00**	**B**
#12	-$6 with 50% chance; $5.00 with 50% chance	**$0.00**	**B**
#13	-$6.5 with 50% chance; $5.00 with 50% chance	**$0.00**	**B**
#14	-$7 with 50% chance; $5.00 with 50% chance	**$0.00**	**B**
#15	-$7.5 with 50% chance; $5.00 with 50% chance	**$0.00**	**B**
#16	-$8 with 50% chance; $5.00 with 50% chance	**$0.00**	**B**
#17	-$8.5 with 50% chance; $5.00 with 50% chance	**$0.00**	**B**
#18	-$9 with 50% chance; $5.00 with 50% chance	**$0.00**	**B**
#19	-$9.5 with 50% chance; $5.00 with 50% chance	**$0.00**	**B**
#20	-$10 with 50% chance; $5.00 with 50% chance	**$0.00**	**B**

Submit Decisions

Figure D.6.: Screenshots of the loss aversion elicitation task.



Please make a choice for **each of the 20 decisions** in this task. Reminder: Uncompleted Project Payoff = [Endowment -$5]; Completed Project Payoff= [Endowment -$5]+[$7.5- Project Completion Cost].

Decision	Completion Cost	Option A	Option B	Your Choice
#1	$0.5	Complete	Do Not Complete	A
#2	$1.0	Complete	Do Not Complete	A
#3	$1.5	Complete	Do Not Complete	A
#4	$2.0	Complete	Do Not Complete	A
#5	$2.5	Complete	Do Not Complete	A
#6	$3.0	Complete	Do Not Complete	A
#7	$3.5	Complete	Do Not Complete	A
#8	$4.0	Complete	Do Not Complete	A
#9	$4.5	Complete	Do Not Complete	A
#10	$5.0	Complete	Do Not Complete	B
#11	$5.5	Complete	Do Not Complete	B
#12	$6.0	Complete	Do Not Complete	B
#13	$6.5	Complete	Do Not Complete	B
#14	$7.0	Complete	Do Not Complete	B
#15	$7.5	Complete	Do Not Complete	B
#16	$8.0	Complete	Do Not Complete	B
#17	$8.5	Complete	Do Not Complete	B
#18	$9.0	Complete	Do Not Complete	B
#19	$9.5	Complete	Do Not Complete	B
#20	$10.0	Complete	Do Not Complete	B

Submit Decisions

Figure D.7.: Screenshots of the sunk cost fallacy elicitation task.

In Task # 20, you will be the sole entrepreneur. You are able to develop up to 10 technologies at the cost of $1 per technology.

Market Summary:

- Number of entrepreneurs: **One**
- Existing market technology: **Known (Shown in red)**
- Cost per technology: **$1**
- Your endowment: **$10**

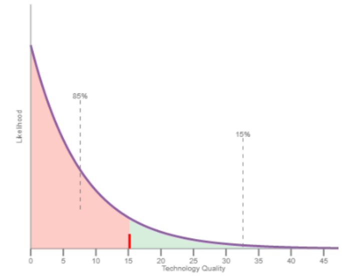
You will make up to 10 decisions. Each decision is a choice between two options:

- Option A: develop another technology at a cost of $1
- Option B: do NOT develop another technology

The summary of the probability that the new technology will be better (or worse) than the existing technology is presented below. The graphical summary is presented to the right.

Decision Summary:

- Decision number: **1**
- Technologies developed by you: **None**
- Incurred cost: **0 x $1= $0**
- Existing market technology: **15.177**
- Probability that technology **#1** will be better than **15.177** is **15%**
- Probability that technology **#1** will be worse than **15.177** is **85%**



Please make your decisions for task # 20.

Option A: Develop Technology #1 for $1. | Option B: Do NOT Develop Technology #1.

Figure D.8.: Screenshots of the individual search task.

