Coordination with Uncertainty

Dean Jens

Abstract

An experiment is conducted in which subjects play simple stochastic games in which they face choices between stage payoffs and random continuation values. In the two-player game, both players receive the same payoff, and thus find it in their interest to coordinate to optimize their payoff, but may face conflict ex-ante if their risk preference and time preferences are different. It is observed that, while even after many periods in the one-player game agents only weakly settle on pure Markov strategies, in the two-player game agents are more reluctant to switch strategies after achieving coordination, so that two-player games do generally converge on equilibria in stationary Markov strategies. Subjects have more trouble coordinating in a state of potential loss than in a state of potential gain, especially where they have different time/risk preferences.

1 Background

1.1 Coordination and strategic uncertainty

While theoretical game theory has made great strides in analyzing normative solution concepts in games,[1] these solution concepts encounter occasional problems in describing the strategic interaction of people. One class of problems stems from the failure of the assumption of common knowledge of rationality, but another class falls under the rubric of "equilibrium selection"; if a game has two theoretically sound equilibria, the agents need a way to coordinate on one equilibrium.

If two agents are driving cars toward each other, they need to decide whether to drive on the right hand side of the street or the left hand side (figure 1). There are three Nash equilibria, all of which are essential equilibria — any small perturbation of the game has a similar equilibrium structure. Any pure action by either player is part of an equlbrium, and thus strategic equilibrium theory provides no guidance to either player as to which action to take. Indeed, the two pure-strategy equilibria survive attempts at equilibrium selection (e.g. Harsanyi and Selten (1988)); while either will be self-fulfilling if both agents expect it to obtain, coordination on an equilibrium requires either luck or some exogenous source of beliefs.

[1] This is reviewed in Hillas and Kohlberg (2002); Van Damme (2002).

Figure 2 represents a game with a Pareto-dominant equilibrium in which players encounter a different problem with coordination. The narrative is that two hunters are individually capable of hunting hares, but may bring down a stag if they coordinate on that choice; they would each prefer their share of a stag to a hare, but the consequences of a coordination failure are more severe for an agent who chooses the bigger game. Harsanyi and Selten (1988) and Harsanyi (1995) come to different conclusions as to which equilibrium is (in some sense) the "right" one.

In a final illustrative example, figure 3 shows the "battle of the sexes", in which each player has the same action set and has a strong preference to coordinate on the same action as the other player, but also has a (weaker) preference for playing one action or the other, with the two players differing in this last preference. The two-player game in this experiment (figure 5 on page 7) seems superficially to resemble figure 1, but if the players have different risk preferences the strategic situation is closer to that of figure 3. Of course, agents with different risk preference may be expected to respond differently to the strategic risk inherent in not knowing what one's partner will choose. There is thus a coupling between an agent's preference between the two (pure) equilibria and its response to the possibility that its partner would choose differently.

In a game with one player, this strategic uncertainty isn't an issue. It is still the case that subjects in the lab, as well as economic agents in the field, exhibit violations of rationality that cannot be explained by unusual transitive preferences. There have been attempts to expand the utility-maximization framework (e.g. Gul and Pesendorfer (2001)), and attempts to more deeply model possible explanations for violations of utility maximization, such as bounded rationality. One popular model for boundedly rational agents is that of a finite-state automaton; the finite number of states limits the ability of the agent to reliably choose the "best" option available according to some underlying preferences. (See, for example, Salant (2011).)

An experimental setting allows us to observe human subjects making decisions in a comparatively informationally sparse environment. In this study, we present the subjects with a simple stochastic game. The explicit information entailed in this environment is well attuned to the finite automaton construct, though there is a great deal of implicit information and room for heuristic decision making remaining.

1.2 Stochastic Games

In 1953, Lloyd Shapley introduced into the literature the concept of a "stochastic game", a game with multiple stages and a "state space" such that the actions available to players in a stage may depend on the state in that stage (Shapley

		player 2	
		left	right
player 1	left	1,1	$-1,-1$
	right	$-1,-1$	1,1

Figure 1: A simple game with multiple equilibria.

		player 2	
		stag	hare
player 1	stag	2,2	$-2,1$
	hare	$1,-2$	1,1

Figure 2: Jean-Jacques Rousseau's "stag hunt"

(1953)). Payoffs to players will be (possibly discounted) sums of stage game payoffs, which in general depend on the state in that stage as well as the action profile played in that state. The state in a period is random, with a distribution determined by the state in the previous stage, as well as the action profile in that stage. Such games can model economic situations with repeated interaction in which actions affect state variables; for example, fishermen drawing fish from a common fishery might affect the quality of the fishery going forward by their fishing practices, subject to some variability that is outside their ability to precisely forecast. Shapley proved the existence of an equilibrium in a context that is somewhat narrower than usual modern interest, but existence results have been generalized substantially in the last six decades (Mertens (2002); Vieille (2002)).

A pure strategy in such a setting could be quite complicated, but of particular interest are stationary Markov strategies, in which a player's action is a function only of the state. For the duration of this paper, we shall only be concerned with stationary Markov strategies.

		player 2	
		ballet	ball game
player 1	ballet	3,2	1,1
	ball game	0,0	2,3

Figure 3: "Battle of the sexes." There are two Pareto-optimal essential equilibria, but the preferences over the equilibria are in conflict, making coordination more difficult.

1.3 Individual Decision Making

Traditional game theory investigates the strategic consequences of rational utility maximization; it has been well established, however, that human agents do not always behave as rational expected utility maximizers. This is especially true in situations that are novel and those that involve probabilities.

For binary lotteries over gains, human agents generally exhibit risk aversion to a degree that seems severe if risk aversion is measured relative to the agent's total wealth. Indeed, there is some evidence of risk-seeking behavior in the face of "losses", with agents preferring risks of larger losses to certainties of the same expected loss. It is also the case that, given series of lotteries, human agents routinely fail to aggregate them according to laws of probability; for example, a lottery presenting a subject with a 1/3 probability of winning the chance to then have a 50/50 chance of winning a $5 gain tends to be treated differently from a single, atomic 1/6 probability of a $5 gain. Even the simplest of stochastic games will necessarily incorporate many features that may distinguish the actual behavior of subjects from expected valuation maximization.

