University of California
Santa Barbara

Three Essays on Experimental Economics and Applied Microeconomics

A dissertation submitted in partial satisfaction
of the requirements for the degree

Doctor of Philosophy
in
Econonomics

by

James Michael Banovetz, III

Committee in charge:

Professor Ryan Oprea , Chair
Professor Ignacio Esponda
Professor Emanuel Vespa

June 2020

The dissertation of James Michael Banovetz, III is approved.

Ignacio Esponda

Emanuel Vespa

Ryan Oprea , Committee Chair

June 2020

Three Essays on Experimental Economics and Applied Microeconomics

by

James Michael Banovetz, III

To my family.

Acknowledgments

I would first like to thank my dissertation committee–Ryan Oprea, Ignacio Esponda and Emanuel Vespa–for their invaluable guidance and encouragement, as well as Randal Rucker for his continued support. I would also like to thank Sevgi Yuksel, Erik Eyster, Cheng-Zhong Qin, and the rest of the experimental and theory group at UCSB, seminar participants at UC Santa Barbara, Montana State University, the 2019 Economic Science Association meeting, and the Brattle Group. I thank Timothy Fitzgerald, Mark Anderson, and Charles Steele, and all of my teachers for helping me advance my education. Finally, I would like to thank my family, friends, and fellow graduate students at UCSB, who provided me with endless love, patience, and support.

Curriculum Vitae
James Michael Banovetz, III

Education

Doctor of Philosophy in Economics, University of California, Santa Barbara	June 2020
Master of Arts in Economics, University of California, Santa Barbara	June 2015
Master of Science in Applied Economics, Montana State University	May 2014
Bachelor of Science in Economics, Hillsdale College	May 2012

Fields of Study

Primary: Experimental Economics, Behavioral Theory
Secondary: Applied Microeconomics, Econometrics and Statistics

Working Papers

"Preferences of Exploration: Simple Bandits in the Lab"
"Complexity and Procedural Choice: Evidence from Experimental Bandits," with Ryan Oprea
"The Economics of the Montana Liquor License System," with Randal Rucker

Awards

Department of Economics Research Quarter Fellowship	Spring 2019
Keith Griffin Memorial Fellowship	2017-2018
Outstanding Graduate Course Teaching Assistant of the Year	2017-2018
Thormahlen Family Fellowship in Economics	2016-2017
Outstanding Graduate Course Teaching Assistant of the Year	2016-2017
Humane Studies Fellowship	2012-2013
Meritorious Graduate Scholarship	2012-2013
Adam Smith Prize for Excellence in Economics	2011-2012

Presentations

North American Meetings of the Economic Science Association	2019
Hillsdale College, "What is Experimental Economics?"	2017
Montana State University, "The Economics of the Montana Liquor License System"	2017

Abstract

Three Essays on Experimental Economics and Applied Microeconomics

by

James Michael Banovetz, III

This dissertation consists of two essays in experimental economics and one in applied microeconomics. The first chapter, *Preferences over Exploration: Simple Bandits in the Lab*, presents an experimental study about how people view the trade-off between exploration (trying a new option) and exploitation (taking a familiar option). This study documents a large and persistent tendency of people to under-explore, even in the simplest possible bandit setting. By simplifying the environment, the study can directly control for risk preferences, abstract from other issues such as ambiguity, Bayesian updating, or failures to reduce compound lotteries. The second chapter, *Complexity and Procedural Choice: Evidence from Experimental Bandits*, presents an experiment about how subjects select procedural rules to make decisions in a simple bandit task. The optimal strategy involves a relatively sophisticated 4-state decision rule. When complexity costs are low, subjects conform to the optimal behavior; when complexity costs are higher, holding other aspects of the problem constant, subjects systematically employ lower-complexity rules. This suggests that aversion to complexity causes subjects to play simpler strategies and earn lower payoffs. The final chapter, *The Economics of the Montana Liquor License System*, is an applied microeconomics paper about the the liquor license system in the state of Montana. The state has a complicated system of quotas for liquor licenses, involving multiple license types, which are tradable within quota area (and occasionally between quota areas). The study offers a simple model of the system, providing testable hypotheses. Empirical results support these hypotheses, indicating that gambling revenue, non-permanent population, and income all positively affect prices, while the emergence of small-scale producers (e.g., craft breweries) negatively affect license prices.

Contents

Chapter 1

Preferences Over Exploration: Simple Bandits in the Lab

James Banovetz

1.1 Introduction

One of the central trade-offs in human decision making is that of exploration versus exploitation. Exploration entails trying something new that offers unknown payoffs–from going to a new restaurant to finding alternate routes to the New World. Exploitation, on the other hand, is simply the act of choosing an option with known payoffs–ordering the usual. How much time and energy should we devote to trying new things, finding new places, or developing new ideas? Picking a novel option offers potentially generous rewards, but comes at the cost of forgoing the benefits of the established option. For example, hunter-gatherers may spend time searching new areas for food and water, but that is time away from extracting resources in familiar territory. Modern society features this trade-off as well: should a worker join a large firm with stable hours and pay, or test a new market as a self-employed entrepreneur?

The canonical model of the exploration/exploitation trade-off is the bandit,[1] a statistical framework that features a profit-maximizing agent making dynamic choices over multiple choices, known as "arms". Each arm can be "pulled," offering a payout drawn from a stationary and independent distribution, the properties of which the agent may or may not know *ex ante*. The agent can "exploit" the environment by pulling an arm which she knows (or believes) to pay well; for example, an arm with which she has enough experience to be confident in its payoffs. Alternatively, she may "explore" by choosing an arm with which she has less experience–it may be riskier, but each pull offers additional information, which she can use to update her beliefs about the profitability of the arm. In the simplest case, the choice is between two arms: a "safe" arm that pays a known constant amount, and a "risky" arm that pays randomly according to a distribution with an unknown mean. By playing the risky arm, the agent would forgo the known payoff, but would learn potentially valuable information (i.e., whether the mean of the risky arm is higher than the safe payoff).

The degree to which people view exploration value in strictly expected-payoff terms, however, is an open question. As typically modeled, the exploration/exploitation trade off is nothing more than the more familiar tension between risk and return. In a literal interpretation of such a model, risk preferences should fully account for people's behavior in bandits. A risk averse person would tend to take the safe arm, while a risk seeker would do the opposite. Other behaviors, however, are plausible. For example, myopia would drive people to take the option with the highest current-period expected value, disregarding the value of future earnings and thus the value of exploration. Something like curiosity, on the other hand, may lead people to explore simply for the joy of edification. This raises several key questions. Do people recognize the value of exploration? Do they

[1] So named for the colloquial term for a slot-machine. Slot machines are sometimes referred to as "one-armed bandits," as they traditionally operate via a single mechanical lever (which looks like an arm) and they take players' money.

systematically respond to changes in incentive to explore? In light of the standard bandit model, do people over- or under-explore relative to their risk preferences?

This paper examines how people value exploration, employing independent lottery tasks as well as simple two-period bandits in controlled laboratory settings. In the experiments, subjects begin by making decisions over single-period lotteries featuring two arms: a "risky" arm that randomly pays either zero or a large amount depending on its state; and a "safe" arm that offers a guaranteed payment. This task serves as a risk-elicitation mechanism, where the values stochastically change after every choice. In the bandit, the game adds a second period, where the state of each arm remains constant. This implies that after a pull of the risky arm, a subject can fully determine the arm's payoff in the next period. In both the lottery and the bandit task, the first-period payoffs are identical, but the bandit also includes the opportunity to explore. Furthermore, all gambles are either fifty-fifty propositions or sure-things, drastically simplifying the computational burden of the problem.

The design employs a Becker-Degroot-Marschak (BDM) mechanism to elicit each subject's willingness-to-pay. Subjects report their valuations of the risky arm, prior to observing the value of the safe arm. The stated valuation operates as a threshold–if the realized value of the safe arm is lower than the threshold, a subject must play the risky arm; otherwise, she must play the safe arm. In the bandit, this threshold only constrains the first-period choice; subjects may always choose freely in the second period. The BDM mechanism serves to convexify the action space, producing more precise data for each subject than a set of binary choices. Furthermore, the value each subject reports corresponds to two theoretically appealing measures: the certainty equivalent and the Gittins Index (Gittins, 1979).

The certainty equivalent is a familiar measure in standard economics problems of risk and uncertainty. Suppose an agent faces a lottery with random payoffs. The certainty equivalent is a non-random payoff such that the agent is indifferent between the fixed payoff and the lottery. Though less familiar, a closely related concept is the Gittins Index. For each risky bandit arm, there is a fixed, non-random stream of payoffs such that an agent is *ex ante* indifferent between (optimally) playing the risky arm and taking the guaranteed stream. In other words, what amount would a safe arm have to pay such that the agent is indifferent between playing the safe and risky arms? The BDM mechanism asks subjects to report their certainty equivalents (CEs) in the lottery task and their individual Subjective Gittins Indices (SGIs) in the bandit task.

We design the treatments around two sets of comparisons: a within-subject comparison between the lottery task and the bandit task, to test if subjects recognize exploration value and, if so, whether subjects over- or under-explore given their risk preferences; and a between-subject comparison across different bandit tasks. The $WEAK$ treatment has relatively weak incentives to explore; we compare this to the behavior of subjects in the $STRONG$ treatment (with stronger incentives to explore) to test if there is a response to the increased value of exploration, holding constant the current-period expected values. These comparisons yield several intriguing results. Subjects do in fact respond systematically to incentives to explore, increasing their stated valuations when exploration is objectively more valuable. They do not, however, display a willingness to explore commensurate with their measured risk aversion–virtually all subjects systematically under-value exploration in bandit tasks relative to their individually-measured risk preferences.

On average, subjects account for less than 20% of exploration value in the bandit problems. A small portion of subjects, however, persistently violate optimality in the bandit problem–they routinely take the safe option early and the risky option late, a stochastically dominated pattern. When the subjects who make multiple mistakes are dropped from the sample, however, the average subject still only accounts for roughly

30% of exploration value. These patterns persist even after 5 practice rounds and 10 incentivized repetitions of the problem with feedback about payoffs. The multiple repetitions allow subjects, having accumulated experience, to settle on a strategy. This makes it implausible that the novelty of the environment or the interface drives our results.

While bandit problems provide a potentially deep and rich environment to model the exploration/exploitation trade off, we focus only on the simplest possible case. The classic bandit problems addressed by Gittins (1979), for example, features multiple risky arms and an infinite time horizon. It is intuitive on its face, but the apparent simplicity of the classic bandit belies the extreme difficulty of finding an optimal strategy. The Gittins Index is the foundation of an elegant solution to the problem, requiring the pull of the arm with the highest index in each period. It is notoriously difficult to calculate, however, in all but the simplest problems. Similarly, the Pandora's Box problem of Weitzman (1979), framed as an optimal stopping problem, weights exploration and exploitation; it too has a difficult solution. Indeed, the Gittins Index, Pandora's Rule, and their calculation provides the foundation for a large theoretical literature on bandit problems that spans statistics, computer science, economics, and other fields.[2] These problems, however, are frequently quite difficult, making it unlikely we could reasonably expect subjects to intuit solutions in a limited time frame; hence the focus on simple bandits.

The existing experimental economics literature on bandits is quite small and does not offer much clarity on the question of exploration versus exploitation. Banks et al. (1997) employ infinite-time bandits, finding that subjects play close to the optimal strategy, but tend to over-explore in one-armed bandits. Conversely, Anderson (2012) employs a series of finite-time bandits with beta priors. The study argues that subjects appear to under-

[2]For example, there are models of bandits with infinite-arms (Banks and Sundaram, 1992) and with switching costs (Bergemann and Välimäki, 2001). There is a large literature on strategic learning in bandits: Bolton and Harris (1999), Keller et al. (2005), Rosenberg et al. (2007), Keller and Rady (2010), Klein and Rady (2011), and Heidhues et al. (2015). Other topics include applications in imitation (Schlag, 1998), job-search ((Miller, 1984; McCall, 1991)), optimal contracting ((Halac et al., 2016, 2017)), resource extraction (Marcoul and Weninger, 2008), among others.

value new information, attributing subjects' tendency to under-exploring to ambiguity aversion. Meyer and Shi (1995), employ simpler, but longer finite-time bandits and find evidence of both under- and over-exploration, depending on the circumstances (with an overall tendency towards under-exploration). The closest study to ours is Hudja and Woods (2019), which also examines people's relative willingness to explore, but does so in a bandit with infinite continuous time. They also find evidence of under-exploration, attributing the phenomenon to a combination of risk aversion, conservatism, and probability mis-weighting. Other bandit experiments focus on other questions, e.g., strategic exploration and free-riding ((Boyce et al., 2016; Hoelzemann and Klein, 2019)), wage search (Cox and Oaxaca, 2000), the testing and fitting of specific behavioral models ((Gabaix et al., 2006; Gans et al., 2007)), and non-parametric learning rules (Hu et al., 2013).

To our knowledge, this is the first study to focus explicitly on the trade-off between exploration and exploitation in the simplest setting. Our focus is not on the bandit model per se; rather, the bandit is the tool by which we can examine this fundamental tension. The data from existing experiments, generating in more complicated environments, is somewhat difficult to interpret in regards to the specific question as to whether subjects understand exploration value. Is the under- or over-exploration due to behavioral phenomena or a novel and difficult setting? By simplifying the environment, we contribute to the literature by targeting a more fundamental set of questions with easier to interpret results.

The remainder of the paper is organized as follows. Section 2 lays out the bandit environment considered in this paper. Section 3 details the design and implementation in the laboratory. Section 4 presents the results of the experiment. Section 5 provides a discussion of the results and concludes.

1.2 The Experimental Environment

To isolate subjects' behavior, we study a simple lottery and the simplest possible bandit problem. First, consider an agent who faces a single-period, binary decision between two arms, $a_1 \in \{A, B\}$. A is the "risky" arm, in the sense that it offers a random, state-dependent payoff $x_1 \in \{L, H_1\}$, where $H_1 > L$. A can be one of two states: "good" (G), which guarantees the high payoff, or "not good" (NG), which guarantees the low payoff. Nature chooses the state of A, which is equally likely to be good or not good (i.e., the prior is $\alpha = 0.5$). The reduced lottery associated with A is then:

$$\tilde{A}_1 = \big(0.5 \circ H_1,\ 0.5 \circ L\big). \tag{1}$$

In comparison, the safe arm B pays a constant amount y, which is known to the agent. She is fully informed of the payoffs and probabilities, but does not know the state of A ex ante. Thus, the relevant valuations for the agent with utility function $U(x)$ are $V_l(A) = U(\tilde{A}_1)$ and $V_l(B) = U(y)$. While the uncertainty is over the state instead of the payoffs directly, the problem is only one period, so the distinction is irrelevant theoretically. Given this setup, we can define a payoff for B, call it $\bar{y}_l$, such that the agent is indifferent between pulling arms A (the risky lottery) and B (the safe option), i.e.,

$$U(\tilde{A}_1) = U(\bar{y}_l). \tag{2}$$

Definition 1. *The value $\bar{y}_l \in [L, H_1]$, such that the agent is indifferent between $\bar{y}_l$ and the lottery $\tilde{A}_1$, is the certainty equivalent (CE).*

Next, consider an agent who faces a similar problem that lasts for two periods, $t = 1, 2$. Nature chooses the state, again according to the prior $p = 0.5$, which remains constant for both periods. The agent then chooses between pulling the safe and risky arms,

$a_t \in \{A, B\}$. A offers payoffs $x_t \in \{L, H_t\}$, where $H_1, H_2 > L$; the payoffs are guaranteed to be high if the state is good, and low otherwise. In this problem, because the state remains constant, the agent can now learn A's state by choosing it in the first period and observing x_1. This allows the agent to update beliefs and make a more informed decision for period two.

Due to the prior and all-or-nothing nature of A, Bayesian updating is trivial and the reduction of compound lotteries is drastically simplified. A pays in the first period according to the reduced lottery $\tilde{A}_1 = \big(0.5 \circ H_1,\ 0.5 \circ L\big)$, identical to the lottery in equation 1. Suppose the agent pulls arm A in period 1. Upon receiving the payoff x_1, the agent updates her prior, to $\tilde{p}_H = 1$ or $\tilde{p}_L = 0$; all uncertainty is resolved by pulling A in period 1. As a result, for high- and low-payoffs in the first period, the second period reduced "lotteries" are no longer random:

$$\tilde{A}_{2H} = \big(1 \circ H_2,\ 0 \circ L\big) \tag{3}$$

$$\tilde{A}_{2L} = \big(0 \circ H_2,\ 1 \circ L\big). \tag{4}$$

If the agent pulls arm B in the first period, then she does not receive a signal about the state of A, and the reduced lottery is still a fifty-fifty gamble:

$$\tilde{A}_{2y} = \big(0.5 \circ H_2,\ 0.5 \circ L\big). \tag{5}$$

Without the outside option B, the agent's valuation of A would simply be the sum of the expected utility of the first-period reduced lottery and the weighted average of the second-period lotteries A_{2H} and A_{2L}. With the addition of the safe arm B, however, the first period choice of A potentially becomes much more valuable. This is due to the fact that the agent can switch from A to B in period two (if y is worth more than A_{2L}). Intuitively, this makes exploration valuable. By playing A in the first period, the subject can learn about the state, informing the second-period decision. The outside option B, moreover, limits the down-side risk of exploration. In the case of a low payoff, the subject can opt to forgo playing A a second time, taking the outside option B instead.

Note that while the agent may pull arm A, then deviate to B, she has no incentive to do the reverse. This is because playing A and resolving the uncertainty in period one stochastically dominates waiting to resolve the uncertainty until period two. Intuitively, by pulling A in period one, the agent faces two possible sets of payoffs with equal probability: (1) earning H_1 then H_2, or (2) earning L then y. Conversely, by pulling arm B and following with A, the agent faces a different set of two equally likely payoffs: (1) y then H_2, or y then L. So long as $y < H_1$, pulling A first (if at all) is state-wise dominant. This allows us to form a prediction over second-period actions:

Prediction 1. *In the second period of the bandit, it is unequivocally more valuable to pull A rather than B after learning the state is G, to pull B rather than A after learning the state is NG, and given that $H_1 \geq y$, to pull A for the first time in period one rather than in period two.*

Assuming separability of preferences across periods one and two, the valuation of A in period 1 of the bandit is given by:

$$V_b(A) = U(\tilde{A}_1) + 0.5U(\tilde{A}_{2H}) + 0.5\max\{U(y), U(\tilde{A}_{2L})\}. \tag{6}$$

Furthermore, assuming that $\min\{L_1, L_2\} \leq y \leq \max\{H_1, H_2\}$ the relevant comparison the agent makes in period one is between $V_b(A)$ and $V_b(B)$, which is given by:

$$V_b(B) = U(y) + U(y). \tag{7}$$

Thus, depending on the value of the outside option B, the agent will either explore in period one (then choose optimally in period two), or choose not to explore at all.

We can define another point of indifference, $\bar{y}_b$, such that the agent is indifferent between pulling arm A (the risky bandit) and arm B (the safe option) in period 1, i.e.,

$$U(\tilde{A}_1) + 0.5U(\tilde{A}_{2H}) + 0.5\max\{U(\bar{y}_b), U(\tilde{A}_{2L})\} = U(\bar{y}_b) + U(\bar{y}_b). \tag{8}$$

Definition 2. *The value $\bar{y}_b$, such that the agent is indifferent between earning $\bar{y}_b$ twice and playing the bandit arm A in period one, then playing optimally in period two, is the Subjective Gittins Index (SGI) for arm A.*

For more general problems, the Gittins Index is essentially the expected value from the optimal weighting of exploration and exploitation from the risky arm A, e.g., how valuable is the arm at any point in time if it is played optimally? We can adapt this measure here to account for risk preferences other than risk-neutrality; hence, "Subjective Gittins Index." The simplicity of the model allows us to isolate two key testable predictions regarding an agent's certainty equivalent for the lottery, her Subjective Gittins Index for the bandit, and the relationship between the two.

Prediction 2: *Holding preferences constant, the certainty equivalent is no greater than the Subjective Gittins Index, i.e., $\bar{y}_l \leq \bar{y}_b$.*

The intuition behind prediction 2 is fairly straight forward. In the lottery problem, the agent faces the lottery $\tilde{A}_1$, where the only relevant concerns are the current-period probabilities and payoffs. In the first period of the bandit problem, the agent also faces $\tilde{A}_1$; the difference, however, is that A contains more value than the lottery $\tilde{A}_1$ in isolation. A not only offers a current-period payoff, but also provides value in revealing useful information about its state. In other words, the lottery problem only has exploitation value, while the bandit provides both exploitation *and* exploration value.

Prediction 3: *Holding preferences constant, the difference between the certainty equivalent $\bar{y}_l$ and the Subjective Gittins Index $\bar{y}_b$ is increasing in H_2.*

This prediction is clear from equations 2 and 8. H_2 has no bearing on the lottery problem, so changes do not affect $\bar{y}_l$. In the bandit problem, the lottery $\tilde{A}_{2H}$ pays out the guaranteed amount H_2. So long as $U(\cdot)$ is increasing, as the high payoff H_2 increases on the left hand side of the equation, so the value $\bar{y}_b$ must increase on the right hand side. Intuitively, the first-period expected value of the lottery remains constant as H_2 changes. The exploration value associated with arm A, however, increases, as it is more valuable to learn the state if H_2 is larger.

1.3 Design

The experiment features a series of repeated lottery and bandit choices to examine how subjects view risky options with and without exploration value, designed around predictions 2 and 3. In the $WEAK$ treatment, subjects face relatively weak incentives to explore; in the $STRONG$ treatment, we increase these incentives by manipulating H_2. In both treatments, the experiment takes place in three stages: (1) a binary lottery stage, to familiarize subjects with making choices between fixed payoffs and random payoffs; (2) a BDM lottery stage, to elicit thresholds to estimate $\bar{y}_l$ for each subject; and (3) a BDM bandit stage, to elicit thresholds to estimate an individual-level Gittins Index $\bar{y}_b$ for each subject. In addition, the second period of the the bandit stage allows to examine the degree to which subjects adhere to prediction 1.[3]

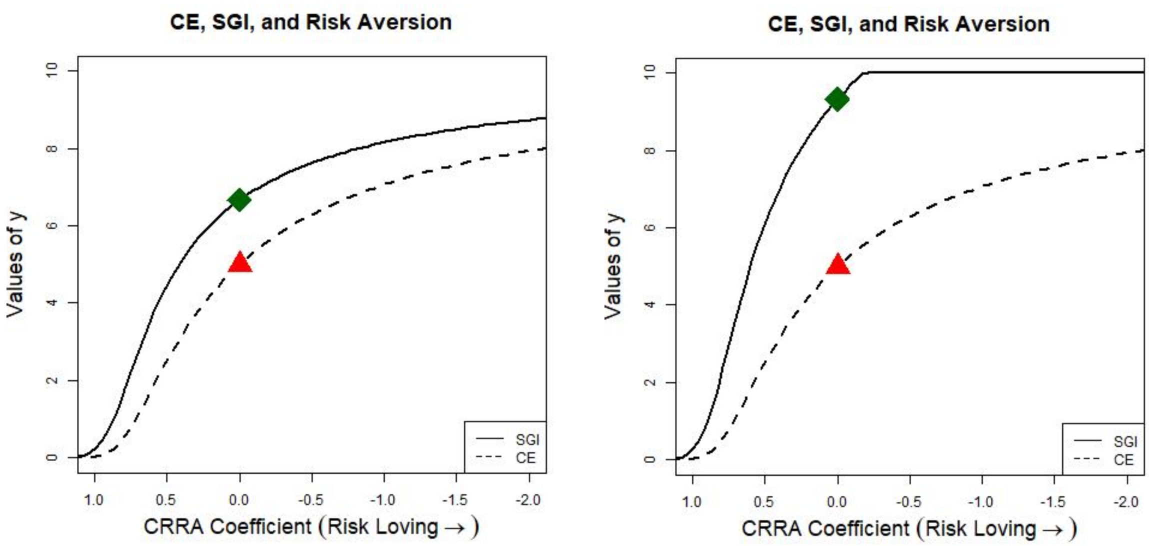
[3]Note that we could have designed the problem to constrain subjects to play the optimal strategy, i.e., to choose B twice or A optimally. Allowing flexibility in the second period, however, allows us to

Table 1: Treatments and Parameter Values

Treatment	L_1	H_1	L_2	H_2	Risk-Neutral $\bar{y}_l$	Risk-Neutral $\bar{y}_b$
WEAK	0.00	10.00	0.00	10.00	5.00	6.67
STRONG	0.00	10.00	0.00	18.00	5.00	9.33

Figure 1: Subjective Gittins Indices as a Function of Risk Aversion

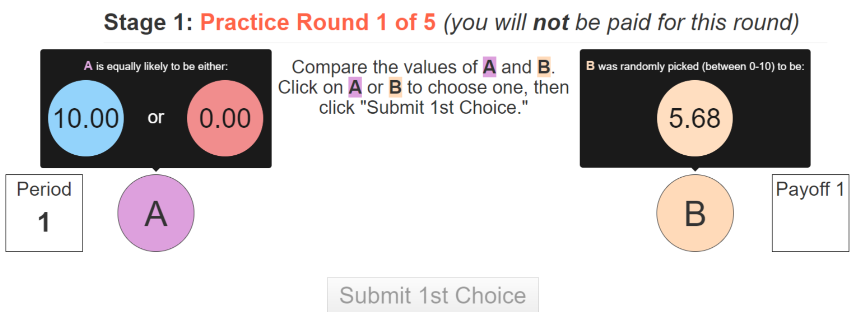
(a) *WEAK* (Full Sample) (b) *STRONG* (Full Sample)



Prior to starting a stage, instructions are read aloud and subject must take a quiz for comprehension. Subjects begin a stage by playing 5 practice rounds (which are not incentivized) to familiarize themselves with the interface and the problem. After completing the practice round, subjects are informed that practice is over, then move on to the first of 10 incentivized rounds per stage. After playing 10 problems, the stage ends and the game pauses until all subjects finish the stage. This serves two purposes: first, to pause the interface and task so that additional instructions may be given; second, to ensure that there is no individual benefit to playing quickly (i.e., so subjects cannot tradeoff between time and payoffs).

address this prediction directly and provides a richer environment to test how subjects view the trade-off.

Figure 2: Stage 1 Interface

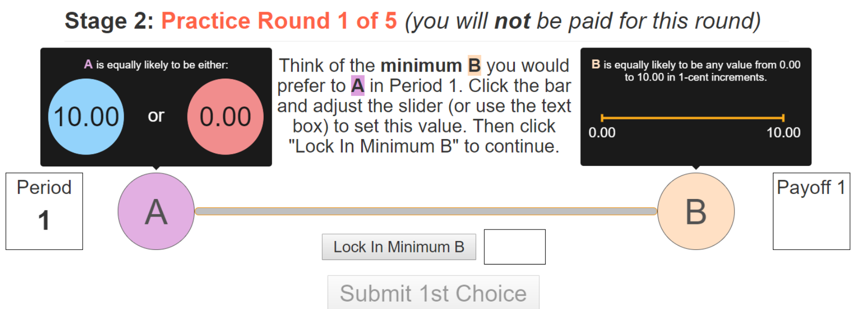


In all three stages, the computer randomly draws values at the beginning of each round. Arm B's guaranteed payoff y is drawn from a uniform [0,10] distribution (rounded to the nearest cent). This is to ensure that subjects see a variety of outside options while maintaining independence across rounds. A's state is equally likely to be G or NG. In both lottery stages, arm A pays $H_1 = \$10.00$ or $L_1 = \$0.00$. In stage 3, the payments for A depend on period and treatment. In the first period of both treatments, the payoffs are $H_1 = \$10.00$ and $L_1 = \$0.00$. In the $WEAK$ treatment, these values remain the same; in the $STRONG$ treatment, the values are $H_2 = \$18.00$[4] and $L_2 = \$0.00$. For a risk-neutral agent, these values imply a certainty equivalent of $\bar{y}_l = 5.00$ for the lotteries, while the risk-neutral Gittins Indices are \$6.67 for $WEAK$ and \$9.33 for $STRONG$. Table 1 summarizes the parameters for both treatments. Note that the values of the CE and SGI vary with risk preferences. These values are plotted in figure 1, emphasizing the importance of eliciting risk preferences to assess how subjects view exploration.

In the first stage, subjects make binary decisions between the risky arm A and the safe arm B, which takes on new values each round. The interface for the problem is presented in figure 2. Subjects see the two potential choices as circles, where the risky arm A is purple and the safe arm B is yellow. The information about each arm appears above the circles: A could be blue (the good state) or red (the bad state), paying \$10.00

[4]The high payoff is set to 18.00, rather than the more intuitive 20.00, to ensure the risk-neutral value $\bar{y}_b$ did not fall on the edge of the action space.

Figure 3: Stage 2 Interface–Initial

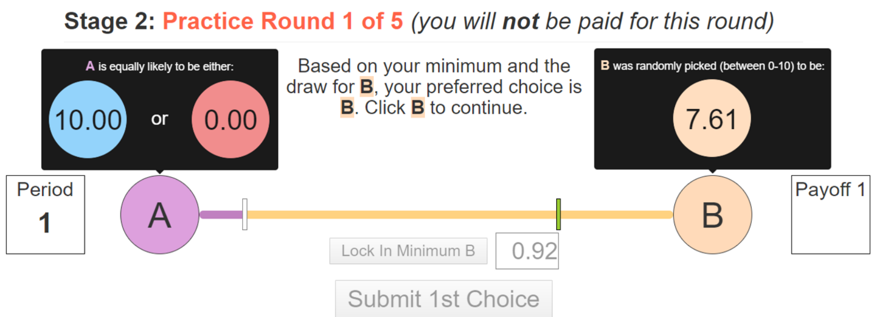


or \$0.00, respectively; B pays a constant value (which the subject knows when making the decision). Subjects simply click on one of two circles to make their choices, but the choice is not finalized until they click "Submit." After a subject confirms a choice and the payoff is realized, the earnings for the period appear in the box on the right. The software informs the subject that the round has ended, then draws new values for the next round. During stage 1, subjects are also asked to form a strategy: how high would y have to be to entice you to choose B instead of A?[5]

In the second stage of the experiment, subject play a BDM lottery task, with the same payoffs and randomization procedure as stage 1. In the BDM lottery, however, subjects must state the "minimum B" that they would accept (the strategy that they were asked to form in stage 1) *before* observing the guaranteed payment y. It is this value, w_l, that allows us to estimate a subject's certainty equivalent $\bar{y}_l$. To report w_l, subjects either manipulate a slider bar, or type in a textbox. The values they may report are limited to \$0.00 to \$10.00 in 1-cent intervals, mirroring the domain of y. The interface is presented in figure 3. The slider bar is initially gray to avoid priming subjects; the information about the outside option B appears as a number line to emphasize that all values between \$0.00 and \$10.00 are equally likely.

[5]The request to begin thinking in terms of a strategy is not incentivized; rather, it is part of the instructions to help subjects think about the problem.

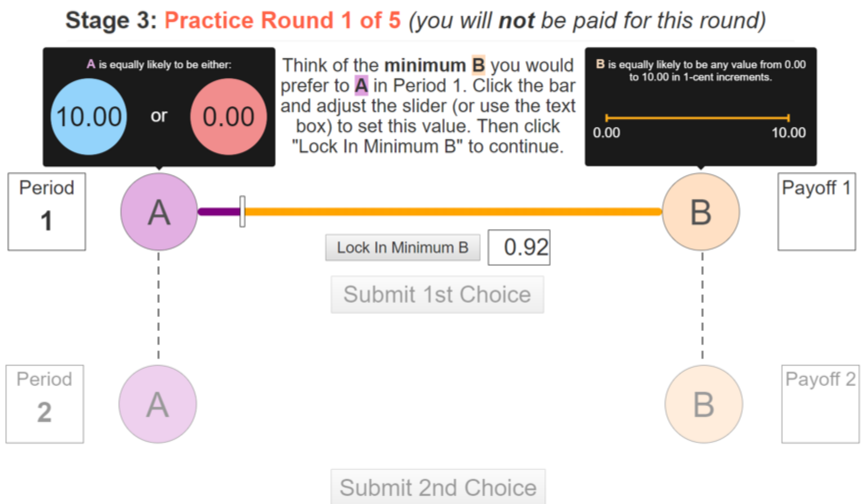
Figure 4: Stage 2 Interface–Post Lock-In



When a subject clicks on the slider bar or types a value in the textbox, a slider appears on the slider bar. Furthermore, the reservation value a subject states, w_l, serves as a threshold. If $y > w_l$, the software constrains the subject to play B; conversely, if $y < w_l$, the subject must play A. Because the value y is randomly selected, the stated value w_l affects the probabilities of playing A or B. The higher the reservation value, the more likely it is that the software constrains the subject to play A. To reflect this, once a subject states her w_l, the areas to the left and right of the slider-bar turn purple and yellow. These colored areas represent the relative probabilities of being constrained to play A or B, respectively.

Upon setting a threshold, the subject needs to actively "lock in" the choice. Once locked in, the software reveals the value of y. This value is represented both by an animated vertical green bar on the slider and a yellow circle appearing above B (replacing the number line). This screen is presented in figure 4. After viewing the value of the outside option, the subject then chooses between playing A and B. Though this is a phantom choice (subjects are constrained by their thresholds), it makes the valuation process more salient to subjects. The software informs them of their required choice, based on the threshold. If a subject tries to choose in violation of the threshold, they receive the message: "Your choice is not consistent with the Minimum B you selected. You will have an opportunity to change your value in future rounds." This is to assist

Figure 5: Stage 3 Interface



subjects in connecting the threshold choice (prior to observing the draw of y) with the set of binary choices made in the first stage.

In stage 3, the game is extended to two periods. The structure is almost identical to the stage 2 BDM lottery, with the simple addition of a second-period binary choice. The third stage interface is presented in figure 5. The first period is functionally the same as the second stage–subjects state the "minimum B" that they would accept. This value, w_b, allows us to estimate a subject's individual Gittins Index $\bar{y}_l$. The higher the stated valuation, the more likely it is the subject plays A; the lower w_b is, the more likely it is the software constrains her to play B.

After stating w_b, the value of the safe arm is revealed, the information about B updates, and the subject makes a constrained choice between A and B. The key difference in stage 3, however, is the presence of the temporally dependent second choice. The safe arm B pays out the same amount in both periods, while the state of A remains constant. Depending on the first-period choice, the information updates for the second-period arms (information appears below the circles). If a subject chooses A, she learns the state with certainty, and the information updates to be a single blue or red circle. If she chooses

B, the information remains the same.[6] After the information updates, the software then instructs subjects to choose between A and B a second time.

The crucial stages of the design are the second and third–the first stage is intended to familiarize subjects with the nature of the problem. In the latter two stages, the principal data of interest are the thresholds set by subjects. It is these thresholds that allow us to test predictions 2 and 3. Using the valuation w_l from each round, we can estimate a subject's certainty equivalent $\bar{y}_l$ for the lotteries. Similarly, if a subject is optimizing, then she states the value w_b such that she would be indifferent between playing the outside option B twice versus taking A in the first period, then playing optimally thereafter. Again, under this assumption,[7] the value w_b is the individual subjective Gittins Index for each subject. These stated valuations are the main point of the third stage. The second period choices serve to create exploration value and also allow us to test whether or not subjects play in a consistent fashion with prediction 1 (e.g., do they stay on good arms, switch off of bad arms, and avoid playing B then A?). To test prediction 2, we can compare the values w_l to w_b for each subject; to test prediction 3, we can compare the changes in these values in the $WEAK$ treatment to the changes in the $STRONG$ treatment.

A key to the design is that in both the lottery task and the bandit task, subjects face first-period (or only-period) lotteries with identical stakes and probabilities. In the lottery, the expected payoff is \$5.00; in the bandit, the expected first-round payoff is \$5.00 (the exploitation value), while the time-dependency and ability to switch off of a not-good risky arm creates exploration value (the downside of exploration is limited by the outside option B). Without the lottery task, it would be impossible to decompose

[6] In the $STRONG$ treatment, information for the second stage is present immediately (although non-interactive). This is to make sure subjects understand that the payoff grows for A if the state of the world is good.

[7] We can test this assumption empirically. A majority of subjects do play arm A optimally; more on this in section 4.

the two sources of value in the bandit. Because the exploitation value is initially the same in both the lottery and the bandit, we can attribute differences in thresholds (if any) to a subject's valuation of exploration.[8] Furthermore, the similarity is not limited to the payoffs. In addition to being the simplest possible bandit, the design minimizes the changes in the interface between the lottery and bandit problems. By making stages two and three as similar as possible, we limit the degree to which the interface causes changes in behavior.

Another important aspect of the design is the relative simplicity of the computations necessary to calculate expected payoffs. As mentioned in section 2, all probabilities subjects face are 0, 0.5, or 1.0. In addition to being fairly easy to understand, these probabilities make the bandit problem much less complicated. These problems ensure subjects need not be able to employ Bayes' rule with any sophistication. All uncertainty can be resolved in a single period, otherwise the probability at hand is 0.5. Similarly, calculating expected payoffs in the the bandit does not require the multiplication of probabilities strictly between zero and one. Reduction of lotteries in the game only requires weighting the two non-random "lotteries" by 0.5. Thus, all gambles under consideration are sure things or fifty-fifty propositions. While a failure to employ Bayes' Rule and/or reduce compound lotteries undoubtedly affects people's decisions in problems of exploration and exploitation, these phenomenon are not the focus of the study.

Sessions took place using undergraduate students in the Experimental and Behavioral Economics Lab (EBEL) at the University of California, Santa Barbara (UCSB) during 2019. Subjects were recruited from the undergraduate and graduate population using the ORSEE platform (Greiner, 2004). Subjects could participate in no more than one session. All sessions were conducted using software written in JavaScript, deployed on the oTree platform (Chen et al., 2016). Subjects were seated at terminals with concealed

[8] Note that these are not independent, as subjects may anchor on their decisions from stage 2. We run a robustness treatment, described in section 4, to assess this potential confound.

screens and did not communicate or interact during sessions; they were also informed that they were participating in an individual-decision problem.

Prior to the beginning of each stage of the experiment, subjects viewed instructions as they were read aloud. After the reading and listening to the instructions, subjects took a short quiz for comprehension.[9] Subjects viewed quiz results, then played five unincentivized practice rounds; this was to ensure that subjects understood both the interface and the structure of the environment. At the end of the practice rounds, subjects were informed that practice had ended, they were reminded of the payoff procedures, and were encouraged not to rush their decisions. Each stage consisted of 10 independent incentivized rounds. At any point, subjects could click buttons to view a summary of instructions and/or their history of choices and payoffs. After completing a stage, subjects were asked to sit quietly until all participants finished the stage and the next set of instructions began. Each session lasted less than one hour.

In addition to a \$5.00 show-up payment, subjects were paid their earnings from a single round. Note that payment was potentially *not* for a single choice; rather, if the random round was from the bandit stage, the accumulated earnings from both periods was paid. This is due to the dependent structure of the problem. Because first-period choices affect valuations in the second-period, it is not incentive-compatible to incentive a single period (Azrieli et al., 2018). Subjects could earn up to \$20.00 or \$28.00, in addition to the show-up fee, for the $WEAK$ and $STRONG$ treatments, respectively. 67 subjects participated in the $WEAK$ treatment; 66 participated in the $STRONG$ treatment. The subjects in the $WEAK$ treatment earned \$11.90 on average, while those in the $STRONG$ treatment earned an average of \$10.85.[10]

[9]The quiz was not incentivized. Subjects were informed that the quizzes were only meant to help them understand the interface and the problem–they were allowed to refer to a hard copy of the instructions as they answered questions.

[10]The discrepancy in payments (that the $STRONG$ treatment paid less than the $WEAK$) arises from the fact that the random round was selected at the session level. If the paying round was selected from one of the lottery tasks, then the maximum possible earnings were \$15 for the session, versus a maximum

1.4 Results

1.4.1 Consistency with the Model

The first prediction of the environment is that subjects play optimally in the second period of the bandit. If subjects violate this prediction, then the elicited values in the bandit task are not individual-level SGIs; recall that the Gittins Index relies on the assumption of future optimal play. We can test this prediction using the data from the second period of the bandit task in both treatments. There are three possible violations of the prediction 1: (1) failure to stay on the risky arm A when the state is good, (2) failure to *switch* off of the risky arm A when the state is not good, and (3) deviating from safe to risky (i.e., resolving uncertainty late rather than early). The first two cases turn out to be relative non-issues. The first mistake is literally non-existent: there is never a case in either treatment where a subject switches to B after finding that A is state G (out of 300 possible decisions). The second mistake is somewhat more common, but still infrequent, with subjects staying on arm A known to have state NG only 3.8% of the time (12 times out of 312 possible).[11] Thus, these mistakes are not a substantial concern for the analysis.

The third mistake, however, is far more common. In the $WEAK$ treatment, subjects switch to the risky arm in period two 22.5% of the time (81 out of 363 possible) when they are constrained to take the safe option B in period one. This is clearly a mistake, given the fact that resolving uncertainty early stochastically dominates the safe-to-risky-arm pattern. Similarly, across the $STRONG$ treatment, slightly more than 35% of choices follow this pattern (124 times out of 355 possible). This behavior is also quite localized in both treatments; pooling the data, 40% of subjects never make this switch, while over

of \$25 in $WEAK$ and \$33 in strong. Note that several subjects made the minimum amount (the \$5.00 show-up fee).

[11]This was also fairly localized; more than half of these mistakes were committed by only two subjects.

63% do not do this more than once. These subjects who do not routinely make these mistakes in the second period are "Gittins Consistent," in that their play satisfies the basic requirements for their stated valuations to be an individual-level SGI. The rate at which the population makes the safe-to-risky mistake, moreover, does not vary much over time, indicating that the sub-optimal behavior is persistent for about a third of the population.

The second prediction addresses the direction of the valuation adjustment as subjects move from the lottery to the bandit. Given that subjects' choices are somewhat noisy, we consider the average lottery and bandit valuations, $\bar{w}_{li}$ and $\bar{w}_{bi}$, respectively, for each subject i.[12] We classify the average thresholds within $0.10 of each other as "no change" in valuation; while changes of greater than $0.10 indicate increases or decreases. Using this criterion, in the $WEAK$ treatment, a plurality (43.3%) of subjects set higher thresholds in the bandit task, while a substantial portion of the subject pool (17.9%) do not change their thresholds. More puzzling is the 32.8% of subjects who set *lower* valuations in the bandit than the lottery. The results are similar in the $STRONG$ treatment: 59.1% increase their valuations; 21.2% do not change; 19.7% decrease them.[13] Again, the reason for decreasing valuations is unclear. Arm A in the bandit offers objectively more value than A in the lottery–the risky arm in the bandit offers the same current-period payoffs, but also provides value in revealing the state. Thus, these decisions are not consistent with monotonic preferences.

We can combine the second-period behavior bandit with subjects' thresholds in the second and third stages of the experiment to form a taxonomy of types. Subjects are classified as a "Decreaser", "No Change", or "Increaser" as described above. To classify a

[12]The results do not change substantially if we focus on other measures, such as the median choice, or the average for the last few periods.

[13]If we take these classifications literally (i.e., only subjects who set *exactly* identical average thresholds, a majority technically increase their valuations: in $WEAK$, 52.2% increase, 10.5% do not change, and 37.3% decrease; in $STRONG$, 59.1% increase, 15.1% do not change, and 25.8% decrease.

Table 2: Classification by Type

Treatment		Lottery-to-Bandit Thresholds: Decreaser	No Change	Increaser	Total
WEAK	Gittins Inconsistent	10	4	6	20
		(14.9%)	(6%)	(9%)	(29.9%)
	Gittins Consistent	12	12	23	47
		(17.9%)	(17.9%)	(34.3%)	(70.1%)
	Total	22	16	29	67
		(32.8%)	(17.9%)	(34.3%)	(100%)
STRONG	Gittins Inconsistent	8	5	8	21
		(12.1%)	(7.6%)	(12.1%)	(31.8%)
	Gittins Consistent	5	9	31	45
		(7.6%)	(13.6%)	(47%)	(68.2%)
	Total	13	14	39	66
		(19.7%)	(21.2%)	(59.1)%	(100%)

This table presents the distribution of subjects in each classification. Subjects who make a 2nd-period mistake more than once after period 3 are classified as "Gittins Inconsistent," and "Gittins Consistent" otherwise. Subjects who set thresholds such that $|\bar{w}_{bi} - \bar{w}_{li}| \leq 0.10$ are considered "No Change". If the difference is greater than 0.10, a subject is labeled as a "Decreaser" or "Increaser" based on the sign of the difference.

subject as Gittins Consistent, she must not make any of the second-period mistakes more than once after the third period. This creates six categories, presented in table 2. In both treatments, around 70% of subjects are classified as Gittins Consistent. Note that setting the cut-off at period 3 does not substantially affect the classification in $WEAK$; moving the cut-off period from 2 to 3 in $STRONG$, however, causes 9.5% more subjects to appear consistent. This is not a major concern, as early-stage mistakes are expected given the increased incentive in period two. If subjects take B in the first period in $STRONG$, the second choice is between y and the lottery $(0.5 \circ 18.00, 0.5 \circ 0)$, with expected payoff

\$9.00. As the stage progresses, some subjects learn out of this sub-optimal behavior.

More importantly, a clear plurality of subjects both increased their thresholds and played consistently in the second period–34% in $WEAK$ and 41% in $STRONG$. These subjects are classified as "Near-Optimal", in that their play qualitatively conforms to the model's predictions for an optimizing agent. The pattern of second-period and threshold choices leads us to our first result:

Result 1. *A majority of subjects' play is consistent with the assumptions of the Gittins Index; a majority of subjects' choices are consistent with monotonicity of preferences. Furthermore, subjects are approximately twice as likely to be Gittins Consistent and increase their thresholds than any other pattern of behavior.*

1.4.2 Threshold Choices

We designed the experiment with two main questions in mind: (1) whether or not subjects recognize the value of exploration and (2) if subjects respond systematically to changes in exploration value. To quantitatively answer these questions, the analysis focuses on the thresholds set by subjects in the lottery task and the bandit task, in light of the heterogeneous behavior of subjects. That is, we can use $\bar{w}_{li}$ and $\bar{w}_{bi}$ to examine the subjects' attitudes towards exploration and exploitation, using the taxonomy to evaluate different subsets of the data. In particular, we are interested in the subset of subjects who are Gittins Consistent, and within that subset, the subjects who increase their valuations.

Table 3 provides summary statistics for the two treatments. In the $WEAK$ treatment, the full sample average $\bar{w}_l$ is \$4.63, while in $STRONG$ is is \$4.24. This points to the importance of using both tasks in the design–in particular, the subject pool in $STRONG$ looks fairly risk averse. Having this control is important in distinguishing subjects' attitudes towards exploration from their risk tolerance. In the bandit tasks, subjects in

WEAK set an average threshold of \$4.83, while in *STRONG*, the average threshold is \$5.01. The increase in *WEAK* is a positive \$0.20, but not significant based on a paired t-test (p-value of 0.219); the increase in *STRONG* is \$0.77 and statistically significant (p-value of less than 0.001).[14] Note that these data contain observations from subjects who make persistent mistakes in the second period, however, which confounds the notion that we are comparing certainty equivalents to Subjective Gittins Indices (recall that consistent second-period behavior is a necessary assumption for values to be considered Gittins Indices).

When we restrict the samples, note that the average bandit thresholds increase substantially. Focusing on Gittins Consistent subjects, the average *WEAK* Subjective Gittins Index is \$5.30, while in *STRONG* it is \$5.46. The differences increase as well, to \$0.51 (p-value of 0.005 based on a paired t-test) and \$1.09 (p-value of less than 0.001), respectively.[15] Restricting the sample further to include only those subjects who set higher SGIs than CEs, these values necessarily increase. It is still instructive, however, to focus on those subjects who most closely adhere to the model's predictions. In *WEAK*, the average difference of \$1.41 is nearly as large as the risk-neutral predicted increase (\$1.67), though the increase of \$1.66 in *STRONG* still falls well short of the risk-neutral benchmark (\$4.33). These data address our second prediction:

Result 2. *On average, subjects recognize that exploration is valuable, setting higher SGIs than CEs, holding current-period expected payoffs constant. There is heterogeneity, however; the difference between $\bar{w}_b$ and $\bar{w}_l$ among subjects who are Gittins Consistent is nearly twice as large as the effect among the full population.*

[14] Employing non-parametric paired Wilcoxon tests, the p-values are 0.200 and 0.000, respectively.

[15] Again, employing non-parametric paired Wilcoxon tests, the p-values are 0.015 and 0.000, respectively.

Table 3: Summary Statistics

Treatment	Subset	$\bar{w}_l$	Med. w_l	$\bar{w}_b$	Med. w_b	$\bar{w}_b - \bar{w}_l$
$WEAK$	Full Sample	4.63	4.50	4.83	5.00	0.20
		(0.191)		(0.224)		(0.161)
	Gittins Consistent	4.80	4.98	5.30	5.30	0.51
		(0.246)		(0.243)		(0.171)
	GC-Increaser	4.46	4.40	5.87	6.10	1.41
		(0.369)		(0.337)		(0.204)
$STRONG$	Full Sample	4.24	4.50	5.01	5.00	0.77
		(0.169)		(0.212)		(0.183)
	Gittins Consistent	4.37	4.76	5.46	5.50	1.09
		(0.212)		(0.254)		(0.240)
	GC-Increaser	4.24	4.51	5.91	6.68	1.65
		(0.237)		(0.285)		(0.296)

This table presents the breakdown of subjects based on second-period decisions and individual changes in threshold values between lottery and bandit tasks. Subjects who more than once play safe-then-risky or stay on a not good arm are classified as "2nd Period Inconsistent," and "Gittins Consistent" otherwise. Subjects who set thresholds such that $|\bar{w}_{bi} - \bar{w}_{li}| \leq 0.10$ are considered "No Change". If the difference is greater than 0.10, a subject is labeled as a "Decreaser" or "Increaser" based on the sign of the difference. Standard errors are in parentheses below each estimate.

In both treatments, there is evidence that most subjects value the bandit more than the lottery, particularly among those who play consistently in period 2. The relative sizes of the changes systematically vary with the incentives to explore–regardless of sample, the average differences between $\bar{w}_b$ and $\bar{w}_l$ are always larger in $STRONG$ than $WEAK$. In the full sample, the $STRONG$ difference of \$0.77 is statistically greater than the \$0.20 in $WEAK$, based on a two-sample t-test (p-value of 0.021). Similar, among the consistent subjects, the differences are further apart and marginally significantly different at \$1.09 and \$0.51 (p-value of 0.051).[16] This gap closes, however, when we restrict the sample to the Near-Optimal subjects; in $STRONG$ the average difference is \$1.66 compared to \$1.41 in $WEAK$. Due to power issues, these are not significantly different. The difference in treatment effects addresses the third prediction of the model.