Table D.1.: Displays the results of regressions. The regression pools the data from the individual search tasks, the private-feedback treatment, and the leaderboard-feedback treatment.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. Var.:	Pooled		Private			Leaderboard	
Draw Decision		All	Leader	Follower	All	Leader	Follower
L-Board	-0.70***	—	—	—	—	—	—
	(0.20)	—	—	—	—	—	—
Priv. x Score	-0.17***	-0.21***	-0.25***	-0.18***	—	—	—
	(0.01)	(0.02)	(0.04)	(0.02)	—	—	—
L-Board x MaxScore	-0.11***	—	—	—	-0.11***	-0.23***	-0.11***
	(0.01)	—	—	—	(0.01)	(0.02)	(0.01)
Period	-0.12***	-0.13***	-0.19***	-0.11***	-0.10***	-0.24***	-0.03
	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.05)
Risk Aversion	-1.10**	-1.61**	-1.71	-1.36*	-0.79*	-0.64	-0.20
	(0.50)	(0.76)	(1.26)	(0.70)	(0.45)	(0.93)	(0.91)
Loss Aversion	-0.13	0.02	1.26	-0.74	-0.22	-1.33	0.07
	(0.63)	(0.78)	(1.11)	(0.65)	(0.66)	(1.11)	(0.96)
Sunk Cost Fallacy	0.11	0.29	-0.93	0.41	-0.15	-0.62	0.01
	(0.63)	(0.97)	(0.85)	(1.00)	(0.51)	(0.87)	(0.96)
Factor 1	0.05	0.17	0.02	0.12	-0.04	-0.05	-0.31***
	(0.09)	(0.13)	(0.14)	(0.13)	(0.09)	(0.18)	(0.12)
Factor 2	—	—	—	—	—	—	—
	—	—	—	—	—	—	—
Factor 3	0.02	-0.03	0.12	-0.07	0.05	-0.23	0.12
	(0.11)	(0.14)	(0.21)	(0.13)	(0.11)	(0.19)	(0.14)
Factor 4	0.06	0.07	-0.01	0.09	0.06	0.08	0.15
	(0.06)	(0.09)	(0.09)	(0.08)	(0.07)	(0.15)	(0.18)
Factor 5	0.16	0.22	0.11	0.15	0.10	0.05	0.20**
	(0.10)	(0.16)	(0.19)	(0.14)	(0.08)	(0.18)	(0.08)
Factor 6	0.18**	0.17	0.09	0.17	0.20**	0.16	0.08
	(0.08)	(0.11)	(0.13)	(0.12)	(0.09)	(0.17)	(0.14)
Factor 7	-0.09	-0.17	-0.22	-0.17	-0.04	-0.06	-0.22
	(0.11)	(0.18)	(0.23)	(0.17)	(0.08)	(0.18)	(0.18)
Factor 8	-0.09	-0.07	0.00	-0.10	-0.12	-0.03	-0.14
	(0.08)	(0.12)	(0.16)	(0.13)	(0.09)	(0.16)	(0.15)
Factor 9	-0.08	-0.08	-0.19	0.02	-0.11	0.04	-0.44**
	(0.12)	(0.16)	(0.17)	(0.16)	(0.10)	(0.22)	(0.22)
Constant	1.69***	1.83**	4.26***	1.17	1.17**	2.17***	1.99***
	(0.57)	(0.81)	(0.77)	(0.86)	(0.51)	(0.74)	(0.76)
Observations	15,360	7,680	3,451	3,451	7,680	3,411	3,411

