ADAPTIVE BEHAVIOR IN CONTINUOUS TIME

A Dissertation

by

DANIEL GRAYDON STEPHENSON

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Alex Brown
Committee Members,	Catherine Eckel
	Daniel Fragiadakis
	Marco Palma
Head of Department,	Timothy Gronberg

May 2017

Major Subject: Economics

ABSTRACT

This research investigates population-level behavioral dynamics, how they affect the emergence of self-enforcing conventions, and how they can aid in the design of mechanisms to better achieve policy goals. It seeks to identify why long-run behavior approaches equilibrium in some environments, and fails to do so in others. This question is important because equilibrium is frequently employed to make policy recommendations, so it is necessary to identify when it provides reliable predictions. Further, many strategic environments only reach equilibrium in the long run, so modeling the short run process from which long run equilibria eventually emerge can help answer important policy-relevant questions. To answer these questions this research experimentally investigates behavioral dynamics in continuous-time strategic environments. We find that adaptive models provide remarkably powerful tools for identifying which strategic environments exhibit convergence to equilibrium and for characterizing disequilibrium dynamics in non-convergent strategic environments.

DEDICATION

To anyone and everyone who ever asked "Why?"

ACKNOWLEDGMENTS

I would like to thank thank Alexander L. Brown, Catherine E. Eckel, Daniel E. Fragidakis, Marco A. Palma, Ryan Oprea, Rodrigo A. Velez, and the experimental research team at Texas A&M for their continual support and feedback on this research. I am also grateful for the insightful comments from attendees of the 2014 North American Meeting of the Economic Science Association, the 2014 European Meeting of the Economic Science Association, and the 2016 Texas Economic Theory Camp. I also appreciate the generous support of this research by Texas A&M University and the National Science Foundation.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professors Alexander L. Brown, Catherine E. Eckel, and Daniel E. Fragidakis of the Department of Economics and Professor Marco A. Palma of the Department of Agricultural Economics. All work for the dissertation was completed independently by the student.

Funding Sources

Graduate study was supported by a fellowship from Texas A&M University, a doctoral dissertation research improvement grant (SES-1458541) from the National Science Foundation in 2014, the 2015 S. Charles Maurice Graduate Fellowship, and the 2016 John Van Huyck Graduate Fellowship.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION: COORDINATION AND CONVERGENCE

1.1 Motivation

Classical game theory describes Nash equilibrium as the outcome of introspective reasoning prior to play by perfectly rational agents.[1] In contrast, evolutionary game theory explains Nash equilibrium as a self enforcing convention that emerges as the long run outcome of dynamic interaction among large populations of boundedly rational agents.[2] While evolutionary models are frequently invoked as equilibrium selection tools, their ability to explicitly characterize disequilibrium dynamics is frequently overlooked. In addition to the identification and classification of equilibria, evolutionary game theory also provides theoretical models that explicitly describe the dynamic process from which Nash equilibrium emerges.[3] These dynamic evolutionary models provide a theoretical framework for addressing aspects of disequilibrium behavior which are inaccessible to classical models that exclusively identify and classify equilibria.

To test theoretical predictions from these evolutionary models, this chapter experimentally investigates dynamic behavior in a class of attacker defender population games that exhibit identical equilibrium predictions but starkly different evolutionary predictions. In the control treatment, subjects play a conventional attacker-defender population game with two classes of equally valuable targets. Here attackers prefer to attack the class of targets that is least likely to be defended, but defenders prefer to defend the class of targets that is most likely to be attacked. Evolutionary dynamics predict global convergence to equilibrium in these conventional attacker defender games.

[1] A thorough discussion of the epistemic conditions for Nash equilibrium in classical game theory can be found in Aumann and Brandenburger [5].

[2] This interpretation of Nash equilibrium as the long run outcome of an adaptive adjustment process is not a recent innovation. Notably, it was employed by Cournot [6] and Nash [7].

[3] See Sandholm [8] for more on the use of dynamic models in evolutionary game theory.

The coordination treatment adds weak intrapopulation coordination incentives, giving attackers an incentive to coordinate their attacks on a class of targets and giving defenders an incentive to coordinate their defenses on a class of targets. In these coordinated attacker-defender games, evolutionary dynamics predict divergence from equilibrium and the emergence of stable limit cycles. Under both experimental treatments, subjects adjusted their strategy continuously and earned continuous flow payoffs. This continuous-time experimental design is consistent with the continuous-time structure of dynamic evolutionary models and allows for the observation of long term behavioral phenomena that may be difficult to observe in a discrete-time setting.[4]

The attacker-defender games investigated by this chapter provide a remarkably powerful test of evolutionary game theory because they cleanly separate evolutionary predictions from those of other models.[5] A variety of solution concepts yield identical predictions under both treatments, including Nash equilibrium, logit quantal response equilibrum [10], level-k [11, 12], cognitive hierarchy models [13], and the time average of the Shaply polygon (TASP) [14]. The limitation of such models is that they exclusively identify and classify strategy profiles satisfying their respective behavioral criteria. In contrast, evolutionary models explicitly describe the dynamic process from which such strategy profiles emerge.

Unlike the aforementioned models, evolutionary dynamics predict markedly different behavior under each treatment. Specifically, evolutionary models predict convergence to equilibrium under the control treatment, but predict the emergence of divergent limit cycles under the coordination treatment. This clear division between theoretical predictions from evolutionary models and those from other behavioral models allows this particular class of attacker-defender games to serve as an efficient testing structure for evolutionary game

[4] See [9] for an example of such phenomena.

[5] See section 1.2.4 for more details regarding these theoretical predictions.

theory.

Consistent with theoretical predictions from evolutionary models, the observed subject behavior was tightly clustered around Nash equilibrium under the control treatment but was widely dispersed from Nash equilibrium under the coordination treatment. In opposition to the Nash equilibrium predictions, these results indicate that coordination incentives can lead to autocorrelated attacks in attacker-defender population games, thus making attacks more predictable.

This chapter tests three widely employed evolutionary models,[6]: the best response dynamic Gilboa and Matsui [15], the Smith dynamic Smith [16] and the logit dynamic Fudenberg and Levine [17], each of which is derived from a distinct set of underlying behavioral assumptions. Both the best response dynamic and the Smith dynamic maintain the conventional assumption of sign-preservation, namely that agents exclusively switch from lower performing strategies to higher performing strategies.[7] In contrast, the logit dynamic is not sign-preserving as it describes agents who sometimes switch from higher performing strategies to lower performing strategies.[8]

Under the coordination treatment, sign-preserving dynamics predict that behavioral limit cycles will approach the boundary of the state space.[9] In contrast, dynamics that violate sign-preservation predict that behavioral limit cycles will remain strictly in the interior of the state space.[10] This distinction between the theoretical predictions from different classes of evolutionary models provides a powerful test for the widely maintained assumption of sign-preservation. In contrast to theoretical predictions from sign-preserving dynamics, the observed cycles remained strictly in the interior of the state

[6] See [8] for more on these models.

[7] See section 1.2.3 for more details regarding the definition of sign-preservation.

[8] The logit dynamic does satisfy the weaker assumption of sign-correlation, namely that agents are more likely to switch from lower performing strategies to higher performing strategies.

[9] Here the state space denotes the set of all possible mixed strategy profiles. See section 1.2.1 for more details.

[10] See section 1.2.4 for a graphical depiction of these theoretical predictions.

space under the coordination treatment. This violation of sign-preservation suggests that the wider class of sign-correlated dynamics deserves further attention and may provide a superior characterization of human behavior over the more conventional class of sign-preserving dynamics.

The remainder of this chapter is organized as follows: Subsection 1.1.1 discusses the related literature. Section 1.2 describes the theoretical framework of dynamic evolutionary models and their predictions in attacker-defender games. Section 1.3 provides a thorough description of our experimental design and the procedures employed. Section 1.4 indicates the hypothesis that are tested in this experiment. Section 1.5 presents the main results and section 1.6 concludes.

1.1.1 Related Literature

Unlike previous experimental investigations of evolutionary game theory, this chapter tests the predictions of continuous sign-preserving evolutionary dynamics against those of discontinuous dynamics, sign-correlated dynamics, and Nash equilibrium. This chapter obtains a clean test of these theoretical predictions by experimentally investigating attacker-defender population games that yield distinct evolutionary predictions but yield the same predictions from a variety of a variety of other widely employed solution concepts. Specifically, this particular class of games yields identical predictions from Nash equilibrium, logit quantal response equilibrum [10], level-k [11, 12], cognitive hierarchy models [13], and TASP [14]. In contrast, evolutionary dynamics yield strikingly different predictions for different games in this class. Further, different types of evolutionary models yield different predictions for the same game. This division between the theoretical predictions of different evolutionary models provides a powerful test of their respective underlying behavioral assumptions.

This chapter contributes to a growing body of experimental research testing evolution-

ary game theory in laboratory experiments. Early experimental investigation of evolutionary game theory focused on testing evolutionary models of equilibrium selection. In particular, [18] observed strong convergence to inefficent pareto dominated equilibria in minimum effort games with multiple equilibria. They suggested that the emergence of these particular equilibria was driven by the presence of strategic uncertainty rather than previously traditional equilibrium refinement methods. [19] later formalized these observations through the use of evolutionary stability criteria.

[20] were among the first to employ continuous-time laboratory procedures in testing evolutionary models. They considered Hawk-Dove population games with with two asymmetric Nash equilibria and one symmetric Nash equilibrium. In accordance with evolutionary selection criteria, they found that subjects converge to an asymmetric equilibrium under two population matching protocols, but converge to a symmetric equilibrium under one population matching protocols. Their experiment provided empirical evidence for the ability of evolutionary models to help solve the equilibrium selection problem in games with multiple Nash equilibria. In contrast, this chapter investigates the ability of explicitly dynamic evolutionary models to predict dynamic disequilibrium behavior in games with unique Nash equilibria.

[21] implemented laboratory procedures where subjects adjusted mixed strategies continuously over time. They observed cyclical behavior in three rock-chapter-scissors population games: one with a stable equilibrium and two with unstable equilibrium. In these games, they tested point predictions from the time average of the Shaply polygon (TASP) [14] against point predictions from Nash equilibrium. In contrast, this chapter tests predictions from sign-preserving evolutionary dynamics against those of Nash equilibrium, discontinuous dynamics, and sign-correlated dynamics.

[22] investigated continuous-time population games with non-linear payoff functions and continuous strategy spaces. They test evolutionary models of imitative behavior

against evolutionary models of optimization behavior in all-pay auction population games exhibiting unique mixed-strategy Nash equilibria. In contrast, this chapter tests predictions from sign-preserving evolutionary dynamics against those of discontinuous dynamics and sign-correlated dynamics.

1.2 Theory

In this chapter we test theoretical predictions from dynamic evolutionary models in attacker-defender population games. Population games provide a theoretical framework for the analysis of repeated strategic interaction between large numbers of agents. In evolutionary game theory, population games are employed to model a wide variety of strategic environments including market competition [23], highway traffic [16], and tax compliance [24].

1.2.1 Population Games

A population game is played by a *society* composed of one or more *populations* $p \in P = \{1, \ldots, \mathfrak{p}\}$. Each population p contains a continuum of agents who choose pure strategies from the set $S^p = \{1, \ldots, n^p\}$. The proportion of population p that employs the pure strategy i is denoted by $x_i^p \in [0, 1]$. Accordingly, a *population state* $x^p = (x_1^p, \ldots, x_{n^p}^p)$ indicates the proportion of population p that employs each pure strategy $i \in S^p$. Further, a *social state* $x = (x^1, \ldots, x^{\mathfrak{p}})$ describes the state of each population $p \in P$. The payoff to an agent in population p who employs pure strategy $i \in S^p$ is given by the payoff function $\pi_i^p(x)$.

A social state x is said to be a *Nash equilibrium* if no agent in any population can increase her payoff by unilaterally adjusting her strategy. More formally, a social state x is a *Nash equilibrium* if, for every population $p \in P$ and every pair of pure strategies $i, j \in S^p$, $x_i^p > 0$ implies $\pi_i^p(x) \geq \pi_j^p(x)$. A Nash equilibrium is said to be *evolutionarily stable* if, whenever any sufficiently small proportion of agents deviates to some alternate

social state, then under the resulting social state, agents who stay at the Nash equilibrium do better, on average, then agents who deviated to the alternate social state. More formally, A Nash equilibrium x is said to be *evolutionarily stable* if there exists some $\varepsilon > 0$ such that for any alternate social state $y \neq x$ and any proportion $\alpha \in (0, \varepsilon)$ we have $x \cdot \pi(z) > y \cdot \pi(z)$ where $z = \alpha y + (1 - \alpha)x$. This evolutionary stability criterion was originally developed for games played by a monomorphic population of agents by [25] and was extended to the case of multiple polymorphic populations by [26].

1.2.2 Attacker-Defender Population Games

This chapter considers attacker-defender population games[11] played by a population of attackers A and a population of defenders D. Each defender chooses to defend one of her two possible targets $S^D = \{1, 2\}$. Each attacker is randomly matched with a defender and chooses to attack one of her two possible targets $S^A = \{1, 2\}$. Attackers prefer to attack the target that is least likely to be defended and defenders prefer to defend the targets that is most likely to be attacked. The expected payoff to a defender is proportional to the share of her chosen target in the attacker population. Conversely, the expected payoff to an attacker is proportional to the share of her unchosen strategy in the defender population. If both populations are evenly divided between target 1 and target 2, then both targets yield equal payoffs and no agent has an incentive to switch targets, so this social state is a Nash equilibrium. Formally, the payoffs functions in this standard attacker-defender game are given by

$$\begin{aligned} \pi_1^A(x) &= Mx_2^D & \qquad \pi_2^A(x) &= Mx_1^D \\ \pi_1^D(x) &= Mx_1^A & \qquad \pi_2^D(x) &= Mx_2^A \end{aligned} \tag{1.1}$$

This chapter also considered attacker-defender population games with weak intrapopulation coordination invectives. In such games, attackers have an incentive to coordinate

[11] See [27] and [28] for more on attacker-defender games.

their attacks on a target and defenders have an incentive to coordinate their defense on a target. Formally, the payoff functions in these coordinated attacker-defender games are given by

$$\begin{aligned} \pi_1^A(x) &= Mx_2^D + Cx_1^A & \qquad \pi_2^A(x) &= Mx_1^D + Cx_2^A \\ \pi_1^D(x) &= Mx_1^A + Cx_1^D & \qquad \pi_2^D(x) &= Mx_2^A + Cx_2^D \end{aligned} \tag{1.2}$$

Here the parameter M denotes the strength of the attacker-defender incentives while the parameter C denotes the strength of the coordination incentives. So long as the coordination incentives remain weaker than the attacker-defender incentives ($M > C > 0$), this game still has a unique Nash equilibrium social state under which both populations are evenly divided between their two pure strategies. Hence the equilibrium predictions in this coordinated attacker-defender game remain unchanged from those of the standard attacker-defender game.

1.2.3 Evolutionary Dynamics

This paper considers theoretical predictions from dynamic evolutionary models describing the diachronic process of behavioral adjustment as a system of differential equations. This method of describing evolutionary dynamics in population games via revision protocols was originally introduced by Bjornerstedtt and Weibull [29] and has since been employed by numerous researchers including Sandholm [30] and Bulò and Bomze [31]. A *revision protocol* $\rho_{ij}^p(x)$ indicates the rate at which agents in population p switch from strategy i to strategy j as a function of the social state x. The induced *evolutionary dynamics* are described by by the non-linear differential equation:

$$\dot{x}_i^p = \sum_{j \in S^p} x_j^p \rho_{ji}^p(x) - x_i^p \sum_{j \in S^p} \rho_{ij}^p(x) \tag{1.3}$$

The first summation in this expression describes the inflow of agents into strategy i from other strategies, while the second summation describes the outflow of agents from

strategy i onto other strategies. The difference between these two summations yields the net rate of change $\dot{x}_i^p$ in the proportion of population p employing the pure strategy i. An evolutionary dynamic is said to be *sign-preserving* when agents exclusively switch from lower performing strategies to higher performing strategies. More formally, an evolutionary dynamic is said to be is said to be *sign-preserving* if $\pi_j^p(x) > \pi_i^p(x) \iff \rho_{ij}^p(x) > 0$. In contrast, an evolutionary dynamic is said to be *sign-correlated* if agents switch from lower performing strategies to higher performing strategies more rapidly than they switch from higher performing strategies to lower performing strategies. More formally, an evolutionary dynamic is said to be *sign-correlated* if $\pi_j^p(x) > \pi_i^p(x) \implies \rho_{ij}^p(x) > \rho_{ji}^p(x)$.

Conventional solution concepts such as Nash equilibrium [7], logit quantal response equilibrum [10], level-k [11, 12], cognitive hierarchy models [13], and TASP [14] exclusively identify and classify the social states that satisfy their respective criteria. In contrast, evolutionary dynamics explicitly describe the dynamic process of behavioral adjustment from which such social states may emerge. Since evolutionary dynamics explicitly model the process from which equilibria emerge, these models are uniquely positioned to characterize behavior out of equilibrium. The remainder of this section will describe three widely employed evolutionary dynamics, identify their underlying behavioral assumptions, and highlight the key differences between their theoretical predictions for the coordinated attacker-defender game.

In their analysis of the social stability of Nash equilibrium, Gilboa and Matsui [15] consider an evolutionary dynamic under which agents myopically switch to their best response under the current social state. This best response dynamic is also closely related to the discrete-time fictitious play dynamic discussed by Brown [32]. Formally, the conditional switch rate from the pure strategy i to the pure strategy j in population p

under the best response dynamic is given by

$$\rho_{ij}^{p}(x) = \begin{cases} 1 & \text{if } j \in \underset{k \in S^p}{\text{argmax}}\ \pi_k^p(x) \\ 0 & \text{otherwise} \end{cases} \tag{1.4}$$

Since the best response dynamic assumes that agents myopically switch to their current best response, any fixed point of the best response dynamic must have every agent simultaneously best responding to the strategies selected by other agents. Hence every Nash equilibrium is a fixed point of the best response dynamic and every fixed point of the best response dynamic is a Nash equilibrium.

Figure 1.1a illustrates the conditional switch rate under the best response dynamic in games where each agent has two pure strategies. The vertical axis indicates the conditional switch rate from the pure strategy i to the pure strategy j and the horizontal axis indicates the difference in payoffs between agents in population p who employ strategy j and those who employ strategy i. Note the abrupt change in the conditional switch rate at the point where both strategies yield equal payoffs. This jump illustrates the discontinuity in the conditional switch rate under the best response dynamic.

In his analysis of disequilibrium highway traffic congestion dynamics, Smith [16] considers an evolutionary dynamic under which agents switch from lower performing strategies to higher performing strategies at a rate proportional to the difference in payoffs. Hence the conditional switch rate from the pure strategy i to the pure strategy j in population p is given by

$$\rho_{ij}^{p}\left(x\right) = \max\left\{0, \pi_j^p\left(x\right) - \pi_i^p\left(x\right)\right\}, \tag{1.5}$$

The Smith dynamic falls into the category of *pairwise comparison* dynamics as the

conditional switch rate from strategy i to strategy j in population p depends exclusively on the pairwise comparison $\pi_j^p(x) - \pi_i^p(x)$. Like the best response dynamic, the Smith dynamic is sign-preserving because the conditional switch rate $\rho_{ij}^p(x)$ positive if and only if $\pi_j^p(x) > \pi_i^p(x)$. Sandholm [30] proves that the fixed points of every sign-preserving pairwise comparison dynamic exactly coincide with the Nash equilibria of the corresponding population game. Hence every fixed point of the Smith dynamic is a Nash equilibrium and every Nash equilibrium is a fixed point of the Smith dynamic.

Figure 1.1b illustrates the conditional switch rate under the Smith dynamic in games with two pure strategies for each population. The vertical axis indicates the conditional switch rate from the pure strategy i to the pure strategy j and the horizontal axis indicates the difference in payoffs between agents in population p who employ strategy j and those who employ strategy i. Unlike the best response dynamic, conditional switch rate is continuous in payoffs under the Smith dynamic, so small changes in payoffs produce small changes in the conditional switch rate. In contrast, a minute change in payoffs can produce a sudden and drastic change conditional switch rates under the best response dynamic.

Fudenberg and Levine [17] describe describe an evolutionary dynamic under which agents switch towards their perceived best response under a noisy perturbation of their payoffs. This logit dynamic is closely related the logit quantal response equilibrium described by McKelvey and Palfrey [10]. Under the logit dynamic, agents are more likely to switch to strategies that yield higher payoffs. Formally, the conditional switch rate from the pure strategy i to the pure strategy j in population p under the logit dynamic is given by

$$\rho_{ij}^p(x) = \frac{\exp(\eta^{-1}\pi_j^p(x))}{\sum_{k \in S^p} \exp(\eta^{-1}\pi_k^p(x))}. \tag{1.6}$$

Here η denotes the noise level in an agent's perception of payoffs. As η becomes small, agents perceive their payoffs more precisely, and the conditional switch rate approaches

that of the best response dynamic. Conversely, as η becomes large, agents become increasingly insensitive to payoff differences and the conditional switch rate approaches uniformly random behavior.

Unlike the Smith dynamic and the best response dynamic, the logit dynamic is not sign-preserving, but it is sign-correlated since agents are more likely to switch from lower performing strategies to higher performing strategies. Consequently, some population games have Nash equilibria that are not fixed points of the logit dynamic. Specifically, the fixed points of the logit response dynamic correspond to the set of logit quantal response equilibria, which do not always coincide with Nash equilibria.

Figure 1.1c illustrates the conditional switch rate under the logit dynamic in games with two pure strategies for each population. The vertical axis indicates the conditional switch rate from the pure strategy i to the pure strategy j and the horizontal axis indicates the difference in payoffs between agents in population p who employ strategy j and those who employ strategy i. Like the Smith dynamic, but unlike the best response dynamic, the conditional switch rate is continuous in payoffs under the logit dynamic.

Each of the aforementioned evolutionary dynamics exhibit markedly different behavioral assumptions, but each dynamic also shares some key assumptions with others. For instance, both the logit dynamic and the Smith dynamic predict that the conditional switch rate will be continuous in payoffs, so small changes in the payoffs will always result in correspondingly small changes in conditional switch rate under these models. In contrast, the best response dynamic is discontinuous in payoffs, so small changes in payoffs can produce disproportionately large changes in the conditional switch rate under the best response dynamic.

Both the Smith dynamic and the best response dynamic are sign-preserving, meaning that agents who follow these dynamics will exclusively switch from lower performing strategies to higher performing strategies. Consequently, the fixed points of the Smith

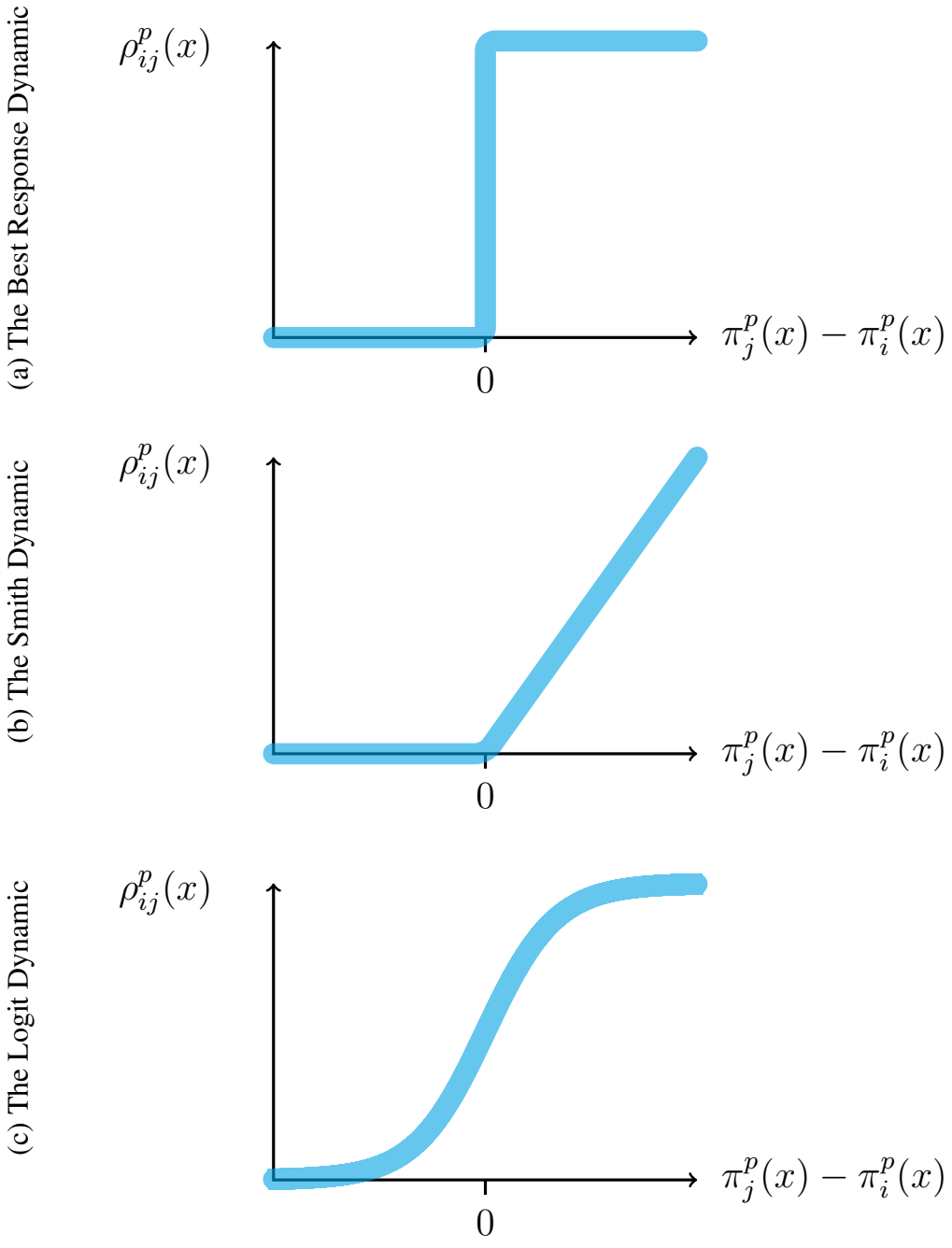

Figure 1.1: Conditional switch rates in populations with two pure strategies

dynamic and the best response dynamic reliably coincide with Nash equilibria. In contrast, the logit dynamic is not sign-preserving but it is sign-correlated, meaning that agents who follow this dynamic are more likely to switch from lower performing strategies to higher performing strategies. Consequently, the fixed points of the logit dynamic frequently fail to coincide with Nash equilibrium.