Result 3. *Subjects systematically respond to changes in exploration value; they increase their valuations more when exploration is more valuable, holding current-period payoffs constant. The differences between $\bar{w}_b$ and $\bar{w}_l$ are larger when the the incentives to explore are stronger.*

While the averages are instructive, note that the predicted change is not identical across all risk preferences. For instance, optimal subjects whose preferences are sufficiently extreme would have CEs and SGIs that are nearly identical, e.g., a perfectly risk-loving agent would set valuations of \$10.00 in both tasks.[17] Thus, it is beneficial to look at the distribution of choices in addition to the averages. While we we have established that behavior is noisy at the individual level, it is useful to examine the distributions of $\bar{w}_{li}$ and $\bar{w}_{bi}$. If most subjects recognize the value of exploration (as opposed to a few outliers), we should see general rightward shift in the distributions of thresholds from lottery to bandit. Figure 6 presents the CDFs of averages for both the full sample and Gittins Consistent subjects.

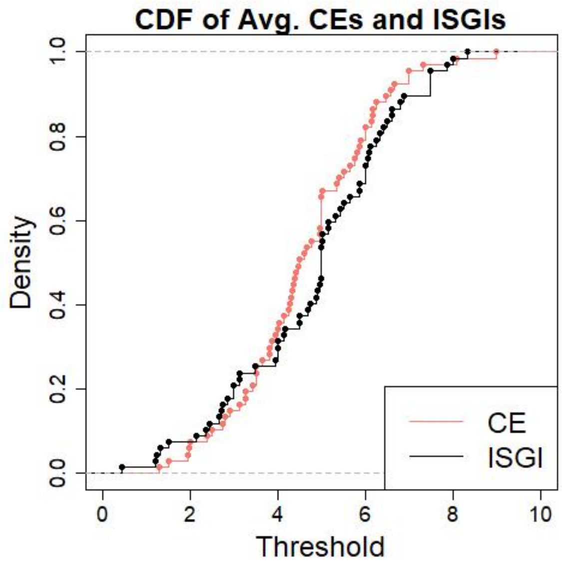
The CDFs for the full samples are in figures 6a and 6b. There is no clear difference in the distributions for the $WEAK$ full sample, consistent with the insignificant average change in table 3. The CDFs are not statistically different, based on a Kolmogorov-

[16] Using two-sample unpaired Wilcoxon tests, the p-value for the full sample is 0.061; it is 0.069 for the Gittins Consistent subjects.

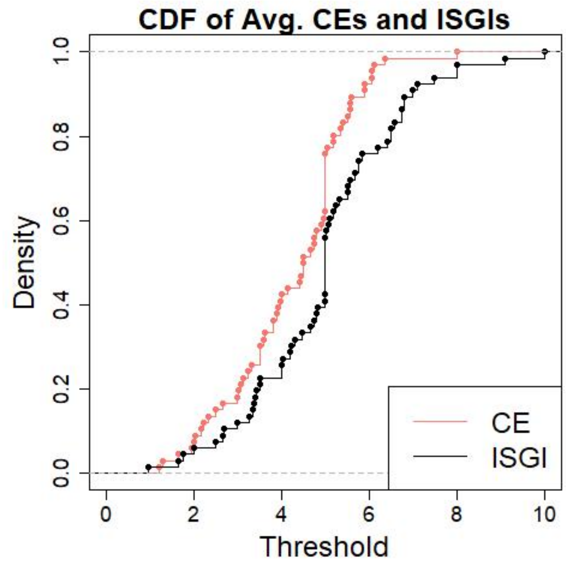
[17] Under the standard assumptions of choice under uncertainty, the predicted differences are maximized in $WEAK$ for a subject who sets a certainty equivalent of around \$3.00 (very risk averse). In $STRONG$, the largest difference is associated with a certainty equivalent of \$5.55 (risk seeking).

Figure 6: CDFs of Average Thresholds

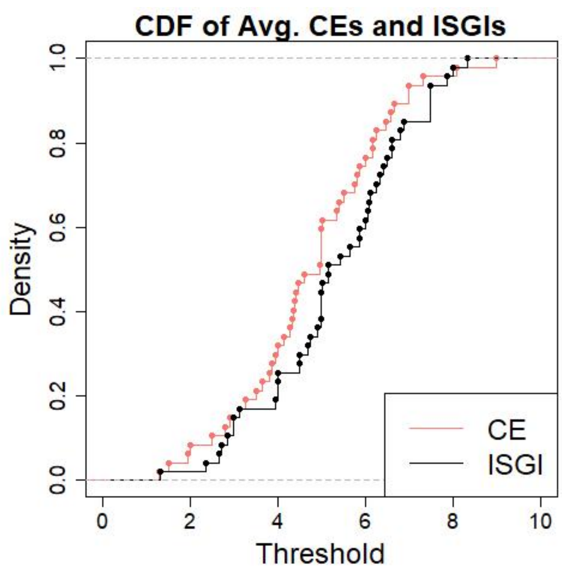
(a) $WEAK$ (Full Sample)



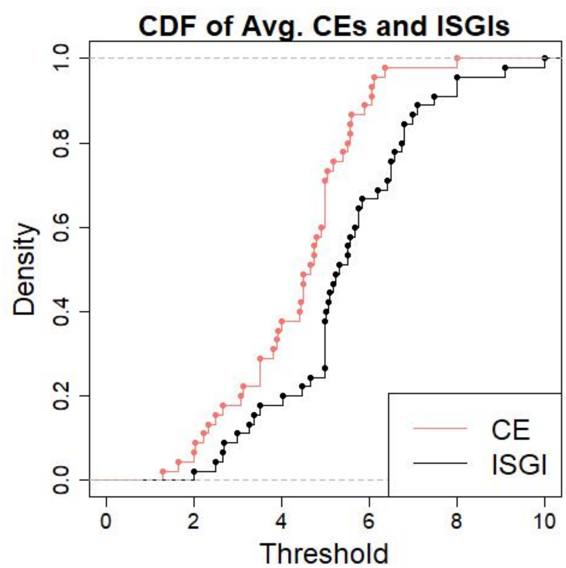
(b) $STRONG$ (Full Sample)



(c) $WEAK$ (Gittins Consistent)



(d) $STRONG$ (Gittins Consistent)



Smirnov (KS) test (p-value of 0.327). Figure 6b presents a clearer pattern, where the CDF of $\bar{w}_{bi}$ is very nearly stochastically dominated by CDF of $\bar{w}_{li}$, save for one observation(although these are not quite significantly different in a KS test, with a p-value of 0.103). The general rightward shift in CDFs of the $STRONG$ treatment, however, is consistent with the result that subjects recognize that there is some value to explore. The incentives are stronger in this treatment, so the more pronounced effect is unsurprising.[18]

We restrict the sample to those subjects who are Gittins Consistent in figures 6c and 6d. In these figures, we can reasonably treat the average valuations $\bar{w}_{bi}$ as individual SGIs, as the subjects adhere to the basic assumptions of optimal play. Given this subsample, the relationship between $\bar{w}_{li}$ and $\bar{w}_{bi}$ is quite clear in both plots. In the $WEAK$ treatment, there is a shift to the right in the bandit valuations, although these are not statistically different (p-value of 0.238 based on a KS test). In the $STRONG$ treatment, the SGIs are clearly stochastically dominated (and statistically different with a p-value of 0.007). These plots indicate that the consistent subjects systematically set higher thresholds in the bandit task, revealing that they value exploration as a group (despite some heterogeneity at the individual level).[19]

Based on these data, subjects appear to adjust their valuations in the direction that theory would predict across most of the action space. Once again, the consistent subjects behave much more closely to the predicted outcomes than the population as a whole. Quantitatively, however, subjects seem to under-weight the value of exploration relative to the risk-neutral prediction, even among those that conform to the qualitative predictions. While risk aversion would push subjects to under-explore, the differences

[18]The CDFs are almost identical if we examine the first half and last half of the data. Furthermore, as we formally analyzed below, there is no evidence that substantial learning takes place. If we focus on the *last round* only, the results are more significant; but these may be tainted by end-of-game effects and are not robust to the inclusion of additional periods near the end of a stage.

[19]Note that we have omitted the CDFs of the Gittins Consistent/Increasers, as by definition the CDF of SGIs would be stochastically dominated by the CEs (these subjects are typed as such based on increasing their thresholds).

are larger than individual risk preferences would predict. To evaluate this quantitatively, however, we need to apply more structure to the problem.

1.4.3 Risk Preferences and Exploration

While most subjects increase their valuations when the risky arm provides exploration, we need a measure of risk aversion to quantify the increase–remember that the Subjective Gittins Index relies on risk preferences. Under the null hypothesis, i.e., standard risk preferences fully account for behavior in bandits, a subject's risk aversion parameter should predict her bandit valuation and vice-versa. To measure subjects' risk preferences, we map the average certainty equivalent to a coefficient of risk aversion. While there are potential issues in measuring risk preferences in one domain to predict behavior in another, the experimental design specifically minimizes the differences in the tasks and interfaces to mitigate these concerns.

To measure risk aversion, standard practice is to apply structure to the utility function. Assuming that subjects have constant relative risk aversion (CRRA) preferences and Von Neumann-Morgenstern (VNM) utility, the utility function is given by:

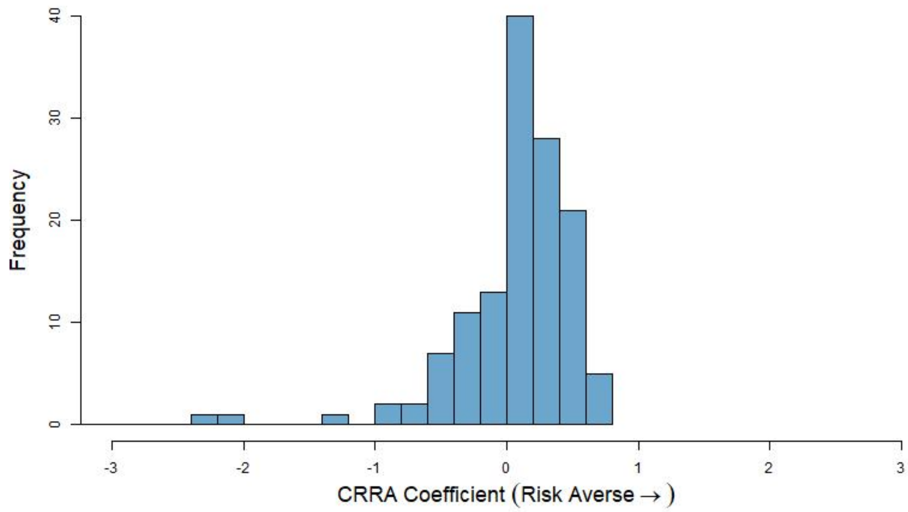
$$u(x) = \frac{x^{1-\sigma}}{1-\sigma}. \tag{9}$$

Using the average threshold from the BDM-lottery, subjects' averages map one-to-one from the action space to the risk-aversion parameter σ. For example, in the case of the lottery task, we find σ_i such that the following equation holds:[20]

$$u(\bar{w}_{li}) = \frac{1}{2}u(0) + \frac{1}{2}u(10). \tag{10}$$

[20]In actuality, we estimate using $\varepsilon = 0.001$ in place of 0, because for $\sigma > 1$, the utility is undefined at 0. The results are not sensitive to small changes in ε.

Figure 7: Distribution of CRRA Coefficients



The empirical distribution of CRRA coefficients is presented in figure 7. Our distribution is slightly shifted towards risk-seekers relative to Holt and Laury (2002), but is broadly consistent with measured risk preferences in other studies. A majority of our subjects are technically risk averse, with a plurality being quite close to risk neutral. The median CRRA coefficient is 0.048, while the mean is -0.035. Note that the transformation from thresholds to risk preferences is non-linear–while the average $\bar{w}_l$ implies slight risk aversion, the average coefficient of risk aversion indicates slight risk-loving attitude. There is a long left tail, however, as one subject (not pictured due to scale) set an average CE consistent with $\sigma_i = -5.69$.

Table 4 presents the coefficients of risk aversion associated with the population averages for each treatment. In $WEAK$, the full-sample average $\bar{w}_l = 4.63$ implies a $\sigma = 0.100$, corresponding to the "risk neutral" range in Holt and Laury (2002). In $STRONG$ treatment, the average is $\bar{w}_b = 4.24$, mapping to $\sigma = 0.192$, implying the average subject is "slightly risk averse." Restricting attention to the Gittins Consistent subjects, the subjects are marginally more risk-neutral: in $WEAK$, the consistent subjects have a coefficient of $\sigma = 0.057$, in $STRONG$ they have a coefficient of $\sigma = 0.163$.

Table 4: Average Risk Preferences

Treatment	Subset	$\bar{w}_l$	σ_l	$\bar{w}_b$	$SGI(\sigma)$	$Exp\%$
$WEAK$	Full Sample	4.63	0.100	4.83	6.37	11.50%
	Gittins Consistent	4.80	0.057	5.30	6.51	29.69%
	GC-Increaser	4.46	0.141	5.87	6.24	79.34%
$STRONG$	Full Sample	4.24	0.192	5.01	8.41	18.45%
	Gittins Consistent	4.37	0.163	5.46	8.57	26.05%
	GC-Increaser	4.24	0.192	5.91	8.41	39.83%

This table presents the average lottery valuations $\bar{w}_l$ (across all subjects), and the implied risk preferences assuming CRRA preferences and VNM utility; it also presents the average bandit valuations $\bar{w}_b$, the implied SGI based on σ, and the percentage of exploration value $Exp\%$, defined as $(\bar{w}_b - \bar{w}_l)/(SGI(\sigma) - \bar{w}_l)$.

For the consistent subjects who set higher thresholds (i.e., those that seem to best understand the problem), both treatments are slightly more risk averse, with $WEAK$ and $STRONG$ having coefficients of $\sigma = 0.141$ and $\sigma = 0.192$, respectively.

Using each subject's measured risk aversion σ_i, we can impute the associated SGI. That is, assuming that risk preferences fully characterize exploration, we can set a benchmark bandit valuation for each subject (recall the plot of SGIs in figure 1). A subject's valuation relative to the imputed SGI will tell us whether she under- or over-weights exploration relative to her risk preferences as measured in the lottery task. Figure 4 also displays the average individual-SGI as well as the imputed SGI as predicted by each σ. On average, it is clear that subjects under-weight exploration, in some cases substantially. In both main treatments, the subjects fall well short of the predicted SGI.

To calculate the proportion of exploration value subjects take into account, we can compare the average increase to the imputed increased using the following equation:

$$Exp\,\% = \frac{\bar{w}_b - \bar{w}_l}{SGI(\sigma) - \bar{w}_l}. \tag{11}$$

where $SGI(\sigma)$ is the imputed Gittins Index associated with the coefficient of risk aversion σ. This allows us to quantify how close subjects' average valuations conform to the theoretical predictions. Using this measure, in the full sample of $WEAK$, subjects appear to account for only 11.5% of the bandit's exploration value on average. Subjects in the $STRONG$ treatment are similar, accounting for slightly more at 18.45%. Gittins Consistent subjects get closer to the model's predictions, but still only come in at 32.2% and 28.9%, respectively. Again, this is a remarkable deviation from the baseline–even among those whose bandit valuation satisfy the Gittins Index assumptions, subjects account for less than a third of exploration value. Finally, we can focus on the Near-Optimal subjects. Given how we purged the samples of subjects for whom $\bar{w}_{bi} < \bar{w}_{li}$, we should anticipate substantial increases for both $WEAK$ and $STRONG$. Indeed, the percentage for $WEAK$ jumps to over 79%. Conversely, the percentage still less than half in $STRONG$: our restricted subject pool only account for 44.2% of exploration value.[21] This leads us to our fourth result:

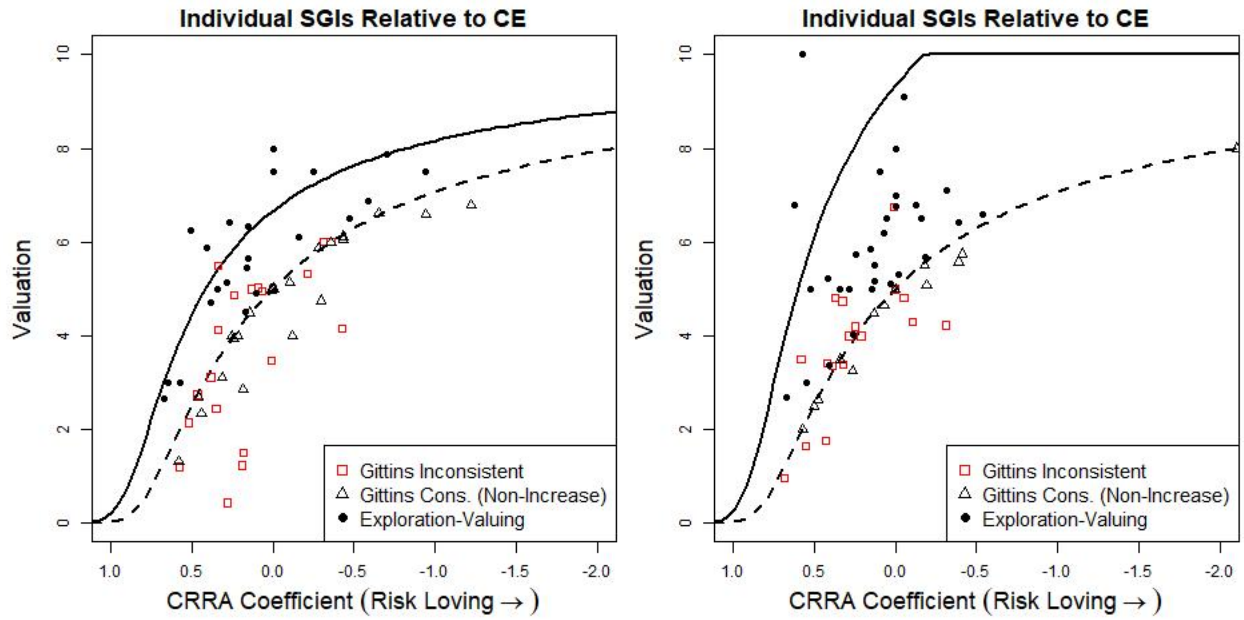
Result 4. *On average, subjects are less willing to explore than predicted by their risk preferences. In the full sample, subjects on average account for less than 20% of the risk-adjusted exploration value in bandits; for Gittins Consistent subjects, they account for less than 33%; among consistent subjects who set higher SGIs than CEs, they account for less than 80% in $WEAK$ and less than 40% in $STRONG$.*

To better understand which subjects drive these results, we can take advantage of the within-subject aspect of the design–we elicit risk preferences and individual SGIs for each subject. This allows us to consider how closely subjects come to fully valuing exploration

[21]Note that this is fairly robust to order effects. In the $REVERSE$ treatment, described below, the subject pool is less risk averse than $STRONG$; in the analogous restricted sample, however, subjects account for only 57.1% of exploration value.

Figure 8: Individual Gittins Index, by Subject

(a) $WEAK$ (b) $STRONG$



at the individual level. Figure 8 plots each subject's risk preferences and individual SGI. On the x-axis, we have subjects' CRRA coefficients of risk aversion; on the y-axis are valuations in dollars. The dotted line plots the certainty equivalents as a function of risk aversion; the solid line plots the predict SGIs. Each point on the plot represents a subject: squares represent Gittins Inconsistent subjects; triangles represent consistent subjects who did not increase their valuations lottery-to-bandit; and solid circles represent subjects who were both Gittins Consistent and set individual SGIs greater than CEs.

Points in figure 8 should be interpreted via vertical distance to the solid and dashed lines. Relative to the dashed certainty equivalent curve, points below represent subjects who set individual SGIs *lower* than their CEs. Points between the dashed CE curve and solid SGI curve represent subjects who under-adjusted their individual SGI relative to their measured risk preferences; points above the solid SGI curve represent subjects who over-adjusted. The plots show that while there is some heterogeneity, the under-weighting of exploration is a remarkably consistent phenomenon. In $WEAK$, only eight subjects (11.94%) set individual SGIs at least as high as the SGI implied by their risk preferences (although a ninth subject is within a few cents). In $STRONG$, the pattern is even more striking: only two subjects (3.03$) set individual SGIs higher than their measured risk preferences would imply. The vast majority of subjects set lower (in some cases much lower) SGIs, indicating that under-exploration is the consistent pattern.

Again, we can focus on the subjects who are Gittins Consistent (triangles and solid circles) and the subset who are consistent and set higher SGIs than CEs (solid circles only). In both treatments, these subjects display the same overall pattern as the full sample. That is, even the subjects who conform to the basic optimality conditions display under-exploration. When we restrict the sample to the consistent increasers, however, the pattern in $WEAK$ looks much closer to the model's predictions–subjects appear somewhat more evenly distributed above and below the imputed SGI curve. Conversely, in $STRONG$, restricting the sample does not substantially affect the distribution of points–the vast majority of consistent subjects fall between the two curves. These patterns form the basis for our next result.

Result 5. *The vast majority of individuals under-explore in bandits relative to their risk preferences, even when restricting the sample to Gittins Consistent subjects. Among consistent subjects who set higher SGIs than CEs, under-exploration is less prevalent when the incentives to explore are weak; when the incentives are stronger, subjects still display a substantial tendency to under-explore.*

This is one of the central findings of the study: subjects persistently under-explore. Despite the relative simplicity of the environment, subjects appear substantially less willing to take the bandit risky arm than implied by their risk preferences. Under the standard expected utility assumptions, the CEs set by subjects should be commensurate with the SGIs. The results here, however, provide evidence that this is not the case, even among subjects who are Gittins Consistent. Note that in $WEAK$, however, the Near-Optimal subjects seem somewhat close to the theoretical benchmark than in $STRONG$. This discrepancy is expected, due to the strength of the incentives–more marginal subjects probably respond to the stronger incentives to explore, whereas those who respond optimally to the weak incentives probably understand the problem better on average. Furthermore, the $STRONG$ treatment may better identify subjects, as it requires a larger adjustment to be consistent–a fixed willingness to pay for information, for example, is less likely to generate an individual SGI consistent with the predicted valuation.

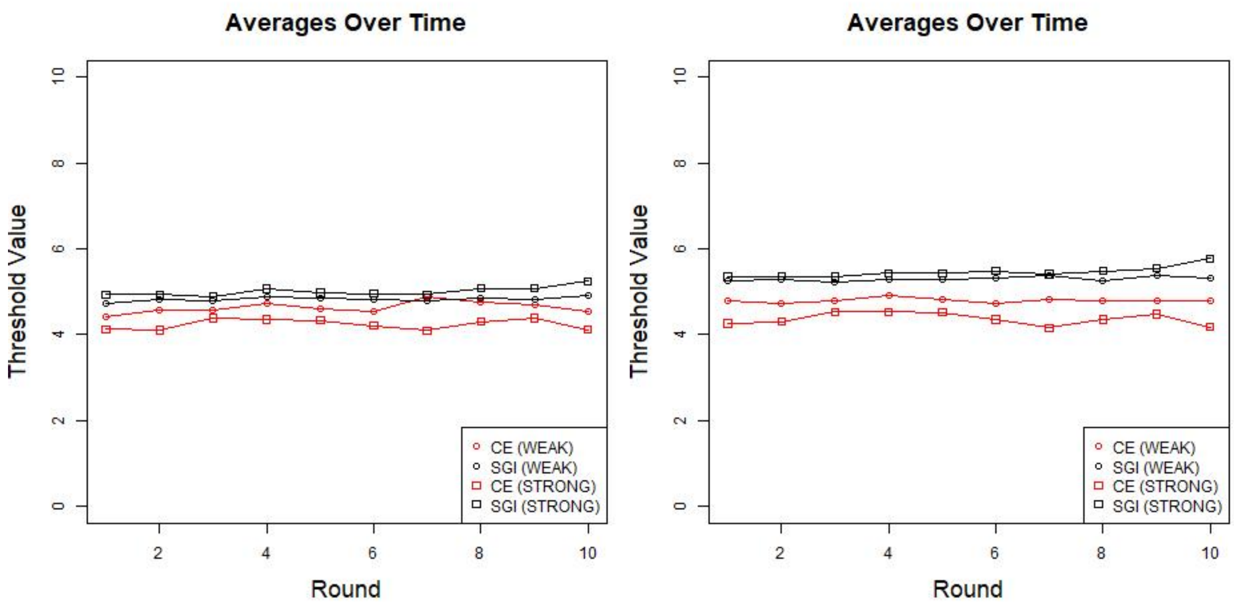
1.4.4 Robustness

There is a potential concern that due to the within-subject nature of the design, the order in which subjects experience the environment–lotteries first, bandits second–may drive the results. This is because the action space (the 0-10 valuation slider/textbox) remains fixed, but the incentives change when we add the second period. Thus, an inattentive subject who reflexively sets the same value in both tasks would appear to display a large under-adjustment. Even if a subject initially bases her SGI on her certainty equivalent (i.e., using the CE as a reference point), it would exacerbate the apparent discrepancy in predicted versus actual SGI. We address this potential issue by running the $REVERSE$ treatment, where subjects play the two tasks with the same incentives as the $STRONG$ treatment; the difference is that they play the BDM-Bandit prior to playing the BDM-Lottery. This treatment is described in Appendix A.1. There is evidence of order effects–subjects who play the bandit task first set higher valuations in both tasks–but the between-treatment comparisons are consistent with our within-subject results.[22]

[22] See the appendix for more detail on the $REVERSE$ treatment and the associated results.

Figure 9: Average CEs and SGIs By Period

(a) All Subjects (b) Gittins Consistent Subjects



Result 6. *The under-exploration in bandits relative to risk preferences measured in the lottery task is not due to order effects. Using between-subject comparisons, subjects appear more willing to explore when playing the bandit first, but still under-explore relative to the distribution of risk preferences for subjects who play the lottery first.*

We can also test to see whether or not subjects learn to optimize over time. If the under-exploration is simply a mistake due to the novelty of the problem, we would expect to see systematic adjustments in the SGIs as subjects gain experience. For example, if subjects gradually increase their individual SGIs as they play (e.g., they learn that exploration is valuable over time), using the averages over all ten rounds may exaggerate the degree to which subjects under-explore. Figure 9 plots the SGIs, averaged over the full sample and the Gittins Consistent subjects by period, for each treatment. From the plots, it is clear that there are no systematic patterns in the averages. Unsurprisingly, the consistent subjects in figure 9b display a slightly larger treatment difference between CEs and SGIs, but even the consistent subjects do not appear to adjust valuations by very much over the rounds. Indeed, for the consistent subjects in $WEAK$, the average CE and average SGI are remarkably stable.

Additionally, we can exploit the fact that each subject engages in ten incentivized repetitions of each bandit and lottery task. In particular, we can measure how previous realizations in the game influence subsequent threshold choices. Even if the averages are stable, since approximately half of subjects are in the good and not good states each round (since both states are equally likely and independent across subjects), it may be that subjects systematically adjust, but tend to offset on average. To further investigate potential learning in the environments, we can use a reduced-form regression, presented in equation 12.

$$w_{it} = \alpha + \beta_1 good_{it-1} + \beta_2 bad_{it-1} + \beta_3 \pi_{it-1} + \beta_4 t + \gamma_i + \varepsilon_{it} \tag{12}$$

The dependent variable is the current-period threshold; the covariates include lagged dummies for finding the state is good, finding the state is bad, and the previous period's earnings.[23] Note that these are not perfectly correlated (even in the lottery), as subjects can earn more or less having taken the "safe" option; in addition, in the bandit, subjects may explore in either the first- or the second-period, which will affect their earnings. Finally, the regression contains a variable tracking the period of the game and individual-level fixed effects.

Table 5 presents the regressions for the two tasks: column (1) contains the estimates for the lottery task, pooled across all three treatments ($WEAK$, $STRONG$, and $REVERSE$) . Column (2) contains the estimates from the pooled $STRONG$ and $REVERSE$ bandit tasks.[24] The standard errors are clustered at the subject level (with

[23]For the dummies, the omitted category is having taken the "safe" arm and not learned about the state

[24]Results from the $WEAK$ treatment are omitted for the bandit, as the full-population average

Table 5: Regressions Evaluating the Effect of Realized Outcomes on Thresholds

Variable	(1) Lottery (Pooled)	(2) Bandit (High-Incentives)	(3) Lottery (*REVERSE*)	(4) Bandit *STRONG*
$good_{it-1}$	0.088	−0.098	0.204*	0.058
	(0.131)	(0.138)	(0.099)	(0.128)
bad_{it-1}	−0.383	−0.060	−0.263	−0.360**
	(0.227)	(0.159)	(0.137)	(0.131)
π_{it-1}	−0.034	0.004	−0.036*	−0.012
	(0.030)	(0.011)	(0.018)	(0.009)
t	0.009	0.020	0.019	0.028
	(0.015)	(0.013)	(0.014)	(0.016)
n	1746	1143	549	594

This table presents fixed-effect regression results evaluating the effect of the realized state, G or NG, on the following period valuation. Column (1) pools lottery threshold data from $WEAK$, $STRONG$, and $REVERSE$ treatments. Column (2) pools bandit threshold data from $STRONG$ and $REVERSE$. Column (3) uses lottery threshold data from $REVERSE$; column (4) uses bandit threshold data from $STRONG$. Standard errors, clustered at the subject level, are presented in parentheses. * denotes significance at the 0.05 level; ** denotes significance at the 0.01 level.

nine observations per cluster on account of the lag). Columns (3) and (4) restrict the sample by treatment, focusing on the task with which subjects had less experience. In the $STRONG$ treatment, subjects play a binary lottery, a BDM-lottery, and a BDM-Bandit (so they have the least experience with the bandit). Conversely, the subjects in the $REVERSE$ treatment have more experience with bandit tasks, as they play both a binary and a BDM version of the bandit. Essentially, there may be less learning in the repeated environment, as behavior can stabilize somewhat in the binary task, leading to less adjustment in the associated BDM task.

In column (1), signs on the coefficients on $good_{it-1}$ and bad_{it-1} are consistent with subjects responding to previous outcomes. Relative to having taken the safe option, finding that the state was good increases subjects' thresholds (by about 9 cents), while

treatment effect was quite small compared to the two treatments with higher incentives.

finding the state was bad decreases them (by about 38 cents). Neither of these is statistically significant, however, with p-values of 0.503 and 0.091, respectively. At most, experience appears to have a negligible effect on subjects' play in the lottery task. In the bandit task, there does not appear to be any systematic effect from the realized state, with point estimates are all very close to zero.

It may also be the case that learning happens during the binary stages, so play has stabilized by the second stage of each treatment. For example, subjects are asked to start forming a threshold value during the binary lottery stage of $STRONG$. If subjects follow that advice, their thresholds may begin to stabilize prior to the BDM-lottery stage; a similar argument holds for the bandit task in $REVERSE$. Columns (3) and (4) focus on subsets of the data where subjects have less experience with the environment. In these regressions, there is more evidence of learning. In the $REVERSE$ treatment, subjects increase their certainty equivalents by roughly 20 cents ($p = 0.040$) following a good realization, while they decrease it by about 26 cents ($p = 0.055$). In column (4), the data suggest that there may be an asymmetric learning effect–subjects *decrease* their thresholds by about 36 cents after finding that the state is bad. Conversely, subjects do not appear to significantly increase their thresholds following a game where they learn the state is good.[25] Thus, it does not appear that subjects learn to increase their SGIs systematically in the bandit. This leads to our final result.

Result 7. *Under-exploration in bandits is robust to learning. There is very limited evidence that subjects adjust their thresholds in response to realizations in previous games. Furthermore, there is no evidence that subjects' differences between certainty equivalents and individual Subject Gittins Indices increase as subjects gain experience.*

[25]Running this regression on the $WEAK$ data yields no evidence of systematic adjustment of any kind.

Because the under-exploration is so robust to learning, it potentially points to a deeper cognitive mechanism at play.

1.5 Conclusion

Taken collectively, the results indicate that subjects do value exploration, although there is substantial heterogeneity in the population. In addition, subjects value exploration less than their risk preferences would predict. Indeed, the vast majority of subjects display this pattern, although it is muted among those who respond to weak incentives to explore. This implies that subjects tend towards considerable under-exploration relative to their risk preferences from a standard risk-elicitation task. This result, moreover, is highly robust to learning–we see no evidence that subjects increase their SGIs as they gain experience with the problem and the interface.

The simplicity of the design limits two of the more common mistakes in lottery problems, i.e., a failure of Bayesian updating and/or a failure to reduce compound lotteries. Again, all updating in the problem is as simple as possible, where all uncertainty can be resolved in a single round. Similarly, the problem does not require a reduction in compound lotteries in the traditional sense, as the first-period reduced lottery is a fifty-fifty weighting of two sure-things; no multiplication of two probabilities strictly between 0 and 1 is required. Given the simplified environment, it is not plausible that these are major drivers of the results.

Part of the discrepancy between choices in the lottery tasks and bandit tasks is due to subjects' apparent misunderstandings or mistakes in the game: non-trivial portions of the subject pool display non-monotonicity of preferences, while a non-mutually exclusive portion fail to recognize that resolving uncertainty early is stochastically dominant. Whatever the cause of these behaviors, subjects who make these mistakes sent consistently lower bandit thresholds relative to their revealed preferences in the lottery.

When we remove the inconsistent subjects, however, the discrepancy remains and is resistant to experience. Even among those who conform to the qualitative predictions of the model, i.e., the consistent subjects who set higher SGIs than CEs, the gap persists. In $WEAK$, subjects were closer to optimal, accounting for roughly 79% of exploration value on average. In the $STRONG$ treatment, however, these subjects only accounted for only 39% of exploration value. What explains the discrepancy between treatments? In part, it is an issue of incentives. Exploration in $WEAK$ is substantially less valuable–subjects who played in line with predictions did so in response to relatively low incentives to explore. In $STRONG$, however, the incentives to explore are much higher, and caused proportionally more marginal subjects to conform to the qualitative predictions–indeed, 59% of subjects increased their thresholds in the $STRONG$ bandit, relative to 34% in $WEAK$. As a result, the refined populations in each treatment are probably not directly comparable; the subjects in the $WEAK$ restrict sample are probably those that fully understand the problem.

How to think about this phenomenon? A possible model of under-exploration is to think about a fundamental misunderstanding of option value. Such models exists–Gabaix and Laibson (2005), for example, develop a model of directed cognition, where subjects cease to consider option value beyond some point in the future. Although closely related to option value, note that exploration value is a distinct phenomenon. While taking the risky arm in period 1 preserves the option to switch to the safe arm, typical option value problems do not have a "learn useful information for the future" aspect to them. That said, we can build a model of option value neglect to potentially rationalize the under-exploration we see in the data.

To do so, consider two important models in behavioral economics: correlation neglect (Enke and Zimmermann, 2017; Eyster and Weisacker, 2010) and analogy-based expectations (Jehiel, 2005). While these are game-theoretic models, we can adapt key features to help understand what it means to neglect option value. Consider an alternative calculation of the expected utility of playing the $WEAK$ risky arm in period 1:

$$V_b(A) = U(\tilde{A}_1) + (1-p)\Big[0.5U(10) + 0.5U(y)\Big] + p\Big[0.5U(\tilde{A}_1) + 0.5U(y)\Big]. \qquad (13)$$

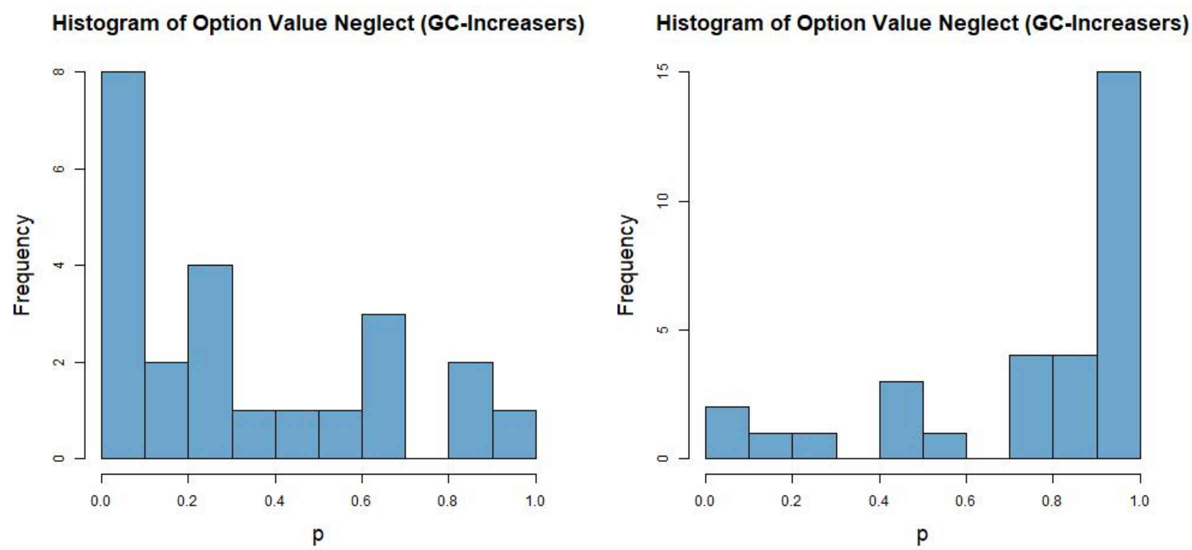
where $\tilde{A}_1 = \big(0.5 \circ 10, 0.5 \circ 0\big)$, i.e., the reduced fifty-fifty lottery in period 1, and p is the weight placed on the calculation that neglects correlations. The first term is the expected utility of playing the risky arm in period 1; the second term is the weighted correct expected value of the risky arm in period 2 (from the perspective of period 1); the third term is the weighted behavioral calculation.

This third term is the key to what it means to neglect option value. Inside the brackets, a subject neglecting correlation still realizes that she plays the risky arm and safe arm with equal probability in period 2 if she leads with the risky arm. In this sense, she is correct in the spirit of analogy-based expectations: her belief, on average, over her future actions is correct. How the outcomes in period 1 relate to her second-period actions, however, is incorrect. In this model, she believes that she will play each arm in period 2 half the time, *regardless* of period 1 outcome. Thus, she neglects the fact that the risky arm payoff in period 1 is perfectly correlated with her action in period 2. Note that if $p = 1$, then the agent fully neglects option value, and she views the risky arm as being independent across rounds.

Figure 10 presents histograms of the values of p that rationalize decisions for Gittins Consistent subjects who set higher SGIs than CEs. In figure 10a, there is a mass at 0; this represents the eight subjects who set valuations as high or higher than the implied SGI. Otherwise, the distribution is relatively uniform–subjects span the space and the model seems to fit the data reasonably well. In figure 10b, however, the pattern puts much more weight on 1. Why the mass at the full-neglect point? It is because it is the edge of the parameter space (it does not make sense to have *negative* weight on understanding

Figure 10: Weights on Option Value Neglect

(a) $WEAK$ (b) $STRONG$



correlations). Many subjects, even among those who are consistent and increase their valuations, actually fail to set SGIs that are as high as predicted with full option value neglect (given their risk preferences).

To illustrate this issue, consider the following example: as discussed above, a risk-neutral agent would set a certainty equivalent of 5.00; the full-exploration value SGI would be 9.33; in the correlation-neglect model, however, a subject who views the risky arm as independent over time would set an SGI of 6.33. We arrive at this value assuming the subject sets the SGI so as to make herself indifferent between taking the safe option twice and earning the expected value of the risky arm twice.[26] If the subject set an individual SGI in between 5 and 6.33, she would technically have increased her valuation moving from CE to SGI, but would have fallen short even of the valuation that neglects correlations. Roughly a third of $STRONG$ subjects in this subsample fit this description (adjusting for risk preferences). Thus, while the model rationalizes the $WEAK$ data

[26]The point of indifference is calculated by solving the following equation for s: $s + s = 5 + 0.5 \times 9 + 0.5 \times s$, (again, assuming risk neutrality)

fairly well, it does a poor job explaining subjects behavior in $STRONG$. It seems that instead, subjects seem to be somewhat myopic: while they recognize that there is more value to playing the risky arm A in period 1 when there is exploration value, they drastically over-weight the period 1 payoffs when making their decisions.

This is our preferred explanation, as most subjects probably do not (or cannot) calculate the exact expected values (despite the relative computational ease of the bandit problems considered here). In "intuiting" their way to a valuation, subjects may under-adjust. This interpretation is consistent with the discrepancy between the two main treatments: in the $WEAK$ case, the problem is (1) computationally easier, so more subjects calculate the exact expected payoffs, and (2) subjects have a much smaller distance to move (in terms of action space) than in $STRONG$. In the higher-incentive treatment, the problem is more difficult, so more subjects may fail to perform exact calculations. They know they should set higher SGIs than CEs, but are unsure of how much to change. The second-period payoff of 18 increases exploration value substantially: the risk-averse consistent SGIs require a near-doubling of the thresholds, so that may contribute to subjects' systematic under-adjustment.

There are other possible explanations for why subjects under-explore. Our experiment, for example, cannot rule out that subjects process risk differently in problems of exploration than in standard lottery tasks. In particular, subjects may behave as though they are more risk averse over state-uncertainty than over payoff-uncertainty. There is some evidence that this may be the case–in Halevy (2007), for example, subjects display marginally more risk aversion over urns that are equally likely to contain all red or all black balls than they are over urns that are equally mixed. The payoff-relevant lottery is a fifty-fifty proposition in both cases (with the same expected value and variance), but the state-uncertainty seems somewhat less attractive to subjects. While we cannot rule out this type of phenomenon outright, it is not our preferred explanation– the disparity in our experiment is much larger in magnitude than would be expected given the existing evidence on state- versus payoff-uncertainty.

Bandit models are an important tool in understanding how people make decisions in environments where people must make a trade off between exploration and exploitation. While the models frequently assume that this trade-off can be fully characterized by risk preferences, this study provides evidence that this may not be the case. Subjects do appear to respond systematically to the value of exploration, but also appear to systematically under-weight exploration relative to current-period expected payoffs. This phenomenon persist despite experience in the simplest possible bandit problem.

If these results were due to simple mistakes due to the novelty of the problem or the interface, we would also expect learning to moderate these effects as subjects gain experience. Subjects' behavior, however, is quite stable over time. This implies that the under-exploration is a persistent phenomenon, pointing to something deeper in how subjects value exploration. Furthermore, there is a large degree of heterogeneity in decision making that is also persistent. In particular, a substantial portion of subjects show a tendency to resolve uncertainty late in the bandit. This is actually commensurate with the idea of correlation neglect; if subjects do not understand the dependency structure of the game, hedging (although technically incorrect) can maximize perceived payoffs. Furthermore, in many lab experiments, subjects display a tendency to randomize, even when repeating choices in identical environments (Agranov and Ortoleva, 2017). This too could drive some of the Gittins Inconsistency. The persistence of this mistake could also point to how people view the relative attractiveness of exploration, i.e., whether or not they feel secure in how much value they have already guaranteed for themselves. That is, people may be more willing to explore after guaranteeing some fixed amount.

More work is also necessary to identify the mechanism that drives the under-exploration, even among those who qualitatively conform to the optimal strategy. While this study documents that subjects appear much less willing to explore than predicted by their risk preferences, we cannot identify the exact source of the discrepancy. The evidence suggests that subjects neglect option value and behave somewhat myopically, but we cannot definitively rule out other mechanisms–it is possible that taking on risk to learn may engage a different cognitive mechanism than taking on risk for payoffs. Designing around identification of the mechanism is an exciting prospect for future research.

There are other promising avenues of research regarding bandit problems. This experiment, which provides some evidence on how subjects behave in bandits, grossly simplifies the problem and abstracts away from issues that complicate the exploration-exploitation trade-off in real economic problems. Issues such as ego, discouragement, or even attachment to particular geography can have deep implications for how people value exploration. For example, job search and learning one's own ability could lead to strategic ignorance (e.g., deliberately *not* exploring to avoid getting bad news). Similarly, emotional attachment to one's hometown could lead to under-exploration of new employment opportunities. Such real-world problems provide interesting possibilities for extending our understanding of the exploration-exploitation trade-off.

Chapter 2

Complexity and Procedural Choice: Evidence from Experimental Bandits

James Banovetz and Ryan Oprea

2.1 Introduction

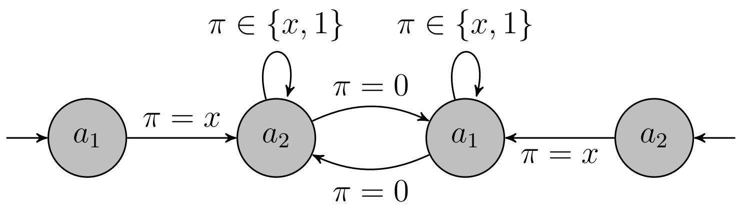
There is a long tradition in economics arguing that humans (i) dislike using complex rules to make decisions (Simon, 1955), and (ii) economize on this "procedural complexity" by substituting to simpler (and often suboptimal) rules instead. In the last few decades, the "automata literature" in economic theory (e.g., Aumann, 1981; Neyman, 1985; Rubinstein, 1986; Abreu and Rubinstein, 1988; Kalai and Stanford, 1988) has formalized this idea by modeling decision procedures (i.e., strategies) as *finite state machines*, models of algorithms developed in computer science. By advancing hypotheses about what makes one procedure/algorithm more "complex" than another, this literature provides a theoretical framework for studying the influence of complexity on human behavior. The original aim of this literature was to ease the indeterminacy of the folk theorem by using complexity to refine the set of feasible strategies in games and thus produce more predictive theory. In the decades since, this approach to bounded rationality has proven equally successful at rationalizing phenomena ranging from the emergence of Walrasian outcomes in markets (e.g., Sabourian, 2004; Gale and Sabourian, 2005) to some of the key anomalies documented in behavioral economics (e.g., Salant, 2011; Kalai and Solan, 2003; Chauvin, 2020; Wilson, 2014).

In this paper we provide a crisp experimental test of the core ideas from this literature. In the process, we evaluate the idea that (i) people use procedures to guide even simple behavior and (ii) that choice among procedures is fundamentally shaped by deliberate efforts to avoid complexity. Instead of studying repeated games (which introduce a number of complications orthogonal to our question), we study the closest analogue individual decision making task: a (particularly simple) two-arm bandit problem. In this task, subjects repeatedly choose between two "arms," each of which offers a fixed payment.The first arm "pulled" (chosen) always pays an intermediate value $x \in (0, 1)$, but the second arm pays 0, x or 1 with equal likelihood. Under our parameters, the subject should always "explore" by switching once from the initially selected arm, returning if the second arm pays 0 and "exploiting" the second arm (playing it forever) otherwise.[1]

To see how optimal behavior in this problem can be formally described as a procedure, we visualize the automata (finite state machine) description of the optimal procedure (in its general form) in figure 1. Each of the circles represents a *state*; each state both (i) specifies an action (a_1 or a_2, shown inside each circle) and (ii) is attached to a set of transitions (shown as arcs) to new states. Next to each arc is the event (here, the payoff π received after each choice) that triggers that transition. Finally, free standing arrows pointing at the states at the extreme left and right of the diagram show the two initial states the decision maker (DM) can choose from (by pulling one of the two initial arms). The procedure instructs the DM to start with either arm (given the payoffs, either a_1 or

[1]This bandit problem and the automata analysis of it discussed below are due to Börgers and Morales (2004).

Figure 1: Automata representation of the optimal 4-state procedure.



a_2 is equally good, and this choice determines whether the initial state is to the far left or far right); immediately switch to the other arm upon receiving the initial payoff of x (shown as transitions inward); and remain on the arm (the reflexive transition looping above each of the second states) if the payoff there is x or 1, but advance to the third state from the initial one if it pays 0. This third state instructs the DM to choose the initially selected arm forever.

The central idea of the automata literature is that people are averse to *complexity*, usually measured by the number of *states* in the procedure ("state complexity" or *s-complexity*). Every state added to a procedure requires the decision maker to pay richer attention to history and to process information in more distinct ways over the course of the problem. As Börgers and Morales (2004) show, a decision maker can economize on complexity (i.e. states) by using simpler procedures in this problem, some of which are quite suboptimal. For instance, by suboptimally choosing not to explore at all, the DM can remove three states from this procedure. By selectively (and again, suboptimally) randomizing in some histories she can remove two states. Finally, and most diagnostically interesting, by establishing the habit of always choosing the same initial arm, she can remove one state from the procedure without sacrificing any earnings.

To directly test this idea, we vary whether subjects must actually track states themselves when deploying procedures. Varying this state-tracking requirement varies precisely the costs hypothesized to drive procedural choice in auotomata models and does so cleanly, without actually changing anything about the payoffs, timing, or informational structure of the game. Each state added to a rule requires the decision makers to more carefully track (i) which arms she has chosen so far in the game and (ii) what payoffs have occurred in the history. In our State Tracking (ST) treatment we naturalistically require subjects to do this tracking themselves, while in our No State Tracking (NST) treatment we remove this burden: we show the subject, at each stage of the game, what arms they have pulled and what payments resulted from each arm. Everything else about the problem (payoffs, risks, and to the degree possible, instructions) are kept identical. Under the hypothesis that procedural choice is governed by aversion to state complexity, we should observe subjects choosing systematically lower-state procedures in State Tracking than in No State Tracking.

We find strong evidence in favor of this hypothesis. Despite the extreme simplicity of this problem, we find significantly suboptimal behavior in our baseline ST condition. The average subject under-explores and suffers significant payoff losses. Conducting a detailed analysis of individual decision making, we show that these averages conceal clear and widespread (though heterogeneous) use of simplified procedures by most subjects. Only 20% of our subjects behave randomly: most instead use consistent rules, albeit rules with fewer than the four states specified in the general optimal procedure. In fact, the modal subject uses an optimal 3-state procedure, with a smaller number using 1-state or 4-state procedures. There are indications in the data that even the minority of subjects who appear to randomize do so deliberately (i.e., choose to use a 0-state procedure).

In NST, where the complexity costs of states are removed, we find dramatically different behavior. Most subjects in NST use maximally complex 4-state procedures and virtually none use simple 1-state procedures (or 0-state randomization). Of particular importance, we find that even among subjects that optimize, NST and ST behaviors are dramatically different. Subjects in NST only rarely use habitual, initial-biased first arm-choices that economize on states, while almost all optimizing subjects do so in ST. This is strong evidence that subjects deliberately economize on complexity, and are capable of doing so in a relatively sophisticated fashion.

