The Determinants of Efficient Behavior in Coordination Games

Pedro Dal Bó — Brown University & NBER

Guillaume R. Fréchette — New York University

Jeongbin Kim* — Caltech

May 20, 2020

Abstract

We study the determinants of efficient behavior in stag hunt games (2x2 symmetric coordination games with Pareto ranked equilibria) using both data from eight previous experiments on stag hunt games and data from a new experiment which allows for a more systematic variation of parameters. In line with the previous experimental literature, we find that subjects do not necessarily play the efficient action (stag). While the rate of stag is greater when stag is risk dominant, we find that the equilibrium selection criterion of risk dominance is neither a necessary nor sufficient condition for a majority of subjects to choose the efficient action. We do find that an increase in the size of the basin of attraction of stag results in an increase in efficient behavior. We show that the results are robust to controlling for risk preferences.

JEL codes: C9, D7.

Keywords: Coordination games, strategic uncertainty, equilibrium selection.

*We thank all the authors who made available the data that we use from previous articles, Hui Wen Ng for her assistance running experiments and Anna Aizer for useful comments.

1 Introduction

The study of coordination games has a long history as many situations of interest present a coordination component: for example, the choice of technologies that require a threshold number of users to be sustainable, currency attacks, bank runs, asset price bubbles, cooperation in repeated games, etc. In such examples, agents may face strategic uncertainty; that is, they may be uncertain about how the other agents will respond to the multiplicity of equilibria, even when they have complete information about the environment.

A simple coordination game that captures the main forces present in the previous examples is the well-known stag hunt game: a two-player and two-choice game with Pareto ranked equilibria. That game features two Nash equilibria in pure strategies, both players selecting stag, or both players playing hare; with stag being socially optimal (payoff dominant). Such a simple game allows us to study the conditions that lead people to coordinate on the efficient equilibrium.

The first experimental study of the stag hunt game, Cooper et al. (1992), focuses on a stag hunt game in which hare is risk dominant (that is, hare is the best response to the belief that the other player is randomizing 50-50 between stag and hare). They find that, absent communication, an overwhelming fraction of choices are in line with the risk dominant choice of hare. This is consistent with the idea that people may coordinate on the action most robust to strategic uncertainty. Relatedly, experiments on the minimum effort game, starting with Van Huyck et al. (1990), find a strong tendency for behavior to quickly settle on the minimum effort, where strategic risk is minimized (as opposed to the maximal effort—payoff dominant—equilibrium). Despite other studies that followed with mixed results, see for example Straub (1995) and Battalio et al. (2001), these early results created a strong notion that risk dominance was the key determinant of behavior in such coordination games.

In this paper, we return to stag hunt games for a systematic assessment of the determinants of efficient behavior (playing stag) using two data sets. First, we study behavior in the metadata from eight previous experiments on stag hunt games. Second, we study behavior from a new experiment that allows us to easily explore more parameter combinations than in the previous experiments. In each round of this experiment, subjects participate in sixteen simultaneous stag hunt games with different payoff parameters. Moreover, in some sessions we use the lottery procedure introduced by Roth & Malouf (1979) to induce risk neutral preferences. This allows us to explore the role of risk preferences on equilibrium selection.

We find that, consistent with the previous experimental literature, people do not necessarily coordinate on the efficient equilibrium. In fact, for some treatments, only a very small minority plays stag. The fact that payoff dominance is not used by the subjects as an equilibrium selection criterion suggests that strategic uncertainty may be important in coordination games. As such, one may believe that agents would choose actions corresponding to the equilibrium most robust to strategic uncertainty, that is, the risk dominant action (Harsanyi & Selten (1988)). While we find that the rate of stag is higher on average when it is risk dominant, it is not always the case that a majority of people coordinate on the risk dominant equilibrium. There are treatments in which only a minority of subjects choose the risk dominant action.

Although risk dominance does not predict equilibrium selection, we do find that a measure of the risk arising from strategic uncertainty can help explain behavior. The share of subjects choosing stag is increasing in the size of its basin of attraction.[1]

Interestingly, although the effect of the size of the basin of attraction of stag on efficient behavior is found both in the metadata from the previous literature and in the experiments using our new design, the exact relation is different. The rate of stag for intermediate sizes of the basin of attraction of stag is lower in the new experiment (where subjects participate in several games simultaneously) than in the earlier experiments (where they participate in only one game at a time). This suggests that behavior in a coordination game may depend not only on the parameters of the game, but also on the context in which it is being played.

Finally, we show that behavior in stag hunt games is not affected by risk attitudes. Using

[1]The size of the basin of attraction of stag is the maximum probability of the other player choosing hare that would still make a player choose stag. There is a connection here with the study of cooperation in repeated games. If the infinitely repeated game is suitably simplified, it can be reduced to a stag hunt game, and thus one can identify parameters for which cooperation can be supported as part of a risk dominant and payoff dominant equilibrium versus others where only defection can be risk dominant, see Blonski & Spagnolo (2015). Dal Bó & Fréchette (2011) and Dal Bó & Fréchette (2018) show that variation in cooperation rates are related to the size of the basin of attraction of the strategies in the simplified game. It has also been found that the basin of attraction is an important determinant of behavior in other games, see Healy (2016), Calford & Oprea (2017), Embrey et al. (2017), Vespa & Wilson (2017), Kartal & Müller (2018), and Castillo & Dianat (2018).

the lottery procedure to induce risk neutral preferences (Roth & Malouf (1979)) does not affect behavior in the games we study. Hence, the failure of payoff or risk dominance to explain coordination in and of themselves is not due to risk attitudes, but instead may be due to fundamental ways in which people respond to strategic uncertainty.

2 Theoretical Background

The stag hunt game is a two-player game with two actions, stag and hare, with the payoffs as shown in Table 1 (Original) with the constraint on payoffs that $T < R > P > S$. Note that $(stag, stag)$ is a Nash equilibrium given that $T < R$, but $(hare, hare)$ is also a Nash equilibrium given that $S < P$. Given that $R > P$, the former equilibrium results in higher payoffs than the latter one. Following Harsanyi & Selten (1988), we say that $(stag, stag)$ is the payoff dominant (or Pareto efficient) equilibrium and stag is the payoff dominant action. There is also a mixed strategy Nash equilibrium in which subjects play hare with probability $\frac{1}{1+\frac{P-S}{R-T}}$, assuming that subjects are risk neutral.

Table 1: Stag Hunt Game - Row Player's Payoffs

Original

	hare	stag
hare	P	T
stag	S	R

Normalized

	hare	stag
hare	$\frac{P-P}{R-P} = 0$	$\frac{T-P}{R-P} = 1 - \Lambda$
stag	$\frac{S-P}{R-P} = -\lambda$	$\frac{R-P}{R-P} = 1$

Under the assumption that behavior is not affected by linear transformation of payoffs, any stag hunt game with four parameters R, S, T, P as in the left panel of Table 1 can be normalized to a game with only two parameters, Λ and λ as in the right panel of Table 1 (Normalized). The parameter Λ denotes the loss arising from an unilateral deviation from the efficient equilibrium, while the parameter λ denotes the loss arising from an unilateral deviation from the inefficient equilibrium. This normalization will allow us to compare behavior across stag hunt experiments while keeping track of only two payoff parameters, Λ and λ, instead of the four original parameters.

How should we expect people to behave in the stag hunt game? Previous authors, see for example Luce & Raiffa (1957), Schelling (1960), and Harsanyi & Selten (1988), have theorized that people would coordinate on the efficient equilibrium, in this case $(stag, stag)$. This is quite intuitive for a game as the one shown in the left panel of Table 2 (example 1), but may be less so in the game shown in the right panel (example 2). The reason is that, in the latter game, a small uncertainty about the action of the other player would make stag a sub-optimal choice. In other words, $(stag, stag)$ is not very robust to strategic uncertainty in the second example.

Table 2: Stag Hunt Games - Row Player's Payoffs

Example 1

	hare	stag
hare	0	-1
stag	-1	1

Example 2

	hare	stag
hare	0	-1
stag	-100	1

The robustness to strategic uncertainty of the equilibrium $(stag, stag)$ can be measured by the maximum probability of other subject playing hare that still makes stag a best response. This number is provided by the probability of hare in the mixed strategy Nash equilibrium and is usually referred to as the size of the basin of attraction of stag. Under normalized payoffs, the size of the basin of attraction of stag is equal to $\frac{\Lambda}{\Lambda+\lambda}$.[2] Note that, intuitively, this number is decreasing in λ and increasing in Λ. Following Harsanyi & Selten (1988), we say that stag is risk dominant if its basin of attraction is greater than one half. If that is the case, $(stag, stag)$ is more robust to strategic uncertainty than $(hare, hare)$. Harsanyi & Selten (1988) proposed risk dominance as an alternative equilibrium selection criterion. The idea that people may coordinate on the risk dominant equilibrium received support from evolutionary theories (see Kandori et al. (1993) and Young (1993)).

While the previous experimental literature on coordination games has shown that subjects do not necessarily coordinate on the efficient equilibrium (see Cooper et al. (1990), Van Huyck et al. (1990), and Cooper et al. (1992)), the literature has not yet provided a clear answer to the issue of when people would coordinate on the efficient equilibrium. In particular, we seek to answer the following questions. Is it the case that people coordinate on the efficient outcome if, and only if, it is risk dominant? Does the prevalence of the efficient

[2] If agents are not risk neutral, then the size of the basin of attraction and the parameter values for which hare is risk dominant are different. In particular, if subjects are risk averse, the size of the basin of stag is smaller.

action, stag, depend on how robust it is to strategic uncertainty (that is, the size of its basin of attraction)? Moreover, are the answers to these questions affected by the subjects' risk attitudes?

3 Determinants of efficient behavior: previous experiments

We have identified nine previous stag hunt experiments with experimental designs that are amenable to be analyzed jointly; of which we were able to obtain the data from eight of them.[3] The collected data satisfies the following conditions: (1) 2x2 stag hunt game, (2) no pre-play communication, and (3) using non-fixed matching across periods.[4]

We refer to this data set as the metadata for simplicity, even though it is a collection of raw data sets rather than a collection of aggregated data sets as in a typical metadata.

Table 3 summarizes the treatments in the previous experiments that satisfy the conditions described above. Some of the papers have treatments, not reported here, that do not fit our criteria, e.g., treatments with pre-play communication or with fixed matching throughout the experiment, and those treatments are not included in our analysis. We have data from eight articles, involving 18 different treatments (combinations of the four payoff parameters T, R, P, and S), with 90 experimental sessions and 970 subjects. The vast majority of treatments are such that hare is risk dominant (14 out of 18 treatments) and in only two treatments stag was risk dominant. That is, in most treatments from previous articles there is a tension between payoff dominance and risk dominance. Moreover, while the basin of attraction of stag goes from $\frac{1}{8}$ to $\frac{2}{3}$, there is limited variation in this dimension as 72% of treatments have a size of the basin of stag in a small interval (between $\frac{1}{5}$ to $\frac{1}{3}$).

[3]The eight articles from which we have data are Cooper et al. (1992), Straub (1995), Battalio et al. (2001), Clark et al. (2001), Duffy & Feltovich (2002), Schmidt et al. (2003), Dubois et al. (2012), and Feltovich et al. (2012). The data from Charness (2000) is no-longer retrievable.

[4]In particular, Clark et al. (2001), Schmidt et al. (2003), and Straub (1995) use the perfect stranger matching. In Cooper et al. (1992), subjects play against every other player twice: once as a row player and once as a column player. Battalio et al. (2001), Dubois et al. (2012) and Feltovich et al. (2012) use random matching across periods. In Duffy & Feltovich (2002), subjects are assigned to the role of a row or a column player which remain fixed throughout the experiment and play with every other subject of the opposite role.