1.2.4 Evolutionary Dynamics in Attacker-Defender Games

Figure 1.2a depicts the best response dynamics for both variations of the attacker defender game. Similarly, figure 1.2b depicts the Smith dynamics, and figure 1.2c depicts logit dynamics. The graph on the left side of each figure illustrates the theoretical predictions of each evolutionary dynamic under the standard attacker-defender game, and the graph on the right side of each figure illustrates the theoretical predictions under the coordinated attacker-defender game. The horizontal axis of each graph indicates the proportion of attackers that attack target 2 and the vertical axis of each graph indicates the proportion of defenders that defend target 2. Each solid line in these figures illustrates the predicted path of the social state starting from a different set of initial conditions.

Recall that defenders in these games have an incentive to defend the target that is most likely to be attacked, while attackers have an incentive to attack the target that is least likely to be defended. Consequently, dynamic evolutionary models consistently predict the presence of clockwise cycles in these games. However, in the conventional attacker-defender game, dynamic evolutionary models predict that the social state will gradually spiral inwards towards equilibrium. Whereas, in the coordinated attacker-defender game, dynamic evolutionary models predict that the social state will converge to a stable limit cycle that indefinitely orbits the Nash equilibrium.

Different evolutionary models yield significantly different predictions regarding the shape of the limit cycles. As illustrated in figure 1.2a, the best response dynamic predicts

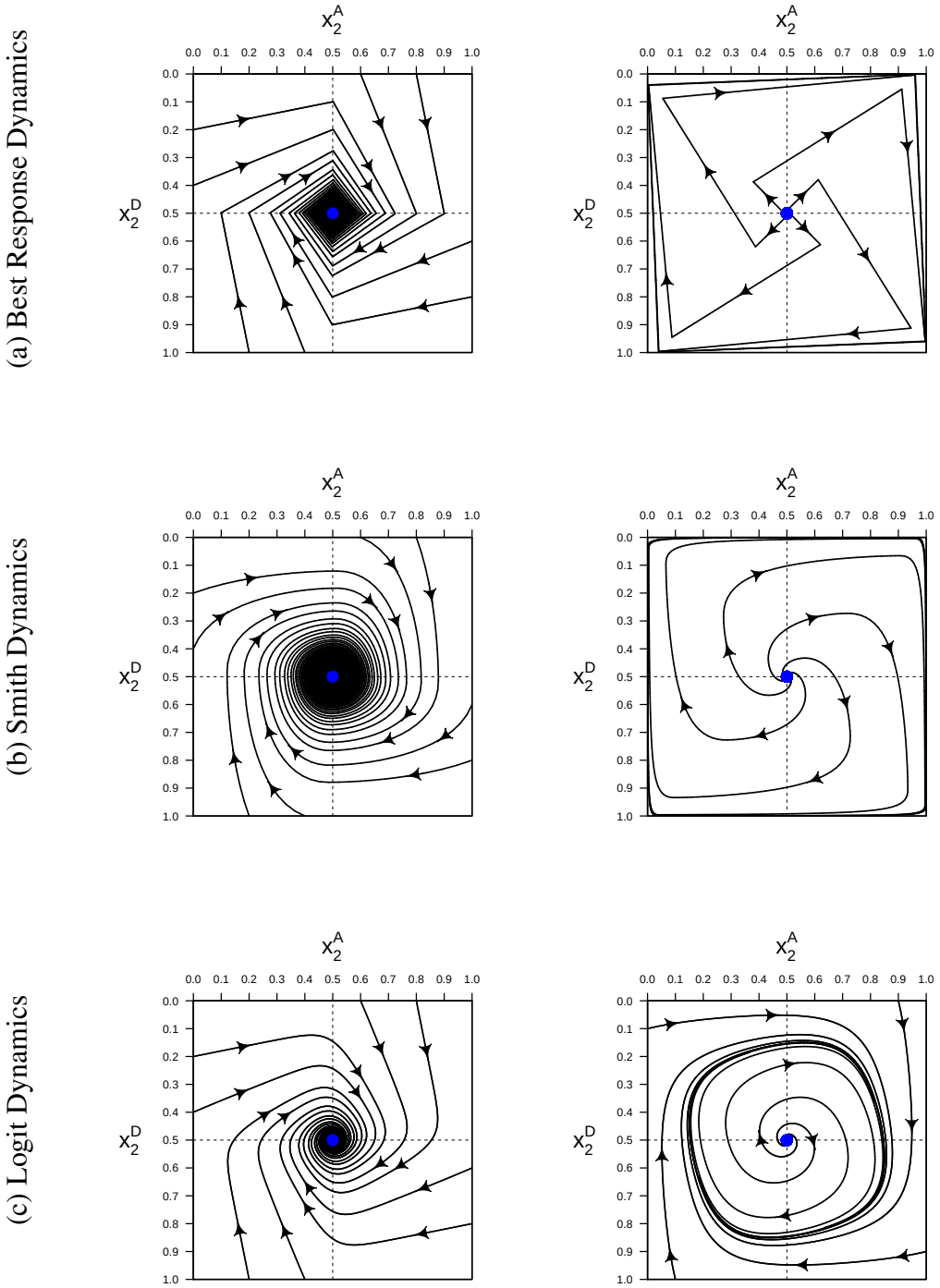

Figure 1.2: Evolutionary dynamics in the attacker defender game. Theoretical predictions for the standard attacker-defender game where $M = 5$ and $C = 0$ are shown on the left. Theoretical predictions for the coordinated attacker-defender game where $M = 2.6$ and $C = 2.4$ are shown on the right. These parameters are selected to reflect those implemented in the experimental treatments.

that agents will exhibit sudden sharp reversals in behavior when the best response changes. Accordingly, the best response dynamic predicts that the stable limit cycle orbiting the Nash equilibrium in the coordinated attacker-defender game will take a rectangular shape. In contrast, as illustrated in figures 1.2b and 1.2c, the Smith dynamic and the logit dynamic predict that agents will exhibit smooth gradual changes in behavior as payoffs change smoothly over time. Figure 1.2c depicts the predictions of the logit dynamic for $\eta = 0.08$. As η becomes small, subjects become increasingly sensitive to small differences in payoffs, so the predicted path of the social state becomes increasingly similar to that of the best response dynamic, with increasingly sharp changes in switching behavior. However, as η increases, agents become more noisy in their behavior and switching behavior becomes less sensitive to relative payoff differences.

When the population of attackers is near equilibrium, the targets are nearly equally to be attacked, so the coordination incentives can outweigh the defense incentives for defenders. Conversely, when the population of defenders is near equilibrium, the targets are nearly equally likely to be attacked, so the coordination incentives can outweigh the attack incentives for attackers. Consequently, as illustrated in figures 1.2a and 1.2b the limit cycles predicted by sign-preserving dynamics in the coordinated attacker defender game approach the boundary of the state space. In contrast, non-sign-preserving dynamics predict that agents exhibit some noise in their behavior, sometimes switching from higher performing strategies back to lower performing strategies. Thus, when any population coordinates on a single target, noisy switching behavior pushes the population state back towards a distribution of strategies in the interior of the population state space. Consequently, as illustrated in figure 1.2c, the limit cycles predicted by non-sign-preserving dynamics remain strictly in the interior of the state space.

1.3 The Experiment

This experiment varies coordination incentives across two attacker-defender population games. Both treatment exhibit identical equilibrium predictions, but dynamic evolutionary models yield strikingly different predictions regarding the dynamic disequilibrium behavior in each treatment.

1.3.1 Experimental Design

This study implements two experimental treatments: a control treatment and a coordination treatment. In the control treatment, subjects played a conventional two-population attacker defender game as described in section 1.2.2. Subjects in the defender group earned \$5.00 per minute times the proportion of attackers that they defended against. Similarly, subjects in the attacker group earned \$5.00 per minute times the proportion of defenders that did not defend againt their attack.[12] Accordingly, under the control treatment, earnings per minute were determined by

$$\begin{aligned} \pi_1^A(x) &= \$5.00 x_2^D & \qquad \pi_1^D(x) &= \$5.00 x_1^A \\ \pi_2^A(x) &= \$5.00 x_1^D & \qquad \pi_2^D(x) &= \$5.00 x_2^A \end{aligned} \tag{1.7}$$

where x_i^g denotes the proportion of group g employing the pure strategy i.

In the coordination treatment, each subject faced intrapopulation coordination incentives in addition to the payoffs from the standard attacker-defender game. As discussed in section 1.2.2, these coordination incentives do not effect the Nash equilibrium predictions, but they do effect the disequilibrium dynamics predicted by evolutionary models. Here defenders earned \$2.60 per minute times the proportion of attackers that they successfully defended against and attackers earned \$2.60 per minute times the proportion of defenders

[12] Both [20] and [22] similarly employ mean-matching protocols and have their subjects earn payoffs continuously over time. The use of mean-matching and continuous flow payoffs are a standard procedures in the experimental literature investigating evolutionary game theory.

who failed to defend against their attack. In addition, attackers earned \$2.40 per minute times the proportion of their fellow attackers that they coordinated with and defenders earned \$2.40 per minute times the proportion of their fellow defenders that they coordinated with. Formally, in the coordination treatment, earnings per minute were determined by

$$\begin{aligned} \pi_1^A(x) &= \$2.60x_2^D + \$2.40x_1^A & \qquad \pi_2^A(x) &= \$2.60x_1^D + \$2.40x_2^A \\ \pi_1^D(x) &= \$2.60x_1^A + \$2.40x_1^D & \qquad \pi_2^D(x) &= \$2.60x_2^A + \$2.40x_2^D \end{aligned} \tag{1.8}$$

These parameters are selected to equalize the equilibrium expected payoff across treatments, so that Nash equilibrium not only predicts identical behavior across both treatments, but also identical payoffs across treatments. Further, these parameters equalize the total strength of incentives across treatments, so that evolutionary dynamics predict equal adjustment speeds in both treatments.

1.3.2 Experimental Procedures

Each experimental session was conducted with twenty subjects and lasted for about thirty minutes. On average, each subject earned a total of \$18.63, including a five dollar show-up payment. We employ a between-subjects design, so each subject participated in one and only one experimental session. Two sessions were conducted with each of the two experimental treatments, for a total of four sessions with 80 distinct experimental subjects. At the beginning of each session, subjects were randomly divided into two equally sized population groups, so each group of attackers consisted of exactly 10 subjects and each group of defenders consisted of exactly 10 subjects. Subjects stayed in the same population group for the duration of a session. Each session consisted of eight identical periods wherein subjects played an attacker-defender population game. Each period lasted for a total of forty seconds. Each session lasted for about thirty minutes, including time to read

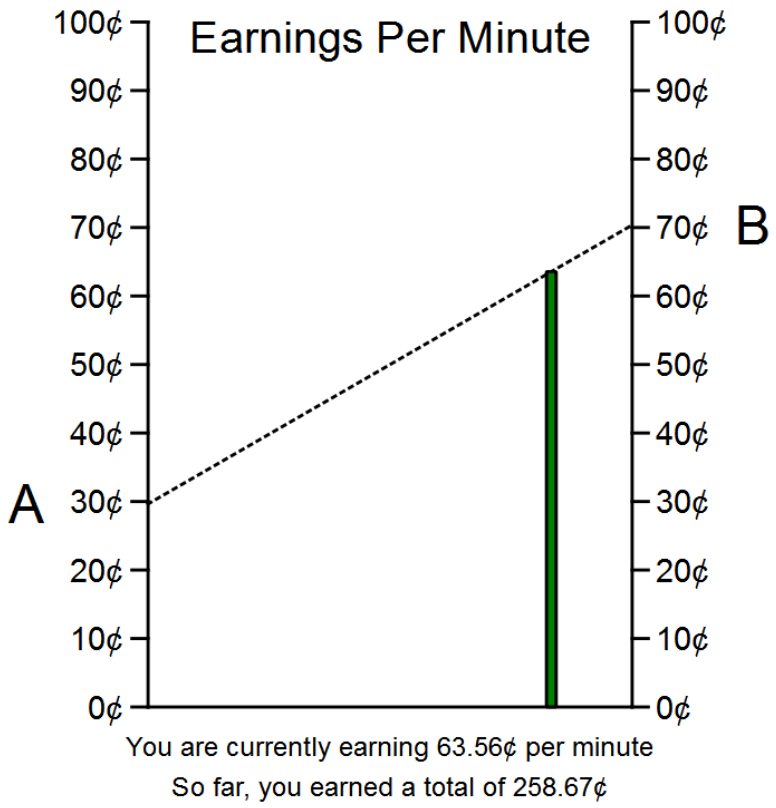

Figure 1.3: Experimental decision interface

the instructions.

Throughout each period, subjects could continuously adjust their probability of employing each strategy.[13] Figure 1.3 illustrates an example of the decision interface used by subjects. The horizontal position of the green bar illustrates the probability of each strategy currently being selected by the subject. When the green bar was all the way on the left side of the graph, the subject employed the pure strategy A with certainty. Conversely, when the green bar was all the way on the right side of the graph, the subject employed the pure strategy B with certainty. When the green bar was in the interior of the graph,

[13] [21] similarly allow their subjects to adjust their mixed strategies continuously over time.

the subject had a positive probability of employing each pure strategy. More precisely, the subject's probability of taking action A was proportional to the distance between the green bar and the right side of the graph and the subject's probability of taking action B was proportional to the distance between the green bar and the left side of the graph.

The dotted black line in figure 1.3 illustrates the current expected earnings rate for each feasible mixed strategy. The height of the dotted line at the left side of the graph indicates the current payoff for the pure strategy A. Similarly, the height of the dotted line at the right side of the graph shows the current payoff for the pure strategy B. During each period subjects earned continuous flow payoffs. In figure 1.3, the height of the green bar indicates the subject's current earnings rate. In addition, the subject's current earnings rate and the subject's current accumulated earnings are listed at the bottom of the decision screen. At the end of each experimental session, subjects received their total accumulated earnings plus a five dollar show up payment in cash.

1.4 Hypotheses

Both the control treatment and the coordination treatments have identical unique Nash equilibria. However, as described in section 1.2.2, different evolutionary models yield markedly different predictions regarding the disequilibrium dynamics of subjects in each treatment group. In particular, evolutionary dynamics consistently predict that the social state will exhibit global convergence to equilibrium under the control treatment, gradually spiraling inwards towards the Nash equilibrium. In contrast, evolutionary dynamics consistently predict that the social state will converge to a stable limit cycle that perpetually orbits the Nash equilibrium under the coordination treatment.

Although a evolutionary models consistently predict non-convergent cyclical behavior in the coordination treatment, different evolutionary models yield sharply different predictions regarding the shape of these cycles. In particular, sign-preserving dynamics

predict that limit cycles will approach the boundary of the state space, while non-sign-preserving dynamics predict that limit cycles will remain strictly in the interior of the state space. Further, continuous dynamics predict that subjects will exhibit small behavioral changes in response to correspondingly small changes in payoffs, but best response dynamics predict that subjects will exhibit drastic behavioral changes in response to small changes in payoffs when their best response changes. This bifurcation between theoretical predictions provides this paper with a uniquely powerful test for the underlying behavioral assumptions of evolutionary models. See figure 1.2 for a graphical illustration of these theoretical predictions.

Accordingly, behavioral deviations from equilibrium are predicted to be self-correcting under the control treatment. Whereas, under the coordination treatment, small behavioral deviations are predicted to produce sustained non-convergence and persistent cycling. Under both treatments, evolutionary dynamics predict that the population state will exhibit clockwise cyclical dynamics. From these theoretical predictions, the following hypotheses are obtained:

H0. Nash Equilibrium: Subjects in both treatments exhibit identical behavior since both treatments share an identical unique Nash equilibrium.

H2. Evolutionary Dynamics: Both treatments will exhibit clockwise cycles but the social state will exhibit greater stability and lower deviation from equilibrium in the control treatment than in the coordination treatment as predicted evolutionary dynamics.

H3. Continuous Dynamics: Subjects will exhibit small behavioral changes in response to correspondingly small changes in payoffs as predicted by continuous evolutionary dynamics.

H4. Sign-Preserving Dynamics: Subjects in the coordination treatment will exhibit behavioral limit cycles that approach the boundary of the state space as predicted by sign-preserving evolutionary dynamics.

The first hypothesis coincides, not only with the predictions of Nash equilibrium, but also with the predictions from a variety of widely employed behavioral models including logit quantal response equilibrum [10], level-k [11, 12], and cognitive hierarchy models [13], as all of these models yield predictions that are identical to Nash equilibrium for the attacker-defender games under consideration.[14] Thus a rejection of our first hypothesis would not only falsify Nash equilibrium, but would also falsify several other widely employed behavioral models.

The second hypothesis is drawn from the theoretical predictions of three widely used dynamic evolutionary models: the best response dynamic [15], the Smith dynamic [16], and the logit dynamic [17], each of which are described in detail in section 1.2.3. Our second hypothesis reflects the starkly contrasting predictions across treatments from these evolutionary models regarding behavioral stability. It should be noted that this bifurcation in theoretical predictions across treatments is present neither in the Nash predictions nor in the predictions from any of the other aforementioned behavioral solution concepts.

The third hypothesis reflects the characterization of disequilibrium behavior from the class of continuous evolutionary models and the fourth hypothesis reflects the characterization of disequilibrium behavior from the class of sign-preserving evolutionary

[14]To see why these these theoretical predictions are identical, recall that, in the attacker defender games under consideration, both targets are equally valuable, so in equilibrium, attackers are equally likely to attack each target and defenders are equally likely to defend each target. This is also the behavior that occurs if agents pick their strategy purely at random, so the Nash equilibrium coincides with the predictions from many solution concepts that describe some form of random behavioral noise. This feature of the attacker-defender game under consideration is part of what allows our experimental design to provide a sharp bifurcation between the theoretical predictions of evolutionary dynamics and those of other solution concepts.

models. To the extent that Nash equilibrium is said to model the long run limiting behavior of experienced agents, it correspondingly fails to explicitly model the dynamic process of convergence from which equilibrium emerges. In this sense, our third and fourth hypotheses consider a prediction from dynamic evolutionary models regarding an aspect of behavior that remains unmodeled by static behavioral solution concepts.

	Nash Prediction	Control Treatment	Coordination Treatment
mean distance from equilibrium to social state	0	0.067	0.226
		(0.012)	(0.026)
mean proportion of defenders who select strategy B	0.5	0.503	0.505
		(0.013)	(0.037)
std. dev. of proportion of defenders who select strategy B	0	0.059	0.166
		(0.014)	(0.017)
mean proportion of attackers who select strategy B	0.5	0.497	0.494
		(0.006)	(0.034)
std. dev. of proportion of attackers who select strategy B	0	0.049	0.166
		(0.009)	(0.026)

Table 1.1: Summary statistics for strategy selection in the coordination treatment and the control treatment. Here distance from equilibrium is measured as the mean euclidean distance between the Nash equilibrium social state and the observed social state, so as to maintain comparability with population proportions. Standard deviations across periods are given in parenthesis.

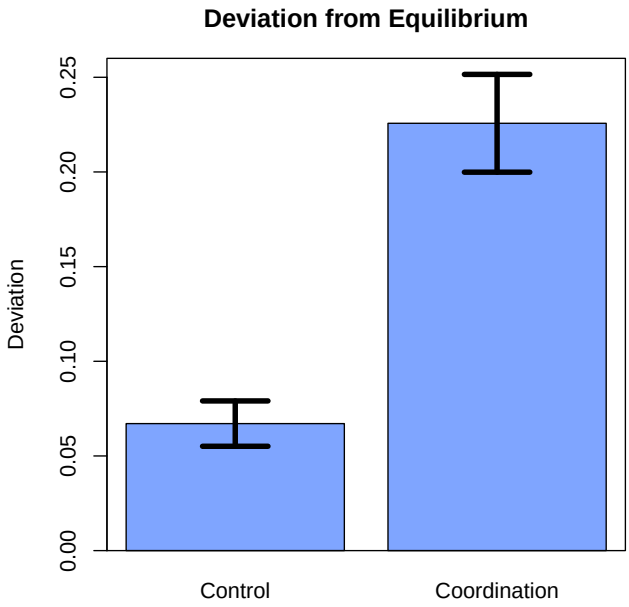

Figure 1.4: Mean deviation from equilibrium across treatments.

1.5 Results

Table 1.1 provides summary statistics[15] for the both coordination treatment and the control treatment and a comparison with the equilibrium predictions. As discussed in section 1.2, Nash equilibrium predictions are identical for both treatments: in equilibrium attackers are equally likely to attack each target and defenders are equally likely to defend each target. While the time-average of the observed behavior is highly consistent with the Nash predictions, it fails to tell the whole story. In particular, time averaging the observed within-period behavior would fail to detect both the striking differences in behavioral dynamics across treatments, which which are discussed below.

[15]Standard deviations for this table are calculated across periods since, as predicted by evolutionary models, aggregate behavior within a period is highly autocorrelated over time. Consequently, calculating the standard deviation of the relevant statistics across periods provides a more conservative estimate than calculating the total standard deviations of the high frequency data, which would largely capture within period variation.

Result 1. Subjects in the coordination treatment persistently exhibited significantly higher deviation from equilibrium than subjects in the control treatment.

The average social state exhibited considerably greater variance under the coordination treatment than under the control treatment. Consequently, as illustrated by figure 1.4, behavior in the coordination treatment is characterized by significantly larger deviations from equilibrium than the behavior observed under the control treatment. Here distance from equilibrium is measured as the mean Euclidean distance[16] between the Nash equilibrium social state and the observed social state in the state space of the population game as described in section 1.2.1. Error bars indicate standard deviations across periods. In the control treatment the social state rapidly converged to a small neighborhood of Nash equilibrium, while, in the coordination treatment, the social state persistently maintained a large distance from the Nash equilibrium.

Result 2. In both treatments, the social state exhibited significantly clockwise cyclical dynamics.

Subjects in both treatments exhibited significant clockwise behavioral cycles that orbit the Nash equilibrium social state. The presence of these clockwise cycles is consistent with the theoretical predictions from dynamic evolutionary models as discussed in section 1.2.4. The continuous-time experimental design employed by this study provides the opportunity to characterize these features of disequilibrium behavior which are described by dynamic evolutionary models, but remain unaddressed in equilibrium models and are less accessible in discrete-time experimental studies. In particular, subjects in the control treatment exhibited clockwise behavioral cycles that rapidly converged to a small neighborhood around the Nash equilibrium social state. In contrast, subjects in

[16]This method of measuring distance maintains comparability with statistical measures characterizing proportions of a population.