Our design also rules out a number of alternative explanations, allowing us to identify this systematic shift in procedural choice as resulting from subjects' distaste for and efforts to avoid state-complexity (as hypothesized by the automata literature). For instance, risk or other uncertainty preferences cannot explain the results, because the risk characteristics of the problem are identical in ST and NST. Likewise, the results do not arise because of confusion: the subset of subjects who score almost perfectly in a comprehension quiz also behave in systematically different fashions in ST and NST. The design also allows us to reject accidental mistakes generated by task difficulty as an explanation of our findings. First, the changes in behavior required to implement simpler procedures require deliberate and consistent sequences of actions that are unlikely to arise by accident. Moreover, subjects do not actually take longer to make decisions in ST than NST, suggesting that the former may not actually be more "difficult" than the latter (state-complexity and difficulty are not quite the same thing). Finally, we ran an additional diagnostic treatment that does not target states, but does make the problem generically more difficult. In our State Tracking + Cognitive Load (ST+CL) treatment, we repeat the ST treatment but require subjects to conduct a simultaneous memory task, causing decision making to become harder (subjects take twice as long to make decisions in this treatment). We find that, in contrast to NST, this has no systematic effect on the procedures subjects use, suggesting that task difficulty and resulting mistakes do not produce the sorts of shift in behavior we observe in our main treatments.

Our findings support the "procedures and complexity" approach to bounded rationality, as formalized in the automata literature, and suggest that this may be a promising source of explanation for both strategy choice in games and departures from neoclassical rationality in experiments and other empirical settings. Despite the promise of this approach as a way of integrating bounded rationality into economic theory, it has received surprisingly little empirical attention thus far. In a companion paper, Oprea (2020) conducts a very different experimental evaluation of the automata approach by exogenously assigning subjects artificial rules (meaningless algorithms specifying sequences of letters that must be typed in order to earn payoffs) and eliciting their distaste for implementing these rules again in the future. That paper finds that several automata characteristics– including, notably, states–generate significant subjective cost for subjects. By contrast, in this study we directly observe the consequences of such costs in naturalistic decision making by observing the procedures subjects *endogenously choose*, identifying the role of complexity in this choice by varying the need to track states. Unlike Oprea (2020), we observe subjects' own formulation of and choice over procedures and thus are able to study the degree to which automata models organize behavior in an actual decision task. Also related is Jones (2014), who compares simple prisoner's dilemmas to multi-state dynamic games; in the latter case the automata states required to cooperate is higher. This study finds evidence that subjects cooperate less often in the latter case, providing indirect evidence of aversion to state-complexity. By studying simpler bandit problems, we are able to directly identify the procedures used and we also are able to more clearly identify state complexity as the source of behavior by using a more targeted intervention that leaves the nature of the decision problem unchanged.[2]

A secondary contribution of our paper is to provide a detailed analysis of how people behave in bandit problems and why. Here, too, there is surprisingly little prior work given the importance and canonicity of the problem. Using a range of different bandit problems Banovetz (2020), Banks et al. (1997), Hoelzemann and Klein (2019), Gans et al.

[2]Although, experiments directly evaluating the complexity responses to automata are few, automata have proved useful in recent experimental work for taxonomizing and estimating strategies in repeated games. This literature uses automata to measure and taxonomize strategy choices in repeated prisoner's dilemmas either by using automata in specifying estimators (e.g., Engle-Warnick et al., 2007; Dal Bó and Fréchette, 2011) or by allowing subjects to choose or program automata directly when playing games (Dal Bó and Fréchette, 2019; Romero and Rosokha, 2018; **?**; Cason and Mui, 2019).

(2007), and Meyer and Shi (1995) find evidence of over-switching among their subjects, a pattern also found in many subjects in our data. Anderson (2012), Hudja and Woods (2019) and Banovetz (2020) find evidence of under-exploration in the average subject just as we do. Finally, like us, Gans et al. (2007) and Meyer and Shi (1995) find evidence that subjects employ suboptimal decision making routines (in their cases, systematic overweighting of recent information and/or myopically discounting the distant future). Unlike this previous literature we use automata to formally benchmark the complexity of the procedures used by subjects and exogenously vary the cost of this complexity to evaluate its causal role in behavior.

The remainder of the paper is organized as follows. In section 2 we describe the bandit problem we implement in the experiment, then characterize optimal behavior and describe a salient suite of procedures (alterantives to the optimal behavior) that economize on state complexity. In section 3 we describe the specifics of the experimental design. In section 4 we provide our main empirical results and we conclude with a discussion in section 5.

2.2 Theoretical Background

In this section, we first introduce the two-armed bandit problem that we implement in the experiment (section 2.2.1) and then provide an overview of finite automata models of decision rules (section 2.2.2), the foundation of how we think about complexity in this context. In section 2.2.3 we describe the set of procedures, modeled as finite automata of various complexity, that are potentially available to decision makers in our bandit problem. Much of the material in this section, particularly regarding the two-armed bandit, relies on modeling work in Börgers and Morales (2004) and we refer the reader there for more details and additional information.

2.2.1 A Bandit Problem

Consider the following simple bandit problem analyzed in Börgers and Morales (2004) and adapted for use in our experiment. In each period t, a decision maker (DM) must choose ("pull") one of two actions ("arms") $a_i \in \{a_1, a_2\}$; each period there is a probability $(1-\delta)$ that the game ends. Each arm a_i yields a fixed payoff $\pi_i \in \{0, x, 1\}$ ($x \in (0,1)$) every time the arm is pulled. The DM faces initial uncertainty about the payoff of each arm: ex ante, each arm takes each of its possible values (0, x or 1) with equal probability.

In the experiment we focus on an artificial but diagnostically valuable variation of this setup: the DM's first action (whichever action it is) is guaranteed to pay the intermediate value x and all uncertainty surrounds the value of the alternative arm. It is optimal to *explore* in this problem by immediately switching away from the initially selected arm, if x is smaller than the value of optimally exploring. If the DM "explores" by switching to the second arm, she should optimally (i) *exploit* the second arm by playing it forever if its payoff is 1, but should instead (ii) return to the initial arm and play it forever if the second arm's payoff is 0 (if the second arm's payoff is x then it does not matter what arm the DM chooses thereafter). Thus it is optimal to *explore* by immediately switching away from the initially selected arm if:

$$\left(\frac{x}{1-\delta}\right) \leq \frac{1}{3}\left(\frac{\delta x}{1-\delta} + \frac{x}{1-\delta} + \frac{1}{1-\delta}\right)$$

or

$$x \leq \frac{1}{2-\delta}.$$

We will assume this condition holds in what follows (and parameterize the experiment so it it holds in our design).

2.2.2 Automata and Rules

A decision maker uses a procedure (a rule) that specifies (i) the actions $A = \{a_1, a_2\}$ the DM takes in response to (ii) payoff *events*, $E = \{0, x, 1\}$, and, crucially, (iii) how these responses depend on history. Following the automata literature, we describe such decision rules formally using *finite automata*, simple models of algorithms imported from computer science (e.g., Hopcroft and Ullman (1979)). Formally, a *finite state machine* is a four-tuple (S, s^0, f, τ) where S is a set of *states*, s^0 is the set of initial states, $f : S \to R$ is an output function, specifying a response in each state, and $\tau : S \times E \to S$ is a *transition* function that specifies a next state as a function of the current state and the current event. Each row of table 1 represents an automata visually in the standard way, as a directed graphs with (i) circles representing states, (ii) arcs representing transitions, (iii) symbols inside of the circles representing actions from a set $A = \{a_1, a_2\}$ and (iv) expressions next to transitions represent events from a set $E = \{0, x, 1\}$. Free standing arrows pointing at states indicate possible initial states for the rule.

The top row of figure 1 shows the automata for the general form of the *optimal rule*, which contains 4 states. The DM first chooses one initial arm, thereby entering either the state to the far left or right of the graph. The arrows pointing inward from either initial state instruct the DM to immediately move to a new state, requiring her to choose the second (non-initial) arm, provided she receives a payoff of $\pi = x$ on the first arm (which, recall, in this version of the problem, she is guaranteed to receive, regardless of which arm she chooses initially). That is, the rule instructs the DM to immediately *explore* after her first choice. If the DM receives a payoff of x or 1 on the second arm, the rule instructs her to remain in this state (this is represented by the reflexive, looping transitions above the each of the two inner states). If, however, she receives a low payoff of 0 on the second arm, she transitions one more state away from her initial state and, given the structure of payoffs, will remain at that state forever.

This rule requires four states. Why? First, the rule must account for what the DM has observed in the past and thus must include two distinct states that both specify choice of the initially selected arm: one state reached prior to exploration (i.e., the state all the way to the left or right), another reached after exploration in case the second arm turns out to have a payoff of 0 (the interior states). Second, the rule must remember which arm the DM initially chose, so it can specify a return to that arm in case of a second-arm payoff of 0. Crucially, the return to the initial arm requires a state distinct from the initial state, to distinguish between pre- and post-experimentation. Since there are two possible initial arms, the general rule must contain states for each.

2.2.3 Complexity and Simple Rules

The automata literature studies how procedures vary in their *complexity* and it generally defines the complexity of a procedure as the number of *states* it contains (sometimes called state complexity or "s-complexity"). A DM that is averse to complexity will *ceteris paribus* prefer rules with fewer states, and may even be willing to sacrifice earnings to use rules that economize on states if her complexity aversion is strong enough.

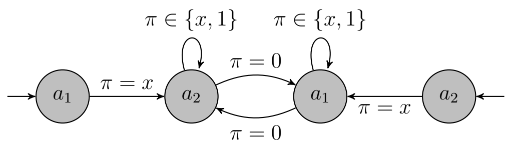
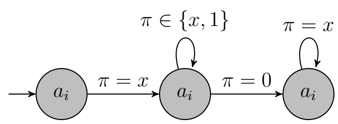
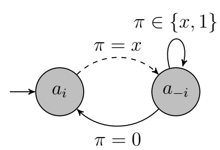
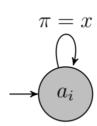
The lowest cost way of reducing complexity in our bandit problem is to establish a long-run habit or cross-game consistency, making use of what we will call (following Börgers and Morales (2004)) *initial bias*: each time the DM faces the decision problem, she uses a rule that always begins with the same initial arm (which can be a_1 or a_2). By doing this, the DM can simplify the procedure she deploys in the problem:. Intuitively, the DM no longer has to track her initial choice, which removes a state from the rule. As the second row of table 1 shows, using such an initial-biased rule allows the DM to reduce the rule's states at no cost. The DM now always plays a_i initially, immediately explores by playing a_{-i}, and transitions to a terminal state specifying a_i again if a_{-i} turns out to pay zero. This rule is simpler than the general version of the optimal procedure, but is still optimal: because a_1 and a_2 are symmetric ex ante, there is no cost to habitually choosing one of the arms initially and building a simplified rule around this habit.

A more costly but more effective "tool" for avoiding complexity is to *neglect history*, by processing information (payoffs) at each arm identically regardless of what has occurred in the past. If the DM is willing to forget whether she has observed a 0 payoff in the past and randomize her arm-switching behavior every time she observes a payoff of x, she can implement the *2-state* procedure shown in the third row of table 1. This rule specifies that the DM start with some action, a_i and transition from that initial action to a_{-i} with some probability p (the dotted transition line from the first state specifies a *stochastic* transition, made with some iid probability). At the second arm, if she receives 0, she returns to the initial action a_i immediately, but does not keep track of this history (note that if she earns 1 with second pull, she will remain on that arm forever). Instead, she transitions once again from the first arm each period with probability p. Randomization allows the DM to reduce states, and if p is well-calibrated, she can do this in a minimally costly way (in our parametrization in which $x = 0.65$ the optimal choice of p in a 2-state rule is 0.19).

More radically still, the DM can simply avoid exploration altogether, employing a *non-exploration* procedure. In our setting, this guarantees the DM the known intermediate payoff of x, but reduces the procedure to a very simple 1-state rule. The final row of table 1 shows the 1-state automata, which simply specifies remaining on the initial action after each realization of its payoff of x.

Finally, a subject can avoid state complexity altogether by simply randomizing–a "0-state" rule that requires no tracking of payoffs, histories, or (unlike the 1-state rule) even recalling actions just taken. This too is a kind of procedure, though one that lowers payoffs relative even to a 1-state rule. Unlike the other procedures, a randomization rule involves strictly interior probabilities of switching arms each period over all histories.

Table 1: Automata representations of four procedures.

Rule	Automata	Expected Payoff
4-state *General Optimal.*		$x + \frac{\delta + \delta(1+\delta)x}{3(1-\delta)}$
3-state *Initial Bias*		$x + \frac{\delta + \delta(1+\delta)x}{3(1-\delta)}$
2-state *History Neglect*		$\frac{\frac{(2+\delta p)x}{(1+\delta p)} + \frac{(1-\delta)x+\delta p}{(1-\delta+\delta p)}}{3(1-\delta)}$
1-state *Non-exploration*		$\frac{x}{1-\delta}$
0-state *Random*		$x + \frac{\delta(4x+1)}{6(1-\delta)}$

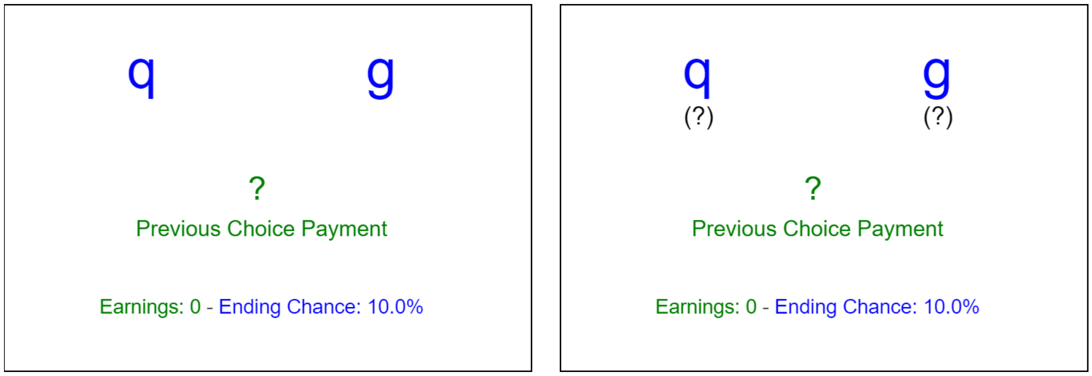
This table presents five procedures. The left had column names and describes each procedure; the center column graphically presents each procedure as an automata; the right hand side gives the expected payoffs from each procedure.

Removing states from procedures–moving between 4-state, 3-state, 2-state, 1-state and 0-state/random–progressively reduces state- complexity costs, but as table 1 shows, it also weakly reduces expected payoffs: 4-state and 3-state are payoff equivalent, with payoffs decreasing as states are shed and minimized at pure randomization. Thus, subjects face a trade-off between complexity costs and monetary earnings.

2.3 Experiment Design

The experiment consists of 20 *tasks*, each of which is an implementation of the bandit problem described in section 2.1. The subject chooses between pulling two arms, 1 and

Figure 2: Screenshots for the two main treatments.



(a) ST (b) NST

2. To make the problem easier to understand, we scale payoffs by a factor of 100. The first arm pulled (note that this could be arm 1 or 2) always pays $x = 65$ points. The second pays 0, 65 or 100 points, each with equal probability. Each period there is a $(1 - \delta) = 0.1$ chance that the period is the last. As we show in Appendix A.2, this choice of (x, δ) guarantees that the optimal procedure is the one described in section 2 (i.e, a payoff-maximizing subject should always explore).

Figure 2 shows a pair of screenshots from the experimental software that illustrate how we implemented the task. In each period, subjects selected an arm of the bandit by typing a letter linked to that arm. The letters on the screen changed randomly each period, but subjects knew that typing the alphabetically earlier letter (of the two on the screen) always selected action a_1 and typing the alphabetically later letter selected a_2 The two letters required for selection were also shown in a random left-right order on the screen each period. We implemented these two features (random letters to select arms and random screen orientation) to reduce the ability of subjects to rely on artificial mnemonic devices otherwise available in any computerized implementation.[3] After pulling an arm

[3]For instance, if the left-right orientation were held constant across periods, the subject could focus her eyes (or rest her finger) on a side of the screen to track actions that have been taken in the past. To remove this prop we randomized which option appeared on the screen. Another mnemonic is to rest one's finger on a key representing the action the subject took in a previous period, aiding memory–changing the letter required for each action in each period prevented this. Together, these force the subjects to

by typing a letter, subjects were shown their latest payoff, their cumulative earnings and were given a new set of letters to choose from to make their next choice.

We ran the experiment using the *block random termination* (BRT) protocol introduced by Frechette and Yuksel (2017). In this protocol, subjects made their choices in blocks of five periods and were only told if and when the final “real” (paying) period of the task had occurred once a block was over. This protocol retains the intertemporal discounting incentives of the model but allows us to gather more periods of incentivized data per task. Once a task was completed, a subject began a new task until the subject had played 20 total tasks.

2.3.1 Treatments: Varying the Costliness of Complexity

The heart of the experimental design is a pair of between-subjects treatments that vary the complexity costs of tracking states without changing the formal properties of the decision problem. In our **State Tracking** treatment (ST), visualized in figure 2a, we implemented the experiment in the natural way: subjects were required to track both their history of actions and their prior payoffs themselves, thus exposing them to two key sources of state-complexity in the problem. In the **No State Tracking** (NST) treatment we drastically removed the need for subjects track states themselves, as figure 2b shows. The display is identical to the one from the ST treatment except for one change: subjects saw question marks below each of the letters, which were replaced with the payoff of each arm once the subject has sampled that arm. Subjects in NST therefore did not have to track past events and choices to implement any of the rules in table 1, as the interface did it for them. This removed the cognitive burdens summarized by state-complexity, making it less costly to implement complex rules, while keeping all other features of the problem constant.

track which arms they have selected in the past and what their payoffs were in each case.

We also included in the design a variation of the ST treatment that is meant to make the problem more *difficult*, but (unlike NST) does not directly target the costliness of tracking states. The **State Tracking + Cognitive Load** (ST+CL) treatment is identical to ST, except that between periods we show the subject a random letter on their screens. At the end of each 5-period block the subject is required to type out these 5 letters as part of their payment. The treatment provides some useful diagnostic information on the mechanism driving results from our main treatment, and we provide motivation for it in section 2.4.6.

2.3.2 Implementation

We ran the experiment using custom Javascript software, deployed with Qualtrics. Participants were recruited using the Prolific platform (prolific.co) in April and May, 2020. We recruited 180 subjects from the United States (60 in each treatment) and subjects were allowed to participate in no more than one session/treatment. Subjects were required to take a comprehension quiz consisting of 7-8 questions that we will use to understand the role of confusion in subjects' decision-making. Subjects earned a $2.50 show-up fee for completing the study ($5.00 for ST+CL) and a bonus based on their *average* earnings across all 20 games. Subjects in ST earned an average of $5.05 ($12.13/hour); subjects in NST earned an average of $7.14 ($17.13/hour)[4,5] Most Sessions lasted in total between 20 and 30 minutes.[6]

[4]Prolific requires researchers to pay subjects at least $6.50/hour. Because the ST+CL treatment was much longer, we were required to pay a higher fixed fee. Our subjects were well-compensated relative to the minimum wage on Prolific.

[5]We paid subjects according to a threshold rule to intensify incentives: if subjects earned fewer than 700 points in the average task, they earned zero and they earned 3 cents for every point greater than 700.

[6]Because the sessions were run online, we could not directly observe subjects. Some subjects took much longer to finish a session. Our data on time supports the notion that they took breaks in between tasks (we did not record time between tasks, but the total time within tasks was typically much less 20 minutes).

Table 2: Exploration rates and efficiencies from each treatment.

Treatment	Exploration Rate	Efficiency	Normalized Efficiency
	All		
No State Tracking	0.749	0.929	0.821
State Tracking	0.526	0.76	0.391
State Tracking+Cognitive Load	0.555	0.751	0.368
	Unconfused		
No State Tracking	0.904	0.979	0.944
State Tracking	0.671	0.848	0.598
State Tracking+Cognitive Load	0.682	0.848	0.599

This table presents exploration rates, efficiency, and normalized efficiency (define as the fraction of the possible earnings improvement relative to random play). The top portion includes the entire sample; the bottom portion includes only subjects who made no more than one mistake on the comprehension quiz during the instructions.

2.4 Results

In this section we report the results from the experiment, focusing on our main ST and NST treatments. In section 2.4.1 we give an overview of behavior at the aggregate level. In section 2.4.2 we describe our methods for identifying empirical markers of procedures, and in section 2.4.3 we report rates of these markers across treatments. In section 2.4.4 we study dominated actions and randomization in some of our subjects subjects and in section 2.4.5 we classify subjects according to the procedures they employ. In section 2.4.6 we evaluate the role of task difficulty and mistakes in driving our findings and in section 2.4.7 we examine the degree to which the opportunity cost of time acts as a driver of our results. In all of our analysis, we focus strictly on behavior from the last half of the session, after subjects have had a chance to experience the environment and formulate a considered response.

2.4.1 Suboptimality

The upper panel of table 2 lists averages of key aggregates for each of our treatments. Efficiency is the fraction of earnings relative to optimal earnings, while Normalized Efficiency is the fraction of the *possible earnings improvement* relative to random play, earned by subjects. On both measures, we find relatively poor performance in our main ST treatment. Subjects leave about 25% of earnings on the table and achieve less than half of the improvement over random play achievable by optimal choice. To help understand this phenomenon, we can look at the Exploration Rate. This is the per-period rate at which subjects switch off of their initially chosen arm, which an optimal decision maker should always do immediately (switching at a rate of 1). Subjects in ST, however, only switch off the initial arm at half of the optimal rate.

Result 1. *Subjects in ST, on average, significantly under-explore and make overall inefficient choices.*

Because our task is extremely simple (indeed, it is among the simplest possible infinite horizon bandit problems), it is tempting to attribute patterns like these to preferences (e.g., risk attitudes) or perhaps to mere confusion about the reward structure of the problem. Indeed, the low rate of exploration would seem to support this idea: by avoiding exploration, subjects avoid the risk of a low second-arm payoff; if confused, subjects also avoid ambiguity about the payoff consequences of deviating from their initial action by declining to explore.

The data, however, allow us to reject such explanations. The NST treatment is identical to ST in risk characteristics and features virtually identical instructions to subjects. Nonetheless, these errors largely disappear in this treatment: both efficiency and explorations increase substantially relative to NST. The lower panel of table 2 goes further, removing confusion as a plausible explanation from either treatment by including only the subjects who made no more than one mistake in the comprehension quiz given during the instructions (these subjects account for more than 50% of our pool). We will refer to these as our "Unconfused" subjects. Performance measures improve for this subsample in both treatments, but the gap in performance between the two remains nearly as large: each of these measures are statistically different across treatments (for both the full sample and the subsample of "Unconfused" subjects) at the < 0.001 level by Wilcoxon tests.[7]

Result 2. *Subjects in NST, on average, explore at near optimal rates and achieve high rates of efficiency.*

This result suggests that uncertainty aversion or confusion are not the primary drivers of suboptimal behavior in this problem. Instead, it suggests, as hypothesized, that subjects may be deliberately substituting to simpler procedures (i.e., those with fewer states) in order to avoid the subjective costs of tracking states. To understand the degree to which this is true, we must examine the extent to which subjects use procedures like those discussed in section 2, and to what extent the choice among these various procedures differs across treatments.

2.4.2 Individual Level Analysis

In order to study procedural choice, we need individual-level estimates of the per-period probability of switching from the current arm (i.e., the arm pulled immediately prior) to the alternative arm in a specific set of histories. Procedures differ precisely in these probabilities. There are four main histories (sequences of events and actions) in the data that, together, produce distinctive fingerprints of the procedures described in section 2. In describing them, we will refer to the initially selected arm as the "first arm" and the other as the "second arm."

[7] In this and all other statistical tests we use subject-wise aggregates, meaning we use only one observation per subject in order to avoid attributing more independent variation to the data than is warranted.

- **1A-I**: where "1A" specifies that the subject most recently selected the first arm and "I " (standing for "initial"), specifies that the subject has not yet sampled the second arm.

- **2A-0**: the subject most recently selected the second arm and received a 0 payment from it.

- **2A-100**: the subject most recently selected the second arm and received a 100 payment from it.

- **1A-0**: the subject most recently selected the first arm but has sampled the second arm in the past, observing a payoff of 0 on that second arm.[8]

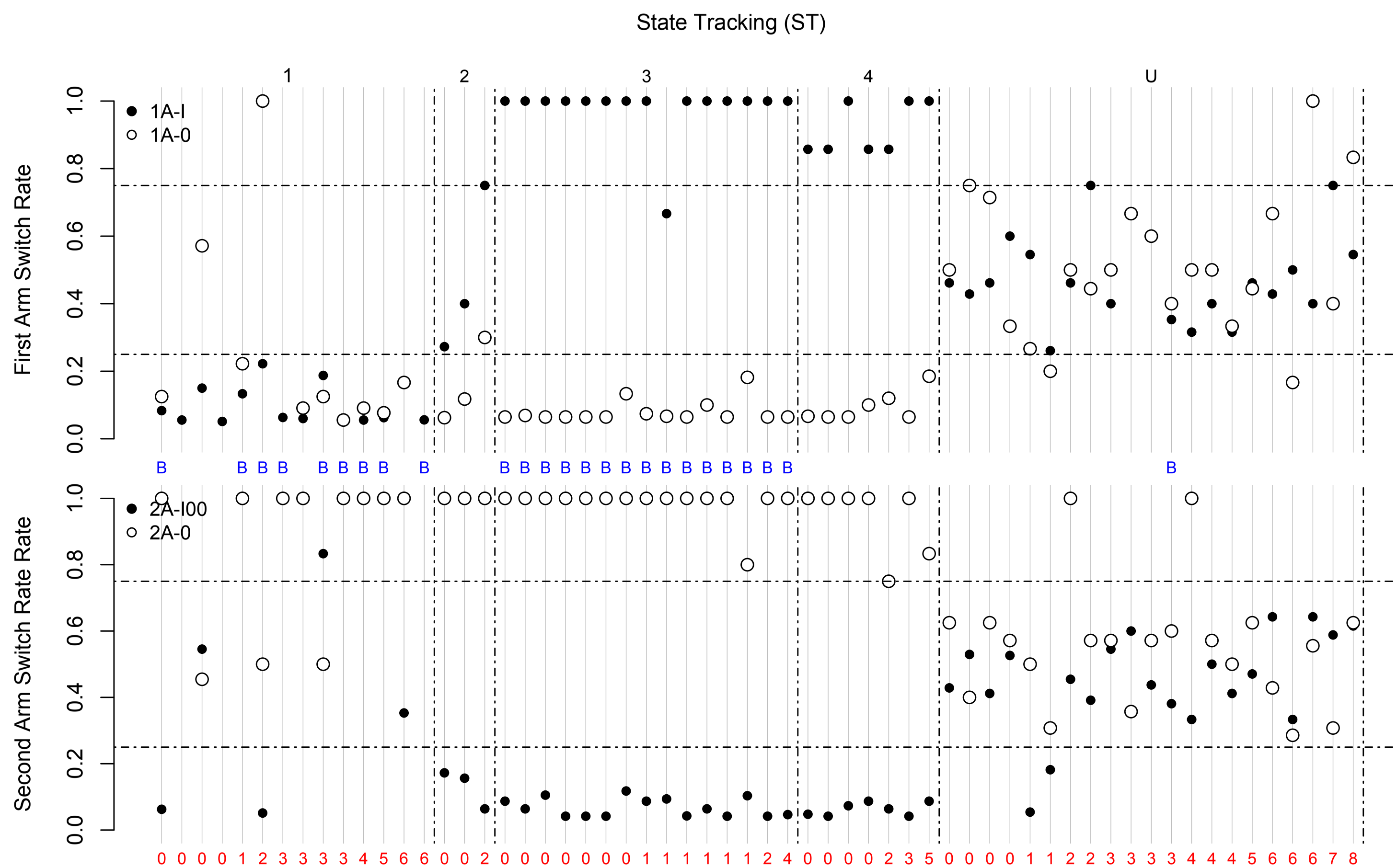
For each of these cases, we can estimate the per-period probability of switching arms p as $\frac{1}{\overline{N}}$, where $\overline{N}$ is the average number of periods the subject spent on the arm for all instances in which we observed her under that history. Because the tasks end with random probability, however, it is important that we are careful to minimize bias due to truncation. For our estimates, we include only cases in the sample in which subjects entered the history (e.g., 1A-I or 2A-0) with at least 10 periods remaining in the period, meaning we can always observe probabilities as low as 0.1 (and sometimes lower).[9]

Figures 3 and 4 plot estimates for each of these histories for every subject in the dataset (each horizontal position is a separate subject) for ST (figure 3) and NST (f:master1.2). The top plot in each figure presents the first-arm switch rates, including 1A-I (black dot) and 1A-0 (hollow dots); the bottom plot presents the second-arm switch rates, including 2A-100 (black dots) and 2A-0 (hollow dots) for the same subjects.

[8]There are three other cases that are not useful for our purposes. 1A-100 (the subject most recently selected the first arm but previously observed a 100 payment on the second arm) should never occur, even for lower state procedures. Indeed, it occurs at a low rate in the data, implying we have small and inconsistent samples for estimation; we exclude this case from the analysis. 1A-65 and 2A-65 are cases that only occur if both arms pay the same amount (both pay $x = 65$) and in this case there are no constraints on behavior under any procedure, making these cases useless for classification.

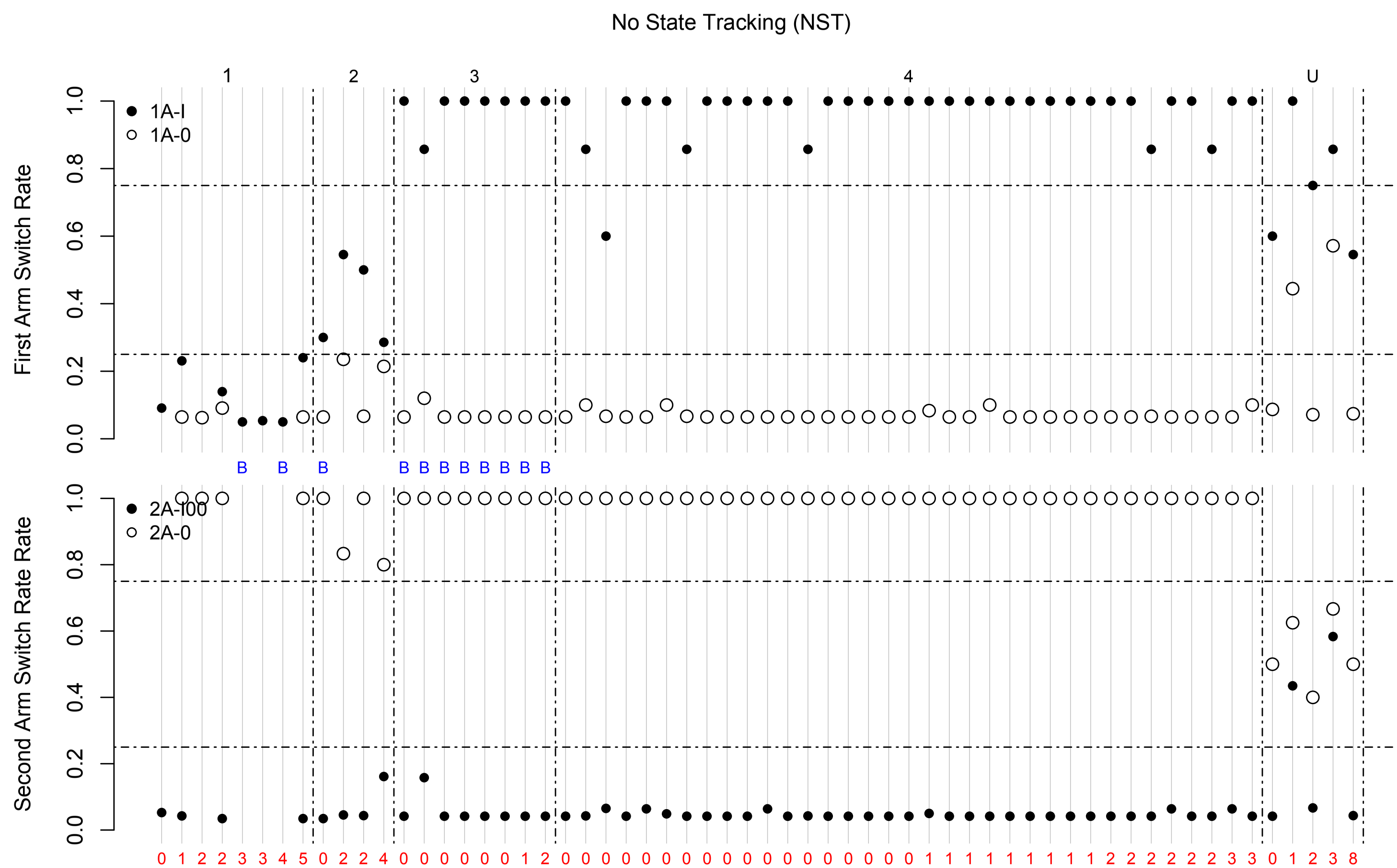
[9]Recall that the subject does not know how many periods remain because of the stochasticity of the ending rule, meaning this decision introduces no bias due to, e.g., self-selection.

Figure 3: Switching Probabilities for Each subject in ST



The upper plot shows switching behavior after play of the initial arm; the lower plot after play of the second arm. Numbers at the bottom indicate how many errors a subject made in a comprehension quiz prior to the experiment. The letter 'B' between plots designates subjects who were initial-biased.

Figure 4: Switching Probabilities for Each Subject in NST



The upper plot shows switching behavior after play of the initial arm; the lower plot after play of the second arm. Numbers at the bottom indicate how many errors a subject made in a comprehension quiz prior to the experiment. The letter 'B' between plots designates subjects who were initial-biased.

A "B" between plots designates a subject who displays *initial bias*, choosing the same initial arm repeatedly across tasks.[10] Along the bottom of each bottom plot we include the number of mistakes the subject made in the comprehension quiz (e.g., 0 means the subject never made a mistake, 7 means the subject submitted 7 mistaken answers).[11] The horizontal ordering of subjects (and the dashed vertical lines grouping them) is based on classifications discussed in the next section.

Most of the predictions in which we are interested (i.e., the procedures described in section 2) require subjects to make switches with very high or very low per-period likelihoods. Theoretically, this usually corresponds to switching with probabilities $p = 1$ or $p = 0$, respectively. We must be more forgiving for two reasons. First, as just discussed, our probability estimates are conducted (unavoidably) on truncated samples, meaning we will have some mechanical bias away from the extremes (for instance it is not possible given our design to estimate probabilities lower than 0.07). Second, any implementation mistakes or behavioral noise will systematically push estimates into the interior.

For classification purposes, we will adopt the convention that an arm switch occurs at a "high rate" if it occurs with an estimated $p \geq 0.75$ per period, a "low rate" if $p \leq 0.25$, and that two rates are "very different" if they are at least as large as the difference between these two boundary rates. These boundaries are visualized as horizontal lines in figures 3 and 4 Results do not change qualitatively if we vary these thresholds. Regardless, we present the data in as transparent a way as possible in figures 3 and 4, such that readers can evaluate the data visually on their own to assess the reasonableness and robustness of this convention.

[10]We say a subject displays initial bias if we can reject the hypothesis at the five percent level that she chooses each arm with equal frequency using a binomial test. In our data this requires the subject to choose the same action in at least 9 out of 10 cases.

[11]Note that subjects can *repeatedly* submit incorrect answers, as the software required a correct answer before subjects could move to the next question. Also note that upon answering incorrectly, the software provided subjects with an explanation that pointed to the correct answer; even so, several subjects produced more than 7 mistakes.

2.4.3 Procedural Markers

As we discuss in section 2.2.3, each of the state-economizing procedures (3-, 2- and 1-state) involves the use of a specific pattern of behavior that allows the decision maker to avoid implementing additional states. We will refer to these as "procedural markers" and we repeat them below along with empirical operationalizations:

- **Non-Exploration**: Subjects implementing a 1-state procedure will explore (switch under history 1A-I, denoted by the black dots in the upper plots of figures 3 and 4) at a low rate, while subjects employing 3- or 4-state procedures will explore at a very high rate.

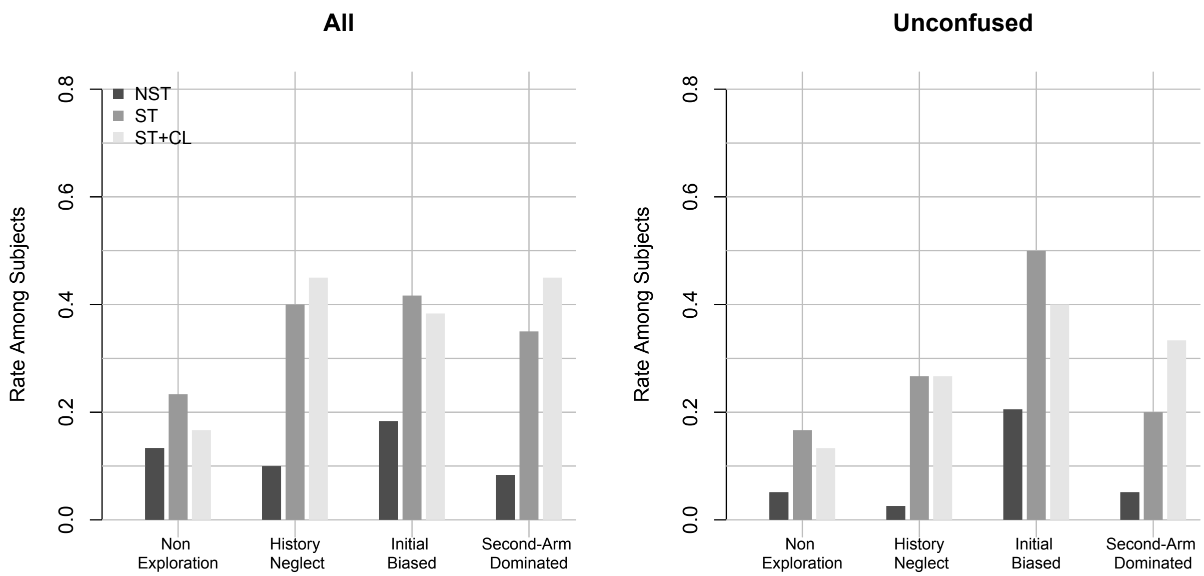
- **History Neglect**: Subjects implementing a 3- or 4-state procedure will switch from the first arm at very different rates depending on whether they have explored in the past. Pre-exploration, a follower of a 3- or 4-state procedure will switch at a very high rate (1A-I, the black dots in the upper upper plots of figures 3 and 4 will be high), while post-exploration (after having observed a 0 payment in the second arm), she will switch at a very low rate (1A-0, the white dots in the upper upper plots of figures 3 and 4, will be very low). By contrast, a subject implementing a 2-state procedure should *not* behave differently across these two histories.

- **Initial Bias**: Subjects implementing a 3-state rule will systematically choose one initial arm consistently across games (shown as B symbols between the plots in figures 3 and 4), while subjects implementing a 4-state rule will not. Initial Bias does not impact the state-complexity of lower state procedures.

In figure 5, we plot the rate at which subjects display each of these markers.[12] The right hand plot includes only subjects who are very unlikely to be confused in the

[12] Again, for this classification we adopt the convention that a subject switches at a "high rate" if $p \geq 0.75$, a "low rate" if $p \leq 0.25$ and that two rates are very different if they at least as large as the distance between these two boundaries.

Figure 5: Incidence Rates of Procedural Markers



This figure presents rates at which subjects: switch off the first arm at a low rate (Non-Exploration); switch off the first arm at the same rate pre- and post-exploration (History Neglect); display bias for one arm for initial choices (Initial Biased); stay an an arm that pays 0 or switch off of an an arm that pays 100 (Second-Arm Dominated). The left panel presents data for all subjects; the right panel for the subset who made no more than one mistake on the comprehension quiz.

experiment–subjects who made no more than 1 mistake on the comprehension quiz in the experiment. For most of our analysis we will focus on this subset, though we will make comparisons across the two throughout.

Beginning with non-exploration: nearly 25% of subjects in ST explore at a low rate, though about a third of this is plausibly driven by confusion: for Unconfused subjects, this drops to 17%. Thus, a substantial minority of subjects with a demonstrated understanding of the game (i.e., those who scored very well on the comprehension quiz) explore at a very low rate, consistent with selection of a highly simplified one-state strategy. By contrast, NST subjects are less than a third as likely (5% likely) to avoid exploration, suggesting that avoidance of state-complexity costs is likely the major motivation for this behavior.

Result 3. *Nearly 20% of subjects in ST use a low-exploration procedure, while only 5% of subjects in NST do the same.*

Similarly 40% of ST subjects show evidence of simplifying by under-conditioning their behavior on history (History Neglect) in the raw sample but this rate shrinks to 25% among subjects with demonstrated understanding of the experiment. By contrast, virtually none (under 3%) of the subjects in NST show evidence of History Neglect. Again, this suggests that almost all of this behavior is caused by state complexity costs.

Result 4. *25% of subjects in ST show evidence of using a history-neglecting procedure, while virtually no subjects do in NST.*

Finally, and perhaps most tellingly, 50% of Unconfused ST subjects use Initial Biased procedures, by systematically selecting the same arm in each task. Doing this enables subjects to use simpler 3-state strategies instead of 4-state strategies without sacrificing payoffs. This choice is clearly driven in large measure due to a response to complexity costs: subjects are less than half as likely to exhibit Initial Bias in the NST treatment. This is perhaps surprising, given that employing an Initial Bias procedure would seem to be quite easy and comes at little cost. The fact that subjects are much less likely to use initially biased strategies in NST, however, suggests that this treatment is (as designed) extremely effective at removing the costs of tracking states. It also suggests that subjects are relatively sophisticated about choosing procedures that economize on complexity costs when these costs become significant.

Result 5. *50% of subjects in ST use Initial Biased procedures, while only 20% of subjects do in NST.*

Overall, Unconfused subjects are more than 3 times as likely (87% vs. 26%) to exhibit one of these procedural markers in ST than in NST ($p < 0.001$ by a Pearson's proportions test). Moreover, the effect of complexity costs (ST vs. NST) is far more important than confusion in driving the emergence of these markers: for the sample overall there is still an extremely wide (88% vs. 37%) difference in incidence of these markers across treatments.

2.4.4 Second-Arm Dominated Behavior

All of the non-random procedures that contain more than one state described in section 2 include an additional prescription in common: subjects should immediately switch from the second arm if it pays 0 and never switch from the second arm if it pays 100. We must include this as an additional marker required of a subject before concluding she is implementing any of the multi-state procedures shown in table 1.

- **Second-Arm Rationality**: In none of the multi-state procedures in table 1 will subjects fail to switch at a high rate from a second arm after a payment of 0 (2A-0, the white dots in the low plot of figures 3 and 4, will be very high) or fail to switch at a low rate from a second arm payment of 100 (2A-100, the black dots in the low plot of figures 3 and 4, should be very low). If a failure of either of these conditions occurs one average we say the subject displays *second-arm dominated* behavior.

In each panel of figures 3 and 4 we group subjects who make at least one of the two possible "second-arm dominated" actions in the category to the far right of the plot (the category labeled "U"). figure 5 shows that 35% of subjects show a systematic tendency to take at least one of the two "second-arm dominated" actions on average, but this number drops almost in half among Unconfused subjects. Only 5% of subjects in NST take such dominated actions, suggesting, again, that this pattern of behavior is largely driven by the burdens associated with states.

Interestingly, ST subjects categorized as second-arm dominated (again visualized on the right hand side of figure 3) are quite consistent in their own way. Although they are grouped here based only on having taken at least *one* second-arm dominated action, most of them actually take *both* dominated actions on average (black and white dots in the lower plot of each panel are usually both in the interval $[0.25, 0.75]$). Moreover, taking second-arm dominated actions *predicts* suboptimal, interior first-arm actions: most of subjects who take second-arm dominated actions also switch from the first arm with strictly interior likelihood regardless of history as well. This strongly consistent randomization suggests that many or most these subjects may in fact be deliberately employing a random procedure: a 0-state procedure that completely removes state-complexity costs. We return to the interpretation of such behavior in section 2.4.6.

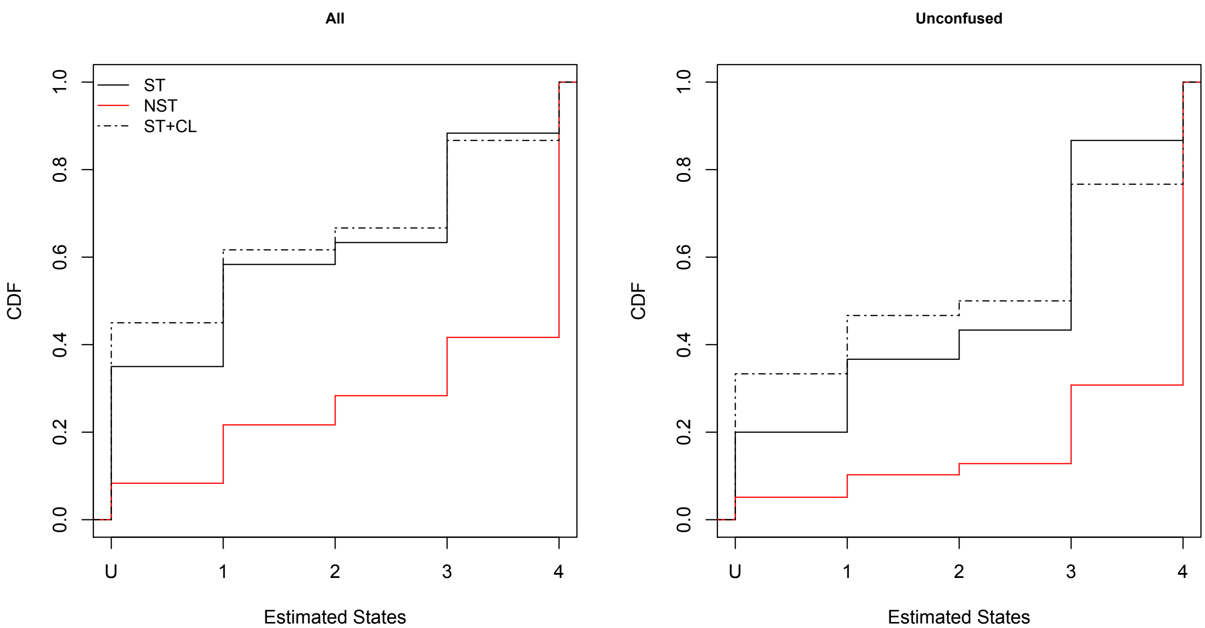
Result 6. *20% of unconfused subjects in ST systematically take dominated actions at the second-arm. Most of these subjects* consistently *switch arms with strictly interior probability in all four of our diagnostic histories, suggesting that this behavior may be due to deliberate randomization. This behavior almost entirely disappears in the NST treatment.*

2.4.5 Complexity and Procedural Choice

By studying the within-subject relationship between these markers (Non-Exploration, History Neglect, Initial Bias and Second-Arm Domination) we can classify subjects according to their use of the four economizing procedures listed in table 1. These produce the following distinctive fingerprints in the data:

- **1-State**: *Non-Exploration*
- **2-State**: *History Neglect* but not *Second-Arm Domination* or *Non-Exploration.*

Figure 6: Empirical Cumulative Density Functions of Procedure-Classifications



This figure plots the empirical cumulative density functions, by treatment, of the number of states subjects employ according to the procedures in table 1 The left panel shows the entire sample; the right shows the subset that made no more than one mistake on the comprehension quiz.

- **3-State**: *Initial Bias* but not *History Neglect*, *Non-Exploration* or *Second-Arm Domination.*
- **4-State**: No *Initial Bias*, *History Neglect*, *Non-Exploration* or *Second-Arm Domination.*

Thus, to classify subjects as users of 2-, 3- or 4-state procedures, we require both strict evidence of rationality in the second arm (no systematic evidence of dominated behavior) and evidence of the simplifying marker characteristic of the procedure. For subjects who rarely explore, we rarely observe behavior on the second arm and expect any resulting estimates to be very noisy. Thus, we tolerate evidence of Second-Arm Domination in classifying subjects as 1-State (this only effects the classification of four subjects total across ST and NST). Subjects who do not fulfill this strict criterion are grouped as:

- **Unclassified**: Subjects who show evidence of marker *Second-Arm Domination* do not strictly fit into the taxonomy in table 1 but may be consistent with a 0-state randomization procedure.

Figures 3 and 4 organize subjects horizontally by this classification from left to right. At the top of the upper plots we list state classifications: 1, 2, 3, 4 and U for Unclassified. 1-state subjects to the far left explore at a very low rate (black dots are below the horizontal line in the upper plot) but otherwise show little evidence of deviations from optimality (there are few dots between the horizontal lines in either the upper or lower plot of either panel). Only a handful of subjects are classified as 2-state (subjects with relatively close-together black and white dots in the upper plot of each panel). 3- and 4-state subjects (distinguished from one another by the presence or absence of a 'B' designating initial Bias between the plots) are particularly non-noisy and virtually always show very high rates of exploration (black dots in the upper plot) and very low rates of switching back to an arm that has paid 0 in the past (white dots in the upper plot).

As the figure makes clear, most subjects who display History Neglect in our data are not 2-state subjects of the sort discussed in section 2, but rather Unclassified subjects, shown on the far right of the plots. As we highlighted in the previous subsection, these subjects both have relatively similar switch rates from the first arm across histories (white dots are not systematically lower than black dots in the upper plot) and take frequent dominated actions on the second arm (white and black dots are largely between the horizontal lines on the lower plot). The average such subject behaves as though she is making random decisions. Many (though not all) of these subjects show possible evidence of confusion, with relatively high error rates on their comprehension quiz (shown as numbers on the underside of the bottom plot). Indeed, in the ST treatment, the median number of errors on the comprehension quiz for Unclassified subjects is three times larger than for classified subjects.

Figure 6 shows empirical CDFs of procedure-types for each treatment and, as in figure 5, includes separate panels for subjects overall and for Unconfused subjects (who made no more than one mistake on their comprehension quizzes). In ST, use of non-random procedures is dominant, rising from 66% to 80% as subject confusion falls away (as we remove subjects who made mistakes in their comprehension quiz). Among Unconfused subjects, 3-state procedures are dominant, accounting for 43% of subjects; there are also smaller but significant modes using 1- and 4-state procedures. Notably, 2-state procedures are quite rare; as we pointed out above, most subjects exhibiting History Neglect are subjects who also make high rates of second-arm dominated actions.

