Behavioral Spillovers in Multiple Games: An Experimental Study*

Jenna Bednar† Yan Chen‡ Tracy Xiao Liu§ Scott Page¶

January 15, 2009

Abstract

We present laboratory experiments that produce behavioral spillovers across strategic contexts. When subjects play two distinct games simultaneously with different opponents, behavior is highly context dependent. When games are paired in ensembles, play differs from the isolated controls. Behavior is also influenced by which other game composes the ensemble. These results suggest that people do not treat each strategic situation in isolation but instead develop heuristics that they apply across games. Our findings imply that the effect of a particular institution on behavior depends upon the full institutional context, leading to disparate performance of identical institutions in different cultures.

Keywords: multiple games, behavioral spillover, experiment
JEL Classification: C72, C91, D03

*This research was sponsored by grants from the National Science Foundation, the James S. McDonnell Foundation, and the United States Air Force Office of Scientific Research MURI. We are indebted to those who have supplied thoughtful reactions to our project, including Vincent Crawford, Rachel Croson, Catherine Eckel, Simon Gächter, Ernan Haruvy, Tom Palfrey, Rahul Sami, Rajiv Sethi, Roberto Weber, Kan Takeuchi and seminar participants at the annual meeting of the Midwest Political Science Association (Chicago, March 2007), ICER (Turin, June 2007), the Third Asia Pacific Regional Meetings of the ESA (Shanghai, July 2007), the Santa Fe Institute (July 2007), USC Gould School of Law (November 2007), Shanghai Jiaotong University (June 2008), UT-Dallas (March 2008), the 2008 DFG-NSF Research Conference (New York, August 2008), the North America Regional Meetings of the ESA (Tucson, AZ, November 2008), the Allied Social Science Association Meetings (San Francisco, CA, January 2009). We would like to thank Andrea Jones-Rooy for excellent assistance.

†Department of Political Science, University of Michigan, Ann Arbor, MI 48109. Email: jbednar@umich.edu.

‡School of Information, University of Michigan, Ann Arbor, MI 48109. Email: yanchen@umich.edu.

§School of Information, University of Michigan, Ann Arbor, MI 48109. Email: liuxiao@umich.edu.

¶Center for the Study of Complex Systems, Departments of Political Science and Economics, University of Michigan, Ann Arbor, MI 48109. Email: scottepage@gmail.com.

1 Introduction

Policies geared toward the economic improvement of developing nations often fail. Interventions that appear efficient on paper play out quite differently on the ground (Easterly 2006). Identical institutions implemented in different environments sometimes produce divergent outcomes. Putnam (1993) chronicles the disparate performance of an identical institutional innovation in northern and southern Italy. In explaining the success or failures of institutions, some scholars have focused on how belief systems or trust relationships support or fail to support incentive structures designed to produce optimal outcomes (e.g. North (2005), Grief (2006), Putnam (1993)). In this paper, we present experimental evidence that supports an alternative explanation: people possess behavioral repertoires that they carry from one strategic situation to the next. In a sense, behavior in one context affects behavior in other contexts (Gigerenzer and Selten (2002), Page (2007), Bednar and Page (2007)). Using a novel experimental design, we find strong evidence of that individuals carry behavioral routines across strategic domains. Our findings suggest that one reason different societies react differently to common strategic situations is that when confronted with a strategic choice people build from or employ existing behavioral rules.[1] We do not deny that diverse population characteristics (such as wealth, education, risk, etc) also contribute to this behavioral variation, but supplement those explanations with one that highlights the institutional context.

Our research design, analyzing multiple games played simultaneously, departs from convention. Almost all game theoretic models and economic experiments consider isolated strategic environments. It is not possible to analyze or empirically test behavioral spillovers when one focuses on a single game. Experiments do exist in which subjects play games sequentially, and, consistent with our theory, those experiments find significant framing (Tversky and Kahneman 1986) and learning transfer effects (Cooper and Kagel 2008).[2] In addition, Huyck, Battalio and Beil (1993) demonstrate that the median-effort coordination game played in isolation leads to inefficient equilibrium, while auctioning off the right to play before the coordination game leads to the

[1]Empirical evidence on diverse population level reactions to institutions spans multiple methodologies, including cross national surveys (Inglehart 1990), laboratory experiments (Henrich et al 2001, 2004), and real world choices (Fernandez and Fogli 2006).

[2]Weber (2006) reports the results of a minimum-effort game experiment where successful coordination is achieved in large groups by starting with small groups and adding entrants who are aware of the group's history. Successful coordination in large groups can be interpreted as learning transfer from small groups that find it easier to coordinate.

payoff-dominant equilibrium. The efficiency-enhancing effect of auctions is analyzed by Crawford and Broseta (1998) using a model of stochastic, history-dependent learning dynamics. Our design differs in that it considers ensembles of games played simultaneously (Samuelson 2001, Bednar and Page 2007). In our experiments, each subject plays two different games with two different subjects. The games are displayed on their screen at the same time. Other than having the two games presented on the screen simultaneously, there is no indication to the subject that the games are linked, and the subject is free to develop distinct responses to each. This design enables us to test whether the the presence of one set of incentives influences outcomes in other contexts by shaping subjects' behavioral repertoires. In effect, we are challenging an implicit assumption that strategic contexts exist in isolation—that people possess the ability to construct cognitive silos around each strategic situation.

Our analysis focuses on two distinct effects of context: behavioral spillovers and cognitive overload. We posit specific hypotheses for each type of effect. The results strongly support our hypotheses: We find that behavior varies significantly from control sessions when multiple games are played simultaneously, and we find that behavior in one game depends upon what other game is included in the two-game ensemble. This trend suggests that variance is not (exclusively) attributable to cognitive overload, but instead indicates the presence of behavioral externalities. Finally, behavior suggestive of cognitive overload is most present where predicted.

Why does this matter? Even if we can produce these behavioral spillovers in the laboratory, why should social scientists take notice? First, as we have already mentioned, this initial evidence of the importance of institutional context for behavior offers an incentive-based explanation for some portion of the variation in institutional performance seen empirically. Second, experiments like ours may help identify which combinations of institutions can coexist successfully. If so, they may help to explain why markets and democracies take off in some societies but do not in others. Finally, experiments such as these might even inform the choice over institutions in designing political and economic transitions.

Institutional contexts have been studied in other experiments. An emergent cross-cultural experimental literature documents the effects of culture on behavior in bargaining, trust, market and public goods games. Rather than a comprehensive survey of this literature, we highlight a few studies whose findings are especially relevant for our study. Roth, Prasnikar, Okuno-Fujiwara and

Zamir (1991) study bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh and Tokyo, and find that while market outcomes converge robustly to equilibrium in all four countries with no significant cross-cultural differences, bargaining outcomes exhibit substantial cross-cultural differences. Furthermore, despite the differences in offers in bargaining games, rejection rates are not substantially different, indicating that subjects in each country were able to best respond to each other. Using a larger sample with more variations in economic development, Henrich, Boyd, Bowles, Camerer, Fehr, Gintis and McElreath (2001) identify market integration and cooperation in production as significant predictors of the difference in average ultimatum offers in 15 small-scale societies. In comparison, using a voluntary contributions public goods game with third party punishments, Herrmann, Thöni and Gächter (2008) identify norms of civic cooperation and the weakness of the rule of law in a country as significant predictors of antisocial punishment in 16 countries around the world. One implication of Herrmann et al. (2008) is that punishment opportunities in public goods provision are socially beneficial "only if complemented by strong social norms of cooperation," i.e., if the context of the mechanism is right.

Another stream of related research investigates whether professionals transfer heuristics developed in the field into the lab. Palacios-Huerta and Volij (2006a) find that 69% of chess players and 100% of chess Grandmasters stop the centipede game immediately, consistent with the hypothesis that they transfer backward induction reasoning in chess to an unfamiliar laboratory game.[3] Similarly, professional soccer players play remarkably consistently with equilibrium predictions in a 2×2 game with a unique mixed strategy equilibrium and the 4×4 zero-sum game developed in O'Neill (1987), again consistent with the hypothesis that they transfer skills developed in penalty kicks, i.e., using mixed strategies, to the laboratory (Palacios-Huerta and Volij 2006b). In comparison, college students play these games far from the equilibrium predictions.

Compared to these two streams of research, our paper simulates the institutional context in the lab and measures the extent to which institutional context influences behavior in a game. In particular, we apply the concept of entropy to measure the strategic uncertainty in each game, and use it to predict the extent to which behavior in a game might influence and be influenced by other games. Therefore, it complements the cross-cultural experiments by developing a new experimental design which enables more precise measurements of the effects of institutional contexts.