1.4 Coordination

There are many experimental results in which subjects are presented with games in which there are many Nash equilibria that are either Pareto ranked (e.g. "minimum effort" games; cf. Van Huyck, Battalio, and Beil (1990)) or such that agents have different preferences over them (e.g. "battle of the sexes"). Outcomes in either case can depend on a variety of details. In genuinely one-shot games with multiple equilibria, there's frequently no particular reason to expect that any equilibrium will be attained if there's no common basis for beliefs as to which equilibrium should obtain.

If an agent has clear, concrete expectations for the actions of other agents, the agent may view the choice of action as a decision problem; if there is one other agent with possible actions C and D, and the agent assigns a 70% chance to action C, then choosing action A is tantamount to choosing a 70% chance of the results of action profile (A,C) and a 30% chance of the outcome of (A,D). In practice, a crisp probability of this sort is likely not to be at hand; a result generally attributed to Ellsberg (1961) shows that people have preferences for "certain uncertainties" over less quantifiable uncertainties, as are exhibited in situations of strategic uncertainty.

1.5 Other Related Literature

Cooper, DeJong, Forsythe, and Ross (1990) investigate equilibrium selection in play of small games with rematching after each round; in each case play does state 1

A	B
0	a

state 2

A	B
b	c

A	B
p_1	p_3

A	B
p_4	p_2

Figure 4: Transition probabilities given are for transition to state 2.

converge to a Nash equilibrium, but which equilibrium obtains can be affected by payoffs that should, from a theoretical standpoint, be irrelevant to strategic stability. Pareto dominated equilibria often result.

Mookherjee and Sopher (1994, 1997) examine learning behavior in repeated constant-sum games; these are similar to a zero-difference game in that the folk theorem has very little "bite". They find that experienced payoffs drive behavior to a greater extent than deep strategizing. Camerer and Ho (1999) examine a more general learning model and find support for some influence of higher-order strategic behavior, but with an important contribution from experienced feedback.

This experiment introduces both an element of (non-strategic) randomness and a time-varying element (i.e. the game can change states). Coordination in this case

The section shouldn't end in the middle of a sentence.

2 A Stochastic Game Experiment

2.1 One player game

Consider the following one-player stochastic "game" (decision problem): there are two states, state 1 and state 2, and two actions in each state, which will be denoted A and B in both states, as in figure 4. State 2 has generally higher payoffs than state 1 ($b > 0$, $c > a$), and action A in either state results in a higher probability that the next stage will be in state 2 rather than state 1; however, action B results in a higher payoff ($a > 0$, $c > b$) in the current stage. Of particular interest is the case with $p_3 = 0$ and $p_4 = 1$; in each state, the agent is asked to choose between a payoff that will leave the state unchanged versus a different payoff with a lottery over continuation values.

Valuations for different states and strategies are (concisely) derived at the top of page 6 for stationary Markov pure strategies with discount factor δ. In particular, for $p_3 = 0$, $p_4 = 1$, and $\delta \to 1$, AB gives an expected average payoff

The pure stationary Markov strategies for the decision problem have expected values that can be expressed recursively

strategy	state 1	state 2
AA	$0 + \delta p_1 c_2 + \delta(1-p_1)c_1$	$b + \delta c_2 + \delta(1-p_4)(c_1 - c_2)$
AB	$0 + \delta p_1 c_2 + \delta(1-p_1)c_1$	$c + \delta p_2 c_2 + \delta(1-p_2)c_1$
BA	$a + \delta c_1 + p_3\delta(c_2 - c_1)$	$b + \delta c_2 + \delta(1-p_4)(c_1 - c_2)$
BB	$a + \delta c_1 + p_3\delta(c_2 - c_1)$	$c + \delta p_2 c_2 + \delta(1-p_2)c_1$

For interior values of δ and p_i, we can solve and multiply by $1-\delta$:

strategy	state 1	state 2
AA	$\delta p_1(1+\delta p_1 - \delta p_4)^{-1}b$	$(1-\delta(1-p_4)(1+\delta p_1 - \delta p_4)^{-1})b$
AB	$c\delta p_1(1+\delta(p_1 - p_2))^{-1}$	$c(1-\delta(1-p_1))(1+\delta(p_1-p_2))^{-1}$
BA	$a + p_3\delta(1-\delta p_4 + p_3\delta)^{-1}(b-a)$	$b - \delta(1-p_4)(1-\delta p_4 + p_3\delta)^{-1}(b-a)$
BB	$a + p_3\delta(1-\delta p_2 + \delta p_3)^{-1}(c-a)$	$c - \delta(1-p_2)(c-a)(1-\delta p_2 + \delta p_3)^{-1}$

For $p_3 \to 0$ and $p_4 \to 1$, these reduce to

strategy	state 1	state 2
AA	$(1-\delta(1-p_1))^{-1}\delta p_1 b$	b
AB	$c\delta p_1(1+\delta(p_1-p_2))^{-1}$	$c(1-\delta(1-p_1))(1+\delta(p_1-p_2))^{-1}$
BA	a	b
BB	a	$c + \delta(a-c)(1-p_2)(1-\delta p_2)^{-1}$

of $cp_1(1+(p_1-p_2))^{-1}$, while the other three strategies end up stuck in one state repeatedly getting the same payoff. For any value of δ, p_1, $p_2 \in (0,1)$, and a, b, and c, there will be at least one strategy that is optimal in both state 1 and state 2.

As $\delta \to 1$, the law of large numbers kicks in, and discounted payoffs converge in probability. Classical treatments of risk-aversion, then, leave asymptotically little room for non-trivial risk-aversion as long as agents compound lotteries and only care about the distribution of total final payoffs. A long strain of results, however, shows that subjects presented with sequences of risky choices may not aggregate them according to the laws of probability (Allais (1953)). Yaari (1987) constructs a model of choice in which he replaces the assumption that compound lotteries are handled in this fashion with a different assumption, leading him to a model that differs from expected utility; applying his model here amounts to allowing subjects to act as though p_1 and p_2 had different values than they do here. If $\min\{c,b\} > \max\{a,0\}$, the value of an agent playing any strategy will be weakly higher in state 2 than in state 1, and strictly so for δ, p_1, and $p_2 \in (0,1)$. Under these circumstances, it seems natural to frame a move from state 1 to state 2 as a gain and a move the other direction as a loss; in this case a risk state 1

	A	B
A	0,0	0,0
B	0,0	a,a

state 2

	A	B
A	b,b	b,b
B	b,b	c,c

	A	B
A	p_1	p_3
B	p_3	p_3

	A	B
A	p_4	p_2
B	p_2	p_2

Figure 5: Transition probabilities given are for transition to state 2.

averse agent will use a lower effective value for both p_1 and p_2; a risk-seeking agent will maximize an expression raising those numbers. If agents are risk averse when faced with potential gains but risk-seeking when faced with losses, as suggested by Kahneman and Tversky, they might maximize an expression that lowers p_1 and raises p_2.