Result 7. *The vast majority (80%) of unconfused ST subjects show evidence of using systematic, non-random procedures. The modal such subject (43%) uses a 3-state procedure, with fewer using 1- and 4- state procedures. 2-state procedures are rare.*

Our key finding is that although the removal of complexity costs in NST does reduce apparently random behavior (the share of Unclassified subjects drops from 20% to 5%), most of its effect is in altering the pattern of subjects' choices among non-random procedures. Subjects use 1-, 2- and 3- state procedures each more than twice as often in ST than in NST. Moreover, subjects are about 5 times more likely to use maximally complex 4-state rules in NST than in ST. The differences between the two distributions are statistically different (for both the full sample and Unconfused subjects) at the < 0.001 level (using both Wilcoxon and Kolmogorov-Smirnov tests). Thus, when complexity costs rise, subjects substitute systematically towards lower-state rules.

Result 8. *Virtually all (95%) of Unconfused NST subjects show evidence of using systematic procedures. Most of these (70%) use maximally complex 4-state strategies.*

Again, the clearest indicator that subjects are deliberately (and in sophisticated fashion) choosing procedures to economize on complexity costs is in the very different ways subjects optimize in ST relative to NST. When the costs of tracking states are low (NST) most (80%) of Unconfused subjects use maximally complex 4-state rules. When the costs rise in ST, this reverses, with most (75%) of subjects electing to use systematically Initial Biased strategies, removing one state from their optimizing procedure:

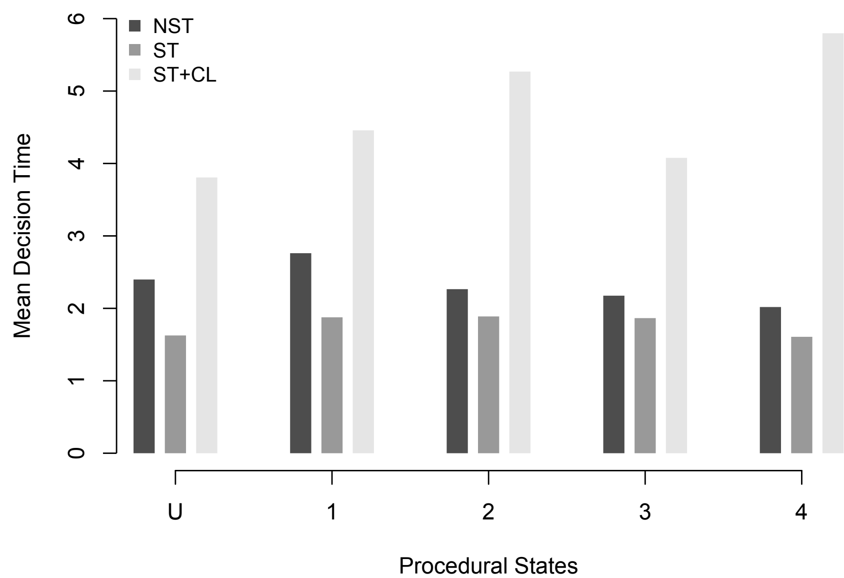
Result 9. *9 Subjects optimize in a systematically different way when complexity costs are increased. In ST 75% of optimizing subjects use simpler 3-state rules, while in NST 80% use more complex 4-state rules.*

2.4.6 Difficulty and Mistakes

Our findings suggest that people choose low-state procedures in ST to avoid subjective costs associated with tracking states. A natural alternative explanation is that subjects simply make more mistakes in ST than NST, because the former is more difficult. Before evaluating this idea in the data, it is important to emphasize that most of the behaviors generating low state procedures are unlikely, ex ante, to arise due to slips of the hand or memory lapses. Two of our three procedural markers–non-exploration and initial bias–require the subjects to make highly consistent choices (in each case, consistently choosing an arm) that are unlikely to arise by accident. It is possible, however, that history neglect and second-arm dominated actions (which, recall, characterize about 20% of unconfused subjects) are generated by mistake.

One way to assess the plausibility of this mechanism is to examine whether subjects actually find the ST treatment more difficult than the NST treatment. A common metric of task difficulty, used widely in economics and psychology, is the time subjects require to make decisions in the task: in a wide range of settings subjects tend to spend more

Figure 7: Mean Decision Time Per-Period



This figure presents mean number of seconds for subjects to choose an arm, by treatment and estimated procedure classification. The sample is limited to subjects who made no more than one mistake on the comprehension quiz.

time on tasks they find more difficult. Figure 7 shows mean decision times for subjects using each type of procedure, for each treatment. The results show that subjects don't spend any more time on decisions in ST than NST (in fact NST subjects spend slightly more time, probably due to the larger amount of information presented on their screens). This casts some doubt on the hypothesis that subjects should be more likely to make accidental errors in ST purely due to increased task difficulty.

This, too, highlights an important point that has been documented in previous experiments: state complexity and task difficulty are not the same thing. For example, Oprea (2020) finds that subjects make mistakes only about 5% of the time when implementing procedures with similar automata characteristics to our optimal 4-state rule.[13] Nonethe-

[13]The closest automata to our optimal 4-state rule is the rule 4S-1Ta from Oprea (2020) which contains 4 states including an absorbing state.

less, the same subjects are willing to pay a significant share of their earnings to avoid having to implement these same rules. Thus, people find tracking states unpleasant even when it is not actually difficult to do so correctly.

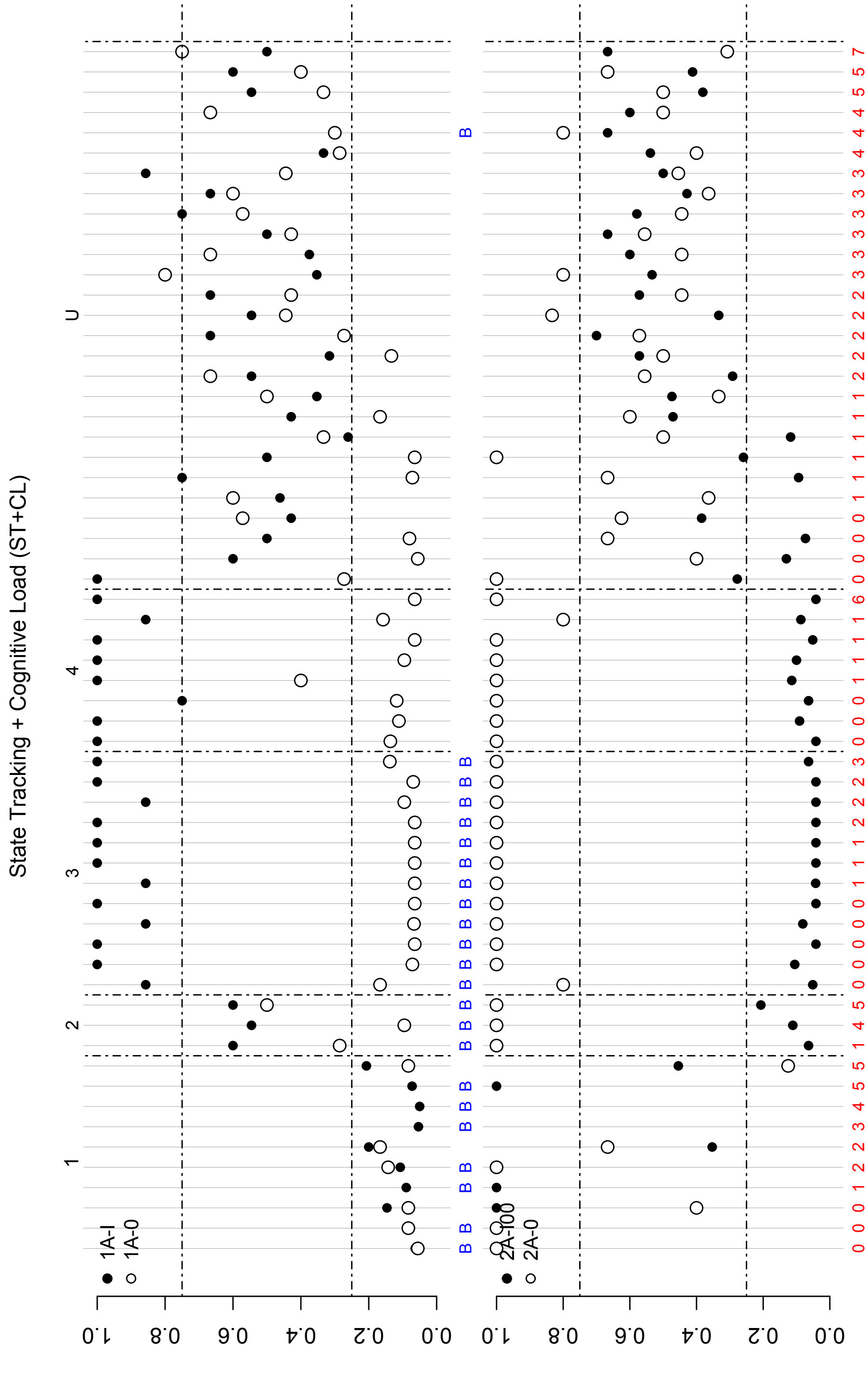
A related way to assess the hypothesis that difficulty might generate procedural simplifications due to an uptick in accidental mistakes is to run a treatment in which we actually increase task difficulty in a measurable way. The State Tracking + Cognitive Load (ST+CL) treatment is a variation on the ST treatment in which we require subjects to hold a string of letters in their working memories while engaged in the main bandit task. While the NST and ST treatments vary the task in a relatively targeted way (aimed specifically at influencing the cost of tracking states), the ST+CL treatment simply adds background difficulty to the task by taxing the subjects' attention and mental resources.

Figure 7 shows that ST+CL is, in fact, likely much more difficult than ST: subjects take nearly *twice* as long to make decisions in ST+CL than they do in ST. Additionally, in figure 8 we plot switching probabilities for subjects in ST+CL (using the same format as figures 3 and 4). Our main finding is that, despite being more difficult, ST+CL has little meaningful effect on behavior relative to ST. Efficiency drops slightly (see table 2), mistake rates increase, and there seem to be slightly more subjects Unclassified, but these effects are not statistically significant. Indeed, none of the incidences of markers from figure 5 differ significantly from ST (by Pearson's Chi-square tests) at conventional levels and the CDFs of state classifications in figure 6 are extremely close together.

Result 10. *Subjects in ST+CL do not behave significantly differently from subjects in ST, suggesting that task difficulty per se does not influence procedure use.*

We interpret this as further evidence in favor of the behavioral hypothesis advanced by the automata literature. It is only when we alter a specific face of task, linked to the

Figure 8: Switching Probabilities for Each Subject in ST+CL



The upper plot shows switching behavior after play of the initial arm; the lower plot after play of the second arm. Numbers at the bottom indicate how many errors a subject made in a comprehension quiz prior to the experiment. The letter 'B' between plots designates subjects who were initial-biased.

underlying automata characteristics, that we observe a systematic change in procedures selected by subjects.

If mistakes are not likely to be a major driver of behavior in our experiment, what accounts to the roughly 20% of subjects who regularly make second-arm dominated choices (classified as 'U' in figures 3-8)? Principally (as we note above and as is evident in the figures), these subjects tend to switch arms with interior probabilities in *all histories*, strongly suggesting that many may be deliberately selecting a maximally simple 0-state randomization procedure. Indeed, some methods for deliberate randomization are easy to see in the data: one easy way to randomize in our implementation is to repeatedly type whichever letter shows up on the left hand side of the screen. Doing this consistently will select each arm with equal likelihood and there is some evidence for this among unclassified ST subjects: one third of unclassified subjects–including subjects who show high comprehension–choose the left or arm at a high (≥ 0.75) rate but *no* classified subjects do this, suggesting that at least some of the random-looking unclassified behavior in the data is deliberate.

2.4.7 Decision Time

Finally, figure 7 and the ST+CL treatments are also diagnostically valuable for interpreting the motivations behind subjects' shift to simpler procedures in ST (and ST+CL) relative to NST. We have interpreted this as a consequence of sheer subjective distaste for implementing state-complex rules. Another possibility, however, is that subjects pursue simpler, lower state rules in order to complete the task more quickly (e.g., driven by the opportunity cost of time).

Figure 7 suggests time is unlikely to be subjects' motivation. First, subjects don't in fact move substantially more quickly (the differences are fractions of seconds) when they use lower state rules, regardless of the procedure they use in the ST treatment.

Of course, assignment to procedures is not random, so this evidence is only suggestive (subjects that would have taken a long time in a higher state procedure might be precisely those who choose lower state procedures). More directly, evidence comes from ST-CL where *all* procedures take significantly longer to complete. If opportunity cost of time were a major motivator of behavior in the experiment, we would expect these costs to be substantially larger in ST+CL than ST, generating stronger shifts to lower procedures. Again, however, we see little evidence of this in the ST+CL treatment. We conclude that subjects likely select lower state procedures in ST/ST+CL than in NST simply because they find it costly (e.g., unpleasant or distasteful) to track states.

2.5 Discussion

How much of human behavior can be explained by the simple idea that people's dislike doing complex things? This idea has a long history in economics, but it has received surprisingly little direct empirical attention and, perhaps as a consequence, it is rarely invoked to explain empirical puzzles in economics. We provide one of the first empirical tests of the idea, framing our investigation using formal automata models of procedural complexity that make predictions which are amenable to empirical study. Our findings provide strong support for this theory: even in our very basic decision problem, aversion to procedural complexity has a first order effect on behavior. Distaste for procedural complexity may therefore be important for interpreting and predicting behavior in a much wider domain of problems.

We study bandit problems that are simple enough that it is possible to identify the procedures subjects use to guide their decisions. By varying the costs of complexity in a way directly informed by automata theory, we can therefore observe how the choice of procedures responds to complexity costs. Moreover, by choosing a very targeted intervention–removing the requirement to track states without changing the timing, uncertainty or payoffs of the problem–we are able to crisply identify the mechanism by which our intervention alters behavior. We find that subjects use maximally complex 4-state procedures when complexity costs are low (i.e., when no cognitive effort is required to track states), but economize on complexity by substituting systematically to simpler, lower-state procedures when complexity costs are higher. Our design rules out confusion, uncertainty aversion, task-difficult, and the opportunity cost of times as likely mechanisms. The results therefore strongly support the central hypothesis offered by automata models of complexity: that decision makers systematically choose to implement simple procedures *because* they dislike implementing complex ones.

In a companion paper, Oprea (2020) asks complementary questions about procedural complexity using very different methods. In that experiment, subjects are *exogenously assigned* artificial rules (of the form 'type x, until you see y, then switch to z') with various automata characteristics (state counts, transition counts etc.) and are paid to implement these rule on their keyboard. Afterwards, the experiment elicits subjects' *willingness to pay* to avoid being assigned each of these rules to implement again in the future. This direct measurement produces evidence that subjects suffer subjective costs from implementing complex rules, and that these costs rise as the number of states in the rule rises.[14]

Our contribution relative to this work is to study the critical linkage between these costs and actual choice, by directly testing the hypothesis that these complexity costs materially influence behavior in a real decision making context. Unlike Oprea (2020), subjects in our experiment not only must choose which procedures to use, they also must formulate these procedures themselves. To study complexity in this very different setting, we introduce a new and different method for assessing the influence of automata characteristics (in this case, states) on complexity costs. Rather than eliciting these costs,

[14]Oprea (2020) also finds that other features of rules like transitions and the presence or absence of an absorbing state also influence subjects' aversion to rules.

we diagnostically alter the decision setting in a way that removes the costs altogether, allowing us to study the causal impact of these costs on procedural choice. Thus, while Oprea (2020) shows that states are costly, our experiment shows that subjects are sensitive enough to and sophisticated enough about these costs to formulate procedures that effectively economize on them.

These results are important because they suggest that there is an “algorithmic layer” lying between primitives like preferences/beliefs and actions: people have preferences directly over the procedures guiding their choices and, in particular, dislike rules that are complex. That is, people have direct preferences not only over the consequences of their choices, but also the process by which they arrive at these choices. Because of this, human behavior is (at least in part) shaped by features of decision problems that are not reducible to the standard suite of primitives usually studied in economic theory.

Importantly, these preferences have structure and are therefore amenable to modeling, prediction and integration into economic theory. For instance, our work shows that people shape their behavior to avoid rules with many states. Oprea (2020) provides evidence that people also dislike rules with many transitions or rules that require perpetual transition between states. Mapping the structure of these preferences empirically (using methods like those in Oprea (2020)) and examining their organizing power over real choice in a broader set of environments (using methods like ours) is therefore a potentially important empirical enterprise.

Because algorithmic structure underlies decision making across social and economic life (in all but the simplest settings), the scope of application for this approach seems great. We highlight two settings within economics in which procedural complexity may be a particularly valuable conceptual and empirical tool. First among these is the study of repeated interaction, where multiplicity of equilibria hampers prediction and estimation in theoretical and applied work. The original aim of automata models was to make repeated game theory more predictive by limiting the set of strategies available to agents in a principled way. Using methods like ours to understand how complexity interacts with other selective forces (e.g., avoidance of strategic risk) may eventually allow for the development of empirically informed predictive theoretical tools, rooted in agents' motivation to economize on complexity. Doing this will require methodological advances that can empirically tease out and separately identify the effects of procedural complexity from alternative forces, such as computational limitations or the distaste for strategic risk.

No less promising an application is behavioral economics, where identification of unified explanations for the large and growing list of departures from neoclassical benchmarks seems particularly valuable. How much of behavioral economics can be parsimoniously understood as avoidance of the complex procedures often required to optimize? Facially, the possibility seems promising: many (if not most) departures from neoclassical benchmarks seem to involve simpler-than-optimal behavior. Moreover, automata models have been used to theoretically explain a range of behavioral phenomena including satisficing, primacy and recency effects, choice overload and status quo bias (Salant, 2011), stochastic choice (Kalai and Solan, 2003), non-Bayesian inference (Chauvin, 2020), biases in information processing (Wilson, 2014), and failures of backwards induction (Neyman, 1985). Use of methods like ours to examine the degree to which patterns in behavioral economics arise due to procedural complexity seems like a particularly promising enterprise for future work.

Chapter 3

The Economics of the Montana Liquor License System

James Banovetz and Randal Rucker

3.1 Introduction

Government regulation of markets takes many forms, and the regulations often have important impacts on producers and consumers. One common type of regulatory system employs policy tools designed to reduce output levels, which results in increased prices and (possibly) increased profits of producers. In programs that restrict output or subsidize production, expected future program-created profits (or rents) are often capitalized into the value of one or more tradeable assets. For example, in agricultural programs that indirectly limit output by restricting the use of an essential input (commonly land), the resulting increased profits are reflected in farmland rental rates and are capitalized into farmland sale prices (e.g., Kirwan, 2009; Goodwin and Ortalo-Magné, 1992; Barnard et al., 1997; Weersink et al., 1999). In farm programs that directly limit output through the use of transferable production or marketing quotas, annual program rents are reflected in lease rates or are capitalized into the sale price of the quota (e.g., Rucker et al., 1995; Rucker and Thurman, 1990; Brown et al., 2007). In other contexts, the program-created profits are capitalized into the prices of such assets as individual tradable fishing quotas or limited vessel entry licenses (e.g., Newell et al., 2005, 2007; Deacon et al., 2013; Grainger and Costello, 2014), tradable pollution permits (e.g., Hahn, 1984; Joskow et al., 1998; Stavins, 1998; Carlson et al., 2000; Tietenberg, 2006), and transferable taxicab medallions (O'Donnell, 2017), as well as into the sales prices of the liquor licenses that are the focus of this analysis.

In the Montana liquor license regulatory system that we study, producer participation in the on-premises retail liquor market requires possession of one of several types of transferable licenses. Quota areas for the licenses are clearly defined and the licenses can be traded within their designated quota areas, but they can only be traded across quota boundaries under very limited and specific circumstances. The system results in differences in license values across quota areas, and in some instances these differences are substantial. For example, licenses have sold for as much as $1 million in Bozeman, while in Butte, a city about 90 miles away with a roughly comparable population, licenses sell for considerably less than $100,000. Within each quota area, the number of licenses can increase over time if population grows, but within any given year the number of licenses is fixed in supply. Proponents of such regulatory programs maintain that they serve the important function of curbing the oversupply of externality-producing activities—in our case, alcohol consumption. These programs also, however, explicitly support rents for incumbent suppliers.[1]

[1]The history of limited entry programs in, for example, commercial fisheries supports this view (e.g., Grainger and Parker, 2013). At least as early as Crutchfield (1961), economists were espousing limits on entry to address the over-capitalization problem of "too many boats chasing too few fish." The first limited entry programs were instituted in Australia, British Columbia, Alaska, and Washington State in the late 1960s and in the early 1970s (e.g., Wilen, 1988; Townsend, 1990). Limited entry programs are now prevalent in U.S., Canadian, and European waters. Their explicit intent is to cap the number of vessels permitted to compete for a fishery-wide quota as a way to keep fishery profits high for incumbents.

Below, we analyze the determinants of liquor license values, focusing on the Montana system. Economic analyses of liquor license programs can be found in microeconomics textbooks (e.g., Nicholson and Snyder, 2014), and the programs have been discussed in mainstream media outlets like the New York Times. There, for example, Rabin and Bookman (2006) argue that limits on liquor licenses in New York City constrain commercial development, tourism, and city nightlife. In addition, an extensive academic literature, which we review in section 3 below, examines a broad spectrum of issues related to the economics of alcohol. There is, however, little or no previous academic research that examines on-premise retailers and the licensing systems that regulate them. Regarding the latter, our review of the literature suggests that no empirical economic analysis of markets for liquor licenses has been undertaken. In fact, empirical economic studies of the determinants of prices for transferable operational licenses in other contexts are scarce—to our knowledge the only other existing studies of markets with features similar to the Montana liquor license market are O'Donnell (2017) who studies markets for taxi medallions and the determinants of their prices in several major U.S. cities, and **?** who analyze the determinants of New Zealand individual fishing quota values.

In our empirical investigation of the determinants of license values, we address several issues of current interest to the alcoholic beverage industry in Montana and elsewhere. One such issue relates to the impacts of small breweries on liquor license values. The number of small breweries has grown dramatically in recent years, from less than 10 at the turn of the century to 80 at present.[2] Whether the restrictions on the operation of these breweries have effectively mitigated their impact on the value of existing liquor licenses is a question that has not previously been addressed. Another interesting issue relates to the fact that, until very recently, when a new license has become available in a quota area as a result of growing population, the mechanism for allocating that license

[2]In addition to these beer breweries, there are also roughly 20 distilleries and a number of wineries (including businesses making cider and mead) operating in the state today.

has typically been a random lottery.[3] Using detailed data on lotteries held since 2007, we discuss and examine several aspects of this allocation mechanism.

In section 2 below, we provide historical detail on the evolution of the Montana retail liquor license system, as well as a description of the three predominant types of licenses in use today. We then review the economic literature on alcohol and also describe existing research that has examined aspects of regulatory programs in other industries with similarities to the Montana liquor licenses system. We develop a theoretical framework and derive comparative statics predictions in section 4, describe our data and empirical methodology in section 5, and present our empirical results in section 6. In the final section of the paper, we conclude and discuss several relevant policy issues and implications of our analysis.

3.2 Background

This section begins with a brief history of legislative changes relevant to the Montana liquor license system, and then continues with a discussion of additional details on the three prominent license types and relevant regulatory features of the system. Near the end of the paper, when the reader better understands important features and details of this complex program, and has been introduced to our theoretical model and seen the results of our empirical analysis, we provide insights into the political economy of the program.

3.2.1 History

Prior to Prohibition, Montana had relatively few regulations governing the issuance of alcoholic beverage licenses—the state required applicants to pay a licensing fee (MT

[3] As of 2019, these lotteries were replaced with auctions (MCA 16-4-430), thereby re-allocating the rents associated with newly issued licenses from lucky lottery winners to the state of Montana.

Revenue Oversight Committee, 1980).[4] Once Congress voted to enact Prohibition, however, Montana moved relatively quickly, becoming the seventh state to ratify the 18th Amendment on February 19th, 1918. As in the rest of the United States, retail sales of alcoholic beverages were illegal in Montana between 1920 and 1933 at which time the 21st Amendment was passed to repeal Prohibition. Montana ratified the amendment in 1934 and then created the Montana Liquor Control Board (LCB) to regulate alcohol sales (Congressional Research Service, 2004).[5]

Modeling its control system after that of Alberta, all the package alcohol stores in Montana were owned and operated by the state. The LCB was granted a monopoly on all liquor sales, with bars and taverns initially limited to serving beer.[6] In 1937 the legislature passed House Bill 196, which permitted liquor-by-the-drink sales for on-premise consumption in licensed establishments statewide. To acquire a license to serve liquor, an establishment also had to hold a beer retail license.[7] In 1947, the limits on entry into the sector were made more explicit when the legislature created a quota system for incorporated communities, which still serves as the basis of Montana's liquor licensing system (Quinn, 1970).

Two major factors motivated the creation of the quota system: (1) state (rather than local) control of alcohol sales and (2) a desire to increase the value of licenses. The authors of the quota bill sought to limit local bureaucratic discretion in the issuance of licenses. By setting the number of establishments and relieving local government of licensing responsibilities, the state could indirectly limit consumption and avoid local corruption. Legislators realized the quota could not be unduly restrictive, however, or

[4] The primary source for the historical information in this section is Quinn (1970).

[5] The timeline in appendix A.3 lists important dates, events, and legislative modifications to the license system from 1933 to the present.

[6] Wine was also under the authority of the LCB, but did not gain popularity in Montana until the 1960s.

[7] During the decade following 1937, new beer retail licenses could be issued at the discretion of local authorities. which effectively regulated entry into the on-premises sector.

bootleggers would re-create the black markets of the Prohibition years.[8] Limiting the future number of establishments also served to protect incumbents and create value for existing licenses. Thus, both the government and the proprietors of existing bars and taverns were supportive of the quota. If licenses acquired considerable value, the state believed it would cut down on corruption because establishment owners who violated the state's liquor laws would face the prospect of losing their valuable licenses. The Association of Retail Liquor Dealers publicly supported the adoption of a quota system, because it would insulate them from new competitors as Montana communities grew and, in the process, their businesses would become more valuable.[9]

In 1963, the state extended the quota system for liquor licenses to include unincorporated communities. Until then, new licenses had been issued at the discretion of the LCB (Quinn, 1970). The issuance of new on-premises retail beer licenses, however, remained under the LCB's control. State liquor stores held a monopoly on all off-premises sales of liquor, beer and wine until 1967, at which time the state created licenses for off-premises beer sales, which were also to be issued at the LCB's discretion (Revised Code of Montana, 1967).[10]

Montana's liquor licensing system underwent substantial changes in the 1970s. In 1972, Montana ratified a new constitution, under which the state government was reorganized. In the reorganization, the Montana Liquor Control Board was dissolved, and its responsibilities were transferred to the newly created Liquor Control Division, which was housed within the state's Department of Revenue (MT Subcommittee on Judiciary, 1974). The state also made major changes to on-premises licenses in 1974-1975,

[8]From the Minutes of Senate Standing Committees: Business Administration, Feb. 6-26, 1947. Montana Legislative Assembly (30th: 1947) records, 1945-1947. Legislative Records 30. Box 1, Folder 22. Montana Historical Society Research Center. Archives. Helena, Montana.

[9]In the fall of 1966, the Association of Retail Liquor Dealers became the Montana Tavern Association, which is the current lobbying group for owners of retail on-premises all-beverage and retail (or city) beer licenses.

[10]Businesses that were eligible for these licenses included grocery stores, drugstores that were licensed as pharmacies, and dedicated beer stores.

when a wine amendment was created for beer-only retail licenses, jointly held liquor and beer licenses were folded into "all-beverage" licenses, and the state created "floater" all-beverage licenses, which, under limited conditions, are allowed to move between communities.[11] Changes were also made to off-premises sales regulations—state liquor stores were banned from selling beer in 1975, while owners of off-premises beer licensees were granted the right to sell wine through a voter initiative in 1978 (MT Revenue Oversight Committee 1980).

The only notable change that appears to have taken place in the 1980s was that in 1983, legislation was passed that instituted a lottery system for newly issued licenses. Prior to that time, when the population-based quota formula dictated that a new license was to be issued in an area, it appears that the process for determining the new license owner was costly and time consuming, with outcomes largely based on subjective considerations.

Between 1995 and 1999, three important changes were made to the liquor license system. The first change was the creation of a new type of license, known as Restaurant, Beer, and Wine (RBW) or "cabaret" licenses. In the early 1990s, a movement gained ground to allow restaurant owners to be able to serve beer and wine without having to acquire a gaming license, which by that time had attained substantial value.[12] During the 1993 legislative session, a bill (Senate Bill 240) was proposed that would have created RBW licenses. Owners of existing all-beverage and beer and wine licenses and their MTA lobbyist objected to this bill for a variety of reasons, with a common concern being that the creation of this new class of licenses would substantially diminish the value of their licenses, for which many of them had paid considerable sums. In the face of this opposition, SB240 was defeated. In 1997, legislation was again proposed (in Senate Bill

[11]We provide more detail on the differences among license types below.

[12]Prices for all-beverage licenses (all of which included gaming rights at that time) in major Montana cities referenced between 1991 and 1993 were in the range of $100,000 - $175,000. See, for example, Stergionis (1991).

354) to create RBW licenses. This bill also drew strong objections from the MTA, at least initially. Over the course of the legislative session, however, a host of changes were made to the bill, and by the end of the session, the MTA had agreed to support the bill and it passed. Lurking in the background at this time was the specter of an initiative to let the public vote whether RBW licenses would be allowed—an option that concerned liquor license owners.[13] The quota for these licenses was set as a percentage of the beer retail license quota, thereby effectively limiting RBW licenses to incorporated communities. In 2007, with the support of the MTA, the legislature voted to double the number of available restaurant beer and wine licenses.

The second change relates to the linkage between liquor licenses and gambling rights in Montana. Prior to 1972, gambling was illegal in Montana. A relatively unusual feature of Montana's alcohol licensing system is the explicit packaging of gambling rights with alcoholic beverage licenses. Gambling was illegal in Montana until 1972, when the new constitution granted the legislature the right to legalize games of chance. In 1976, the Montana Supreme Court legalized video keno, but in 1984 ruled that video poker was illegal. In 1985, the legislature passed the Video Poker Machine Act, which allowed five video poker machines per alcoholic beverage license, with gaming rights being automatically attached to all existing all-beverage and on-premises beer licenses.[14] In 1991, the legislature replaced the five-machine limit on video poker machines with an overall cap for each license of 20 gaming machines that is still in effect today.[15] In 1997, when RBW licenses were created, the linking of gambling rights to RBW licenses was expressly prohibited. This restriction reduced the MTA's objections to the creation of

[13]A reasonable characterization of the changes made to SB354 was that they made operating with an RBW license more costly and therefore the new licenses became less of a threat to the value of existing licenses. The threat of a voter initiative was credible because a similar initiative allowing wine to be sold by licensed off-premise vendors was approved by voters in 1978.

[14]The tying of gambling rights to RBW licenses was explicitly prohibited when those licenses were created in 1997.

[15]See *History of Gambling in Montana* at `https://doj.mt.gov/gaming/history-of-gambling`.

RBW licenses. Also in 1997, the legislature voted to discontinue the linking of gaming rights to on-premises beer and wine licenses issued after that date, thereby dramatically limiting the issuance of new gambling rights since then to the relatively small number of licenses issued in unincorporated areas.[16] A decade later, the legislature revoked gambling rights on all-beverage floater licenses transferred after 2007.

The third change made during this period was to alter the status of microbreweries. As of 1995, in the context of the growing nationwide popularity of microbreweries, Montana had a dozen active microbreweries, each constrained to produce less than 15,000 barrels annually and to sell their products only through restaurants, grocery stores and bars. A bill to allow microbreweries to sell beer on their premises was introduced in 1995, but in the face of opposition from the Montana Tavern Association (MTA), was defeated. In 1999, another bill proposing that Montana's 20 small breweries be allowed to sell up to 48 ounces of their beer to individual consumers between the hours of 10:00 A.M. and 8:00 P.M. was enacted, this time with the support of the MTA.[17] Liquor license owners at the time clearly did not foresee the future expansion of Montana's microbreweries (both in numbers and the scale of their operations), which we discuss in more detail below.

In 2005, two other notable changes were made. First, until that year, state law limited ownership of alcohol retail licenses to state residents. In response to a lawsuit challenging that restriction, the Federal District court ruled the law was unconstitutional, thereby allowing ownership by non-residents.[18] Also, the state legislature passed a state-wide indoor smoking ban in 2005, although bars were exempted from complying with the law

[16]All licenses issued prior to 1997 were allowed to keep their gambling rights. The near-simultaneous timing of the creation of RBW licenses and the restrictions on new gambling rights was not coincidental, with the latter being part of the political concessions that convinced the MTA to support the creation of RBW licenses.

[17]The MTA's long-time lobbyist Mark Staples indicated that although bar (and liquor license) owners originally opposed the bill, they then agreed to support it after being convinced that breweries did not want to directly compete with bars, who were a primary purchaser of beer (Curliss, 1999).

[18]See *McGrath: State will Not Appeal Liquor-License Decision* at `https://doj.mt.gov/2005/06/mcgrath-state-will-not-appeal-liquor-license-decision`.

until October of 2009.[19] In 2013, the limit on the number of all-beverage licenses in which an individual could have an ownership share was increased from one to three. In 2017, legislation was passed to discontinue the lottery system for new licenses and replace it with an auction process, to be implemented in 2019.

3.2.2 License Descriptions and Regulatory Provisions

The state of Montana regulates—and requires licenses for—almost all activities related to the manufacturing, marketing, and selling (at both wholesale and retail) of alcohol.[20] Under Montana's three-tiered regulatory system, separate and distinct licenses are issued for manufacturers (and importers), wholesalers, and retailers. In addition, limitations are placed on the financial linkages between the three tiers. Manufacturers and wholesalers are prohibited from having a financial interest in any business that sells alcohol at retail. Similarly, manufacturers and importers are not allowed to have a financial interest in a wholesaler. State law also limits the amount of advertising tools, paraphernalia, and gifts that manufacturers and wholesalers may provide to retailers. Beer imported into the state must be sold to wholesalers, who store, then sell and distribute the beer to both off- and on-premises retailers, who sell to consumers. Imported liquor, on the other hand must be sold to the Department of Revenue, who then sells only to licensed liquor wholesalers (referred to as agency or franchise liquor stores), which are limited in number by the state.[21] These agency stores, in turn, sell to bars and restaurants that have all-beverage licenses; the agency stores also sell directly to walk-in consumers for off-premises consumption.

[19] See *Talking about the Montana Clean Indoor Air Act* at `http://www.dphhs.mt.gov/mtupp//cleanairact/talkingaboutciaa.pdf`

[20] For information on state-issued alcohol-related licenses, see Montana Department of Revenue, Alcoholic Beverage Licenses at `https://mtrevenue.gov/liquor-tobacco/liquor-licenses/`.

[21] The Bozeman quota area, for example, has two licensed liquor wholesalers who are allowed to purchase packaged liquor from the state. There are currently 95 agency liquor stores in Montana. See `https://mtrevenue.gov/liquor-tobacco/agency-liquor-stores/store-list/`. The agency stores originated as state-owned enterprises, but were quasi-privatized as state franchises in 1995.

The Montana Department of Revenue (DOR) issues new licenses to businesses in each of the tiers through an application process and also reviews all license transfers. Applicants for both new and transferred licenses are required to satisfy the same ownership requirements. In particular, the department carefully vets potential licensees to ensure that the state's constraints on financial linkages between tiers are observed.

For some activities in the liquor businesses, there are few if any limits on the number of licenses that the state issues. For other activities, the state imposes limitations on entry that are sufficiently restrictive that the rights to engage in those activities have substantial value. As an example of the former, the number of off-premises licenses to, e.g., grocery stores (who, in Montana, can sell beer and wine, but not liquor for off-premises consumption), is not limited by statute. The two categories of licenses that often have substantial value as a result of state-created limitations on entry are those for (1) the agency or franchise liquor stores, and (2) the on-premises licenses, with the latter being the focus of the following discussion.

Numerous license types exist for on-premises sales. Historically, the three predominant license types, which are the focus of our analysis below, have been retail all-beverage licenses, retail (or city) beer licenses, and restaurant beer and wine licenses.[22] Retail beer licensees can serve wine simply by adding a wine amendment, and each of the license types is eligible for an optional catering endorsement, which allows alcohol to be served at off-site events held within 100 miles of the licensee's establishment.

The issuance or transfer of any of the three major license types is subject to a ten-day waiting period. If the DOR receives letters of protest during that time, a public hearing is held to determine whether there is sufficient cause to deny an application. Until 2017,

[22]Other on-premises license types include resort all-beverage, airport all-beverage, golf course beer and wine, tribal and military, veterans/fraternal organization, non-profit arts, West-Yellowstone Airport, Montana Heritage Preservation and Development, passenger carrier, and special (temporary). Sale for on-premises consumption of alcohol is also allowed for manufacturers who have small (as defined by state law) brewery, winery, or distillery licenses.

when there were available licenses and the number of applicants exceeded the number of available licenses, the department allocated them through a lottery.[23] The DOR will only issue one retail on-premises license per building.[24] Retail licensees may, however, operate other businesses (that do not serve liquor) simultaneously in the same building, provided that the on-premises retail alcohol sales area is physically separated. Owners of on-premises retail licenses may not hold an interest in an agency liquor store.

All three of the predominant on-premises license types—each of whose provisions we discuss in more detail below—are subject to quotas, which are determined solely by the population within a quota area.[25] For both incorporated communities and unincorporated areas, the quota is re-calculated annually by the DOR using the Bureau of the Census's annual estimates of population within each quota area. Quotas for an incorporated community apply for a radius of five miles beyond the boundaries of the community. For the time period spanned by our empirical analysis, if two communities had overlapping five-mile radii, they were combined into a single quota area, with the area's quota of licenses determined based on the total population of the combined area.[26]

A complicating feature of the Montana quota system is that the formula-based quota numbers are "soft" in the sense that communities are allowed to have more licensed establishments than indicated by the quota formulas. The excess of licenses may occur for at least three reasons. First, the law passed in 1947 that established the quota system stipulated that existing businesses were to be grandfathered into the new system.

[23]In November of 2017, the Montana legislature voted to replace the lottery system with a competitive auction process, to be implemented in 2019.

[24]A given license can be used in more than one retail business that serves alcohol, so long as the businesses are all under the same roof. We know of at least four such arrangements in Bozeman, the area with which we are most familiar

[25]The formulas used for each type of license are presented and discussed in appendix A.3.

[26]During the time span of most of our data, there were six combined quota areas: Bozeman/Belgrade, Helena/East Helena, Whitefish/Columbia Falls, Red Lodge/Bearcreek, Hamilton/Pinesdale, and Eureka/Rexford. Additionally, Butte/Silver Bow County and Anaconda/Deer Lodge County are each single quota areas (with the entirety of both counties being incorporated). In November of 2017, Montana SB5 mandated the re-separation of the quota areas in these communities.

Existing establishments in excess of the quota at the time of the new law were allowed to continue operation. Second, in those instances where population has fallen over time (as it did, for example, in Butte), the state does not automatically rescind any "excess" licenses.[27] Third, for all-beverage licenses, under certain limited, but well-defined conditions, communities may "float" additional all-beverage licenses into their quota areas, even though the initial number of licenses exceeds the formula-based quota number.[28]

3.2.3 Ownership and Transfer Regulations

Retail all-beverage licensees are permitted to sell beer, wine, and liquor by the drink for on-premises consumption between the hours of 8:00 A.M. and 2:00 A.M. Licensees may also have self-service shelves or coolers exclusively to sell alcohol for off-premises consumption, provided that these are contiguous to, but physically walled-off from, the on-premises area.[29] All-beverage licenses with gambling rights attached may offer card games, sports pools, and video gaming machines. The number of gaming machines is limited to 20 per establishment, and there is a 15 percent state tax on the net revenue from gaming machines (defined to be the difference between the money put into the machine and the machine's payouts). There are quotas for all-beverage licenses in both incorporated and unincorporated areas. Currently, an all-beverage licensee may own an interest in no more than three licenses, or half the licenses in a quota area, whichever is fewer.[30]

[27] The state may revoke licenses, but only for violation of state law or if a licensee does not operate a retail establishment for a period of 90 days, although the period may be extended beyond 90 days for extenuating circumstances.

[28] As explained in more detail below, for all-beverage licenses, additional licenses will be allowed in a community any time the number of licenses issued falls below 1.43 times the number of licenses allowed by the quota (so long as an additional license does not increase the number beyond 1.43 time the allowed number of licenses).

[29] Licensees may not provide accommodations for self service on-premises consumption. Relatively small liquor stores that are contiguous to bars and lounges with all-beverage licenses are quite common.

[30] An ownership share in three licenses represents the allowable total across all quota areas. Prior to 2013, an individual could hold an interest in only one all-beverage license. Note that neither an individual who owns 1 percent of three different licenses nor one who owns 100 percent of three licenses

All-beverage licenses may be sold with few restrictions within a quota area, subject to receiving approval for the sale from the DOR. In addition, if the number of active all-beverage licenses in an incorporated quota area exceeds the area's formula-based license quota by less than 43 percent (33 percent if the population is less than 10,000), and if an additional license would not put it over that threshold, the state will allow another all-beverage license in that area. When these conditions hold, the state announces the availability of an additional license and a deadline for the submission of applications. During the time span of the data we use in our empirical analysis, if the number of applications exceeded the number of additional licenses offered, a lottery was held. The winner of the lottery was then allowed to purchase an all-beverage license from another incorporated area and "float" it into the quota area where the lottery was held. The quota area from which the "floater" license was acquired had to be "over-quota" by more than 25 percent, and the loss of the floater license could not reduce the number of active licenses below the 25 percent threshold.

Since 2007, floater licenses have been gaming-restricted (i.e., licensees may not offer gambling in the quota area to which they are moved), may not be re-transferred, mortgaged, or pledged as security on a loan for five years, and must remain permanently in their new quota area. Additional fees required by the state (associated with actually obtaining the floater license and opening a retail operation) include a one time Liquor Control processing fee of $400 and annual license renewal fees that range from $400 - $800, depending on the population of the quota area. The lottery winner must also prepare and submit an application, which is then investigated by the Montana Department of Justice (MCA 16-4-402, 2017).

Retail on-premises beer licenses (often referred to as city beer and wine licenses) permit the sale of beer for on-premises or off-premises consumption. With a wine amendment, a licensee may also sell wine. To obtain a wine amendment, licensees must demon-

is allowed to acquire ownership in a fourth license.

strate that wine would complement their food service, and pay a one-time processing fee of $400 as well as an annual license renewal fee of $400. For off-premises sales, licensees are regulated by the same rules as all-beverage establishments (i.e., beer for off-premises consumption must be sold from a contiguous, but physically separate space). All licenses issued in unincorporated communities may offer gambling. Licenses issued in incorporated communities before 1997 may offer gambling, but those issued after 1997 are gaming-restricted. The number of beer retail licenses is subject to the quota schedule within incorporated communities, but is left to the DOR's discretion in unincorporated areas. The number of retail beer licenses an individual may hold is not limited by state law. All transfers of retail beer licenses must take place within the licensee's quota area.

The third type of on-premises license is the restaurant beer and wine (RBW) license (also known as a cabaret license). RBW licensees face a more restrictive set of constraints than the two types of on-premises retail licensees just described. Establishments holding an RBW license may sell beer and wine, but only for on-premises consumption from 11:00 A.M. to 11:00 P.M.[31] Similarly, licensees may not offer gambling under any circumstances. To qualify for a restaurant beer and wine license, establishments must provide individually-priced meals, prepared and served on site, and at least 65 percent of their total revenue must come from food sales. In addition, the DOR may issue no more than 25 percent of all RBW licenses to large restaurants, i.e., those with seating capacities of over 100.

The quota and allocation processes are also somewhat different for RBW licenses. Rather than being a direct function of population, state law sets the quota for RBW as a percentage of the retail on-premises beer quota (the exact percentage varies by population size, as can be seen in appendix B). In unincorporated areas, where there is no retail beer quota, the department does not issue RBW licenses. Until November of 2017, like

[31]No off-premises sales of any kind are allowed. This restriction applies, for example, to unfinished bottles of wine.

all-beverage and retail beer licenses, the department held lotteries (f necessary) for newly issued RBW licenses. The lottery for restaurant licenses, however, was a priority lottery in which applicants with preference were awarded licenses before those without it. The department granted preference to all applicants who operated restaurants in the quota area, but did not possess an alcohol license. Additionally, state law limits the ability of a business to "trade down" across license types—in particular, if a business sells its retail beer and wine license, it must wait a year before it is allowed to obtain an RBW license.

3.2.4 Related Literature

Although an extensive literature exists regarding the economics of alcohol, the majority of it does not directly relate to the issue of supply-side licensure. Rather, most studies tend to focus on the demand side of the retail market and the effects of consumption. The literature is well developed regarding the effects of advertising and legal changes on alcohol sales, along with the associated (potentially negative) effects of drinking.

Given that alcohol consumption has the potential to create negative externalities, examining how advertising and institutions affect consumption has definite policy implications. Tremblay and Okuyama (2001) find that although there is little evidence that advertising increases market demand, consumption may still increase. Advertising may lead to greater price competition among producers, increasing the quantity supplied. On the other side of the issue, Saffer and Dave (2002) examine advertising bans and find that they have a significant negative impact on alcohol consumption.

Along with limiting advertising, government may institute other regulations to potentially reduce the consumption of alcohol. Taxation of alcohol, for example, is universal across U.S. states. A less common regulation is the Sunday sales ban (also known as a blue law). Cook and Tauchen (1982) examine how liquor excise taxes affect consumption, using liver cirrhosis as a proxy for long-term heavy drinking. All else equal, they find that a one dollar increase in liquor excise taxes reduces heavy drinking by 5.4 percent, with a possible larger long run decline. Sunday sales restrictions, however, have a less profound effect. They find that allowing alcohol to be sold on Sundays may increase consumption on Sundays, but does not have a significant effect on total consumption.

Health effects related to alcohol consumption extend well beyond direct mortality for the person drinking. Another substantial portion of the literature attempts to measure the impact of alcohol consumption on traffic-related accidents and deaths. In particular, the literature examines the impacts of changes in regulations regarding the purchase of alcohol. McMillan and Lapham (2006) find that lifting the Sunday off-premises sales ban in New Mexico leads to an increase in Sunday traffic accidents by 29 percent. They employ a time series for a single state, however, making it hard to attribute causality. Indeed, with a state-level panel, Lovenheim and Steefel (2011) find that lifting blue laws results in no measurable impact on traffic accidents or fatalities. Similarly, Miron and Tetelbaum (2009) determine that previously estimated reductions in alcohol-related deaths from increasing the minimum drinking age to 21 are probably overstated.

In addition to the effects on traffic accidents, another large portion of the literature looks at the connection between alcohol and crime. Markowitz (2005) finds that amidst other drug regulations, beer taxes have a negative relationship with some violent crime—for example, assaults decrease with higher beer taxes, but rape and robbery do not. Similarly, zero-tolerance drunk driving laws do not have a measurable impact on violent crime, but may lead to a slight decrease in petty criminal offenses (Carpenter, 2007). Using a county panel in Virginia, Heaton (2012) re-examines the effect of blue law repeals within the context of crime. Employing difference-in-difference and triple-difference approaches, the evidence suggests that allowing package sales on Sundays increases both petty and violent crime. Anderson et al. (2018) estimate that as Kansas voted to legalize the sale of alcohol to the general public in recent decades, a 10 percent increase in the number of drinking establishments was associated with a 3 to 5 percent increase in violent crimes.

Of the studies that explicitly treat the licensing of retail liquor sales, off-premises establishments are the predominant focus. Smith (1982) examines the political economy of liquor store laws, evaluating the effects that special interests have on the licensing structure across the United States. Similarly, Seim and Waldfogel (2013) examine the Pennsylvania liquor control system. The state is the sole owner and operator of off-premises liquor stores, leading to different numbers and locations of stores relative to a license-to-operate system. While Toma (1988) examines retail license quotas, the analysis is limited to the context of dry/wet counties—states with more restrictive licensing are found to have higher numbers of dry counties.

Within the literature on the economics of alcohol, there is a contribution to be made by an examination of the factors that determine the market values of tradable on-premises retail licenses. Broadly speaking, relatively little has been done examining on-premises retailers. More specifically, no literature exists looking at license prices in the context of liquor quota systems. This topic also has relatively wide applicability, because almost half the states allow transfers of retail licenses (although not all have license quotas, and we could find no others from whom sufficient data for analysis of transaction prices are available).[32]

Although issues related to tradable alcohol retail licenses have not been examined in the literature on the economics of alcohol, there are several other regulatory contexts in which various issues related to transferable permits and production quotas have been examined. Tradable production quotas have been used in a number of agricultural pro-

[32]Based on an examination of all 50 states' online statues individually, it appears that just over twenty states permit the transfer of retail licenses. Until recently, Montana was the only state we found with accessible data on enough license transactions to accommodate statistical analysis of the determinants of prices. Although we recently found what may be sufficient transaction price data for New Mexico, a detailed analysis of that system and the determinants of license prices within it are beyond the scope of the present paper.

grams, as well as in the regulation of certain natural resources. For example, both the United States tobacco and peanut programs have employed production quotas (Rucker and Thurman, 1990; Rucker et al., 1995; Brown et al., 2007). The federal quota systems for both peanuts and tobacco were similar along certain dimensions to the Montana alcoholic beverage license quota system. Tobacco and peanut quota could be transferred between producers within counties, but could not be sold or leased across county lines—similar to the transfer restrictions in Montana on retail alcohol licenses. Such regulatory limitations create divergent prices (for poundage quota or liquor licenses) between quota areas. Although such price differences suggest there are potential social gains from allowing trade across county lines, not all industry participants would gain from that elimination of trade restrictions. In counties where, for example, liquor license values are high, license owners would lose because those values would fall with the elimination of trade restrictions. [33]

Another market with similarities to the Montana alcohol licensing system is the market for taxicab medallions in New York City.[34] The NYC regulatory system requires taxis to possess a medallion to operate and has imposed limits on the number of medallions issued since 1947. Medallions are tradable, and in New York City have occasionally fetched prices upwards of $1 million (O'Donnell, 2017). Like alcohol retail licenses, possession of a medallion confers the right to operate, but does not directly limit output.

This system is distinct from tradable individual fishing and agricultural production quotas, which strictly limit output. Although the imposition of a quota system may confer monopoly rents on those currently operating, entrants into the market face the expectation of earning normal profits because the medallion or quota price rises to equal the present value of expected future profits (Tullock, 1975). The existence of positive

[33]Rucker et al. (1995) investigate these issues (and quantify the potential gains from trade) in the context of inter-county trade restrictions in the federal tobacco program. A substantial literature also exists on tradable permits for air and water, beginning with Crocker (1966) and Dales (1968).