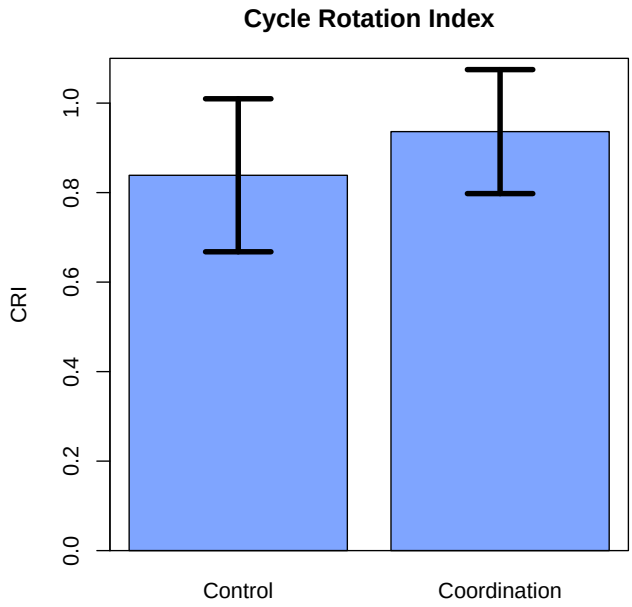

Figure 1.5: Mean value of the cycle rotation index across treatments.

the coordination treatment exhibited wider clockwise behavioral cycles that persistently diverge from the Nash equilibrium social state.

Figure 1.5 compares the mean value of the cycle rotation index [21] under the control treatment and the coordination treatment. Over all sixteen periods that implemented the control treatment, the mean value of the cycle rotation index was 0.84 and the standard deviation was 0.17. Similarly, over all sixteen periods that implemented the coordination treatment, the mean value of the cycle rotation index was 0.93 and the standard deviation was 0.14. Both treatments have significantly positive values of the cycle rotation index, indicating that behavior in both treatments was characterized by significantly clockwise cycles.

Figure 1.6 illustrates the observed behavior under each experimental treatment. The left column depicts observed behavior under the control treatment. The right column depicts the observed behavior under the coordination treatment. The upper row depicts

the observed behavior during the first half of each period. The lower row depicts behavior during the second half of each period. The horizontal axis in each graph represents the proportion of attackers that choose to attack target 2 and the vertical axis in each graph depicts the proportion of defenders that choose to defend target 2.

Result 3. The empirical limit cycles and smooth and remain strictly in the interior of the state space.

As illustrated by figure 1.6, the empirical limit cycles remain strictly in the interior of the state space, which contradicts theoretical predictions from sign preserving evolutionary dynamics,[17] but is consistent with theoretical predictions from sign-correlated evolutionary dynamics. In addition, the empirical limit cycles are smooth, which contradicts theoretical predictions from the best response dynamic and the wider class of discontinuous evolutionary dynamics, but is consistent with theoretical predictions from continuous evolutionary dynamics, such as the Smith dynamic and the logit dynamic.

Result 4. The empirical switch rates strongly violate the widely employed assumption of sign-preservation, that subjects exclusively adjust from lower performing strategies to higher performing strategies.

Figure 1.7 illustrates the observed rate at which subjects switch from one strategy to another, conditional on the difference in payoffs between the two strategies. Note that this observed conditional switch rate changes smoothly in response to changes in relative payoffs, supporting the assumption of continuity upheld by the Smith dynamic and the logit dynamic but contrasting with the sharp reversal in switch rates predicted by the best response dynamic. Further, note that subjects usually switch from lower performing strategies to higher performing strategies, but they also frequently switched from higher

[17] See section 1.2.4 for more details regarding the theoretical predictions from sign preserving evolutionary dynamics regarding the shape of limit cycles in coordinated attacker-defender population games

performing strategies to lower performing strategies, strongly violating the assumption of sign-preservation upheld by both the Smith dynamic and the best response dynamic, but conforming to the weaker sign-correlation assumption maintained by the logit dynamic.

The observed violation of sign-preservation suggests that the wider class of sign-correlated dynamic deserves further attention and may provide a superior characterization of human behavior in dynamic environments. Moreover, the violation of sign-preservation indicates that the fixed points of the empirical evolutionary dynamics may frequently fail to correspond with Nash equilibria in some games. Further research is needed to determine the extent to which sign-correlated evolutionary dynamics that violate sign-preservation can predict and explain persistent deviations from equilibrium other strategic environments.

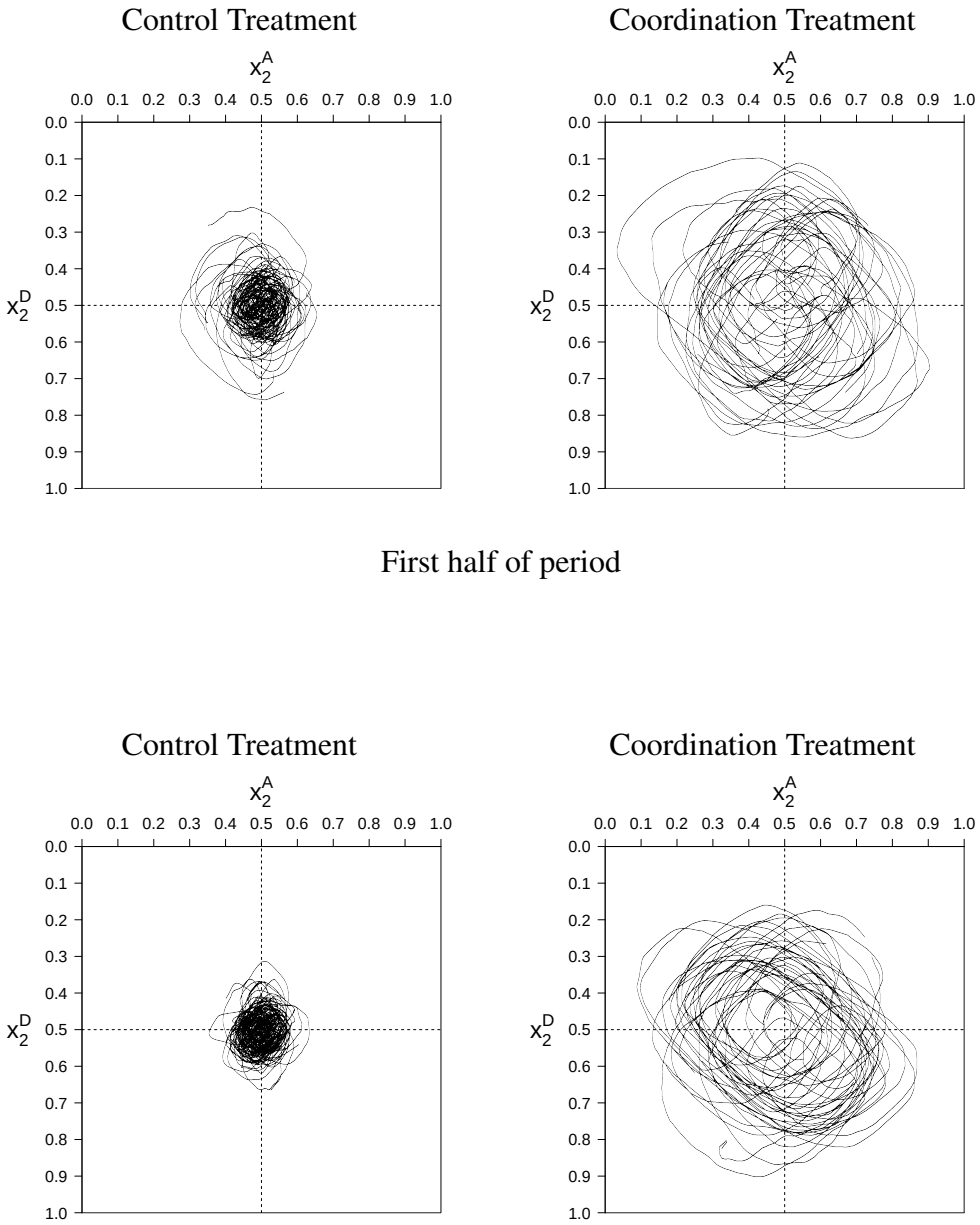

Figure 1.6: The observed path of the social state. The left column depicts observed behavior under the control treatment. The right column depicts the observed behavior under the coordination treatment. The upper row depicts the observed behavior during the first half of each period. The lower row depicts behavior during the second half of each period.

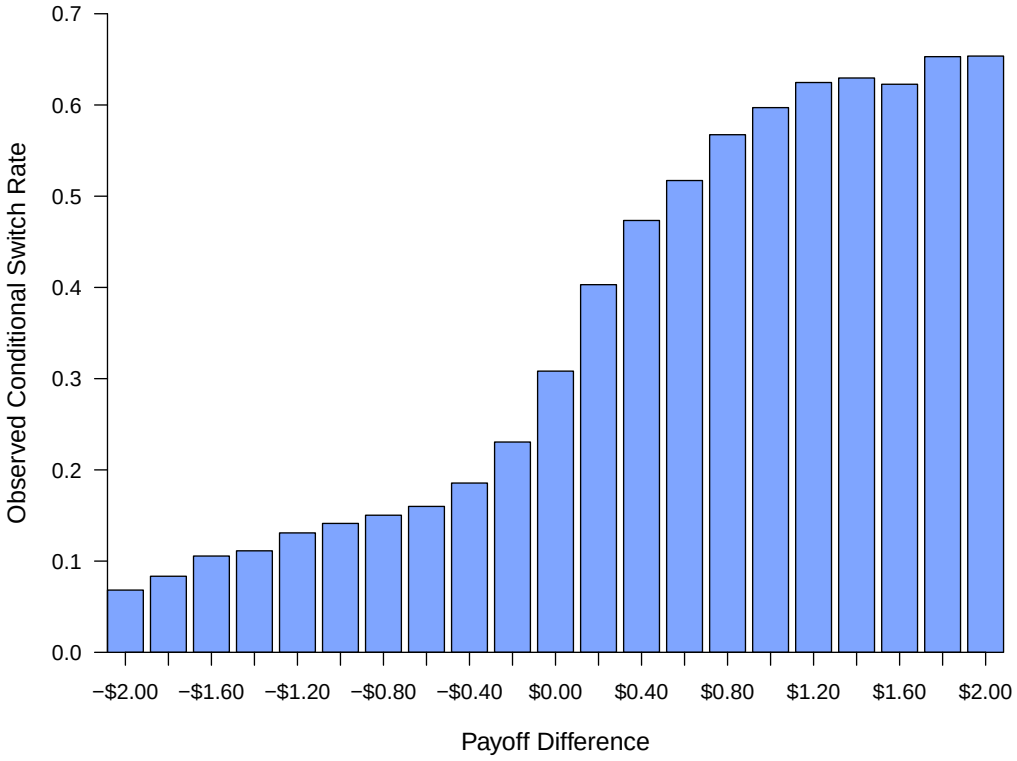

Figure 1.7: The observed conditional switch rate. The vertical axis indicates the percentage ρ_{ij}^p of subjects that choose to switch from strategy i to strategy j. The horizontal axis indicates the difference in payoffs $\pi_j^p - \pi_i^p$.

1.6 Conclusion

In this study, we experimentally test the empirical validity of dynamic evolutionary models across population games exhibiting identical Nash equilibria but starkly different evolutionary predictions. In particular, we vary coordination incentives across two attacker-defender population games. Crucially, both Nash equilibrium and several other behavioral models yield identical predictions for both of these experimental treatments. In particular, Nash equilibrium [7], logit quantal response equilibrum [10], level-k [11, 12], and cognitive hierarchy models [13] all yield identical predictions for both treatments.

In contrast, different evolutionary dynamics predict strikingly different behavior across treatments. In particular, we consider three widely employed evolutionary dynamics: the best response dynamic [15], the Smith dynamic [16], and the logit dynamic [17]. Although both treatments exhibit identical Nash equilibrium predictions, evolutionary dynamics generally predict that the social state will only converge to Nash equilibrium in our control treatment, and will exhibit persistent non-convergent limit cycles in the coordination treatment.

The best response dynamic and the Smith dynamic are both sign-preserving but the logit dynamic falls into the wider class of sign-correlated evolutionary dynamics. Consequently, the best response dynamic and the Smith dynamic predict that limit cycles in the coordination treatment will approach the boundary of the state space, while the logit dynamic predicts that these limit cycles will remain strictly in the interior of the state space. Further, the Smith dynamic and the logit dynamic are both continuous, but the best response dynamic is discontinuous, so the Smith dynamic and the logit dynamic predict that the social state will follow a smooth path, while the best response dynamic predicts that the path of the social state will exhibit sudden directional changes when the best response changes. This bifurcation between theoretical predictions across evolutionary

dynamics and other behavioral models allows this chapter to provide a clean experimental test of theoretical predictions from evolutionary game theory regarding aspects of dynamic disequilibrium behavior that remain largely unaddressed by standard models.

Subjects in both both treatments adjusted their strategies continuously and earned continuous flow payoffs, providing fine-grained behavioral data and allowing for the observation of long term behavioral phenomena that may be difficult to observe in a discrete-time setting. In accordance with theoretical predictions from evolutionary dynamics, subject behavior was tightly clustered around the Nash equilibrium in the control treatment but widely dispersed from the Nash equilibrium in the coordination treatment. In further agreement with evolutionary predictions, subjects also exhibited persistent clockwise cyclic behavior under both experimental treatments. In contradiction to theoretical predictions from the best response dynamic and the wider class of discontinuous evolutionary dynamics, the empirical path of the social state was smooth and did not exhibit sudden directional reversals when the best response changed, suggesting the continuous evolutionary dynamics may provide a more accurate depiction of human behavior than discontinuous models such as the best response dynamic. In contradiction to the predictions from the conventional class sign-preserving dynamics the empirical limit cycles remained strictly in the interior of the state space, suggesting that the wider class of sign-correlated dynamics deserves further attention and may provide a superior characterization of human behavior in dynamic environments.

While classical game theory focuses primarily on modeling equilibrium behavior, our experimental results suggest that evolutionary models of disequilibrium behavior can yield valuable insights. In contrast to the equilibrium predictions, our results suggest that the introduction of coordination incentives can prevent convergence to Nash equilibrium in attacker defender games, leading to autocorrelated attacks and making the behavior of attackers more predictable. These results may have important policy implications for

strategic environments that are modeled by attacker-defender games such as network security [27], property crime [33], and counter-terrorism [28]. In particular, these results suggest that policy makers ought not to rely exclusively on equilibrium models, or even on the wider class of static behavioral models, but should also consider predictions from evolutionary dynamics. Further, these results indicate that some classes of dynamic evolutionary models may provide a superior characterization of human behavior, namely continuous sign-correlated evolutionary dynamics. Finally, it should be noted that these results do not invalidate the use of equilibrium solution concepts. Rather, they suggest that evolutionary models can help identify where equilibrium models are most reliable and can characterize disequilibrium behavior where equilibrium models are less reliable.

2. DESIGNING MECHANISMS THAT RELIABLY CONVERGE

2.1 Motivation

Children in the United States are traditionally assigned to public schools based exclusively on where they live. However, a growing number of public school districts now allow parents to indicate their preferences over schools. Since each school can support only a limited number of students, it is often impossible to give every student her top choice of schools. To resolve these shortages, policy makers frequently employ student assignment mechanisms that assign each student to a school based on both reported student preferences and legally determined student priorities.

Under some of these mechanisms, participants have an incentive to strategically misreport their preferences. Some parent groups have even explicitly recommended particular misreporting strategies.[1] Misreported preferences prevent policymakers from accurately evaluating mechanism efficiency and make it difficult to reliably achieve policy goals. To encourage truthful preference reports, mechanism designers typically recommend the use of strategy-proof assignment mechanisms under which participants never have an incentive to misreport their preferences. However, previous studies[2] have found that even strategy-proof mechanisms fail to reliably induce truthful preference revelation from boundedly rational participants.

Standard implementations of school choice mechanisms only reveal assignments at the end of the reporting period, after all preference reports have been finalized. In contrast, this paper considers school choice mechanisms that provide participants with continuous assignment feedback. Under such mechanisms, each subject is shown her

[1] See Abdulkadiroglu et al. [34] for more details regarding these misreporting strategies.

[2] For example, Chen and Sönmez [35] find that subjects misrepresent their preferences 50% of the time under a top trading cycles mechanism. Similarly, Pais and Pintér [36] find that subjects misrepresent their preferences 33% of the time under a full information deferred acceptance mechanism.

current assignment throughout the preference reporting period, before preference reports are finalized. To the best of our knowledge, this study is the first to experimentally investigate such mechanisms. By providing increased opportunity for learning and adjustment, continuous assignment feedback is hypothesized to reduce confusion and promote rational preference revelation. To test this hypothesis, this study implements both discrete feedback and continuous feedback treatments for three widely employed school choice mechanisms: the deferred acceptance mechanism, the top trading cycles mechanism, and the Boston mechanism.

At present, barriers to implementing continuous assignment feedback are largely computational. Hence as computational power increases, this type of feedback will become increasingly feasible. Computationally simpler forms of continuous feedback are already being employed by some school districts. Specifically, continuous feedback regarding the first choices of other participants under the boston mechanism has been provided by the Wake County Public School System [37]. Similarly, Inner Mongolia provides continuous feedback in a dynamic queuing mechanism where subjects exclusively report their first choices [38].

Both top trading cycles and deferred acceptance are strategy proof mechanisms, so both have a Nash equilibrium in weakly dominant strategies under which participants truthfully report their preferences. Under the top-trading cycles mechanism, this truthful equilibrium always yields a Pareto optimal assignment. Under the deferred acceptance mechanism, the truthful equilibrium always yields an assignment that eliminates justified envy. Unlike the other two mechanisms under consideration, the Boston mechanism is manipulable. It has no dominant strategy equilibrium and participants frequently have an incentive to misreport their preferences. Nevertheless, as discussed by Ergin and Sönmez [39], the set of equilibrium assignments under the Boston mechanism coincide exactly with the set of assignments that eliminate justified envy.

This study connects two distinct strands of literature: the school choice literature in mechanism design theory, and experimental literature investigating dynamic behavior in continuous-time games. The school choice mechanism design literature provides an axiomatic analysis of rational preference revelation behavior under school choice mechanisms. Abdulkadiroglu and Sönmez [40] describe the school choice problem and discuss the fundamental tradeoff between Pareto efficiency and the elimination of justified envy. They also describe a variation of the top trading cycles mechanism introduced by Shapley and Scarf [41] which we investigate in this study. A powerful characterization of the Nash equilibria of the Boston mechanism was provided by Ergin and Sönmez [39] and the student optimal deferred acceptance mechanism was described by Gale and Shapley [42].

Previous experimental studies, such as Chen and Sönmez [35], conducted school choice mechanisms in discrete periods, which is ideal for the study of static one-shot mechanisms. In contrast, continuous-time experimental studies have successfully investigated dynamic behavior in a variety strategic settings involving continuous-time interaction. For example, Cason et al. [21] conduct a experimental investigation of dynamic behavior in continuous-time rock-paper-scissors games and Oprea et al. [20] conduct a continuous-time experimental study of evolutionary dynamics in the Hawk-Dove game. Both studies provide subjects with continuous feedback and allow subjects to adjust their strategies asynchronously. Unlike these experimental studies, this study employs continuous time experimental methodology to investigate dynamic preference revelation behavior in widely employed school choice mechanisms under continuous feedback.

This study finds that the provision of continuous assignment feedback helps school choice mechanisms to achieve equilibrium assignments significantly more often than conventional discrete feedback implementations. Discrete feedback mechanisms persistently

fail to reach equilibrium assignments due to the presence of behavioral noise in preference reports. In contrast, participants exhibited significantly less behavioral noise and significantly stronger convergence to equilibrium under continuous assignment feedback. Accordingly, the top trading cycles mechanism achieved greater efficiency under continuous feedback while the deferred acceptance mechanism and the Boston mechanisms eliminated significantly more justified envy under continuous feedback. These results suggest that the implementation of continuous feedback mechanisms can provide participants with greater opportunity for learning and adjustment, leading to more rational preference reports and significantly increasing the ability of policy makers to achieve policy goals in their school district.

2.2 Theory

2.2.1 The School Choice Environment

This study experimentally investigates a simple school choice environment that illustrates the fundamental tradeoff between Pareto efficiency and the elimination of justified envy.[3] Each school can accept up to n students and each student can be assigned to only one school. Students have strict preferences over schools and schools have strict priority rankings over students. Here there are three types of students and there are n students of each type. Student preferences over schools are given by

Student	1	2	3
	b	a	a
Preference	a	b	b
	c	c	c

where higher vertical position indicates a higher preference ranking, so type 1 students prefer school b over school a and prefer school a over school c. Similarly, school priority

[3] A similar example with only one student of each type was discussed by Abdulkadiroglu and Sönmez [40] and Roth [43].

rankings over students are given by

School	a	b	c
	1	2	2
Priority	3	1	1
	2	3	3

where priorities rankings between students of the same type are determined by lottery. A student x is said to have justified envy towards a student y if the student x prefers the school that is assigned to y and x also ranked higher at this school than y. If no student has justified envy under an assignment we say that the assignment eliminates justified envy. In general, several distinct assignments may eliminate justified envy. However, in this environment, the only assignment that eliminates justified envy is given by

$$\mu = \begin{pmatrix} 1 & 2 & 3 \\ a & b & c \end{pmatrix}$$

where all type 1 students are assigned to school a, all type 2 students are assigned to school b, and all type 3 students are assigned to school c. However, this assignment is Pareto dominated by the assignment

$$\lambda = \begin{pmatrix} 1 & 2 & 3 \\ b & a & c \end{pmatrix}$$

where all type 1 students are assigned to school b, all type 2 students are assigned to school a, and all type 3 students are assigned to school c. The assignment λ Pareto dominates the assignment μ because types 1 and 2 prefer the schools they receive under λ to the schools they receive under μ and student 3 receives the same school under λ as she does under μ. However, λ fails to eliminate justified envy because type 3 students have justified envy towards type 2 students under λ. Specifically, student 3 would prefer school a over school c, and school a gives student 3 a higher priority than student 2.

Since the unique assignment μ that eliminates justified envy is Pareto dominated by λ, no Pareto optimal assignment can eliminate justified envy in this environment. Furthermore, in general, no student assignment mechanism can guarantees both the elimination justified envy and Pareto optimality. Consequently, policy makers face a fundamental tradeoff between Pareto efficiency and the elimination of justified envy in school choice environments.

2.2.2 Student Assignment Mechanisms

Student assignment mechanisms select an assignment of students to schools based on the priority rankings of each school and the preferences reported by students. Since it is impossible for any student assignment mechanism to ensure both Pareto optimality and the elimination of justified envy, the optimal student assignment mechanism for a particular school district depends partly on the particular goals of the policy maker. Hence different school districts might reasonably employ different student assignment mechanisms. In particular, this paper considers three widely employed assignment mechanisms: the Boston mechanism, the top trading cycles mechanism, and the deferred acceptance mechanism.

A student x is said to have justified envy towards a student y if the student x prefers the school s that is assigned to student y and the student x also has higher priority at school s than student y. If no student has justified envy under an assignment then we say that the assignment eliminates justified envy. A mechanism is said to eliminate justified envy if it always selects an assignment that eliminates justified envy under the reported preferences. Similarly, a mechanism is Pareto optimal if it always selects an assignment that is Pareto optimal under the reported preferences. We say that a mechanism is strategy proof if no student can ever benefit by unilaterally misreporting her preferences.

Under the Boston mechanism, each student initially applies to her top choice of schools

according to her reported preferences. Each school accepts applicants in priority order until it runs out of seats. The remaining students apply to their second choice of schools according to their reported preferences. Again, each school accepts students in priority order until it runs out of seats. This process repeats until every student is assigned to a school.

When students truthfully report their preferences, the Boston mechanism selects a Pareto optimal assignment. Yet the Boston mechanism is not strategy proof, so students can often benefit by misreporting their preferences. Ergin and Sönmez [39] show that the set of Nash equilibrium assignments under the Boston Mechanism exactly coincide with the set of matchings that eliminate justified envy under the true preferences. However, these equilibrium assignments may be Pareto dominated. In this school choice environment,[4] the sole assignment μ that eliminates justified envy is the unique Nash equilibrium assignment for the Boston mechanism. However, no students receive their most preferred school under this assignment. Moreover, it is Pareto dominated by the Pareto optimal assignment λ where type 1 students and type 2 students receive their most preferred schools.