[3]Professional chess players' taking probability is about ten times higher than that of college students.

We have organized the paper as follows. In Section 2, we summarize the relevant theoretical literature and describe the specific games included in this study. Section 3 describes our experimental design. Somewhat unusually, we first present the results from the control sessions, where agents play a single game, in Section 4, and then develop our multiple-game hypotheses in Section 5, which are based on theory as well as results from the control sessions. Section 6 reports our findings on the ensemble effects. In Section 7, we discuss what these findings might mean and comment on potential future directions.

2 Theoretical Literature and the Games

In this section, we review the theoretical literature on multiple games, and describe the specific games included in our study.

While most game theory papers analyze specific games in isolation, Samuelson (2001) first formally modeled behavioral spillovers and cognitive load when people play multiple games. He assumes that people are unable to study every strategic interaction in its full complexity. Instead, they maintain a costly stock of analogies to organize their reasoning. In the analysis of three different bargaining games, the ultimatum game, the Rubinstein (1982) alternating offer bargaining game, and a tournament, he characterizes two equilibria, one in which the two bargaining games are played separately, and one in which they are played jointly. In the latter, players apply common analogies to disparate bargaining situations.

While Samuelson analyzes common analogies used in bargaining games, Bednar and Page (2007) examine behavioral spillovers and cognitive load effects in six 2×2 games. They prove conditions for the existence and efficiency of behavioral externalities, using computational agent based models (Miller and Page 2007). Their agent based models show that simple learning rules could locate the proposed equilibria when played in isolation. When agents needed to solve multiple games simultaneously, the agents often created routines that they applied across strategic domains. The agent based model generates behavioral spillovers; agents employed identical strategies in distinct games. The model also shows evidence of cognitive load: some ensembles of games outstrip the capacity of the agents to play each game optimally. In those cases, they find especially strong ensemble effects.

Bednar and Page (2007) provide a theoretical foundation for the current paper: Here we test whether the phenomena derived within models and generated by artificial agents can be produced in a laboratory with real people. We focus on four 2×2 games: the Prisoner's Dilemma (PD), Strong Alternation (SA), Weak Alternation (WA), and a Self Interest game (SI). The individual games belong to a class of two-person two-action games that contain a self-regarding action S and an alternative, C; in three of the games (PD, SA and WA), this alternative is cooperative. In these three games, cooperation lowers a player's own payoff and raises the payoff of the other and being selfish does the opposite, so in the one shot game, the unique dominant strategy equilibrium involves both players choosing selfish. In the fourth game, Self-Interest (SI), selfish behavior is both the stage game dominant strategy and Pareto dominant.

The first game is a standard Prisoner's Dilemma, where the stage game has a dominant strategy equilibrium, (S, S), which is Pareto dominated by (C, C). Note that (C, C) also maximizes the joint payoff of the two players.

Prisoner's Dilemma: (PD)

	C	S
C	7, 7	2,10
S	10,2	4,4

In the second and third games, Strong Alternation (SA) and Weak Alternation (WA), while (S, S) remains the dominant strategy equilibrium for the stage game, agents do best (i.e., maximize joint payoff) in repeated play by alternating between the off-diagonals, (C, S) and (S, C). In Strong Alternation, the incentives to alternate are much stronger than in Weak Alternation. The alternation games are a distant cousin to the conventional Battle of Sexes and Game of Chicken, where agents are rewarded for coordinating their behavior. In our alternation games, four outcomes are rewarded with positive payoffs: CC, SS, and the alternating strategies of CS then SC and SC then CS. Coordinating on CC or SS is much less taxing than working out an alternating behavior, and the positive payoffs for each reduce the focality of an alternating equilibrium.

Strong Alternation: (SA)

	C	S
C	7, 7	4,14
S	14,4	5,5

Weak Alternation: (WA)

	C	S
C	7, 7	4,11
S	11,4	5,5

In the final game, Self Interest (SI), the dominant strategy equilibrium, (S, S), also Pareto dominates all other outcomes. Furthermore, in the stage game, S uniformly dominates C.

Self Interest: (SI)		C	S
	C	7, 7	2,9
	S	9,2	10,10

These four games are variants of the six studied by Bednar and Page (2007). Their findings echo those demonstrated here. As the first experiment with simultaneous play of multiple games, we start with four games whereas they consider six possible games, and we consider only two game ensembles, whereas they also consider three game ensembles.

3 Experimental Design

Our experiments consist of four control sessions, each of which consists of a single game, and 14 treatment sessions, each of which consists of a pair of games. This experimental design enables us to determine the effects of ensemble on behavior by comparing the ensemble with the corresponding control sessions and to compare behavior across ensembles.

The control sessions follow the standard protocol of infinitely repeated games in the laboratory. We have one 12-player session for each of the single games. Participants are randomly matched into pairs at the beginning of each session, and play the same match for the entire experiment. In each session, participants first play the game for 200 rounds. After round 200, whether the game will continue to the next round depends on the "throw of the die" that is determined by the computer's random number generator. At the end of each round after round 200, with 90% chance, the game will continue to the next round. With 10% chance, the game stops. In other words, we implement an infinitely repeated game, with a discount factor of 1 for the first 200 rounds, and 0.9 thereafter. With the chosen discount factors, (C, C) can be sustained as a repeated game equilibrium in PD, SA and WA. With 12 players in each control session, we have 6 independent observations for each single game.

In the ensemble treatment, we again use twelve players in each session. Within each session, at the beginning, each player is randomly matched with two other participants, both of whom will be her matches for the entire experiment. She plays two distinct games with each of these people. This design allows us to analyze whether or not behavior in one game is influenced by the nature of the other game. As in the control sessions, we implement an infinitely repeated game, with a discount factor of 1 for the first 200 rounds, and 0.9 thereafter. Within each session, the 12 players are partitioned into independent groups of 4 each,[4] yielding 3 independent observations. As the two games are displayed side by side, we conduct two independent sessions for each game ensemble, changing the order of the display to avoid the order effect within each round. For example, for the game ensemble of SA and WA, we display SA as the left game in one session, and WA as the left game in another session. This way, if a player always makes decisions from left to right, we have a balanced number of observations for each order.

We used z-Tree (Fischbacher 2007) to program our experiments. As z-Tree does not record the mouse movements within each stage, we ran two additional sessions with ensembles, (SI, WA) and (WA, SI), where we use the software Morae to record the mouse movement. These two sessions will allow us to determine the order of decisions within each round. The (SI, WA) session has 12 subjects, while the (WA, SI) has only eight subjects.[5]

[Table 1 about here.]

Table 1 reports the features of experimental sessions, including the name of the game, the number of players in each session, the number of independent pairs for each control session, the ensemble of games, the number of players in each session, as well as the number of independent groups in each ensemble session.

Overall, 18 independent computerized sessions were conducted in the RCGD lab at the University of Michigan from March to October 2007, yielding a total of 212 subjects. Our subjects were students from the University of Michigan, recruited by email from a subject pool for economic experiments.[6] Participants were allowed to participate in only one session. Each ensemble treatment session lasted approximately 90 minutes, whereas each control session lasted about 45 minutes. The exchange rate was set to 100 tokens for $1. In addition, each participant was paid a $5 show-up fee. Average earnings per participant were $37.49 for those in the treatment sessions and $22.77 for those in the control sessions. Data are available from the authors upon request.

[4]The matching protocol is the following: $\underbrace{4-2-1-3}$, $\underbrace{6-5-7-8}$, $\underbrace{10-9-11-12}$ form three independent groups, each with four participants positioned on a circle, and each participant plays her left and right match.

[5]We recruited for twelve subjects, however, only eight showed up.

[6]Graduate students from the Economics Department are excluded from the list.

4 Results: Control Sessions

In this section, we report the results from the control sessions. We first use entropy (Shannon 1948) to measure the strategic uncertainty in the outcomes of individual games. We then analyze repeated game strategies emerged in each game. The analysis of outcomes as well as strategies provides a benchmark from which we can interpret the ensemble results in Section 6.

To formally measure and rank the strategic uncertainty in each game, we apply the concept of entropy to the analysis of the outcomes.[7] The *entropy* of a random variable X with a probability density function, $p(x) = Pr\{X = x\}$, is defined by

$$H(X) = -\sum_{x} p(x) \log_2 p(x),$$

which is used to measure the average uncertainty in the random variable. When the logarithms to base 2 is used, entropy is expressed in bits. It is the average number of bits required to describe the random variable. Note that the random variable can be vector valued. For the analysis of two-person games, we model individual stage game strategies as a discrete random variable, X, with realizations in one of the four cells. Throughout the analysis, we use the convention that $0 \log 0 = 0$.[8] The entropy in a generic 2×2 game is in the interval, $[0, 2]$, with the lower bound indicating certainty, i.e., all outcomes are in one cell, and the upper bound indicating a uniform distribution among the four cells.