[34]Similar systems with tradable medallions exist in, e.g., Chicago, Philadelphia, and Boston.

prices for medallions, however, does not necessarily imply outcomes that are socially inefficient in all contexts. The regulatory barrier to entry on taxis may, for example, reduce externalities associated with congestion, and the limits on liquor licenses may reduce externalities associated with alcohol consumption. The entry barrier on taxis may also secure a premium for honest drivers so that licensing may increase the quality of taxi services provided (Demsetz, 1982; Cairns and Liston-Heyes, 1996). Similarly, high values for liquor licenses may induce license owners to closely comply with regulations rather than violate the rules and face the prospect of losing a valuable asset.

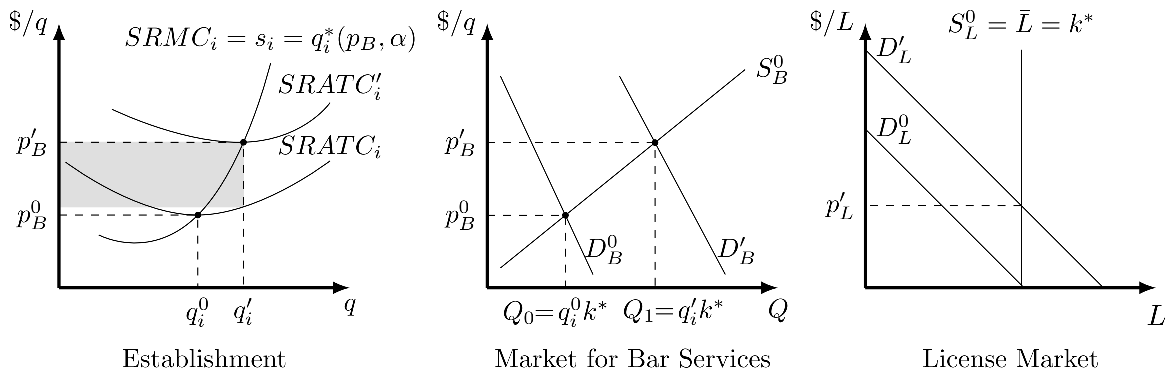
Although there is not a literature that specifically examines the issuance and trade of retail alcohol licenses, similar issues arise in other industries and existing analyses of these provide insights into developing tools for analysis. Individual tradable fishing quotas, agricultural production quotas and taxicab medallions have particular relevance for Montana's licensing system, given the tradable nature of the licenses. Even in those other contexts, however, very little has been done in terms of examining how transferable operational licenses are priced—to our knowledge the only other existing studies of these markets are O'Donnell (2017) and Roistacher and White (1982) who study taxi medallion prices and Newell et al. (2007) who analyze the determinants of New Zealand individual fishing quota values. In light of with the lack of research in the area of alcoholic beverage retailers, our empirical analysis below provides a clear contribution to the literature.

3.3 Theory

3.3.1 The Competitive Market Model

In this section, we develop a theoretical model of regulation of the on-premises retail liquor industry in which a quota-area license program like the Montana system described above is implemented. Intuitively, the regulatory framework reduces the number of establishments that provide bar services below the number that would operate in an unrestricted

Figure 1: Firm Costs, Market for Bar Services, and License Market



competitive market.[35] The corresponding reduction in the supply of bar services causes an increase in the price of those services, which results in positive profits for the firms that are allowed to operate in the regulated market. Liquor licenses are transferable (within, but not between quota areas) in our model, which leads to the present value of expected profits being capitalized into the market price of the licenses. The model we develop here is for a single quota area and is a one-period model. At the end of the section, we discuss the extension of the model to an infinite period framework.

In the center panel of figure 1, the initial demand for bar services is labeled D_B^0. Under competitive conditions with identical price-taking firms, the ith individual establishment maximizes its short run profits:

$$\max_{q_i} \; \pi_i = p_B^0 \cdot q_i - c(q_i, \boldsymbol{\alpha}) - F \tag{1}$$

where p_B^0 is the initial market price of bar services (assumed exogenous to the i^{th} firm), q_i^0 is the quantity of bar services produced by firm i, $c(q_i, \boldsymbol{\alpha})$ is the short run variable cost function, $\boldsymbol{\alpha}$ is a vector of cost shifters, and fixed costs are F. The first order condition

[35] Bar services can be thought of as an index of goods and services, such as alcoholic beverages, food, gambling, ambiance, etc.

for profit maximization is

$$\frac{\partial \pi_i}{\partial q_i} = p_B^0 - c'(q_i, \boldsymbol{\alpha}) = 0, \tag{2}$$

i.e., that marginal revenue equals short-run marginal cost. Solving the first order condition for the optimal output as a function of price and cost shifters yields the firm's short-run supply function, $q_i = q_i^*(p_B, \boldsymbol{\alpha})$, which will be upward sloping by the second order conditions ($s_i = SRMC_i$, above p_B^0 in the left hand panel of figure 1). The short-run market supply for bar services (S_B^0 in the middle panel of figure 1) is the horizontal summation of the individual firms' supply functions, given by:

$$Q_S^B = S_B(p_B, \boldsymbol{\alpha}) = \sum_{i=1}^{k^*} q_i^*(p_B, \boldsymbol{\alpha}) \tag{3}$$

where k^* is the equilibrium number of firms in the market, and the market-clearing quantity is $Q_0 = q_i^0 \cdot k^*$. In the long run, positive or negative profits will induce entry or exit until the market supply curve shifts to the point where profits are zero and price equals the minimum of long-run average cost, which is assumed to be at p_B^0 in figure 1.[36] The market demand function is:

$$Q_D^B = D_B(p_B, \boldsymbol{\beta}) \tag{4}$$

where $\boldsymbol{\beta}$ is a vector of demand shifters. Assuming the industry is constant-cost, the long-run market supply is perfectly elastic at price p_B^0, and market demand determines both the long-run equilibrium quantity (Q_0) and the number of zero-profit firms in the industry, k^* ($= Q_0/q_i^0$).

[36]To show that long-run profits are zero at p_B^0, the first panel in figure 1 would also include a long-run average cost curve that lies everywhere below the short-run average total cost curve except at q_i^0, where the short- and long-run average cost curves are tangent. We mention this for clarity, but to reduce clutter in the figure, do not include the long-run average cost curve.

3.3.2 Quota for Licenses

Now, suppose the regulatory authorities implement a license system and set the initial number of licenses at $k^* = \bar{L}$.[37] The third panel of figure 1 shows the market for licenses, where they are transferable within a quota area and their market price reflects the profits earned from owning a license. The supply of licenses, whose number is determined by regulators, is assumed to be vertical, and the demand is downward sloping.[38] Because the initial quantity of licenses issued by regulators is assumed to equal the number of bar-service-providing firms operating in the competitive equilibrium, the initial marginal willingness to pay for a license is \$0.

Now, suppose the demand for bar services increases to D'_B in figure 1, say, because of increased incomes or tourism. In the short run, the price of bar services will rise to p'_B. Ignoring $SRATC'_i$ for the moment, each firm will move up along its marginal cost curve, increasing production to q'_i and profits to the shaded area in the first panel. The market-clearing quantity increases to $Q_1 = k^* \cdot q'_i$. With no license quota program, in the long run, profits would be competed away by entry of firms into the bar service industry.

[37] As indicated in our discussion above, in Montana, the quota system was introduced in 1947—into a market that was previously competitive—by giving all operating establishments a license. Assuming zero profits in the pre-quota-system competitive environment, licenses had little or no value when the quota system was introduced. Over time, various factors caused the demand for bar services (and, as a result, the demand for licenses) to increase. Simultaneously, population in some Montana cities increased, with concomitant increases in license numbers. The fact that license values have increased over time suggests that the price-increasing effects of increasing demand in the license market have outweighed the price-decreasing effects of increases in supply.

[38] That the demand curve for licenses is downward sloping can be demonstrated by realizing that the change in profits with respect to an exogenous change in the number of licenses (L) can be written as

$$\frac{\partial p_L^*}{\partial L} = \frac{\partial \pi_i^*(\cdot)}{\partial L} = \frac{\partial \pi_i^*(\cdot)}{\partial p_B} \cdot \frac{\partial p_B^*(\cdot)}{\partial L} = q_i^*(\cdot) \cdot \frac{\partial p_B^*(\cdot)}{\partial L} < 0,$$

where the first two terms indicate that the market clearing price of licenses, P_L^*, is equal to the firm's maximum attainable profits, or the individual firm's indirect profit function, $\pi_i^*(\cdot) = \pi_i^*(p_B, \boldsymbol{\alpha}, F)$. In the second two expressions, $q_i^*(\cdot) = q_i^*(P_b, \boldsymbol{\alpha})$ is the firm's supply function of bar services. By the Envelope theorem, $\frac{\partial \pi_i^*(\cdot)}{\partial p_B} = q_i^*(\cdot)$ (which is positive), and $\frac{\partial p_B^*(\cdot)}{\partial L} < 0$ because an increase in the number of licenses causes the supply of bar services (in the middle panel of figure 1) to shift to the right, resulting in a lower market price.

In the middle panel of figure 1, the market supply would shift to the right until price returned to p_B^0. With the quota program and transferable licenses, however, profits are competed away in the market for licenses, where the demand curve in the third panel of figure 1 shifts from D_L^0 to D_L^1.[39] In equilibrium, the market price of licenses in the third panel (p'_L) will equal the shaded area of profits in the first panel. A firm entering the market will have to pay the shaded area (or, equivalently, p'_L) for a license, so its average costs are represented by $SRATC'_i$ rather than by $SRATC_i^0$.[40] Similarly, existing establishments must bear the opportunity cost of holding a license—it could be sold for p'_L and the sale proceeds invested—which causes the firm's average costs to increase to $SRATC'_i$.

In the market for licenses in figure 1, supply is perfectly inelastic, set exogenously by a regulatory agency. Thus, the supply function for licenses is

$$Q_S^L = \bar{L} = k^*. \tag{5}$$

Each establishment is limited to a single license and a business's willingness to pay for a license is equal to the profits it will earn. An important determinant of profits is the price of bar services. Accordingly, the demand function for licenses (denoted by D_L^0 in the third panel of figure 1) depends on these prices, as well as on the price of licenses (p_L):

$$Q_D^L = D_L(p_L, p_B). \tag{6}$$

[39]Note that in the scenario we describe here, neither the individual firm's, nor the market supply of bar services, is affected by the implementation of a license quota program—firms in the industry when the quota program is introduced continue to operate as before. Moreover, unlike with a production quota, the supply curve does not become vertical. When the price increases to p'_B, firms may adjust along various margins to increase quantity, e.g., they may increase available seating and staff. In the long run, we would predict an increase in the capacity of retail on-premise establishments in markets where license prices are higher—a prediction supported by our casual empiricism, at least in the Bozeman market.

[40]The price of the license represents an increase in fixed costs, F, but not in marginal costs, because the cost of the license does not vary with the level of bar services provided.

The horizontal intercept of the initial license demand curve is at k^*—the free-entry equilibrium number of firms.

The market price for licenses, p_L, rather than entry or exit, is the mechanism through which profits are competed away. To enter the market, a business must first purchase a license, the price of which reflects the profits to be earned. Similarly, existing establishments must bear the opportunity cost of holding a license—it could be sold for p_L (and the proceeds invested). In either case, average cost will increase until establishments no longer earn positive profit. In the first panel of figure 1, this is the shift from $SRATC_i^0$ to $SRATC_i'$, where p_L is equal to profits (the shaded area). Additionally, each licensed establishment increases output, from q_i^0 to q_i'.

Additionally, profit (or expected profit) may differ based on license attributes. In the context of Montana's licensing system, there are several types of licenses. All-beverage, beer and wine, and RBW licenses may each be thought of as having a distinct market within quota areas, each following the model above. The licenses are imperfect substitutes, with premiums for certain license types. *Ceteris paribus*, an all-beverage license is predicted to yield higher expected profits than a beer and wine license, because it imposes the fewest constraints on the selection of alcoholic beverages a license holder can offer. Similarly, (1) a beer and wine license will sell for more than an RBW license, because the hours of operation permitted with the latter are more restricted and there is no minimum required proportion of revenues from food generated under a beer and wine license, and (2) a license (either all-beverage or beer and wine) with gambling rights attached is predicted to have a higher market price than a gaming-restricted license of the same type because of the extra revenues (and profits) generated with gambling operations.

The model to this point is a one-period model, but market-clearing license prices are determined in a multi-period reality as the present value of expected future profits. Referring back to figure 1, p_B^0 is the current price of bar services in a one-period world.

Suppose this also represents the expected price of services into the indefinite future. Then, p_B^0 can be relabeled $\mathbb{E}[p_B^0]$ and demand, supply and cost curves can be similarly relabeled (e.g., $\mathbb{E}[D_B^0]$, $\mathbb{E}[S_B^0]$, and so forth). The profit rectangle in the first panel of figure 1 can now be thought of as the expected profits in period t ($t = 0, 1, 2, \ldots$). Assuming expected profit in each period is $\mathbb{E}[Profits_0]$, the expected present value of those profits in an infinite period model with interest rate i would be equal to

$$V(Profits) = \frac{\mathbb{E}[Profits_0]}{(1+i)} + \frac{\mathbb{E}[Profits_0]}{(1+i)^2} + \cdots = \frac{\mathbb{E}[Profits_0]}{i}. \tag{7}$$

This value will be the price of licenses (p_L) in the third panel of figure 1. Alternatively, suppose bar-service owners' expectations are that the market-clearing price of bar services will be changing over time. Then, for any sequence of expected future prices, there is a price that is constant over time and yields the same discounted value as the non-constant future sequence of prices.[41]In this scenario, we could replace the bar service price in figure 1 with $\mathbb{E}[p_B^c]$ where c indicates that constant-value-flow-equivalent (in present value terms) of the non-constant actual sequence of expected prices. Corresponding to the constant expected price, of course, would be corresponding "constant" expected demand and supply curves. There would also be constant annual expected profits ($Profits_c$) that would yield a license price of

$$\text{License Price} = V(Profits) = \frac{\mathbb{E}[Profits_c]}{i} \tag{8}$$

which is the same price that would be observed with the corresponding non-constant flow of future prices, costs, annual profits, etc. Thus, although the graphical model figure 1

[41]To clarify, suppose (in a three-period model) that expected bar-service prices in the three periods are \$10, \$15, and \$18. At an interest rate of 5 percent, the discounted value of one unit of bar services in each period is \$36.68. A constant bar service price of \$14.20 yields the same discounted value of rides. In an infinite period model, the equivalent perpetuity flow to a non-constant flow of expected future values will be $i \cdot V(Profits)$, where $V(Profits)$ is the expected present value of profits associated with the non-constant flow of future income.

is a single period representation of the determination of license prices, it is relatively straightforward to modify it to be useful for conceptualizing the multi-period reality in which liquor license prices are actually determined.

3.3.3 Comparative Statics Predictions

Several testable implications are generated by the competitive market model with a binding quota discussed above. To derive these, consider the four equation system comprised of equations 2-6. Begin by setting $Q_D^B = Q_S^B$ and $Q_D^L = Q_S^L$ (or, equivalently, $D_B((p_B, \beta))$ $= S_B(p_B, \alpha)$ and $D_L(p_L, p_B) = \bar{L})$), and solving for the equilibrium prices of bar services and licenses as a function of the model's parameters:

$$p_B = p_B^*(\bar{L}, \alpha, \beta) \tag{9}$$

$$p_L = p_L^*(\bar{L}, \alpha, \beta) \tag{10}$$

Plugging these equilibrium values of the two prices back into the two market equilibrium conditions yields the two-identity system:

$$S_B(p_B^*(\bar{L}, \boldsymbol{\alpha}, \boldsymbol{\beta}), \bar{L}, \boldsymbol{\alpha}) \equiv D_B(p_B^*(\bar{L}, \boldsymbol{\alpha}, \boldsymbol{\beta}), \boldsymbol{\beta}) \tag{11}$$

$$\bar{L} \equiv D_L(p_L^*(\bar{L}, \boldsymbol{\alpha}, \boldsymbol{\beta}), p_B^*(\bar{L}, \boldsymbol{\alpha}, \boldsymbol{\beta})) \tag{12}$$

Comparative statics predictions from the model can be obtained by differentiating the two-equation system with respect to the model's parameters and then using Cramer's rule to solve for the effects of interest. In the context of Montana's licensing system, an important result is the effect of a positive demand shifter, denoted by β_1, on the market price for licenses. Differentiating equations (10) and (11) with respect to β_1 and solving for the impact on the equilibrium bar-services and license prices yields:

$$\frac{\partial p_B^*(\cdot)}{\partial \beta_1} \equiv \frac{-\frac{\partial D_B(\cdot)}{\partial \beta_1} \cdot \frac{\partial D_L(\cdot)}{\partial p_L}}{\left(\frac{\partial S_B(\cdot)}{\partial p_B} - \frac{\partial D_B(\cdot)}{\partial p_B}\right) \cdot \frac{\partial D_L(\cdot)}{\partial p_L}} > 0 \tag{13}$$

$$\frac{\partial p_L^*(\cdot)}{\partial \beta_1} \equiv \frac{-\frac{\partial D_B(\cdot)}{\partial \beta_1} \cdot \frac{\partial D_L(\cdot)}{\partial p_B}}{\left(\frac{\partial S_B(\cdot)}{\partial p_B} - \frac{\partial D_B(\cdot)}{\partial p_B}\right) \cdot \frac{\partial D_L(\cdot)}{\partial p_L}} > 0 \tag{14}$$

The interpretation of equation (11) is that an increase in the demand for bar services (from D_B^0 to D_B^1 in the second panel of figure 1) leads to an increase in the price of bar services (from p_B^0 to p_B^1). Regarding equation (12), intuitively, in the second panel of figure 1, an increase in the demand for bar services induces an increase in both the market quantity (from Q_0 to Q_1) and the price of bar services (from p_B^0 to p_B^1). In the first panel, the increase in price (holding constant the number of firms) results in an increase in the quantity produced by each firm (from q_i^0 to q_i^1) and in profits (from \$0 to the shaded area). In the third panel of figure 1, the increase in firm-level profits causes an increase in the demand for licenses (from D_L^0 to D_L^1) and thus an increase in their price (from \$0 to p_L^1).

For a given number of licenses, their price should therefore increase (decrease) as the demand for retail alcohol services increases (decreases). Demand shifters, including variables such as income, should have a positive effect on the prices of licenses.[42] Furthermore, population variables that do not affect the number of licenses–namely student population and tourism–should similarly be positively related to prices. Finally, the income from gambling should be priced directly into license values.

Consider next the effects of a change in, say, α_1, one of the parameters in the vector of supply shifters in equation 1 above. Suppose α_1 is the unit price of a variable input, so that an increase in α_1 causes production costs to increase. Differentiating equations (10) and (11) with respect to α_1 and solving for the resulting impact on the equilibrium

[42] *Ceteris paribus*, an increase in population is also predicted to increase the demand for bar services and the price of licenses. Within the Montana liquor license system, however, such an increase in population also triggers a rightward shift in the supply of licenses and a reduction in the price of licenses.

bar services and license prices yields the predictions:

$$\frac{\partial p_B^*(\cdot)}{\partial \alpha_1} = \frac{-\frac{\partial S_B(\cdot)}{\partial \alpha_1} \cdot \frac{\partial D_L(\cdot)}{\partial p_L}}{\left(\frac{\partial S_B(\cdot)}{\partial p_B} - \frac{\partial D_B(\cdot)}{\partial p_B}\right) \cdot \frac{\partial D_L(\cdot)}{\partial p_L}} > 0 \tag{15}$$

$$\frac{\partial p_L^*(\cdot)}{\partial \alpha_1} = \frac{\frac{\partial S_B(\cdot)}{\partial \alpha_1} \cdot \frac{\partial D_L(\cdot)}{\partial p_B}}{\left(\frac{\partial S_B(\cdot)}{\partial p_B} - \frac{\partial D_B(\cdot)}{\partial p_B}\right) \cdot \frac{\partial D_L(\cdot)}{\partial p_L}} < 0 \tag{16}$$

In equation 15, an increase in an input's cost causes the marginal and average cost curves in the first panel of figure 1 to shift up, which causes the market supply function in the middle panel to shift up and to the left, thereby increasing the price of bar services. Regarding equation 16, an increase in an input's price increases firms' costs thereby causing profits, thus also license prices, to fall.[43]

Finally, the model produces testable predictions regarding the effects of changes in the quantity of licenses available under the quota system ($\bar{L}$). Differentiating equations (10) and (11) with respect to $\bar{L}$ and solving for the resulting impacts on the equilibrium prices yields:

$$\frac{\partial p_B^*(\cdot)}{\partial \bar{L}} = \frac{-\frac{\partial S_B(\cdot)}{\partial \bar{L}} \cdot \frac{\partial D_L(\cdot)}{\partial p_L}}{\left(\frac{\partial S_B(\cdot)}{\partial p_B} - \frac{\partial D_B(\cdot)}{\partial p_B}\right) \cdot \frac{\partial D_L(\cdot)}{\partial p_L}} < 0 \tag{17}$$

$$\frac{\partial p_L^*(\cdot)}{\partial \bar{L}} = \frac{\left(\frac{\partial S_B(\cdot)}{\partial p_B} - \frac{\partial D_B(\cdot)}{\partial p_B}\right) - \left(-\frac{\partial S_B(\cdot)}{\partial \bar{L}} \cdot \frac{\partial D_L(\cdot)}{\partial p_L}\right)}{\left(\frac{\partial S_B(\cdot)}{\partial p_B} - \frac{\partial D_B(\cdot)}{\partial p_B}\right) \cdot \frac{\partial D_L(\cdot)}{\partial p_L}} < 0 \tag{18}$$

Equations 17 and 18 indicate that an increase in the number of licenses causes the market

[43]Note that an increase in costs causes both firm and market level supply to decrease, which leads to increased bar services prices, which seemingly renders ambiguous the prediction for the effects of an increase in input prices on the market price of licenses. We dismiss this ambiguity by noting that if an increase in input prices caused profits to increase (after the ensuing rise in output price takes place), firms could always increase their profits by agreeing that they all would pay more than the market price for their inputs. We have heard of no accounts suggesting firms in this industry lobby to increase the list prices for, e.g., wholesale liquor or beer and wine or for mandates to increase wage rates of their workers.

price of both bar services and licenses to fall. Intuitively, for the former, an increase in the number of licenses (and firms) induces a rightward shift in the market supply of bar services in the second panel of figure 1, thereby reducing the price of bar services. Similarly, an increase in the number of licenses and the resulting reduction in bar service prices causes firms' profits (and license prices) to fall.

Although our model of Montana's liquor license program yields clear predictions related to the impacts of changes in several parameters on prices, not all the predictions are actually testable. There are, for example, no data collected on "bar service" prices, partly because bar services is an amorphous concept comprised of liquor, beer, and wine as well as ambiance, entertainment, etc., and is highly heterogeneous.[44] We also determined that undertaking a comprehensive field survey would be prohibitively costly. Accordingly, we are not able to empirically test the model's predictions regarding those prices. Similarly, while we have relatively good transfer price data for liquor licenses, we do not have annual data on any input prices at either the quota area or county level. Labor prices vary over time, but any differences between areas are likely relatively constant over time (so their effects will be absorbed in county fixed effects). The same is mostly likely true for costs such as construction and maintenance.

Regarding the impacts of changes in the number of licenses, there are limited circumstance under which licenses can move from one quota area to another. As described above, when an area's population grows, at established thresholds, a new license is allowed in the area. The additional license in an area typically is acquired ("floated in") from another area (where there are excess licenses and prices are lower). The final comparative statics prediction above suggests this transfer should cause a decrease in license prices in the area that floats in a license and an increase in price in the area from which

[44]Efforts to acquire data on alcohol prices by telephoning individual bars and restaurants and asking for prices of a few simple drinks were met with sufficient reticence (or worse) on the part of the individuals who answered our calls, that we quickly discontinued those efforts.

the floater license is acquired. Although there are a number of instances during the time span of our data where an additional license was allowed in an area, given the substantial noise in license transfer prices (see discussion below), the impacts of one more (or one fewer) license are likely not measurable.[45] That said, there are instances where we may be able to estimate the impacts of a change in the number of licenses. In 2007, the quota for RBW licenses was increased dramatically, doubling in some areas (we discuss this event in more detail in the empirical results section below). In addition, there has been recent growth in the number of manufacturers who sell directly to consumers (e.g., small breweries).

A final point to make regarding our comparative statics predictions is that a caveat to the license price results is that they only apply to cases where the quota for licenses is binding. If an area has a non-binding quota, the competitive, free-market number of firms operate in the market. In figure 1, this scenario would be represented in the third panel by a license quota in excess of $\bar{L}$ ($= k^*$), firms would make zero economic profits, and the license price would be \$0. In this situation, relatively small changes in demand or supply parameters in the bar services market, or license numbers, will have no impact on the price of licenses.

3.4 Data and Methodology

3.4.1 Data

The Montana Department of Revenue must approve all transfers of alcohol licenses. As part of the approval process, the department records data on all transfers, including the sale price, license type, buyer's quota area, transaction type, application date, and approval date. The information is public record, and is published on the department's

[45]For example, adding one additional license to a 98 all-beverage licenses in Billings constitutes a mere 1.02 percent increase. As a result, it is unlikely that we will be able to obtain precise/credible estimates of these effects from our transfer price data.

website for all transactions since 2007. Additionally, we also obtained earlier price data from 2002-2007 from the DOR.[46] The two data sets partially overlap for 2006-2007. Our analysis employs observations from 2005-2017; in total, there are 1,773 observed transactions from forming pooled cross-sections over time. Included in the data are 1,065 all-beverage, 352 beer retail, and 234 restaurant beer and wine transactions.[47] Both the earlier and later observations in the sample are incomplete, causing them to be dropped (i.e., observations before 2005 and after 2017). Similarly, non-market transactions (such as transfers of location and corporate structure changes) and those listing a price of zero (or other implausibly low values) are omitted.[48] Other observations have potential clerical errors, either between our pre- and post-2007 data sets or within one or the other data set. For each type of potential problem, the observation is flagged with a binary variable. The most commonly flagged issue—with 71 occurrences—is for a missing transaction type.[49] Further, all observations missing an application date fall into the "missing" transaction type category, with 53 of these observations in 2006.[50]

The supply of licenses in a quota area is perfectly inelastic for any given year, with the quantity based on the area's population, as laid out in title 16 of the MCA. The

[46] We obtained these data prior to the implementation by the DOR of a new computer system in 2007, at which point almost all data from earlier years were apparently destroyed.

[47] The infrequency of the other license types precludes analysis. Only a handful of transactions across all other license types involved trades between private parties.

[48] Transfers of location should list a price of zero. In the data, however, several transactions recorded as transfers of location list non-zero (and sometimes substantial) prices. To identify which transactions are mislabeled, all observations flagged as a transfer of location listing a price of over $800.00 are changed to "transfer of ownership and location." Several transfers of location appear to list the annual licensing fee as the price; $800 is the highest annual fee, and serves as our line of demarcation.

[49] Other errors we found include incomplete entries in the 2002-2007 data; inconsistent dates, prices, and/or license types between our two data sets for observations that appear in both; and unusually long times between reported application and approval dates. In all cases in which there are discrepancies between transactions that appear in both data sets, the newer value supersedes the older. A binary variable for each potential error was initially included in the regressions. The estimated coefficients on these variables were insignificant, however, and so these results are not reported or discussed further in the results section below.

[50] Because there is a concentration of observations with missing information, these observations cannot be treated as random. The missing application dates for these data are imputed using the average time between application and approval, and a dummy variable is included in preliminary regressions. In effect, "missing" is treated as a distinct transaction type.

DOR updates the license quotas annually, using the previous year's city-level population estimates from the U.S. Census Bureau. Quota levels and the number of licenses issued in each area are currently available from the DOR for all-beverage, beer retail, and restaurant beer and wine for 2008-2017, along with an existing report from 2005.[51] For 2006 and 2007, updated city-level population estimates from the U.S. Census Bureau are used to impute quotas.[52] Similarly, the number of licenses issued for 2006 and 2007 are interpolated using the 2005 and 2008 quota reports.[53]

In examining the determinants of license prices, several characteristics of each quota area are included. Median household incomes, reported at the county level, are obtained from the U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE). Unemployment, also reported at the county level, are obtained from the U.S. Bureau of Labor Statistics. Additionally, city-level population estimates and data on several demographic characteristics are obtained from the U.S. Census Bureau Intercensal Estimates.

Other important demand-side factors include university enrollment, tourism, and gambling revenue. The National Center for Education Statistics provides annual fall enrollments for all two and four year institutions in Montana through the Integrated Postsecondary Education Data System (IPEDS).[54] Data on the number of visitors and tourists to any particular community in a given year are not generally available. As a proxy for this factor, we use county-level tax revenues from the state's hotel and lodging tax. The supply of hotel rooms will likely be quite inelastic over any relatively short

[51] The 2005 quota spreadsheet was obtained before the update to the DOR's computer system in 2007 and is no longer available from the department.

[52] These imputed quotas are not perfect replicas of the original (unavailable) quota spreadsheets. The department uses initial population estimates, which the Census Bureau subsequently updates.

[53] The interpolated number of all-beverage and beer and wine licenses issued in 2006 is one-third the difference between 2008 and 2005, added to the number of licenses issued in 2005; 2007 values are assigned using the 2008 numbers less one-third the difference. Both values are rounded to the nearest whole number. For RBW licenses, the 2006 licenses issued take on the 2005 value, while the 2007 interpolated number is the average of 2005 and 2008.

[54] We do not distinguish among different types of universities in our analysis. Preliminary analysis indicated that our results are robust to separating enrollment by type of institution (e.g., private/public/tribal, or two-year/four-year).

period of time. As a result, differences in the number of tourists should be a primary determinant of the observed changes in tax revenue within a given county.[55] Additionally, as a measure of gambling revenues, we obtain data from the Montana Department of Justice on annual revenue collected from the 15-percent tax on net revenues from gambling machines (the money fed into the machines, less the amount paid out in winnings).[56]

Given that the data span more than a decade, we deflate prices to net out the effects of inflation over time. The Producer Price Index (PPI) for Beer, Wine, and Liquor is used to deflate license prices in 2018 dollars. Similarly, the Mountain-Plains CPI is used to deflate income, gambling revenue, and hotel revenue. Both indices are obtained from the Bureau of Labor Statistics.

3.4.2 Empirical Methodology

We estimate the empirical impacts of various factors on alcohol retail on-premise license prices with a multivariate linear regression. The basic econometric specification takes the form:

$$\begin{aligned} Price_{ijt} = \beta_0 + \beta_1 Gambling_{jt} + \beta_2 Gambling_{jt} \times Restricted_i + \beta_3 Hotel_{jt} + \beta_4 Pop_{jt} \\ + \beta_5 PctYoung_{jt} + \beta_6 Female_{jt} + \beta_7 Enroll_{jt} + \beta_8 Inc_{jt} + \beta_9 Unemp_{jt} \\ + \boldsymbol{LicenseType}_{ijt}\boldsymbol{\gamma} + \boldsymbol{TransactionType}_{ijt}\boldsymbol{\delta} + \boldsymbol{Quota}_{jt} + \varepsilon_{ijt} \end{aligned} \quad (19)$$

$Price_{ijt}$ is the natural log of the license price for transaction i, in quota area j, in year

[55]Several Montana counties have jointly reported revenue from this tax, including Carter/Golden Valley/Treasure, Prairie/Wibaux, Garfield/McCone, and Judith Basin/Liberty/Petroleum. We impute county-level revenue numbers for the few jointly reported counties as the total revenue divided by the number of hotels in each county. Because the jointly reported areas had very few liquor license transactions (all with very low prices), how we imputed these data points does not influence our results.

[56]The pre-2007 data are city level; the post-2007 data are for larger cities with the county aggregate reported for smaller locales. Post-2007 city-level values are imputed for the smaller cities, using the historical average split from the pre-2007 data. Again, the estimates are not sensitive to how these values are imputed.

t. $Gambling_{jt}$ is total gambling revenue in the relevant quota area and year, while $Restricted_i$ indicates whether a transaction involves a gaming-restricted license (i.e., for licenses that do not permit gambling). $Hotel_{jt}$ is the total county-level hotel revenue attributed to quota area j in year t.[57] Pop_{jt} is the estimated population, while $PctYoung_{jt}$ and $Female_{jt}$ are the proportion of the population between 20 and 44 and the proportion that is female, respectively. $Enroll_{jt}$ measures university/college enrollment (taking on a value of zero for most of the data), while Inc_{jt} and $Unemp_{jt}$ control for quota-area and time economic characteristics. $\boldsymbol{LicenseType}_{ijt}$ is a vector of indicator variables for floater, beer and wine, and RBW licenses.[58] Similarly, $\boldsymbol{TransactionType}_{ijt}$ is a vector of indicator variables for the transaction type: transfer of ownership, corporate structure change, or missing (transfer of ownership and location is the omitted category).[59]

The main controls we employ for supply-side variation is the vector $\boldsymbol{Quota_{jt}}$. For beer and wine licenses and RBW licenses, we include indicator variables for whether or not the quota is binding. For these licenses, "binding" is defined as at or above quota.[60] Unincorporated quota areas have neither beer retail nor RBW quotas and take on a value of zero for each indicator variable by default. For all-beverage licenses, we include two terms: percentage all-beverage quota and percentage all-beverage quota squared. Both the linear and quadratic terms are included for a more flexible functional form.[61] Note

[57]The county level hotel revenue estimate is used for every quota area in a county. For example, the 2005 hotel revenue value from Gallatin County is used for both the Bozeman and Three Forks quota areas in 2005.

[58]For the purposes of our analysis, the ten resort all-beverage licenses in the data are separated from the typical all-beverage licenses. Although issued under different circumstances, resort all-beverage license prices are not statistically distinguishable from all-beverage license prices. Whether or not we pool these license types does not affect results.

[59]The specifications do not include license/transaction type interactions, because there are not enough observations for every combination for the interaction term to be meaningful. The results, however, are robust to their inclusion.

[60]A relatively small number of communities are above their beer retail quota. Areas appear to be above quota for one of two reasons. First, several areas appear to have had historic decreases in population (e.g., Butte and Anaconda). Second, the U.S. Census Bureau may have over-estimated population in their initial intercensal estimates. If a population estimate is subsequently revised to a lower value, the relevant quota may decrease.

[61]If a quota area is well below or well above their established quota, prices are expected to be low;

that there is the potential for simultaneity bias, as population potentially drives both the supply of licenses as well as the demand. We argue, however, that the structure of our data limits this potential effect. For much of the state, there is very little variation in the quotas and number of licenses issues in our time frame.[62] Year on year, there is little change in most areas, allowing us to treat supply as exogenous.

Finally, we also include quota-area fixed effects in most specifications, to account for unobserved heterogeneity. Note that certain variables, such as local university enrollment, have relatively low within-area variation. While enrollment undoubtedly plays a role in determining license prices, including fixed effects will capture most of this cross-sectional variation (and thus we would expect a null effect with their inclusion). All standard errors are calculated clustering at the quota area level. Clustering by quota area allows for a flexible dependency structure spatially and temporally, as opposed to clustering by area-time, which would not account for potential autocorrelation. There are a relatively small number of transactions per quota area per year and during transaction negotiation, historical license prices from the quota area are almost certainly used in determining current-period price. Additionally, license prices within a quota area are far more likely to be correlated (over time and space) than at a broader geographic level (e.g., by county– for example, Bozeman's prices will not correlate highly with Three Fork's). Given these considerations, we choose to cluster at the quota area level.[63]

license values will tend to peak at roughly 1.43 times the quota (e.g., Bozeman floats a license in whenever they drop below the threshold). Unlike the other two license types, there is significant variation in all-beverage licenses issued compared to the quota. Our results are not sensitive to how we specify controls for supply of licenses.

[62] Exceptions to this are Bozeman (which has seen a relatively large increase in licenses) and Anaconda (which as seen a relatively large decrease in licenses). Indeed, it appears that when licenses float into Bozeman, they frequently come from Anaconda.

[63] The number of observations per quota-area, however, differs fairly substantially. To address this issue, we run specifications where broader ranges are clustered, e.g., multiple contiguous counties with relatively low populations (whereas larger communities, such as Billings, Bozeman, and Missoula, are treated as their own clusters). These alternate clusters do not affect our results. Similarly, when we focus exclusively on high-transaction areas, the results are qualitatively similar to using the full sample.

3.5 Results

3.5.1 Baseline Specifications

The baseline results are reported in table 1. For each column, regressions include all usable observations.[64]Column 1 omits the quota-area fixed effects, allowing for cross-sectional variation. Column 2 reports the same estimates including location fixed effects. Column 3 restricts the sample, dropping all observations for which the license prices are more than two standard deviations below the quota-area mean; it also drops RBW and resort licenses, quota areas with non-binding quotas, and transactions other than full license sales.[65] Column 4 further restricts the sample, including observations only for the 15 areas with large numbers of transactions and binding quotas.[66] All regressions presented here include the demographic controls and the controls for license supply. For ease of interpretation, the regression measures gambling revenue and hotel revenue in millions of dollars. Along the same lines, population, enrollment, and household median income are measured in thousands.

In the first column of table 1, the specification provides strong evidence that gambling revenue, hotel revenue (used as a proxy for tourism), college enrollment, and income are statistically and economically significant positive determinants of license price. The coefficient on gambling revenue, for example, indicates an $8,000 increase in license price for a million-dollar increase in annual gambling revenue. Interacted with a indicator variable

[64]The dummies for potential clerical errors are omitted throughout the section. The results across specifications are robust to the inclusion of these dummy variables, and are robust to the inclusion of the flagged observations, which occur exclusively in the pre-2008 data.

[65]There are a non-trivial number of license transactions that list prices that are extremely low (e.g., $10,000 in an area with an average price of $750,000). These transactions may be clerical errors, or may be indicative of unobserved characteristics (e.g., transactions between friends or relatives). Furthermore, Corporate Structure Changes (and missing transaction type) include transactions ranging from a near-complete buyout of partner (e.g., a 90% stakeholder selling his or her share of a business) to a restructuring for tax purposes (e.g., from a partnership to an LLC).

[66]These include Anaconda, Billings, Bozeman, Butte, Great Falls, Havre, Helena, Kalispell, Lewistown, Livingston, Laurel, Miles City, Missoula, Whitefish, and Wolf Point.

Table 1: Regression Results

VARIABLES	(1) Price	(2) Price	(3) Price	(4) Price
Gambling Revenue	8,273***	6,679***	11,351***	9,371**
	(819.4)	(1,929)	(3,893)	(3,940)
Gambling Rev.×Restricted	-4,960***	-5,021***	-471.6	-313.8
	(875.5)	(1,109)	(401.5)	(366.1)
Hotel Revenue	6,354**	9,022***	7,949**	7,041*
	(2,818)	(3,080)	(3,220)	(3,940)
Fall Enrollment	17,884***	-6,747	-7,884	-7,776
	(4,281)	(5,814)	(7,099)	(9,606)
Population	-2,904***	-885.3	-4,335*	-5,797*
	(610.7)	(1,378)	(2,312)	(2,724)
Real Median Income	6,306***	6,641***	7,616***	12,549***
	(1,410)	(1,503)	(2,251)	(3,950)
Unemployment	7,368***	2,423	-3,421	-5,758
	(2,203)	(3,710)	(5,738)	(10,828)
All-Bev. Restricted	-3,737	15,180	-277,496***	-306,350***
	(51,835)	(59,640)	(34,936)	(28,942)
All-Beverage Floater	-196,787***	-195,249**	-531,480***	-553,091***
	(67,682)	(86,892)	(38,685)	(28,683)
Beer & Wine	-110,939***	-124,702***	-161,748***	-198,632***
	(29,099)	(31,094)	(41,723)	(50,943)
Beer & Wine Restricted	-150,560***	-179,862***	-445,439***	-476,486***
	(41,452)	(56,329)	(41,829)	(38,459)
Resort	-81,085	-115,341		
	(68,800)	(110,595)		
RBW	-239,599***	-251,592***		
	(39,798)	(48,436)		
Transfer of Ownership	-20,152**	-12,323	-10,214	-7,767
	(8,968)	(9,511)	(10,719)	(12,873)
Corporate Structure Change	-9,214	-21,612*		
	(13,013)	(12,049)		
Missing Transaction Type	-50,007***	-49,083***		
	(13,582)	(12,328)		
Supply Controls	Yes	Yes	Yes	Yes
Demographic Controls	Yes	Yes	Yes	Yes
Fixed Effects	No	Yes	Yes	Yes
Observations	1,330	1,330	852	492
R-squared	0.658	0.735	0.826	0.785

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

for being gaming restricted, moreover, indicate that gambling revenue does not have a large impact on licenses that do not permit gambling.[67] These licenses are substitutes, however, so we would expect gambling revenue to have some a positive effect even on restricted licenses. Hotel revenue is somewhat less precisely estimated, though still of economic significance. This is not unexpected compared to gambling revenue, as tourism is a less direct channel (gambling revenue should be directly priced into license values). Similarly, estimates indicate an increase of 1,000 college students would increase prices by close to $18,000, while the estimated effect of an additional $1,000 in median income is a $6,000 increase.

The fact that population is estimated to have a negative effect is somewhat perplexing. Population, however, drives both the supply of licenses (as well as demand). That said, if we hold the percentage of quota constant, it is not clear whether an increase in population would have a positive or negative effect on licenses ex ante. The number of licenses increases with population, but the formula may over- or under-shoot the increase in demand. When the government imposed the system in 1947, it is unlikely that the formulas were well-calibrated to keep values stable.

Of additional interest are the coefficients assigned to the license type and transaction type indicator variables in column 1.[68] As expected, the results indicate that all-beverage licenses trade at a large premium over each of the the other license types. Although floater licenses allow the same privileges as all-beverage licenses (except for gambling on floaters issued after 2007), the results indicate that they are sold for nearly $200,000 less on average. Note that the prices on floaters reflect the market conditions from the license seller's area, not the buyer's. The state issues floaters one at a time, allocating them via lottery. Whoever wins the lottery seeks a lower-price market (which is eligible to float a

[67] Note that statistical tests for Gambling Revenue + Gambling Revenue × Restricted are not significant for column 2, but are significant for the other three columns.

[68] All-beverage and transfer of ownership and location are the omitted variables.

license out) to purchase a license. The buyer acts as a monopsonist—prices will tend to reflect the seller's market conditions.

The disparity between all-beverage and the other two license types (beer and wine retail and RBW licenses), on the other hand, reflects the differences in the privileges associated with each license type. Beer retail licensees are allowed to offer the same services as all-beverage licensees, with the notable exception of selling liquor. The discount for beer and wine licenses is estimated at $100,000 on average. RBW licensees may not sell for off-premises consumption, nor may they offer gambling under any circumstances. Additionally, the more stringent constraints (e.g., the revenue requirement for food sales) on RBW licenses should push prices down even further. The results provide evidence to this effect, estimating that RBW licenses sell for nearly $250,000 less than all-beverage licenses.

Given that state law allows gambling exclusively in establishments licensed to sell alcohol, gambling revenue appears to affect the price of licenses in the expected manner. Licenses that preclude gambling should sell for lower values. The coefficient on gaming restrictions measures the relative difference in price between licenses with and without gambling attached. In the first column, however, there does not appear to be a significant difference for all-beverage gaming restricted licenses, although the coefficients for beer/wine and beer/wine (restricted) are significantly different according to an F-test (p-value less than 0.001).

Column 2 adds fixed effects for quota areas. The estimates remain remarkably similar–both gambling tax revenue and hotel tax revenue are within a standard error of the pooled cross-section results, while the estimate on median income is virtually identical. A major exception to this is the estimate on fall enrollment, which becomes insignificant. This is not surprising, as there is relatively little variation in the size of universities year to year within a given quota area[69]. Furthermore, the estimates for the different license types

[69]Note that Montana State University (Bozeman) steadily grew during our time frame, while the

are extremely consistent with the addition of fixed effects, as are the negative coefficients on the different transaction types.

The similarity in the qualitative results indicates that there is probably not enough evidence to reject the idea that simultaneity bias is limited with respect to the variables other than population. Treating license supply as exogenous (with respect to everything but population) is consistent with the institutional details of the Montana licensing system. This is particularly true for the supply of all-beverage licenses, at least within any given year. For the other license types, exogeneity is somewhat less likely to hold, as beer and wine licenses are not always limited by a quota. Although simultaneity bias is not necessarily expected ex ante, it is still possible that the specification could suffer from bias if the supply of licenses is endogenous with respect to the other variables.

3.5.2 Subsamples

As robustness checks, columns 3 and 4 of table 1 report output for regression specifications with restricted samples. First, these data omit RBW and resort all-beverage licenses, as these are sold less and have far more restrictions than all-beverage and retail beer/wine licenses. Second, the data omit missing transactions and corporate structure changes, as these may include transactions other than full sales. Focusing on full sales may provide better estimates, as the effects of market forces may be attenuated using the other transaction types. In addition, observations missing the transaction type are also omitted, because there is no way to identify if they are full sales or not.

Third, the specifications drop additional low-priced transactions (beyond the $800 minimum used in the main specifications). While not as extreme as $1.00 sales, a price of $10,000 in a market that routinely assigns six-figure values to licenses raises concerns about the nature of such transactions. Relative to a more typical, higher-priced sale, the

University of Montana (Missoula) grew for the first half of the data, then started to slowly contract. This variation, however, is quite small relative to the cross-sectional variation accross quota areas.

low price probably reflects unobserved factors in the transaction (e.g., a clerical error). To address this issue, the restricted sample omits observations listing a license price more than two standard deviations below the quota-area/transaction type mean (i.e., the mean of license prices in 2018 dollars, taken over time).

Finally, we restrict the data to include only observations for areas that are near their all-beverage quotas.[70] If the license quota is non-binding in an area, the expected price should be quite close to zero.[71] If a quota area employs fewer licenses than permitted by law, marginal changes in demand factors should have little to no impact on license price.[72]. Column 4 further restricts the sample along these lines, focusing on the 15 areas with the highest number of observations.

The results in column 3 identify the same significant effects as the baseline specifications, although the magnitudes of several coefficients are considerably larger. This is expected, as the restricted sample focuses on areas with binding quotas—we would expect null effects in non-binding areas. In particular, the discounts for the various licenses are much larger. With the restricted sample, gambling rights (for both all-beverage and beer and wine licenses) are worth more than $250,000. Floater licenses, moreover, sell for more than $500,000 less in the restricted sample. The discount for a beer and wine license is over $150,000, a larger difference than in columns 1 and 2 (though smaller than the differences for other license types).

When we restrict the data to the 15 areas with the largest number of trades, the estimates are virtually identical, with the exception of median income (which has a

[70]For the purposes of the regression, "near" means either above or within one license of meeting the quota. The one license allowance is to account for administrative lag—when a quota is increased, it can take several months to fill the license.

[71]This is frequently borne out in the data, with prices in many non-binding areas mirroring the annual licensing fee. The license fees tend to be $400-$650, depending on population and license type. For a full listing of license fees, see M.C.A. 16-4-501.

[72]A caveat to the non-binding quotas is the licenses attached to businesses near incorporated areas. As communities expand, so too does the five-mile buffer around city limits. If a license is absorbed into the quota area, it potentially gains a great deal of value.

substantially larger estimate). Overall, it appears that much of the variation driving the results comes from the non-low price, full transfer transactions in binding quota areas. These results are possibly better estimates of marginal effects, given that the full sample includes observations where we would not expect to be able to detect effects (i.e., in areas that are below quota). Furthermore, it is unsurprising that the results are driven by the quota areas with the largest numbers of transactions.

3.5.3 Small Scale Producers

Another important development in the on-premises liquor license market is the rise of small-scale producers in Montana, including breweries, wineries, and distilleries. The state licenses producers of alcoholic beverages separately from on-premises retailers, with no quota system in effect for production licenses. Producers located within Montana, however, may sell their own products to consumers directly, subject to restrictions, without a retail on-premises license. Despite quantity and hours-of-operation restrictions, small breweries in particular have become a popular alternative to the traditional on-premises establishments. Given the high cost and limited number of retail licenses, the rise of microbreweries has caused political push-back from on-premises licensees.[73]

The state of Montana granted brewers the right to serve their own beer on-premises in 1999, and since then the industry has experienced rapid growth. As of 2011, the Brewers Association recognized 33 breweries in Montana, which grew to 68 by 2016.[74] In 2017, moreover, the Montana Department of Revenue listed 80 active brewery licenses, in addition to 21 distilleries and 20 wineries.[75] Given that these licensees may provide services that are close substitutes for those of the retail on-premises licensees, the producers may have a negative effect on the retail license prices.

[73]See "A Montana Loophole Leaves a Bitter Taste With Bar Owners," New York Times, April 20th, 2013.

[74]See `https://www.brewersassociation.org/statistics/by-state/`.

[75]See `https://tap.dor.mt.gov`.

Table 2: Regression Results - Small Scale Producers

VARIABLES	(1) Price	(2) Price	(3) Price
Gambling Revenue	8,417***	6,401***	6,547***
	(763.6)	(2,123)	(2,097)
Gambling Rev.×Restricted	-4,965***	-5,098***	-4,623***
	(879.3)	(1,137)	(1,054)
Hotel Revenue	5,575*	13,053***	12,834***
	(3,308)	(3,932)	(4,116)
Fall Enrollment	16,679***	-1,558	-243.8
	(3,764)	(5,268)	(5,497)
Population	-3,117***	4,391**	4,203**
	(571.9)	(2,065)	(2,086)
Real Median Income	6,356***	6,155***	5,943***
	(1,397)	(1,384)	(1,386)
Unemployment	7,902***	333.9	-45.52
	(2,438)	(3,664)	(3,753)
All-Bev. Restricted	-9,750	23,688	-12,075
	(52,056)	(61,028)	(54,564)
All-Beverage Floater	-196,831***	-194,820**	-215,878**
	(67,443)	(86,918)	(85,299)
Beer & Wine	-111,425***	-124,868***	-125,545***
	(29,407)	(31,247)	(31,354)
Beer & Wine Restricted	-152,159***	-176,213***	-196,214***
	(42,446)	(55,292)	(60,224)
Resort	-76,908	-118,203	-120,497
	(69,890)	(110,476)	(108,687)
RBW	-239,685***	-252,201***	-225,516***
	(39,567)	(48,711)	(49,918)
Small Scale Producers	2,479	-9,428***	-7,514***
	(2,159)	(2,642)	(2,393)
RBW×Producers			-8,408**
			(4,131)
Transaction Type	Yes	Yes	Yes
Supply Controls	Yes	Yes	Yes
Demographic Controls	Yes	Yes	Yes
Fixed Effects	No	Yes	Yes
Observations	1,330	1,330	1,330
R-squared	0.659	0.736	0.738

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

We run a second set of regressions looking at the effect of the increase in small-scale producers on values of licenses. Table 2 presents these results. Columns 1 and 2 of table 2 replicate the baseline results from table 1, with the addition of a control for the number of small producers (i.e., the sum of breweriers, distilleries, and wineries licenses that permit on-premises consumption). In column 1, which does not include quota-area fixed effects, the effect of small producers appears to be negligible. Note, however, that this value is almost certainly biased upwards due to selection effects. Breweries and other producers will tend to locate in communities with high license values, such as Bozeman and Missoula. Futhermore, many quota areas in Montana do not have any producers.