A model closer to expected utility maximization would allow for attempted "smoothing" between stage-game payoffs, even where actual "consumption" may not be tied (temporally) to those payoffs. Agents would seek to optimize the expected net present value of a function of the payoffs. This, too, is a simple change from the derived expressions; risk-aversion in this framework would amount, for example, to using lower effective values of b and c.

While a setup with $p_3 = 0$ and $p_4 = 1$ is attractive in that the choice at each decision stage is one between a pure value and a lottery, it is unattractive in that some of the strategies do not induce ergodic outcomes, in which case it becomes hard to distinguish between full strategies. Accordingly, in this study we use interior values for these probabilities, allowing the agents to compare different lotteries instead of a lottery with a certainty.

2.2 Two player game

Now expand this to the two-person coordination game in figure 5 with $p_1 > p_3$, $p_4 > p_2$, $c > b > a > 0$; in each state, if the two players coordinate, always playing the same action, then they face the same decision problem as before, choosing between a safe immediate payoff or a chance at higher future payoffs in state 1, and between a safe payoff now or a higher payoff with a risk of lower payoffs in the future.[2] However, if they make different choices of action (one

[2]There's a bit of framing in the phrasing here, with continuation values measured against what would be earned from remaining in a given state. A perfectly rational player might no more view a transition from state 2 to state 1 as a loss than staying in state 1 when there

choosing A and the other choosing B), they get the worst of both: in state 1, they get no payoff now nor any chance of moving up to the higher continuation value, while in state 2, they get the lower available payoff now and the risk of moving to the lower continuation value.

As with many coordination games, each player needs to make an assessment of the other player's likely action, based on beliefs of that player's own preferences as well as that player's strategizing and higher order beliefs. In this game the decisions depend explicitly on preferences over exogenous risk factors, instead of just on uncertainty created by the challenge of coordination.

By construction of the game, if one player could publicly commit to an action, that same action would dominate[3] the other action for the other player. Two players with different risk preferences who fail to coordinate, then, are failing to coordinate for strategic reasons — either because they failed to anticipate each other's actions, or because they hope to affect the other player's choice of action in future rounds.

3 Procedure

The experiment was programmed and conducted with the software z-Tree (which is described in Fischbacher (2007)) and conducted in the Gregory Wachtler Experimental Economics Laboratory (room 107 in Scott Hall at Rutgers University). After being given instructions, the 12 subjects in the first session were presented with the one-player game, which they played for four different values of the parameters, which were chosen randomly; the parameters are given in figure 6. Each of the four rounds used a random stopping rule that imposed a 5% chance of ending the round at each stage. After those four rounds, subjects were then put in pairs, in which they played the two-player version, allowing us to study the ability of subjects to coordinate. In each round the players were randomly matched by the software with a partner, with different values of parameters; the same 5% stopping rule was used as in the one-player game.

In the second session, 20 subjects were first exposed to the two-player game, which they played 4 times, and then to the one-player game, which they played 3 times.

---

was a positive chance of moving to state 2, especially where $p_1 + p_2 = 1$. I expect, and even intend, for subjects to frame the decision as I have in the text, and for any differences between lotteries over gains and lotteries over losses to generate a difference in behavior between state 1 and state 2.

[3]In a first order stochastic sense, not a game theoretic sense.

round	a	b	c	p1	p2	p3	p4	B in st. 1	B in st. 2	stages
1	4	9	9	0.76	0.77	0.29	0.29	0.5909	0.1250	14
2	4	8	13	0.72	0.76	0.28	0.28	0.3803	0.2623	11
3	4	8	13	0.77	0.76	0.24	0.29	0.2927	0.2903	6
4	7	8	13	0.78	0.79	0.28	0.29	0.8387	0.5517	5
5	6	8	12	0.85	0.70	0.26	0.24	0.7411	0.1071	21
6	7	7	12	0.80	0.69	0.28	0.19	0.6923	0.6481	11
7	9	8	12	0.83	0.71	0.29	0.17	0.9712	0.5385	47
8	8	11	11	0.84	0.74	0.32	0.21	0.7284	0.0108	29
9	7	10	12	0.79	0.66	0.33	0.18	0.8529	0.1333	37
10	9	9	12	0.77	0.75	0.28	0.25	0.9867	0.6800	23
11	6	11	13	0.75	0.72	0.35	0.20	0.5682	0.3846	7
12	7	9	12	0.77	0.67	0.29	0.18	0.6303	0.3517	31
13	5	7	11	0.85	0.68	0.28	0.16	0.3957	0.2552	26
14	7	8	14	0.77	0.67	0.34	0.22	0.6506	0.5410	29
15	7	8	10	0.73	0.69	0.30	0.31	0.7852	0.1518	24
16	7	10	10	0.75	0.68	0.22	0.31	0.5000	0.0660	10
17	7	7	13	0.73	0.67	0.25	0.27	0.8212	0.6287	28

Figure 6: The parameters for each round, along with the fraction of choices in each state in which the subject chose B in that round.

previous stage	state 1		state 2	
action	count	B%	count	B%
A	227	34.80	444	7.66
B	604	87.58	181	78.45

Figure 7: Descriptive statistics for one-player games

previous stage		state 1		state 2	
own	partner	count	B%	count	B%
A	A	344	9.59	1018	1.57
A	B	134	34.33	87	21.84
B	A	134	51.49	87	73.56
B	B	1326	99.40	392	97.45