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

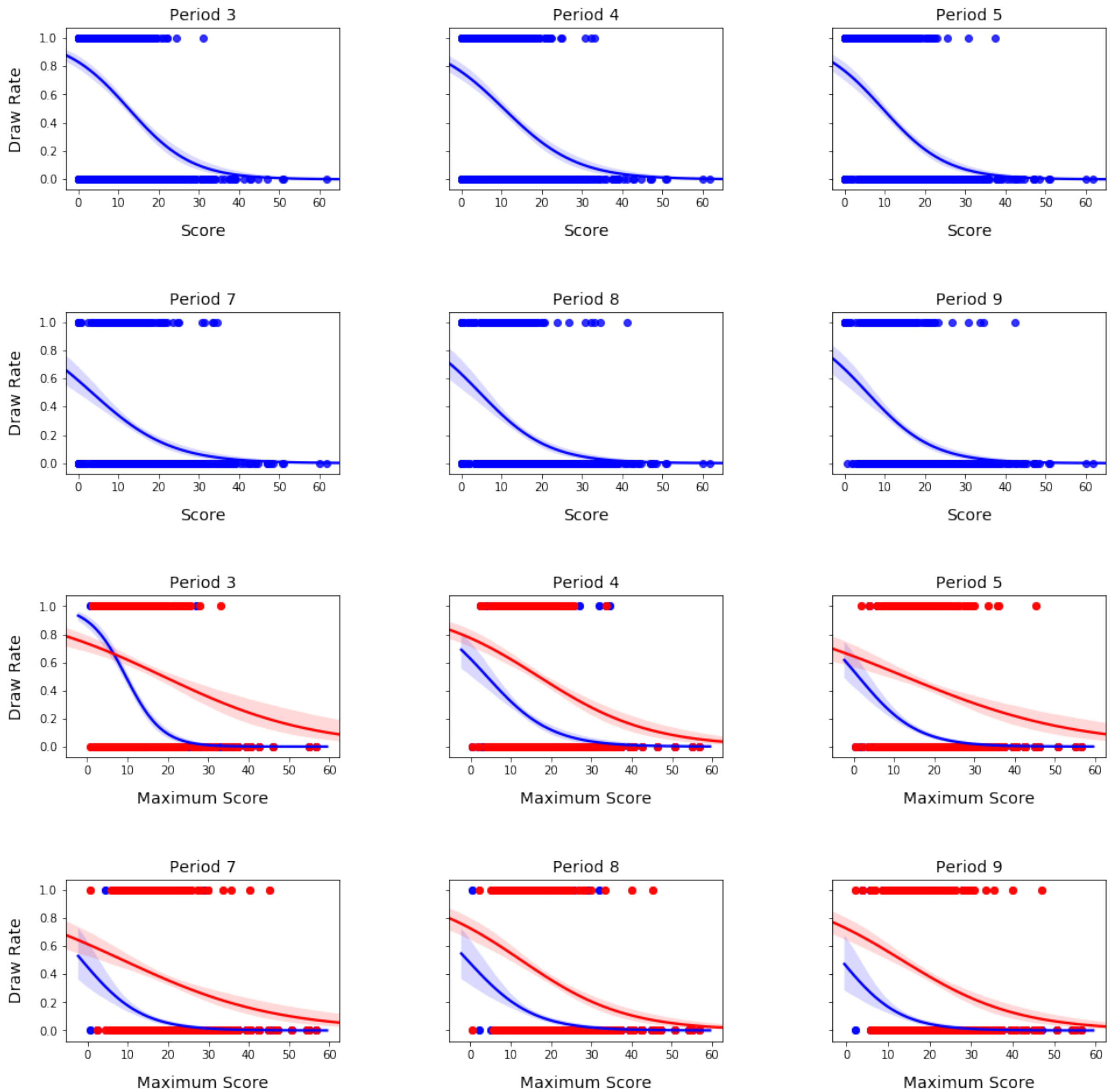


Figure D.9.: Displays the decision to draw in the Leaderboard-Feedback treatment. This figure displays two sets of graphs. The first set of graphs display logistic regressions of the decision to draw in the private-feedback treatment for periods 3, 4, 5, 7, 8, and 9. The second set of graphs display logistic regressions of the leader's decision (blue) to draw and the follower's decision (red) to draw in the leaderboard-feedback treatment for periods 3, 4, 5, 7, 8, and 9.

Table D.2.: Displays the results of the contests. Priv. Draws refers to the mean number of draws in a contest in a session in the pravate-feedback treatment. LB Draws refers to the mean number of draws in a contest in a session in the leaderboard-feedback treatment. Priv. Innovation refers to the mean value of the winning innovation in a session in the pravate-feedback treatment. LB Innovation refers to the mean value of the winning innovation in a session in the leaderboard-feedback treatment.

	Priv. Draws	LB Draws	Priv. Innovation	LB Innovation
Session 1	6.53	7.16	24.02	19.20
Session 2	7.78	8.00	21.81	23.46
Session 3	8.97	7.89	19.34	23.16
Session 4	7.28	6.22	22.44	19.24
Session 5	7.41	7.25	21.06	20.84
Session 6	7.22	7.50	20.40	21.48
Session 7	9.19	7.09	26.54	19.82
Session 8	8.59	6.69	24.82	21.21
Session 9	9.28	7.72	21.62	23.91
Session 10	10.16	9.75	22.92	20.54
Session 11	9.03	7.00	24.18	21.58
Session 12	10.59	8.28	25.33	23.18