We now present the time series data for each pair in each of the control sessions, with the entropy for each pair presented at the bottom of each graph.

[Figure 1 about here.]

Figure 1 presents outcomes in the Self Interest game. In this game, all six pairs converged to the Pareto dominant equilibrium quickly and stayed there. The entropy for each pair ranges from 0 to 0.04, indicating very little strategic uncertainty. We speculate that this is because of the uniform dominance property of the dominant strategy equilibrium in the stage game. Additionally, participants take an average of 0.62 seconds per round to make a decision in SI, significantly

[7] Shannon (1948) is credited with the development of the concept of entropy and the birth of information theory. Many basic concepts and findings in this field are summarized in Cover and Thomas (2006).

[8] This convention is easily justified by continuity, since $x \log x \to 0$ as $x \to 0$.

less than in any other game ($p \leq 0.01$, one-sided permutation tests). Based upon the uniform dominance property of the unique Pareto efficient stage game equilibrium, its low entropy, and the length of time it took for participants to complete the game, we posit that SI would be the easiest to play efficiently.

[Figure 2 about here.]

Figure 2 presents behavior in the Prisoner's Dilemma game. In this game, over half of the pairs established CC, the efficient outcome, which is consistent with findings from previous experiments (Andreoni and Miller 2002). Curiously, one pair also alternated for a fair number of rounds. The entropy for each pair ranges from 0.08 to 1.90, indicating varied amount of strategic uncertainty. In addition, participants take an average of 1.00 second per round to make a decision in PD, significantly longer than SI, but shorter than SA ($p \leq 0.01$, one-sided permutation tests). As a "context" this game does not establish as strong a behavioral norm as the Self Interest game. Based upon this finding, we anticipate that PD will have a weaker behavioral pull than SI. The difficulty of learning to cooperate in the PD game may limit its spillover effects on play in other games.

[Figure 3 about here.]

Figure 3 presents behavior in the Strong Alternation game, where 5/6 of the pairs successfully established the alternation outcomes. Pair 2 also tempted alternation on and off during the experiment. The entropy for each pair ranges from 1.29 to 1.84, indicating a moderate amount of strategic uncertainty.[9] In addition, participants take an average of 2.72 seconds per round to make a decision in SA, significantly longer than in any other game ($p \leq 0.01$, one-sided permutation tests). We interpret the longer decision time in SA as evidence that coordinated alternation requires more mental activities to establish. Since successful alternation is established in five out of six pairs, this game also provides a strong context.

[Figure 4 about here.]

Lastly, Figure 4 presents the dynamics from the Weak Alternation game. In this game, only two out of six pairs developed an alternating behavior, two pairs cooperated, one (pair 4) converged to

[9] Perfect coordinated alternation results in an entropy of 1.

SS, and the last pair (pair 6) did not seem to have converged to a stable outcome. The entropy for each pair ranges from 0.44 to 1.91, with the highest aggregate entropy among all four games. In addition, participants take an average of 1.24 seconds per round to make a decision in WA, significantly longer than SI, shorter than SA ($p \leq 0.01$, one-sided permutation tests), not significantly different from PD ($p = 0.138$, one-sided permutation test). As WA results in higher strategic uncertainty, we speculate that, while subject behavior in WA is more likely to be influenced by the other game in an ensemble, when paired with other games, it might increase the subjects' cognitive load.

[Table 2 about here.]

To summarize our findings, Table 2 reports the aggregate distribution of outcomes in each of the four games in the control sessions, and the respective entropy for each game in the last line. The strategic uncertainty measured by entropy is the lowest in the Self Interest game (0.02), followed by Prisoner's Dilemma and Strong Alternation (1.68), and the highest in Weak Alternation (1.98). We postulate that games with lower strategic uncertainty might have stronger behavioral spillovers than those with higher strategic uncertainty.

Compared to outcomes, strategies are more complex to analyze. As the games are implemented as infinitely repeated games, there are many repeated game strategies and equilibria. In the analysis, we focus on *simple* strategies. When we represent a repeated game strategy as an automaton, the simplest strategy is defined as one with the least number of states (Kalai and Stanford 1988, Baron and Kalai 1993). On both theoretical and empirical grounds, the simplest equilibria among the set of Pareto efficient ones are more likely to be chosen in infinitely repeated games (Schelling 1960, Baron and Kalai 1993). In our set of games, the simplest Pareto efficient equilibrium is (1) for both players to always play selfish each round (SS) in the Self Interest game; (2) for both players to always cooperate (CC) in Prisoner's Dilemma; and (3) coordinated alternation between S and C (ALT) in the Strong and Weak Alternation games, respectively. In subsequent analysis, we focus on three simple repeated-game strategies, SS, CC, and ALT. Table 3 reports the proportion of simple strategies in each game over the entire series. Boldfaced numbers are the mode of the distribution.

[Table 3 about here.]

Result 1 (Simple Pareto Efficient Strategies in Control Sessions)**.** *In the control sessions, the proportion of simple Pareto efficient strategy play (SS in SI, CC in PD, ALT in SA and WA) is significantly higher than any other simple strategies in SI, PD and SA. However, in WA, there is no significant difference in the proportions of simple strategies played.*

Support. *Table 3 reports the proportion of simple strategy play in each of the four games, and the corresponding p-values for the one-sided permutation tests for each pairwise comparison. For SI, PD and SA, $p < 0.05$ for each pairwise comparison between the simplest Pareto efficient equilibrium play and another strategy, while for WA, none of the pairwise comparisons is significant at the 5% level.*

Result 1 establishes a benchmark which will be used to examine behaviors in ensembles. In sum, three distinct behaviors emerge in the control sessions: in the SI game, selfishness, given its uniform dominance; in the PD game, cooperation; and in SA and WA an alternating form of cooperation, where subjects alternate between the cooperative and selfish actions. Weak Alternation has weaker incentives, so the coordinated alternation is not as prominent as with Strong Alternation.

We next compare the efficiency generated in each game. As each game has different maximum and minimum joint payoffs, we use a normalized efficiency measure defined as follows.

$$\text{(1)} \qquad \text{Efficiency} = \frac{\text{Actual joint payoffs} - \text{Minimum joint payoffs}}{\text{Maximum joint payoffs} - \text{Minimum joint payoffs}}$$

[Table 4 about here.]

Table 4 presents the average efficiency achieved in each game in the control, as well as the p-values for pairwise comparisons using one-sided permutation tests. In the analysis, the average efficiency achieved by each pair in a session is treated as an independent observation. We find that the SI game generates significantly higher efficiency than any other game ($p < 0.05$), while pairwise efficiency comparison is not significant between the PD, SA and WA games ($p > 0.10$). In the SI sessions, the Pareto efficient stage game Nash equilibrium is played 99.86 percent of the times overall, and 100 percent in the last 100 rounds.

5 Hypotheses of Ensemble Effects

In this section, we present a set of hypotheses testing the null of game independence against our two posited ensemble effects, behavioral spillovers and cognitive load. Our alternative hypotheses are based on the theoretical results from Bednar and Page (2007), as well as our empirical results from the control sessions presented in Section 4.

The general null hypothesis in our investigations is of game independence: play in one game is not be affected by the existence of another game to play. If the independence hypothesis is correct then we should see no difference between behaviors in the control sessions (games played in isolation) and when games are presented to subjects as part of ensembles.

Our experimental design tests the existence of two types of ensemble effects: *behavioral spillovers* and *cognitive load*. If games presented within ensembles create *behavioral spillovers*, subjects will respond as if they are developing heuristics that they apply across games. In particular, dominant behavior in one game will influence choice in another. With cognitive load effects, subjects might resort to the dominant strategy in the stage game more often when a game is paired with another game with higher strategic uncertainty, such as the Weak Alternation game.

These effects may be manifested in two ways: (1) control/ensemble, where we compare behavior in a game played in isolation against the behavior in that game when it is paired with other games, and (2) ensemble/ensemble, comparing behavior in a game when it is matched with different games. Our investigations center on two questions: first, does behavior differ between the control—isolated game play—and ensemble, where the game is paired with others? Second, does the ensemble play depend upon *which* other game is in the ensemble? A positive answer to either of these questions would fail to support the null hypothesis of independent play, and support instead the hypothesized contextual dependence of game play.