Under the student optimal deferred acceptance mechanism, each student initially applies to her top choice of schools according to her reported preferences. Each school tentatively accepts applicants in priority order until it runs out of seats. The remaining applications are rejected. Students whose applications were rejected then apply to their next highest choice of schools. Next, each school considers its new applicants along side those it has already tentatively accepted. It then tentatively accepts its top priority students among this group until it run out of seats and rejects the remaining students. This process repeats until every student is assigned to a school.

Unlike the Boston mechanism, the deferred acceptance mechanism is strategy proof,

[4] See section 2.2.1 for details regarding the school choice environment under consideration

so students never have an incentive to misreport their preferences. When sudents truthfully report their preferences, the deferred acceptance mechanism always selects an assignment that eliminates justified envy under the true preferences. Yet even when students report their preferences truthfully, the deferred acceptance mechanism does not always select a Pareto optimal assignment. In the this school choice environment, the deferred acceptance mechanism selects the sole assignment μ that eliminates justified envy when students report their preferences truthfully. However, it is Pareto dominated by the Pareto optimal assignment λ where type 1 students and type 2 students receive their most preferred schools. Furthermore, none of students receive their most preferred school under the dominant strategy Nash equilibrium assignment μ.

The top trading cycles mechanism constructs a directed graph based the reported preferences and priorities. Each student points to her most preferred school according to her reported preferences and each school points to it's highest priority student. Since there are a finite number of schools and students, the resulting directed graph will have at least one cycle. Students who are part of a cycle are assigned to the school they point to and removed from the directed graph. Next, each of the remaining students point to their most preferred school according to their reported preferences among those schools that still have open seats. Each school points to their highest priority student among those that remain unassigned. Any students who are part of a cycle are assigned to the school they point to. This process repeats until every student is assigned to a school.

The top trading cycles mechanism strategy proof, so students have no incentive to misreport their preferences. However, unlike the deferred acceptance mechanism, the top trading cycles mechanism always selects a Pareto optimal assignment when students truthfully report their preferences. However, when students report truthfully report their preferences, the top trading cycles mechanism does not always select an assignment that eliminates justified envy. In this school choice environment,4 the Pareto optimal

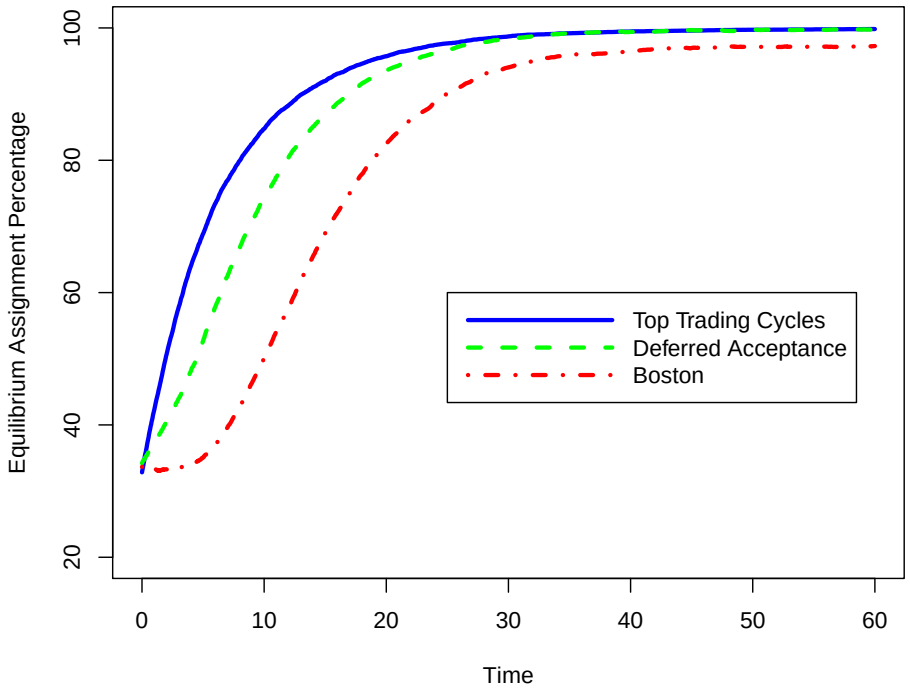

Figure 2.1: Equibrium Assignment Percentages under the Best Response Dynamic

assignment λ is the dominant strategy Nash equilibrium assignment for the top trading cycles mechanism. Under this assignment, type 1 students and type 2 students receive their most preferred school, while Type 3 students receive their least preferred school. Hence two thirds of the student population receive their most preferred school. However, this assignment does not eliminate justified envy since it gives type 3 students justified envy towards type 1 students.

2.2.3 Adaptive Dynamics Under Continuous Feedback

Previous experimental studies[5] find that, under discrete feedback, even strategy-proof mechanisms may fail to reliably induce truthful preference revelation when participants exhibit bounded rationality. These failures to induce truthful preference revelation may result from bounded rationality, confusion, or disbelief regarding the incentives presented by strategy proof mechanisms. To ameliorate these problems, this study considers the implementation of school choice mechanisms with continuous feedback where participants receive information regarding their tentative assignments throughout the preference reporting period.

The provision of continuous feedback has no effect on the Nash equilibria of the school choice mechanisms under consideration because assignments remain exclusively determined by the preference reports selected at the end of the reporting period. However, by allowing for adaptive learning and adjustment, the provision continuous feedback may induce boundedly rational agents to exhibit increasingly rational preference revelation behavior, thus helping school choice mechanisms to achieve their respective equilibrium assignments. Adaptive models can describe the behavior of boundedly rational agents in continuous time strategic environments, such as school choice mechanisms that provide continuous feedback. Here, we consider the best response dyanamic described by Gilboa and Matsui [15] and Matsui [44], under which agents asynchronously switch to one of their myopic best responses. This adaptive dynamic is also closely related to the fictitious play dynamic discussed by Brown [32].

Figure 2.1 depicts the predictions of the best response dynamic under school choice mechanisms with continuous feedback. Here the horizontal axis denotes time over the

[5]For example, Chen and Sönmez [35] find that subjects misrepresent their preferences 50% of the time under a top trading cycles mechanism with discrete feedback. Similarly, Pais and Pintér [36] find that subjects misrepresent their preferences 33% of the time under a deferred acceptance mechanism with decrete feedback.

course of a reporting period and the vertical axis denotes the percentage of participants receiving their equilibrium assignment. Each line depicts the mean path of a particular school choice mechanism under the best response dynamic in the school choice environment.[6] Note that both the deferred acceptance mechanism and the top trading cycles mechanism rapidly converge to equilibrium, while the Boston mechanism converges more slowly and exhibits persistent deviation from equilibrium, suggesting that the manipulability of the Boston mechanism can lead to dynamic instability under continuous feedback.

2.3 Experimental Design

This study implements a 2x3 experimental design with six experimental treatments which are illustrated by table 2.1. Each column of this table denotes one three widely employed school choice mechanisms: the deferred acceptance mechanism, the top trading cycles mechanism, and the Boston mechanism. For each mechanism, the study implements one treatment with continuous feedback and one treatment with discrete feedback. Three experimental sessions were conducted for each of the six treatment blocks. Each session was conducted with twenty-four subjects. Each subject participated in only one experimental session. All sessions were conducted at the Texas A&M Economic Research Laboratory.

	Top Trading Cycles	Deferred Acceptance	Boston
Discrete Feedback	3 Sessions	3 Sessions	3 Sessions
Continuous Feedback	3 Sessions	3 Sessions	3 Sessions

Table 2.1: 3x2 Experimental Design with Three Sessions Per Block

[6] See section 2.2.1 for details regarding the school choice environment under consideration

2.4 Experimental Procedures

During each experimental session, subjects were divided into three groups of eight participants. Members of each group were assigned one of the three student types described in section 2.2.1. Each experimental session consisted of twelve periods, and each period lasted for exactly one minute. At the beginning of each period, subjects were informed about the earnings that they could receive from being assigned each of the three options: *a*, *b*, or *c*. This information remained visible to subjects for the duration of the experimental session. To avoid the possibility of introducing any psychological ordering or labeling bias, the labeling for each school and the order in which the options were presented was randomly reassigned at the beginning of each period.

Throughout each reporting period, subjects were free to adjust their preference reports as frequently as desired. At the end of each period, all preference reports were finalized and assignments were made based on these finalized preference reports. Figure 2.2 depicts the experimental interface under continuous feedback and figure 2.3 illustrates the experimental interface under discrete feedback. Under the discrete feedback treatment, subjects could only observed their assignments at the end of each reporting period, after all preference reports were finalized. In contrast, under the continuous feedback treatment, subjects could also observe their tentative assignments under the currently selected preference reports throughout the one minute reporting period. At the end of each session, subjects were paid the average of their earnings over all periods plus a five dollar participation bonus.

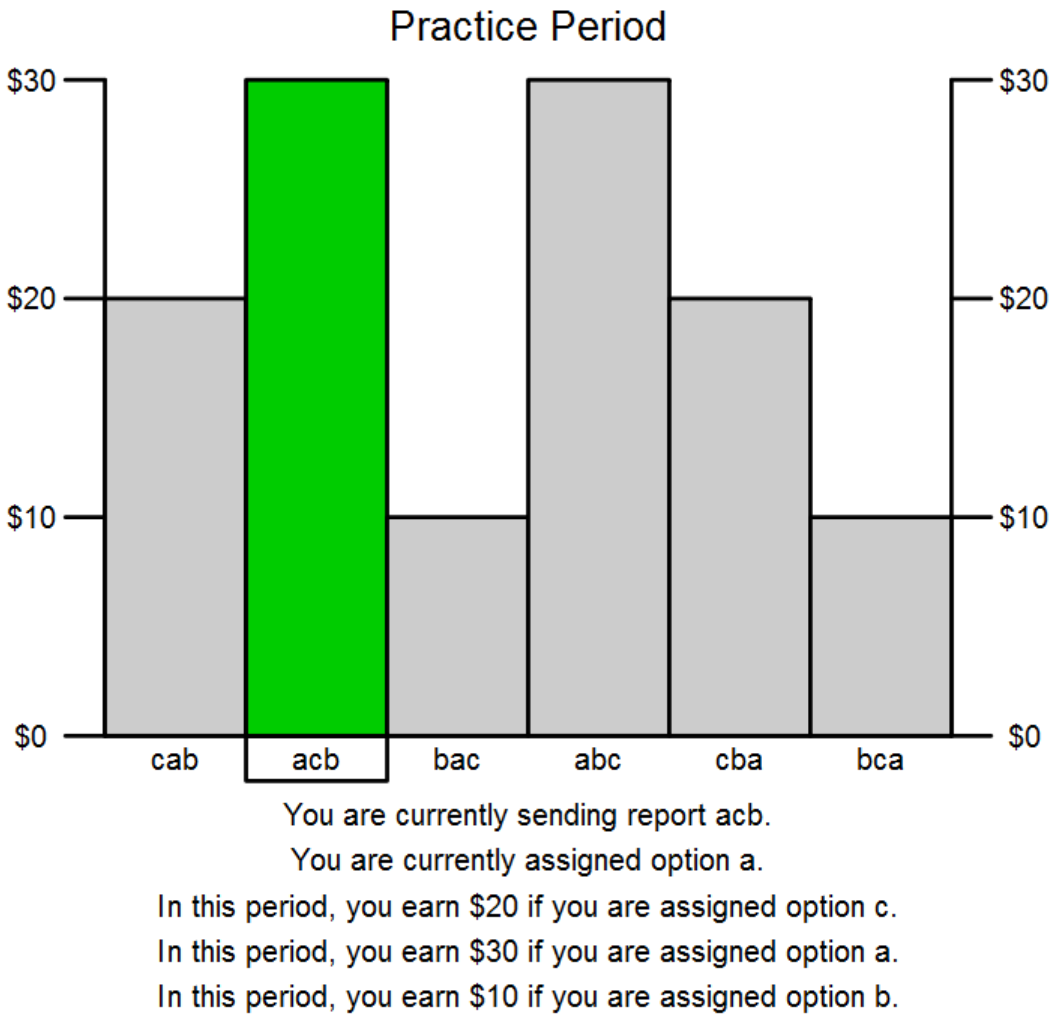

Figure 2.2: Experimental Interface under Continuous Feedback

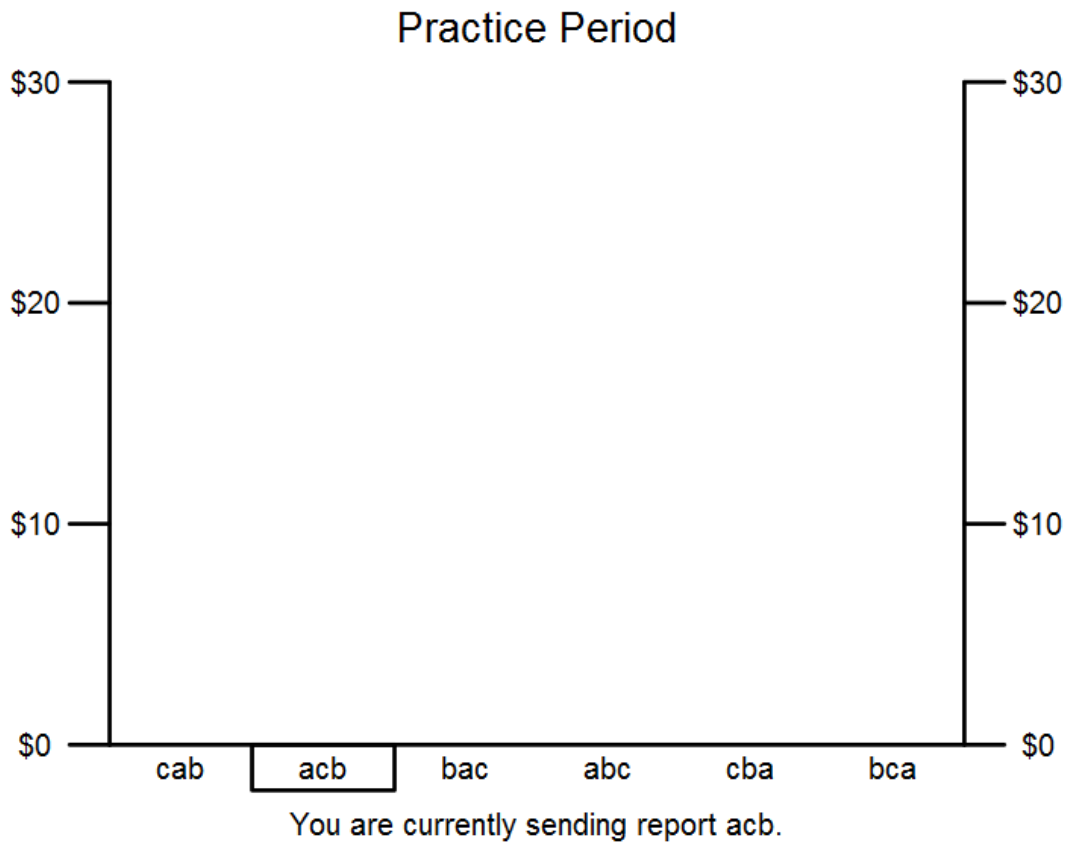

Figure 2.3: Experimental Interface under Discrete Feedback

2.5 Hypotheses

Hypothesis 1. *School choice mechanisms will achieve equilibrium assignments more often when they provide continuous feedback than when they only provide discrete feedback.*

Although the provision of continuous feedback has no effect on the Nash equilibria of the school choice mechanisms under consideration, it can help school choice mechanisms to achieve equilibrium assignments even in the presence of bounded rationality by giving participants more opportunity for learning and adjustment. When provided with continuous feedback, boundedly rational agents can converge towards equilibrium assignments by asynchronously adjusting towards their myopic best response.[7] However, this asynchronous process of myopic adjustment can not occur period if subjects only receive discrete feedback, so the provision of continuous feedback is expected to significantly increase the proportion of participants that receive their equilibrium assignments.

Hypothesis 2. *The top trading cycles mechanism will assign more students their most preferred school when it provides continuous feedback than when it only provides receive discrete feedback.*

In this school choice environment,[8] the dominant strategy Nash equilibrium of the top trading cycles mechanism yields the Pareto efficient assignment λ which assigns two thirds of the student population receives their most preferred school. Moreover, school a is the favorite of both type 2 students and type 3 students, so it is not possible to assign more than two thirds of students their most preferred option. Hence if the provision of continuous feedback makes the top trading cycles mechanism more likely to achieve equilibrium assignments, then it also increases the proportion of students that receive their favorite school. However, it should be noted that λ does not eliminate justified envy. In particular,

[7] See section 2.2.3 for more details regarding the theoretical predictions of the best response dynamic.
[8] See section 2.2.1 for details regarding the school choice environment.

it gives type 3 students justified envy towards type 2 students.

Hypothesis 3. *The deferred acceptance mechanism will eliminate more justified envy when it provides continuous feedback than when it only provides discrete feedback.*

The deferred acceptance mechanism is strategy proof and eliminates justified envy, so its dominant strategy Nash equilibrium yields the unique assignment μ which completely eliminates justified envy under the true preferences in this school choice environment. Thus if the provision of continuous feedback makes the deferred acceptance mechanism more likely to achieve equilibrium assignments, then it will also increase the elimination justified envy. However, the assignment μ is not Pareto optimal and it does not give any of the students their most preferred school, so we do not expect continuous feedback to increase the proportion of of students who are assigned their favorite schools under the deferred acceptance mechanism.

Hypothesis 4. *The Boston mechanism will eliminate more justified envy when it provides continuous feedback than when it only provides discrete feedback.*

Although the Boston mechanism is not strategy proof, all of its Nash equilibria eliminate justified envy under the true preferences. Thus, in this school choice environment, its Nash equilibria yield the aforementioned assignment μ, which eliminates justified envy but does not give any participant their most preferred school. Hence if the provision of continuous feedback makes the Boston mechanism more likely to achieve equilibrium assignments, then it also increases the elimination justified envy. However, as noted in section 2.2.3, the best response dynamic predicts that the Boston mechanism will exhibit greater dynamic instability than the other two mechanisms, so it may achieve less elimination of justified envy than the deferred acceptance mechanism in the presence of continuous feedback.

	Feedback		t-test	Joint F-test
	Discrete	Continuous	p-value	p-value
Top Trading Cycles	0.76273	0.92130	<0.001	
Deferred Acceptance	0.69097	0.98148	<0.001	<0.001
Boston	0.20255	0.89236	<0.001	

Table 2.2: Hypothesis tests regarding the proportion of equilibrium assignments. The unit of observation is one period.

2.6 Results

Result 5. All three school choice mechanisms achieved equilibrium assignments significantly more often when they provided subjects with continuous feedback than when they provided subjects with discrete feedback.

Figure 2.4 illustrates the proportion of equilibrium assignments under each of the six experimental treatments. The vertical axis denotes the percentage of subjects who received their dominant strategy equilibrium assignment. The three mechanisms under consideration, are listed along the horizontal axis. For each of these mechanisms, the height of the left-hand bar denotes the percentage of equilibrium assignments under discrete feedback and the height of the right-hand bar denotes the percentage of equilibrium assignments under continuous feedback. Error bars indicate 95% confidence intervals on the percentage of equilibrium assignments.

A t-test finds that each of these mechanisms achieve equilibrium assignments significantly more often under continuous feedback than under discrete feedback at the one percent level. An F-test rejects the joint hypotheses of equal equilibrium assignment percentages across feedback treatments at the one percent level. Although the provision of continuous feedback has no effect on the Nash equilibria of these school choice mechanisms, this result is consistent with the theoretical predictions of adaptive models.[9] By

[9]See section 2.2.3 for more details regarding the predictions of adaptive models.

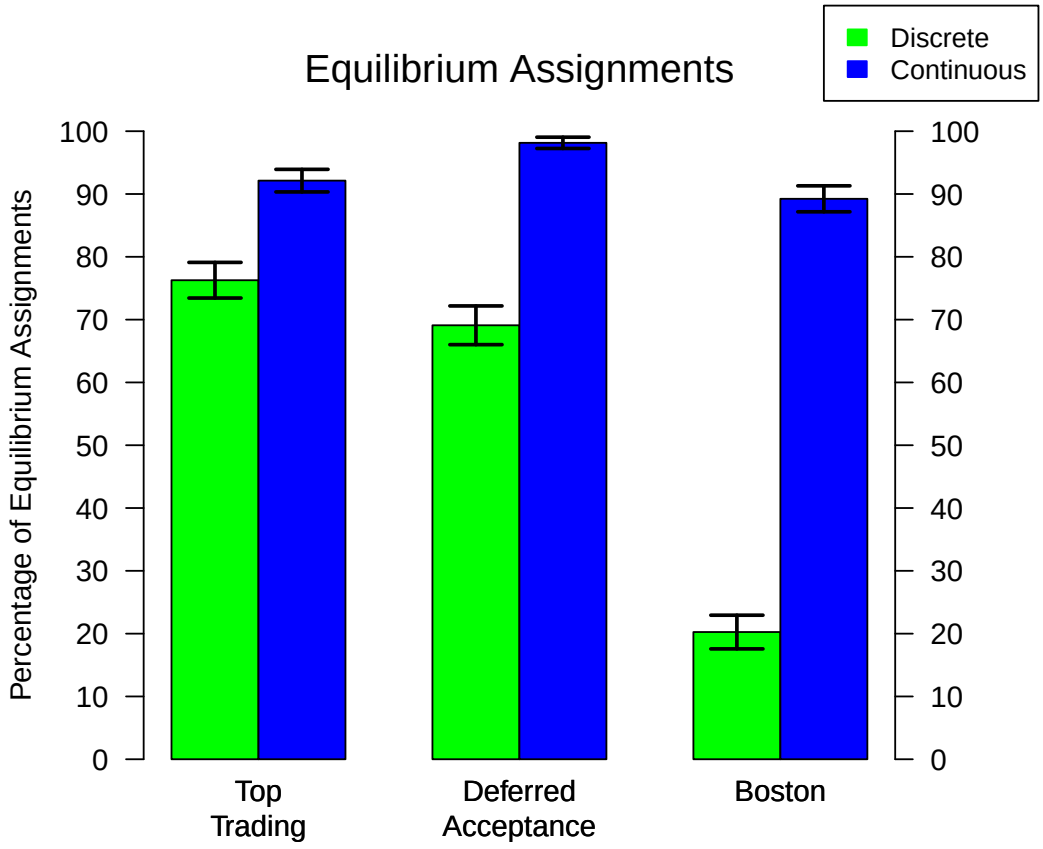

Figure 2.4: Proportion of Equilibrium Assignments by Treatment

allowing for adaptive learning and adjustment, the provision continuous feedback may induce boundedly rational agents to exhibit more rational preference revelation behavior, thus helping these school choice mechanisms to achieve their respective equilibrium assignments.

Result 6. All three school choice mechanisms eliminated significantly more justified envy when they provided subjects with continuous feedback than when they provided subjects with discrete feedback.