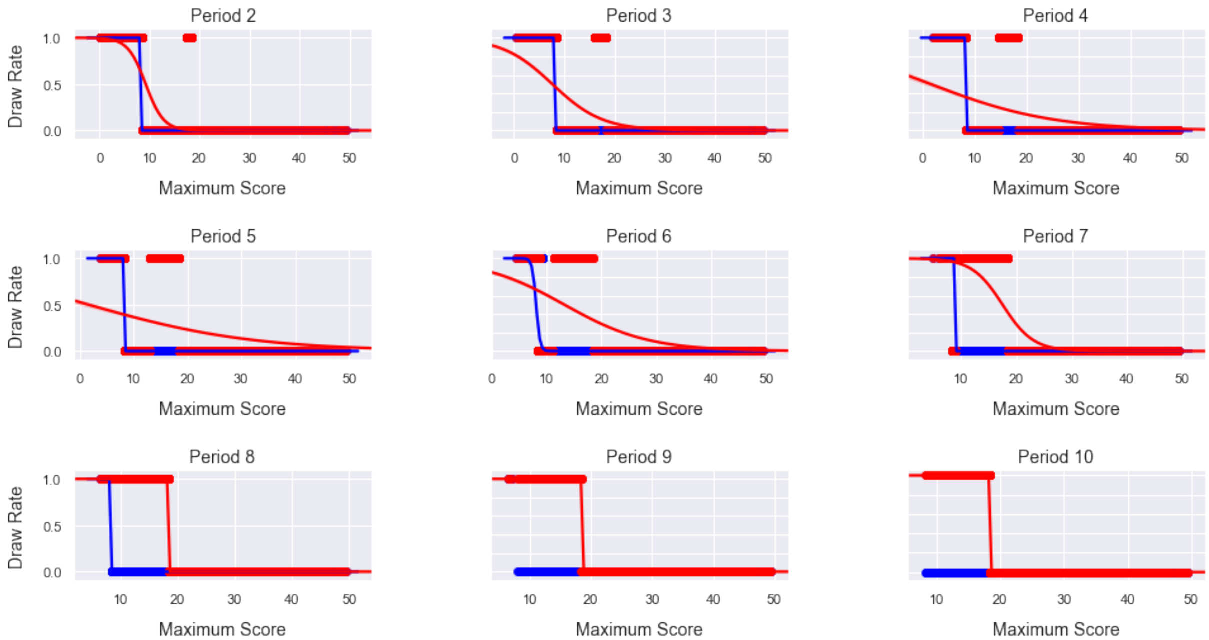


Figure D.10.: Displays the decision to draw in the simulated Leaderboard-Feedback contests. These graphs display logistic regressions of the leader's decision (blue) to draw and the follower's decision (red) to draw in the simulated leaderboard-feedback treatment contests for periods 2, 3, 4, 5, 6, 7, 8, 9, 10.