Figure 8: Descriptive statistics for two-player games

round	A/A	A/B	B/A	B/B	B in st. 1	B in st. 2	stages
1	6.203	4.322	**6.292**	5.303	0.5909	0.1250	14
2	5.472	6.107	5.828	**6.306**	0.3803	0.2623	11
3	5.536	**6.325**	5.648	6.052	0.2927	0.2903	6
4	5.558	6.249	7.450	**8.496**	0.8387	0.5517	5
5	5.951	6.365	6.941	**7.541**	0.7411	0.1071	21
6	5.371	6.223	7.000	**8.369**	0.6923	0.6481	11
7	6.308	6.254	8.434	**9.826**	0.9712	0.5385	47
8	8.380	5.660	**9.648**	8.863	0.7284	0.0108	29
9	7.725	6.309	**8.760**	8.583	0.8529	0.1333	37
10	6.461	5.876	9.000	**9.776**	0.9867	0.6800	23
11	8.228	6.403	**8.904**	8.182	0.5682	0.3846	7
12	6.912	6.190	8.110	**8.432**	0.6303	0.3517	31
13	5.599	5.908	6.137	**6.659**	0.3957	0.2552	26
14	5.908	7.222	7.555	**9.240**	0.6506	0.5410	29
15	5.345	4.957	7.453	**7.863**	0.7852	0.1518	24
16	6.741	5.059	**8.133**	7.693	0.5000	0.0660	10
17	4.854	6.533	7.000	**8.542**	0.8212	0.6287	28

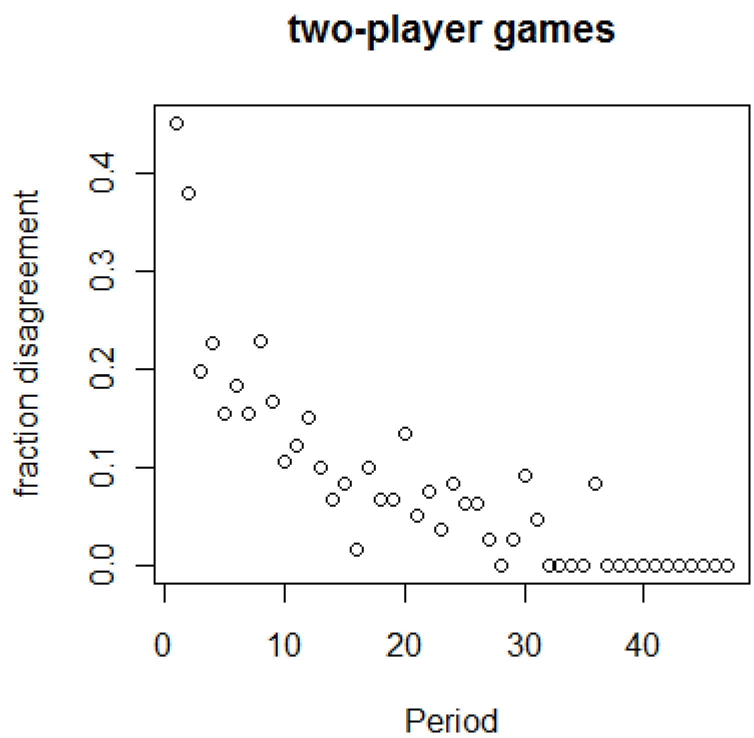
Figure 9: Expected payoff from different pure stationary Markov strategies for the realized parameters of each round.

4 Results

Some basic statistics are presented in figures 7 and 8; player responses in the tables exclude the first stage at which a state is encountered in a round; for example, if the first three stages of a round are in states 1, 1, and 2, respectively, stages 1 and 3 are excluded. This is so that the results can be broken down into categories by which action was played in the previous occurence of that state. Subject in one-player games rarely switch from B to A in state 1 or from A to B in state 2; switches in the other direction are somewhat more common.

In the two-player game, switches out of coordination are even less likely; of the four possible switches, the only one that happens in more than 3% of the opportunities is the attempt to switch to B when the players previously coordinated on A in state 1; note that A to B in state 1 is also the most common switch in the one-player game.

There are 134 stages in which the players are in state 1, will be in state 1 at least one more time in that round, and play different actions; there are 87 similar occurrences in state 2. In state 1 the player who played B switches almost half of the time at the next opportunity, but the player who played A attempts to



	state 1, Estimate	one-player Std. Error	state 2, Estimate	one-player Std. Error
(Intercept)	-1.5729	0.5873	-1.0321	0.8446
A/A−7	-0.5682	3.1863	2.2610	3.5780
A/B−7	-0.8367	2.2355	-0.9049	2.5752
B/A−7	-0.5401	5.2249	-4.6501	5.9522
B/B−7	0.9531	3.9823	3.8065	4.5279
	state 1, Estimate	**two-player Std. Error**	**state 2, Estimate**	**two-player Std. Error**
(Intercept)	-1.8443	0.2504	-2.3733	0.2358
A/A−7	-2.4520	0.7642	-0.4217	0.6047
A/B−7	-0.3261	0.3533	-0.3414	0.3089
B/A−7	2.9508	1.0319	-0.7747	0.8202
B/B−7	-0.6482	0.6264	1.1683	0.5062

Figure 10: Logit coefficients, regressing on theoretical values of pure strategies, subtracting 7 (approximately the mean value) to reduce the standard error on the intercept term. Negative intercepts indicate a preference for A when expected payoffs for different strategies are comparable.