Figure 2.5 illustrates the elimination of justified envy under each of the six treatments. The vertical axis denotes the percentage of subjects who had no justified envy towards

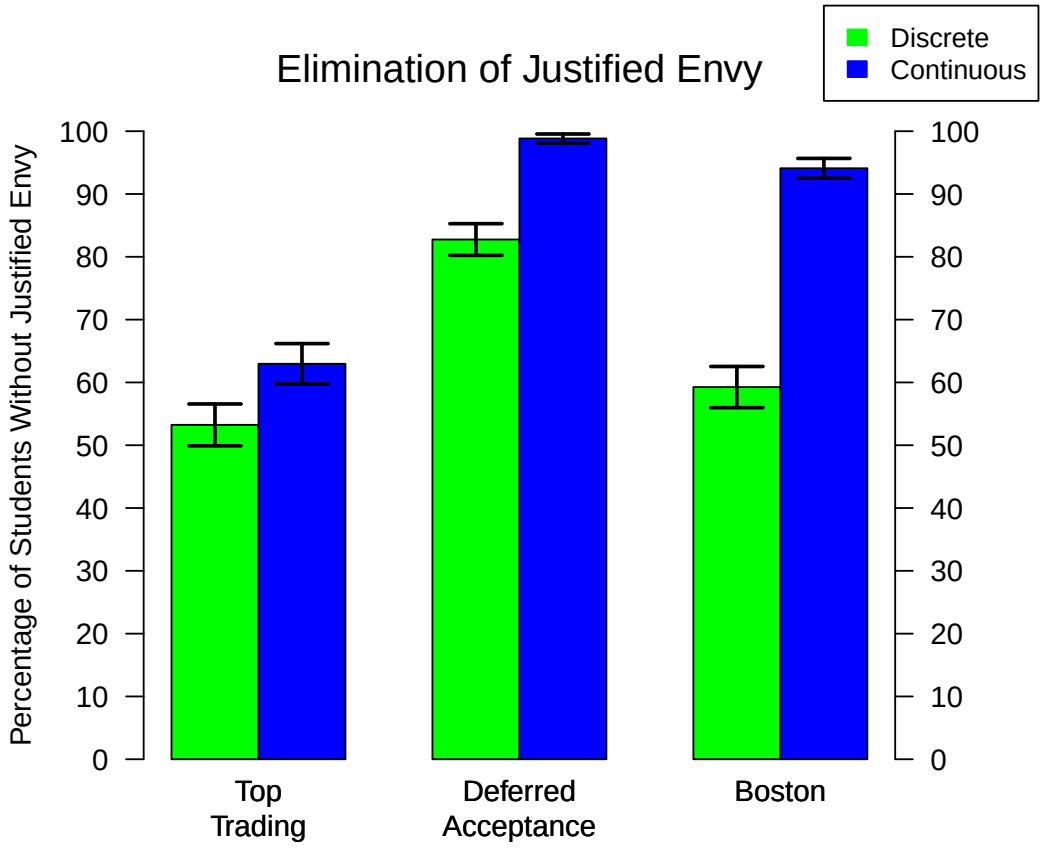

Figure 2.5: Elimination of Justified Envy by Treatment

others under their true preferences.[10] The three mechanisms under consideration, are listed along the horizontal axis. For each of these mechanisms, the height of the left-hand bar denotes the elimination of justified envy under discrete feedback and the height of the right-hand bar denotes the elimination of justified envy under continuous feedback. Error bars indicate 95% confidence intervals on the percentage of students without justified envy.

A t-test finds that all three of these school choice mechanisms eliminate significantly more justified envy under continuous feedback than under discrete feedback at the one percent level. An F-test rejects the joint hypotheses of equal justified envy elimination

[10] A formal definition for the concept of justified envy can be found in section 2.2.2

	Feedback		t-test	Joint F-test
	Discrete	Continuous	p-value	p-value
Top Trading Cycles	0.53241	0.62963	<0.001	
Deferred Acceptance	0.82755	0.98843	<0.001	<0.001
Boston	0.59259	0.94097	<0.001	

Table 2.3: Hypothesis tests regarding the elimination of justified envy. The unit of observation is one period.

rates across feedback treatments at the one percent level. As discussed in section 2.2.2, the assignment μ that uniquely eliminates justified envy in this environment is the equilibrium outcome for both the deferred acceptance mechanism and the Boston mechanism. Hence this result is consistent with the increase in equilibrium assignments from continuous feedback under these two mechanisms.

In contrast, the increase in the elimination of justified envy under the top trading cycles mechanism occurs because it eliminated so little justified envy under discrete feedback. In equilibrium, the top trading cycles mechanism eliminates justified envy from only two thirds of the student population, which is roughly consistent with the observed elimination of justified envy under the top trading cycles mechanism. Yet under discrete feedback, the top trading cycles mechanism eliminated even less justified envy, so the increase in equilibrium assignments from continuous feedback increased the elimination of justified envy despite the presence of justified envy in the equilibrium assignment of the top trading cycles mechanism.

Result 7. The top trading cycles mechanism gave subjects their most preferred option significantly more often when it provided subjects with continuous feedback than when they provided subjects with discrete feedback.

Figure 2.6 illustrates the proportion of most preferred assignments under each of the

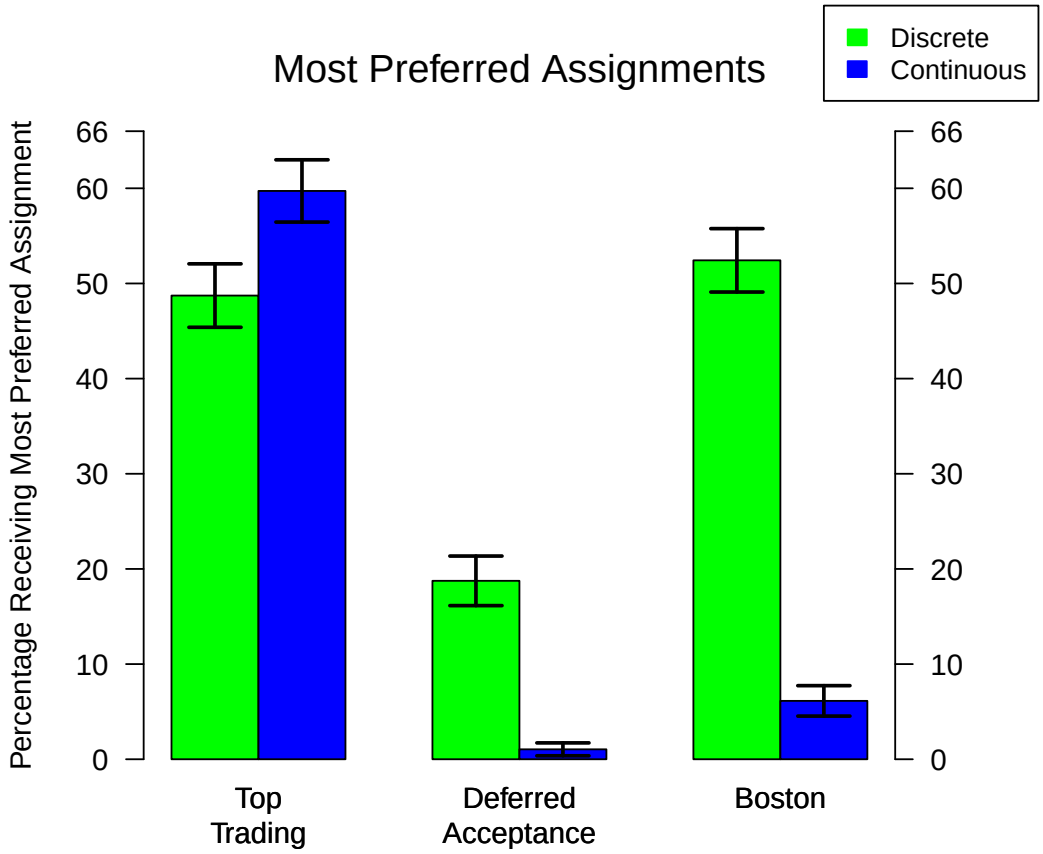

Figure 2.6: Proportion Assigned most Preferred Option by Treatment

six treatments. The vertical axis here denotes the percentage of subjects who who were assigned their favorite option under their true preferences. The three mechanisms under consideration, are listed along the horizontal axis. For each of these mechanisms, the height of the left-hand bar denotes the percentage of most preferred assignments under under discrete feedback and the height of the right-hand bar denotes the percentage of most preferred assignments under continuous feedback. Error bars indicate 95% confidence intervals on the percentage of students receiving their most preferred assignments.

A t-test finds that the top trading cycles mechanism assigned subjects their most preferred option significantly more often under continuous feedback than under discretfeed-

	Feedback		t-test	Joint F-test
	Discrete	Continuous	p-value	p-value
Top Trading Cycles	0.48727	0.59722	<0.001	
Deferred Acceptance	0.18750	0.01042	<0.001	<0.001
Boston	0.52431	0.06134	<0.001	

Table 2.4: Hypothesis tests regarding the proportion of most preferred assignments. The unit of observation is one period.

back at the one percent level. An F-test rejects the joint hypotheses of equal percentages of most preferred assignments across feedback treatments at the one percent level. In contrast, the deferred acceptance mechanism and the Boston mechanism assigned subjects their most preferred option significantly less often under continuous feedback than under discrete feedback at the one percent level. These results are consistent with the increase in equilibrium assignments under continuous feedback since the top trading cycles mechanism assigns two thirds of subjects their most preferred option in equilibrium while the other two mechanisms do not assign any students their most preferred option in equilibrium.

2.7 Conclusion

Classical mechanism design theory predicts that strategy proof mechanisms will reliably induce truthful preference reports and achieve equilibrium outcomes. However, these theoretical predictions are difficult to verify in the field because real world preferences are unobservable and real world school choice mechanisms rarely satisfy the exact assumptions of theory. Hence experimental data can provide valuable information regarding the empirical properties of school choice mechanisms and allow more conclusive testing of predictions from mechanism design theory.

Previous studies find that strategy-proof student assignment mechanisms fail to reliably achieve equilibrium outcomes or induce truthful preference revelation in laboratory

experiments. We hypothesize that these findings result from bounded rationality on the part of participants in these strategy-proof mechanisms. Moreover, we suspect that similar types of bounded rationality play an important role in the field. To reduce confusion and increase understanding, we consider the implementation of school choice mechanisms that provide participants with continuous feedback regarding their school assignments throughout the reporting period. To the best of our knowledge, this study is the first to experimentally investigate such mechanisms.

To investigate the empirical properties of widely employed school choice mechanisms under continuous assignment feedback, this study conducts laboratory experiments comparing assignment outcomes and preference revelation behavior across continuous feedback and discrete feedback school choice mechanisms. In these experiments, all three school choice mechanisms achieved equilibrium assignments significantly more often when they provided subjects with continuous feedback than when they only provided discrete feedback. Consistent with the theoretical predictions from adaptive models, these experimental results suggest that the provision of continuous feedback in school choice mechanisms can help promote rational preference revelation behavior by giving participants more opportunity for learning and adjustment.

Student assignment mechanisms impact the well being of children in many school districts throughout the world. The Boston mechanism was originally used in Boston's school choice system. In 2012, the New Orleans recovery school district used an algorithm based on the top trading cycles assignment mechanism [45]. In 2008, a variation of the student optimal deferred acceptance mechanism was employed in New York City [46]. By investigating continuous assignment feedback, this study can help policy makers to design better school choice mechanisms to achieve the policy goals of their school district. Furthermore, the analysis of continuous feedback may to continue to help researchers design new mechanisms that are more robust the presence of bounded rationality.

3. CONCLUSIONS: CYCLICAL BEHAVIOR IN THE ALL PAY AUCTION

3.1 Motivation

Nash Equilibrium is a powerful tool for understanding strategic behavior. Even when agents fall short of perfect rationality, simple adaptive processes can often drive long-run behavior towards equilibrium predictions. For this reason, adaptive models have long been employed to justify the application of equilibrium solution concepts in the presence of bounded rationality [6, 47]. In many strategic environments, the long run predictions of adaptive models closely resemble the predictions of Nash equilibrium. However, in other strategic environments, adaptive models fail to converge, leading to persistent disequilibrium behavior. In such environments, adaptive models characterize both the degree of deviation from equilibrium and the disequilibrium behavioral dynamics.

To test these theoretical predictions from adaptive models, we conduct laboratory experiments on continuous-time all-pay auctions where subjects adjust their bids asynchronously and earn flow payoffs continuously over time. In one treatment, the single highest bidder receives a prize. In another treatment, the top two bidders each receive an equally valuable prize. Nash equilibrium predicts identical payoff distributions under both of these treatments, but adaptive models predict predict disequilibrium bidding cycles, running contrary to the equilibrium predictions and disrupting the equality of payoffs across treatments. In contrast to the Nash equilibrium predictions, but consistent with the predictions of adaptive models, we observe lower behavioral stability and higher payoffs in all-pay auctions with two winners than in all-pay auctions with a single winner.

The all-pay auction has been examined extensively in experimental environments [see 48, for a survey]. Previous experimental studies [e.g., 49, 50, 51] of the all-pay auction conduct a sequence of discrete rounds in which subjects secretly select their bids and the

highest bidder receives a price. The paper also contributes to the small, but burgeoning area of literature that studies the properties of disequilibrium dynamics in continuous-time games [i.e., 20, 21]. Consistent with the observation of cyclical behavior in rock-paper-scissors by Cason et al., we observe bidding cycles in continuous-time all-pay auctions.

This paper proceeds as follows: Section 3.2 presents the structure of the game and two different equilibrium models. It also describes and the various adaptive models that will be used to characterize the experimental data. Section 3.3 presents the full design and procedures of the experiment. Section 3.4 provides our hypotheses. Section 3.5 presents the main results and Section 3.6 concludes.

3.2 Theory

In all-pay auctions, multiple agents expend costly effort to compete over a limited number of prizes. Prizes are awarded to the agents who expend the most effort, but every agent bears the cost of her own effort, even if she does not win a prize. All-pay auctions often model strategic environments that involve both conflict and non-recoverable costs such as political lobbying [52], patent races [53], biological competition [54], and international warfare [55].

The all-pay auction involves three players who compete over two prizes. Each player i starts with a endowment w and selects her bid b_i from the closed interval $[0, w]$. The top two bidders each receive a prize with value v and the lowest bidder receives no prize. Every player must pay her bid, regardless of whether or not she won a prize. In the case of a tie, the winner is determined randomly. Accordingly, the payoff function for player i

is given by:

$$\pi_i\left(b_i, b_j, b_k\right) = \begin{cases} w - b_i + v & \text{if } b_i > \min\left\{b_j, b_k\right\} \\ w - b_i + 2v/3 & \text{if } b_i = b_j = b_k \\ w - b_i + v/2 & \text{if } b_i = \min\left\{b_j, b_k\right\} < \max\left\{b_j, b_k\right\} \\ w - b_i & \text{otherwise} \end{cases} \tag{3.1}$$

3.2.1 Equilibrium Models

We consider equilibrium models including the Nash equilibrium and the logit quantal response equilibrium. Nash equilibrium assumes that each agent selects a best response to the strategies selected by others. In contrast, the logit quantal response equilibrium assumes that agents make probabilistic errors in their payoff evaluations.

The all-pay auction investigated here with three bidders and two winners has no pure strategy Nash equilibrium, but it does have a unique symmetric mixed strategy Nash equilibrium. First derived by [56],[1] the corresponding probability density function for the bid of player i given by

$$f\left(b_i\right) = \frac{1}{2v}\left(1 - \frac{b_i}{v}\right)^{-1/2} \qquad \text{for all } b_i \in [0, v]. \tag{3.2}$$

The black line in Figure 3.1 illustrates this equilibrium density function. Note that that Nash equilibrium probability density function approaches infinity as bids approaches the value of the prize and remains at zero for any bid above the value of the prize. Thus, in equilibrium players are likely to bid near the value of the prize but never above it.

While the Nash equilibrium describes the behavior of perfectly rational and perfectly precise agents, the logit quantal response equilibrium described by [10] and [57] allows

[1] Appendix Section 4.1 contains an alternate derivation of the equilibrium.

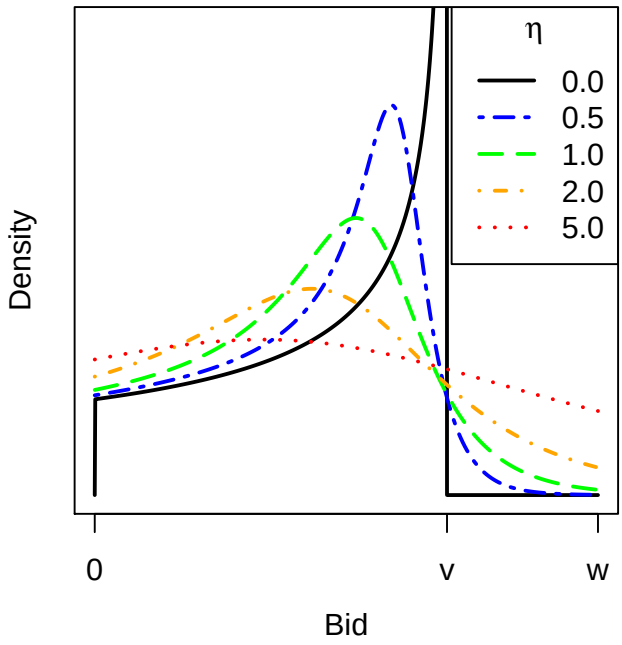

Figure 3.1: **Nash Equilibrium and Logit Quantal Response Equilibria**

us to model the behavior of imprecise boundedly rational agents. Such agents make probabilistic errors in their evaluation of alternate strategies. Although they typically fail to select a best response, they are more likely to select strategies that yield higher payoffs.

Unlike the perfectly rational agents described by Nash equilibrium, agents in logit quantal response equilibrium may place positive probability on dominated strategies, since their behavior is fundamentally stochastic. In the case of a continuous strategy space, the probability density function for the logit quantal response equilibrium mixed strategy σ_i satisfies:

$$f(b_i) = \frac{\exp\left(\eta^{-1}\pi_i\left(b_i, \sigma_{-i}\right)\right)}{\int \exp\left(\eta^{-1}\pi_i\left(x, \sigma_{-i}\right)\right) dx} \tag{3.3}$$

Here η denotes the level of behavioral noise in an agent's evaluation of payoffs. When η is small, agents make small errors, and the strategy distribution approaches the Nash

equilibrium. When η is large, the logit quantal response equilibrium approaches uniformly random play. To illustrate this tendency, Figure 3.1 depicts the logit quantal response equilibrium under alternate values of η.

A closed form solution for the logit quantal response equilibrium of an all-pay auction with a single prize is provided by [58]. To the best of our knowledge, no closed form solution is currently available for the logit quantal response equilibrium of an all-pay auction with two prizes. Accordingly, a formal derivation of the logit quantal response equilibrium in this case is found in Appendix Section 4.2. The logit quantal response equilibrium probability density function for the bid of player i is

$$f\left(b_i\right)=\frac{\eta G\left(\sqrt{\eta v}\right)\exp\left(-\eta b_i\right)}{\sqrt{\eta v}\left[1-\exp\left(-\eta w\right)\right]}\left[\exp\left(G^{-1}\left(G\left(\sqrt{\eta v}\right)\left[1-\frac{1-\exp\left(-\eta b_i\right)}{1-\exp\left(-\eta w\right)}\right]\right)^2\right)\right]^{-1}, \tag{3.4}$$

where $G\left(x\right)=\int_0^x \exp\left(u^2\right)du=\frac{\sqrt{\pi}}{2}\text{erfi}\left(x\right)$.

3.2.2 Evolutionary Game Theory

The experiment in this paper involves continuous-time, two good, three-bidder all pay auctions. These auctions are conducted in groups of three, but subjects' rewards are calculated using mean matching, so essentially every subject plays every other subject all the time. In situations like these, it is useful to consider models from evolutionary game theory, the study of "large populations of agents who repeatedly engage in strategic interactions," [8].

By design, this all-pay auction has evolutionary dynamics that make it a prime candidate for persistent disequilibrium. Accordingly, it has no evolutionary stable strategy. The idea of an evolutionarily stable strategy was introduced by [25], who employed it to identify the stability of biological phenotypes in large populations under the pressures of mutation and natural selection. More recently, game theorists and social scientists have

employed evolutionary stability criteria to model the behavioral stability of Nash equilibria in a wide variety of strategic settings.[2]

A strategy is evolutionarily stable if it induces a self-enforcing convention. That is, a strategy x is evolutionarily stable if no other strategy y can invade it when the entire population initially employs strategy x. More formally, in a in a symmetric normal form game, a strategy x is evolutionarily stable if there exists some $C \in (0, 1)$ such that for all $\varepsilon \in (0, C)$ and for any other strategy y

$$\pi\left(x \mid \epsilon y + (1 - \epsilon)\, x \right) > \pi\left(y \mid \epsilon y + (1 - \epsilon)\, x \right) \tag{3.5}$$

Thus, if x is evolutionarily stable and a sufficiently small proportion of the population deviates to an alternate strategy y, then agents who employ x will earn a strictly higher payoff than agents who employ y.

The Nash equilibrium strategy for the all-pay auction is not evolutionarily stable. Too see why, suppose that a small proportion ϵ of the population deviates from the Nash equilibrium strategy x to an alternate strategy y under which agents always bid the full value of the prize. Since the support of the equilibrium bid distribution is given by the closed interval $[0, v]$, agents who employ the invading strategy y will win the prize with probability one whenever they are matched against an agent who employs the equilibrium bidding strategy. In this case, the invading strategy y earns a higher expected payoff than the equilibrium mixed strategy x, so the equilibrium mixed strategy for the all-pay auction with three bidders and two prizes is not evolutionarily stable. A formal derivation of this result is found in Appendix Section 4.2.1. Since mixed strategy Nash equilibrium fails to induce a self enforcing convention in this all-pay auction, we expect to observe dynamic instability in experimental bidding behavior.

[2]These settings include price competition [23], linguistics [59], and corporate investment [60].

As we expect this experimental environment to be rife with instability, we have ample opportunity to examine the adaptive behavior of subjects. In particular, we will specifically examine noisy optimization dynamics and a noisy imitation dynamics from evolutionary game theory [8]. In these adaptive models, agents make asynchronous strategy adjustments over time. The timing of these adjustments follows a homogeneous Poisson process with a rate of δ adjustments per second. The value b_{it} here denotes the bid employed by agent i at time t. To determine the relative strengths of these models, we also develop a multi-parameter model that nests each as a special case.

Under deterministic optimization models, such as those described by Gilboa and Matsui [15] and Golman [61] agents switch precisely to their best response. In contrast, the logit dynamic is a noisy optimization model [17, 62], predicting that agents will be more likely to select bids that yield higher payoff. Under this model, the likelihood that an agent i who adjusts her bid at time t will select a particular bid b is given by:

$$f_{i,t}(b) = \frac{\exp(\beta \pi_i(b, b_{-i,t}))}{\int_0^w \exp(\beta \pi_i(x, b_{-i,t}))\, dx} \tag{3.6}$$

Purely imitative models [e.g., 61, 63, 64, 65] predict that agents will exclusively imitate the strategies of other agents they encounter. Such models predict that agents will never innovate by playing a strategy that was not previously employed by others in the population. In an experiment such as this one, with a continuous strategy space and finite number of subjects, this prediction will almost certainly fail as subjects pick new strategies that were not previously employed by others. To increase the flexibility of these imitative models, we consider a noisy imitation model under which agents are *more likely* to select bids that are *close* to the bid that is currently employed by the agent with the highest earnings rate. Let b_t^H denote the bid employed by the agent who has the highest earning rate at time t. Under this noisy imitative model, the likelihood that an agent who

adjusts her bid at time t will select a particular bid b is given by:

$$f_{i,t}(b) = \frac{\exp\left(\gamma \left| b - b_t^H \right|\right)}{\int_0^w \exp\left(\gamma \left| x - b_t^H \right|\right) dx}. \tag{3.7}$$

It is important to note that the noiseless imitate-the-best dynamic is a special case of this model where $\gamma \to \infty$.

To examine the relative strength of the noisy imitation and optimization dynamics, we develop a combined model that includes each both imitation and optimization as a special case. In this combined model, the attraction of a bid x for an agent i at time t is given by

$$A_{it}(b) = \alpha \pi_i (b, b_{-i,t}) - \beta \left| b - b_{i,t} \right| - \gamma \left| b - b_t^H \right|. \tag{3.8}$$

Here b_{it} denotes the bid employed by agent i at time t and b_t^H denotes the bid employed by the agent who is earning the highest payoff at time t. The parameter α denotes the extent to which agents are more likely to select strategies that yield higher payoffs. The parameter β denotes the degree to which bids are autocorrelated, that is, the extent to which agents tend to select bids that are close to their previous bids.[3] The parameter γ captures denotes the tendency to imitate success, that is, the extent to which agents tend to pick bids which are close to the bid employed by the highest earning player. Accordingly, the likelihood that agent i will select a bid x when she makes an adjustment at time t is given by

$$f_{it}(b) = \frac{\exp\left(A_{it}(b)\right)}{\int_0^w \exp\left(A_{it}(x)\right) dx}. \tag{3.9}$$

[3]Results show the removal of this autocorrelation term does not affect the relative explanatory power of the imitative or logit terms (see Table 3.3).

3.3 Experimental Design and Procedures

3.3.1 Design

To implement the game discussed in Section 3.2, subjects were endowed with $w = \$10$ and competed for prizes with value $v = \$7$. Subject bids were bounded on the interval $[0, w]$. As this game takes place in continuous time, each session consisted of one continuous 40 minute period. During this period, subjects could adjust their bids as frequently as desired with the click of the mouse. Whenever a subject clicked, her bid instantaneously changed to the level corresponding to the horizontal position of her mouse, and the corresponding payoff rates were immediately recalculated.

The experiment consisted of two informational treatments. Under the global information treatment, each subject received real-time information regarding the bids and payoffs of every participant in her cohort. Under the local information treatment, subjects only observed their own bid and payoff. In both treatments, bids and payoffs were recorded at a rate of four times per second.

Figures 3.2 and 3.3 illustrate the experimental interface under the local-information and global-information treatments, respectively. The subject's current bid and payoff is represented by a blue line. The horizontal position of the blue line indicates the subject's current bid and the height of the blue line indicates the subject's current payoff. The subject's current bid and payoff are displayed numerically at the bottom of the screen. Under the global-information treatment, the current bid and payoff of each other subject is represented by a red line.

To provide random rematching in continuous time, we employ a mean matching protocol [e.g. 20, 21]. Each subject's instantaneous payoff is given by the expected value of her payoff from being randomly matched into a group of three agents. By the law of large numbers, high frequency mean-matching provides a superior approximation to truly

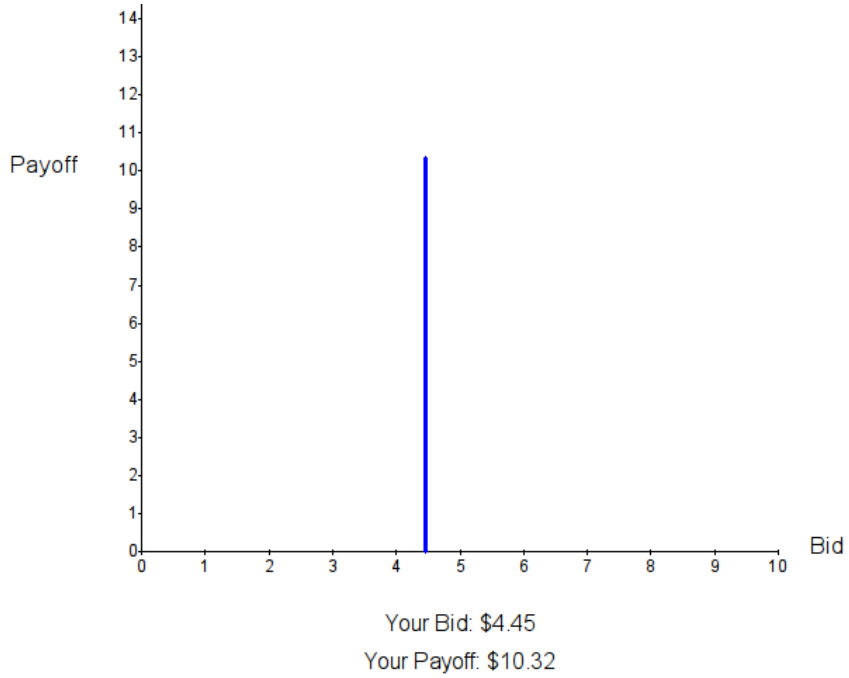

Figure 3.2: **User Interface Under Local Information**

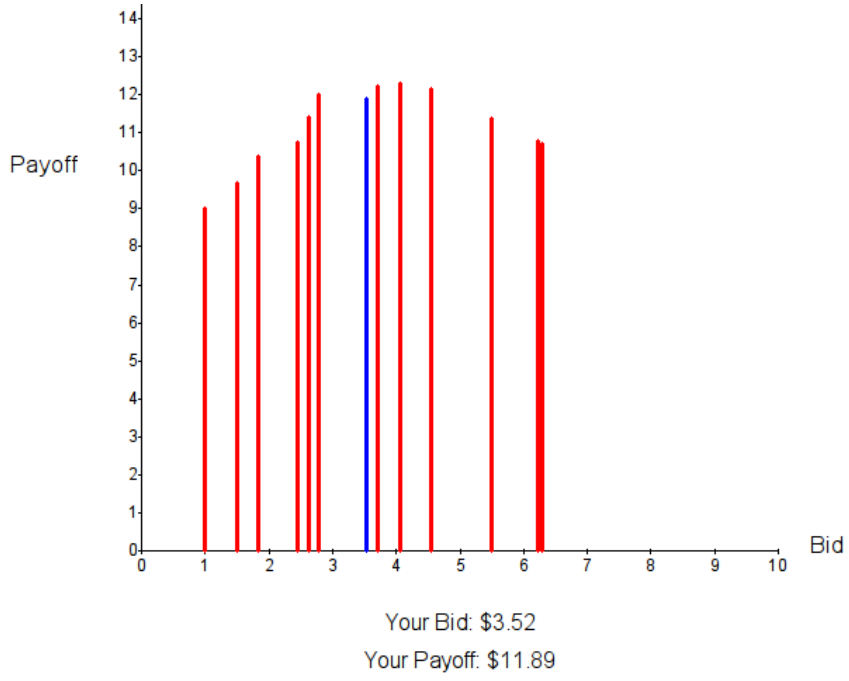

Figure 3.3: **User Interface Under Global Information**

continuous random matching than does high frequency random matching.

3.3.2 Procedures

Thirty subjects participated in one session of the global-information treatment and 27 subjects participated in one session of the local-information treatment. Subjects were recruited from the Texas A&M undergraduate population using an ORSEE database [66]. All sessions were run in the Texas A&M Economic Research Laboratory using z-Tree [67].

At the end of every session, each subject received the time average of their instantaneous payoff plus a five dollar show-up payment. Subject earnings averaged \$15.20 in the global-information treatment and \$16.09 in the local-information treatment, including the five dollar show-up payment. In equilibrium, average subject earnings would equal \$15.00, so subjects received slightly above equilibrium earnings under both treatments. All sessions lasted less than one hour.

3.4 Hypotheses

The game utilized in our experiment is not evolutionary stable, so adaptive models predict persistent disequilibrium rather than convergence to equilibrium. Figure 3.4 shows a heat map illustrating the changes in the predicted distribution of bidding behavior over time in our experiment for a population of 30 simulated agents. We employ the adaptive dynamics in equation 3.8 where $\alpha = 3$, $\beta = 0.3$, and $\gamma = 0.3$. In contrast to the static predictions of Nash and Quantal Response Equilibrium model (see section 3.2.1), the adaptive model predicts persistent bidding cycles.[4] This theoretical prediction results in the following hypothesis.

[4]The intuition for bidding cycles is as follows. When bids are sufficiently low agents can benefit by slightly outbidding their competitors, so competition gradually drives bids upwards. As bids gradually increase towards the value of the prize, average profits decrease. When profits became sufficiently low, agents can effectively opt-out of the auction by bidding close to zero, thus reinitializing the bidding cycle.

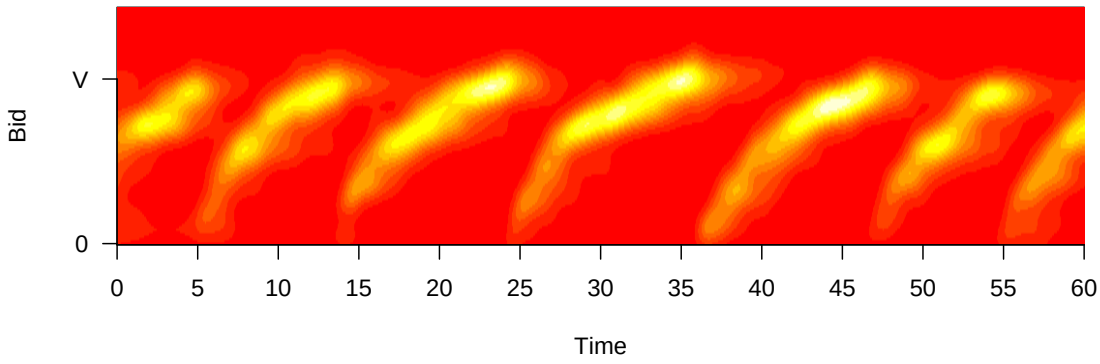

Figure 3.4: **Changes in the distribution of bidding behavior over time in a population of 30 simulated agents under adaptive dynamics obtained from a nonparametric conditional density estimator with a bid bandwidth of 0.5 and a time bandwidth of 0.3 seconds**

Hypothesis 1. *Subjects will exhibit persistent bidding cycles under both the local information treatment and the social information treatment.*

Throughout these disequilibrium cycles, adaptive models of imitation and optimization predict very different behavioral dynamics. Figure 3.5 provides an example of the predicted probability density of a new bid selected by player 1 at time t. The upper figure illustrates the predicted density under noisy imitative models and the lower figure illustrates the predicted density under noisy myopic optimization models. The horizontal position of each vertical line indicates the current bid of one player and the height of the line indicates the current payoff to this player. The shaded area under the curve indicates the probability density for a new bid selected by player 1.

Under the bidding profile depicted in figure 3.5, player 2 is has the highest payoff with bid b_2, so noisy imitation models predict that player 1 is likely to imitate this successful strategy by selecting a new bid that is close to b_2. In contrast, noisy optimization models

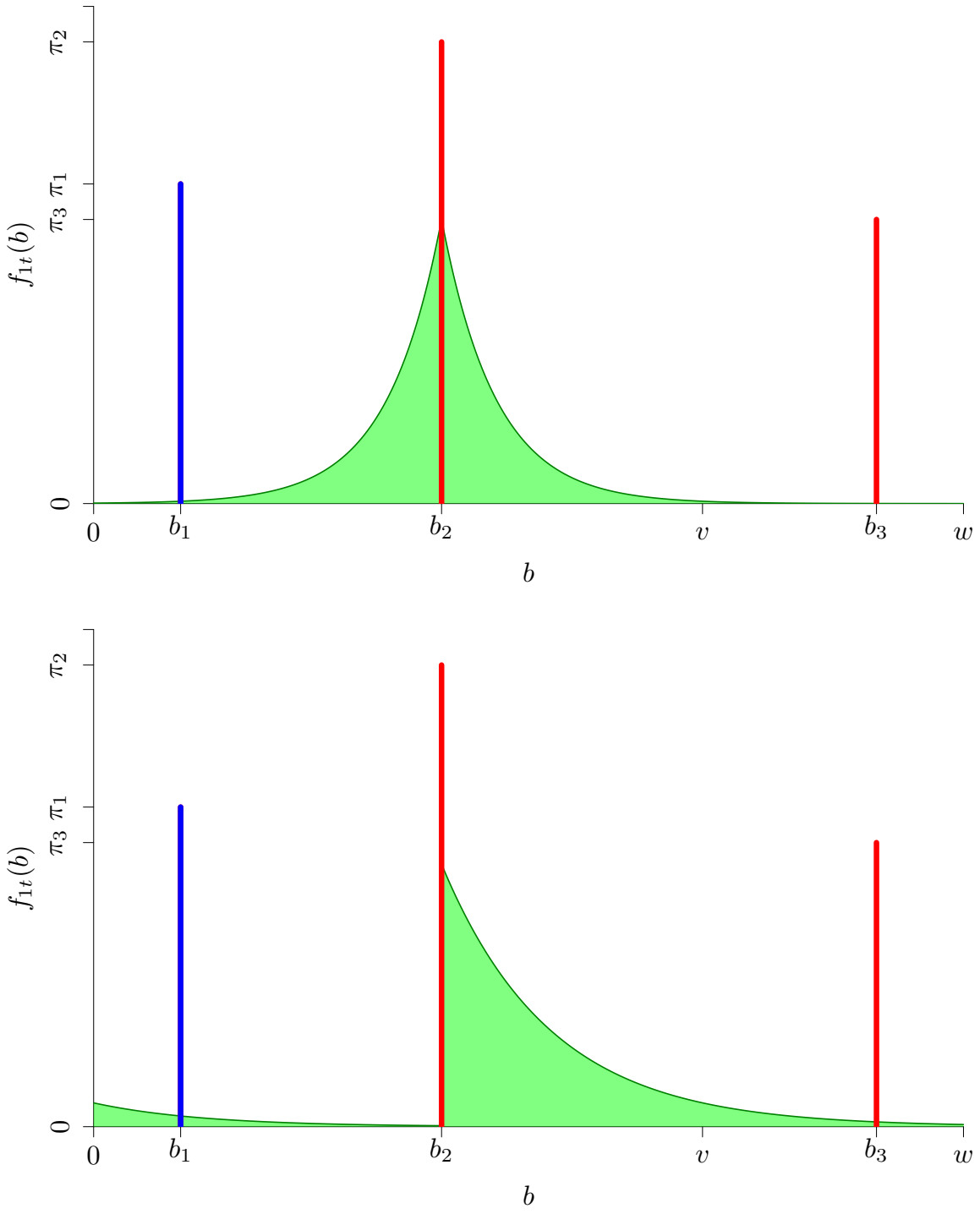

Figure 3.5: **The predicted probability density of a new bid selected by player 1 at time t. The top panel illustrates the predicted density under imitative models and the lower panel illustrates the predicted density under myopic optimization models.**

predict that player 1 is only likely to select bids that are slightly higher than b_2 as selecting a bid slightly lower than b_2 would not improve the payoff to player 1. This sharp contrast between the theoretical predictions from imitation and optimization in this game is part of what allows our experimental design to provide a uniquely powerful test of these adaptive models.

Since our experimental design allows us to disentangle imitation from optimization in observed subject behavior, we next ask how these models will perform under our different information treatments. Adaptive imitation models place a strong informational requirement on agents. As [68] notes, "imitation is a simple behavior that has two basic ingredients. One needs to be able to observe what others have done and one needs to be capable of doing what they have done." Hence adaptive imitation models require agents to observe the behavior of their peers. In contrast, adaptive optimization models describe agents who attempt to directly maximize their own payoff, so adaptive optimization models require agents only to observe their own payoffs.

Our local information treatment provides each subject with information regarding their own bids and payoffs, so we hypothesize that noisy optimization behavior (depicted in the lower panel of figure 3.5) will be observed in that treatment. In contrast, information regarding the bids and payoff of others is only provided in our global information treatment, so we hypothesize that noisy imitative behavior (depicted in the top panel of figure 3.5) will be observed in that treatment. Moreover, our global information treatment is designed to make implementing adaptive imitation as easy as possible; subjects only need to click on the highest bar on a computer screen (see figure 3.3 for a depiction of the interface) to implement adaptive imitation. In contrast, implementation of adaptive optimization is more computationally demanding for subjects, since it requires them to compute counterfactual payoffs from their information regarding the bids of others. Accordingly, we hypothesize that

Hypothesis 2. *Imitative models will outperform optimization models in the social information treatment, but not in the local information treatment.*

Finally, a growing literature [e.g., 69, 70, 71, 72] suggests that the provision of social information to economic agents helps agents to behave more rationally and come closer to the predictions of traditional economic theory and those of Nash equilibrium. For this reason, we speculate that the additional information provided by our social information treatment may help subjects to learn their way out of disequilibrium cycles and behave more consistently with the theoretical predictions of Nash equilibrium. This reasoning leads to our final hypothesis.

Hypothesis 3. *Behavior in the social information treatment will exhibit greater stability and greater consistency with Nash equilibrium.*

3.5 Results

Table 3.1 provides summary statistics for both the local information and global information treatments and a comparison with the equilibrium predictions. Recall that the Nash equilibrium of this game predicts that subjects will employ a mixed strategy with bids distributed according to the probability density function described in Section 3.2.1. On average, subjects in both of our treatments exhibit lower bidding than the equilibrium prediction. Consequently, the average earnings in both treatments are higher than the equilibrium prediction. In both treatments, we also observe instances of dominated bidding—bids above 7—which are never predicted to occur in equilibrium. Consistent with hypothesis 1, we do not observe a convergence to equilibrium in either treatment; the last 10 minutes of the experiment are not noticeably closer to equilibrium play than the first 10 minutes of the experiment. In general, there are only minor differences between the first and last 10 minutes of the experiment. Of those differences that exist (e.g., average bid

	Private Information			Global Information			Equilibrium Prediction
	initial 10 minutes	last 10 minutes	overall	initial 10 minutes	last 10 minutes	overall	
mean bid	3.73	3.54	3.57	4.57	4.31	4.47	4.67
bids above 7[a]	1.95%	0.45%	1.22%	5.57%	4.58%	6.96%	0.00%
minimum bids[b]	1.63%	3.77%	3.03%	0.72%	0.81%	0.90%	0.00%[c]
mean earnings	10.94	11.13	11.11	10.20	10.36	10.21	10.00

a. In this game, bids above 7 are always dominated by bidding 0.

b. The minimum bid is 0.

c. A bid of 0 is in the support of the mixed equilibrium strategy. However, the predicted occurrence of such bids by the equilibrium model is 0, because the strategy space is continuous.

Table 3.1: **Summary Statistics for Bids and Earnings in Local Information Treatment, Global Information Treatment, and Equilibrium Predictions.** The treatment statistics include groupings by the first 10 minutes and last 10 minutes to provide more detail about initial and final play.

decreasing, earnings increasing), most are moving away, rather than toward, equilibrium predictions.

Result 8. Both average payoffs and average bids differ significantly across treatments.

i. Subjects bids are higher and closer to the Nash equilibrium predictions in the global-information treatment.

ii. Payoffs in the global-information treatment are lower and closer to equilibrium than those in the local-information treatment.

Bids in the global-information treatment are significantly higher than those in the local information treatment. Table 3.1 shows that the mean bid in the global information treatment is \$0.90 higher than the mean bid in the local information treatment, so the former is closer to the equilibrium prediction. Subjects in the global information treatment are also more likely to select dominated bids above 7 and less likely to make 0 bids than those in the local information treatment. Figure 3.6 provides nonparametric estimates of the aggregate bid density under each treatment, showing that bids in the global

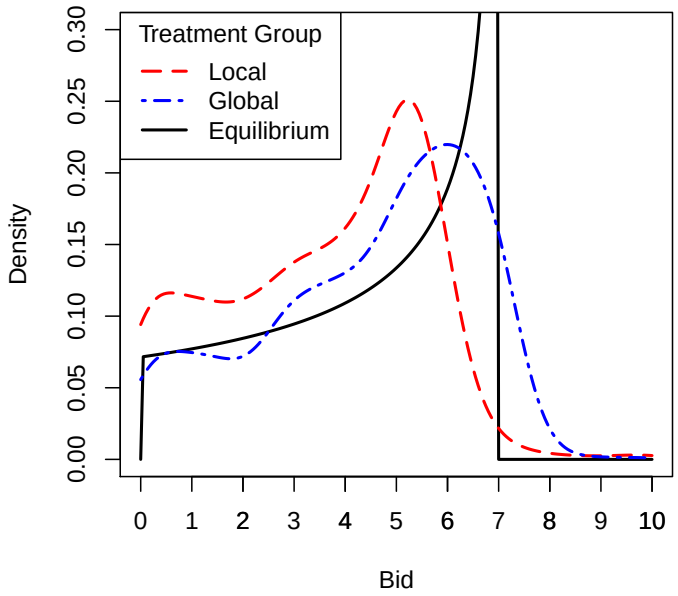

Figure 3.6: **Empirical Bid Distributions.** The density function is estimated using the local constant kernel density estimator of [1] and [2] with a normal kernel and a bandwidth of 0.5.

information treatment are generally larger than those in the local information treatment. A non-parametric Kolmogorov-Smirnov test finds the empirical bid distributions to be significantly different $(p < 0.01)$.

Figure 3.6 illustrates nonparametric estimates of the aggregate bid density under each treatment alongside the symmetric Nash equilibrium density function. As Table 3.1 implies, both of the observed bid distributions are generally lower than the equilibrium distribution, but with longer right tails. A non-parametric Kolmogorov-Smirnov test finds both empirically observed bid-distributions to be significantly different from the Nash equilibrium distribution $(p < 0.01)$. Moreover, neither bid distribution is consistent with logit quantal response equilibrium. A Kolmogorov-Smirnov test finds both the

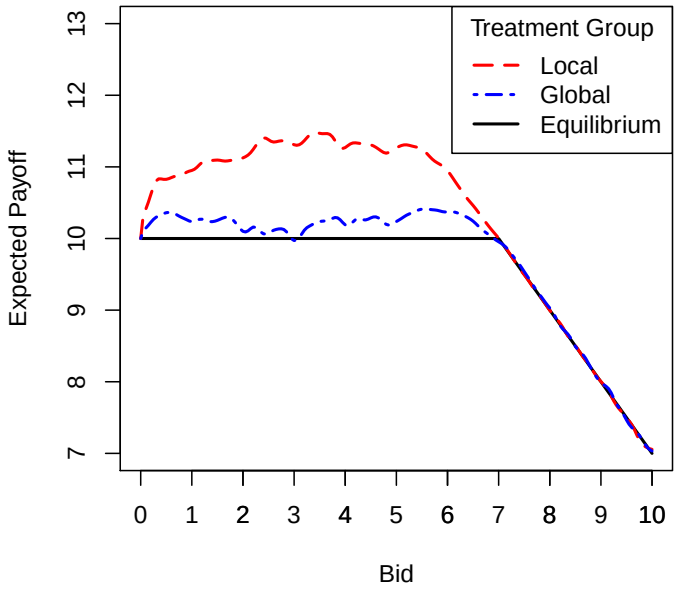

Figure 3.7: **Empirical Expected Payoff Functions under Alternate Treatments.** The expected payoff function is estimated using the local constant kernel regerssion estimator of [3] and [4] with a normal kernel and a bandwidth of 0.5

local-information-treatment bid distribution and the global-information-treatment bid distribution to be significantly different from their corresponding maximum-likelihood logit quantal response equilibrium predictions ($p < 0.01$; local information: $\eta = 0.792$, global information: $\eta = 0.505$).

Payoffs also differ significantly across treatments. Table 3.1 shows that mean payoffs in the local information treatment are \$0.90 higher than the global information treatment. Subjects in the global information treatment earned an average of \$15.21 while subjects in the local information treatment, subjects earned an average of \$16.11; significantly higher earnings at the one percent level. A non-parametric Kolmogorov-Smirnov test finds the empirical payoff distributions to be significantly different $(p < 0.01)$.

In equilibrium, every bid between zero and the value of the prize should yield the same expected payoff since rational agents must be indifferent between pure strategies over which they mix. Figure 3.7 shows that both treatments violate this indifference property. However, this violation is much more severe in the local information treatment than in the global information treatment, suggesting that subjects in the global information treatment are more precisely maximizing their payoffs. A non-parametric Kolmogorov-Smirnov test finds both empirically observed earnings distributions to be significantly different from the Nash equilibrium distribution ($p < 0.01$). Similarly, the empirical earnings distributions are also inconsistent with quantal response equilibrium. A Kolmogorov-Smirnov test finds both the local-information-treatment earnings distribution and the global-information-treatment earnings distribution to be significantly different from their corresponding maximum-likelihood, quantal response equilibrium predictions ($p < 0.01$; local information: $\eta = 0.792$, global information: $\eta = 0.505$).

Result 9. Throughout the 40 minute session, subject behavior in both treatments is characterized by a state of disequilibrium, resembling neither the Nash-equilibrium nor the logit quantal response equilibrium. There is no convergence to equilibrium; rather behavior in both treatments is characterized by persistent cycling.

Consistent with hypothesis 1, the bidding behavior observed in each treatment is characterized by persistent, identifiable, disequilibrium cycling. Figures 3.8 and 3.9 illustrate "heat maps" for the same periods as Figures 3.10(a-f). Moreover, note that the observed cycles in bidding behavior are consistent with the theoretical predictions from adaptive models illustrated by figure 3.4. In contrast to Hypothesis 3, these cycles are more rapid in the global-information treatment than in the local information treatment, so bidding behavior is actually less stable under the presence of social information.

One quantitative way to analyze the observed bidding behavior is to examine the time series of the mean bid employed by subjects. Figures 3.10(a-c) illustrate the dynamics of

the average bid in the global information treatment for the first, middle, and last minute, respectively, of the session. There is also a strong cyclical pattern to the mean bid in all three phases. Figures 3.10(d-f) provide the dynamics of the average bid in the local information treatment for the first, middle, and last minute, respectively, of the session. If subjects employ an equilibrium mixed strategy, then future changes in the mean bid should be uncorrelated with past changes in the mean bid. To test this hypothesis, we conduct the Ljung–Box test on the differenced time series of the mean bid. We find that the Ljung–Box test rejects the null hypothesis of uncorrelated changes in the mean bid at the one percent level under both treatments, suggesting that subjects exhibit significant disequilibrium dynamics in both treatments.

Result 10. Observed bidding dynamics differ significantly across treatments.

i. Bidding cycles have higher frequency in the global information treatment than in local information treatment.

ii. Bidding dynamics under the global information treatment exhibit far less behavioral noise than under the local information treatment.

The cycles observed in the aggregate bid data also differ across treatments. Figures 3.8 and 3.9 in the global information treatment are characterized by frequent cycles that appear to be about 5 seconds in length. The cycles are more noisy in Figures 3.10 (b,d,f) in the local information treatment and the distance from the peak of one cycle to the next can be as large as 20 seconds. The heat maps in Figures 3.8 and 3.9 also confirm these results. While cyclical patterns repeat about every 5 seconds in the global information treatment, they repeat about every 20-40 seconds in the private information treatment. Consistent with these results, the maximum likelihood estimates reported below indicate greater precision and less autocorrelation in the global information treatment, suggesting that the underlying adaptive processes are significantly different across treatments.

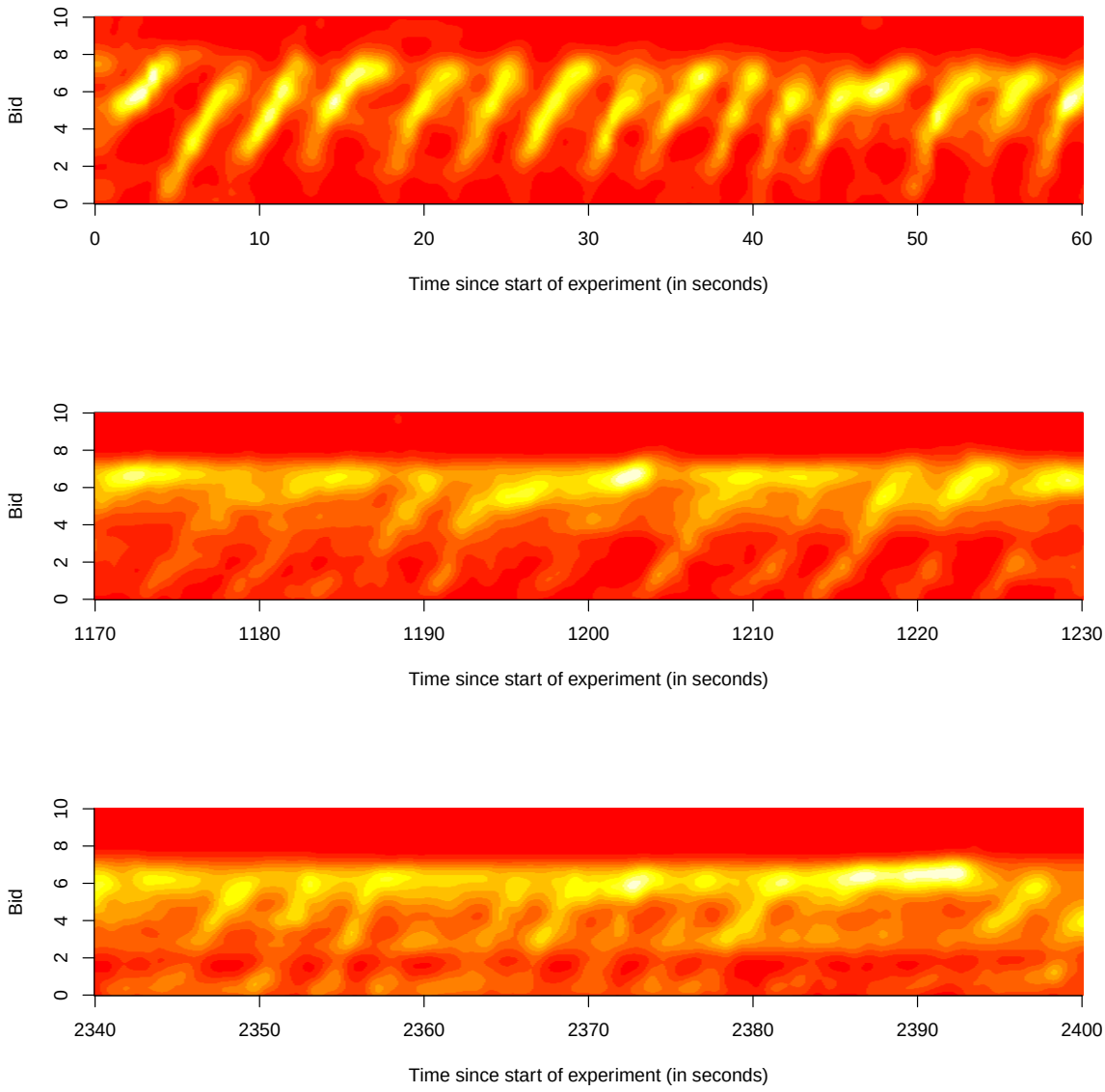

Figure 3.8: **Changes in the empirical distribution of bids over time under global information obtained from a nonparametric conditional density estimator with a bid bandwidth of 0.5 and a time bandwidth of 0.3 seconds.** Figures (a-c) depict the global information treatment for the first, median, and last minute, respectively.

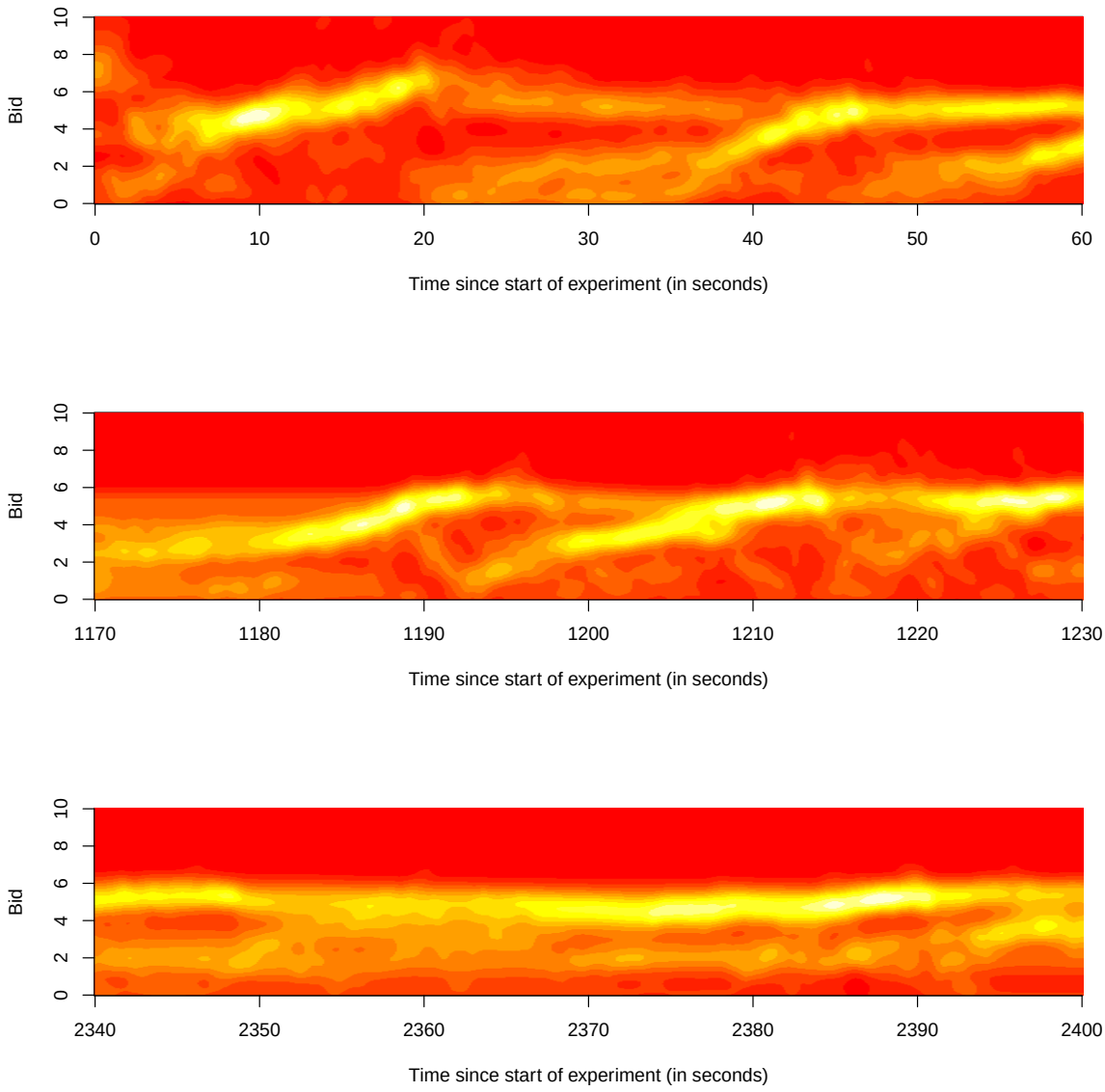

Figure 3.9: **Changes in the empirical distribution of bids over time under local information obtained from a nonparametric conditional density estimator with a bid bandwidth of 0.5 and a time bandwidth of 0.3 seconds.** Figures (a-c) depict the local information treatment for the first, median, and last minute, respectively.

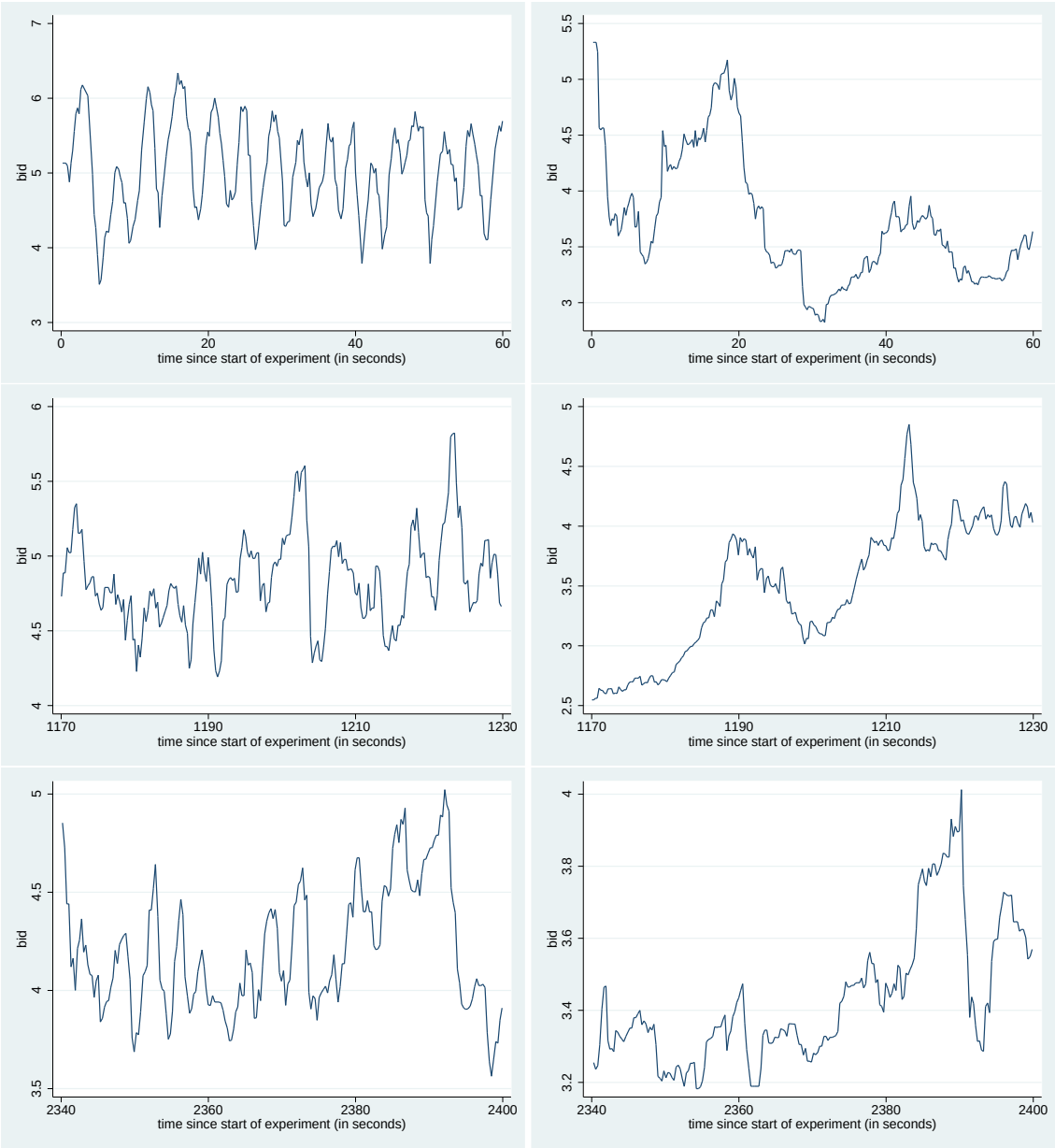

Figure 3.10: **Movement of the average bid in the first (top row), median (middle row), and last (bottom row) minutes of the global and local information treatments.** Figures (a, c, e) (left) show the global information treatment. Figures (b, d, f) (right) show the local information treatment.

	Global Information		Local Information	
	Optimization	Imitation	Optimization	Imitation
precision parameter (β)	1.54 (0.15)	0.60 (0.07)	0.68 (0.05)	0.33 (0.12)
observations	32634	32634	27225	27225
total log-likelihood	-58808.48	-66379.35	-57155.04	-58812.47

Table 3.2: **Maximum-Likelihood Models of Noisy optimization and Imitation Dynamics, Global-Information and Local-Information Treatments.** The noisy optimzation model outperforms the noisy imitation-response model. Both models perform better in the global information treatment than in the local information treatment. All parameters are significant at the 1% level. Standard errors are obtained via subject clustered bootstrap estimation.

Result 11. Behavior in both treatments is far more consistent with optimization than imitation.

It is useful to examine which factors best explain bidding behavior in the observed data. To that end, we estimate adaptive models of both noisy optimization and noisy imitation (see Section 3.2.2 for a full description of these models). Each model provides a continuous probability distribution that gives the likelihood $f_{it}(b)$ that a given bid b will be selected by a subject i at time t, based on a single precision parameter β (see equations 3.6 and 3.7 and figure 3.5 for more details). For a given β, the total likelihood for each of these models is the product of all observed $f_{i,t}(b_{i,t})$. The β^* that maximizes this likelihood function is the maximum likelihood precision parameter. Table 3.2 provides results for both models under both treatments of the experiment.

There is a clear ranking of these single parameter models in terms of how they explain the data. In contrast to hypothesis 2, optimization dynamics are a better predictor of disequilibrium subject behavior than noisy imitation dynamics under both treatments. Both models have more explanatory power in the global information treatment than in local information treatment, which is consistent with our finding of greater behavioral

noise in the local-information treatment than the global-information treatment (see Result 10 for details).

After examining simple one-parameter models separately, it is desirable to combine the noisy optimization, and noisy imitation models in a combined model. We also consider models that include an autocorrelation term accounting for the tendency of subjects to select new bids close to their previous bid. This term is especially relevant in the local information treatment where subjects are unable to observe the bids of others, and hence, tend to employ a trial and error strategy.[5]

Table 3.3 provides parameter estimates for the combined model in the local-information and global-information treatments. In the global-information treatment, subject bidding behavior is primarily driven by payoff incentives following the noisy optimization dynamic. In addition, there is also some degree to which individuals tend to select bids close to their own previously used bid. Under local information, bidding behavior is largely driven by autocorrelation with previously selected bids, with the payoffs under the noisy imitation dynamic as a secondary factor.

The difference in the explanatory power of the noisy optimization dynamic across treatments is not surprising. Subjects have the ability to directly maximize their payoff only when they have information regarding the bids employed by others. Thus, in the global-information treatment, they can directly respond to their payoff incentives. In the local-information treatment, they only receive information about the payoff they earn from the strategy they currently employ. Without further information about the strategies employed by others, subjects cannot easily determine how their payoff would change if they were to adjust their strategy. In this case, it makes sense that subjects would

[5]To make this point more salient, if the bid autocorrelation term were part of our one-parameter model comparison in Table 3.2, it would provide the greatest explanatory power in the local-information treatment. In the global information treatment, it would still outperform the imitative response model, but not the optimization response model.

	Global Information		Local Information	
logit (α)	1.37	1.48	0.56	0.48
(payoffs)	(0.11)	(0.14)	(0.07)	(0.05)
previous bids (β)	-	0.56	-	1.39
(subject specific)		(0.07)		(0.11)
imitation response (γ)	0.12	0.07	0.10	0.08
(the highest earning bid)	(0.03)	(0.03)	(0.02)	(0.02)
observations	32634	32634	27225	27225
log-likelihood	-58345.36	-49805.06	-56954.91	-33030.53
mean log-likelihood	-1.79	-1.52	-2.09	-1.21
typical bid likelihood	0.17	0.22	0.12	0.30

Table 3.3: **Multiple-Parameter Models of Bidding Dynamics.** A multi-parameter model including terms for logit and imitative dynamics is estimated on both the local and global information treatments. An additional specification includes a term for the tendency of subjects to make bids close to their previous bids. All parameters are significant at the 1% level. Standard errors are obtained via subject clustered bootstrap estimation.

experiment by making trial adjustments and then return to the strategies that provided the highest payoffs. This trial-and-error approach is consistent with the high autocorrelation of current bids and previous bids in the local-information treatment and it also explains how subjects are able to approximately best respond to their opponents' strategies without directly observing them.

Imitation explains very little of the observed behavior in either treatment. In the local information treatment, subjects see neither the payoff nor the strategy of any subject other than themselves, so the lack of imitation is unsurprising, since subjects can not directly implement the imitative model. However, in the global-information treatment, subjects need only to click on the highest bar to perfectly follow the imitative model, but the data indicate that subjects use something more complex than a simple imitation heuristic; they perform noisy myopic payoff optimization.

The autocorrelation in subject bids may explain why the imitative dynamic appeared

to have some explanatory power in a one-parameter model. Since bids tend to bunch together, as illustrated in Figure 3.8, the autocorrelation with a subject's own previous bid produces similar predictions to imitation of the bids employed by others. Consequently, a simple one-parameter imitation model with no autocorrelation parameter can misidentify autocorrelation in a subject's own bids for imitation of others. Results from the combined model presented in Table 3.3 suggest that much of the explanatory power attributed to imitation under the global informaton treatment actually results from autocorrelation with a subjects own previous bid.

3.6 Conclusion

This study experimentally investigates dynamic bidding behavior in continuous-time, all-pay auctions. In contrast to previous experimental studies of the all-pay auction, our subjects earned continuous flow payoffs and could adjust their bids asynchronously throughout the experiment. By permitting this type of asynchronous adjustment, we obtain a remarkably fine-grained picture of the empirical bidding behavior, allowing a close examination of behavioral bidding dynamics.

Consistent with theoretical predictions from adaptive models, but in contrast with both Nash and quantal response equilibrium predictions, subjects in our experiment exhibited persistent cyclical bidding behavior. This sustained disequilibrium behavior, along with the markedly discontinuous nature of payoff functions in the all-pay auction, allows us to closely investigate the predictive power of imitative and optimization dynamics. Surprisingly, behavior in the global-information treatment, which provides each subject with the information to easily employ imitative dynamics, is characterized by increased precision of optimization behavior but very little imitative behavior, resulting in higher bids, lower payoffs, and more rapid behavioral cycles.

Our results suggest a general failure of imitative models to adequately describe

human cognition in strategic settings. Subjects in the global information treatment could easily imitate the highest performing subject by selecting the highest line on a computer screen. However, instead of merely imitating successful strategies, subjects followed more sophisticated optimization methods, responding to the structure of their payoff incentives. In the local-information treatment, subjects do not have the necessary information to imitate other subjects. In the absence of social information, subjects employ trial-and-error strategies, selecting strategies near those that gave them higher payoffs. Subjects in the global information treatment compete more vigorously, their bidding cycles are far more rapid, and they exhibit far less behavioral noise. As a result, both average bids and average earnings are significantly closer to equilibrium predictions in the global information treatment than in the local information treatment.

While this experiment is primarily concerned with testing theoretical predictions, it also provides some interesting policy implications. In particular, these results suggest that policy makers may want to promote the distribution of social information in strategic environments where effort expenditure has positive externalities, such as patent races or competition for research grants. In contrast, policy makers may want to discourage the distribution of social information in strategic environments where effort expenditure is wasteful or has negative externalities, such as political lobbying or international warfare. Naturally, further research will be needed to verify the extent to which these experimental results carry over to other strategic environments.

3.7 Mathematical Appendix

3.7.1 Nash Equilibrium Derivation

Consider the following auction with three bidders and two prizes. Each player starts with an endowment w selects her bid from the closed interval $[0, w]$. After all three bids have been selected, the top two bidders each receive a prize with value $v < w$. However,

every player must pay her bid, regardless of whether or not she won a prize. In the case of a tie, the remaining prizes are randomly assigned among the tying players. Accordingly, the payoff function for player i is given by:

$$\pi_i(s_i, s_j, s_k) = \begin{cases} w - s_i + v & \text{if } s_i > \min\{s_j, s_k\} \\ w - s_i + 2v/3 & \text{if } s_i = s_j = s_k \\ w - s_i + v/2 & \text{if } s_i = \min\{s_j, s_k\} < \max\{s_j, s_k\} \\ w - s_i & \text{otherwise} \end{cases}$$

Suppose that there exists a continuous symmetric mixed strategy Nash equilibrium with support over the closed interval $[0, v]$. Let $F(z) = P(b_j < z)$ denote the corresponding cumulative distribution function. Let W_i denote the event that bidder i wins and receives and item. Let L_i denote the event that bidder i loses and does not receive and item. If bidder j and bidder k follow this Nash equilibrium mixed strategy, then the probability that bidder i loses the auction is given by:

$$\begin{aligned} P(L_i) &= P(b_i < b_j \text{ and } b_i < b_k) \\ &= P(b_i < b_j) P(b_i < b_k) = P(b_i < b_j)^2 \\ &= [1 - P(b_j < b_i)]^2 = [1 - F(b_i)]^2 \\ &= 1 - 2F(b_i) + F(b_i)^2 \end{aligned}$$

Now the probability that bidder i wins the auction is given by:

$$\begin{aligned} P(W_i) &= 1 - P(L_i) \\ &= 1 - \left[1 - 2F(b_i) + F(b_i)^2\right] \\ &= 2F(b_i) - F(b_i)^2 \end{aligned}$$

Hence bidder i's expected payoff, conditional on her bid, is given by:

$$\begin{aligned}
\pi_i(b_i) &= w + vP(W_i) - b_i \\
&= w + v\left[2F(b_i) - F(b_i)^2\right] - b_i \\
&= w + 2vF(b_i) - vF(b_i)^2 - b_i
\end{aligned}$$

If the mixed strategy F is a best response for agent i, then she must be indifferent between all of the bids in the support of F. Hence all of the bids in the closed interval $[0, v]$ must yield the same expected payoff for bidder i. Moreover, since bidding zero will certainly yield an expected payoff of zero, every bid between zero and v must yield an expected payoff of zero. Accordingly, we can write:

$$\begin{aligned}
\pi_i(b_i) &= 0 \qquad \text{for all } b_i \in [0, v] \\
2vF(b_i) - vF(b_i)^2 - b_i &= 0 \\
-b_i &= -2vF(b_i) + vF(b_i)^2 \\
-\frac{b_i}{v} &= -2F(b_i) + F(b_i)^2 \\
1 - \frac{b_i}{v} &= 1 - 2F(b_i) + F(b_i)^2 \\
\sqrt{1 - \frac{b_i}{v}} &= 1 - F(b_i) \\
F(b_i) &= 1 - \sqrt{1 - \frac{b_i}{v}} \qquad \text{for all } b_i \in [0, v]
\end{aligned}$$

Differentiating this cumulative distribution function obtains the Nash equilibrium probability density function:

$$f(b_i) = \frac{1}{2v}\left(1 - \frac{b_i}{v}\right)^{-1/2} \qquad \text{for all } b_i \in [0, v]$$

3.7.2 Logit Quantal Response Equilibrium Derivation

Consider the following auction with three bidders and two prizes. Each player starts with an endowment w selects her bid from the closed interval $[0, w]$. After all three bids have been selected, the top two bidders each receive a prize with value $v < w$. However, every player must pay her bid, regardless of whether or not she won a prize. In the case of a tie, the remaining prizes are randomly assigned among the tying players. Accordingly, the payoff function for player i is given by:

$$\pi_i\left(s_i, s_j, s_k\right) = \begin{cases} w - s_i + v & \text{if } s_i > \min\left\{s_j, s_k\right\} \\ w - s_i + 2v/3 & \text{if } s_i = s_j = s_k \\ w - s_i + v/2 & \text{if } s_i = \min\left\{s_j, s_k\right\} < \max\left\{s_j, s_k\right\} \\ w - s_i & \text{otherwise} \end{cases}$$

Suppose that there exists a continuous symmetric logit quantal response equilibrium with support over the closed interval $[0, w]$. Let $F(z) = P(b_j < z)$ denote the corresponding cumulative distribution function. Let W_i denote the event that bidder i wins and receives and item. Let L_i denote the event that bidder i loses and does not receive and item. If bidder j and bidder k follow this mixed strategy, then the probability that bidder i loses the auction is given by:

$$\begin{aligned} P(L_i) &= P(b_i < b_j \text{ and } b_i < b_k) \\ &= P(b_i < b_j) P(b_i < b_k) = P(b_i < b_j)^2 \\ &= [1 - P(b_j < b_i)]^2 = [1 - F(b_i)]^2 \\ &= 1 - 2F(b_i) + F(b_i)^2 \end{aligned}$$

Accordingly, the probability that bidder i wins the auction is given by:

$$\begin{aligned} P(W_i) &= 1 - P(L_i) \\ &= 1 - \left[1 - 2F(b_i) + F(b_i)^2\right] \\ &= 2F(b_i) - F(b_i)^2 \end{aligned}$$

Hence bidder i's expected payoff, conditional on her bid, is given by:

$$\begin{aligned} \pi_i(b_i) &= w + vP(W_i) - b_i \\ &= w + v\left[2F(b_i) - F(b_i)^2\right] - b_i \\ &= w + 2vF(b_i) - vF(b_i)^2 - b_i \end{aligned}$$

Under a logit quantal response equilibrium, agents do not always select their best response, but they are more likely to select bids that yield higher payoffs. Here the level of behavioral noise is indexed by the parameter η. As η approaches infinity, the logit quantal response equilibrium approaches uniformly random behavior. As η approaches zero, the logit quantal response equilibrium approximates a Nash equilibrium. Formally, we can write:

$$\begin{aligned} f(b) &= \frac{\exp(\eta\pi_i(b))}{\int_0^w \exp(\eta\pi_i(x))\,dx} \\ f(b) &= \frac{\exp\left(\eta\left(2vF(b) - vF(b)^2 - b\right)\right)}{C_0} \\ C_0 f(b) &= \exp\left(2\eta vF(b) - \eta vF(b)^2 - \eta b\right) \\ C_0\frac{dF}{db} &= \exp\left(2\eta vF(b) - \eta vF(b)^2 - \eta b\right) \end{aligned}$$

Integrating both sides of this differential equation obtains

$$\begin{aligned}
C_0 \int \exp\left(\eta v F^2 - 2\eta v F\right) dF &= \exp\left(-\eta b\right) db \\
\eta C_0 \int \exp\left(\eta v F^2 - 2\eta v F\right) dF &= \frac{1}{\eta} - \frac{1}{\eta}\exp\left(-\eta b\right) \\
C_1 \exp\left(\eta v\right) \int \exp\left(\eta v \left(F-1\right)^2\right) dF &= \frac{1}{\eta} - \frac{1}{\eta}\exp\left(-\eta b\right) \\
\int \exp\left(\eta v (F-1)^2\right) dF &= C_3 - C_4 \exp\left(-\eta b\right)
\end{aligned}$$

We can solve for the cumulative distribution function F in terms of the imaginary error function by introducing the function $G\left(x\right) = \int_0^x \exp\left(u^2\right) du = \frac{\sqrt{\pi}}{2}\text{erfi}\left(x\right)$.

$$\begin{aligned}
G\left(\sqrt{\eta v}\left(F-1\right)\right) &= C_3 - C_4 \exp\left(-\eta b\right) \\
\sqrt{\eta v}\left(F-1\right) &= G^{-1}\left(C_3 - C_4 \exp\left(-\eta b\right)\right) \\
F - 1 &= \frac{1}{\sqrt{\eta v}} G^{-1}\left(C_3 - C_4 \exp\left(-\eta b\right)\right) \\
F\left(b\right) &= 1 - \frac{1}{\sqrt{\eta v}} G^{-1}\left(C_3 - C_4 \exp\left(-\eta b\right)\right)
\end{aligned}$$

Now since bids are restricted to be non-negative, we have

$$\begin{aligned}
F\left(0\right) &= 0 \\
1 - \frac{1}{\sqrt{\eta v}} G^{-1}\left(C_3 - C_4 \exp\left(0\right)\right) &= 0 \\
\frac{1}{\sqrt{\eta v}} G^{-1}\left(C_3 - C_4\right) &= 1 \\
G^{-1}\left(C_3 - C_4\right) &= \sqrt{\eta v} \\
C_3 - C_4 &= G\left(\sqrt{\eta v}\right) \\
C_3 &= C_4 + G\left(\sqrt{\eta v}\right)
\end{aligned}$$

Similarly, since bids cannot exceed the endowment w, we have

$$\begin{aligned}
F(w) &= 1 \\
1 - \frac{1}{\sqrt{\eta v}} G^{-1}\left(C_3 - C_4 \exp(-\eta w)\right) &= 1 \\
\frac{1}{\sqrt{\eta v}} G^{-1}\left(C_3 - C_4 \exp(-\eta w)\right) &= 0 \\
G^{-1}\left(C_3 - C_4 \exp(-\eta w)\right) &= 0 \\
C_3 - C_4 \exp(-\eta w) &= G(0) = 0 \\
C_4 + G(\sqrt{\eta v}) - C_4 \exp(-\eta w) &= 0 \qquad \text{since } C_3 = C_4 + G(\sqrt{\eta v}) \\
G(\sqrt{\eta v}) &= C_4(\exp(-\eta w) - 1) \\
C_4 &= \frac{G\left(\sqrt{\eta v}\right)}{\exp(-\eta w) - 1}
\end{aligned}$$

We can use these solutions for C_3 and C_4 to obtain a closed form solution for the cumulative distribution function F

$$\begin{aligned}
F(b) &= 1 - \frac{1}{\sqrt{\eta v}} G^{-1}\left(C_3 - C_4 \exp(-\eta b)\right) \\
F(b) &= 1 - \frac{1}{\sqrt{\eta v}} G^{-1}\left(C_4 + G(\sqrt{\eta v}) - C_4 \exp(-\eta b)\right) \qquad \text{since } C_3 = C_4 + G(\sqrt{\eta v}) \\
F(b) &= 1 - \frac{1}{\sqrt{\eta v}} G^{-1}\left(G(\sqrt{\eta v})\left[1 - \frac{1 - \exp(-\eta b)}{1 - \exp(-\eta w)}\right]\right) \qquad \text{since } C_4 = \frac{G\left(\sqrt{\eta v}\right)}{\exp(-\eta w) - 1}
\end{aligned}$$

Differentiating the cumulative distribution function F obtains the corresponding

probability density function

$$
\begin{aligned}
f(b) &= F'(b) = -\frac{1}{\sqrt{\eta v}}\frac{\partial}{\partial b}\left[G^{-1}\left(G\left(\sqrt{\eta v}\right)\left[1-\frac{1-\exp(-\eta b)}{1-\exp(-\eta w)}\right]\right)\right] \\
f(b) &= -\frac{1}{\sqrt{\eta v}}\frac{\partial}{\partial b}\left[G^{-1}(H(b))\right] \\
f(b) &= -\frac{1}{\sqrt{\eta v}}\frac{\partial G^{-1}(H(b))}{\partial H(b)}H'(b) \\
f(b) &= -\frac{1}{\sqrt{\eta v}}G'(G^{-1}(H(b)))^{-1}H'(b) \\
f(b) &= -\frac{1}{\sqrt{\eta v}}\exp(G^{-1}(H(b))^2)^{-1}H'(b) \qquad \text{since } G'(x) = \exp\left(x^2\right) \\
f(b) &= -\frac{1}{\sqrt{\eta v}}\exp(-G^{-1}(H(b))^2)H'(b) \\
f(b) &= -\frac{1}{\sqrt{\eta v}}\exp(-G^{-1}(H(b))^2)H'(b) \\
f(b) &= \frac{\eta G\left(\sqrt{\eta v}\right)\exp(-\eta b)}{\sqrt{\eta v}\left[1-\exp(-\eta w)\right]}\exp(-G^{-1}(H(b))^2) \qquad \text{since } H'(b) = -\frac{\eta G\left(\sqrt{\eta v}\right)\exp(-\eta b)}{\left[1-\exp(-\eta w)\right]} \\
f(b) &= \frac{\eta G\left(\sqrt{\eta v}\right)\exp(-\eta b)}{\sqrt{\eta v}\left[1-\exp(-\eta w)\right]}\exp\left(-G^{-1}\left(G\left(\sqrt{\eta v}\right)\left[1-\frac{1-\exp(-\eta b)}{1-\exp(-\eta w)}\right]\right)^2\right)
\end{aligned}
$$

3.7.3 Evolutionary Instability of the Nash Equilibrium

Intuitively, a strategy is evolutionarily stable if it induces a self-enforcing convention. In other words, a strategy x is evolutionarily stable if no other strategy y can invade it when the entire population initially employs strategy x. More formally, in a in a symmetric normal form game, a strategy x is evolutionarily stable if there exists some $C \in (0,1)$ such that for all $\varepsilon \in (0, C)$ and for any other strategy y

$$\pi\left(x \mid \epsilon y + (1-\epsilon)x\right) > \pi\left(y \mid \epsilon y + (1-\epsilon)x\right) \tag{3.10}$$

Thus, if x is evolutionarily stable and a sufficiently small proportion of the population deviates to an alternate strategy y, then agents who employ x will earn a strictly higher

payoff than agents who employ y.

The Nash equilibrium strategy for the all-pay auction is not evolutionarily stable. To see why, suppose that a small proportion ϵ of the population deviates from the Nash equilibrium strategy x to an alternate strategy y under which agents always bid the full value of the prize. Since the support of the equilibrium bid distribution is given by the closed interval $[0, v]$, agents who employ the invading strategy y will win the prize with probability one whenever they are matched against an agent who employs the equilibrium bidding strategy. So the expected payoff to an agent who deviates to strategy y is given by

$$\begin{aligned}
\pi\left(\, y \mid \varepsilon y + (1-\varepsilon)x \,\right) &= \varepsilon^2 \pi_1(y,y,y) + 2\varepsilon(1-\varepsilon)\pi_1(y,y,x) + (1-\varepsilon)^2 \pi_1(y,x,x) \\
\pi\left(\, y \mid \varepsilon y + (1-\varepsilon)x \,\right) &= \varepsilon^2 \pi_1(y,y,y) \qquad \text{since } \pi(y,y,x) = \pi_1(y,x,x) = 0 \\
\pi\left(\, y \mid \varepsilon y + (1-\varepsilon)x \,\right) &= -\frac{\varepsilon^2 v}{3} \qquad \text{since } \pi_1(y,y,y) = -\frac{v}{3}
\end{aligned}$$

On the other hand, the expected payoff to an agent who employs the equilibrium mixed strategy is given by

$$\begin{aligned}
\pi\left(\, x \mid \varepsilon y + (1-\varepsilon)x \,\right) &= \varepsilon^2 \pi_1(x,y,y) + 2\varepsilon(1-\varepsilon)\pi_1(x,y,x) + (1-\varepsilon)^2 \pi_1(x,x,x) \\
\pi\left(\, x \mid \varepsilon y + (1-\varepsilon)x \,\right) &= \varepsilon^2 \pi_1(x,y,y) + 2\varepsilon(1-\varepsilon)\pi_1(x,y,x) \qquad \text{since } \pi_1(x,x,x) = 0 \\
\pi\left(\, x \mid \varepsilon y + (1-\varepsilon)x \,\right) &< -\varepsilon^2 \pi_1(x,y,y) \qquad \text{since } \pi_1(x,y,x) < 0 \\
\pi\left(\, x \mid \varepsilon y + (1-\varepsilon)x \,\right) &< -\varepsilon^2 E\{bid|x\} \qquad \text{since } \pi_1(x,y,y) = -E\{bid|x\} \\
\pi\left(\, x \mid \varepsilon y + (1-\varepsilon)x \,\right) &< -\frac{2\varepsilon^2 v}{3} \qquad \text{since } E\{bid|x\} = \frac{2v}{3} \\
\pi\left(\, x \mid \varepsilon y + (1-\varepsilon)x \,\right) &< \pi_1\left(y, (\varepsilon y + (1-\varepsilon)x)^2\right) \qquad \text{since } -\frac{2\varepsilon^2 v}{3} < -\frac{\varepsilon^2 v}{3}
\end{aligned}$$

Thus the invading strategy y earns a higher expected payoff than the equilibrium mixed strategy x, so the equilibrium mixed strategy for the all-pay auction with three bidders and

two prizes is not evolutionarily stable. Hence the mixed strategy Nash equilibrium does not induce a self enforcing convention in this all-pay auction. Accordingly, we expect to observe dynamic instability in experimental bidding behavior.

REFERENCES

[1] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *Annals of Mathematical Statistics*, vol. 27, no. 3, pp. 832–837, 1956.

[2] E. Parzen, "On estimation of a probability density function and mode," *Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076, 1962.

[3] E. A. Nadaraya, "On estimating regression," *Theory of Probability & Its Applications*, vol. 9, no. 1, pp. 141–142, 1964.

[4] G. S. Watson, "Smooth regression analysis," *Sankhyā: The Indian Journal of Statistics, Series A*, pp. 359–372, 1964.

[5] R. Aumann and A. Brandenburger, "Epistemic conditions for nash equilibrium," *Econometrica*, vol. 63, no. 5, pp. 1161–80, 1995.

[6] A.-A. Cournot, *Researches into the Mathematical Principles of the Theory of Wealth*. chez L. Hachette, 1838.

[7] J. Nash, "Equilibrium points in n-person games," *Proc. Nat. Acad. Sci. USA*, vol. 36, no. 1, pp. 48–49, 1950.

[8] W. H. Sandholm, *Population games and evolutionary dynamics*. MIT press, 2010.

[9] D. Friedman, S. Huck, R. Oprea, and S. Weidenholzer, "From imitation to collusion: Long-run learning in a low-information environment," *Journal of Economic Theory*, vol. 155, pp. 185–205, 2015.

[10] R. D. McKelvey and T. R. Palfrey, "Quantal response equilibria for normal form games," *Games and Economic Behavior*, vol. 10, no. 1, pp. 6–38, 1995.

[11] D. O. Stahl and P. W. Wilson, "Experimental evidence on players' models of other players," *Journal of economic behavior & organization*, vol. 25, no. 3, pp. 309–327, 1994.

[12] R. Nagel, "Unraveling in guessing games: An experimental study," *The American Economic Review*, vol. 85, no. 5, pp. 1313–1326, 1995.

[13] C. F. Camerer, T.-H. Ho, and J.-K. Chong, "A cognitive hierarchy model of games," *The Quarterly Journal of Economics*, pp. 861–898, 2004.

[14] M. Benaïm, J. Hofbauer, and E. Hopkins, "Learning in games with unstable equilibria," *Journal of Economic Theory*, vol. 144, no. 4, pp. 1694–1709, 2009.

[15] I. Gilboa and A. Matsui, "Social stability and equilibrium," pp. 859–867.

[16] M. J. Smith, "The stability of a dynamic model of traffic assignment-an application of a method of lyapunov," *Transportation Science*, vol. 18, no. 3, pp. 245–252, 1984.

[17] D. Fudenberg and D. K. Levine, *The theory of learning in games*. MIT press, 1998, vol. 2.

[18] J. B. Van Huyck, R. C. Battalio, and R. O. Beil, "Tacit coordination games, strategic uncertainty, and coordination failure," *The American Economic Review*, vol. 80, no. 1, pp. 234–248, 1990.

[19] V. P. Crawford, "An evolutionary interpretation of van huyck, battalio, and beil's experimental results on coordination," *Games and Economic behavior*, vol. 3, no. 1, pp. 25–59, 1991.

[20] R. Oprea, K. Henwood, and D. Friedman, "Separating the hawks from the doves: Evidence from continuous time laboratory games," *Journal of Economic Theory*, vol. 146, no. 6, pp. 2206–2225, 2011.

[21] T. Cason, D. Friedman, and E. Hopkins, "Cycles and instability in a rock-paper-scissors population game: A continuous time experiment," *Review of Economic Studies*, vol. 81, no. 1, pp. 112–136, 2013.

[22] D. Stephenson and A. L. Brown, "Characterizing disequilibrium dynamics: Imitation and optimization in continuous-time all-pay auctions," *Working Paper*, 2016.

[23] C. Alos-Ferrer and A. B. Ania, "The evolutionary stability of perfectly competitive

behavior," *Economic Theory*, vol. 26, no. 3, pp. 497–516, 2005.

[24] A. Antoci, P. Russu, and L. Zarri, "Tax evasion in a behaviorally heterogeneous society: An evolutionary analysis," *Economic Modelling*, vol. 42, pp. 106–115, 2014.

[25] J. M. Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, p. 15, 1973.

[26] P. D. Taylor, "Evolutionarily stable strategies with two types of player," *Journal of applied probability*, pp. 76–83, 1979.

[27] E. Ibidunmoye, B. Alese, and O. Ogundele, "Modeling attacker-defender interaction as a zero-sum stochastic game," *Journal of Computer Sciences and Applications*, vol. 1, no. 2, pp. 27–32, 2013.

[28] R. Powell, "Defending against terrorist attacks with limited resources," *American Political Science Review*, vol. 101, no. 03, pp. 527–541, 2007.

[29] J. Bjornerstedtt and J. Weibull, "Nash equilibrium and evolution by imitation," 1994.

[30] W. H. Sandholm, "Pairwise comparison dynamics and evolutionary foundations for nash equilibrium," *Games*, vol. 1, no. 1, pp. 3–17, 2009.

[31] S. R. Bulò and I. M. Bomze, "Infection and immunization: a new class of evolutionary game dynamics," *Games and Economic Behavior*, vol. 71, no. 1, pp. 193–211, 2011.

[32] G. W. Brown, "Iterative solution of games by fictitious play," *Activity analysis of production and allocation*, vol. 13, no. 1, pp. 374–376, 1951.

[33] C. Bruni, J. C. Nuño, and M. Primicerio, "What if criminals optimize their choice? optimal strategies to defend public goods," *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 4, pp. 840–850, 2013.

[34] A. Abdulkadiroglu, P. Pathak, A. E. Roth, and T. Sonmez, "Changing the boston school choice mechanism," National Bureau of Economic Research, Tech. Rep., 2006.

[35] Y. Chen and T. Sönmez, "School choice: an experimental study," *Journal of Economic theory*, vol. 127, no. 1, pp. 202–231, 2006.

[36] J. Pais and Á. Pintér, "School choice and information: An experimental study on matching mechanisms," *Games and Economic Behavior*, vol. 64, no. 1, pp. 303–328, 2008.

[37] U. Dur, R. G. Hammond, and T. Morrill, "Identifying the harm of manipulable school-choice mechanisms," mimeo, Tech. Rep., 2015.

[38] B. Gong and Y. Liang, "A dynamic college admission mechanism in inner mongolia: Theory and experiment," *Working Paper*, 2016.

[39] H. Ergin and T. Sönmez, "Games of school choice under the boston mechanism," *Journal of public Economics*, vol. 90, no. 1, pp. 215–237, 2006.

[40] A. Abdulkadiroglu and T. Sönmez, "School choice: A mechanism design approach," *The American Economic Review*, vol. 93, no. 3, pp. 729–747, 2003.

[41] L. Shapley and H. Scarf, "On cores and indivisibility," *Journal of mathematical economics*, vol. 1, no. 1, pp. 23–37, 1974.

[42] D. Gale and L. S. Shapley, "College admissions and the stability of marriage," *The American Mathematical Monthly*, vol. 69, no. 1, pp. 9–15, 1962.

[43] A. E. Roth, "The economics of matching: Stability and incentives," *Mathematics of operations research*, vol. 7, no. 4, pp. 617–628, 1982.

[44] A. Matsui, "Best response dynamics and socially stable strategies," *Journal of Economic Theory*, vol. 57, no. 2, pp. 343–362, 1992.

[45] A. Vanacore, "Centralized enrollment in recovery school district gets first tryout," *Times-Picayune, April*, vol. 16, 2012.

[46] A. E. Roth, "Deferred acceptance algorithms: History, theory, practice, and open questions," *international Journal of game Theory*, vol. 36, no. 3-4, pp. 537–569, 2008.

[47] J. Nash, "Non-cooperative games," *Annals of Mathematics*, vol. 54, no. 2, pp. 286–295, 1951.

[48] E. Dechenaux, D. Kovenock, and R. M. Sheremeta, "A survey of experimental research on contests, all-pay auctions and tournaments," *Experimental Economics*, vol. 18, no. 4, pp. 609–669, 2014.

[49] U. Gneezy and R. Smorodinsky, "All-pay auctions: an experimental study," *Journal of Economic Behavior & Organization*, vol. 61, no. 2, pp. 255–275, 2006.

[50] V. Lugovskyy, D. Puzzello, and S. Tucker, "An experimental investigation of overdissipation in the all pay auction," *European Economic Review*, vol. 54, no. 8, pp. 974–997, 2010.

[51] C. Ernst and C. Thöni, "Bimodal bidding in experimental all-pay auctions," *Games*, vol. 4, no. 4, pp. 608–623, 2013.

[52] M. R. Baye, D. Kovenock, and C. G. De Vries, "Rigging the lobbying process: an application of the all-pay auction," *American Economic Review*, vol. 83, no. 1, pp. 289–294, 1993.

[53] M. Marinucci and W. Vergote, "Endogenous network formation in patent contests and its role as a barrier to entry," *Journal of Industrial Economics*, vol. 59, no. 4, pp. 529–551, 2011.

[54] K. Chatterjee, J. G. Reiter, and M. A. Nowak, "Evolutionary dynamics of biological auctions," *Theoretical Population Biology*, vol. 81, no. 1, pp. 69–80, 2012.

[55] R. Hodler and H. Yektaş, "All-pay war," *Games and Economic Behavior*, vol. 74, no. 2, pp. 526–540, 2012.

[56] D. J. Clark and C. Riis, "Competition over more than one prize," *American Economic Review*, vol. 88, no. 1, pp. 276–289, 1998.

[57] G. Lopez, "Quantal response equilibria for models of price competition," *Unpublished Ph. D. dissertation, University of Virginia*, 1995.

[58] S. P. Anderson, J. K. Goeree, and C. A. Holt, "Rent seeking with bounded rationality: An analysis of the all-pay auction," *Journal of Political Economy*, vol. 106, no. 4, pp. 828–853, 1998.

[59] S. Demichelis and J. W. Weibull, "Language, meaning, and games: A model of communication, coordination, and evolution," *American Economic Review*, vol. 98, no. 4, pp. 1292–1311, 2008.

[60] R. Parayre and D. Hurry, "Corporate investment and strategic stability in hypercompetition," *Managerial and Decision Economics*, vol. 22, no. 4-5, pp. 281–298, 2001.

[61] R. Golman, "Why learning doesnŠt add up: equilibrium selection with a composition of learning rules," *International Journal of Game Theory*, vol. 40, no. 4, pp. 719–733, 2011.

[62] E. Hopkins, "A note on best response dynamics," *Games and Economic Behavior*, vol. 29, no. 1, pp. 138–150, 1999.

[63] P. Duersch, J. Oechssler, and B. C. Schipper, "Unbeatable imitation," *Games and Economic Behavior*, vol. 76, no. 1, pp. 88–96, 2012.

[64] P. D. Taylor and L. B. Jonker, "Evolutionary stable strategies and game dynamics," *Mathematical Biosciences*, vol. 40, no. 1, pp. 145–156, 1978.

[65] J. Oechssler and F. Riedel, "Evolutionary dynamics on infinite strategy spaces," *Economic Theory*, vol. 17, no. 1, pp. 141–162, 2001.

[66] B. Greiner, "Subject pool recruitment procedures: organizing experiments with orsee," *Journal of the Economic Science Association*, vol. 1, no. 1, pp. 114–125, 2015.

[67] U. Fischbacher, "z-tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, vol. 10, no. 2, pp. 171–178, 2007.

[68] K. Schlag, "Imitation and social learning," in *Encyclopedia of the Sciences of Learning*, N. Seel, Ed. Springer Science & Business Media, 2011, pp. 1489–1493.

[69] A. Merlo and A. Schotter, "Learning by not doing: an experimental investigation of observational learning," *Games and Economic Behavior*, vol. 42, no. 1, pp. 116–136, 2003.

[70] O. Armantier, "Does observation influence learning?" *Games and Economic Behavior*, vol. 46, no. 2, pp. 221–239, 2004.

[71] E. Cardella, "Learning to make better strategic decisions," *Journal of Economic Behavior & Organization*, vol. 84, no. 1, pp. 382–392, 2012.

[72] A. L. Brown, Z. E. Chua, and C. F. Camerer, "Learning and visceral temptation in dynamic saving experiments," *The Quarterly Journal of Economics*, pp. 197–231, 2009.