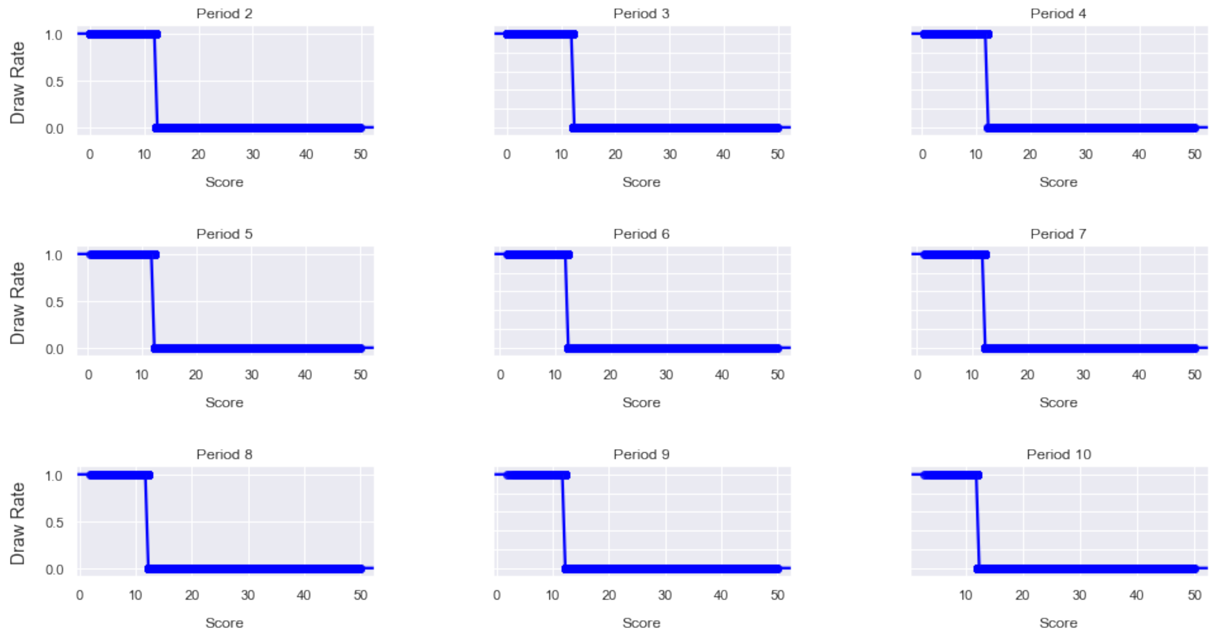


Figure D.11.: Displays the decision to draw in the simulated Private-Feedback contests. The first set of graphs display logistic regressions of the decision to draw in the simulated pravate-feedback treatment contests for periods 2, 3, 4, 5, 6, 7, 8, 9, 10.

Table D.3.: Displays regression results for the individual search task. The regression pools the data from the individual search tasks, the pravate-feedback treatment, and the leaderboard-feedback treatment.

	(1)	(2)
Dep. Var.: Draw Decision	Individual	Individual
Individual Score	-0.04***	-0.04***
	(0.01)	(0.01)
Period	-0.15***	-0.15***
	(0.02)	(0.02)
Risk Aversion	-2.70**	-2.72***
	(1.17)	(1.00)
Loss Aversion	-0.80	-0.84
	(1.00)	(1.02)
Sunk Cost Fallacy	-0.13	-0.10
	(0.59)	(0.72)
Grit/Factor 1	-0.57**	-0.08
	(0.23)	(0.14)
Competitiveness/Factor 2	0.05	—
	(0.43)	—
Achievement Striving/Factor 3	0.02	0.27**
	(0.66)	(0.11)
Extraversion/Factor 4	0.17	0.06
	(0.11)	(0.12)
Agreeableness/Factor 5	0.27	0.19**
	(0.24)	(0.08)
Neuroticism/Factor 6	-0.18	-0.10
	(0.23)	(0.12)
Openness/Factor 7	0.05	-0.06
	(0.18)	(0.19)
Conscientiousness/Factor 8	0.09	-0.19**
	(0.24)	(0.08)
Factor 9	—	0.07
	—	(0.22)
Constant	-1.35	-1.48**
	(0.99)	(0.67)
Observations	7,680	7,680

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

Table D.4.: Displays the results of the demographics regressions. The regressions analyze how demographics influence the decision to draw. Gender is a dummy variable for male. There are multiple race dummy variables, major dummy variables, school year dummy variables, and high school location dummy variables that are in these regressions, but not included in the tables.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. Var.:	Pooled		Private			Leaderoard	
Draw Decision		All	Leader	Follower	All	Leader	Follower
L-Board	-0.70***	—	—	—	—	—	—
	(0.20)	—	—	—	—	—	—
Priv. x Score	-0.17***	-0.21***	-0.25***	-0.18***	—	—	—
	(0.01)	(0.02)	(0.04)	(0.02)	—	—	—
L-Board x MaxScore	-0.11***	—	—	—	-0.11***	-0.23***	-0.11***
	(0.01)	—	—	—	(0.01)	(0.02)	(0.01)
Period	-0.12***	-0.13***	-0.19***	-0.11***	-0.10***	-0.24***	-0.03
	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.04)	(0.05)
Risk Aversion	-0.74**	-1.15**	-1.64**	-0.95	-0.60	-0.22	-0.58
	(0.35)	(0.47)	(0.80)	(0.61)	(0.38)	(1.05)	(0.65)
Loss Aversion	-0.72	-0.67	0.52	-1.01*	-0.79	-1.32	-1.00
	(0.50)	(0.59)	(0.72)	(0.56)	(0.62)	(1.06)	(0.91)
Sunk Cost Fallacy	0.40	0.66	-0.47	0.77	0.07	-0.29	0.51
	(0.49)	(0.75)	(0.67)	(0.81)	(0.32)	(0.80)	(0.62)
Gender	-0.15	-0.17	-0.11	-0.03	-0.14	-0.36	0.33
	(0.14)	(0.20)	(0.25)	(0.22)	(0.17)	(0.30)	(0.28)
Age	-0.11***	-0.18**	-0.23***	-0.16**	-0.07	-0.04	-0.22
	(0.04)	(0.07)	(0.08)	(0.07)	(0.05)	(0.13)	(0.13)
Constant	3.09***	3.93**	7.35***	2.78	2.05**	3.56	4.13*
	(0.96)	(1.75)	(1.48)	(1.83)	(0.85)	(2.25)	(2.22)
Observations	15,360	7,680	3,451	3,451	7,680	3,411	3,411