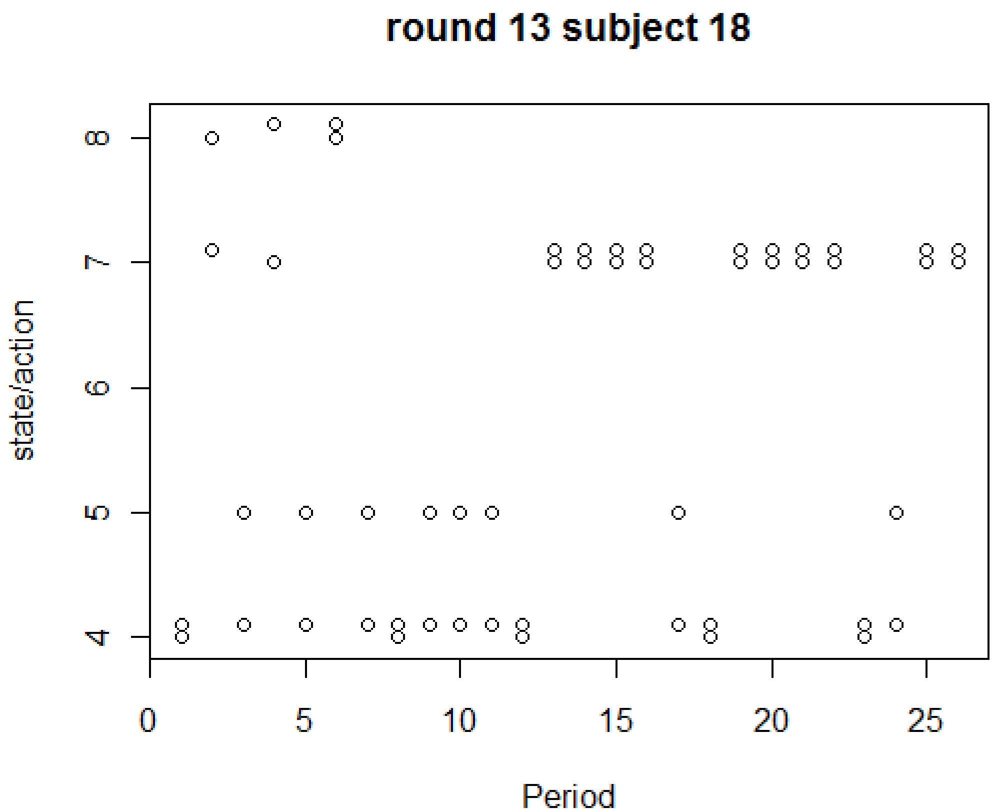


Figure 11: Round 13, subjects 18 and 25.

switch more than 1/3 of the time as well. In state 2, the players both tend to "dig in their heels" a bit more, switching about a quarter of the time. This can result in a sustained period of non-coordination; see figure 13 for an illustration.

Figure 11 shows one of the only three situations in which partners managed to switch directly from playing the same action in one stage to both playing the other action in the next occurrence of the same state. The subjects switch from playing B in stage 6 to, when they finally get back into state 2 in stage 13, both playing A, on which they coordinate for the remainder of that round. This game is an outlier in most of the logit models, and motivates the inclusion of a variable that captures the number of stages subjects spend in state 1 between recurrences of state 2; the significance of the coefficient on this variable demonstrates that agents are at least in part responding to experienced probabilities and not just the numbers presented.

4.1 Stationary Markov strategies

It was hypothesized that subjects in one-player games would play stationary Markov strategies. A rational agent basing its play solely on the data presented and not on experience in the game would have no reason to change its mind during play. It was also hypothesized that, in two-player games, both subjects would play the same stationary Markov strategy after a "coordinating period".

Of 1456 occasions in which subjects played a one-player game having been in the same state previously, the subject played the same thing 1229 times, viz. 84.4% of the time. Even beyond 20 periods, however, this only goes up to 88.9%. On the other hand, of 3522 stages in which subjects played a two-player game having been in the same state previously, the subject played the same action as the previous occasion 3302 times, viz. 93.8% of the time; beyond 20 periods, this goes up to 1178/1204, or 97.8%. (In the first 8 periods, the ratio is 87.1%.) This suggests that in fact the two-player games are converging to a pure stationary Markov equilibrium more quickly than the one-player games are.

In another, more formal test of these hypotheses, likelihood ratio tests were performed. A logit is fit against an indicator variable indicating either the previous action (for the one-player game) or the previous pair of actions (for the two-player game; this results in 3 dummy variables in addition to the intercept term), along with

- indicator variables for the subjects
- the 7 parameters describing the particular game
- both of these sets of explanatory variables.

These fits are done separately for each state, and separately for games with one player and games with two players excluding the first 20 stages of each round; because this excludes so many one-player games altogether, it is also done for one-player games excluding only the first 10 stages. If the subjects are playing stationary Markov strategies at this point, actions should be entirely predictable from the previous action in that state; other explanatory variables should not improve the prediction. Results are in figure 12. While subjects do not deviate at a significant level from two-player equilibrium, perhaps because coordination is valuable and difficult to achieve, they do, even after 20 periods, switch their actions, even in predictable ways.

4.2 Hypotheses on differing risk preferences

It was hypothesized that subjects with different risk preferences would take longer to coordinate than subjects whose risk preferences were similar. It was

		all 38	no dummies	no 7 params	just prev A
one player	log-likelihood	-6.821	-43.134	-16.599	-49.153
Period> 20	p-values	2.08×10^{-5}	0.0993	3.23×10^{-4}	NA
one player					
Period> 10	p-values	3.55×10^{-15}	.127	1.69×10^{-10}	NA
two player	log-likelihood	-21.714	-42.610	-24.234	-45.262
Period> 20	p-values	0.1479	0.6228	0.0889	NA
1p, > 20	p-values	2.38×10^{-3}	2.49×10^{-3}	.1214	(state 2)
1p, > 10	p-values	1.65×10^{-7}	3.63×10^{-7}	.0508	(state 2)
2p, > 20	p-values	.256	.242	.413	(state 2)

Figure 12: Likelihood ratio tests as described in the text. For one-player games, even after 20 periods, the previous action by itself is not the best predictor of the subsequent action, but for two-player games the data are more consistent with the assertion that convergence to a stationary markov strategy has taken place. This is true both in state 1 and in state 2.

of general interest to see how coordination took place, and whether there was an asymmetry in the resolution of disagreements, i.e. whether subjects who preferred B (the immediate option) were more or less likely to defer to subjects who preferred A than vice versa.

In state 2, there were 16 occasions when a disagreement was followed by A/A, and 6 when it was followed by B/B. In state 1, the respective figures were 31 and 8.[4] For a symmetric binomial distribution, these have two-tailed p-values of 0.0525 and 2.94×10^{-4}, respectively.

Figure 13 shows no discernible tendency to take longer to coordinate in state 1 on the basis of mismatched preferences, but state 2 does appear to exhibit such a tendency. In periods 2–6, the mismatched pairs disagree on 21 out of 61 occasions, while the similar pairs disagree 13 times out of 93. A random sample of 61 items from a population of 154 will include at least 21 out of 34 specially designated items with a probability of 8.30×10^{-13}.

Figure 10 demonstrates a generic tendency by subjects to prefer A in situations in which an expected payoff maximizing agent would be indifferent.

[4] See page 19 for an exhaustive breakdown on which action pairs followed each other in what quantity.