The following alternative hypotheses look at behavioral spillovers. The general alternative hypothesis is that subject's choice of action in a particular game will be influenced by the other game in the ensemble, particularly biasing choice toward the other game's simple Pareto optimal strategy. Specifically, we expect:

Hypothesis 1 (Effect of SI)**.** *Compared to the corresponding control (or other ensembles), games paired with Self-Interest will exhibit more selfishness.*

Hypothesis 2 (Effect of PD)**.** *Compared to the corresponding control (or other ensembles), games paired with the Prisoner's Dilemma will exhibit more cooperation.*

Hypothesis 3 (Effect of SA)**.** *Compared to the corresponding control (or other ensembles), games paired with Strong Alternation will exhibit more alternation.*

As the Weak Alternation game only generates 36% of alternation in the control sessions, we do not expect that games paired with Weak Alternation will exhibit more alternation. Instead, we expect that it might have an effect in terms of cognitive load.

In Section 4, we develop a partial ordering of the four games based upon the entropy of each game in the control sessions, i.e., the strategic uncertainty follows the order of SI $<$ PD $\sim$ SA $<$ WA. Based on the entropy, decision time and the payoff structure of each game, we posit that Self Interest is the easiest game to play and Weak Alternation the most difficult, and this relative difficulty will be reflected in subjects' choice of action. In particular, we conjecture that cognitive load effects will be most prevalent in games played with more difficult games.

Hypothesis 4 (Effect of WA: Cognitive Load)**.** *Compared to the corresponding control sessions (or other ensembles), subjects will exhibit more selfish behavior (stage game dominant strategy) in a game when it is paired with WA.*

Hypothesis 4 conveys our interest in the experiment's ability to reveal limitations in the cognitive processing of subjects playing multiple games simultaneously. While there are no design features to the experiments that would preclude the subjects from optimizing in each game, we believe that when subjects are asked to solve two games simultaneously they will not be as efficient as they are when playing an isolated game. In particular, when an ensemble contains WA, we predict more *selfish* behavior in PD or SA, a stage-game dominant strategy but inefficient in PD or SA.

Lastly, we are interested in the dynamic effect of behavioral spillovers. It would be interesting to see whether behavioral spillover increases over time or whether subjects grope their way to optimize separately in each game. The latter alternative is related to the independence hypothesis. If spillover effects increase through the series, we may interpret this as evidence of a growing prediliction to apply a successful strategy elsewhere. This effect may be particularly strong if the strategy requires orchestrated coordination, such as with WA or SA. However, this tendency may be overcome as the rounds progress, and the players gain experience: initial cross-application of one game's successful strategy may diminish as the subject is able to turn attention to solving the other game in the ensemble.

6 Results: Ensemble Effects

In this section, we present our results comparing both control and ensemble play (subsection 6.1) and behavioral differences in particular games compared across ensembles (subsection 6.2).

6.1 Comparing Control and Ensemble

Our anticipation was that subjects would play particular games differently between the control sessions, where they played a single game, and when that game appeared as part of an ensemble. This prediction emerges from the two core hypotheses: both behavioral spillovers and cognitive taxes will affect play in ensembles.

In the control sessions, the simple Pareto-efficient strategies (SS in SI, CC in PD, ALT in SA and WA) emerge as the mode among all simple strategies in SI, PD and SA played in isolation. We now examine the likelihood of simple Pareto-efficient strategy play when a game is played in an ensemble.

[Table 5 about here.]

Table 5 reports the proportion of simple Pareto efficient strategy play in the control and ensembles, as well as four probit specifications. The dependent variable is: (1) SS in SI, (2) CC in PD, (3) ALT in SA, and (4) ALT in WA, while the independent variables include a dummy variable, Control, which equals one for an observation in a control session and zero otherwise, and Round to capture any effect of learning. We summarize the results below.

Result 2 (Simple Pareto Efficient Strategies in Ensembles)**.** *While the likelihood of SS play in SI does not change from control to ensembles, the likelihood of simple Pareto efficient strategy play for PD, SA and WA decreases in ensembles compared to the corresponding control. The decrease is substantial and statistically significant for ALT in SA. Furthermore, the likelihood of simple Pareto efficient strategy play increases significantly over time in all ensembles.*

Support. *In Table 5, the coefficients for Control are positive and significant in specifications (1) and (3), indicating a 0.4% decrease in the likelihood of SS in SI, and a 29% decrease of the likelihood of ALT in SA, respectively, in the ensembles compared to the control. The coefficients for Round are positive and significant in specifications (1), (3) and (4), but the magnitude is small.*

Result 2 indicates that, in three out of four games, behavior in a game played in an ensemble is different from the same game played in isolation. In particular, in Strong Alternation, subjects alternate significantly less ($p = 0.029$) when it is in an ensemble than when it is played in isolation. The Self Interest game, however, is not affected by the presence of other games. It is the easiest game to play. Whether in control or ensemble treatments, subjects quickly converged to SS play, with over 99% of selfishness across the rounds.

We next investigate how behavior in a game is influenced by the other game in the ensemble. We first look at the influence of the Self Interest game on other games.

Result 3 (Effects of SI on other games)**.** *In (SI, **PD**) and (SI, **SA**), participants are significantly more likely to play SS in PD and SA, compared to the corresponding PD and SA control groups. The proportion of SS increases from 17.82% (control) to 45.71% (ensemble) in PD, and from 14.81 (control) to 32.45% (ensemble) in SA.*

[Table 6 about here.]

Support. *Table 6 reports six probit specifications. The dependent variable is SS in PD (1 and 2), SA (3 and 4) and WA (5 and 6), while the independent variables include a dummy variable, Control, which equals one for an observation in a control session and zero otherwise, and Round to capture any effect of learning. Specifications (2), (4), and (6) investigate behavior beyond the first 100 rounds. The coefficients for the variable, Control, are negative for the first five specifications, indicating that participants are less likely to play SS in the control compared with the corresponding game in the ensemble ($p < 0.10$ for PD (all rounds) and SA (all rounds), $p < 0.05$ for SA (round $\geq$ 100)). None of the coefficients of Round is significant, indicating no increase of SS play over time.*

By Result 3, we reject the general null hypothesis of game independence in favor of Hypothesis 1 in the (SI, PD) and (SI, SA) ensembles. The proportion of selfish behavior exhibited in both PD and SA more than double when these games are paired with SI. Compared to Result 2 where we find a significant decrease of ALT in SA, Result 3 indicates that the decrease in ALT is accompanied by a sizeable increase in the proportion of SS when Strong Alternation is paired with Self Interest. The effect on the Prisoner's Dilemma is particularly salient given the game's ubiquity in the social sciences: it appears highly vulnerable to selfish contextual influence, and subjects do not recover in later rounds but instead maintain the selfish behavior. The change in behavior in WA is not significant.

Next, we consider the effects of the Prisoner's Dilemma on other games. While cooperation rate in the Weak Alternation game is not affected by the presence of the Prisoner's Dilemma ($p > 0.10$), in the Strong Alternation game, the proportion of CC increases significantly when SA is played in the ensemble with PD compared to when SA is played alone.

Result 4 (Effects of PD on other games)**.** *In Strong Alternation, the proportion of CC increases significantly from 5% when SA is played in isolation to 15% when SA is played in the ensemble with PD.*

Support. *Using similar probit specifications as those reported in Table 6, the coefficient of Control is -0.097 (*$p = 0.029$*).*

By Result 4, we reject the general null hypothesis in favor of Hypothesis 2 in the (PD, SA) ensemble. However, the effect of PD on SI, SA and WA are not significant.

When SA is paired with another game, we find that the proportion of ALT or SS does not change significantly in PD or WA. Therefore, we fail to reject the general null hypothesis in favor of Hypothesis 3.

Finally, we note the effect of cognitive load when a game is paired with WA. We hypothesized that when either PD or SA was played with WA, subjects would play PD and SA suboptimally, by resorting to the dominant strategy in the stage game. Findings confirm these expectations, as demonstrated in Table 7. When PD is paired with WA, subjects played selfish nearly as often as when PD was paired with SI; for SA, the effect of pairing SA and WA was even greater than with SA and SI. Again, we want to draw particular attention to the vulnerability of the Prisoner's Dilemma; in this case cognitive load appears to make cooperative behavior less likely to emerge.

[Table 7 about here.]

Result 5 (Effects of WA on Other Games)**.** *In (**PD**, WA), the proportion of SS is significantly higher than that in PD alone. Similarly, in (**SA**, WA), the proportion of SS is significantly higher while the proportion of ALT is significantly lower that in SA alone. Lastly, in (**SA**, WA), the proportion of ALT increases significantly over time.*

[Table 8 about here.]

Support. *Table 7 presents the proportion of simple strategies in the control sessions of SA and PD as well as in the ensemble, (**SA**, WA), (**PD**, WA). The last three column reports p-value for the one-sided permutation tests. Furthermore, the proportion of ALT is significantly lower in (SA, WA) compared to SA alone. To investigate the effects of learning, we report four probit specifications in Table 8, which largely confirms results from non-parametric tests reported in Table 7. In addition, the coefficient for Round in specification (3) is positive and significant, indicating increased play of ALT over time.*

By Result 5, we reject the general null hypothesis in favor of Hypothesis 4. The effect of WA on other games is increased proportion of SS, rather than ALT. Furthermore, the dynamics evolved in (**SA**, WA) is particularly interesting. The proportion of ALT increases significantly over time, which suggests that subjects initially resorted to SS, a dominant strategy of the stage game, but learned the more difficult alternating form of cooperation as the rounds progressed.

[Tables 9 and 10 about here.]

Lastly, we compare the efficiency in the control and ensemble sessions. Table 9 presents the average efficiency in the control and the ensemble sessions. For each ensemble, we present the efficiency of each game in the ensemble, as well as the overall ensemble efficiency. Table 10 presents the p-values of one-sided permutation tests comparing efficiency between the control and the corresponding game in the ensemble sessions. Consistent with Result 2, the efficiency in SI, PD and SA control sessions is higher than that of the corresponding games in ensemble sessions. In particular, the following comparisons are significant at the 5% level: SI $>$ (**SI**, SA), SI $>$ (**SI**, WA), and SA $>$ (**SA**, WA).

6.2 Comparing Behavior Between Ensembles

A second method for investigating the presence of behavioral spillovers and cognitive overload is to compare behavior between ensembles. The effect of behavioral spillovers on play in the PD is most significantly seen between ensembles. We also highlight key results in WA and SA play.

Result 6 (Behavioral Spillover to PD)**.** *Comparing (**PD**, SA) and (**PD**, SI), subjects alternated more in PD when paired with SA (21% versus 5%), and played selfishly more in SI (46% versus 23%). Comparing (**PD**, SA) and (**PD**, WA), subjects alternated in PD more with SA (21% versus 9%) and played the PD selfishly significantly more with WA (39% versus 23%).*

Support. *One-sided permutation test comparing the proportion of ALT in PD in ensembles (**PD**, SA) and (**PD**, SI) yields* $p = 0.10$*, while one-sided permutation test comparing the proportion of SS in PD between the two ensembles yield* $p = 0.10$*. Similarly, one-sided permutation test comparing the proportion of ALT in PD in ensembles (**PD**, SA) and (**PD**, WA) yields* $p = 0.18$*, while one-sided permutation test comparing the proportion of SS in PD between the two ensembles yields* $p = 0.04$*.*

Result 6 indicates that the Prisoner's Dilemma is susceptible to the institutional context it is situated in. While participants alternate more in PD when they also play Strong Alternation, they play PD selfishly more when they also play Self Interest or Weak Alternation. We next examine how behavior in the Strong Alternation game is influenced by other games in the ensemble.

Result 7 (Behavioral Spillover to SA)**.** *Comparing (**SA**, PD) and (**SA**, SI), subjects played CC more often in SA when paired with PD (14% versus 7%). The difference is more pronounced and significant in the last 100 rounds (13% versus 1%).*

Support. *One-sided permutation tests comparing the proportion of CC in in SA in ensembles (**SA**, PD) and (**SA**, SI) yield* $p = 0.11$ *for the entire series, and* $p = 0.021$ *for the last 100 rounds.*

In later rounds, the effect grows even more pronounced: while they continue to play CC in SA when it is paired with PD, where SA is paired with SI subjects shift from CC to ALT, so that the CC percentages are 13% in (**SA**, PD) versus 1% in (**SA**, SI).

We next examine how behavior in the Weak Alternation game is influenced by other games in the ensemble. Recall from Result 1 that, unlike Self Interest, no dominant behavior emerges in the control sessions of Weak Alternation, making it susceptible to influence from other co-existing institutions.

Result 8 (Behavioral Spillover to WA)**.** *Comparing (**WA**, PD) and (**WA**, SA), subjects alternated more in WA when also playing SA (37% versus 18%) and cooperated significantly more when WA was paired with PD (31% versus 11%).*

Support. *One-sided permutation tests comparing the proportion of ALT in (**WA**, PD) and (**WA**, SA) yield* $p = 0.12$*. Similarly, one-sided permutation tests comparing the proportion of CC in WA between the two ensembles yield* $p = 0.022$*.*

Result 8 indicates that behavior in Weak Alternation is indeed susceptible to the influence of the other games in the ensemble. In particular, participants alternated more in WA when also playing SA and cooperated more when WA was paired with PD. While the former is not significant at the 10% level, the latter is significant at the 5% level.

[Table 11 about here.]

Lastly, we examine the efficiency comparisons between ensembles. While the last column of Table 9 presents the average efficiency in ensembles, Table 11 reports the p-values of one-sided permutation tests of pairwise comparison between ensembles which have at least one game in common. We find that a game played together with SI generates significantly higher efficiency than the corresponding game paired with any other game. Specifically, the following comparisons are significant at the 5% level, except for (SI, PD) $\geq$ (SA, PD), which is significant at the 10% level:

- (SI, PD) $>$ (WA, PD) and (SI, PD) $\geq$ (SA, PD);
- (SI, SA) $>$ (PD, SA) and (SI, SA) $>$ (WA, SA);
- (SI, WA) $>$ (PD, WA) and (SI, WA) $>$ (SA, WA).

The efficiency comparison across ensembles follows from several behavioral regularities. First, unlike any other game, the Self Interest game generates consistently high efficiency regardless of whether it is play alone or with other games. Second, SI takes significantly less time to play than any other game, indicating lower cognitive load, which implies that participants can devote more time to optimize in the other game which causes overall high efficiency in the ensemble.

In general, the experimental results agree with our alternative hypotheses: game independence is not supported, but instead subjects are influenced by contextual effects of behavioral spillovers and cognitive load. In addition, we find that behavioral spillovers decrease over time, while subjects learn (at a slow but significant rate) to separately optimize in the Strong Alternation game.

7 Discussion

In this paper, we present an experimental study to test for ensemble effects in game playing behavior. Our study reveals strong evidence of behavioral spillovers that depend in predictable ways on features of the games in the ensemble. In particular, if subjects play one game in an ensemble that encourages selfishness or cooperation, then they are more likely to exhibit that behavior in the other game in their ensemble. We also see evidence of cognitive overload when ensembles include the Weak Alternation game.

Our findings, demonstrating how a person's behavior in a given game depends on the ensemble of strategic situations that the person faces, calls into question both the theoretical and empirical analysis of isolated games as well as the standard mechanism design assumption that incentives can be considered independent of the broader behavioral context. They also provides a possible path toward greater understanding of cross organizational and cross national differences in the performance of specific institutions or game forms. Our results encourage a context sensitivity in designing new policies or developmental programs, and we show several forms that contextual influence is likely to take.

In future work, we hope to consider an experimental design that adds games to the ensemble sequentially. If we can learn how people play one in one game depends upon they other games they face, then if we know the current ensemble of strategic choices that a population confronts, we can gain insights into how they might behave when confronted with a new institutionalized game form, such as a market or democratic mechanism. Moreover, evidence of sequential behavioral spillovers would imply the potential for a theory of institutional path dependence based on behavioral spillovers (Page 2006).

As with any laboratory experiment, our results may not translate to the larger world. People often rely on contextual clues to behave differently in distinct situations. Thus, people can act altruistically toward their children but competitively at work. We do not deny the human capacity to bracket contexts and act accordingly. However, we believe that such contextual bracketing requires cognitive effort and that, in general, people will seek out consistent behaviors that apply across multiple settings. Our experiments support that hypothesis.

A second potential criticism pertains to the simplicity of the games we consider. Would these effects continue to hold for more complex games embedded in a richer institutional and cultural context? We cannot answer that question in a laboratory. But the fact that the game ensembles can influence behaviors in individual games in a relatively sterile laboratory would seem to suggest that such effects might also exist in the real world.

To summarize, these experiments demonstrate that significant ensemble effects can be produced in the lab. And, more importantly, these ensemble effects can be predicted based upon the the attributes of the games. Subjects with incentives to behave cooperatively (resp. selfishly) in one game, tend to behave similarly in another game even if that behavior is neither efficient nor an equilibrium. This finding suggests that when we consider the performance of an institution, we should look not only at the outcome that it produces but also at the behaviors that produce that outcome as those behaviors may well influence play in other settings. The creation of a cooperative culture may well depend on creating multiple institutions that create strong incentives for cooperation so that cooperative behavior can spill over into other contexts as well.

Appendix. Experimental Instructions

(Not For Publication)

We present the instructions for the (PD, WA) ensemble. Instructions for other ensemble treatments are identical except for the specific game forms. Instructions for the control sessions are standard. They are identical to the ensemble instructions except that two games and two other participants are replaced with one game and one other participant everywhere. Hence we omit them here.

Name: PCLAB: Total Payoff:

Introduction

- You are about to participate in a decision process in which you will play two games with two other participants. Each game will be played with a different participant and will be played for many rounds. This is part of a study intended to provide insight into certain features of decision processes. If you follow the instructions carefully and make good decisions, you may earn a considerable amount of money. You will be paid in cash at the end of the experiment.
- During the experiment, we ask that you please do not talk to each other. If you have a question, please raise your hand and an experimenter will assist you.

Procedure

- Matching: At the beginning of the experiment, you will be matched randomly with two other participants, both of whom will be your matches for the entire experiment. You will be matched with these same two people in all rounds. You will play a different game with each of these people.
- Roles: Throughout the game, you will be designated as the "row" player and your matches will be the "column" players. You will be a "row" player in all rounds, and your matches will be "column" players in all rounds

- Actions: In each round, you and your two matches will simultaneously and independently make decisions in two different games. One is the left game and the other is the right game. You will play the left game with one of your matches (Left Game Match) and play the right game with the other match (Right Game Match). In each game, the row player (you) will click either the Top (A) or the Bottom (B) button. The column player (your Left or Right Game Match) will choose either the Left (A) or Right (B) button. These choices determine which part of the matrix is relevant (Top Left, Top Right, Bottom Left, Bottom Right).

- Interdependence: A player's earnings depend on the decision made by the player and on the decision made by his or her two matches as shown in the matrix below. In each cell, the row player's payoff is shown in red and the column player's payoff is shown in blue.

		Column Player			Column Player	
		Left (A)	Right (B)		Left (A)	Right (B)
Row	Top (A)	7, 7	2,10	Top (A)	7, 7	4,11
Player	Bottom (B)	10,2	4,4	Bottom (B)	11,4	5, 5

  For example, if the row player (you) chooses Top (A) and the column player (your left game match) chooses Right (B) in the left game, then the row player (you) will get 2 points, while the column player (your left game match) will get 10 points in this game. Meanwhile, if the row player (you) chooses Bottom (B) and the column player (your right game match) chooses right (B) in the right game, then the row player (you) will get 5 points, and the column player (your right game match) will also get 5 points in this game. So as the row player in both games, you will get 7 points in this round totally.

- Rounds: You will first play the two games for 200 rounds. After round 200, whether the games will continue to the next round depends on the "throw of a die" that is determined by the computer's random number generator. At the end of each round after round 200, with 90% chance, the games will continue to the next round. With 10% chance, the games stop.

- Earnings: Your earnings are determined by the choices that you and your two matches make in every round. Your total earning is the sum of your earnings in all rounds.

  The exchange rate is $1 for 100 points.

You can round up your total earning to the next dollar. For example, if you earn $15.23, you can round it up to $16.

- History: In each round, your and your two matches' decisions in all previous rounds will be displayed in a history window.

We encourage you to earn as much money as you can. Do you have any questions?

Table 1: Features of Experimental Sessions

Control		Ensemble Treatment		
Game: n	Pairs	(Left, Right):	n	Groups
		(PD, WA):	12	3
		(WA, PD):	12	3
		(PD, SI):	12	3
PD: 12	6	(SI, PD):	12	3
		(SA, WA):	12	3
		(WA, SA):	12	3
		(SA, PD):	12	3
SA: 12	6	(PD, SA):	12	3
		(SI, WA):	12 + 12	6
		(WA, SI):	12 + 8	5
		(SI, PD):	12	3
SI: 12	6	(PD, SI):	12	3
WA: 12	6			
Total: 48	24		164	41

Table 2: Distribution of Outcomes and Entropy in Control Sessions

	SI		PD		SA		WA	
	C	S	C	S	C	S	C	S
C	0.00	0.14	55.68	11.67	5.02	39.81	33.18	21.57
S	0.00	99.86	14.82	17.82	40.37	14.81	22.74	22.51
Entropy	0.02		1.68		1.68		1.98	

Table 3: Average Proportion of Simple Strategies in Control Sessions

	% Simple Strategies			P-value of Permutation Tests		
Games	SS	CC	ALT	CC v. SS	CC v. ALT	SS v. ALT
SI	**99.86**	0.00	0.00	0.000	0.500	0.000
PD	17.82	**55.68**	15.44	0.039	0.031	0.389
SA	14.81	5.02	**71.12**	0.040	0.000	0.001
WA	22.51	33.18	**36.14**	0.308	0.430	0.317

Table 4: Average Efficiency in Control Sessions

Games	Average Efficiency	Permutation Tests H_1	P-value
		SI > PD	0.002
		SI > SA	0.002
SI	99.86	SI > WA	0.005
PD	73.35	SA > PD	0.242
SA	82.69	SA > WA	0.248
WA	70.86	PD > WA	0.452

Table 5: Simple Pareto Efficient Strategies in Ensemble vs. Control

	SS in SI	CC in PD	ALT in SA	ALT in WA
% in Control	0.999	0.557	0.711	0.361
% in Ensemble	0.991	0.410	0.449	0.244
	(1)	(2)	(3)	(4)
Control	0.004	0.151	0.290	0.117
	(0.002)**	(0.169)	(0.121)**	(0.194)
Round	0.000	0.001	0.002	0.001
	(0.000)***	(0.000)*	(0.000)***	(0.000)***
Observations	21984	18644	18252	21680

Notes:

1. Coefficients are probability derivatives.
2. Robust standard errors in parentheses are clustered at the group (pair) level in the treatment (control).
3. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Effects of SI on Other Games in the Ensemble

	SS in PD		SS in SA		SS in WA	
	(1)	(2)	(3)	(4)	(5)	(6)
	All Rounds	Round≥100	All Rounds	Round≥100	All Rounds	Round≥100
Control	-0.279	-0.285	-0.181	-0.259	-0.050	0.020
	(0.162)*	(0.182)	(0.100)*	(0.111)**	(0.154)	(0.176)
Round	-0.000	-0.000	-0.001	-0.001	-0.001	-0.000
	(0.000)	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)
Obs.	7824	4260	7428	3864	11580	6036

Notes:
1. Coefficients are probability derivatives.
2. Robust standard errors in parentheses are clustered at the group (pair) level in the treatment (control).
3. * significant at 10%; ** significant at 5%.

Table 7: Effects of WA: (PD, WA) and (SA, WA)

	% Simple Strategies				% Simple Strategies			Control vs. Ensemble		
Control	SS	CC	ALT	Ensemble	SS	CC	ALT	CC	SS	ALT
PD	17.82	**55.68**	15.44	(PD,WA)	39.20	**40.32**	8.69	0.182	0.027	0.401
SA	14.81	5.02	**71.12**	(SA,WA)	**39.39**	9.90	37.76	0.106	0.045	0.038

Table 8: Effects of WA and Dynamics

	(1)	(2)	(3)	(4)
	ALT in PD	SS in PD	ALT in SA	SS in SA
Control	0.070	-0.216	0.356	-0.250
	(0.108)	(0.092)*	(0.168)**	(0.119)**
Round	-0.000	0.000	0.002	-0.001
	(0.000)	(0.000)	(0.001)***	(0.001)
Observations	7652	7652	7524	7524

Notes:
1. Coefficients are probability derivatives.
2. Robust standard errors in parentheses are clustered at the group (pair) level in the treatment (control).
3. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 9: Efficiency in Ensemble

Control	Efficiency in Control	Ensemble (Game1, Game2)	Average Efficiency in Ensemble		
			Game1	Game2	Ensemble
		(SI,PD)	99.32	51.50	80.19
		(SI,SA)	99.23	64.24	82.76
SI	99.86	(SI,WA)	98.99	63.06	86.16
PD	73.35	(PD,SA)	63.59	68.10	66.17
SA	82.69	(PD,WA)	53.97	53.97	53.97
WA	70.86	(SA,WA)	55.66	57.89	56.52

Table 10: P-values of Pairwise Efficiency Comparisons between Control and the Corresponding Game in the Ensemble

Control	Ensemble					
	(SI,PD)	(SI,SA)	(SI,WA)	(PD,SA)	(PD,WA)	(SA,WA)
SI	0.112	0.000	0.004			
PD	0.115			0.230	0.059	
SA		0.053		0.087		0.033
WA			0.296		0.141	0.219

Table 11: P-values of Pairwise Efficiency Comparisons between Ensembles

Ensemble	(SI,SA)	(SI,WA)	(PD,SA)	(PD,WA)	(SA,WA)
(SI,PD)	0.366	0.114	0.079	0.004	
(SI,SA)		0.770	0.039		0.022
(SI,WA)				0.000	0.003
(PD,SA)				0.079	0.227
(PD,WA)					0.416

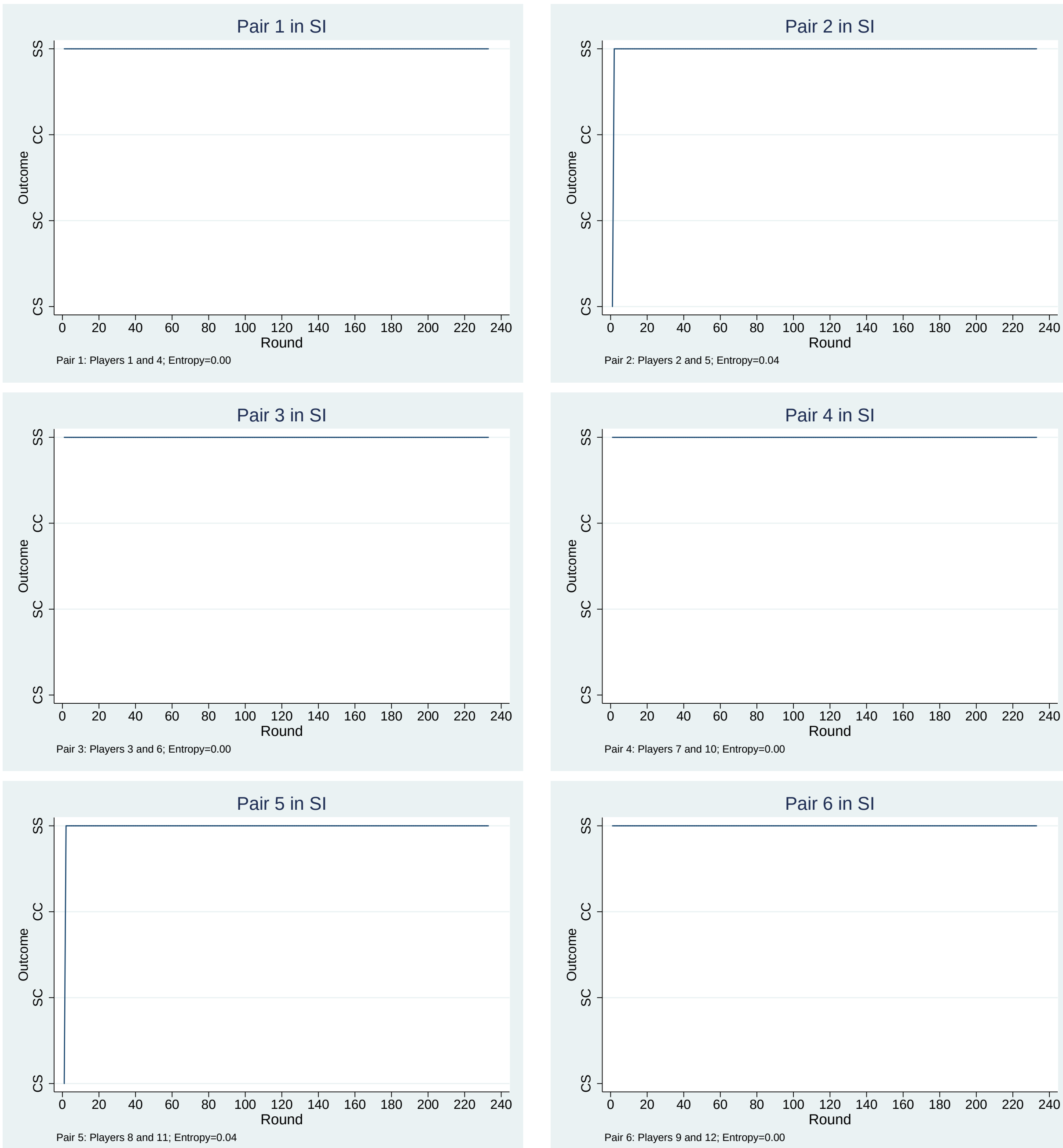

Figure 1: Control: Behavior in Self Interest

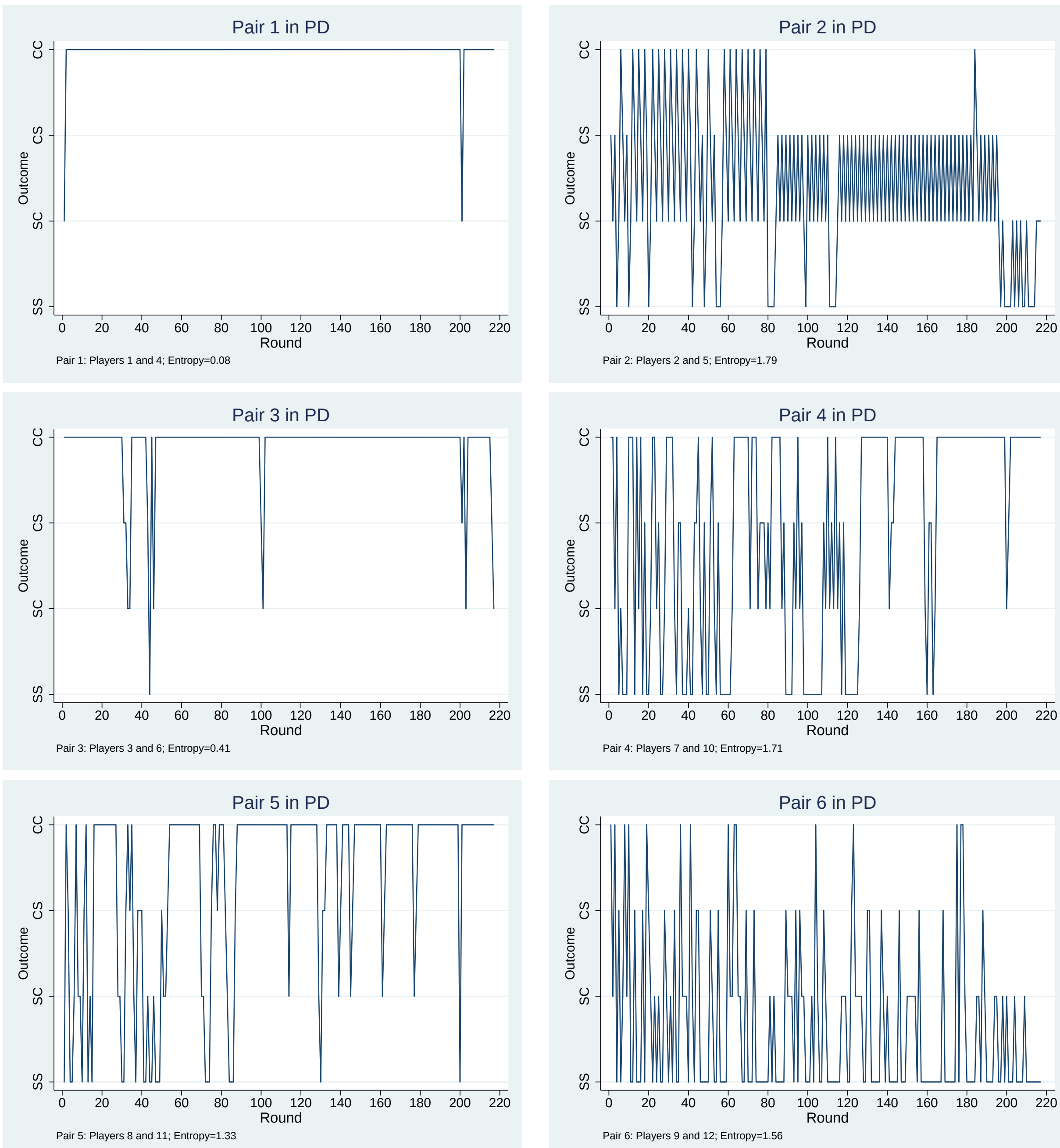

Figure 2: Control: Behavior in Prisoner's Dilemma

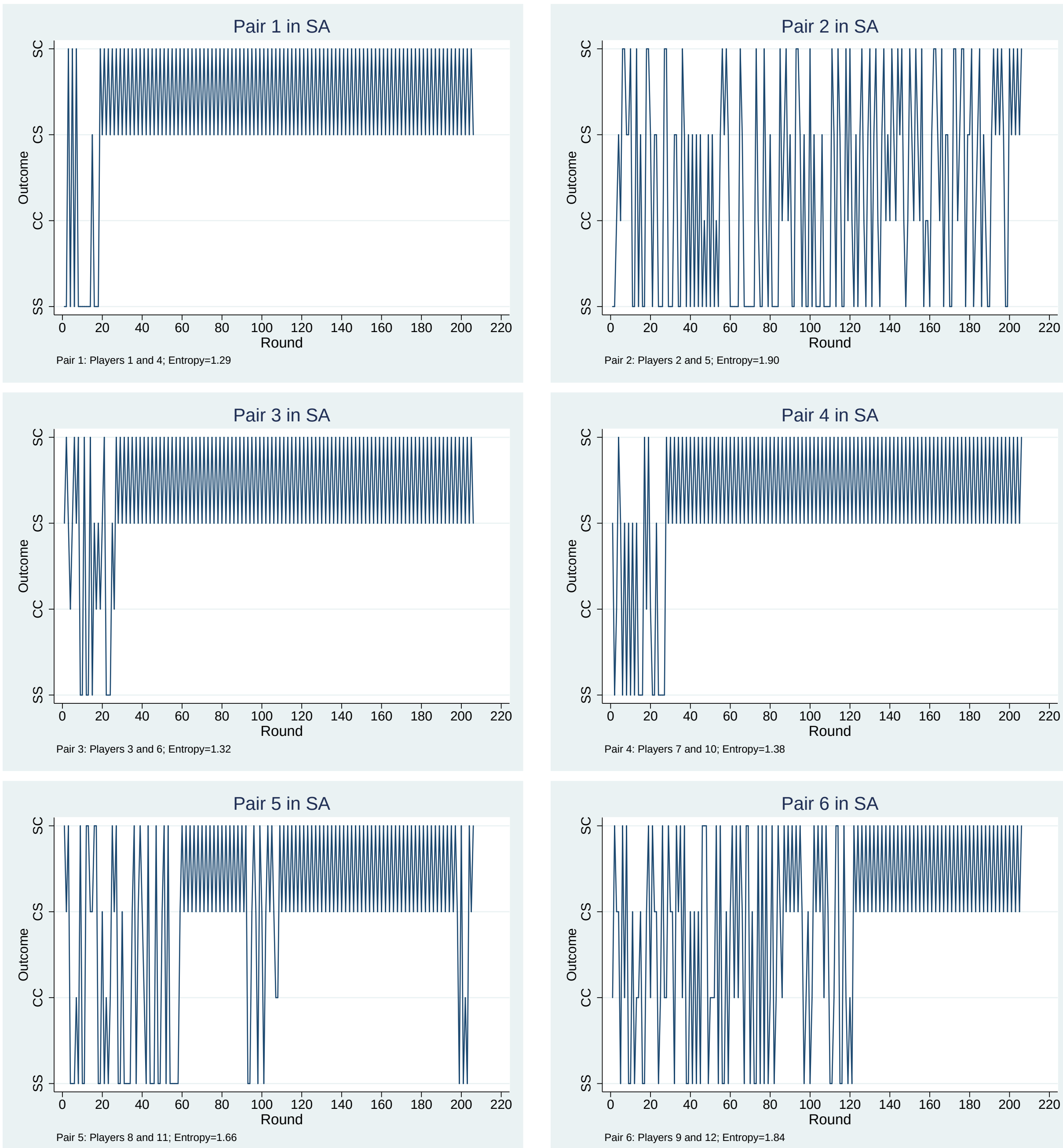

Figure 3: Control: Behavior in Strong Alternation

Figure 4: Control: Behavior in Weak Alternation

References

Andreoni, James and John Miller, "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism," *Econometrica*, 2002, *70* (2), 737–753.

Baron, David and Ehud Kalai, "The Simplest Equilibrium of a Majority-Rule Division Game," *Journal of Economic Theory*, 1993, *61*, 290–301.

Bednar, Jenna and Scott E Page, "Can Game(s) Theory Explain Culture?: The Emergence of Cultural Behavior in Multiple Games," *Rationality and Society*, 2007, *19* (1), 65–97.

Cooper, David J. and John H. Kagel, "Learning and transfer in signaling games," *Economic Theory*, March 2008, *34* (3), 415 – 439.

Cover, Thomas M. and Joy A. Thomas, *Elements of Information Theory*, Wiley-Interscience, 2006.

Crawford, Vincent and Bruno Broseta, "What Price Coordination? The Efficiency-Enhancing Effect of Auctioning the Right to Play," *The American Economic Review*, 1998, *88* (1), 198–225.

Easterly, William, *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done SO Much Ill and SO Little Good*, New York: Penguin Press, 2006.

Fernandez, Raquel and Alessandra Fogli, "Fertility:The Role of Culture and Family Experience," *Journal of the European Economic Association*, 2006, *4* (2-3), 552–561.

Fischbacher, Urs, "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 2007, *10* (2), 171–178.

Gigerenzer, Gerd and Reinhard Selten, *Bounded Rationality: The Adaptive Toolbox*, Cambridge: MIT Press, 2002.

Grief, Avner, *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*, New York: Cambridge University Press, 2006.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin F. Camerer, Ernst Fehr, and Herbert Gintis, *Foudations of Human Sociality:Economic Experiments and Ethnographic Evidence from Fifteen Small-scale Societies*, New York: Oxford University Press, 2004.

__ , __ , __ , __ , __ , __ , and Richard McElreath, "In search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies," *American Economic Review*, 2001, *91*

(2), 73–78.

Herrmann, Benedikt, Christian Thöni, and Simon Gächter, "Antisocial Punishment Across Societies," *Science*, 2008, *319* (5868), 1362–1367.

Huyck, John B. Van, Raymond C. Battalio, and Richard O. Beil, "Asset Markets as an Equilibrium Selection Mechanism: Coordination Failure, Game Form Auctions, and Tacit Communication," *Games and Economic Behavior*, 1993, *5* (3), 485 – 504.

Inglehart, Ronald, *Culture Shift in Advanced Industrial Society*, Princeton: Princeton University Press, 1990.

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

Miller, John H. and Scott E. Page, *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*, Princeton, NJ.: Princeton University Press, 2007.

North, Douglass C., *Understanding the Process of Economic Change*, Princeton, NJ: Princeton University Press, 2005.

O'Neill, Barry, "Nonmetric Test of the Minimax Theory of Two-Person Zero-Sum Games," *Proceedings of the National Academy of Sciences*, April 1987, *84* (7), 2106–2109.

Page, Scott E., "Path Dependence," *Quarterly Journal of Political Science*, 2006, *1* (1).

Page, Scott E, *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*, Princeton, NJ.: Princeton University Press, 2007.

Palacios-Huerta, Ignacio and Oscar Volij, "Experientia Docet: Professionals Play Minimax in Laboratory Experiments," 2006. Manuscript.

__ and __ , "Field Centipedes," 2006. Manuscript.

Putnam, Robert D., *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton, NJ.: Princeton University Press, 1993.

Roth, Alvin E., Vesna Prasnikar, Masahiro Okuno-Fujiwara, and Shmuel Zamir, "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study," *The American Economic Review*, 1991, *81* (5), 1068–1095.

Rubinstein, Ariel, "Perfect equilibrium in a bargaining model," *Econometrica*, 1982, *50*, 97–109.

Samuelson, Larry, "Analogies, Adaptations, and Anomalies," *Journal of Economic Theory*, 2001, *97* (2), 320–366.

Schelling, Thomas, *The Strategy of Conflict*, Cambridge,MA: Harvard University Press, 1960.

Shannon, Claude E., "A Mathematical Theory of Communication," *Bell System Technical Journal*, 1948, *27*, 379–423.

Tversky, Amos and Daniel Kahneman, "Rational Choice and the Framing of Decisions," *The Journal of Business*, 1986, *59* (s4), S251.

Weber, Roberto A., "Managing growth to achieve efficient coordination in large groups," *American Economic Review*, 2006, *96* (1), 114 – 126.