Table 3: Treatment Parameters in Prior Experiments

	Λ	λ	**Basin S**	**Sessions**	**Subjects**	**Periods**
Battalio et al. (2001)				**24**	**192**	
	0.091	0.364	0.2	8	64	75
	0.2	0.8	0.2	8	64	75
	2	8	0.2	8	64	75
Clark et al. (2001)				**5**	**100**	
	0.333	2.333	0.125	2	40	10
	1	4	0.2	2	40	10
	3	9	0.25	1	20	10
Cooper et al. (1992)	1	4	0.20	**3**	**30**	22
Dubois et al. (2012)				**24**	**192**	
	0.091	0.364	0.2	8	64	75
	0.375	1.5	0.2	8	64	75
	0.375	1.5	0.2	8	64	75
Duffy and Feltovich (2002)	1	3	0.25	**3**	**60**	10
Feltovich et al. (2012)				**10**	**186**	
	1	2.2	0.313	6	90	20
	2	1	0.667	4	96	40
Schmidt et al. (2003)				**16**	**160**	
	0.5	1.5	0.25	4	40	8
	1	1	0.5	4	40	8
	1	1	0.5	4	40	8
	1	3	0.25	4	40	8
Straub (1995)				**5**	**50**	
	0.2	0.4	0.333	1	10	9
	1	0.5	0.667	1	10	9
	1	1	0.5	1	10	9
	1	3	0.25	1	10	9
	1	4	0.2	1	10	9
Total				**90**	**970**	

We study behavior in period 1 as well as in period 8. The latter period is the largest period with observations in every treatment, as the experiment with the smallest number of periods is Schmidt et al. (2003) with 8 periods. Focusing on period 8 allows us to study behavior across treatments after subjects have gained some experience.

Figure 1 shows the average rate of stag for each combination of Λ and λ in the metadata for periods 1 and 8. The percentage of stag increases with Λ and decreases with λ in both periods. The impact of these parameters on behavior increases as subjects gain experience, as shown by the greater differences in period 8 than in period 1. The first two columns in Table 4 confirm these results. These columns show the estimates of the marginal effects of Λ and λ in a Probit analysis of subject choices where the dependent variable is an indicator variable equal to 1 if the subject chose stag. The estimated effect of Λ on the probability of choosing stag is positive and significant at the 1% level, while the effect of λ is negative and significant at the 1% level. Note that the magnitude of the effects increase with experience.

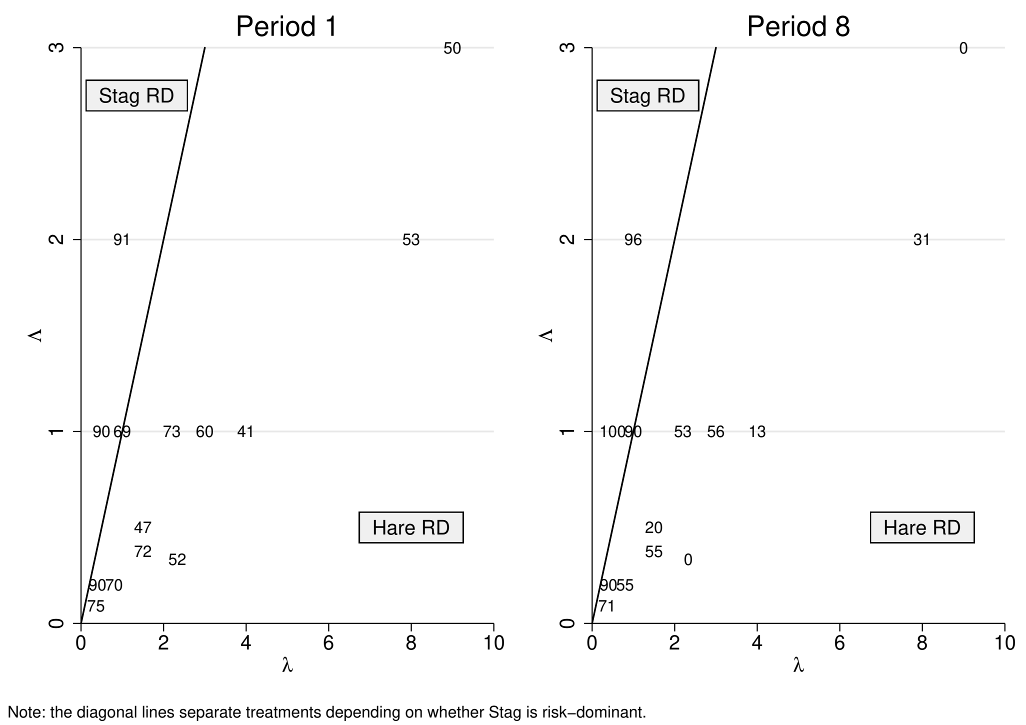


Figure 1: Meta-analysis: Stag % by Λ and λ combination

The last two columns in Table 4 show that the results are robust to whether the experiment used the method in Roth & Malouf (1979) to induce risk neutral preferences.[5] We find that the percentage of stag does not depend on a consistent or significant way on whether risk neutral preferences are induced.

Table 4: Meta-analysis. Determinants of Stag (Probit Analysis—Marginal Effects)

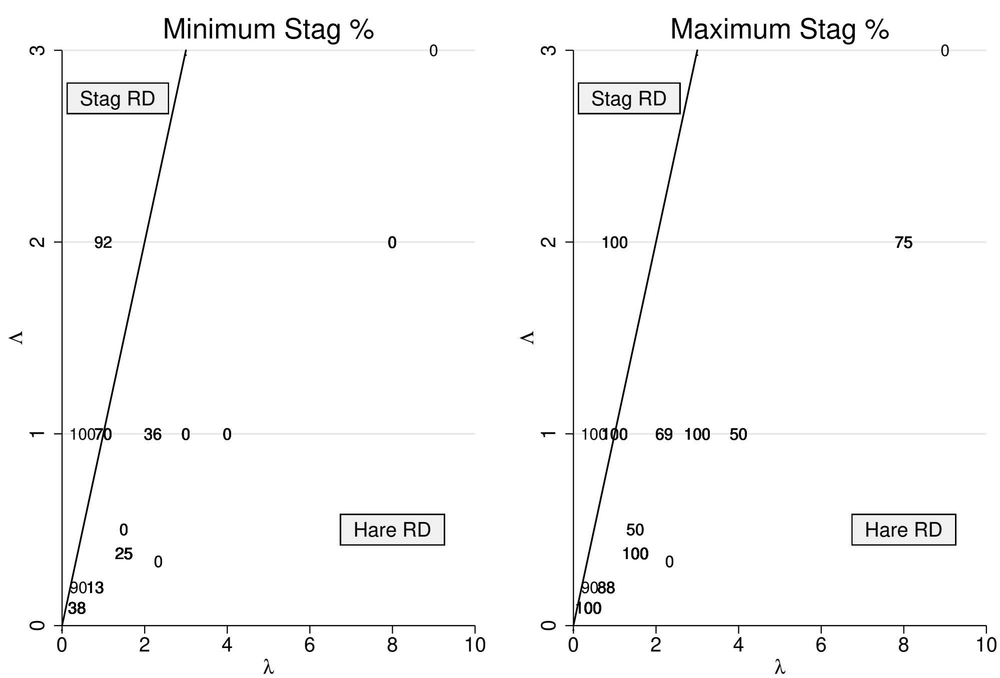
	Period 1	**Period 8**	**Period 1**	**Period 8**
Λ	0.15***	0.41***	0.15***	0.41***
	(0.031)	(0.070)	(0.031)	(0.069)
λ	-0.07***	-0.18***	-0.07***	-0.18***
	(0.011)	(0.026)	(0.011)	(0.027)
Lottery (d)			-0.02	0.04
			(0.052)	(0.090)
Observations	970	970	970	970

Lottery denotes Roth-Malouf lottery was used.
(d) denotes dummy variable, effect of change from 0 to 1 is reported.
Standard errors clustered at paper level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Consistent with the previous literature, Figure 1 shows that subjects do not necessarily coordinate on the efficient equilibrium. In many treatments, we observe very low stag rates, with two treatments reaching 0% of stag by period 8.

Since payoff dominance does not work as an equilibrium selection devise, let us consider risk dominance. As Figure 1 and Table 5 show, subjects are significantly more likely to choose stag when it is risk dominant. In fact, for the two treatments with stag being risk dominant we observe that most subjects choose stag in period 8. This is consistent with the idea that stag being risk dominant may be a sufficient condition for subjects to coordinate on the efficient equilibrium. However, is stag being risk dominant also a necessary condition for efficient coordination? That does not seem to be the case. There is great variation in behavior for treatments in which stag is not risk dominant; we see the rate of stag going from 0% to 100% across sessions (see Figure 2 which shows the minimum and maximum percentage of stag across sessions in each treatment). That is, there are several sessions

[5] Among the papers included in our metadata, Cooper et al. (1992), Straub (1995), and Duffy & Feltovich (2002) used the lottery method proposed by Roth & Malouf (1979) to induce risk neutral preferences.



Note: the diagonal lines separate treatments depending on whether Stag is risk–dominant.

Figure 2: Meta-analysis. Maximum and Minimum Stag % by Session in a Treatment (period 8)

in treatments in which stag is not risk dominant which reach perfect coordination on the efficient equilibrium in period 8. That is, based on the metadata from the previous articles, risk dominance does not seem to be a necessary condition for efficient coordination.

Hence, neither payoff nor risk dominance on their own predict efficient coordination. However, the payoff parameters (Λ and λ) may not impact behavior linearly, as assumed in Table 4, nor discontinuously as a function of whether stag is risk dominant, as assumed in columns 1 and 2 of Table 5. Instead, we study next whether the percentage of stag in a treatment can be explained by the robustness of the efficient equilibrium $(stag, stag)$ to strategic uncertainty. As discussed in section 2, we measure robustness to strategic uncertainty by the size of the basin of attraction of stag, which is equal to $\frac{\Lambda}{\Lambda+\lambda}$.

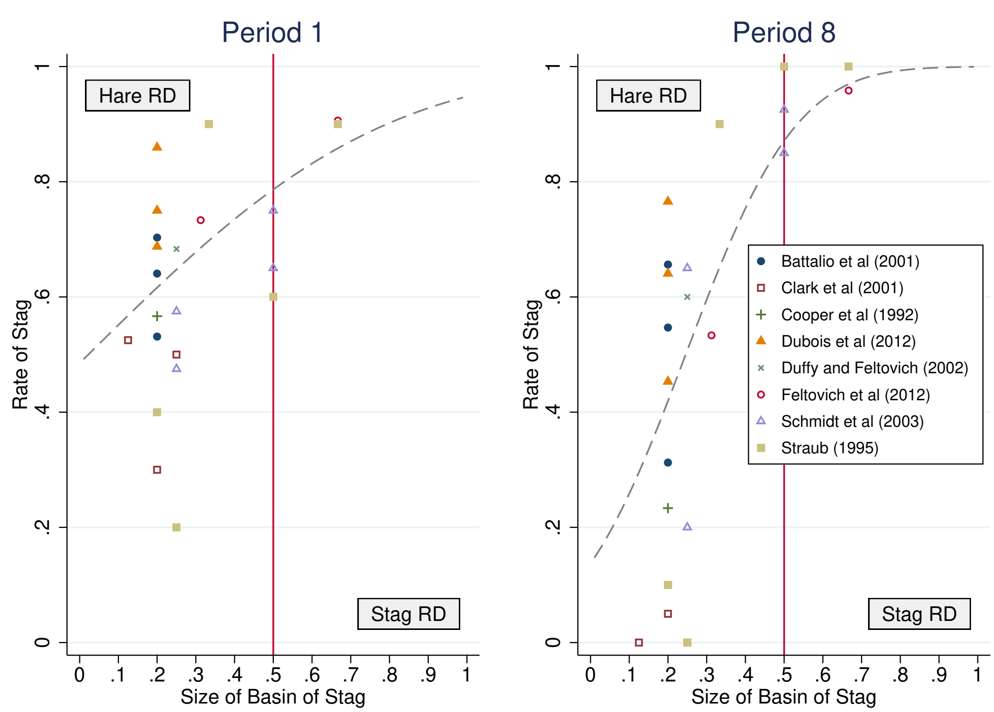


Figure 3: Meta-analysis. Rate of Stag and its Basin

Figure 3 shows the average rate of stag in each treatment and article in the metadata for periods 1 and 8 as a function of the size of the basin of attraction of stag. The doted line is a simple Probit linear fit through these points.

Overall, the rate of stag is positively correlated with its basin of attraction: as the basin of attraction increases, the rate of stag also increases. This relation is present from the onset, but becomes more pronounced with experience (see columns 3 and 4 in Table 5). This result is robust to controlling for whether risk neutral preferences were induced (see columns 5 and 6 in Table 5). Note that the induction of risk neutral preferences does not affect the prevalence of stag as a function of the size of the basin of attraction.

Figure 3 also suggests that risk dominance on its own cannot account for all of the observed variation, as the rate of stag amongst treatments with a basin of stag between 0.1 and 0.5 varies from 0% to 90% in period 8. Hence, it is not surprising that previous studies reached different conclusions in this regard. For instance, Cooper et al. (1992) conclude that "coordination failures always occur" while Schmidt et al. (2003) found "players selecting the payoff dominant strategy more often than not, [...] supporting Harsanyi and Selten's original assertion [...]" Although some of these differences could be accounted for by variation in the basins of attraction, others happen at a given value of the basin—consider for instance the variability in results when the basin of stag is 0.2. Part of this variation for a given basin of attraction may be explained by differences in experimental designs. For example, different implementations resulted in differences in the number of times the same pair of subjects was expected to play together (as a function of the matching protocol and the total number of periods in a session). This seems to explain part of the observed variation in behavior for the treatments with a basin of attraction equal to 0.2, with higher rates of stag in experiments in which the expected number of interactions for the same pair of subjects was higher. Clearly, as there are many elements of experimental design that may vary across experiments, this poses a challenge for a meta-study.

Another limitation of this meta-study is that, although the original (non-normalized) payoffs are quite different across previous experiments, they involve a narrow range for the basin of attraction of stag: as mentioned before, 72% of treatments have a size of the basin of stag between $\frac{1}{5}$ to $\frac{1}{3}$. This narrow range in the available treatments limits the study of how strategic uncertainty affects behavior based on the metadata.

Therefore, to study the determinants of efficient coordination more systematically, we turn to a new experimental design, which will allow to consider more variation in the variables

Table 5: Meta-analysis. Determinants of Stag (Probit Analysis—Marginal Effects)

	Period 1	**Period 8**	**Period 1**	**Period 8**	**Period 1**	**Period 8**
Stag RD (d)	0.27***	0.46***	0.19**	-0.08	0.19**	-0.08
	(0.029)	(0.041)	(0.076)	(0.143)	(0.077)	(0.145)
Basin of stag			0.28	1.82***	0.29	1.83***
			(0.229)	(0.318)	(0.230)	(0.319)
Lottery (d)					-0.04	-0.04
					(0.054)	(0.094)
Observations	970	970	970	970	970	970

Lottery denotes Roth-Malouf lottery was used.
(d) denotes dummy variable, effect of change from 0 to 1 is reported.
Standard errors clustered at the paper level in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

of interest. In addition, unlike the meta-study, the comparisons will not involve variation in methods, eliminating the possibility of confounders.

4 The New Experimental Design

The main design innovation is to allow many more comparisons across parameters by presenting multiple stag hunt games simultaneously on the subjects' screen. More specifically, each session consists of 15 periods in which subjects participate anonymously through computers in the coordination games presented in Table 6. Parameter T take values in the set $\{25, 45, 65, 85\}$ and parameter S take values in the set $\{10, 20, 30, 40\}$. The relevant T and S for each stage games are known to subjects.[6] This results in 16 coordination games in each period—see the decision screen in Figure 15 of the Appendix.[7]

Table 6: Stag Hunt Game - Row Player's Payoffs

Original

	hare	stag
hare	60	T
stag	S	90

Normalized

	hare	stag
hare	0	$-\Lambda$
stag	$1-\lambda$	1

[6]In terms of normalized payoffs, the 16 games have Λ in the set $\{\frac{1}{6}, \frac{5}{6}, \frac{3}{2}, \frac{13}{6}\}$ and λ in the set $\{\frac{2}{3}, 1, \frac{4}{3}, \frac{5}{3}\}$
[7]The actions were simply described as "1" and "2" in the experiment.

Table 7: Size of Basin of Attraction of Stag

	λ			
Λ	2/3	1	4/3	5/3
1/6	0.2	0.143	0.111	0.091
5/6	**0.556**	0.455	0.385	0.333
3/2	**0.692**	**0.6**	**0.529**	0.474
13/6	**0.765**	**0.684**	**0.619**	**0.565**

Note: Bold font denotes Stag is Risk Dominant.

The set of possible values of T and S were chosen to reach two objectives. First, we want to have large and systematic variation in the parameters of the games. In particular we want to have large variations in the size of the basin of attraction of stag. The new experiments have the size of the basin of stag going from 0.091 to 0.765, with many intermediate values. Second, we want to have many treatments for which stag is risk dominant so as to be able to study if that condition is sufficient for subjects to coordinate on the efficient equilibrium. Half the treatments have stag being risk dominant in the new experiment.

Subjects were randomly matched in each period to another subject, with the same pair not matched twice (perfect strangers). We have two main treatments, *Baseline* and *Lottery*, which differ by how subjects are paid. In *Baseline*, one randomly chosen game in one randomly chosen period is used to pay subjects at the exchange rate of \$35 per 100 points. In *Lottery*, one randomly chosen game in one randomly chosen period is used to pay subjects following the lottery procedure introduced by Roth & Malouf (1979) to induce risk neutral preferences. The points of the chosen game are the probabilities (in percentages) of earning \$35. In addition, in both treatments, subjects are paid a \$5 participation fee and a \$5 show-up fee. Note that for any given outcome of the game, the expected payoff is equal across the two treatments, thus facilitating comparisons.

The experiment was programmed using z-Tree (Fischbacher 2007). We conducted four experimental sessions for each of these two treatments with a total of 140 subjects. See Table 13 in the Appendix for the number of subjects per session and treatment. The experimental sessions lasted less than an hour. The subjects were Brown University undergraduates recruited through advertisement in university web pages, leaflets, and signs posted on campus. Subjects earned \$34.35 on average, with a minimum of \$10 and a maximum of \$45, including the participation fee and the show-up fee of \$10.[8]

5 Results from the New Experiments

We start this section by focusing on the *Baseline* treatment and study the conditions under which subjects coordinate on the Pareto efficient equilibrium $(stag, stag)$.

Averaging across games, 61% of subjects chose stag in period 1 and 51% in period 15. As shown in Table 8, it is not the case that subject coordinate on stag regardless of the payoff matrix. Note that in period 1, a majority of subjects chooses stag in only 11 of the 16 games and this number is reduced to 7 in period 15. Consistent with the previous literature, this shows that payoff dominance is not an appropriate equilibrium selection criterion.

Moreover, behavior in these games depends on the parameters Λ and λ as it did in previous experiments: the percentage of stag increases with Λ and decreases with λ—see Table 8. These effects are significant at the 1% level and increasing with experience—see Table 9.

Is risk dominance an appropriate equilibrium selection criterion? It is the case that the rate of stag is significantly higher in games in which it is risk dominant—see Tables 8 and 10. However, the rate of stag is far away from 100% for most of these games and, even, one of them (the game with $\Lambda = \frac{3}{2}$ and $\lambda = \frac{4}{3}$) has a lower rate of stag than for a game in which stag is not risk dominant (the game with $\Lambda = \frac{1}{6}$ and $\lambda = \frac{2}{3}$). Moreover, note that even for the same game in which stag is risk dominant, different sessions may exhibit very different behavior. For example, as shown in Figure 4, in the game with $\Lambda = \frac{13}{6}$ and $\lambda = \frac{5}{3}$, one session reaches levels of stag above 90% while another session falls below 20% as subjects gain experience. Thus, it is not the case that stag being risk dominant necessarily leads to high rates of stag. This is in contrast to what we find for the two treatments with stag being risk dominant in the metadata. In section 6 we will further discuss the differences of observed behavior between our new experiments and the metadata.

We study next whether the prevalence of stag can be explained by the robustness of

[8]The minimum of \$10 and the maximum of \$45 were both reached in the *Lottery* treatment. In the *Baseline* treatment the minimum and maximum earnings were \$17 and \$41.50.

Table 8: Percentage of Stag by Game in *Baseline*

Panel A: Period 1

	λ			
Λ	2/3	1	4/3	5/3
1/6	60.00	51.43	38.57	40.00
5/6	**72.86**	55.71	47.14	44.29
3/2	**84.29**	**74.29**	**58.57**	42.86
13/6	**84.29**	**81.43**	**72.86**	**62.86**

Panel B: Period 15

	λ			
Λ	2/3	1	4/3	5/3
1/6	50.00	21.43	11.43	8.57
5/6	**84.29**	35.71	20.00	10.00
3/2	**95.71**	**80.00**	**40.00**	18.57
13/6	**97.14**	**95.71**	**91.43**	**57.14**

Panel C: All Periods

	λ			
Λ	2/3	1	4/3	5/3
1/6	57.90	35.05	23.81	20.57
5/6	**78.57**	47.43	34.10	25.33
3/2	**94.38**	**73.90**	**49.14**	32.38
13/6	**96.67**	**93.71**	**84.95**	**63.90**

Note: Bold font denotes Stag is Risk Dominant.

Table 9: *Baseline*: Determinants of Stag (Probit Analysis - Marginal Effects)

	Period 1	**Period 8**	**Period 15**
Λ	0.15***	0.32***	0.45***
	(0.018)	(0.042)	(0.057)
λ	-0.30***	-0.57***	-0.83***
	(0.042)	(0.029)	(0.087)
Observations	1120	1120	1120

Clustered at session level standard errors in parentheses

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

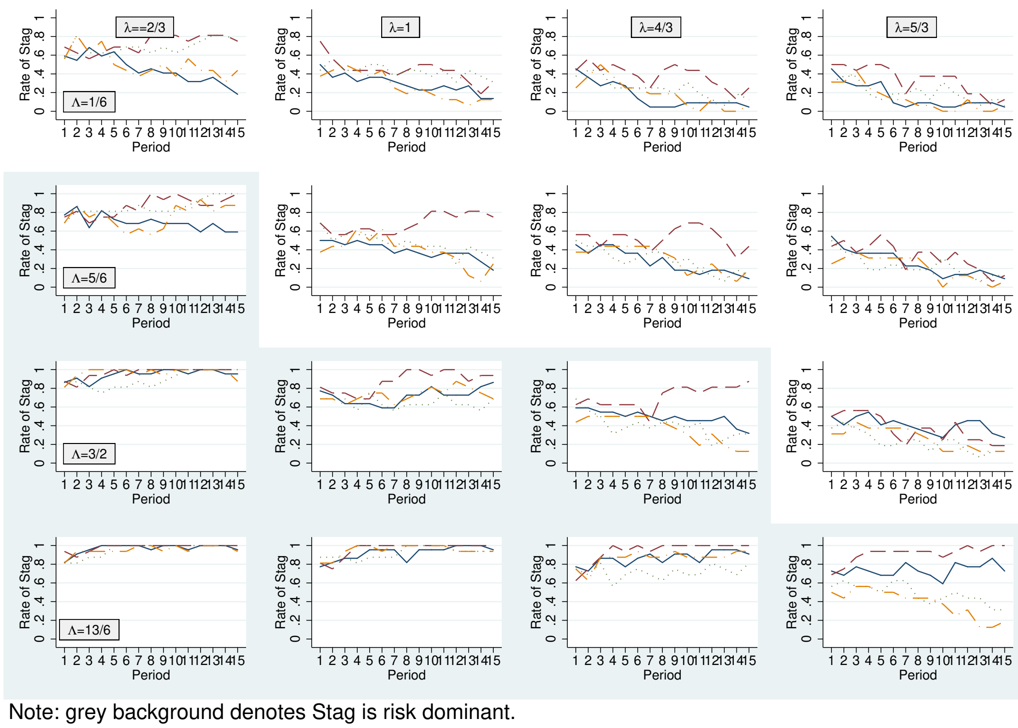


Figure 4: Evolution of Behavior in the *Baseline* Experiment by Session

the efficient equilibrium to strategic uncertainty, as measured by the size of the basin of attraction. Figure 5 shows the rate of stag as a function of the basin of attraction for periods 1, 8, and 15. As we find for the metadata, the correlation between the size of the basin and the rate of stag is positive and increases with experience in the new experiment as well. The last 3 columns in Table 10 show that the size of the basin of attraction of stag has a small effect on behavior if stag is not risk dominant, while it has a large and significant effect if stag is risk dominant. This is consistent with the findings regarding the effect of the size of the basin of attraction of Always Defect on cooperation in repeated games—see Dal Bó & Fréchette (2018).

In conclusion, from the previous analysis it is clear that neither payoff dominance nor risk dominance work as perfect equilibrium selection criteria. However, the robustness of the efficient action to strategic uncertainty (as measured by the size of the basin of attraction of stag) affects the likelihood of efficient behavior.

Table 10: *Baseline*: Determinants of Stag (Probit Analysis - Marginal Effects)

	Period 1	**Period 8**	**Period 15**	**Period 1**	**Period 8**	**Period 15**
RD (d)	0.26***	0.45***	0.58***	-0.48***	-0.93***	-0.97***
	(0.026)	(0.030)	(0.036)	(0.108)	(0.077)	(0.034)
RD × Basin				1.30***	3.33***	3.99***
				(0.212)	(0.679)	(0.480)
Not RD × Basin				0.10	0.35*	0.24
				(0.086)	(0.200)	(0.181)
Observations	1120	1120	1120	1120	1120	1120

(d) denotes dummy variable, effect of change from 0 to 1 is reported.
Standard errors clustered at the session level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As mentioned before, the calculations to determine whether stag is risk dominant and the size of the basin of attraction are done assuming that subjects are risk neutral, which may not be the case. The *Lottery* treatment allows us to study whether this reliance on the assumption of risk neutrality may be problematic.

Figure 6 displays the evolution of the rate of choices of stag for all games for the *Lottery* treatment and includes the data from the *Baseline* treatment for comparison. It is clear that both the levels and evolution of behavior are very similar in these two treatments. In period 1, behavior is significantly different at the 5% level in three out of 16 games, and there are no significant differences by period 15. Moreover, as shown in Appendix Table 14, the results on the impact of risk dominance and the size of the basin of attraction on behavior are robust to including the *Lottery* treatment with no significant differences between *Lottery* and *Baseline*. This suggests that the results from this section are not driven by the risk attitudes of the subjects.

6 Is the New Experimental Design Neutral?

The design novelty in the new experiments presented in this paper is to let subjects participate in several coordination games simultaneously in each period. This allows us to gather data on a greater number of games than it would be possible if subjects only played one game

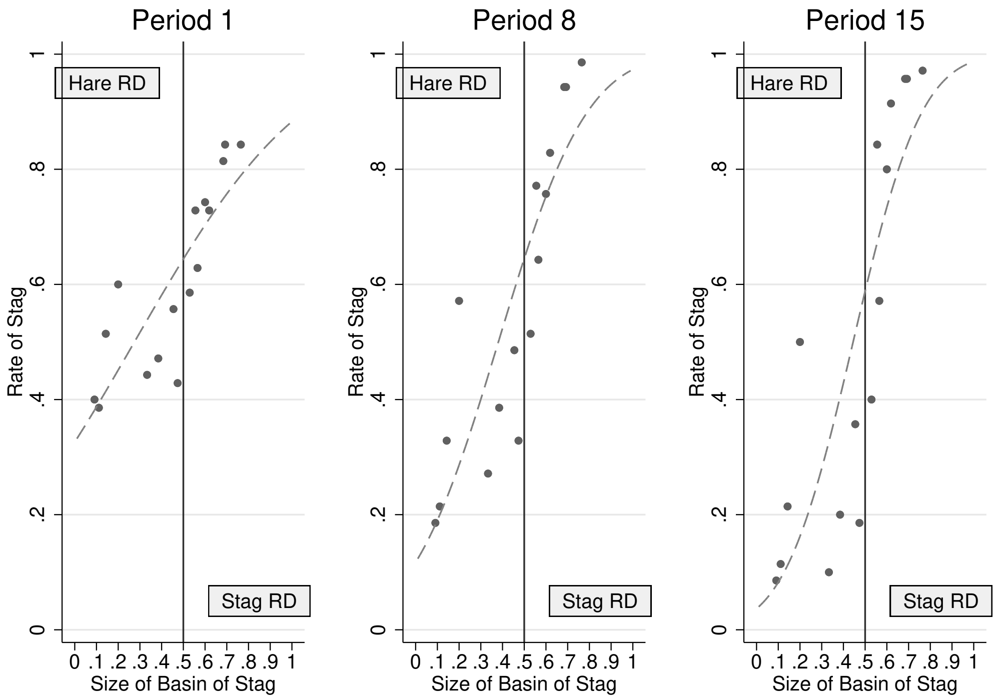
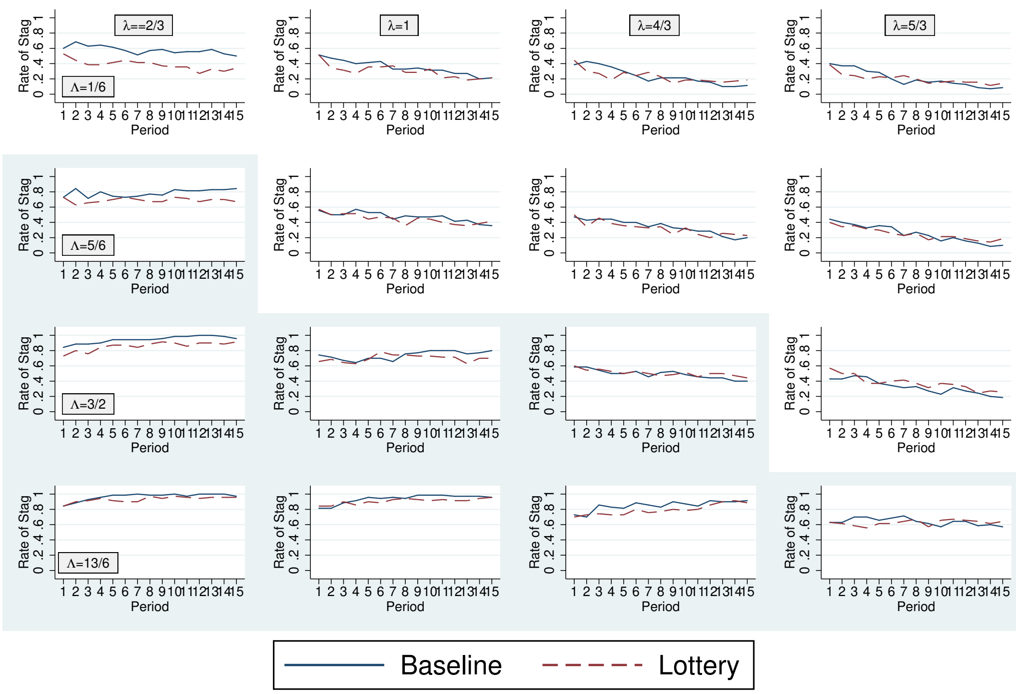


Figure 5: *Baseline* Experiment: Basin of Stag and Behavior

per period. But is this design neutral? Is it possible that behavior is affected by subjects playing several games simultaneously?

To answer these questions we present results from two additional treatments that differ from *Baseline* in that subjects play only one stag hunt game in every period. One of the treatments considers the stag hunt game with $\Lambda = \frac{3}{2}$ and $\lambda = 1$ (in the non-normalized payoffs seen by the subjects that is $T = 45$ and $S = 30$), and the other treatment considers the stag hunt game with $\Lambda = \frac{5}{6}$ and $\lambda = \frac{4}{3}$ ($T = 65$ and $S = 20$). These two treatments, *One Game* $\frac{3}{2}$*&*1 and *One Game* $\frac{5}{6}$*&*$\frac{4}{3}$ allow us to compare behavior with the same game in the *Baseline* treatment.

We conducted four experimental sessions for each of these treatments with a total of 134 subjects. See Table 13 in the Appendix for the number of subjects per session and treatment. The experimental sessions lasted less than an hour. Subjects earned \$36.20 on average, with a minimum of \$17 and a maximum of \$41.5, including the participation and the show-up fee



Note: grey background denotes Stag is risk dominant.

Figure 6: Evolution of Behavior in the *Baseline* and *Lottery* Experiments

of $10.

As can be seen in Figure 7, behavior is quite different between the *Baseline* treatment and the *One Game* treatments. For $\Lambda = \frac{3}{2}$ and $\lambda = 1$, the rate of stag is greater under *One Game* than under the *Baseline* in period 1 (but this difference is not statistically significant at the 10% level).[9] The difference increases after the first two periods and remains statistically significant at the 1% level until the end. For $\Lambda = \frac{5}{6}$ and $\lambda = \frac{4}{3}$, the rate of stag is greater under *One Game* than under the *Baseline* in period 1 (this difference is statistically significant at the 5% level). The difference increases after the first period and becomes statistically significant at the 1% level until the end. Interestingly, the results from *One Game* $\frac{5}{6}$ *&* $\frac{4}{3}$ suggest that the rate of stag can increase with experience and reach high levels even when stag is not risk dominant. This is consistent with what is found in the analysis of the metadata from previous experiments and under the *Baseline* treatment for some sessions that reach high rates of stag even when it is not risk dominant (see game with $\Lambda = \frac{1}{6}$ and

[9]All standard errors in this section are calculated clustering at the session level.

$\lambda = \frac{2}{3}$ in Figure 4).

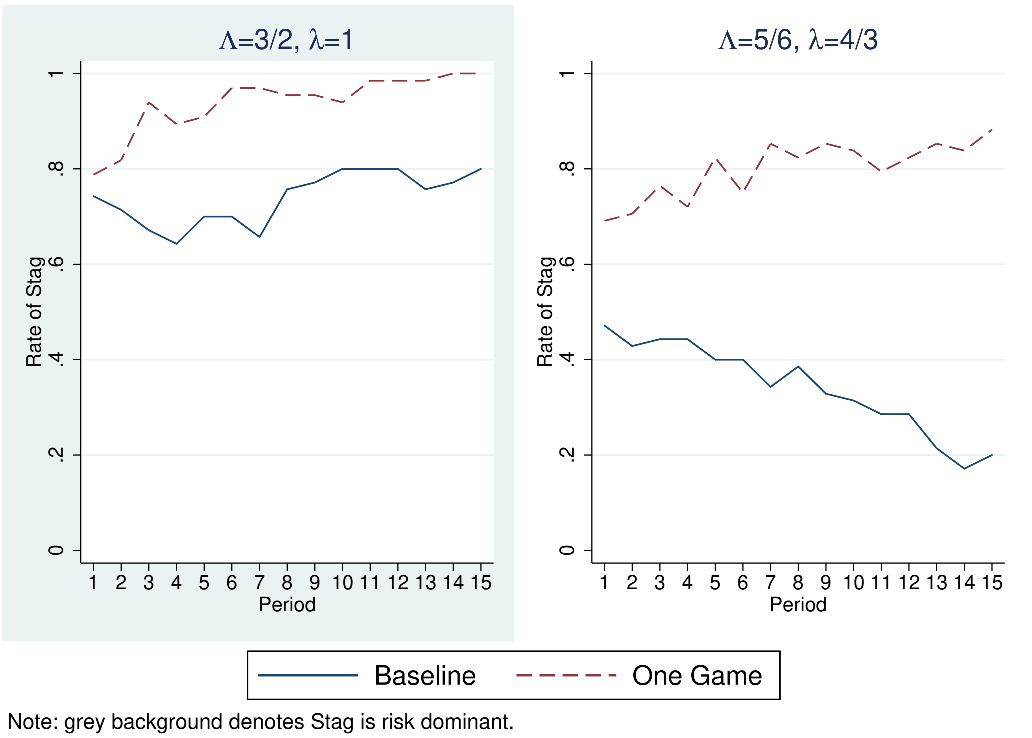


Figure 7: Evolution of Behavior in One Game Experiments

The significant difference between our *One Game* treatments and the *Baseline* is consistent with the differences in behavior between the *Baseline* and prior studies (in which subjects also played one game at a time). Figure 8 shows the relation between the rate of stag and the size of the basin of attraction of stag for all the treatments studied in this paper. Behavior in the two *One Game* treatments is consistent with the observations from the previous studies which reach high rates of stag even when it is not risk dominant. As such, the comparison across treatments and papers shows that behavior in stag hunt games may depend on whether subjects play one game in isolation or several games simultaneously. Regardless of this difference, the fact that the prevalence of stag increases with its robustness to strategic uncertainty, as measured by the size of its basin of attraction, is robust to playing in isolation or multiple games at the same time.

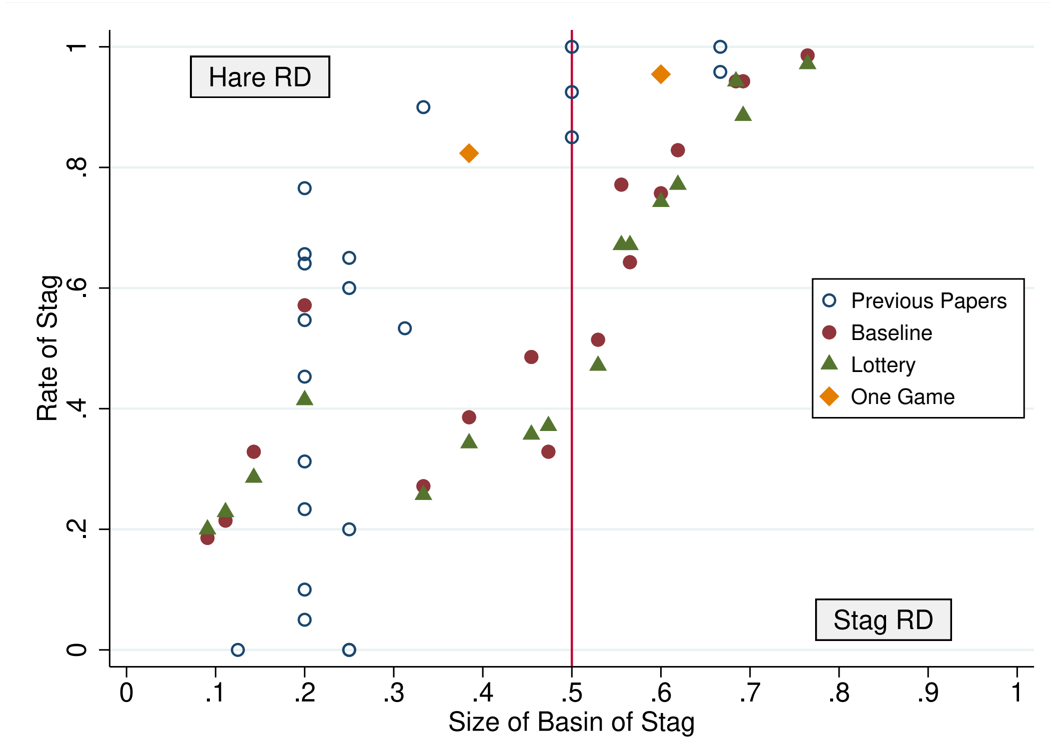


Figure 8: Basin of Attraction and Behavior over All Studies (Period 8)

7 Can Complexity and Context Explain the Non-Neutrality of the New Design?

In this section, we investigate some possible reasons for the significant differences in behavior between *One Game* treatments (including the experiments in previous articles) and the two treatments with several games per period introduced in this paper (*Baseline* and *Lottery* treatments). The intuition for one possible reason comes from the literature studying the determinants of cooperation in the infinitely repeated games experiments. Dal Bó & Fréchette (2018) conducts a meta-analysis using the data from infinitely repeated prisoner' dilemma game experiments from fifteen articles to study how cooperation depends on its robustness to strategic uncertainty. As a measure of the robustness of cooperation to strategic uncertainty they consider the size of the basin of attraction of the strategy grim against the strategy always defect. Cooperation is referred to as risk dominant if grim is risk dominant in the repeated game, see Blonski & Spagnolo (2015). A summary of the results are shown in Figure 9.

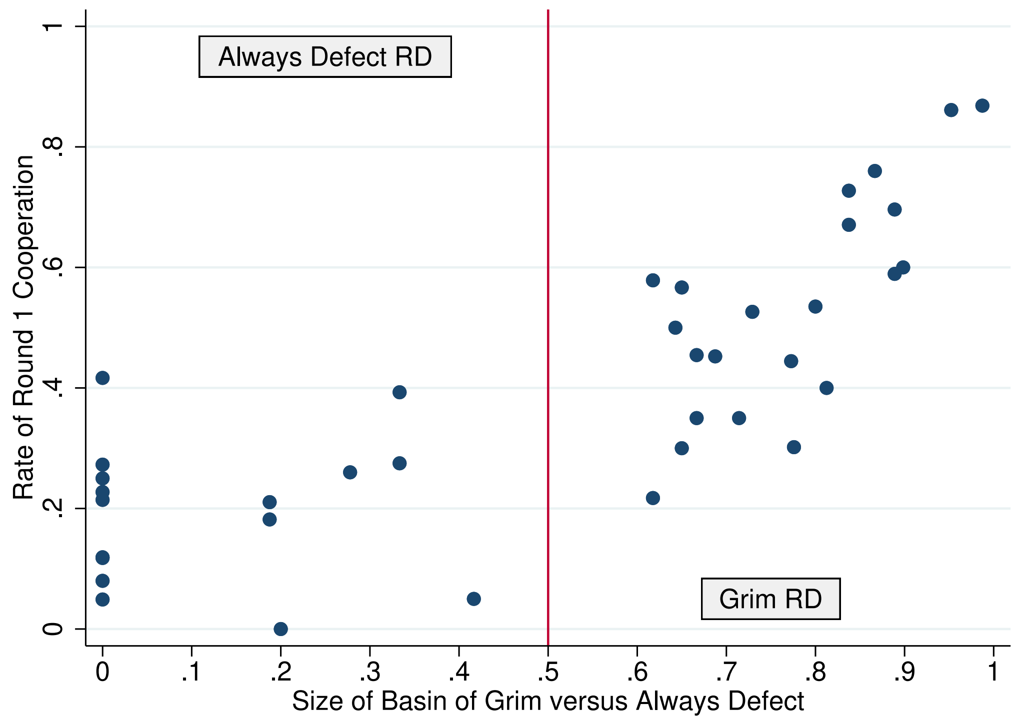


Figure 9: Basin of Attraction and Cooperation (Dal Bó and Fréchette, 2018)

Interestingly, the relation between the basin of attraction of grim and the rate of cooperation is very similar to that observed for the size of the basin of attraction of stag and the rate of stag in the *Baseline* and *Lottery* treatments: When the efficient behavior is not risk dominant, its rate is low and does not seem to depend on the size of the basin of attraction. Instead, when the efficient behavior is risk dominant, the prevalence of efficient behavior increases with the size of its basin of attraction.

One possible factor that could explain why the relation is similar for the *Baseline* and *Lottery* treatments and infinitely repeated PD games, but different for the *One Game* treatments, is that the impact of strategic uncertainty may be mediated by *complexity*. Roughly speaking, we refer to *complexity* as the amount of cognitive load that is required to make decisions—a more complex environment requires higher cognitive load. Compared to oneshot games, making decisions in an infinitely repeated game may need a greater amount of cognitive load to evaluate the tension between immediate benefits from opportunistic behavior and long-run benefits from cooperation. In the same vein, playing multiple games together requires more cognitive load as subjects process more information than what is needed for playing one game (Bednar & Page (2007), also see Proto et al. (2020) where strategy choice and cognitive load are related in infinitely repeated games).

To explore this possibility, we conduct two additional treatments using the *One Game* paradigm, but with a more complex situation in that each game has five actions and five pure-strategy equilibria. Table 11 represents the payoff matrices for a games with five actions which extends the payoff matrix used in the *One Game* $\frac{3}{2}$&1 and *One Game* $\frac{5}{6}$&$\frac{4}{3}$ treatments.

Table 11: Five Action Game

Λ=3/2,λ=1 (T=45,S=30)					
	hare	A	B	C	stag
hare	60, 60	56, 53	53, 45	49, 38	45, 30
A	53, 56	68, 68	64, 60	60, 53	56, 45
B	45, 53	60, 64	75, 75	71, 68	68, 60
C	38, 49	53, 60	68, 71	83, 83	79, 75
stag	30, 45	45, 56	60, 68	75, 79	90, 90

Λ=5/6,λ=4/3 (T=65,S=20)					
	hare	A	B	C	stag
hare	60, 60	61, 50	63, 40	64, 30	65, 20
A	50, 61	68, 68	69, 58	70, 48	71, 38
B	40, 63	58, 69	75, 75	76, 65	78, 55
C	30, 64	48, 70	65, 76	83, 83	84, 73
stag	20, 65	38, 71	55, 78	73, 84	90, 90

Our aim was to make the payoffs of the five action games as close to those of the corresponding games with 2 actions. As presented in Table 11, hare and stag are placed in each corner of the table so that the salience of hare and stag is affected by the introduction of other actions in a minimal manner.[10] For these additional treatments, *Five Actions* $\frac{3}{2}$&1 and *Five Actions* $\frac{5}{6}$&$\frac{4}{3}$, the experimental design differs from that of the *One Game* treatments only in the different payoff matrices.

We conducted four experimental sessions for each of these two additional treatments with a total of 134 subjects. See Table 13 in the Appendix for the number of subjects per session and treatment. The experimental sessions lasted less than an hour. Subjects earned \$38.56 on average, with a minimum of \$17 and a maximum of \$41.5, including the participation and the show-up fee of \$10.

Figure 10 shows the evolution of behavior in the *One Game* and *Five actions* treatments

[10]The actions were simply described as "1" to "5" in the experiment.

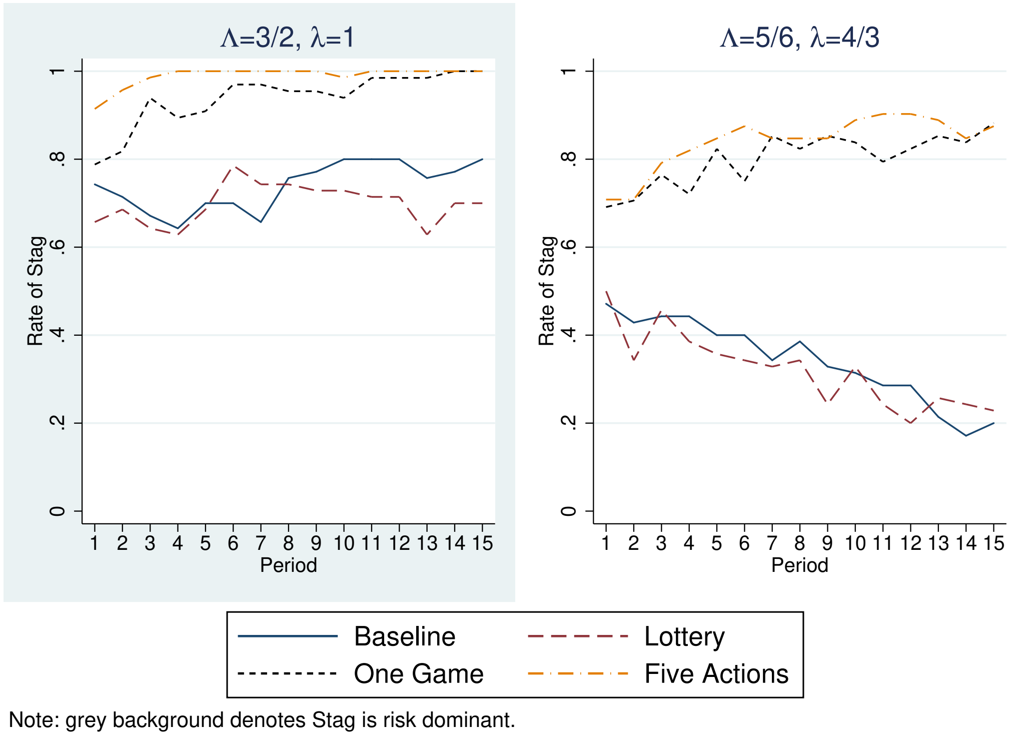


Figure 10: Evolution of Behavior in One Game and Five Actions Treatments

and the corresponding games in the *Baseline* and *Lottery* treatments. For the games with five actions, the rate of stag is computed by the relative frequency over hare and stag.[11] It is clearly the case that the additional treatments with five actions reveal a pattern of behavior that is similar to the one observed for the *One Game* treatments with two actions. Hence, either the impact of strategic uncertainty may not be mediated by *complexity*, or what affects *complexity* may not be captured by these last two additional treatments with greater number of actions and equilibria.

Another possible factor that may affect coordination is *context*. We refer to *context* as the setting and circumstances in which the game under study is being played. For example, part of the context of a game may be the other games that the subject is playing during the same experiment. Cooper & Kagel (2008), Cooper & Kagel (2009), and Rick & Weber (2010) study sequential spillover effects, or order effects, and Bednar & Page (2007), Huck

[11]Only a minority of subjects chose actions besides stag and hare. The distribution of actions is shown in Figure 14 in the Appendix.

Table 12: Effect of Neighbors' Size of Basin of Attraction on Stag (Marginal Effects from Probit - *Baseline* and *Lottery* Treatments)

	Period 1	**Period 15**	**Period 1**	**Period 15**
RD (d)	-0.40***	-0.93***	-0.40***	-0.94***
	(0.061)	(0.104)	(0.063)	(0.088)
RD × Basin	0.98***	3.14***	1.08***	3.57***
	(0.107)	(1.074)	(0.120)	(0.988)
Not RD × Basin	0.04	-0.00	0.13*	0.27**
	(0.089)	(0.190)	(0.067)	(0.119)
4 Neighbors' Basin	0.25***	0.58		
	(0.072)	(0.488)		
8 Neighbors' Basin			0.16**	0.15
			(0.072)	(0.408)
Observations	2240	2240	2240	2240

Marginal effects; Clustered at session level standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

et al. (2011), and Bednar et al. (2012) study simultaneous spillover effects. That literature, and our results, suggest that it may not be enough to explain behavior in a coordination game only as a function of the payoff parameters of that game. The context in which people interact may matter as well.

We provide evidence that context matters in the *Baseline* and *Lottery* treatments by using the characteristics, in terms of the size of the basin of attraction, of neighboring games on the screen with 16 games to explain behavior. If there are spillovers across neighboring games, we should expect that the prevalence of stag should increase in the size of the basin of attraction of stag in neighboring games. We construct two measures of the characteristics of the neighbors. First, we focus on the four neighbors to the left and right and to the top and bottom of each treatment and calculate their average size of the basin of attraction of stag.[12] Second, we focus on the eight neighbors surrounding the game in consideration.[13]

As shown in Table 12, regardless of the definition of neighbors, it is the case that the larger the average size of the basin of attraction of stag for the neighbors, the greater the

[12]For games on the border of the combinations of S and T, this may consist of the average of only two or three numbers.

[13]For games on the border of the combinations of S and T, this may consist of the average of only three or five numbers.

share of subjects choosing stag. The effect is statistically significant for the first period but not for the last one. Note, however, that the magnitude of the effect does not decreases significantly as subjects gain experience.

Hence, the results presented in this section suggests that one aspect of complexity, the number of choices and equilibria, is unlikely to be the driving force for the difference between our *Baseline* and *One Game* treatments. Our observation that the behavior in one coordination game is affected by the characteristics of neighboring games suggests, on the other hand, that there are spillovers across games. These spillovers across neighbors strengthen the idea that behavior in coordination games depends on other elements beyond the payoff structure of that coordination game considered in isolation.

8 Conclusions

We use the metadata from previous experiments and data from a new experiment to study the determinants of efficient behavior in stag hunt games. We find that the prevalence of stag is not simply determined by risk dominance or payoff dominance. The failure of these equilibrium selection criteria to explain behavior cannot be attributed to misspecified risk attitudes, as the results are robust to an implementation that induces risk neutral preferences.

While risk dominance does not perfectly explain behavior, we do find that the prevalence of stag increases with the robustness of the efficient equilibrium to strategic uncertainty: as the basin of stag increases, the rate of stag tends to increase. However, the exact relation depends on the number of games subjects play in a given period. For intermediate values of the size of the basin of attraction of stag, the rate of stag is higher when subjects play only one game per period. These are the values of the basin of attraction of stag for which prior experiments had found more variable results (across papers and authors). This suggests that behavior in coordination games, at least for certain levels of strategic risk, depends not only on the payoff parameters of the game but also on the context.

Although it may seem sensible that behavior depends on context in games with multiple equilibria such as coordination games. This comes against the backdrop of surprisingly strong regularities: the fact that the behavior in our *One Game* treatments and prior experiments display very similar patterns, and that the *Baseline* and *Lottery* reveal almost identical result. This shows that, even in coordination games, there are aspects of the implementation of the game that do not affect behavior. Furthermore, our findings reveal a strong and stable qualitative relationship: as strategic risk increases, as measured by the size of the basin of attraction, subjects become less likely to chose the efficient action, independent of the details of the implementation. It remains for future work to study the dimensions of context that affect behavior in coordination games.

References

Battalio, R., Samuelson, L. & Van Huyck, J. (2001), 'Optimization incentives and coordination failure in laboratory stag hunt games', *Econometrica* **69**(3), 749–764.

Bednar, J., Chen, Y., Liu, T. X. & Page, P. (2012), 'Behavioral spillovers and cognitive load in multiple games: An experimental study', *Games and Economic Behavior* **74**(1), 12–31.

Bednar, J. & Page, S. E. (2007), 'Can game(s) theory explain culture?: The emergence of cultural behavior in multiple games', *Rationality and Society* **19**(1), 65–97.

Blonski, M. & Spagnolo, G. (2015), 'Prisoners' other dilemma', *International Journal of Game Theory* **44**(1), 61–81.

Calford, E. & Oprea, R. (2017), 'Continuity, inertia, and strategic uncertainty: A test of the theory of continuous time games', *Econometrica* **85**(3), 915–935.

Castillo, M. & Dianat, A. (2018), Strategic uncertainty and equilibrium selection in stable matching mechanisms: Experimental evidence. Working Paper.

Charness, G. (2000), 'Self-serving cheap talk: A test of aumann's conjecture', *Games and Economic Behavior* **33**(2), 177–194.

Clark, K., Kay, S. & Sefton, M. (2001), 'When are nash equilibria self-enforcing? an experimental analysis', *International Journal of Game Theory* **29**(4), 495–515.

Cooper, D. J. & Kagel, J. H. (2008), 'Learning and transfer in signaling games', *Economic Theory* **34**(3), 415–439.

Cooper, D. J. & Kagel, J. H. (2009), 'The role of context and team play in cross-game learning', *Journal of the European Economic Association* **7**(5), 1101–1139.

Cooper, R. W., DeJong, D. V., Forsythe, R. & Ross, T. W. (1990), 'Selection criteria in coordination games: Some experimental results', *American Economic Association* **80**(1), 218–233.

Cooper, R. W., DeJong, D. V., Forsythe, R. & Ross, T. W. (1992), 'Communication in coordination games', *The Quarterly Journal of Economics* **107**(2), 739–771.

Dal Bó, P. & Fréchette, G. R. (2011), 'The evolution of cooperation in infinitely repeated games: Experimental evidence', *The American Economic Review* **101**(1), 411–429.

Dal Bó, P. & Fréchette, G. R. (2018), 'On the determinants of cooperation in infinitely repeated games: A survey', *Journal of Economic Literature* **56**(1), 60–114.

Dubois, D., Willinger, M. & Van Nguyen, P. (2012), 'Optimization incentive and relative riskiness in experimental stag-hunt games', *International Journal of Game Theory* **41**(2), 369–380.

Duffy, J. & Feltovich, N. (2002), 'Do actions speak louder than words? an experimental comparison of observation and cheap talk', *Games and Economic Behavior* **39**(1), 1–27.

Embrey, M., Fréchette, G. R. & Yuksel, S. (2017), 'Cooperation in the finitely repeated prisoner's dilemma', *The Quarterly Journal of Economics* **133**(1), 509–551.

Feltovich, N., Iwasaki, A. & Oda, S. H. (2012), 'Payoff levels, loss avoidance, and equilibrium selection in games with multiple equilibria: an experimental study', *Economic Inquiry* **50**(4), 932–952.

Fischbacher, U. (2007), 'z-tree: Zurich toolbox for ready-made economic experiments', *Experimental Economics* **10**(2), 171–178.

Harsanyi, J. C. & Selten, R. (1988), *A General Theory of Equilibrium Selection in Games*, MIT Press, Cambridge, MA.

Healy, P. J. (2016), Epistemic experiments: Utilities, beliefs, and irrational play. Working Paper.

Huck, S., Jehiel, P. & Rutter, T. (2011), 'Feedback spillover and analogy-based expectations: A multi-game experiment', *Games and Economic Behavior* **71**(2), 351–365.

Kandori, M., Mailath, G. J. & Rob, R. (1993), 'Learning, mutation, and long run equilibria in games', *Econometrica* **61**(1), 29–56.

Kartal, M. & Müller, W. (2018), A new approach to the analysis of cooperation under the shadow of the future: Theory and experimental evidence. Working Paper.

Luce, R. D. & Raiffa, H. (1957), *Games and Decisions: Introduction and critical survey*, John Wiley & Sons, New York, NY.

Proto, E., Rustichini, A. & Sofianos, A. (2020), Intelligence, errors and strategic choices in the repeated prisoners' dilemma. Working Paper.

Rick, S. & Weber, R. A. (2010), 'Meaningful learning and transfer of learning in games played repeatedly without feedback', *Games and Economic Behavior 68 (2010)* **68**(2), 716–730.

Roth, A. E. & Malouf, M. W. (1979), 'Game-theoretic models and the role of information in bargaining', *Psychological Review* **86**(6), 574–594.

Schelling, T. C. (1960), *The Strategy of Conflict*, Harvard University Press, Cambridge, MA.

Schmidt, D., Shupp, R., Walker, J. M. & Ostrom, E. (2003), 'Playing safe in coordination games:: the roles of risk dominance, payoff dominance, and history of play', *Games and Economic Behavior* **42**(2), 281–299.

Straub, P. G. (1995), 'Risk dominance and coordination failures in static games', *The Quarterly Review of Economics and Finance* **35**(4), 339–363.

Van Huyck, J. B., Battalio, R. C. & Beil, R. O. (1990), 'Tacit coordination games, strategic uncertainty, and coordination failure', *The American Economic Review* **80**(1), 234–248.

Vespa, E. & Wilson, A. J. (2017), Experimenting with equilibrium selection in dynamic games. Working Paper.

Young, H. P. (1993), 'The evolution of conventions', *Econometrica* **61**(1), 57–84.

9 Appendix

9.1 Additional Tables and Figures

Table 13: Number of subjects per session and treatment

Experiment	Session				Total
	1	2	3	4	
Baseline	22	16	16	16	70
Lottery	16	16	20	18	70
One Game $\frac{3}{2}$&1	16	16	18	16	66
One Game $\frac{5}{6}$&$\frac{4}{3}$	16	16	20	16	68
Five Actions $\frac{3}{2}$&1	18	18	18	16	70
Five Actions $\frac{5}{6}$&$\frac{4}{3}$	16	16	20	20	72

Table 14: *Baseline* and *Lottery*: Determinants of Stag

	Period 1	**Period 8**	**Period 15**	**Period 1**	**Period 8**	**Period 15**
RD (d)	0.25***	0.46***	0.55***	-0.43***	-0.92***	-0.95***
	(0.029)	(0.036)	(0.046)	(0.067)	(0.078)	(0.070)
RD × Basin				1.21***	3.22***	3.70***
				(0.137)	(0.593)	(0.747)
Not RD × Basin				0.18***	0.34**	0.32*
				(0.062)	(0.139)	(0.171)
Lottery (d)	-0.01	-0.04	-0.00	-0.01	-0.05	-0.00
	(0.048)	(0.117)	(0.123)	(0.048)	(0.123)	(0.130)
Observations	2240	2240	2240	2240	2240	2240

(d) denotes dummy variable, effect of change from 0 to 1 is reported.
Clustered at session level standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

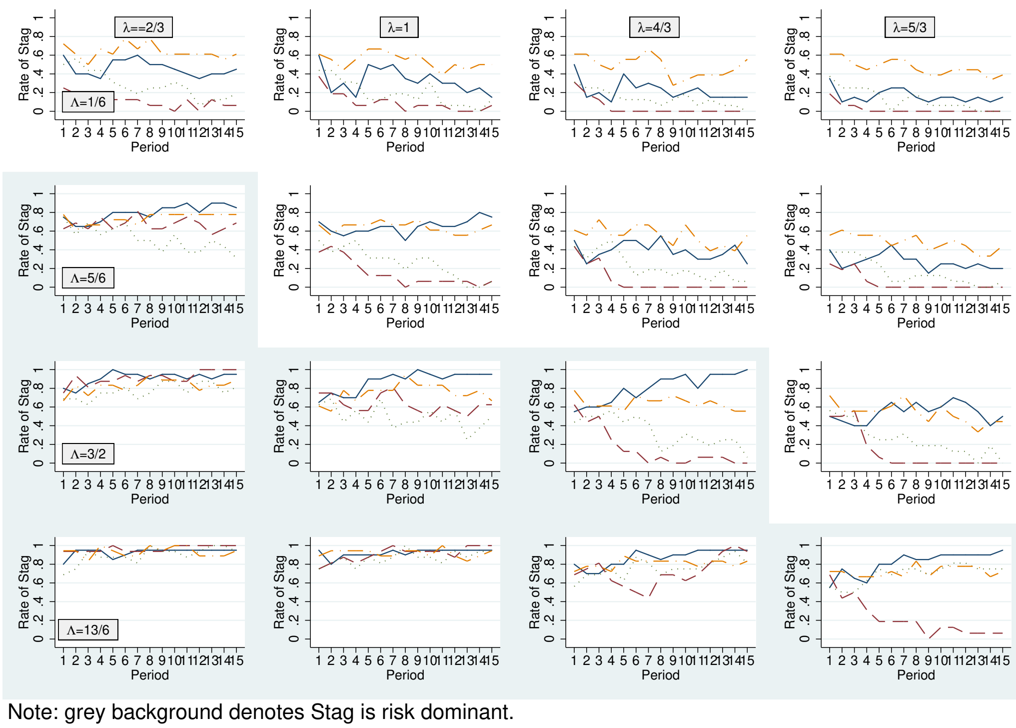


Figure 11: Evolution of Behavior in *Lottery* Treatment by Session

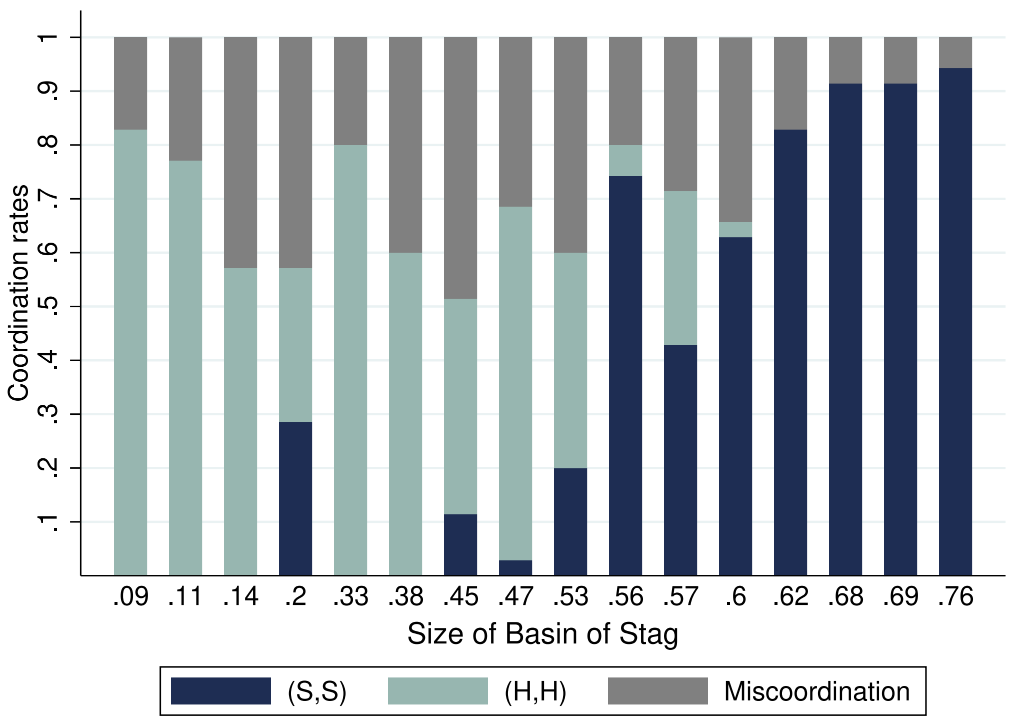


Figure 12: Coordination and Basin of Stag (Period 15)

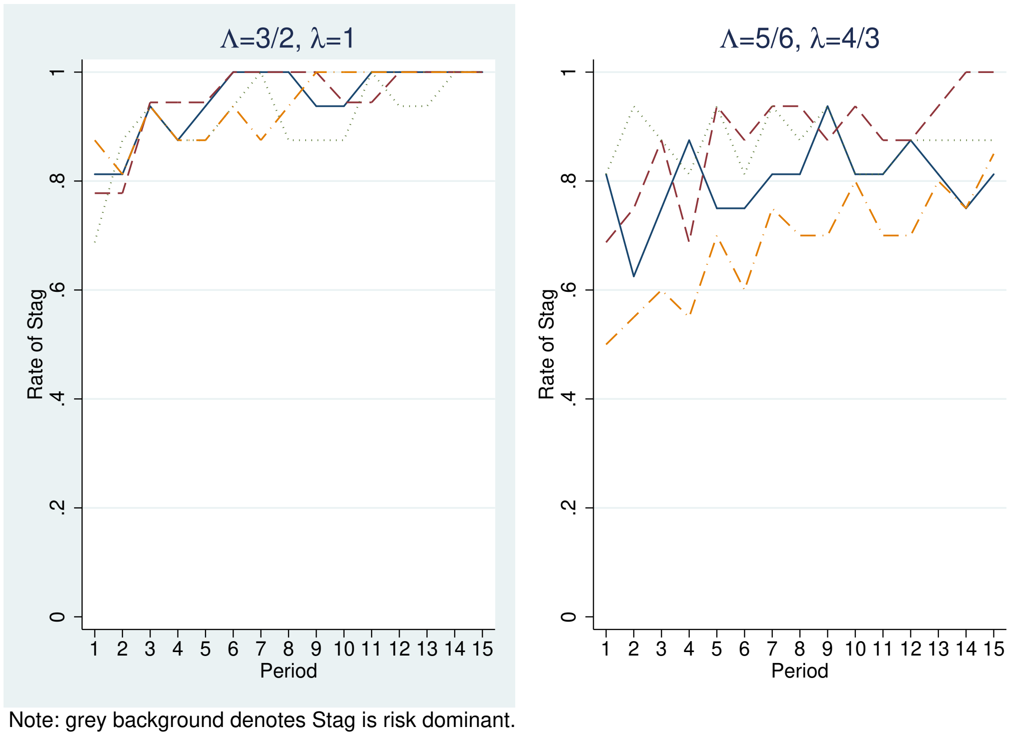


Figure 13: Evolution of Behavior in *One Game* Treatments by Session

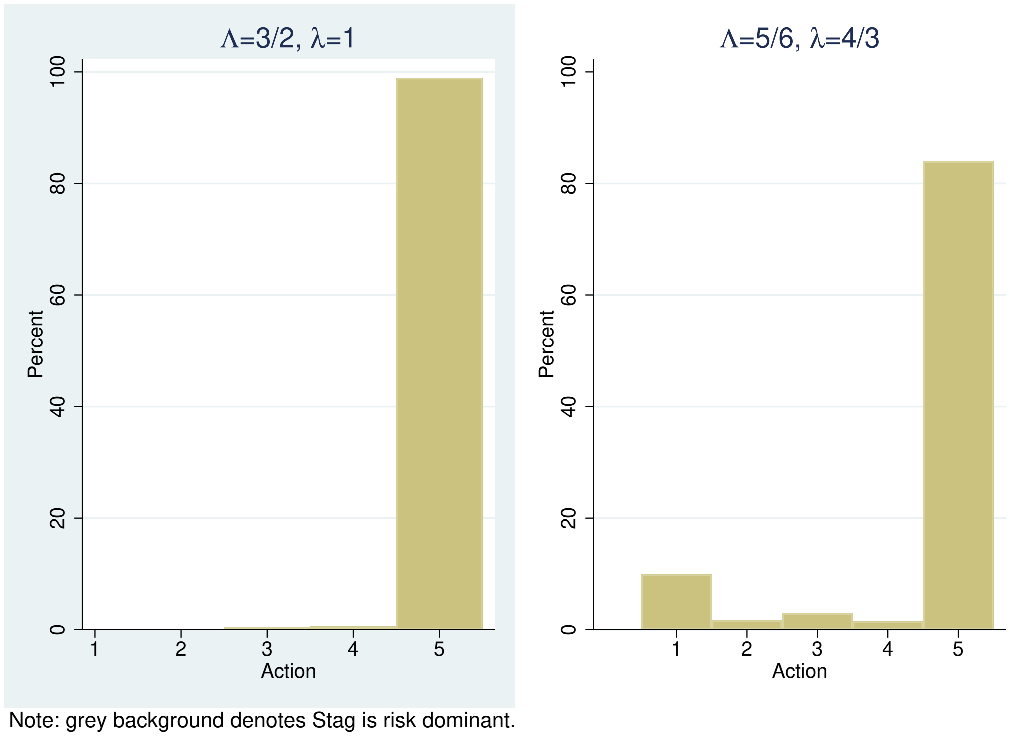


Figure 14: Distribution of Actions in *Five Actions* Treatment (all periods)

9.2 Instructions

9.2.1 Baseline Treatment (No lottery and 16 games)

Instructions

Welcome

You are about to participate in a session on decision-making, and you will be paid for your participation with cash, privately at the end of the session. What you earn depends partly on your decisions, partly on the decisions of others, and partly on chance.

Please turn off cellular phones now.

The entire session will take place through computer terminals, and all interaction between the participants will take place through the computers. Please do not talk or in any way try to communicate with other participants during the session.

We will start with a brief instruction period. During the instruction period you will be given a description of the main features of the session. If you have any questions during this period, raise your hand. Your question will then be answered so everyone can hear.

General Instructions

1. In this experiment, you will be asked to make decisions in each of 15 periods. You will be randomly paired with another person for each period. No pair of participants will interact together more than once.

2. In each period, you will be asked to make one decision in each of 16 different environments. That is, you will be asked to make 16 decisions in each period. For each environment, you will be asked to choose either action 1 or action 2. As an example, the choices and the points you may earn in one environment are as follows:

	The other's choice	
Your choice	1	2
1	$60, 60$	$65, 30$
2	$30, 65$	$90, 90$

The first entry in each cell represents your points, while the second entry represents the points of the person with whom you are matched. That is, in this particular environment, if:

You select 1 and the other selects 1, you each make 60 points.

You select 1 and the other selects 2, you make 65 points while the other makes 30 points.

You select 2 and the other selects 1, you make 30 points while the other makes 65 points.

You select 2 and the other selects 2, you each make 90 points.

Note that, within a period, you are paired with the same person for all 16 environments. These environments will differ in points you may earn.

3. Once the second period begins, and for every period after that, you can see the history of your decisions and the decisions of the participants that were paired with you by clicking the "Feedback" button.

Payment

1. At the end of the experiment, one environment in one period will be randomly selected for payment. Your payment consists of two parts. You will receive a $5 participation fee. On top of this, your earned points will be converted into dollars with the exchange rate of 0.35, that is, 100 points are worth $35.

2. In addition, you will receive a $5 show-up fee.

- Are there any questions?

Before we start, let me remind you that:

- There are 15 periods in each of which you will be asked to make decisions in 16 different environments.

- Within a period you will be paired with the same person for all 16 environments.

- You will be paired with a different person in every period.

- Only one environment in one period will be randomly selected for payment.

9.2.2 Lottery Treatment

Instructions

Welcome

You are about to participate in a session on decision-making, and you will be paid for your participation with cash, privately at the end of the session. What you earn depends partly on your decisions, partly on the decisions of others, and partly on chance.

Please turn off cellular phones now.

The entire session will take place through computer terminals, and all interaction between the participants will take place through the computers. Please do not talk or in any way try to communicate with other participants during the session.

We will start with a brief instruction period. During the instruction period you will be given a description of the main features of the session. If you have any questions during this period, raise your hand. Your question will then be answered so everyone can hear.

General Instructions

1. In this experiment, you will be asked to make decisions in each of 15 periods. You will be randomly paired with another person for each period. No pair of participants will interact together more than once.

2. In each period, you will be asked to make one decision in each of 16 different environments. That is, you will be asked to make 16 decisions in each period. For each environment, you will be asked to choose either action 1 or action 2. As an example, the choices and the points you may earn in one environment are as follows:

	The other's choice	
Your choice	1	2
1	$60, 60$	$65, 30$
2	$30, 65$	$90, 90$

The first entry in each cell represents your points, while the second entry represents the points of the person with whom you are matched.

That is, in this particular environment, if:

You select 1 and the other selects 1, you each make 60 points.

You select 1 and the other selects 2, you make 65 points while the other makes 30 points.

You select 2 and the other selects 1, you make 30 points while the other makes 65 points.

You select 2 and the other selects 2, you each make 90 points.

Note that, within a period, you are paired with the same person for all 16 environments. These environments will differ in points you may earn.

3. Once the second period begins, and for every period after that, you can see the history

of your decisions and the decisions of the participants that were paired with you by clicking the "Feedback" button.

Payment

1. At the end of the experiment, one environment in one period will be randomly selected for payment. Your payment consists of two parts. You will receive a $5 participation fee. On top of this, your earned points will be converted into dollars depending on the draw of a random number between 1 and 100. If the randomly chosen number is less than or equal to your points, you will earn $35; otherwise, you will earn $0. That is, if you earned X points, then you will earn $35 with X percentage chance, and $0 with 100-X percentage chance.

2. In addition, you will receive a $5 show-up fee.

- Are there any questions?

Before we start, let me remind you that:

- There are 15 periods in each of which you will be asked to make decisions in 16 different environments.

- Within a period you will be paired with the same person for all 16 environments.

- You will be paired with a different person in every period.

- Only one environment in one period will be randomly selected for payment.

9.2.3 One game per session (T=45, S=30)

Instructions

Welcome

You are about to participate in a session on decision-making, and you will be paid for your participation with cash, privately at the end of the session. What you earn depends partly on your decisions, partly on the decisions of others, and partly on chance.

Please turn off cellular phones now.

The entire session will take place through computer terminals, and all interaction between the participants will take place through the computers. Please do not talk or in any way try to communicate with other participants during the session.

We will start with a brief instruction period. During the instruction period you will be given a description of the main features of the session. If you have any questions during this period, raise your hand. Your question will then be answered so everyone can hear.

General Instructions

1. In this experiment, you will be asked to make decisions in each of 15 periods. You will be randomly paired with another person for each period. No pair of participants will interact together more than once.

2. In each period you will be asked to make one decision, choosing either action 1 or action 2. The choices and the points you may earn in each period are as follows:

	The other's choice	
Your choice	1	2
1	$60, 60$	$45, 30$
2	$30, 45$	$90, 90$

The first entry in each cell represents your points, while the second entry represents the points of the person with whom you are matched.

That is, if:

You select 1 and the other selects 1, you each make 60 points.

You select 1 and the other selects 2, you make 45 points while the other makes 30 points.

You select 2 and the other selects 1, you make 30 points while the other makes 45 points.

You select 2 and the other selects 2, you each make 90 points.

3. Once the second period begins, and for every period after that, you can see the history of your decisions and the decisions of the participants that were paired with you by clicking the "Feedback" button.

Payment

1. At the end of the experiment, one period will be randomly selected for payment. Your payment consists of two parts. You will receive a $5 participation fee. On top of this, your earned points will be converted into dollars with the exchange rate of 0.35, that is, 100 points are worth $35.

2. In addition, you will receive a $5 show-up fee.

- Are there any questions?

Before we start, let me remind you that:

- There are 15 periods in each of which you will be asked to make decisions.

- You will be paired with a different person in every period

- Only one period will be randomly selected for payment.

9.3 Screen Shoots

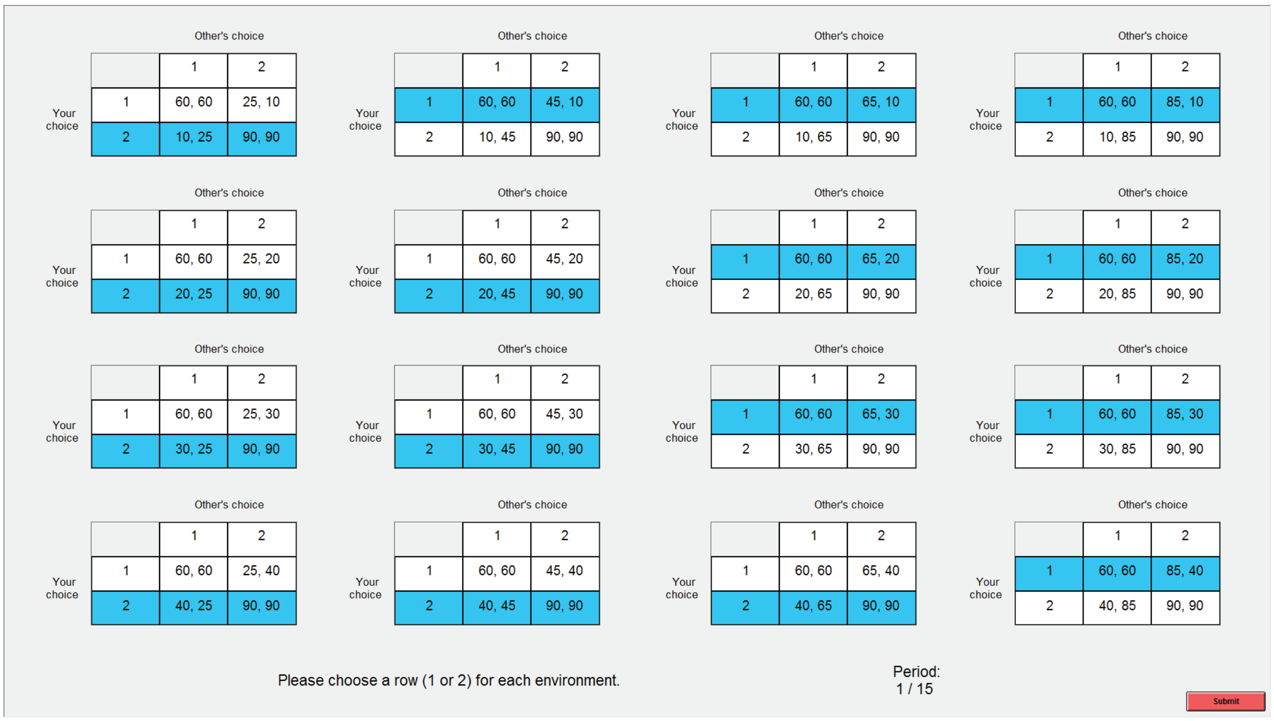


Figure 15: Decision Screen in *Baseline* and *Lottery* Treatments

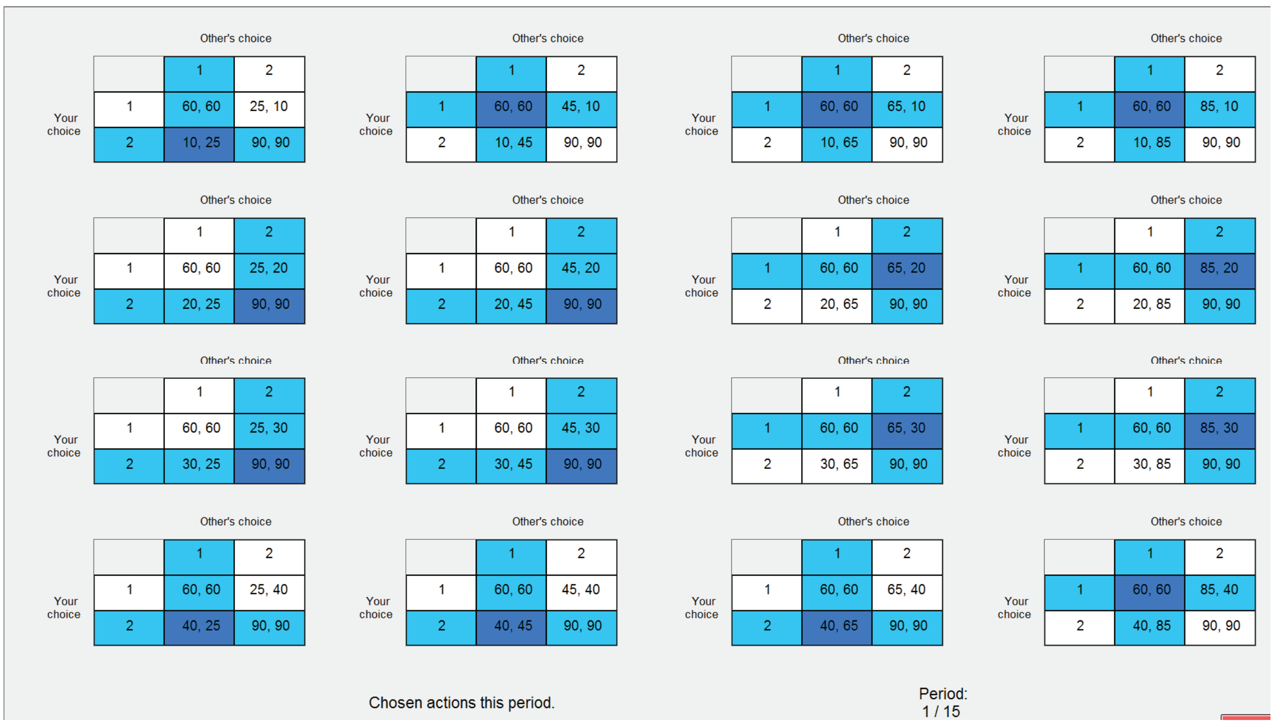


Figure 16: Period Feedback Screen in *Baseline* and *Lottery* Treatments

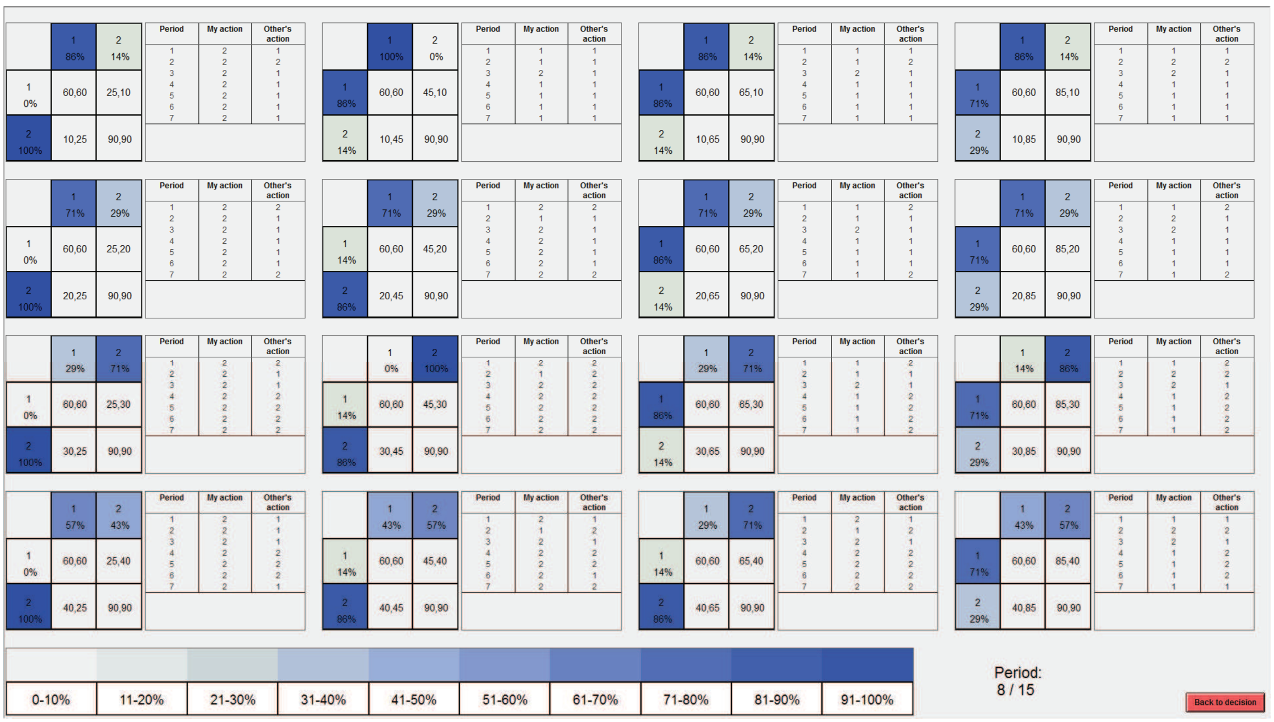


Figure 17: Previous Periods Feedback Screen in *Baseline* and *Lottery* Treatments