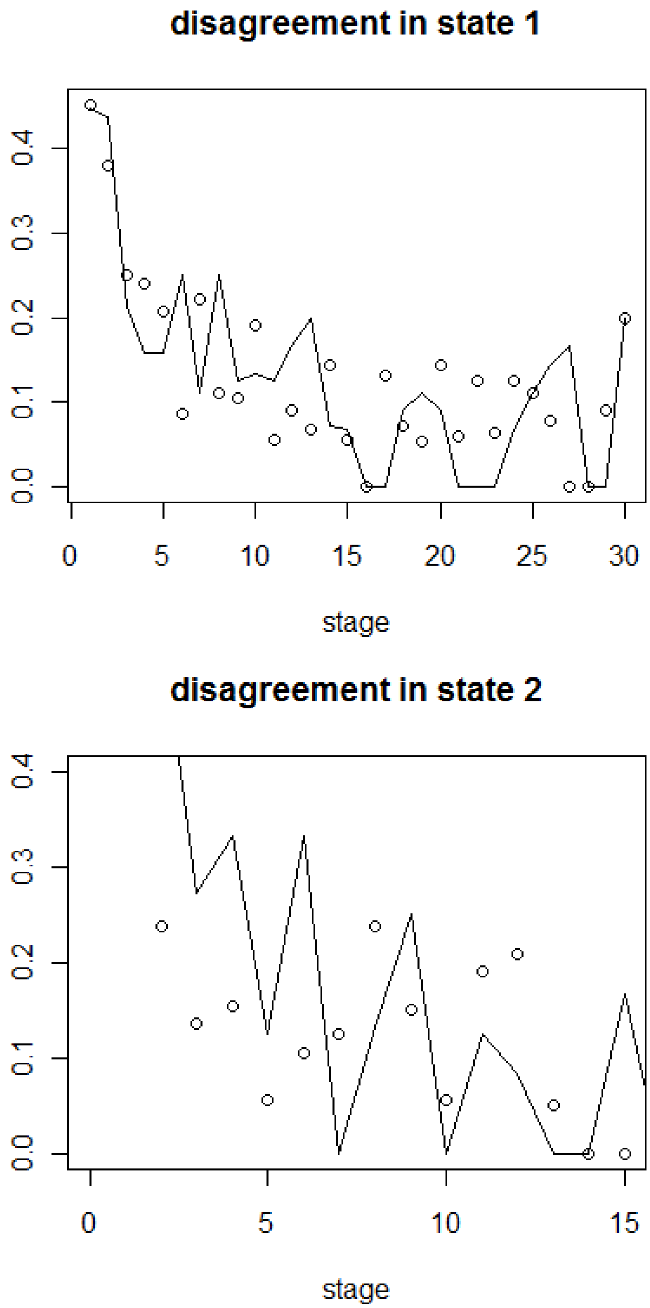


Figure 13: The circles represent the fraction of subject pairs that disagree in a given stage in a two-person game in which two subjects with similar behavior in the one-person game are matched. The lines represent games in which the partners exhibit different one-person game behavior.

5 Discussion and Conclusion

We have presented subjects with a fairly simple environment in which we can study the ability of subjects to coordinate on whether to take a large payoff now or take a chance of receiving a higher payoff in the future. Subjects had a tendency to prefer to enhance their expected continuation value, especially in the two-player game, in which differences tended to be resolved in that direction; if one agent wanted to take the higher payoff immediately, especially in state 2, but the other agent wanted to "play it safe", increasing the likelihood of staying in state 2, the former agent would typically defer to the latter's choice.

In the context of single-agent decision theory, a reluctance to risk moving from a profitable state to an unprofitable state may be related to the concept of "loss aversion" Kahneman and Tversky (1979)). It is interesting to note that an apparent loss aversion seems to be more pronounced in this case when agents are attempting to coordinate than when they do not face that challenge; the subject who is less loss-averse tends to defer to the subject who is more loss-averse.

The challenge of coordination also makes it more costly for a subject to experiment; the subject pairs thus lock more quickly into a pure stationary Markov strategy, while individual players exhibit more of a learning behavior, responding to experience by maintaining a willingness to change strategies even late in a round. Economic agents in the field may experience similar kinds of lock-in, in which agents find it difficult to make even Pareto-improving changes if the cost of coordinating the change seems likely to exceed the benefit from the switch.

While it would be interesting to attempt to infer effective one-player game parameters that would lead to the behavior in the two-player games, and perhaps follow that back to effective beliefs about what one's partner is likely to play, this would seem to require more data than have been gathered in this experiment.

References

Allais, M. (1953): "Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américane," *Econometrica*, 21(4), 503–546.

Camerer, C., and T.-H. Ho (1999): "Experience-weighted Attraction Learning in Normal Form Games," *Econometrica*, 67(4), 827–874.

Cooper, R. W., D. V. DeJong, R. Forsythe, and T. W. Ross (1990): "Selection Criteria in Coordination Games: Some Experimental Results," *American Economic Review*, 80(1), 218–233.

Ellsberg, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75(4), 643–669.

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

Gul, F., and W. Pesendorfer (2001): "Temptation and Self-Control," *Econometrica*, 69(6), 1403–1435.

Harsanyi, J. C. (1995): "A New Theory of Equilibrium Selection for Games with Complete Information," *Games and Economic Behavior*, 8, 91–122.

Harsanyi, J. C., and R. Selten (1988): *A general theory of equilibrium selection in games*. MIT Press, Cambridge, Mass.

Hillas, J., and E. Kohlberg (2002): "Foundations of strategic equilibrium," in *Handbook of Game Theory with Economic Applications*, ed. by R. J. Aumann, and S. Hart, vol. 3, chap. 42, pp. 1597–1663. Elsevier, 1 edn.

Kahneman, D., and A. Tversky (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263–291.

Mertens, J.-F. (2002): "Stochastic games," in *Handbook of Game Theory with Economic Applications*, ed. by R. J. Aumann, and S. Hart, vol. 3, chap. 47, pp. 1809–1832. Elsevier, 1 edn.

Mookherjee, D., and B. Sopher (1994): "Learning Behavior in an Experimental Matching Pennies Game," *Games and Economic Behavior*, 7(1), 62–91.

——— (1997): "Learning and Decision Costs in Experimental Constant Sum Games," *Games and Economic Behavior*, 19(1), 97–132.