In column 2, when we include quota-area fixed effects, this coefficient becomes negative and significant. By restricting variation to within license markets, we can mitigate some of the selection bias. In this specification, each additional producer appears to lower license values by roughly $9,000. Note that this is probably a lower bound on the effect—again, there is a selection effect here. For example, we would expect to see craft breweries open in locations where there is high demand relative to the number of existing establishments, i.e., where license values are quite high. That we find a negative effect, despite a bias in the other direction, provides strong evidence that an increase in small producers has a significant economic effect on the values of licenses.

Producers (particularly breweries and distilleries), however, are not perfect substitutes for the bars and restaurants that hold all-beverage, beer and wine, or restaurant beer and wine licenses. In particular, producers face stricter quantity and hours-of-operation restrictions than the on-premises licensees. Specifically, as of 2017, breweries producing fewer than 10,000 barrels could sell up to 48 ounces of beer to an individual between the hours of 10:00 AM and 8:00 PM; distilleries could sell up to 2 ounces to be consumed between 10:00 AM and 8:00 PM; and wineries could sell without quantity or time restrictions.[76]

[76]M.C.A. 16-3-213; M.C.A. 16-4-312; M.C.A. 16-3-411

Given these requirements, producers probably are closer to being perfect substitutes for restaurant beer and wine licensees than for the other license types. Breweries (or other producers) may serve food, but more importantly, restaurant beer and wine licensees must also abide by time regulations: licensees must stop serving alcohol by 11:00 P.M. [77] In addition, gambling is always prohibited for small-scale producers, though they may sell for off-premise consumption (e.g., in the form of "growlers" of beer or bottles of wine). As such, the growth in producers may have a larger impact on the prices of RBW licenses than on other license types. This prediction is supported by the results in column 3, where we add an interaction term between the indicator for RBW licenses and the number of small scale producers. In this alternate specification, the average estimate is somewhat smaller than in column 2 (of roughly \$7,500); for RBW licenses, the point estimate is closer to \$8,500 (note that these are both within a standard error of the estimate in column 2). Furthermore, the estimates in column 3 (for all-beverage/retail beer licenses versus RBW licenses) are statistically distinct (with an F-test p-value of 0.002).

3.6 Conclusion and Discussion

In this paper, we provide a brief history of the Montana liquor license system, discuss relevant details regarding its complexities, develop a theoretical model of the system, and then use data on license transfer prices to test the model's predictions. The empirical analysis focuses on two broad categories of the determinants of license values. First, we estimate the impacts of several factors that cause shifts in the demand for bar services, which in turn cause changes in firms' profits that translate into changes in liquor license prices. Second, we estimate the price impacts of differences in the attributes of, and restrictions associated with, the three most prominent types of licenses—all-beverage, beer and wine, and RBW. We also estimate the increase in license values attributable to

[77] See ARM 42.13.1103

the attachment of gambling rights, a factor that is unique to only a subset of the license types in the state of Montana.

Our empirical results generally provide strong support for the predictions of the model—that is, markets for a valuable government-created asset work well and perform as economic theory predicts. Demand factors that have significant impacts on license prices include median income and revenues from tourism and gambling. In addition, our estimates of differences in prices among license types are substantial and robust, and the impacts align with those suggested by our theory—all-beverage licenses sell for significantly more than city beer and wine licenses, which in turn sell for substantially more than RBW licenses. Moreover, all-beverage and beer and wine licenses with gambling rights attached sell for significantly more than without the rights. In addition to being statistically significant, these differences are economically significant. Possibly the most important innovation in Montana's retail on-premises markets to date in the 21st century is the expanding role of micro-breweries, distilleries, and wineries, which we combine into a category we label "small producers." Our empirical estimates suggest that an additional small producer reduces all-beverage and city beer and wine license values by about $7,500, and that the impact on an RBW license is $8,500. In Bozeman, Billings, and Missoula, each of which has seen a substantial number of small producers open up operations in the past decade, the total impact of these small producers on all-beverage and city beer and wine licenses is in the neighborhood of $100,000.

Although liquor license have sold for as much as $1 million, given the manner in which licenses were allocated in the 1930s and 1940s, they likely had little value at that time. The earliest mention of license prices that we have found is from late 1964 when they were reported to be selling for $20,000 (which is equivalent to about $126,000 in today's dollars). The political economy of the liquor license program is a story of rent seeking and rent protection in which the Montana Tavern Association (MTA) has proven to be one of the most effective lobbying groups in the state. Statements by legislators and by the media objecting to the liquor license system and arguing for its elimination can be found before virtually every legislative session in the last five decades. At no time, however, has the fundamental structure of the quota-based liquor license system been seriously threatened. Time and again the MTA successfully lobbied against retail on-premises regulatory changes that would have decreased license values and in favor of changes that increased license values. The right to offer gambling activities was first linked to the ownership of liquor licenses in 1976. In 1991, after several amendments, each establishment with gambling rights was allowed to have up to 20 gaming machines, a limit that persists to the present. A number of industry observers in the 1990s commented that the linking of gambling to all-beverage and beer and wine licenses substantially increased the value of those licenses. Moreover, the lobbying efforts to protect the rents reflected in license values became even more effective when the MTA's interests were aligned with those of the gambling industry.

Even though the MTA has been a formidable lobbying presence, there have been two notable instances in which the liquor license lobby acquiesced to pressure for change, the first such instance occurred in 1997 when the MTA supported the creation of a new type of alcohol license (the RBW license) that had the potential to negatively impact the value of all-beverage and city beer and wine licenses. Restaurant owners who did not have liquor licenses argued for the creation of a license that was more affordable than licenses with gambling rights attached[78]. The MTA successfully opposed a bill to create RBW licenses in 1993 and again at the beginning of the 1997 session. When the threat of letting the creation of RBW license be determined through a voter initiative was raised, however, the MTA altered its position, although not without extracting a substantial quid pro quo from RBW supporters—the initial version of the 1997 bill was

[78]Between 1993 and 1997, we found reference to liquor license values as high as $300,000. See, for example, Chronicle Editorial Board (1997).

altered in numerous ways that substantially decreased the value of the rights associated with RBW licenses. In fact, a year after 283 RBW licenses were made available, it was reported that "...a complicated application process and even more troublesome operating rules appear to have taken the fizz out of the opportunity" (Chaney, 1998). At that time, only 33 restaurants had received an RBW license and only 58 others reported still being interested in applying.

The policy change that created RBW licenses was outside the time span of our data set, so we are not able to reliably estimate its impact on all-beverage and city beer and wine license values. When a movement emerged in the mid-2000s to essentially double the number of RBW licenses, however, the MTA supported the bill, commenting that its members did not believe the creation of RBW licenses in 1997 had diminished their license values. In fact, today, RBW licenses in a relatively small number of markets have considerable value, and the lobbying interests for holders of these licenses generally align with the MTA to resist changes that might diminish license values (or support changes that increase their values).

The second instance in which the MTA succumbed to pressure from a number of sources and supported legislation that it initially opposed was in 1999 when House Bill 442 passed, and breweries were granted the right to sell beer directly to retail customers. The amount of beer that could be sold to an individual customer was 48 ounces and the breweries were constrained to stop beer sales at 8:00 p.m. The view expressed at the time by MTA lobbyist Mark Staples was that bar owners realized that "breweries don't want to compete with them—the bars that buy their beer" (Curliss, 1999). In response to this seemingly-highly constrained opportunity to access the on-premises retail alcohol market, since 1999, the number of breweries in Montana has roughly quintupled—from 20 breweries in 1999 to over 100 in 2020. Moreover, these breweries, many of which have substantial food menus to complement their beer offerings have become a major presence in retail on-premises food service markets in several of the larger cities in Montana.[79] Although the breweries, wineries, and distilleries have established a firm foothold in the Montana landscape, their efforts to expand their operations further by, e.g., extending their operating hours from 8:00 p.m. to, say, 10:00 p.m. have repeatedly met with rigorous opposition from the MTA and have not yet come close to garnering approval from the legislature.

Observers continually express outrage at the rents that have been created under the Montana liquor license system, pointing to the ever-increasing value of licenses as evidence of the system's absurdity. Calls to dismantle the program are made at the start of nearly every biennial legislative session. Although there has been no instance to date in which the program came close to being terminated, it is useful to contemplate the economic impacts of a few potential changes to Montana's quota-based liquor license regulatory program.

Possibly the most peculiar aspect of the Montana liquor license system over the time span of most of our data was the fact that the state allocated new licenses, which were often worth hundreds of thousands of dollars, via a lottery to a lucky individual who may or may not have any inkling about operating a bar or restaurant. The lottery appears to have been implemented in the early 1980s when license values were relatively low, and we speculate that a populist justification for using a lottery was to prevent wealthy individuals from being the only ones who obtained the newly issued licenses. From a political economy perspective, there was no obvious interest group supporting the lottery. Given the large number of applicants to many of the lotteries for all-beverage and retail beer licenses, the expected gains to each applicant was probably not large enough to warrant lobbying to continue the system. Neither would existing bar and restaurant

[79]The contentious relationship between Montana's small breweries and its bars and restaurants with the three primary types of liquor licenses attracted the attention of the New York Times several years ago. See Healy (2013) for a discussion of that conflict and also a description of the form into which brewery "taprooms" quickly evolved.

owners looking for liquor licenses be strong supporters of the lottery because if they did not win the lottery, they would have to compete with others in similar situations to convince the lucky lottery winner to enter a concession agreement with them. In 2017, as a means of increasing revenues in a tight budget period, the Governor's office proposed changing the lottery system to an auction process in which the state government obtained the value of new licenses from the winners of the auction. Not surprisingly in our minds, there was little opposition to this change.

Consider next, the most extreme modification—the simple and immediate elimination of Montana's quota-based liquor license system. Arguments in favor of such a change might be that individuals (bar owners) would not benefit from artificial limits on entry, and that community development would be enhanced by removing an important entry barrier for restaurants contemplating opening locations in Montana. A primary counter-argument would be that eliminating the program would render hundreds of millions of dollars in asset values worthless.[80]To put such a loss in perspective, it is important to realize that most current owners of beer and wine and all-beverage licenses purchased them sometime between the 1940s and today. And those owners who purchased licenses in the larger Montana communities in recent years, paid a high price for those licenses, based on expectations that the "rules of the game" in the retail on-premises liquor industry would maintain their status quo. From a political perspective, when the magnitude of asset losses to hundreds of individuals is made known, important equity considerations would become apparent. Moreover, given the influence of the MTA, the simple elimination of the liquor license program seems highly unlikely. Should it happen, however, the state would be surely tied up in costly litigation for years.[81]

[80]In 2007, MTA lobbyist Mark Staples, in a statement to a state legislative committee indicated "there are now, by my calculations, some two billion dollars' worth of investments in the licensing system."

[81]Powell (2015) examines the legal basis for takings claims in the context of peanut quota, taxi medallions, and liquor licenses. She argues that quota-based programs in these contexts are diverse and complicated and concludes that in response to regulatory changes allowing new licenses, there is "little legal support for a takings claim relating to the value of the current licensee's license on a secondary

A suggestion that involves a more gradual phasing out of the liquor license program is to gradually increase the number of liquor licenses issued. So, suppose the number of liquor licenses is increased by, say, 10 percent per year until the long-run competitive number of firms is operating, and the value of the licenses falls to zero. In this proposal, current license owners would continue to earn positive profits for a period of time before their licenses lose their value, thereby (it might be argued) mitigating the negative impacts on license owners. To think about the impacts of such a phase-out on current license owners, consider the following scenario. Suppose the current value of a license is $1 million. At a 5 percent discount rate this license value reflects annual economic profits of $50,000. As additional licenses are issued, prices of bar services fall, as do profits. Suppose, for the sake of concreteness, that over a span of ten years, annual profits are reduced by $5,000 per year, falling to $0 in the tenth year. So, immediately, the value of the licenses fall from $1 million (the present value of $50,000 in profits annually in perpetuity) to the present value of $45,000 in year 1, plus $40,000 in year 2, and so forth. The present value of the expected future flow of profits under the phase-out program is $189,218. In other words, asset values would fall immediately by 81 percent upon (or perhaps before) the adoption of the suggested phase-out. After conducting similar calculations, the MTA would surely object to this or similar phase-out proposals.

A third possibility is to terminate the Montana liquor license system in a manner similar to that used in the past couple decades to bring two major federal agricultural programs to an end. Until fairly recently, under both the U.S. peanut and the tobacco programs growers owned poundage quota that provided them with the right to sell a limited amount of the crop for the program.[82] Both those programs ran into political problems (associated with unacceptably large taxpayer costs) and, in the early 2000s,

market."

[82]See Rucker and Thurman (1990) for a discussion of the details of the U.S. peanut program, and Rucker, Thurman, and Sumner (1995) and Brown, Rucker, and Thurman (2007) for discussions of the tobacco program and various impacts of its elimination in 2004.

were eliminated with the approval of growers and quota owners. The mechanism by which these programs were eliminated was to buy out the federal programs' peanut and tobacco poundage quota. The peanut buyout cost taxpayers about $1.3 billion. The tobacco buyout resulted in payments to quota owners and tobacco growers of about $10 billion, the costs of which were borne by cigarette companies.

An important feature of these buyouts was that it appears quota owners were compensated for their quota at rates of roughly 30 to 40 percent above the market value of quota (prior to the point in time at which buyout discussions were initiated). Support among peanut and tobacco quota owners for these buyouts was strong. At first blush, compensation in excess of quota market values seems excessive, but (as illustrated in figure 1) those market values represent the value of program rents to the marginal firm/farmer/bar service provider. The present value of lost profits to inframarginal operators may therefore be considerably greater than the market values of the quota.

Suppose a similar buyout program was offered to Montana's liquor license owners. Assuming the MTA's lobbyist's estimates of investments in license values of $2 billion is reasonably accurate, then with, say, a 30 percent premium to garner support from inframarginal operators, a liquor license buyout would cost about $2.6 billion. Although license owners may be willing to accept buyout payments of that amount, it seems unlikely that Montana's taxpayers would be willing to fund such a plan.

A less disruptive suggestion for modifying the Montana liquor license system might be made based on a desire to increase the economic efficiency with which the program operates. Briefly, the observation that license values vary dramatically between quota areas suggests that there are social gains to allowing quota to be transferred across quota area boundaries. Such a change, however, would almost certainly result in numerous small Montana communities losing all their liquor licenses—a change that would almost certainly hasten the ongoing demise of Montana's rural communities. A modified approach might be to allow license transfers among Montana's largest, say, 10 or 15 quota areas. License buyers and sellers would both gain from this modification, and licenses would be distributed "efficiently" across the quota areas.[83] Without some insights into the elasticity of the demand for licenses, however, it is not possible to make more specific statements about how licenses would be reallocated. It is noteworthy that the current liquor license system does allow limited transfer among quota areas and, that it also has provisions that assure no community will lose all its licenses. The provisions of the system, however, cause these adjustments to be agonizingly slow.[84]

Based on the preceding, a reasonable assessment of the likelihood of substantial changes being made to correct perceived problems with the Montana liquor license system is that the political economy of the program makes changes unlikely. To the degree that policy makers pursue institutional reforms, however, or simply re-evaluate the liquor license program, we hope that our analysis can provide useful information to assist in policy discussions. In addition, our work can inform policy discussions surrounding the unintended consequences of licensure systems more broadly, as governments seek to adapt or create new regulatory systems.

[83]See Rucker, Thurman, and Sumner (1995) for an analysis of similar issues in the context of the federal tobacco program, in which quota could be transferred within counties, but not across county lines. Their analysis found that the relaxation of transfer restrictions would result in considerable reallocation of quota, but more modest reductions in deadweight losses. They also pointed out that although unrestricted transfer would have resulted in unambiguous net social gains, there would also be losers who would (and, in fact, did) object to such transfer.

[84]A sense for the glacial rate of adjustment in the Montana system is gained from observing that in 2005 (the earliest year for which we have detailed data on license allocations in individual quota areas), the Butte/Silverbow County quota area had 55 licenses available to be floated out, and in 2018 (the most recent year for which we have detailed allocation data), it had 46 licenses available.

Appendix A

Supplementary Materials

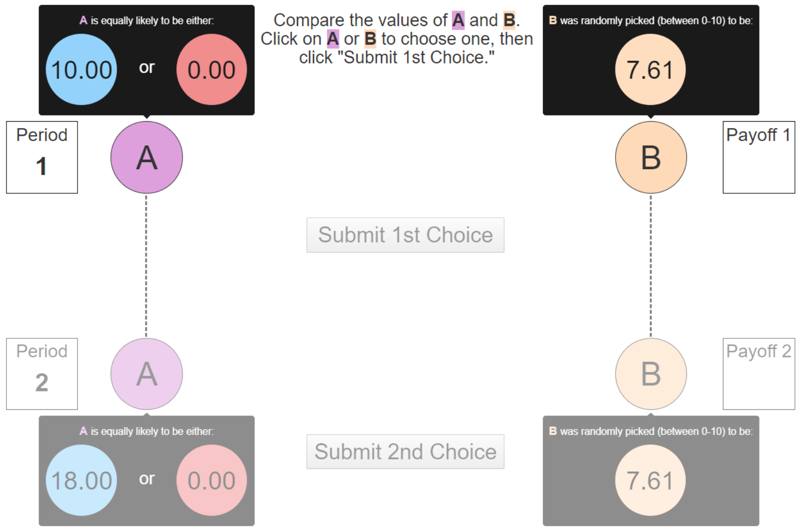
A.1 Preferences over Exploration: Simple Bandits in the Lab

A.1.1 Order Effects

In this appendix, we go through the specifics of the design for the $REVERSE$ treatment and offer comparisons between it and the $STRONG$ treatment. Because much of the key results rely on within-subject variation, order effects are a natural concern. In particular, subjects may anchor on their initial thresholds, i.e., a subject is satisfied with her CE in the lottery task, so uses a similar value for her individual SGI in the bandit. If this is the case, our data would tend to produce CEs and individual SGIs that are closer together in the within-subject measures than they would be if we had designed the two tasks as between-subject. If subjects are anchoring, our data would tend to overestimate the degree of under-exploration.

To understand the impact of order effects, additional subjects participated in the $REVERSE$ treatment. The payoffs and parameters are the same as the $STRONG$ treatment, but subjects play the BDM-bandit prior to the BDM-lottery. In the $REVERSE$ stage 1, they play a binary two-period bandit (as opposed to the binary lottery in $WEAK$

Figure 1: Stage 1 Interface: $REVERSE$



and $STRONG$) to familiarize themselves with the software and environment. Figure 1 presents the $REVERSE$ stage 1 interface. Note that this is not a perfect control, in that subjects will have more experience with bandits in the $REVERSE$ treatment than they do in the $STRONG$ treatment. Furthermore, the progression of the interface is somewhat less intuitive. In the initial treatments, the interfaces grow progressively more complicated; in $REVERSE$, the problem becomes simpler in the last stage. Given the nature of our design, however, this is unavoidable.

The stage 2 and stage 3 $REVERSE$ problems and interfaces are identical to $STRONG$, but subjects see them in a reverse order (BDM-bandit then BDM-lottery). Otherwise, the instructions and protocol for the $REVERSE$ treatment are as close as possible to the $WEAK$ and $STRONG$ treatments. Sessions were conducted during 2019 at the UCSB EBEL lab. Subjects earned a \$5.00 show up fee and were paid for a single randomly selected round (and payments were the sum of earnings for both periods). 61 subjects participated in the $REVERSE$ treatment and earned an average of \$14.94. The minimum payment earned for a session was \$5.00; the maximum earned was \$33.00. Sessions lasted less than one hour.

Table 1: Classification by Type

Treatment		Lottery-to-Bandit Thresholds: Decreaser	No Change	Increaser	Total
REVERSE	Gittins Inconsistent	9	2	8	19
		(14.8%)	(3.3%)	(13.1%)	(31.2%)
	Gittins Consistent	7	3	32	42
		(11.5%)	(4.9%)	(52.4%)	(68.8%)
	Total	16	5	40	61
		(26.2%)	(8.2%)	(65.6%)	(100%)
STRONG	Gittins Inconsistent	8	5	8	21
		(12.1%)	(7.6%)	(12.1%)	(31.8%)
	Gittins Consistent	5	9	31	45
		(7.6%)	(13.6%)	(47%)	(68.2%)
	Total	13	14	39	66
		(19.7%)	(21.2%)	(59.1)%	(100%)

This table presents the distribution of subjects in each classification. Subjects who make a 2nd-period mistake more than once after period 3 are classified as "Gittins Inconsistent," and "Gittins Consistent" otherwise. Subjects who set thresholds such that $|\bar{w}_{bi} - \bar{w}_{li}| \leq 0.10$ are considered "No Change". If the difference is greater than 0.10, a subject is labeled as a "Decreaser" or "Increaser" based on the sign of the difference.

Table 1 presents the taxonomy of types for $REVERSE$ and $STRONG$ treatments. Nearly 70% of subjects are Gittins Consistent in $REVERSE$ (which is extremely similar to both of the main treatments). The breakdown across how subjects adjust individual SGIs to CEs is also quite consistent with $STRONG$. A strict majority of subjects are both consistent and increase their valuations in $REVERSE$, as marginally more subjects set higher SGIs than CE. The taxonomy, however, is remarkably consistent with the $STRONG$ treatment. Furthermore, subjects in $REVERSE$ account for the risk-adjusted exploration value at rates of 20.2% (full sample), 30.2% (Gittins Consistent)

Table 2: Summary Statistics

Treatment	Subset	$\bar{w}_l$	Med. w_l	$\bar{w}_b$	Med. w_b	$\bar{w}_b - \bar{w}_l$
REVERSE	Full Sample	4.88	5.00	5.76	5.73	0.87
		(0.196)		(0.245)		(0.242)
	Gittins Consistent	5.09	5.00	6.40	6.35	1.31
		(0.228)		(0.268)		(0.300)
	GC-Increaser	4.80	5.00	6.81	6.89	2.01
		(0.218)		(0.284)		(0.277)
STRONG	Full Sample	4.24	4.50	5.01	5.00	0.77
		(0.169)		(0.212)		(0.183)
	Gittins Consistent	4.37	4.76	5.46	5.50	1.09
		(0.212)		(0.254)		(0.240)
	GC-Increaser	4.24	4.51	5.91	6.68	1.65
		(0.237)		(0.285)		(0.296)

This table presents the breakdown of subjects based on second-period decisions and individual changes in threshold values between lottery and bandit tasks. Subjects who more than once play safe-then-risky or stay on a not good arm are classified as "2nd Period Inconsistent," and "Gittins Consistent" otherwise. Subjects who set thresholds such that $|\bar{w}_{bi} - \bar{w}_{li}| \leq 0.10$ are considered "No Change". If the difference is greater than 0.10, a subject is labeled as a "Decreaser" or "Increaser" based on the sign of the difference. Standard errors are in parentheses below each estimate.

and 46.9% (consistent increasers). Thus, subjects in the $REVERSE$ treatment also systematically set lower individual SGIs than their risk preferences would predict.

Table 2 presents the average CEs and SGIS for the $WEAK$ $STRONG$ and $REVERSE$ treatments. The order of tasks does appear to influence the thresholds that subject set, at least to some degree: subjects who played in the $REVERSE$ treatment (and engaged in the bandit task first) set higher average thresholds. In $REVERSE$, the average lottery threshold is \$4.89, statistically higher than the \$4.24 set in $STRONG$ (p-value of 0.014 based on a two-sample t-test). Similarly, the average bandit valuations in $REVERSE$

and $STRONG$ are a statistically different \$5.76 and \$5.01 (p-value of 0.023). The results are similar for the Gittins consistent subjects, with $REVERSE$ subjects setting slighly risk-seeking average CE of \$5.09, which is statistically greater than \$4.37 in strong (p-value of 0.023); similarly, the average SGIs are different at \$6.40 and \$5.46 (p-value of 0.13). The treatment effects, however, are *not* statistically different between $REVERSE$ and $STRONG$ in the full sample or among consistent subjects (p-values of 0.737 and 0.569, respectively). While there is evidence of anchoring in levels, the treatment effect is consistent

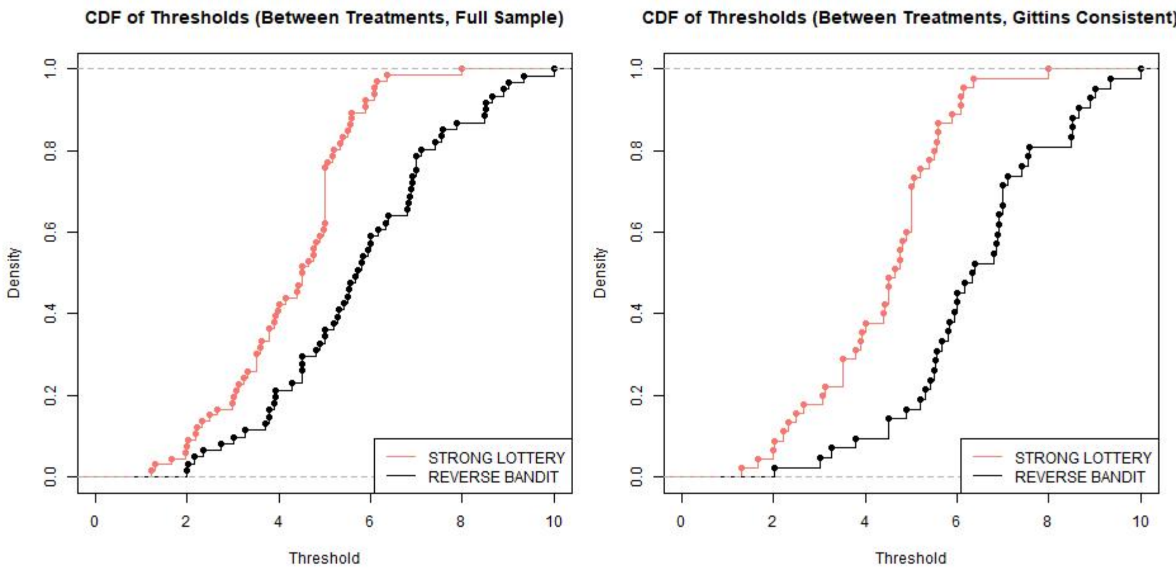
Furthermore, we can use the between comparison in the $REVERSE$ and $STRONG$ treatments to re-evaluate the treatment effect by comparing the the bandit valuations from the former and the lottery valuations from the latter (the first BDM-measures we elicit in each treatment). This will have the tendency to (1) increase the treatment effect between CEs and SGIs, and (2) decrease the discrepancy in how much subjects value exploration. Indeed, the average changes are \$1.52 for the full sample and \$2.03 for consistent subjects–nearly double the initial estimates from $STRONG$.

We can also break out th data to replicate parts of figure 6 which plotted CDFs of CEs and SGIs. Figures 2a and 2b replicate figures 6b and 6d, respectively. In both plots the treatment effect in thresholds is more pronounced–the between-effect indicates a larger premium placed on exploration than the within-estimates. For the full sample and the Gittins Consistent sub-sample, the CDF of CEs clearly stochastically dominates the CDF of SGIs. This adds to the evidence that subjects do recognize that exploration is valuable. The between-subject CEs and SGIs display a larger increase than the within-estimate.

The average differences of \$1.52 and \$2.03, however, are still well short of the risk-neutral prediction of \$4.33. While we cannot reproduce the individual plot akin to figure 8 (which relied on within-subject estimates), we can replicate the percentage of exploration value from the last column of table 4 using between estimates. That is, we

Figure 2: Comparing CEs from $STRONG$ and SGIs from $REVERSE$

(a) Thresholds, Comparing 1st Stages (b) CDF CRRAs, Comparing 1st Stages



can take the imputed SGI using the average CE from $STRONG$, and compare to the actual average SGI from $REVERSE$. When we do so, we find that the full sample takes into account 36.4% of the risk-adjusted exploration value; the Gittins consistent subjects take into account 48.4%; the the consistent increasers come in at 61.7%. Thus, even in the between estimates, we have substantial under-weighting (on average) of the value of exploration.

A.1.2 Instructions to Subjects

In this appendix we present the instructions to subjects. These were deployed via PowerPoint; each numbered point appeared on the screen as instructions were read aloud. We present the instructions from the $WEAK$ and $STRONG$ treatments. Instructions for $REVERSE$ were extremely similar.

Introduction

1. You are about to participate in an experiment in the economics of decision-making. If you follow these instructions carefully and make good decisions, you can earn a **CONSIDERABLE AMOUNT OF MONEY**, which will be **PAID TO YOU IN CASH** at the end of the experiment.

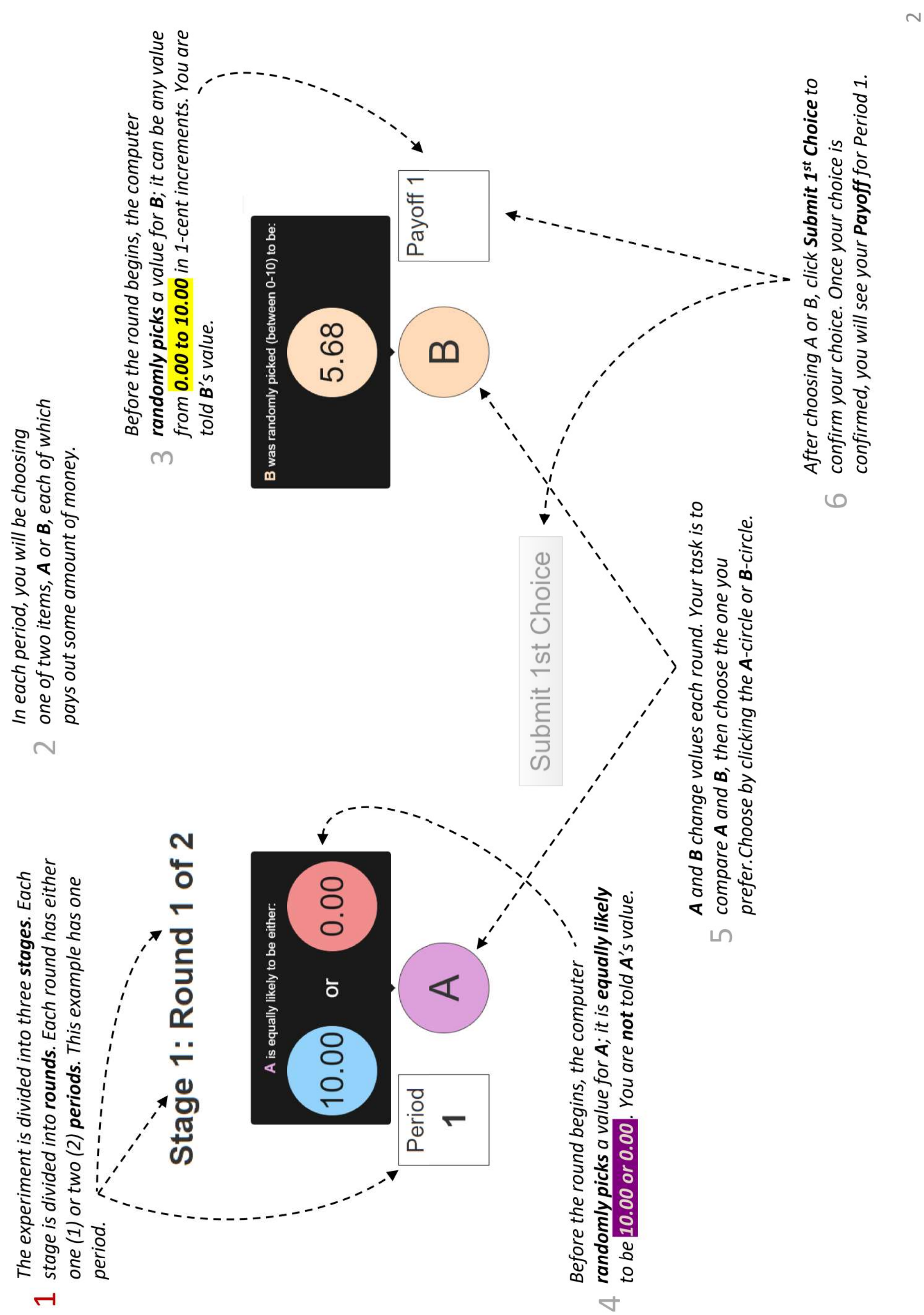
2. Your computer screen will display useful information. Remember that the information on your computer screen is **PRIVATE**. To ensure the best results for yourself and accurate data for the experimenters, please **DO NOT COMMUNICATE** with the other participants at any point during the experiment. If you have any questions, or need assistance of any kind, please **RAISE YOUR HAND** and the experimenter will come to you.

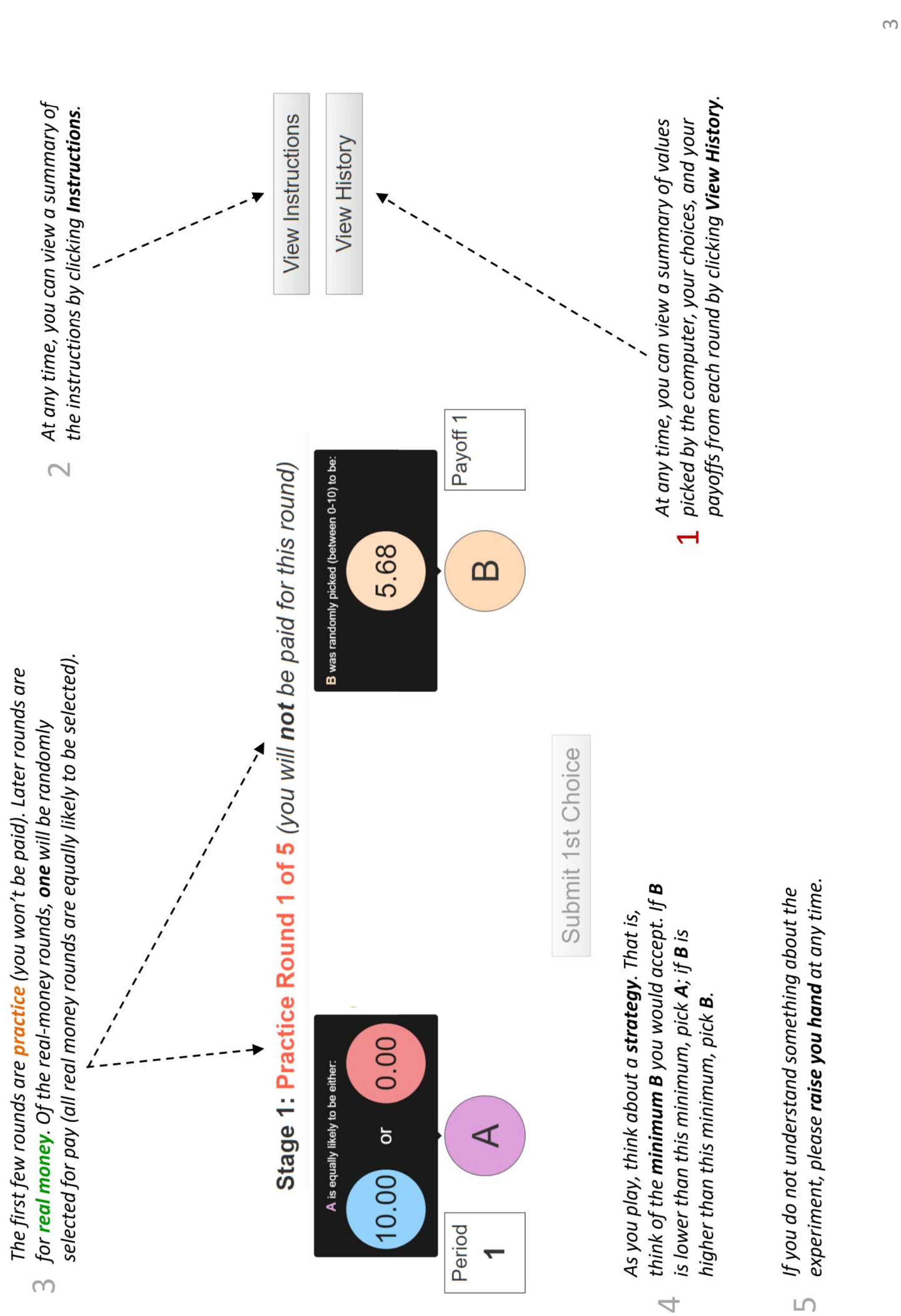
3. Please **TURN OFF YOUR CELLPHONE** and avoid opening any other browsers or programs on your computer. You are being paid; please give your full attention to ensure the best results for yourself.

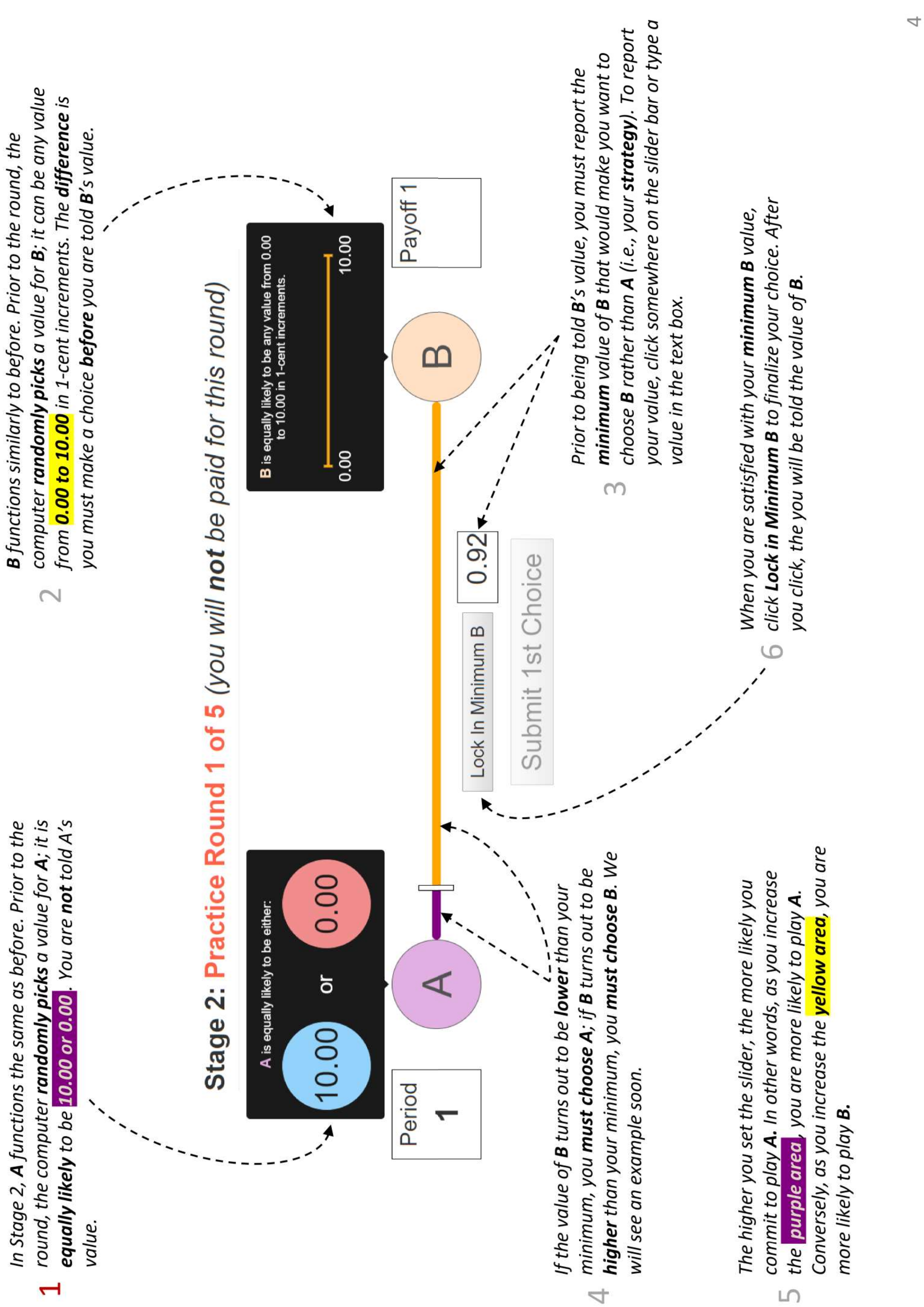
4. Economics experiments have a strict **POLICY AGAINST DECEPTION**. If you feel deceived or mistreated in any way, please contact the **UCSB HUMAN SUBJECTS COMMITTEE (HSC)**. These instructions are meant to clarify how the experiment actually works and how you earn money in the experiment. Our interest is in seeing how people with an **ACCURATE UNDERSTANDING** of how their decisions influence their outcomes and earnings make economic decisions.

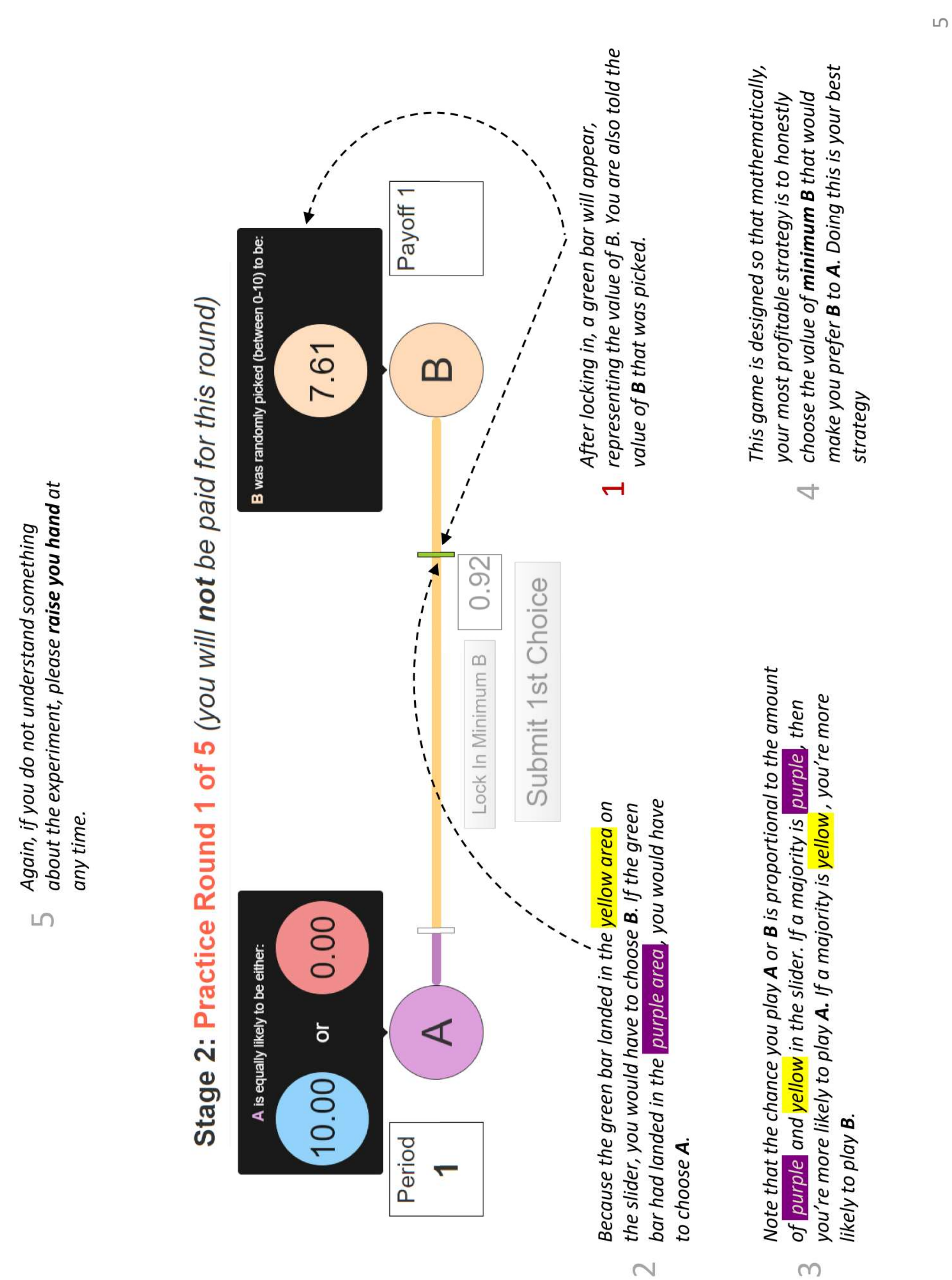
5. In the following instructions we will show you a number of screenshots from the computer interface. The text of the instructions is distributed around each page and is **NUMBERED IN ORDER**. Please read the numbered instructions in order (1,2,3. . .) on each page. You will be **QUIZZED** on the information.

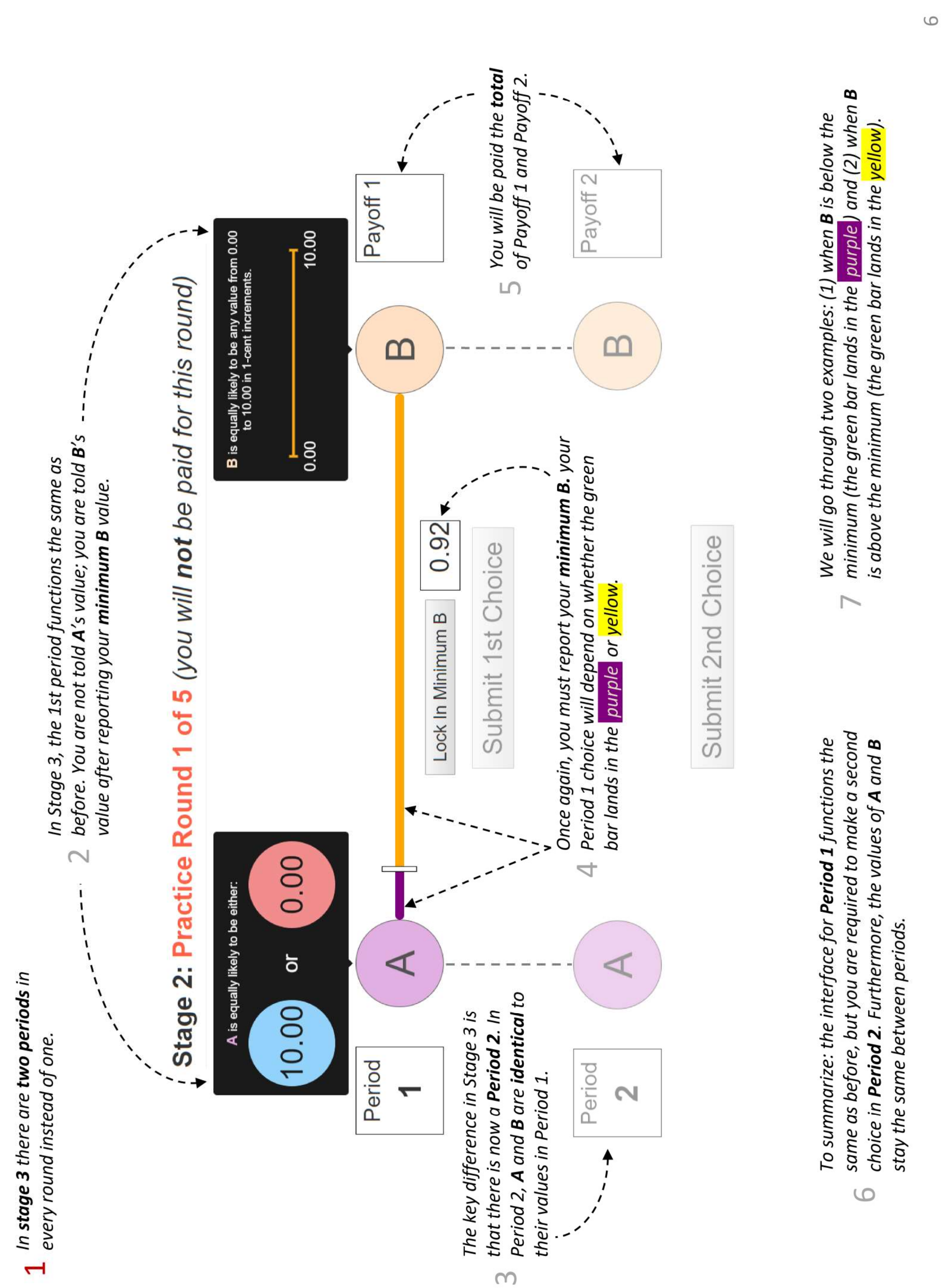
WEAK Treatment Instructions



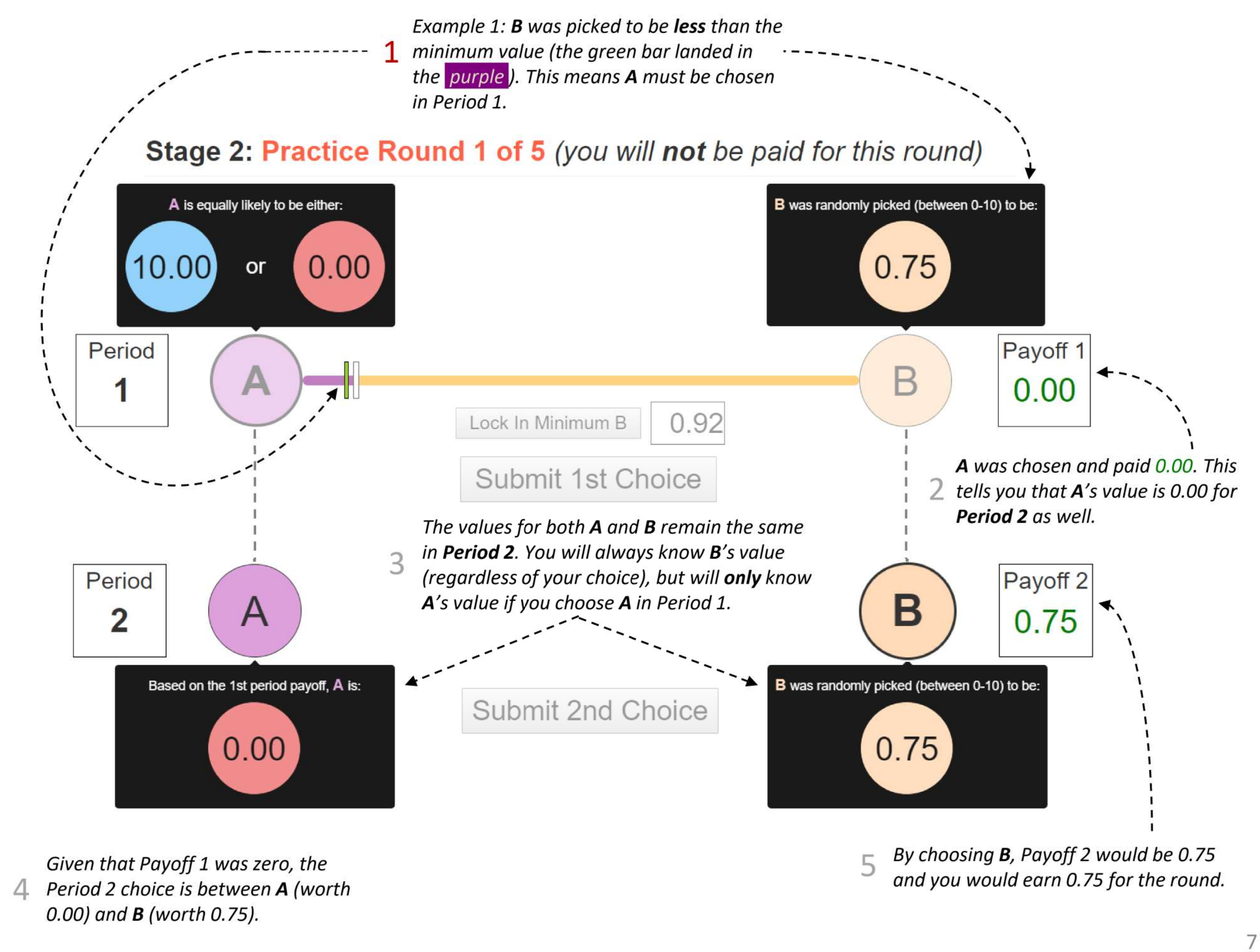

3 The first few rounds are practice (you won't be paid). Later rounds are for real money. Of the real-money rounds, one will be randomly selected for pay (all real money rounds are equally likely to be selected).
2 At any time, you can view a summary of the instructions by clicking Instructions.
Stage 1: Practice Round 1 of 5 (you will not be paid for this round)
A is equally likely to be either:
10.00
or
0.00
B was randomly picked (between 0-10) to be:
5.68
Period 1
A
B
Payoff 1
View Instructions
View History
Submit 1st Choice
4 As you play, think about a strategy. That is, think of the minimum B you would accept. If B is lower than this minimum, pick A; if B is higher than this minimum, pick B.
1 At any time, you can view a summary of values picked by the computer, your choices, and your payoffs from each round by clicking View History.
5 If you do not understand something about the experiment, please raise you hand at any time.
3

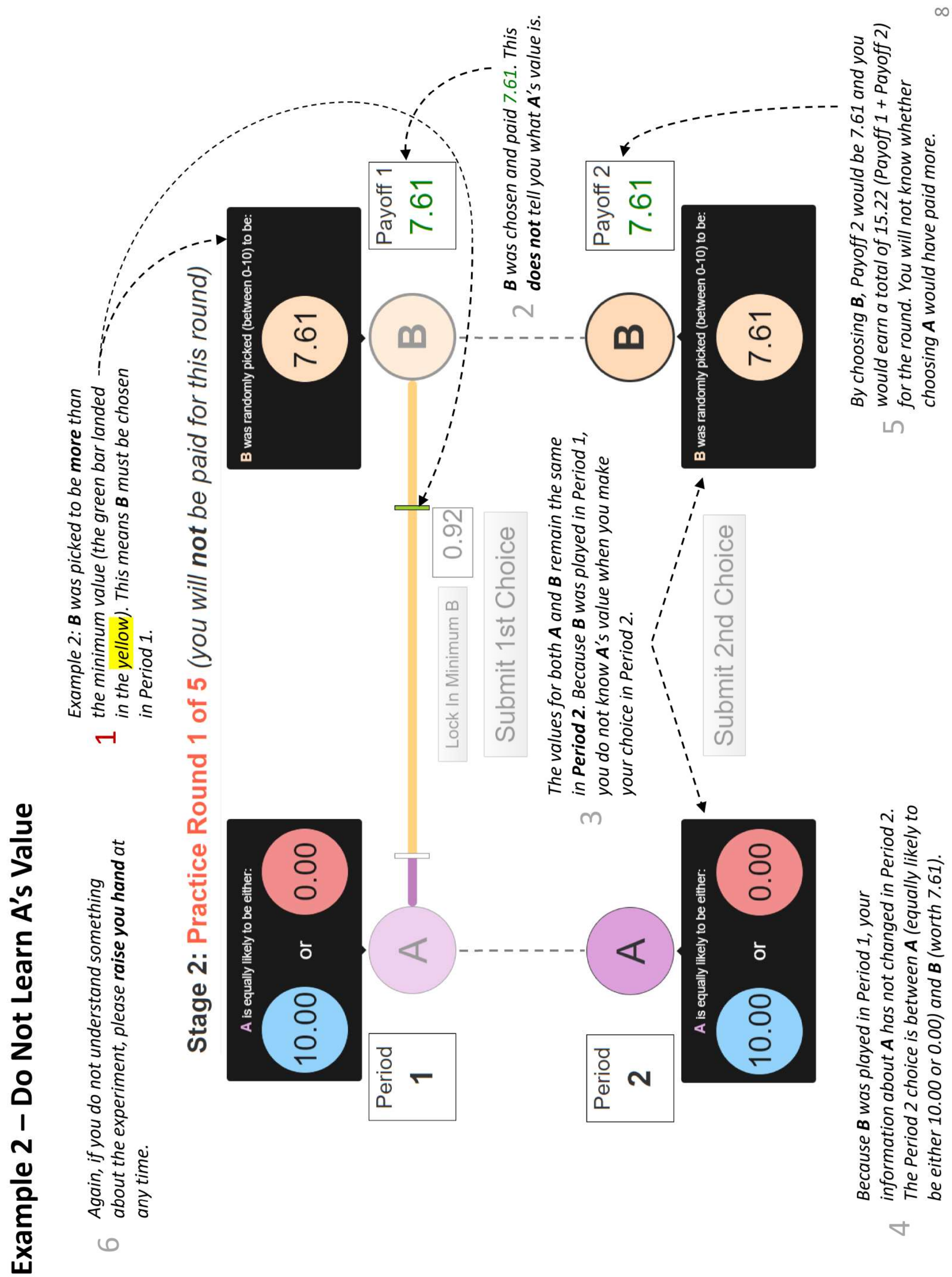

1 In Stage 2, A functions the same as before. Prior to the round, the computer randomly picks a value for A; it is equally likely to be 10.00 or 0.00. You are not told A's value.
2 B functions similarly to before. Prior to the round, the computer randomly picks a value for B; it can be any value from 0.00 to 10.00 in 1-cent increments. The difference is you must make a choice before you are told B's value.
Stage 2: Practice Round 1 of 5 (you will not be paid for this round)
A is equally likely to be either:
10.00
or
0.00
B is equally likely to be any value from 0.00 to 10.00 in 1-cent increments.
0.00
10.00
Period
1
A
B
Payoff 1
Lock In Minimum B
0.92
Submit 1st Choice
3 Prior to being told B's value, you must report the minimum value of B that would make you want to choose B rather than A (i.e., your strategy). To report your value, click somewhere on the slider bar or type a value in the text box.
4 If the value of B turns out to be lower than your minimum, you must choose A; if B turns out to be higher than your minimum, you must choose B. We will see an example soon.
5 The higher you set the slider, the more likely you commit to play A. In other words, as you increase the purple area, you are more likely to play A. Conversely, as you increase the yellow area, you are more likely to play B.
6 When you are satisfied with your minimum B value, click Lock in Minimum B to finalize your choice. After you click, the you will be told the value of B.
4


Stage 2: Practice Round 1 of 5 (you will not be paid for this round)
Period 1
A is equally likely to be either:
10.00 or 0.00
A
B was randomly picked (between 0-10) to be:
7.61
B
Payoff 1
Lock In Minimum B
0.92
Submit 1st Choice
1 After locking in, a green bar will appear, representing the value of B. You are also told the value of B that was picked.
2 Because the green bar landed in the yellow area on the slider, you would have to choose B. If the green bar had landed in the purple area, you would have to choose A.
3 Note that the chance you play A or B is proportional to the amount of purple and yellow in the slider. If a majority is purple, then you're more likely to play A. If a majority is yellow, you're more likely to play B.
4 This game is designed so that mathematically, your most profitable strategy is to honestly choose the value of minimum B that would make you prefer B to A. Doing this is your best strategy
5 Again, if you do not understand something about the experiment, please raise you hand at any time.
5


1 In **stage 3** there are **two periods** in every round instead of one.
2 In Stage 3, the 1st period functions the same as before. You are not told **A**'s value; you are told **B**'s value after reporting your **minimum B** value.
Stage 2: Practice Round 1 of 5 (you will **not** be paid for this round)
A is equally likely to be either:
10.00
or
0.00
B is equally likely to be any value from 0.00 to 10.00 in 1-cent increments.
0.00
10.00
Period 1
A
B
Payoff 1
Lock In Minimum B
0.92
Submit 1st Choice
Period 2
A
B
Payoff 2
Submit 2nd Choice
3 The key difference in Stage 3 is that there is now a **Period 2**. In Period 2, **A** and **B** are **identical** to their values in Period 1.
4 Once again, you must report your **minimum B**. Your Period 1 choice will depend on whether the green bar lands in the purple or yellow.
5 You will be paid the **total** of Payoff 1 and Payoff 2.
6 To summarize: the interface for **Period 1** functions the same as before, but you are required to make a second choice in **Period 2**. Furthermore, the values of **A** and **B** stay the same between periods.
7 We will go through two examples: (1) when **B** is below the minimum (the green bar lands in the purple) and (2) when **B** is above the minimum (the green bar lands in the yellow).
9

Example 1 – Learn A's Value




Example 2 – Do Not Learn A's Value
6 Again, if you do not understand something about the experiment, please raise you hand at any time.
1 Example 2: B was picked to be more than the minimum value (the green bar landed in the yellow). This means B must be chosen in Period 1.
Stage 2: Practice Round 1 of 5 (you will not be paid for this round)
A is equally likely to be either:
10.00
or
0.00
B was randomly picked (between 0-10) to be:
7.61
Period 1
A
B
Payoff 1
7.61
Lock In Minimum B
0.92
Submit 1st Choice
2 B was chosen and paid 7.61. This does not tell you what A's value is.
3 The values for both A and B remain the same in Period 2. Because B was played in Period 1, you do not know A's value when you make your choice in Period 2.
Period 2
A
B
Payoff 2
7.61
A is equally likely to be either:
10.00
or
0.00
Submit 2nd Choice
B was randomly picked (between 0-10) to be:
7.61
4 Because B was played in Period 1, your information about A has not changed in Period 2. The Period 2 choice is between A (equally likely to be either 10.00 or 0.00) and B (worth 7.61).
5 By choosing B, Payoff 2 would be 7.61 and you would earn a total of 15.22 (Payoff 1 + Payoff 2) for the round. You will not know whether choosing A would have paid more.
8

Stage 3 Summary

1 *The interface in Period 1 functions the same as before. The difference is that there are now **TWO PERIODS**.*

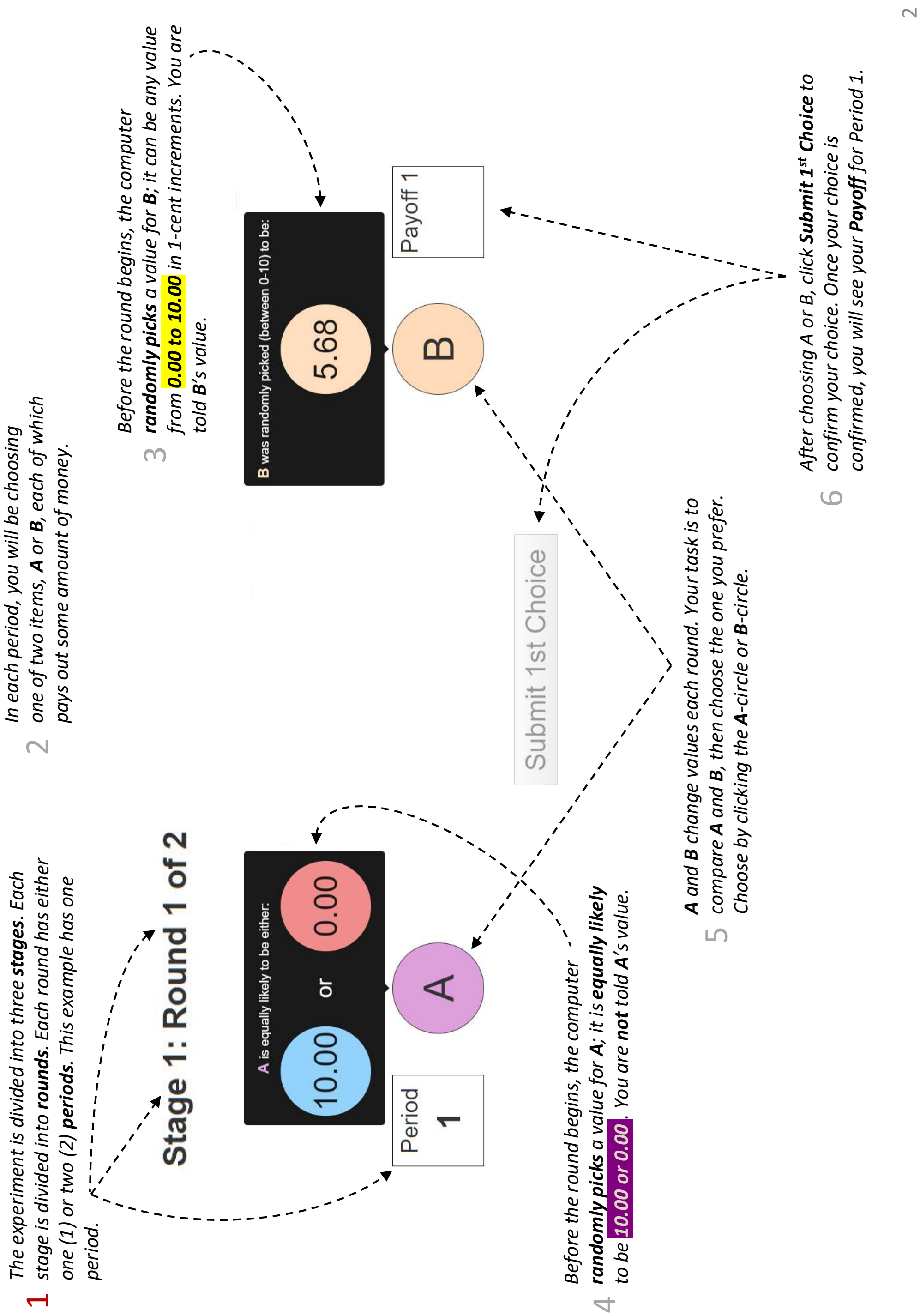
2 *The values of **A** and **B** both **REMAIN THE SAME** in Periods 1 and 2. You are always told the value of **B** (after reporting your minimum). If you play **A** in Period 1, you can **LEARN** its value for Period 2. You do not learn **A**'s value if you choose **B**.*

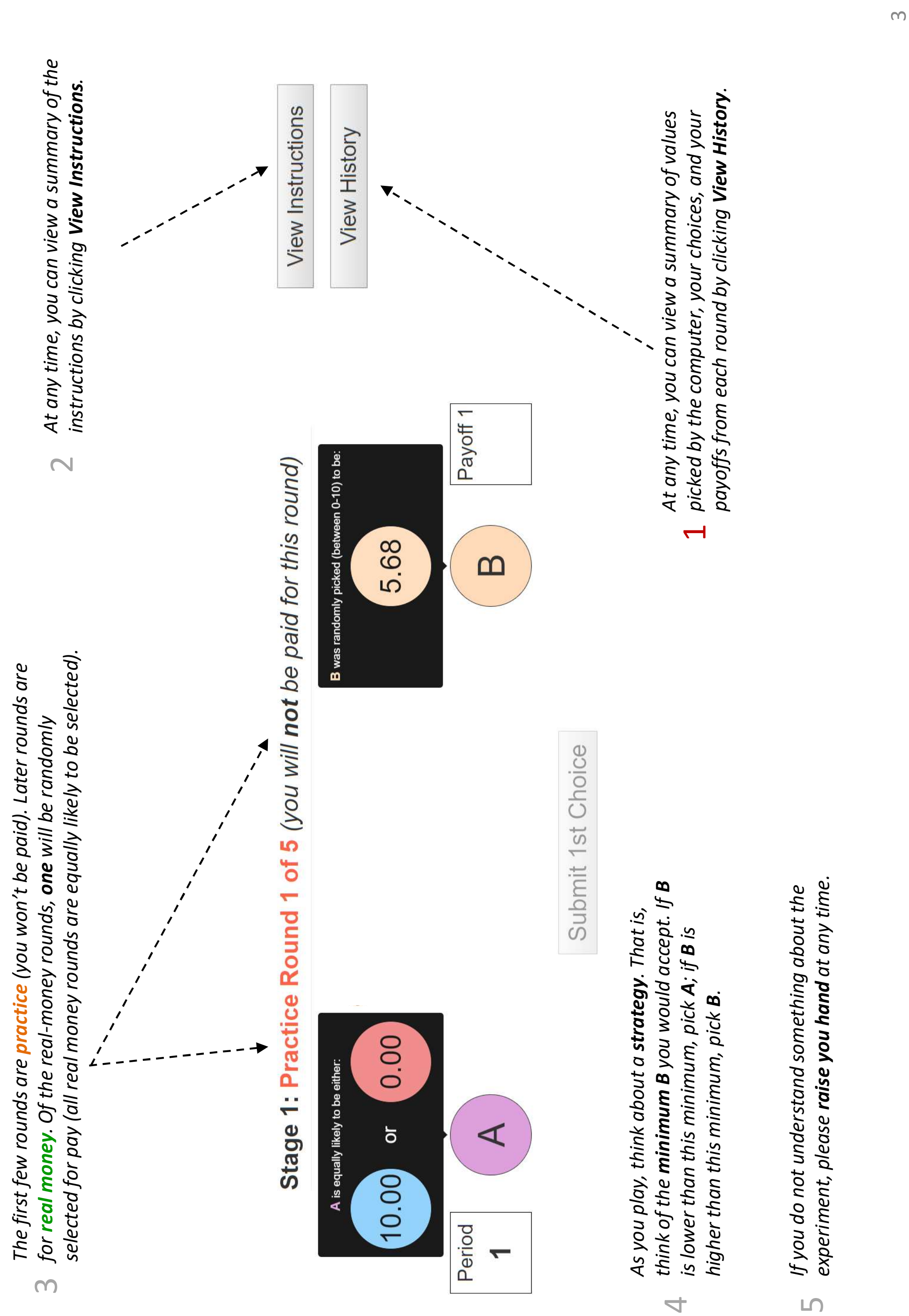
3 *Think of the **MINIMUM B** that would make you want to choose **B** rather than **A** in Period 1 (given that the problem is now two periods). Your Period 1 choice still depends on whether the green bar lands in the yellow or purple area on the slider.*

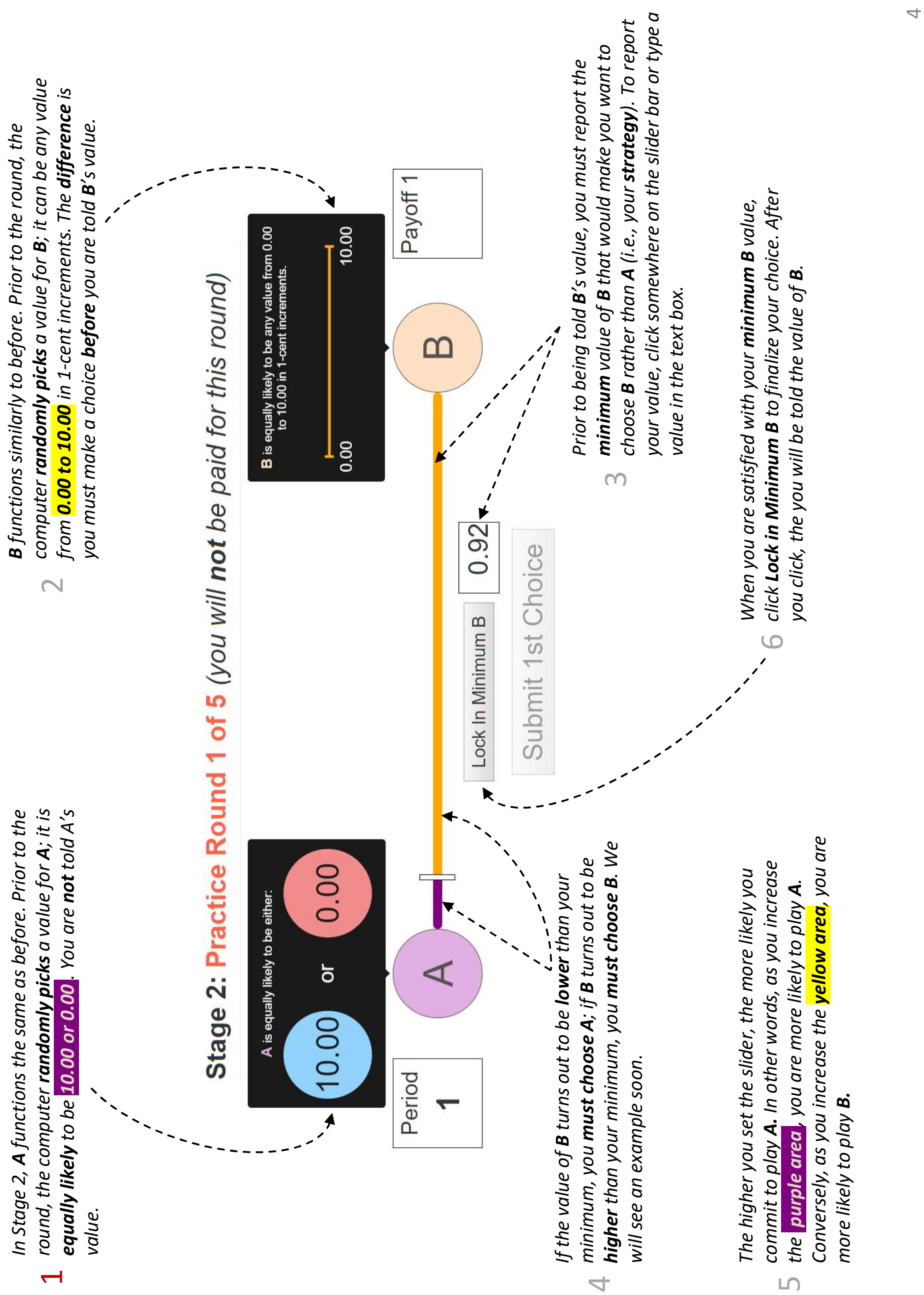
4 *You are free to choose **A** or **B** in Period 2; that is, your Period 2 choice is not limited by where the green bar lands on the slider.*

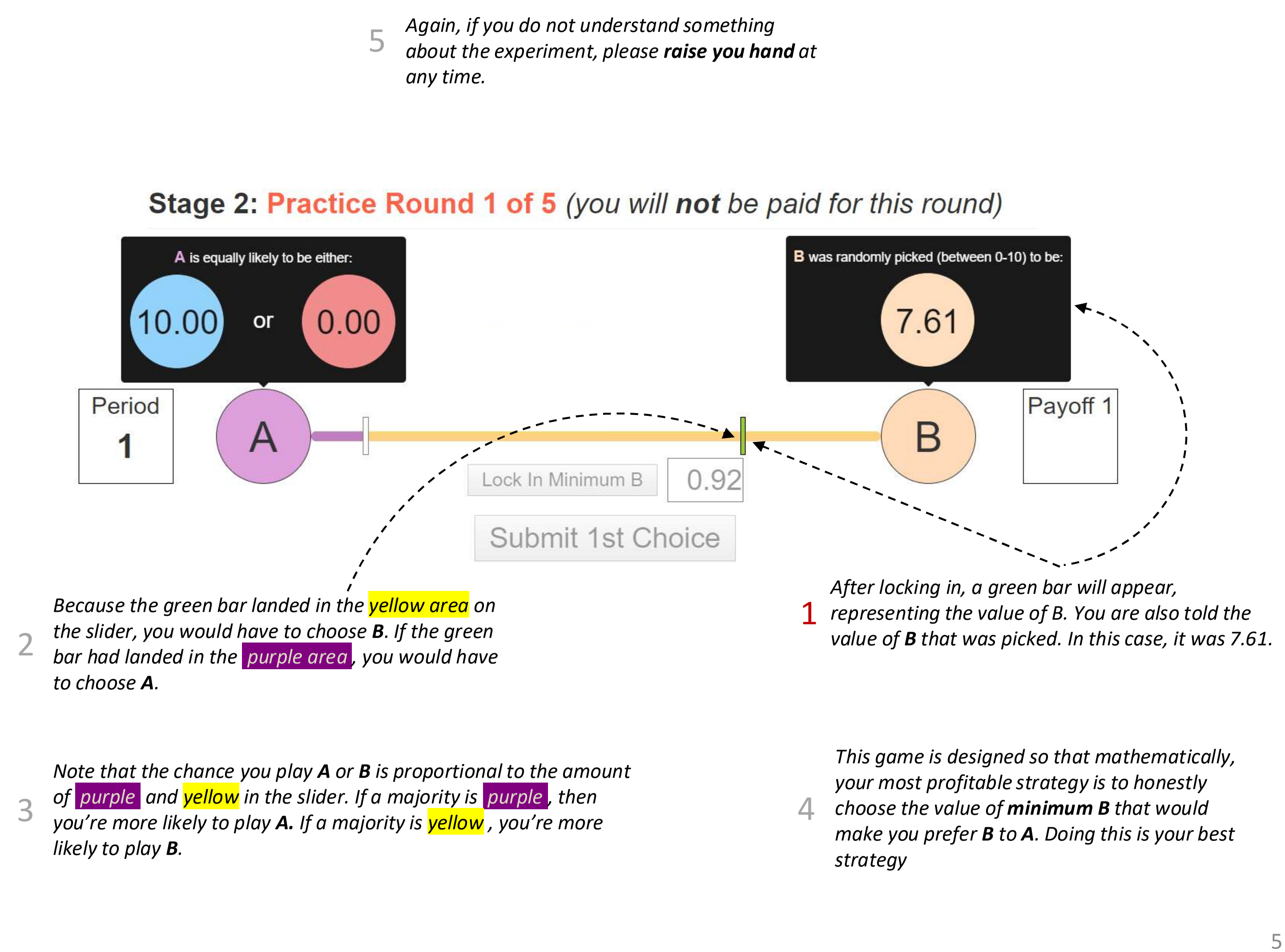
9

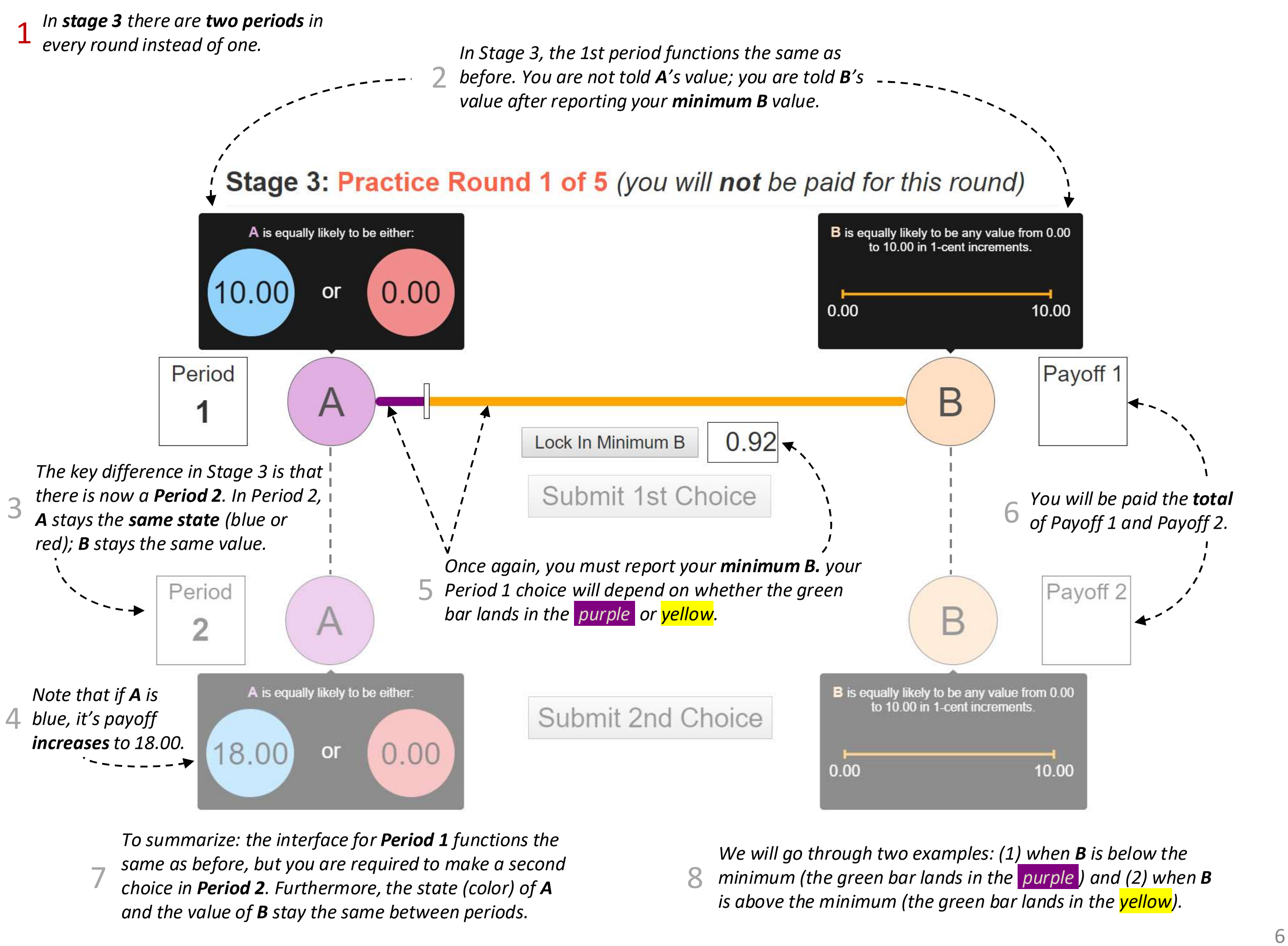
STRONG Treatment Instructions

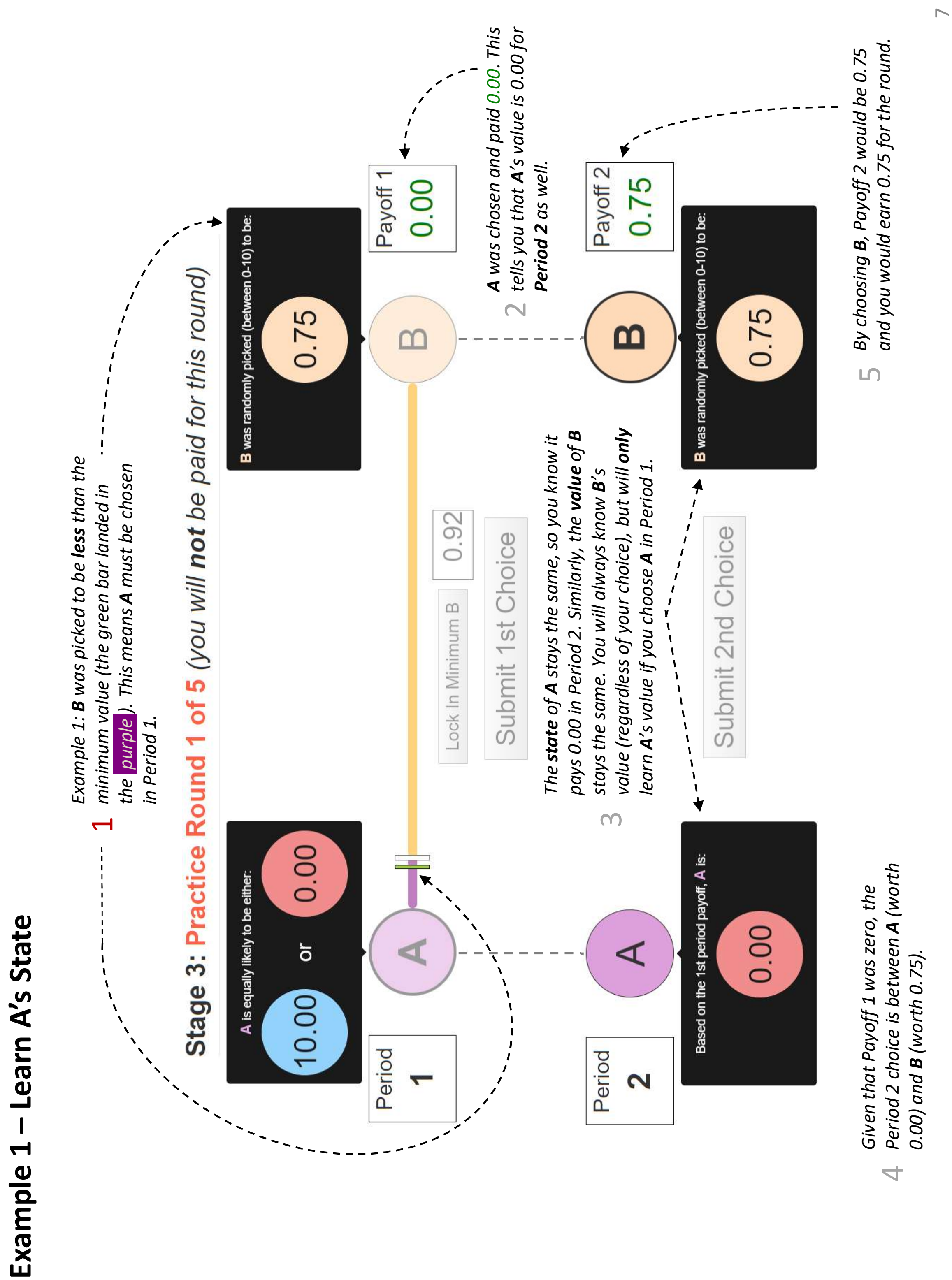



3 The first few rounds are practice (you won't be paid). Later rounds are for real money. Of the real-money rounds, one will be randomly selected for pay (all real money rounds are equally likely to be selected).
2 At any time, you can view a summary of the instructions by clicking View Instructions.
Stage 1: Practice Round 1 of 5 (you will not be paid for this round)
A is equally likely to be either:
10.00
or
0.00
B was randomly picked (between 0-10) to be:
5.68
Period
1
A
B
Payoff 1
View Instructions
View History
Submit 1st Choice
4 As you play, think about a strategy. That is, think of the minimum B you would accept. If B is lower than this minimum, pick A; if B is higher than this minimum, pick B.
1 At any time, you can view a summary of values picked by the computer, your choices, and your payoffs from each round by clicking View History.
5 If you do not understand something about the experiment, please raise you hand at any time.
3


1
In Stage 2, A functions the same as before. Prior to the round, the computer randomly picks a value for A; it is equally likely to be 10.00 or 0.00. You are not told A's value.
2
B functions similarly to before. Prior to the round, the computer randomly picks a value for B; it can be any value from 0.00 to 10.00 in 1-cent increments. The difference is you must make a choice before you are told B's value.
Stage 2: Practice Round 1 of 5 (you will not be paid for this round)
A is equally likely to be either:
10.00
or
0.00
B is equally likely to be any value from 0.00 to 10.00 in 1-cent increments.
0.00
10.00
Period
1
A
B
Payoff 1
Lock In Minimum B
0.92
Submit 1st Choice
4
If the value of B turns out to be lower than your minimum, you must choose A; if B turns out to be higher than your minimum, you must choose B. We will see an example soon.
3
Prior to being told B's value, you must report the minimum value of B that would make you want to choose B rather than A (i.e., your strategy). To report your value, click somewhere on the slider bar or type a value in the text box.
5
The higher you set the slider, the more likely you commit to play A. In other words, as you increase the purple area, you are more likely to play A. Conversely, as you increase the yellow area, you are more likely to play B.
6
When you are satisfied with your minimum B value, click Lock in Minimum B to finalize your choice. After you click, the you will be told the value of B.
4


5 Again, if you do not understand something about the experiment, please **raise you hand** at any time.
Stage 2: Practice Round 1 of 5 (you will **not** be paid for this round)
A is equally likely to be either:
10.00
or
0.00
B was randomly picked (between 0-10) to be:
7.61
Period
1
A
B
Payoff 1
Lock In Minimum B
0.92
Submit 1st Choice
1 After locking in, a green bar will appear, representing the value of B. You are also told the value of **B** that was picked. In this case, it was 7.61.
2 Because the green bar landed in the yellow area on the slider, you would have to choose **B**. If the green bar had landed in the purple area, you would have to choose **A**.
3 Note that the chance you play **A** or **B** is proportional to the amount of purple and yellow in the slider. If a majority is purple, then you're more likely to play **A.** If a majority is yellow, you're more likely to play **B.**
4 This game is designed so that mathematically, your most profitable strategy is to honestly choose the value of **minimum B** that would make you prefer **B** to **A**. Doing this is your best strategy
5

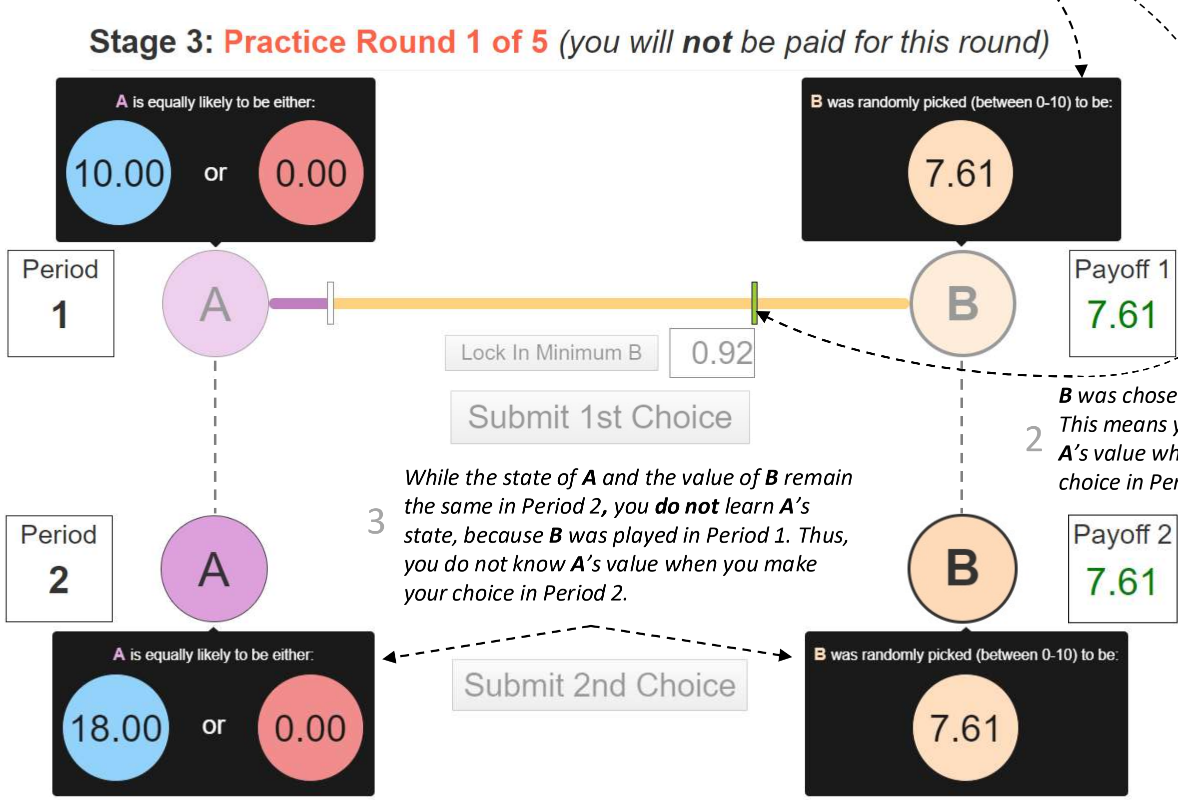

1 In stage 3 there are two periods in every round instead of one.
2 In Stage 3, the 1st period functions the same as before. You are not told A's value; you are told B's value after reporting your minimum B value.
Stage 3: Practice Round 1 of 5 (you will not be paid for this round)
A is equally likely to be either:
10.00
or
0.00
B is equally likely to be any value from 0.00 to 10.00 in 1-cent increments.
0.00
10.00
Period 1
A
B
Payoff 1
Lock In Minimum B
0.92
Submit 1st Choice
3 The key difference in Stage 3 is that there is now a Period 2. In Period 2, A stays the same state (blue or red); B stays the same value.
5 Once again, you must report your minimum B. your Period 1 choice will depend on whether the green bar lands in the purple or yellow.
6 You will be paid the total of Payoff 1 and Payoff 2.
Period 2
A
B
Payoff 2
4 Note that if A is blue, it's payoff increases to 18.00.
A is equally likely to be either:
18.00
or
0.00
Submit 2nd Choice
B is equally likely to be any value from 0.00 to 10.00 in 1-cent increments.
0.00
10.00
7 To summarize: the interface for Period 1 functions the same as before, but you are required to make a second choice in Period 2. Furthermore, the state (color) of A and the value of B stay the same between periods.
8 We will go through two examples: (1) when B is below the minimum (the green bar lands in the purple) and (2) when B is above the minimum (the green bar lands in the yellow).
6


Example 1 – Learn A's State
1 Example 1: B was picked to be less than the minimum value (the green bar landed in the purple). This means A must be chosen in Period 1.
Stage 3: Practice Round 1 of 5 (you will not be paid for this round)
Period 1
A is equally likely to be either:
10.00
or
0.00
A
B
B was randomly picked (between 0-10) to be:
0.75
Lock In Minimum B
0.92
Submit 1st Choice
Payoff 1
0.00
2 A was chosen and paid 0.00. This tells you that A's value is 0.00 for Period 2 as well.
3 The state of A stays the same, so you know it pays 0.00 in Period 2. Similarly, the value of B stays the same. You will always know B's value (regardless of your choice), but will only learn A's value if you choose A in Period 1.
Period 2
A
Based on the 1st period payoff, A is:
0.00
Submit 2nd Choice
B
B was randomly picked (between 0-10) to be:
0.75
Payoff 2
0.75
4 Given that Payoff 1 was zero, the Period 2 choice is between A (worth 0.00) and B (worth 0.75).
5 By choosing B, Payoff 2 would be 0.75 and you would earn 0.75 for the round.
7

Example 2 – Do Not Learn A’s State

6 *Again, if you do not understand something about the experiment, please **raise you hand** at any time.*

1 *Example 2: **B** was picked to be **more** than the minimum value (the green bar landed in the yellow). This means **B** must be chosen in Period 1.*



2 ***B** was chosen and paid 7.61. This means you will **not** know **A**’s value when you make your choice in Period 2.*

3 *While the state of **A** and the value of **B** remain the same in Period 2**,** you **do not** learn **A**’s state, because **B** was played in Period 1. Thus, you do not know **A**’s value when you make your choice in Period 2.*

4 *Because **B** was played in Period 1, your information about **A** has not changed in Period 2. The Period 2 choice is between **A** (equally likely to be either 18.00 or 0.00) and **B** (worth 7.61).*

5 *By choosing **B**, Payoff 2 would be 7.61 and you would earn a total of 15.22 (Payoff 1 + Payoff 2) for the round. You will not know whether choosing **A** would have paid more.*

8

Stage 3 Summary

1 *The interface in Period 1 functions the same as before. The main difference is that there are now **TWO PERIODS** and **A** potentially pays more in the second period (**18.00**) than in the first period (**10.00**).*

2 *The state (color) of **A** and the value of **B** both **REMAIN THE SAME** in Periods 1 and 2. You are always told the value of **B** (after reporting your minimum). If you play **A** in Period 1, you can **LEARN** whether it pays 18.00 or 0.00 in Period 2. You do not learn **A**'s state if you choose **B**.*

3 *Think of the **MINIMUM B** that would make you want to choose **B** rather than **A** in Period 1 (given that the problem is now two periods). Your Period 1 choice still depends on whether the green bar lands in the yellow or purple area on the slider.*

4 *You are free to choose **A** or **B** in Period 2; that is, your Period 2 choice is not limited by where the green bar lands on the slider.*

9

A.2 Complexity and Procedural Choice: Evidence from Experimental Bandits

A.2.1 Theoretical Benchmarks

In this appendix, we derive two benchmarks of interest for our experiment. The first is a derviation of the critical value for exploration in our bandit problem. The second is a derivation of the the optimal random transition probability for the 2-state partially randomized procedure in section 2. These derivations closely follow Börgers and Morales (2004) and we refer the reader there for further analysis.

Critical value for exploration

We can find a critical value such that a decision maker would never explore. Intuitively, if x is close to 1, the cost of exploration (the risk of earning a zero) outweighs benefit (the chance of earning one). Conversely, if x is close to 0, the opportunity cost of exploration is extremely low, so the reverse is true. The condition is:

$$\left(\frac{x}{1-\delta}\right) \geq \frac{1}{3}\left(\frac{\delta x}{1-\delta} + \frac{x}{1-\delta} + \frac{1}{1-\delta}\right) \tag{1}$$

On the left-hand side, we have the present discounted value of staying on the first arm forever. On the right-hand side are three equally likely cases: (i) the DM earns zero, then returns to the first arm forever, (ii) the DM earns x and their decisions do not matter for payoffs, or (iii) the DM earns one and stays on the second arm forever. Setting these two equal, we can solve for the critical value.

$$\overline{x} = \frac{1}{1 + (1-\delta)} \tag{2}$$

For values larger than $\overline{x}$, it is optimal to remain on the initial arm and never explore.

A.2.2 Optimal Transition Probability for Partially Random 2-State Procedure

Consider first the value function when the arms are $(x, 0)$ and the DM's strategy dictates a pull of the initial arm:

$$V_0 = x + \delta\big[(1-p)V_0 + p\delta V_0\big] \tag{3}$$

The DM's first choice yields a payoff of x, and we discount future payoffs by δ. In the next period, with probability $(1-p)$ she pulls the initial arm again (leaving her state unchanged). With probability p she pulls the other arm, which pays zero; she then returns to the initial arm. Solving for V_0:

$$V_0 = \frac{x}{(1-\delta)(1+\delta p)} \tag{4}$$

Next, consider the value function when the arms are (x, x):

$$V_x = \frac{x}{1-\delta} \tag{5}$$

If both arms pay x, then the DM's pattern of choices do not affect payoffs. Finally, consider the value function when the arms are $(x, 1)$ and the DM's strategy dictates a pull of the initial arm:

$$V_1 = x + \delta\left[(1-p)V_1 + p\left(\frac{1}{1-\delta}\right)\right] \tag{6}$$

The DM pulls the initial arm, earning x. She then stays on the initial arm with probability $(1-p)$; with probability p, she pulls the other arm and stays forever. Solving for V_1:

$$V_1 = \frac{(1-\delta)x + \delta p}{(1-\delta)(1-\delta+\delta p)} \tag{7}$$

In the optimal 2-State strategy, when the DM plays the initial arm, she cannot remember if she has tried the other arm (or what it pays). When her strategy dictates a pull of the initial arm, her value function is the average of V_0, V_x, and V_1.

$$V = \frac{1}{3}\left[\frac{x}{(1-\delta)(1+\delta p)} + \frac{x}{1-\delta} + \frac{(1-\delta)x+\delta p}{(1-\delta)(1-\delta+\delta p)}\right] \tag{8}$$

We can maximize this function with respect to p to find the optimal experimentation probability. Note that maximizing V with respect to p is the same as maximizing:

$$M = \frac{x}{1+\delta p} + \frac{(1-\delta)x+\delta p}{(1-\delta+\delta p)} \tag{9}$$

To find the maximum, we consider the first and second derivatives.

$$\frac{dM}{dp} = -\frac{\delta x}{(1+\delta p)^2} + \frac{\delta(1-\delta)(1-x)}{(1-\delta+\delta p)^2} \tag{10}$$

$$\frac{d^2M}{dp^2} = \frac{2\delta^2 x}{(1+\delta p)^3} - \frac{2\delta^2(1-\delta)(1-x)}{(1-\delta+\delta p)^3} \tag{11}$$

The second derivative is negative when:

$$p < \frac{(1-\delta)^{\frac{1}{3}}(1-x)^{\frac{1}{3}} - (1-\delta)x^{\frac{1}{3}}}{\delta(x^{\frac{1}{3}} - (1-\delta)^{\frac{1}{3}}(1-x)^{\frac{1}{3}})} \tag{12}$$

Recall that p must be greater than zero. Consider the numerator and the denominator separately. The numerator is positive when:

$$(1-\delta)x^{\frac{1}{3}} < (1-\delta)^{\frac{1}{3}}(1-x)^{\frac{1}{3}} \tag{13}$$

$$x < \frac{1}{1+(1-\delta)^2} \tag{14}$$

Note that his value is strictly greater than the critical value in (2). As a result, the numerator will be positive for all values of x such that an DM would consider exploration. The denominator is positive when:

$$x^{\frac{1}{3}} > (1-\delta)^{\frac{1}{3}}(1-x)^{\frac{1}{3}} \tag{15}$$

We can set these equal to solve for a second critical value.

$$\underline{x} = \frac{1-\delta}{2-\delta} \tag{16}$$

For values $x \geq \overline{x}$, we know that $p = 0$. For values $\underline{x} < x < \overline{x}$, we can solve for the optimal p by setting the first derivative in 10 equal to zero and solving for p:

$$p = \frac{\sqrt{1-\delta}\sqrt{1-x} - (1-\delta)\sqrt{x}}{\delta\sqrt{x} - \delta\sqrt{1-\delta}\sqrt{1-x}} \tag{17}$$

What about values of x less than the critical value in (16)? Consider a case where $p = 1$, but the first derivative is still positive (i.e., the maximum of M is at the upper boundary of p). In this situation, the following inequality holds:

$$\frac{x}{(1+\delta)^2} < \frac{\delta(1-\delta)(1-x)}{(1-\delta+\delta)^2} \tag{18}$$

$$x < \frac{1-\delta}{\frac{1}{(1+\delta)^2} + 1 - \delta} \tag{19}$$

The value in (19) is strictly greater than $\underline{x}$. Thus, for all values of x less than the critical value in (16), the optimal experimentation probability is $p = 1$. Note, however, that (19) also established that there is a range of x such that $\underline{x} < x < \overline{x}$ and $p = 1$, i.e., p as define in (17) may be greater than 1. Thus, the optimal experimentation probability must be defined piece-wise:

$$p = \begin{cases} 0 & \text{if } x \geq \overline{x} \\ \min\left\{ \frac{\sqrt{1-\delta}\sqrt{1-x}-(1-\delta)\sqrt{x}}{\delta\sqrt{x}-\delta\sqrt{1-\delta}\sqrt{1-x}}, 1 \right\} & \text{if } \underline{x} < x < \overline{x} \\ 1 & \text{if } x \leq \underline{x} \end{cases} \tag{20}$$

A.2.3 Instructions to Subjects

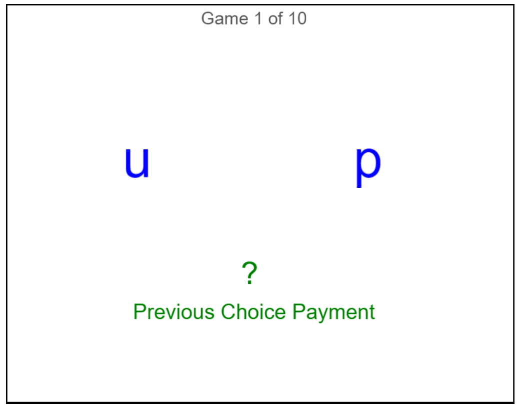
In this appendix we reproduce the instructions to subjects. These were deployed in HTML in the experiment and unfolded progressively interspersed with comprehension quiz questions. We first reproduce instructions from our ST treatment, then again for our NST treatment. Instructions for ST+CL are similar except we also include instructions for the operation span memory task simultaneously deployed in the treatment.

State Tracking Treatment

1. **Introduction**

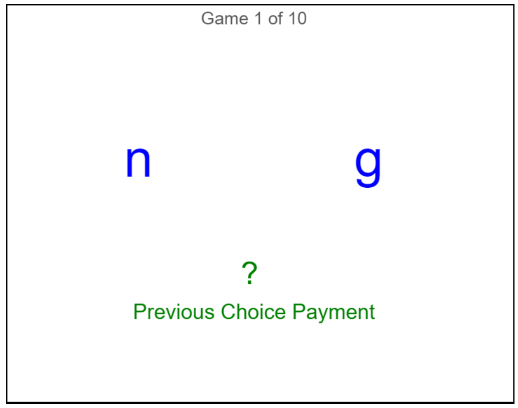
- We will start by providing you with **INSTRUCTIONS** for the study.
- We will ask you **QUESTIONS** to check that you understand the instructions. You should be able to answer all of these questions correctly.
- Please read and follow the instructions closely and carefully.
- If you **COMPLETE** the main parts of the study, you will receive a **GUARANTEED PAYMENT** of **$2.50**.
- In addition, your **CHOICES** in the GAME portion of the study will result in **PERFORMANCE-BASED EARNINGS**. You will play in **TWENTY (20) GAMES** worth **REAL MONEY**. Your **AVERAGE** points from **ALL TWENTY GAMES** will be converted into an additional payment.
- After you finish the instructions, you will have a chance to play several **PRACTICE GAMES** before you play for real money.

2. **Two Options to Choose Between**



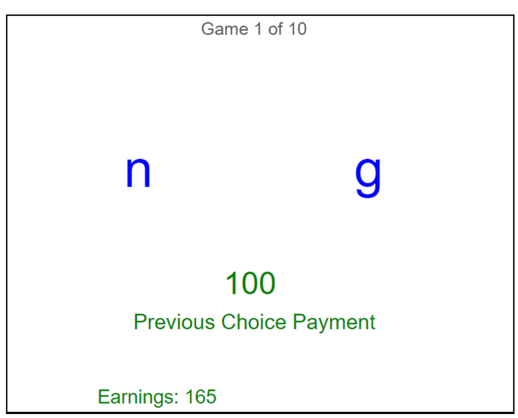
- The experiment is divided into twenty **GAMES**, each of which is is divided into several **CHOICES**.
- In each game, you will repeatedly choose between **TWO OPTIONS** that we will call **EARLIER-LETTER** and **LATER-LETTER**, represented by two letters on your screen.
  - In the example above, the Earlier-Letter option is represented by '**p**' and the later-letter by '**u**' (because '**p**' comes **earlier** than '**u**' in the alphabet).
- Your **FIRST** choice in a game will **ALWAYS PAY 65** points. This is true whether you choose Earlier-Letter or Later-Letter first.
  - For example, suppose your **first** choice were **Earlier-Letter**. Then you would know that Earlier-Letter would pay **65 points every time** you choose it for the rest of the game.
- After your first choice, the **OTHER OPTION** will pay a **VALUE** of either **0**, **65**, or **100**. This value is **INDEPENDENTLY** and **RANDOMLY** determined by the computer before the game begins. Each value (**65**, or **100**) is **EQUALLY LIKELY** to be selected. It remains the **SAME WITHIN A GAME**.
  - For example, suppose your **first** choice were **Later-Letter**. Then **Earlier-Letter** would be equally likely to pay (**65**, or **100**). If you choose Earlier-Letter and it pays **100**, then you would know that Earlier-letter would pay **100 points every time** you choose it for the rest of the game.
- However, the value of each option **CHANGES BETWEEN GAMES**: once a game ends, payments reset. Your first choice (either Earlier-Letter or Later-Latter) will pay 65, and the other option will get a new random value.

3. **Typing Letters to Make Choices**



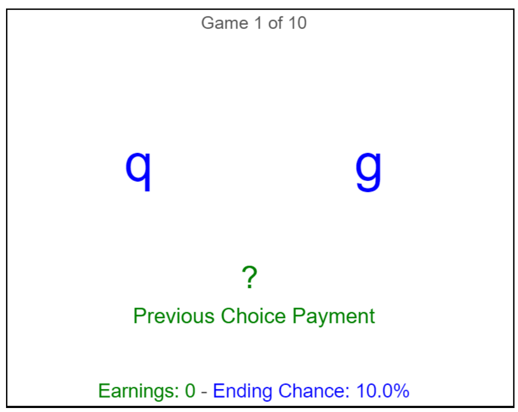
- For each **CHOICE**, each of the **TWO OPTIONS** will require you to **TYPE A LETTER**.
  - In the example above, you would need to type '**n**' (lower case 'N') or '**g**' (lower case 'G') to make a choice.
- These **LETTERS** will **RANDOMLY CHANGE** ('a' to 'z') for each choice and be shown in a **RANDOM ORDER** (left or right) on your screen.
- The **EARLIER LETTER** in the alphabet will always represent the Earlier-Letter option; the **LATER LETTER** in the alphabet will always represent the Later-Letter option.
  - In the example above, typing '**g**' will select the Earlier-Letter option (and give you the Earlier-Letter payment) while typing '**n**' will select the Later-Letter option (and give you the Later-Letter payment).

4. **Tracking Payments**



- After you choose an option, the **PAYMENT** you earned (**0**, **65**, or **100**) for that choice appears in the middle of the screen in OliveGreen. This value always represents your **PREVIOUS CHOICE'S** payment.
  - In the example above, the previous choice paid **100** points.
- Your **EARNINGS** are cumulative for **ALL YOUR CHOICES** so far in the game and appear at the bottom in OliveGreen.
  - This number is the **sum** of all your payments so far in the game.
  - In the example above, the choices have paid a total of **165** points so far.

5. **Blocks of Choices**



- The **NUMBER OF CHOICES** you're allowed to make in any game is **RANDOM**. Every time you make a choice, there is a **10% CHANCE** that the computer will make it the **LAST** (paying) choice of the game.
  - A 10% chance that the game ends each choice means that on **average** there will be **10 choices** in the game.
  - Many games will be **shorter**, but others will be **much longer**.
  - The probability each choice is the last **does not depend** on how many choices you have already made. Every choice is equally likely to be the last one that counts.
- In each game, you will always make your choices in **BLOCKS OF FIVE**.
- After every block of five the computer will tell you whether the game actually **RANDOMLY ENDED** during that block. If the computer randomly ended the game during the block (before the last

choice of the block), any choices you made **AFTER THE LAST** choice in the block **WON'T COUNT** for payment.

  - Example: If you make **five choices** in a block and the computer randomly **ended** the game on the **third choice** of the block, choices **1, 2 and 3 in the block will count** for payment and choices **4 and 5 in the block will not count** for payment.

- When you have made the **FINAL CHOICE** in a game, the computer will inform you that this has happened and you will start a **NEW GAME**. When a new game starts, the **VALUES WILL CHANGE** for the Earlier-Letter and Later-Letter options. There is no connection between games – each game will be brand new.

6. **Other Instructions**

- You will play in two practice games to familiarize yourself with the software before you play games for real money.
- Because part of the experiment tests your memory, please do not use external tools (e.g., pencil and paper) to assist during the experiment.
- Please do not unnecessarily refresh your browser, as doing so can make the software unstable. Qualtrics tracks your refreshes–excessive refreshes will void your bonus payment.

7. **Cash Payments**

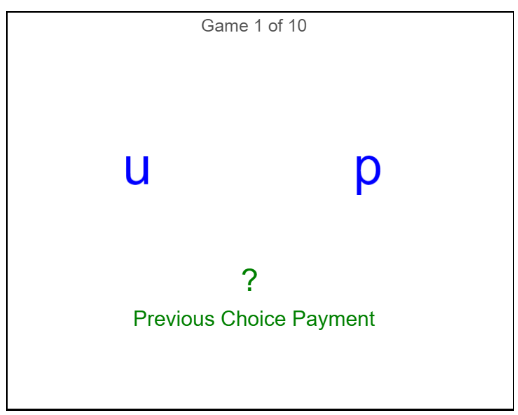
- You will be paid $2.50 for finishing the experiment. If you decide to leave before finishing, you will forfeit this amount.
- In addition, you will potentially earn a **PERFORMANCE-BASED BONUS**.
- Your **POINTS** from **ALL TWENTY (20) GAMES** will be **AVERAGED**.
- If your **AVERAGE** point total is **GREATER THAN 700**, you will earn a **BONUS**.
  - For every point you earn (on average) greater than 700, you will be paid $0.03 (three cents)
  - For example, if you average **950** points, your bonus would be: (**950** - **700**) * $0.03 = $7.50
  - For example, if you average **600** points, you would not earn a bonus.

No State Tracking Treatment

1. **Introduction**

   - We will start by providing you with **INSTRUCTIONS** for the study.
   - We will ask you **QUESTIONS** to check that you understand the instructions. You should be able to answer all of these questions correctly.
   - Please read and follow the instructions closely and carefully.
   - If you **COMPLETE** the main parts of the study, you will receive a **GUARANTEED PAYMENT** of **$2.50**.
   - In addition, your **CHOICES** in the GAME portion of the study will result in **PERFORMANCE-BASED EARNINGS**. You will play in **TWENTY (20) GAMES** worth **REAL MONEY**. Your **AVERAGE** points from **ALL TWENTY GAMES** will be converted into an additional payment.
   - After you finish the instructions, you will have a chance to play several **PRACTICE GAMES** before you play for real money.

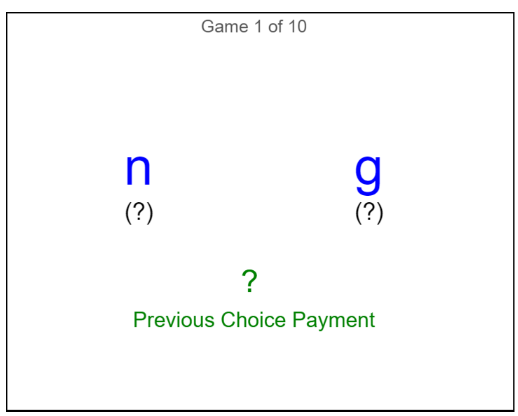
2. **Two Options to Choose Between**



   - The experiment is divided into twenty **GAMES**, each of which is is divided into several **CHOICES**.
   - In each game, you will repeatedly choose between **TWO OPTIONS** that we will call **EARLIER-LETTER** and **LATER-LETTER**, represented by two letters on your screen.
     - In the example above, the Earlier-Letter option is represented by '**p**' and the later-letter by '**u**' (because '**p**' comes **earlier** than '**u**' in the alphabet).

- Your **FIRST** choice in a game will **ALWAYS PAY** **65** points. This is true whether you choose Earlier-Letter or Later-Letter first.
  - For example, suppose your **first** choice were **Earlier-Letter**. Then you would know that Earlier-Letter would pay **65** **points every time** you choose it for the rest of the game.
- After your first choice, the **OTHER OPTION** will pay a **VALUE** of either **0**, **65**, or **100**. This value is **INDEPENDENTLY** and **RANDOMLY** determined by the computer before the game begins. Each value (**65**, or **100**) is **EQUALLY LIKELY** to be selected. It remains the **SAME WITHIN A GAME**.
  - For example, suppose your **first** choice were **Later-Letter**. Then **Earlier-Letter** would be equally likely to pay (**65**, or **100**). If you choose Earlier-Letter and it pays **100**, then you would know that Earlier-letter would pay **100** **points every time** you choose it for the rest of the game.
- However, the value of each option **CHANGES BETWEEN GAMES**: once a game ends, payments reset. Your first choice (either Earlier-Letter or Later-Latter) will pay 65, and the other option will get a new random value.

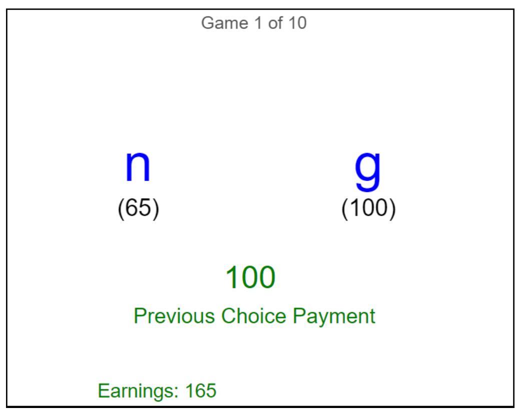
3. **Typing Letters to Make Choices**



- For each **CHOICE**, each of the **TWO OPTIONS** will require you to **TYPE A LETTER**.
  - In the example above, you would need to type '**n**' (lower case 'N') or '**g**' (lower case 'G') to make a choice.
- These **LETTERS** will **RANDOMLY CHANGE** ('a' to 'z') for each choice and be shown in a **RANDOM ORDER** (left or right) on your screen.

- The **EARLIER LETTER** in the alphabet will always represent the Earlier-Letter option; the **LATER LETTER** in the alphabet will always represent the Later-Letter option.
  - In the example above, typing '**g**' will select the Earlier-Letter option (and give you the Earlier-Letter payment) while typing '**n**' will select the Later-Letter option (and give you the Later-Letter payment).
- The **QUESTION MARKS** (in parentheses) under each letter indicates that you have not yet tried that option.
  - In the example above, the '**(?)**' under '**g**' indicates Earlier-Letter has not been tried
  - In the example above, the '**(?)**' under '**n**' indicates Later-Letter has not been tried.

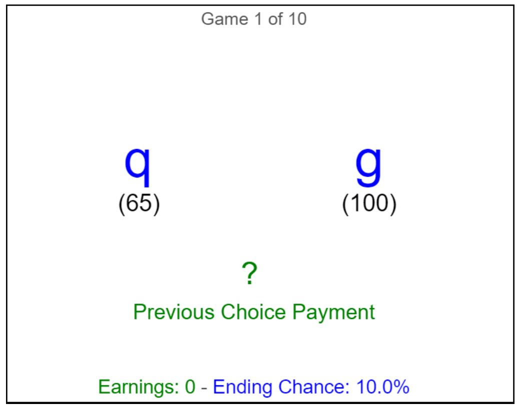
4. **Tracking Payments**



- After you choose an option, the **PAYMENT** you earned (**0**, **65**, or **100**) for that choice appears in the middle of the screen in OliveGreen. This value always represents your **PREVIOUS CHOICE'S** payment.
  - In the example above, the previous choice paid **100** points.
- Your **EARNINGS** are cumulative for **ALL YOUR CHOICES** so far in the game and appear at the bottom in OliveGreen.
  - This number is the **sum** of all your payments so far in the game.
  - In the example above, the choices have paid a total of **165** points so far.
- After you have **CHOSEN** an option, the **AMOUNT** it pays appears **BELOW**. This number will remain until the game ends and a new one begins

- In the example above, Earlier-Letter was tried and paid **65**
- In the example above, Later-Letter was tried and paid **100**

5. **Blocks of Choices**



- The **NUMBER OF CHOICES** you're allowed to make in any game is **RANDOM**. Every time you make a choice, there is a **10% CHANCE** that the computer will make it the **LAST** (paying) choice of the game.
  - A 10% chance that the game ends each choice means that on **average** there will be **10 choices** in the game.
  - Many games will be **shorter**, but others will be **much longer**.
  - The probability each choice is the last **does not depend** on how many choices you have already made. Every choice is equally likely to be the last one that counts.
- In each game, you will always make your choices in **BLOCKS OF FIVE**.
- After every block of five the computer will tell you whether the game actually **RANDOMLY ENDED** during that block. If the computer randomly ended the game during the block (before the last choice of the block), any choices you made **AFTER THE LAST** choice in the block **WON'T COUNT** for payment.
  - Example: If you make **five choices** in a block and the computer randomly **ended** the game on the **third choice** of the block, choices **1, 2 and 3 in the block will count** for payment and choices **4 and 5 in the block will not count** for payment.
- When you have made the **FINAL CHOICE** in a game, the computer will inform you that this has happened and you will start a **NEW GAME**. When a new game starts, the **VALUES**

WILL CHANGE for the Earlier-Letter and Later-Letter options. There is no connection between games – each game will be brand new.

6. **Other Instructions**

- You will play in two practice games to familiarize yourself with the software before you play games for real money.
- Please do not unnecessarily refresh your browser, as doing so can make the software unstable. Qualtrics tracks your refreshes–excessive refreshes will void your bonus payment.

7. **Cash Payments**

- You will be paid $2.50 for finishing the experiment. If you decide to leave before finishing, you will forfeit this amount.
- In addition, you will potentially earn a **PERFORMANCE-BASED BONUS**.
- Your **POINTS** from **ALL TWENTY (20) GAMES** will be **AVERAGED**.
- If your **AVERAGE** point total is **GREATER THAN 700**, you will earn a **BONUS**.
  - For every point you earn (on average) greater than 700, you will be paid $0.03 (three cents)
  - For example, if you average **950** points, your bonus would be: (**950** - **700**) * $0.03 = $7.50
  - For example, if you average **600** points, you would not earn a bonus.

A.3 The Economics of the Montana Liquor License System

A.3.1 Timeline of Major Legal Changes

- **1933**: End of Prohibition; MT Liquor Control Board created; beer by-the-drink legalized
- **1937**: Liquor by-the-drink legalized; MT Liquor License Dealers Association formed
- **1947**: Incorporated license quota system created
- **1963**: Unincorporated license quota implemented
- **1966**: MT Liquor License Dealers Association renamed Montana Taverns Association
- **1967**: Private beer off-premise sales legalized
- **1973**: Department of Revenue assumes the MLCB's duties
- **1974**: Wine amendment added to beer retail licenses
- **1975**: All-beverage, floater, and resort licenses created; retail beer license quota implemented
- **1976**: MT Supreme Court legalizes video keno
- **1978**: Voter initiative legalizes wine sales in grocery stores
- **1979**: Catering endorsements and airport licenses created
- **1983**: Lottery system for new permits implemented
- **1984**: MT Supreme Court rules video poker is illegal
- **1985**: Video Poker Machine Act allows 5 video poker machines per licensee
- **1987**: 15% income tax implemented on gambling machines
- **1991**: Licensees limited to 20 total machines per licensed premises
- **1995**: State liquor stores privatized
- **1997**: Restaurant beer and wine licenses created
- **1999**: Direct sales by small breweries legalized
- **2005**: U.S. District Court rules resident requirement unconsitutional
- **2007**: Floaters gaming-restricted, RBW licenses doubled
- **2009**: Smoking ban in effect

A.3.2 Quota Formula and Examples

Table 3: All-Beverage License Quota

Population	Quota
1 - 499	2
500 - 1,000	3
1,001 - 2,000	4
2,001 - 3,000	5
3,001 - 4,500	6
4,051 - 6,000	7
6,001 - 7,500	8
7,501 - 9,000	9
9,001 - 10,500	10
10,501 - 12,000	11
12,001 - 13,500	12
13,501 - 15,000	13
Every additional 1,500 residents	+ 1

The discussion in this appendix provides a more detailed explanation of the quota system for all-beverage licenses using the scenarios laid out in the tables 3-6 below. Table 3 provides a schedule that dictates the relationship between the population of a quota area and its quota of all-beverage licenses. The remaining tables provide illustrative examples of three different hypothetical communities that provide insights into the subtleties of the quota system.

A fundamental set of rules for the quota system relates to the issuance of additional all-beverage licenses in response to increases in population. Briefly, when the population of a quota area increases, thresholds are crossed that allow the issuance of additional licenses. More specifically, suppose the population of an area increases annually. When the number of licenses issued falls far enough below 1.43 times the number of licenses the quota formula allows that an additional license would not push the number of licenses above the 1.43 threshold, an additional license is issued for that area. Over the time span of most of our data, when the number of applicants for the additional license exceeded the number of additional licenses available, a lottery was held to determine the recipient of the new license. The winner of the additional license then went through an application process in which the Montana Departments of Revenue and Justice both vetted the applicant. If the applicant cleared this hurdle, then she was allowed to go to another quota area, buy a license, and “float” that license into her quota area. A license could be “floated out” of an area only if (1) the number of licenses in that area was at last 25 percent greater than the area’s formula-based number of licenses and (2) the removal of a license from that quota area does not cause its number of licenses to fall below the 25 percent threshold. Some of the impacts of these rules are clarified in the examples below.

A point of clarification is that the city beer and wine licenses have a similar population/quota schedule as in table 3. Suppose the population of a quota area increases to a level that calls for the issuance of a new beer and wine license. Then a lottery is held (assuming the number of applicants exceeds the number of new licenses issued), the lottery winner is vetted in an application process similar to the process for an all-beverage license, and if she is approved, she receives a new license. There are no 25 or 43 percent thresholds, and there is no floating of licenses across quota area boundaries. In addition, no beer and wine quota exists for unincorporated areas of counties; licenses in these areas are issued at the discretion of the Department of Revenue. The quota for restaurant beer and wine (RBW) licenses is linked closely to the quota for city beer and wine licenses and the process for obtaining a new RBW license is closely parallel to the process for beer and wine licenses. The hypothetical examples below help to clarify some of the impacts of the regulatory provisions

Table 4: Example: A Growing Community

Year	Population	Quota	Lic. Issued	1.25×Quota	1.43×Quota	New Lic.
1	11,000	11	16	13.75	15.73	0
2	12,200	12	17	15	17.16	1
3	14,000	13	18	16.25	18.59	1
4	16,000	14	20	17.5	20.02	2
5	16,499	14	20	17.5	20.02	0
6	17,000	15	21	18.75	21.45	1

A Growing Community

Consider first, a growing community—quota area (QA1), presented in table 4. In year one, the quota area has a population of 11,000 and the population/quota schedule dictates that the quota of licenses for that community is 11. The community, however, possibly for historical reasons, has 16 licenses. This number is greater than 1.43 times the schedule-based quota of licenses ($1.43 \times 11 = 15.73$), so QA1 does not qualify for an additional license. In year two, the population increases by enough that the quota formula dictates that its quota of licenses increases to 12. Now with its 16 licenses, QA1 has less than 1.43 times the schedule-based quota ($17.16 = 1.43 \times 12$). Moreover, the addition of one more license would still leave QA1 with fewer than 1.43 times the quota ($17 < 17.16$ licenses). Accordingly, the community is granted the right to float in a 17th license. In year three, QA1's population again increases by enough that the quota area is granted an additional license (its 13th). Following the same logic as in year two, QA1 can again float in a license—its 18th. In year four, population again increases by enough that the quota area's schedule-based quota of licenses increases to 14. This year, however, QA1 is allowed to acquire two new licenses ($18 + 2 = 20$) and still have fewer than 1.43 times its quota ($1.43 \times 14 = 20.2$) of licenses. In year five, population increases again, but not by enough for the area's quota of licenses to increase, so QA1 does not get to acquire an additional license. In year six, however, with additional population growth, the area's schedule-based license quota does increase, and it is allowed to float in an additional license.

A Shrinking Community

Table 5: Example: Shrinking Community

Year	Population	Quota	Lic. Issued	1.25×Quota	1.43×Quota	Lic. Floated
1	14,000	13	19	16.25	18.59	0
2	14,000	13	18	16.25	18.59	1
3	14,000	13	17	16.25	18.59	1
4	14,000	13	17	16.25	18.59	0
5	13,000	12	15	15	17.16	2
6	12,000	11	14	13.75	15.73	1

Consider next a shrinking community (QA2), presented in table 5. Initially, the quota area's population is 14,000, and it is allotted a quota of 13 licenses. Again, possibly for historical reasons, suppose there are 19 licenses issued in the community. Because the area's number of issued licenses is greater than 1.25×13, it could float out two licenses and still not fall below the $1.25 \times 13 = 16.25$ threshold. Assuming there are only two quota areas (QA1 and QA2), however, in year 1, there is no demand to float a license. In year two, however, QA1 is allowed to acquire a new license. Assuming there are several (possibly many) applicants for the new license, a lottery will be held, and the winner of the lottery will acquire the right to float a new license into QA1. If the value of licenses is substantially higher in QA1 than in QA2, the lottery winner will have an incentive to purchase a license in QA2 and float it into QA1. Suppose, for the present example, that this is the case. Note that license values in QA1 will fall and license values in QA2 will rise (see the third panel in figure 1), thereby reducing the differential in license values between the two quota areas. In year three, QA1 again is allowed to float in a license, and QA2 has another license that it is allowed to float out. In year four, the population growth in QA1 makes it eligible to float in two more licenses. If a license if floated out of QA2, however, its number of licenses issued would fall below 1.25 times its quota (16.25), so no licenses are transferred. In year five, the population of QA2 falls, its quota-based allocation of licenses falls, and it can float out two licenses and not fall below 1.25 times its quota-based allocation of quota. In this case, again depending on the relative values of licenses in QA1 and QA2, it is possible that two licenses are transferred from QA2 to QA1. In year six, the population of QA2 falls again, its quota-based allocation of licenses falls, and it is allowed to float out another license. Again, it is possible for a license to move from QA2 to QA1 (which is allowed to float in one more license in year six).

A Stable Community

Table 6: Example: Stable Community

Year	Population	Quota	Lic. Issued	1.25×Quota	1.43×Quota	New Lic.
1	11,000	11	11	13.75	15.73	0
2	11,000	11	15	13.75	15.73	4
3	11,000	11	15	13.75	15.73	0
4	11,000	11	15	13.75	15.73	0
5	11,000	11	15	13.75	15.73	0
6	11,000	11	15	13.75	15.73	0

Finally, consider QA3, a community with a stable population. Initially, the population/quota schedule allots the area 11 licenses and 11 are issued. The area is allowed to have up to 1.43 times its quota (15.73 licenses), implying that four licenses could float into the area. With no population growth, in the hypothetical shown, the area could maintain its 15 licenses indefinitely. One issue with this scenario is that if there are only three quota areas, then it is not clear four licenses will be floated into QA3 in year one. If, however, license values are higher in QA3 than in the other two areas, two licenses could be floated out of the other two areas without either area's license count falling below 1.25 times its quota. Conversely, if license values in QA3 are low enough, the four licenses may never be floated into QA3. Moreover, because QA3 has less than 1.25 times its schedule-based quota, licenses cannot be floated out of the area.

What do the hypotheticals in tables 4-6 suggest about the impacts of Montana's quota system? One feature of the system is that it allows for some movement of all-beverage licenses from low- to high-value areas, which leads to a more economically efficient distribution of the system's licenses across quota areas. At the same time, the system protects small communities (where license values are likely to be quite low) from losing all their licenses to larger communities where license values may be much higher. Note that this was in issue, for example, in the towns of Bozeman and Belgrade. Bozeman is a larger community, but the two town are located close enough to join quota areas. The state split the quota areas by statute, ostensibly to the benefit of Belgrade.[1].

In Montana, communities with the highest valued licenses include Billings, Missoula, and Bozeman. Of these, Bozeman is the community that is growing most quickly. Until the lottery system was replaced with an auction process in 2019, whenever the population-based schedule indicated that Bozeman was eligible for another all-beverage license, there were many applicants for the license. When the lottery winner floated a license into the Bozeman quota area, the license often came from the Butte or Anaconda quota areas, where (for historical reasons) there were substantial numbers of licenses in excess of the number allowed by the system[2]. In addition, whereas all-beverage licenses have sold for as much as $1,000,000 in recent years in Billings, Missoula, and Bozeman, a license can likely be purchased from a Butte license owner for less than $50,000.[3]

[1]See `https://www.bozemandailychronicle.com/news/belgrade-cheers-changes-to-liquor-license-laws/article_c5c391d8-b562-5b1e-93fa-72c17af34e38.html`

[2]In the most recent quota schedule available from the DOR, Butte has a schedule-based license quota of 26, with 78 licenses issued in the area. This implies that 45 licenses ($78–1.25 \times 26 = 45.5$, rounded down to 45) licenses are eligible to be floated out of the area. Bozeman, Missoula, and Billings, on the other hand all have (roughly) 1.43 times their schedule-based number of licenses.

[3]According to the Purchase Price Report posted by the DOR in June of 2020, the last three licenses sold in Butte had transaction prices of $30,000, $50,000, and $65,000.

Bibliography

Abreu, D. and Rubinstein, A. (1988), "The Structure of Nash Equilibrium in Repeated Games with Finite Automata", *Econometrica* **56**(6), 1259–1281.

Administrative Rules of Montana, Chapter 42.12: Liquor Licenses and Permits (2018).

Agranov, M. and Ortoleva, P. (2017), "Stochastic Choice and Preferences for Randomization", *Journal of Political Economy* **125**(1), 40–68.

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

Anderson, D. M., Crost, B. and Rees, D. I. (2018), "Wet Laws, Drinking Establishments and Violent Crime", *The Economic Journal* **128**(611), 1333–1366.

Aumann, R. J. (1981), Survey of Repeated Games, *in* "Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern", Bibliographisches Institut, pp. 11–42.

Azrieli, Y., Chambers, C. P. and Healy, P. J. (2018), "Incentives in Experiments: A Theoretical Analysis", *Journal of Political Economy* **126**(4), 1472–1503.

Banks, J., Olson, M. and Porter, D. (1997), "An experimental analysis of the bandit problem", *Economic Theory* **10**(1), 55–77.

Banks, J. and Sundaram, R. K. (1992), "Denumerable-Armed Bandits", *Econometrica* **60**(5), 1071–1096.

Banovetz, J. (2014), "An Economic Analysis of the Determinants of Montana Alcohol Retail License Prices", *Master's Thesis, Montana State University* .

Banovetz, J. (2020), "Preferences Over Exploration: Simple Bandits in the Lab", *Unpublished Manuscript* .

Barnard, C. H., Whittaker, G., Westenbarger, D. and Ahearn, M. (1997), "Evidence of Capitalization of Direct Government Payments into U.S. Cropland Values", *American Journal of Agricultural Economics* **79**(5), 1642–1650.

Becker, G. M., Degroot, M. H. and Marschak, J. (1997), “Measuring utility by a single-response sequential method”, *Behavioral Science* **9**(3), 226–232.

Bergemann, D. and Välimäki, J. (2001), “Stationary multi-choice bandit problems”, *Journal of Economic Dynamics and Control* **25**(10), 1585–1594.

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

Börgers, T. and Morales, A. (2004), “Complexity Constraints in Two-Armed Bandit Problems: An Example”, *Unpublished Manuscript* .

Boyce, J. R., Bruner, D. M. and McKee, M. (2016), “Strategic Experimentation in the Lab”, *Managerial And Decision Economics* **37**(6), 375–391.

Bradt, R. N., Johnson, S. M. and Karlin, S. (1956), “On Sequential Designs for Maximizing the Sum of n Observations”, *Annals of Mathematical Statistics* **27**(4), 1060–1074.

Brown, A. B., Rucker, R. R. and Thurman, W. N. (2007), “The End of the Federal Tobacco Program: Economic Impacts of the Deregulation of U.S. Tobacco Production”, *Review of Agricultural Economics* **29**(4), 635–655.

Buschena, D. E., Anderson, T. L. and Leonard, J. L. (2001), “Valuing Non-Marketed Goods: The Case of Elk Permit Lotteries”, *Journal of Environmental Economics and Management* **41**(1), 33–43.

Cairns, R. D. and Liston-Heyes, C. (1996), “Competition and regulation in the taxi industry”, *Journal of Public Economics* **59**(1), 1–15.

Carlson, C., Burtraw, D., Cropper, M. and Palmer, K. L. (2000), “Sulfur Dioxide Control by Electric Utilities: What Are the Gains from Trade?”, *Journal of Political Economy* **108**(6), 1292–1326.

Carpenter, C. (2007), “Heavy Alcohol Use and Crime: Evidence from Underage Drunk-Driving Laws”, *The Journal of Law & Economics* **50**(3), 539–557.

Cason, T. N. and Mui, V.-L. (2019), “Individual versus group choices of repeated game strategies: A strategy method approach”, *Games and Economic Behavior* **114**, 128–145.

Chakravorty, J. and Mahajan, A. (2014), Multi-Armed Bandits, Gittins Index, and Its Calculation, *in* “Methods and Applications of Statistics in Clinical Trials: Planning, Analysis, and Inferential Methods”, Vol. 2, Wiley, pp. 416–435.

Chaney, R. (1998), “Cabaret Licenses”, *The Missoulian* (July 16).

Chauvin, K. P. (2020), "Euclidean Properties of Bayesian Updating", *Unpublished Manuscript* .

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.

Chronicle Editorial Board (1997), "Legislature Should Pass Cabaret Bill", *Bozeman Daily Chronicle* (March 18).

Cohen, P., Hahn, R., Hall, J., Levitt, S. and Metcalfe, R. (2016), "Using Big Data to Estimate Consumer Surplus: The Case of Uber", *NBER Working Paper No. 22627* .

Cook, P. J., Ostermann, J. and Sloan, F. A. (2005), "Are Alcohol Excise Taxes Good For Us? Short and Long-Term Effects on Mortality Rates", *NBER Working Paper No. 11138* .

Cook, P. J. and Tauchen, G. (1982), "The effect of liquor taxes on heavy drinking", *The Bell Journal of Economics* **13**(2), 379–390.

Cooper, J. C. and Wright, J. D. (2012), "Alcohol, antitrust, and the 21st Amendment: An empirical examination of post and hold laws", *International Review of Law and Economics* **32**(4), 379–392.

Costello, G. A., Killian, J. H., and Thomas, K. R. (ed.) (2004), The Constitution of the United States of America: Analysis and Interpretation, Congressional Research Service, Washington, DC: Government Printing Office.

Cox, J. C. and Oaxaca, R. L. (2000), "Good News and Bad News: Search from Unknown Wage Offer Distributions", *Experimental Economics* **2**, 197–225.

Crocker, T. D. (1966), The Structuring of Atmospheric Pollution Control Systems, *in* "The Economics of Air Pollution", Norton, New York.

Crutchfield, J. A. (1961), "An Economic Evaluation of Alternative Methods of Fishery Regulation", *The Journal of Law & Economics* **4**, 131–143.

Curliss, E. (1999), "Beer Sales Could Boost Economy", *Great Falls Tribune* (March 11), 16.

Dal Bó, P. and Fréchette, G. R. (2011), "The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence".

Dal Bó, P. and Fréchette, G. R. (2019), "Strategy Choice in the Infinitely Repeated Prisoner's Dilemma", *American Economic Review* **109**(11), 3929–3952.

Dales, J. H. (1968), *Pollution, Property, and Prices*, University of Toronto Press, Toronto.

Deacon, R. T., Parker, D. P. and Costello, C. (2013), "Reforming Fisheries: Lessons from a Self-Selected Cooperative", *The Journal of Law & Economics* **56**(1), 83–125.

DeMay, D. (2015), "A Brewing Compromise?", *Bozeman Daily Chronicle* (January 25), D1.

Demsetz, H. (1982), "Barriers to Entry", *American Economic Review* **72**(1), 47–57.

Engle-Warnick, J., McCausland, W. J. and Miller, J. H. (2007), "The Ghost in the Machine: Inferring Machine-Based Strategies from Observed Behavior", *Unpublished Manuscript* .

Enke, B. and Zimmermann, F. (2017), "Correlation Neglect in Belief Formation", *The Review of Economic Studies* **86**(1), 313–332.

Eyster, E. and Weisacker, G. (2010), "Correlation Neglect in Financial Decision-Making", *DIW Berlin Discussion Paper No. 1104* .

Frankena, M. W. and Pautler, P. A. (1984), An Economic Analysis of Taxicab Regulation, Bureau of Economics, Federal Trade Commission, Washington, DC.

Frechette, G. and Yuksel, S. (2017), "Infinitely repeated games in the laboratory: four perspectives on discounting and random termination", *Experimental Economics* **20**(2), 279–308.

Gabaix, X. and Laibson, D. (2005), "Bounded Rationality and Directed Cognition", *Unpublished Manuscript* .

Gabaix, X., Laibson, D., Moloche, G. and Weinberg, S. (2006), "Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model", *American Economic Review* **96**(4), 1043–1068.

Gale, D. and Sabourian, H. (2005), "Complexity and Competition", *Econometrica* **73**(3), 739–769.

Gans, N., Knox, G. and Croson, R. (2007), "Simple Models of Discrete Choice and Their Performance in Bandit Experiments", *Manufacturing & Service Operations Management* **9**(4), 383–408.

Genesove, D. and Mayer, C. (2001), "Loss Aversion and Seller Behavior: Evidence from the Housing Market", *The Quarterly Journal of Economics* **116**(4), 1233–1260.

Gittins, J. C. (1979), "Bandit Processes and Dynamic Allocation Indices", *Journal of the Royal Statistical Society, Series B* **41**(2), 148–177.

Gokhale, J. and Tremblay, V. J. (2012), "Competition and Price Wars in the U.S. Brewing Industry", *Journal of Wine Economics* **7**(2), 226–240.

Goodwin, B. K. and Ortalo-Magné, F. (1992), "The Capitalization of Wheat Subsidies into Agricultural Land Values", *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie* **40**(1), 37–54.

Grainger, C. A. and Costello, C. J. (2014), "Capitalizing property rights insecurity in natural resource assets", *Journal of Environmental Economics and Management* **67**(2), 224–240.

Grainger, C. A. and Parker, D. P. (2013), "The Political Economy of Fishery Reform", *Annual Review of Resource Economics* **5**(1), 369–386.

Greiner, B. (2004), The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics, Working Paper Series in Economics 10, University of Cologne, Department of Economics.
URL: *https://ideas.repec.org/p/kls/series/0010.html*

Hahn, R. W. (1984), "Market Power and Transferable Property Rights", *The Quarterly Journal of Economics* **99**(4), 753–765.

Halac, M., Kartik, N. and Liu, Q. (2016), "Optimal Contracts for Experimentation", *The Review of Economic Studies* **83**(3), 1040–1091.

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

Halevy, Y. (2007), "Ellsberg Revisited: An Experimental Study", *Econometrica* **75**(2), 503–536.

Harrison, M. (2011), "Local Knowledge: The Spirit of Montana", *Big Sky Journal* .
URL: *http://www.bigskyjournal.com/articles/big-sky-journal/arts-2011/136/local-knowledge-the-spirit-of-montana.html*

Healy, J. (2013), "A Montana Loophole Leaves a Bitter Taste with Bar Owners", *New York Times* (April 20).
URL: *www.nytimes.com/2013/04/21/us/montana-loophole-leaves-bitter-taste-among-bar-owners.html*

Heaton, P. (2012), "Sunday liquor laws and crime", *Journal of Public Economics* **96**(1), 42–52.

Heidhues, P., Rady, S. and Strack, P. (2015), "Strategic experimentation with private payoffs", *Journal of Economic Theory* **159**, 531–551.

Hoelzemann, J. and Klein, N. (2019), "Bandits in the Lab", *Unpublished Manuscript* .

Holt, C. A. and Laury, S. K. (2002), "Risk Aversion and Incentive Effects", *American Economic Review* **92**(5), 1644–1655.

Hopcroft, J. E. and Ullman, J. D. (1979), *Introduction to Automata Theory: Languages and Computation*, Addison-Wesley.

Hu, Y., Kayaba, Y. and Shum, M. (2013), "Nonparametric learning rules from bandit experiments: The eyes have it!", *Games and Economic Behavior* **81**, 215 – 231.

Hudja, S. and Woods, D. (2019), "Behavioral Bandits: Analyzing the Exploration Versus Exploitation Trade-Off in the Lab", *Unpublished Manuscript* .

Jehiel, P. (2005), "Analogy-based expectation equilibrium", *Journal of Economic Theory* **123**(2), 81–104.

Jones, M. T. (2014), "Strategic complexity and cooperation: An experimental study", *Journal of Economic Behavior & Organization* **106**, 352–366.

Joskow, P. L., Schmalensee, R. and Bailey, E. M. (1998), "The Market for Sulfur Dioxide Emissions", *The American Economic Review* **88**(4), 669–685.

Kalai, E. and Solan, E. (2003), "Randomization and simplification in dynamic decision-making", *Journal of Economic Theory* **111**(2), 251–264.

Kalai, E. and Stanford, W. (1988), "Finite Rationality and Interpersonal Complexity in Repeated Games", *Econometrica* **56**(2), 397–410.

Keller, G. and Rady, S. (2010), "Strategic experimentation with Poisson bandits", *Theoretical Economics* **5**, 275–311.

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

Kirwan, B. (2009), "The Incidence of U.S. Agricultural Subsidies on Farmland Rental Rates", *Journal of Political Economy* **117**(1), 138–164.

Klein, N. and Rady, S. (2011), "Negatively Correlated Bandits", *The Review of Economic Studies* **78**(2), 693–732.

Kleiner, M. M. (2000), "Occupational Licensing", *Journal of Economic Perspectives* **14**(4), 189–202.

Kleiner, M. M. (2006), *Licensing Occupations: Ensuring Quality or Restricting Competition?*, W.E. Upjohn Institute for Employment Research, Kalamazoo, MI.

Larcombe, J. E. (1997), "If the Bill Doesn't Pass, the Public Will be Asked to Decide the Liquor License Quota Issue", *Great Falls Tribune* (February 16), 60.

Lovenheim, M. F. and Steefel, D. P. (2011), "Do blue laws save lives? the effect of sunday alcohol sales bans on fatal vehicle accidents", *Journal of Policy Analysis and Management* **30**(4), 798–820.

Marcoul, P. and Weninger, Q. (2008), "Search and active learning with correlated information: Empirical evidence from mid-Atlantic clam fishermen", *Journal of Economic Dynamics and Control* **32**(6), 1921–1948.

Markowitz, S. (2005), "Alcohol, drugs and violent crime", *International Review of Law and Economics* **25**(1), 20–44.

McCall, B. P. (1991), "A dynamic model of occupational choice", *Journal of Economic Dynamics and Control* **15**(2), 387–408.

McMillan, G. P. and Lapham, S. (2006), "Effectiveness of bans and laws in reducing traffic deaths: legalized sunday packaged alcohol sales and alcohol-related traffic crashes and crash fatalities in New Mexico", *American Journal of Public Health* **96**(11), 1944–1948.

Meyer, R. J. and Shi, Y. (1995), "Sequential Choice Under Ambiguity: Intuitive Solutions to the Armed-Bandit Problem", *Management Science* **41**(5), 817–834.

Miller, R. A. (1984), "Job Matching and Occupational Choice", *Journal of Political Economy* **92**(6), 1086–1120.

Miron, J. A. and Tetelbaum, E. (2009), "Does the Minimum Legal Drinking Age save Lives?", *Economic Inquiry* **47**(2), 317–336.

Montana Legislature: Revenue Oversight Committee (1980), Montana's Liquor License Quota System: A Report to the Forty-Seventh Legislature, Montana Legislative Council, Helena, MT.

Montana Legislature: Subcommittee on Judiciary (1974), Alcoholic Beverage Control Statutes: Interim Study, Montana Legislative Council, Helena, MT.

Montana Code Annotated, Title 16: Alcohol and Tobacco (2019).

Montana Code Annotated, Title 23: Parks, Recreation, Sports, and Gambling (2019).

Newell, R. G., Papps, K. L. and Sanchirico, J. N. (2007), "Asset Pricing in Created Markets", *American Journal of Agricultural Economics* **89**(2), 259–272.

Newell, R. G., Sanchirico, J. N. and Kerr, S. (2005), "Fishing quota markets", *Journal of Environmental Economics and Management* **49**(3), 437–462.

Newhouse, R. (2013), *Montana Beer: A Guide to Breweries in Big Sky Country*, American Palate, Charleston, SC.

Neyman, A. (1985), "Bounded complexity justifies cooperation in the finitely repeated prisoners' dilemma", *Economics Letters* **19**(3), 227–229.

Nicholson, W. and Snyder, C. (2014), *Microeconomic Theory: Basic Principles and Extensions*, 12 edn, South-Western, Cengage Learning, Mason, OH.

Nickerson, P. H. (1990), "Demand for the Regulation of Recreation: The Case of Elk and Deer Hunting in Washington State", *Land Economics* **66**(4), 437–447.

O'Donnell, M. K. (2017), "An Economic Analysis of the Effect of Uber on Taxi Medallions", *Master's Thesis, Montana State University* .

Oprea, R. D. (2020), "What Makes a Rule Complex", *Unpublished Manuscript* .

Powell, K. (2015), "How Taxis, Peanuts and Assault Rifles Get You a Martini for Dinner: Examining Peanut Quotas and Taxi Medallions in Consideration of Whether a Fifth Amendment Takings Claim is a Red Herring when Eliminating Alcohol License Quota Systems - Alcohol Licenses under a Quota System: Non-Compensable Regulatory License or Compensable Property Right in an Entrenched Legacy Market", *DePaul Business and Commercial Law Journal* **13**(3), 433–477.

Quinn, L. D. (1970), *Politicians in Business: A History of the Liquor Control System in Montana*, University of Montana Press, Missoula, MT.

Rabin, D. and Bookman, R. (2006), "Is It Last Call for the City?", *New York Times* (May 21).
URL: *www.nytimes.com/2006/05/21/opinion/nyregionopinions/21CIrabin.html*

Revised Code of Montana, Title 4: Alcoholic Beverages (1967).

Roistacher, E. and White, L. J. (1982), The Pricing of Taxicab Medallions in New York City, Mayor's Committee on Taxi Regulatory Issues, New York.

Romero, J. and Rosokha, Y. (2018), "Constructing strategies in the indefinitely repeated prisoner's dilemma game", *European Economic Review* **104**, 185–219.

Rosenberg, D., Solan, E. and Vieille, N. (2007), "Social Learning in One-Arm Bandit Problems", *Econometrica* **75**(6), 1591–1611.

Rubinstein, A. (1986), "Finite Automata Play the Repeated Prisoner's Dilemma", *Journal of Economic Theory* **39**(1), 83–96.

Rucker, R. R. and Thurman, W. N. (1990), "The Economic Effects of Supply Controls: The Simple Analytics of the U. S. Peanut Program", *The Journal of Law & Economics* **33**(2), 483–515.

Rucker, R. R., Thurman, W. N. and Sumner, D. A. (1995), "Restricting the Market for Quota: An Analysis of Tobacco Production Rights with Corroboration from Congressional Testimony", *Journal of Political Economy* **103**(1), 142–175.

Sabourian, H. (2004), "Bargaining and markets: complexity and the competitive outcome", *Journal of Economic Theory* **116**(2), 189–228.

Saffer, H. and Dave, D. (2002), "Alcohol consumption and alcohol advertising bans", *Applied Economics* **34**(11), 1325–1334.

Salant, Y. (2011), "Procedural Analysis of Choice Rules with Applications to Bounded Rationality", *American Economic Review* **101**(2), 724–748.

Sass, T. R. and Saurman, D. S. (1993), "Mandated Exclusive Territories and Economic Efficiency: An Empirical Analysis of the Malt-Beverage Industry", *The Journal of Law & Economics* **36**(1), 153–177.

Schlag, K. H. (1998), "Why Imitate, and If So, How?: A Boundedly Rational Approach to Multi-armed Bandits", *Journal of Economic Theory* **78**(1), 130–156.

Scrogin, D., Berrens, R. P. and Bohara, A. K. (2000), "Policy Changes and the Demand for Lottery-Rationed Big Game Hunting Licenses", *Journal of Agricultural and Resource Economics* **25**(2), 501–519.

Seim, K. and Waldfogel, J. (2013), "Public monopoly and economic efficiency: Evidence from the pennsylvania liquor control board's entry decisions", *American Economic Review* **103**(2), 831–62.

Simon, H. A. (1955), "A Behavioral Model of Rational Choice", *The Quarterly Journal of Economics* **69**(1), 99–118.

Smith, J. K. (1982), "An Analysis of State Regulations Governing Liquor Store Licensees", *The Journal of Law & Economics* **25**(2), 301–319.

Stavins, R. N. (1998), "What Can We Learn from the Grand Policy Experiment? Lessons from SO_2 Allowance Trading", *The Journal of Economic Perspectives* **12**(3), 69–88.

Stergionis, M. (1991), "Allowing Liquor: Alcohol-Permit Process Brings Out Fight in Some", *Great Falls Tribune* (June 2), 63.

Tietenberg, T. H. (2006), *Emissions Trading: Principles and Practice*, Resources for the Future, Washington, DC.

Toma, E. F. (1988), "State Liquor Licensing, Implicit Contracting, and Dry/Wet Counties", *Economic Inquiry* **26**(3), 507–524.

Townsend, R. E. (1990), "Entry Restrictions in the Fishery: A Survey of the Evidence", *Land Economics* **66**(4), 359–378.

Tremblay, V. J. and Okuyama, K. (2001), "Advertising Restrictions, Competition, and Alcohol Consumption", *Contemporary Economic Policy* **19**(3), 313–321.

Tullock, G. (1975), "The transitional gains trap", *The Bell Journal of Economics* **6**(2), 671–678.

Weersink, A., Clark, S., Turvey, C. G. and Sarker, R. (1999), "The Effect of Agricultural Policy on Farmland Values", *Land Economics* **75**(3), 425–439.

Weitzman, M. L. (1979), "Optimal Search for the Best Alternative", *Econometrica* **47**(3), 641–654.

Wilen, J. E. (1988), "Limited Entry Licensing: A Retrospective Assessment", *Marine Resource Economics* **5**(4), 313–324.

Wilson, A. (2014), "Bounded Memory and Biases in Information Processing", *Econometrica* **82**(6), 2257–2294.