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

VITA

EDUCATION

Ph.D. Economics, Department of Economics, Purdue University	May 2020
M.S. Economics, Johns Hopkins University	May 2014
B.A. Economics, B.S. Mathematics, University of Maryland	May 2013

FIELDS OF RESEARCH

Experimental Economics, Industrial Organization, Game Theory, Behavioral Economics

PUBLICATIONS

"Dispositional Negativity in the Wild: Social Environment Governs Momentary Emotional Experience" (with Alexander Shackman, Jennifer Weinstein, Conor Bloomer, Matt Barstead, Andrew Fox, and Edward Lemay Jr.), *Emotion* (2018), 18, 707-724.

"Hypothetical Purchase Task Questionnaires for Behavioral Economic Assessments of Value and Motivation" (with Peter Roma and Steven Hursh), *Managerial and Decision Economics* (2016), 37, 306-323.

"The Effect of Crossing the $100 Million Jackpot Threshold on Ticket Sales" *Undergraduate Economic Review* (2013), 10, Article 7.

WORKING PAPERS

"Behavioral Bandits: Analyzing the Exploration Versus Exploitation Trade-off in the Lab" (With Daniel Woods)

"Public Leaderboard Feedback in Innovation Contests: A Theoretical and Experimental Investigation" (with Brian Roberson and Yaroslav Rosokha) [**Under Review**]

"Voting for Experimentation: A Continuous Time Analysis" [**Under Review**]

"Is Experimentation Invariant to Group Size? A Laboratory Analysis of Innovation Contests" [**Under Review**]

CONFERENCES AND PRESENTATIONS

2019: Southern Economic Association Annual Meeting; Krannert Ph.D. Research Symposium; Jordan-Wabash Conference

2018: Young Economists Symposium; Southern Economic Association Annual Meeting; Jordan-Wabash Conference

2017: Southern Economic Association Annual Meeting; Jordan-Wabash Conference; Midwest Economics Association Annual Meeting; Krannert Ph.D. Research Symposium

2016: Southern Economic Association Annual Meeting; AERUS Midwest Graduate Student Summit; Krannert Ph.D. Research Symposium

SCHOOLS AND WORKSHOPS

2018: Spring School in Behavioral Economics (UC San Diego)

2016: IFREE Graduate Student Workshop in Experimental Economics (Chapman University)

2015: Workshop on Behavioral Game Theory (University of East Anglia)

HONORS AND AWARDS

Purdue Research Foundation Grant	Spring 2018, Summer 2018
Best Presentation, Krannert Ph.D. Research Symposium	November 2017
Krannert Certificate for Distinguished Teaching	Summer 2017
Honorable Mention, Krannert Ph.D. Research Symposium	November 2016
Purdue Doctoral Fellowship, Purdue University	August 2014 – July 2016
Senior Scholar Award in Economics, University of Maryland	May 2013
Department Scholar of Economics, University of Maryland	May 2011

TEACHING EXPERIENCE

Intermediate Microeconomics	Spring 2020
Introduction to Microeconomics	Summer 2017
Introduction to Microeconomics (Online)	Summer 2016

REFEREEING EXPERIENCE

International Tax and Public Finance, Defence and Peace Economics, European Journal of Operational Research, Journal of Economic Behavior and Organization, Southern Economic Journal; American Economic Journal: Microeconomics; Social Choice and Welfare

SKILLS

Python, Stata, R, Matlab, Redwood, z-Tree, oTree