Salant, Y. (2011): "Procedural Analysis of Choice Rules with Applications to Bounded Rationality," *American Economic Review*, 101(2), 724–748.

Shapley, L. S. (1953): "Stochastic games," *PNAS*, 39(10), 1095–1100.

Van Damme, E. (2002): "Strategic equilibrium," in *Handbook of Game Theory with Economic Applications*, ed. by R. J. Aumann, and S. Hart, vol. 3 of *Handbook of Game Theory with Economic Applications*, chap. 41, pp. 1521–1596. Elsevier.

Van Huyck, J., R. Battalio, and R. Beil (1990): "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *American Economic Review*, 80, 234–248.

Vieille, N. (2002): "Stochastic games: Recent results," in *Handbook of Game Theory with Economic Applications*, ed. by R. J. Aumann, and S. Hart, vol. 3 of *Handbook of Game Theory with Economic Applications*, chap. 48, pp. 1833–1850. Elsevier.

Yaari, M. (1987): "The Dual Theory of Choice under Risk," *Econometrica*, 55(1), 95–115.

action	previous	count
state 1, one-player		
A	-	57
B	-	51
A	A	148
B	A	79
A	B	75
B	B	529
state 2, one-player		
A	-	65
B	-	43
A	A	410
B	A	34
A	B	39
B	B	142

action	partner	prev A	prevPA	count
state 1, two-player				
A	A	-	-	26
A	B	-	-	32
B	A	-	-	32
B	B	-	-	52
A	A	A	A	280
A	B	A	A	31
B	A	A	A	31
B	B	A	A	2
A	A	A	B	41
A	B	A	B	47
B	A	A	B	24
B	B	A	B	22
A	A	B	A	41
A	B	B	A	24
B	A	B	A	47
B	B	B	A	22
A	B	B	B	8
B	A	B	B	8
B	B	B	B	1310
state 2, two-player				
A	A	-	-	54
A	B	-	-	22
B	A	-	-	22
B	B	-	-	44
A	A	A	A	986
A	B	A	A	16
B	A	A	A	16
A	A	A	B	20
A	B	A	B	48
B	A	A	B	3
B	B	A	B	16
A	A	B	A	20
A	B	B	A	3
B	A	B	A	48
B	B	B	A	16
A	A	B	B	4
A	B	B	B	6
B	A	B	B	6
B	B	B	B	376

Figure 14: Extended descriptive statistics.

Appendix: Instructions for Subjects

Introduction

You are about to participate in an experiment in the economics of decision making. Various research foundations have provided the funding for this research. The research is designed to study how people make decisions when facing uncertainty, both individually and in small groups. At the end of the experiment you will be paid for your participation, as outlined in the following instructions.

Background

Many decisions we make, especially financial decisions, require long-term considerations, but the future is inherently uncertain. While you may not have complete control over the future, however, you can often have some influence; for example, your car might break down even if you have maintenance done on it, or it might not break down even if you don't, but it's more likely to have a breakdown if you are behind on your regular maintenance.

In this experiment, you will be making a series of simple choices about a machine; you can make light use of it, in which case it produce lower profits for you in the short-run but is more likely to be more productive in the future, or you can make heavy use of it, making it produce more now. Your final payoff will depend on its total production.

Instructions

This experiment will be divided into several rounds; in each round you will make a series of binary choices (A or B). For each choice you will be awarded a certain number of an in-experiment currency unit called shillings, but your decision will also affect the number of shillings that each choice will be worth in later stages of the same round.

At the beginning of each round you will learn something about your machine for that round. It will have two possible conditions: it may be in good condition or bad condition. The round will be divided into stages; the condition of the machine may change from one stage to the next. If the machine is in good condition, it might wear down to bad condition for the next stage, but, because some regular maintenance is being done on the machine, it is also possible for a machine in bad condition to move back into good condition. In each stage you will be given two options as to how to use the machine. Choice A will generally produce less in the current stage than choice B, but it is also generally the case that a machine is more likely to go from bad condition to good condition with choice A, and is more likely to go from good condition to bad condition with choice B.

Figure 15 gives an example. In this example if the machine is in bad condition, choosing A will produce 0 shillings in this stage, but creates a 70% chance that the machine will be in good condition in the next stage; choosing B will produce 2 shillings in this stage, but creates only a 25% chance that the machine will be in good condition in the next stage. If the machine is in good condition, choosing A will produce 3 shillings in this stage, and creates a 40% chance that the machine will fall into bad condition for the next stage; choosing B will produce 4 shillings in this stage, but creates a 65% chance that the machine will fall into bad condition for the next stage.

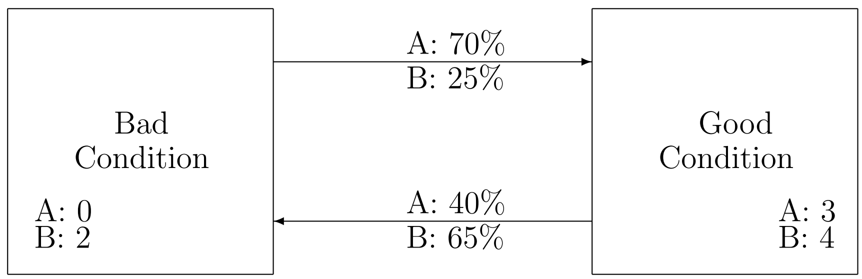
In each round the immediate payouts (i.e. productivity) will be given in whole numbers; there is a payout level for choice A when the machine is in bad condition, choice B when the machine is in bad condition, choice A when the machine is in good condition, and choice B when the machine is in good condition. Similarly, for each choice in each condition, you are provided with the probability that that choice would cause the machine to change conditions in the next stage, i.e. change from bad condition to good condition or from good condition to bad condition.

Throughout the round, all eight of these numbers will always be displayed near the top left of the screen. During the experiment, when the machine is in bad condition, the four numbers related your immediate options will be in red, while the numbers related to the machine in good condition will be in gray. When the machine is in good condition, the four numbers related your current choice will be in green, while the numbers related to the machine in bad condition will be in gray.

Different machines are different. Some produce a lot more when they are in good condition than when they are in bad condition, while for others the difference is small. Some are more sensitive than others to how hard they are run; some might be a lot more likely to deteriorate under high levels of production than others. It is important for you to decide how to trade off these factors; it may make sense to treat some machines differently from others.

The first screen capture on page 24 shows an example; in this example, the machine is in bad condition, so the payoffs (and some of the labels) are in red; those payoffs are 0 for A and 2 for B, as noted both in the main screen and above it; if the player chooses A there is a 75% chance that the machine will be in good condition for the next stage, which would offer 2 for A and 4 for B (as shown in gray above the main screen), but if the player chooses B there is only a 25% chance of switching.

At some point, the machine will abruptly stop producing profits. It is, unfortunately, prone to being hit by lightning, which happens with a 5% probability in each stage. This probability is independent of whether the machine is in good condition or bad condition. When that happens, the round ends; your total payoff from the round will be the total of the payoffs received in each stage of



A	B		A	B
0	2	payouts	3	4
0.7	0.25	switch prob	0.4	0.65

Figure 15: An example of the table you could see at the top of the screen to tell you about your machine, along with a diagram demonstrating what the numbers mean.

the round up to that point.

Partnership machines

In some rounds, you will be matched with one other person who is also participating in the experiment. In each such round, you will be randomly matched; your matching in one round is not related to your matching from another round. In these rounds you have to operate the machine in coordination with your partner. In each stage you will have to make your choices without knowledge of your partner's choice. As with the single-user machine, different short-term profits and probabilities will be associated with the two choices, so that if you both choose A you will tend to get a lower payoff than if you both choose B, but will be more likely to have a machine in good condition in the next round. If you choose A when your partner chooses B, or you choose B when your partner chooses A, the machine will produce the lower possible amount in this stage (as though you chose A) but will be more likely to be in bad condition in the next stage (as though you chose B).

As with the single-user machine, the 8 numbers that describe the machine will appear on near the top left of the screen. The numbers associated with the current condition of the machine will also be indicated to you in each stage in tables as explained in figure 16. The second screen capture on page 24 shows an example; in this case, the players are offered the green payoffs, which are 2 if either player chooses A and 4 if both choose B, as noted both in the grid to

	partner's choice	
	A	B
your choice A	2	2
your choice **B**	**2**	4

	partner's choice	
	A	B
your choice A	0.25	0.75
your choice **B**	**0.75**	0.75

Figure 16: For this set of tables, which would appear on the left side of the screen for a partnership machine, if you choose B and your partner chooses A, you get a payoff of 2 and a 75% chance of switching to the other set of payoffs. In the experiment, the entries in the table may be different from the numbers here, but in each case your payoff and switching probability is determined by the row of the choice you make and the column of the choice your partner makes.

the left and in the bar on top; if both players choose A there is a 25% chance of switching to the red payoffs in the next stage, while if either player chooses B there is a 75% chance of switching to the red payoffs for the next stage, which would offer payoffs of 0 and 2 (as shown in gray above the main screen).

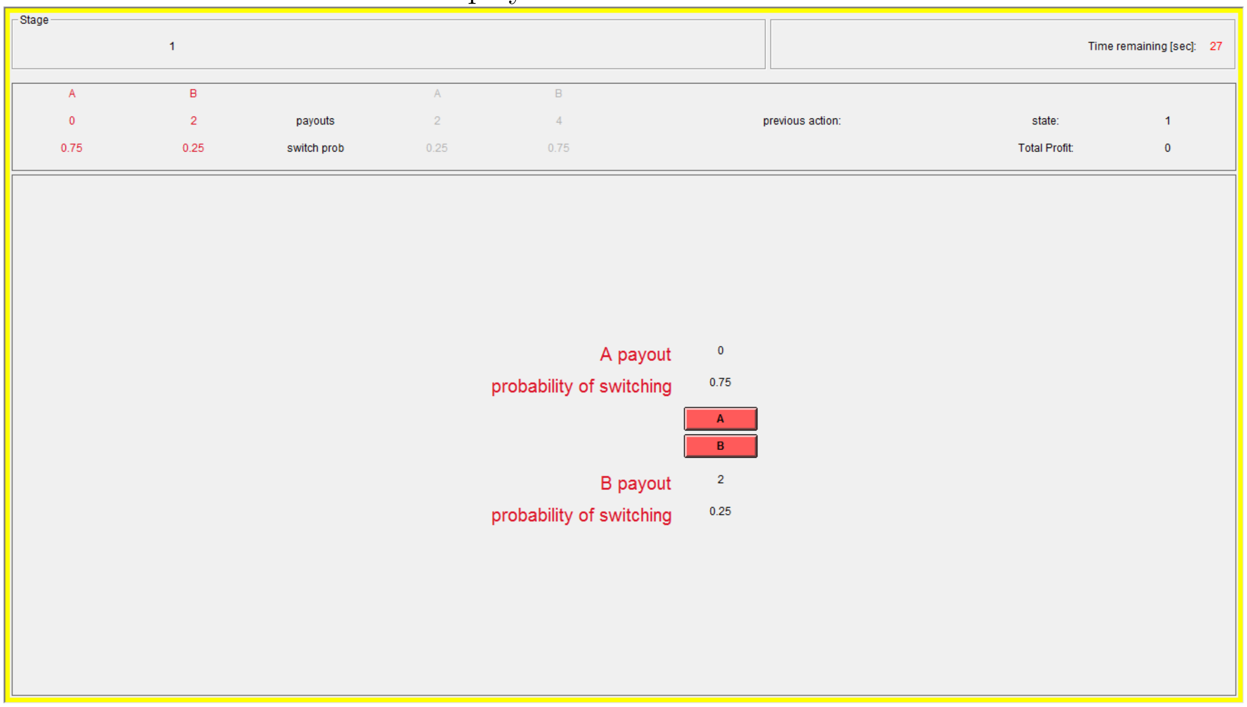
Payment

At the end of the experiment, your total earnings from each round (i.e. the total profits from all stages of all rounds) will be tallied up; you will be paid at a rate of $1 for every 100 shillings you have accumulate. (This is in addition to the $5 participation fee.) You will be paid your earnings in cash before leaving the experiment.

Summary

In the experiment you will make a series of choices between immediate payoffs and possible later payoffs, either alone or with a partner. Whether you are successful in earning profits in each round will depend both upon the decisions that you and the person you are paired with make, as well as on some degree of chance.

one-player machine:



partnership machine: