# Sustaining Coordination in a Dynamic Stag-Hunt Game*

Yuting Chen †

This version: November 14, 2026

**Abstract**

Many collective goals are secured only after sustained coordination, and a single failure can erase accumulated progress. In my dynamic stag-hunt game, consecutive successful coordination advances a public state toward a protection threshold; any failure resets it, and an outside shock before the threshold imposes a loss. I characterize Markov Perfect Equilibria and develop a state-dependent dynamic basin of attraction measuring how much strategic uncertainty agents can tolerate while coordination remains optimal. In an experiment varying group size and the safe payoff and eliciting choice probabilities, subjects choose the risky action more as groups approach the threshold, and coordination persists beyond it. A lower safe payoff increases coordination; smaller groups coordinate more only when the safe payoff is high. After a failure caused by others, subjects coordinate less, especially in four-person groups facing a high safe payoff. Individual mixing is common initially and fades with progress and experience.

**Keywords:** *Coordination, Markov Perfect Equilibrium, Dynamic Basin of Attraction*

**JEL Classification:** *C73, C92, D91*

*I thank my advisors, Prof. Timothy Cason and Prof. David Gill, for encouragement and research advice; workshop presentation participants for valuable comments and discussion; the Vernon Smith Experimental Economics Laboratory at Purdue University for providing the experimental equipment; and the Daniels School of Business for providing internal funding. The experimental design was preregistered at AsPredicted #308239: `https://aspredicted.org/5e24da.pdf`.

†Purdue University, Department of Economics, chen4442@purdue.edu

# 1 Introduction

Many coordination problems are not resolved in a single interaction. A collective goal is often secured only after a group coordinates successfully over a sustained period: each success brings the group closer to the goal, while a single breakdown can erase the accumulated progress. Communities fighting an infectious disease, for example, can eliminate it only if transmission is halted everywhere for a sustained period, and a lapse in any one community can restart the effort. Teams adopting a new practice, such as a safety protocol or a new technology, benefit only if every member follows it, and an early lapse can send the team back to old habits before the practice becomes routine. Minority communities seeking to preserve their language or culture face a similar challenge: preservation requires costly effort from many members, assimilation offers a safer alternative, and progress can be lost when participation breaks down.[1] In each case, the effort is also exposed to outside shocks, such as a new importation of the disease, a change in leadership, or a wave of out-migration, that can end it before the goal is secured. The continuation value of doing one's part today therefore depends on how much progress the group has already made.

The stag hunt is the canonical model of such coordination problems. Each player chooses between a risky action, which yields the highest payoff if others also choose it but the lowest payoff if they do not, and a safe action, which yields a lower payoff that is less sensitive to others' choices. Both universal risky actions and universal safe actions are equilibria, and which one players select depends on how much strategic uncertainty, that is, uncertainty about whether others will do their part, they can tolerate (Harsanyi and Selten, 1988; Cooper et al., 1990; Van Huyck et al., 1990). A standard measure of this tolerance is the size of the *basin of attraction* of the risky action: the largest probability that the other player chooses the safe action at which the risky action remains a best response. The concept originates in evolutionary game theory (Kandori et al., 1993; Young, 1993; Ellison, 1993) and was brought to experimental economics by Dal Bó and Fréchette (2011) in repeated prisoner's dilemmas; Dal Bó et al. (2021) show that it predicts efficient play in one-shot stag hunt games.[2] Existing basin measures, however, assign a single number to each game, computed either for a one-shot game or for a repeated game whose stage game never changes.

In the coordination problems above, the incentive to coordinate is not fixed. Because each success moves the group closer to its goal, the amount of strategic uncertainty a member

[1] Related economic models have studied incentives for minority assimilation and language acquisition, network externalities in language use, and the intergenerational transmission of cultural traits; see Church and King (1993); Lazear (1999); Bisin and Verdier (2000).

[2] Related experimental work studies optimization incentives, risk and payoff dominance, relative riskiness, and history dependence in coordination games (Battalio et al., 2001; Schmidt et al., 2003; Dubois et al., 2012; Kendall, 2022; Jagau, 2024; Silva, 2024).

can tolerate while still doing their part changes as progress accumulates, and a single static measure cannot describe how the robustness of coordination evolves over the course of a group's effort. The dynamic setting also raises a question that one-shot games cannot pose. Because the same level of progress can be reached through different histories, players may condition their choices only on current progress (state-dependent behavior) or also on how the group got there, for example by giving up after a failure (history-dependent behavior). This paper asks three questions. First, does coordination become more robust as the group gets closer to protection? Second, after a failure resets progress, do groups simply start over, or does the failure itself lower coordination? Third, is coordination less likely when the safe action is more attractive and when the group is larger?

I study these questions in a dynamic stag-hunt game with three key features. First, coordination in each round is fragile: an agent who chooses the risky action $X$ receives the benefit only if all other group members also choose $X$, whereas the safe action $Y$ guarantees a lower payoff regardless of the group outcome. Second, progress accumulates but can be lost: a publicly observed state $k$ counts consecutive all-$X$ rounds, in which every group member chooses $X$, and a single $Y$ choice resets $k$ to zero. Third, progress matters because of an outside shock: at the end of each round, the game ends with a constant probability, and if it ends before $k$ reaches a threshold, every member's final-round payoff is overwritten to a loss. Once $k$ reaches the threshold, the group is protected. In terms of the examples above, $k$ measures how long the group has sustained its effort, a reset is a lapse, and the shock is an outside disruption. Because the stage payoffs are the same in every state, the state affects incentives only through continuation values: the closer the group is to the threshold, the fewer all-$X$ rounds it still needs to secure protection, and the more progress a failure destroys.

I analyze the game in two complementary ways. First, I characterize its Markov Perfect Equilibria (MPE), in which each agent's action depends only on the current public state $k$, independent of the history (Maskin and Tirole, 1988a,b, 2001; Dutta, 1995). Every symmetric pure Markov strategy is an MPE, so the game has many equilibria. I focus on *monotone* equilibria, in which the probability of choosing $X$ is weakly increasing in $k$. Along the path of play, the monotone pure MPEs predict either full coordination on $X$ or permanent coordination on $Y$. For thresholds of one to three rounds, the unique on-path monotone mixed MPE, which serves as the benchmark for the experiment, prescribes mixing at $k = 0$ and choosing $X$ for sure once the group has achieved one all-$X$ round, so strategic uncertainty is concentrated at the initial state. As is typical of mixed equilibria in coordination games, however, its mixing probability has counterintuitive comparative statics: it rises with the safe payoff and with group size.

Second, to obtain predictions on how the robustness of coordination varies with progress and with the parameters of the game, I define a *dynamic basin of attraction*. Following Dal Bó and Fréchette (2011), I compare Grim Trigger, under which an agent chooses $X$ as long as every previous round has been all-$X$ and switches permanently to $Y$ after the first round that is not, with All-$Y$. Because coordination requires every other group member to choose $X$, I measure the basin per opponent (Boczoń et al., 2025): the dynamic basin at state $k$ is the largest probability that each other member plays All-$Y$ such that Grim Trigger remains a best response. The dynamic basin has three properties that the experiment tests. (1) The basin expands as the group makes progress, but the expansion stops once protection is one round away; (2) It is larger when the safe action pays less. (3) It is larger when the group is smaller. At the initial state, the dynamic basin coincides with the belief cutoff implied by the monotone mixed MPE, so the Markov and Grim Trigger approaches deliver the same measure of strategic uncertainty at the state where that uncertainty is greatest. Finally, the two approaches yield a direct test of state versus history dependence: under a Markov strategy, an agent's behavior at $k = 0$ should be the same whether the game has just started or the state has just been reset, whereas Grim Trigger prescribes a permanent switch to $Y$ after a reset.

I test these predictions in a preregistered laboratory experiment with 164 subjects in 12 sessions. In every match, the threshold is two consecutive all-$X$ rounds, the match continues after each round with probability 0.8, and if it ends before the threshold is reached, every member's final-round payoff is overwritten to $-100$ points. An all-$X$ round pays each member 100 points, choosing $X$ otherwise pays 0, and choosing $Y$ pays a safe payoff of 15 or 85 points ($u_Y \in \{0.15, 0.85\}$ in normalized terms). Group size varies between subjects ($n \in \{2, 4\}$) and the safe payoff within subjects, yielding four conditions: $2l$, $2h$, $4l$, and $4h$, where the number is the group size and $l$ and $h$ denote the low and high safe payoff. Because larger groups require more optimistic beliefs about each opponent and a higher safe payoff makes $Y$ more attractive, the dynamic basin ranges from close to one at all states in $2l$ to below 0.5 at all states in $4h$. In every condition, it expands from $k = 0$ to $k = 1$ but not from $k = 1$ to $k = 2$, so the probability assigned to $X$ should rise more in the first step than in the second. In each round, subjects choose a probability of $X$ that the computer then implements (Romero and Rosokha, 2023). This elicitation separates individual mixing from population heterogeneity and allows a direct comparison with the mixed MPE benchmark. Two baseline sessions with $n = 4$ remove the public state and the terminal loss and provide a descriptive benchmark for a standard repeated stag hunt.

The experiment yields four main findings. Together, they show that the robustness of coordination depends systematically on accumulated progress, as the dynamic basin predicts,

but that the current state does not fully summarize behavior. The first three findings answer the three research questions in turn, and the fourth concerns individual mixing.

First, coordination depends systematically on progress. In the pooled sample, the probability assigned to $X$ increases as the group moves toward the threshold, with the largest increase between $k = 0$ and $k = 1$, and coordination is highly persistent once groups reach the threshold. The predicted gap between the first and second increments has the expected sign but falls just short of statistical significance at the 5% level. Because the state is generated by past play, I interpret these state gradients as conditional associations rather than causal effects.

Second, groups do not simply start over after a failure. Holding the state fixed at $k = 0$, subjects whose group failed because another member chose $Y$ assign a 5.0 percentage-point lower probability to $X$ than at the start of the match. The decline is concentrated in $4h$, the condition with the smallest dynamic basin, where it reaches 20.1 percentage points, so a coordination failure appears to discourage coordination most where strategic uncertainty is greatest. Still, post-reset behavior rarely resembles the permanent switch to $Y$ prescribed by Grim Trigger.

Third, consistent with the comparative statics of the dynamic basin, coordination is less robust when the safe action is more attractive and, when the safe payoff is high, in larger groups. Lowering the safe payoff increases the probability assigned to $X$ for both group sizes. Group size matters only when the safe payoff is high, where groups of two assign substantially higher probabilities to $X$ than groups of four. When the safe payoff is low, coordination is already near the upper bound in both group sizes.

Fourth, probability elicitation reveals that individual mixing is common, particularly at the initial state. Interior probabilities are concentrated at $k = 0$, pure-$X$ choices dominate at higher states, and mixing declines with experience over repeated matches. This state pattern is in line with the monotone mixed MPE, which prescribes mixing at $k = 0$ and $X$ at higher states, although behavior does not exactly implement that benchmark.

This paper contributes to the literature on strategic uncertainty and equilibrium selection by extending basin-of-attraction reasoning to a game in which successful coordination itself changes the state. Three closely related papers study basin measures in dynamic settings. Vespa and Wilson (2016) use basin-of-attraction reasoning to compare competing equilibrium predictions in a dynamic game; Gill et al. (2026) develop basin measures for equilibrium selection over investment and task choice decisions in dynamic team skill formation; and Rosokha and Wei (2024)study a dynamic stochastic queueing game, using the basin of attraction of Always Defect against Grim Trigger to measure strategic uncertainty and showing that the visibility of the state determines whether subjects use history-contingent,

state-contingent Markov, or state-and-history-contingent strategies. None considers a dynamic stag hunt in which successful coordination advances a publicly observed state toward a threshold. My dynamic basin also adopts the per-opponent formulation that Boczoń et al. (2025) develop for $N$-player repeated games, although their basin is not state-dependent. The dynamic basin developed here delivers state-by-state predictions and comparative statics in group size and the safe payoff, and it coincides with the mixed-MPE belief cutoff at the initial state, linking the Markov and history-dependent approaches. The experimental evidence largely supports these predictions.

This paper also relates to experiments on stag-hunt coordination over time (Rankin et al., 2000; Al-Ubaydli et al., 2013; Duffy and Fehr, 2018; Brocas and Carrillo, 2025). In these environments, repeated interaction allows behavior and coordination to evolve, but current coordination does not itself generate the payoff-relevant progress state studied here. Closer to my setting, Gerlagh and van der Heijden (2024) experimentally study a dynamic three-player stag-hunt game in which groups can move over several rounds from one stable state to a higher-payoff stable state, and they compare neutral and environmental framing. In a related sequence of papers, Rosokha et al. (2024, 2025, 2026) study an indefinitely repeated multiplayer stag hunt in which the benefits of the collective action are divided according to political power won in a contest, so that the distribution of power evolves endogenously with play. They examine how contest rules, an incumbency advantage, and spending limits affect cooperation. In contrast, in this paper the stage game is identical in every state, and the state, generated by consecutive successful coordination, matters only because it determines protection against an outside shock. I study equilibrium selection in this environment through a state-dependent basin of attraction. My baseline sessions, which remove the public state and the terminal loss, suggest that this structure matters most when coordination is otherwise difficult to sustain.

This paper further relates to experimental work on whether behavior in dynamic games conditions on the payoff-relevant state, as Markov strategies prescribe. Evidence from dynamic public-investment, common-pool, cooperation, and team-production environments shows that subjects often respond systematically to the current state, although their behavior need not coincide exactly with Markov equilibrium predictions (Battaglini et al., 2012, 2016; Agranov et al., 2016; Vespa and Wilson, 2019; Vespa, 2020; Salz and Vespa, 2020; Rosokha and Wei, 2024; Gill et al., 2026). In coordination games, past play and precedents are known to shape subsequent equilibrium selection (Van Huyck et al., 1991; Rankin et al., 2000; Schmidt et al., 2003). In my game, the stage game is identical in every state, and a coordination failure returns the game to its initial state, so comparing behavior at the initial $k = 0$ with behavior after a reset holds the payoff-relevant state fixed and cleanly

tests state versus history dependence. Comparing initial behavior with behavior after resets caused by other members, I find that subjects assign a lower probability to $X$ after a failure even though the payoff-relevant state is unchanged, and that this history dependence is concentrated in the $4h$ treatment, the condition with the smallest dynamic basin.

Finally, this paper extends the probability elicitation of Romero and Rosokha (2023) from the indefinitely repeated prisoner's dilemma to dynamic coordination. It shows that individual mixing is a common feature of behavior that is concentrated at the initial state and declines with both progress and experience, so that heterogeneity in observed actions cannot be attributed exclusively to heterogeneity in pure decisions. The elicited probabilities also allow a direct comparison with the state-contingent mixing pattern of the mixed MPE.

The paper is organized as follows. Section 2 presents the model, characterizes the Markov Perfect Equilibria, and develops the dynamic basin of attraction. Section 3 describes the experimental design and hypotheses, Section 4 presents the results, and Section 5 discusses and concludes.

# 2 Theoretical Framework

## 2.1 Model Settings

The model is an $n$-player dynamic stag-hunt game with a public state that records consecutive all-$X$ coordination and an exogenous risk shock. In each round, agent $i \in \{1, \ldots, n\}$ chooses between two actions: $X$ and $Y$. To normalize the payoffs, let the payoff of action $X$ be 1 if the group outcome is all-X (i.e., every group member chooses $X$), and 0 otherwise; let the payoff of action $Y$ be $u_Y \in (0, 1)$. See Table 1 for the stage payoffs of Agent $i$.

| | | *Agent j* | |
|---|---|---|---|
| | | X | Y |
| *Agent i* | X | $(1, 1)$ | $(0, u_Y)$ |
| | Y | $(u_Y, 0)$ | $(u_Y, u_Y)$ |

*(a) Group size $n = 2$*

| | | *Group outcome* | |
|---|---|---|---|
| | | all-X | not all-X |
| *Agent i* | X | 1 | 0 |
| | Y | | $u_Y$ |

*(b) General group size $n$*

Table 1: Normalized Stage Payoffs

An outside shock occurs with probability $1 - \delta \in (0, 1)$ at the end of each round, after all group members have made their choices, where $\delta$ is the continuation probability of this game. Let $t \leq T$ be the round number, where $T$ marks the last round. Let $k_t$ denote the public state at the end of round $t$. The state is a capped counter of consecutive all-$X$ rounds.

Given $k_0 = 0$, the counter $k_t$ updates as

$$k_t = \begin{cases} \min\{k_{t-1} + 1, \bar{K}\} & \text{if all-}X\text{ in round } t, \\ 0, & \text{otherwise} \end{cases}$$

where $\bar{K}$ is the state threshold. The group is *protected* once $k$ reaches $\bar{K}$. Thus, at the beginning of round $t$, agents observe the inherited state $k_{t-1}$ before making their decisions. When the time index is unnecessary, I use $k$ to denote the state observed at the beginning of a decision round. When the shock is realized (in round $T$) and $k_T < \bar{K}$, the payoffs in round $T$ are overwritten to $-d$ for all members. Otherwise, payoffs are not overwritten. Although the within-round stage payoffs are identical across states, the public state $k$ is payoff-relevant through continuation values: if the game ends before the state reaches $\bar{K}$, every member's final-round payoff is overwritten to $-d$. Hence, the environment is a dynamic stochastic game rather than a standard indefinitely repeated game with a fixed stage game.

In a standard repeated stag hunt, the same stage game is repeated over time, but current actions do not typically change an endogenous payoff-relevant state of the game. In the present environment, by contrast, coordination on $X$ affects both current payoffs and the public state $k$. The current state therefore summarizes the group's progress toward protection and directly affects continuation incentives. At the same time, the environment is stationary because the continuation probability $\delta$ is constant. Conditional on the same current state $k$, the continuation problem is therefore the same regardless of how many rounds have already elapsed or how that state was reached. In particular, after any reset to $k = 0$, the incentives to begin coordinating toward the threshold are the same as at any other visit to $k = 0$.

Before characterizing equilibrium behavior, I establish the socially efficient benchmark. At any state $k$, unanimous choice of $X$ yields a total stage payoff of $n$, whereas any action profile containing at least one $Y$ yields a total stage payoff no greater than $nu_Y < n$. Moreover, when $k < \bar{K}$, unanimous $X$ moves the group toward protection, while any outcome containing $Y$ resets the state to zero. Thus, unanimous $X$ weakly improves both the current group payoff and the continuation state.

*Remark.* At every state $k \in \{0, \ldots, \bar{K}\}$, it is socially efficient for all agents to choose $X$. Consequently, the socially efficient path reaches the protected state after $\bar{K}$ consecutive all-$X$ rounds.

Importantly, the socially efficient action is the same in all experimental treatments. Changes in $u_Y$ and $n$ therefore affect the strategic difficulty of implementing the efficient outcome rather than the efficient benchmark itself. The equilibrium analysis below studies how this difficulty varies with the current state and treatment parameters.

## 2.2 Markov Perfect Equilibrium

Because the state $k$ summarizes everything that is payoff-relevant about the past, a natural benchmark is a *Markov strategy*, in which an agent's choice depends only on the current state and not on the detailed history of play. A *Markov Perfect Equilibrium* (MPE) is a profile of Markov strategies in which each agent's strategy is a best response to the others' strategies in every state (Maskin and Tirole, 1988a,b, 2001). It refines Subgame Perfect Equilibrium (SPE), the standard equilibrium concept for dynamic games, which requires best responses after every possible history. I focus on Markov strategies that are *symmetric*, so that all agents use the same strategy, and *stationary*, so that the strategy does not change over time (Dutta, 1995). Such a strategy is summarized by $\sigma = (\sigma_0, \sigma_1, \ldots, \sigma_{\bar{K}})$, where $\sigma_k \in [0,1]$ is the probability that an agent chooses $X$ in state $k$. A strategy is *pure* if $\sigma_k \in \{0,1\}$ in every state and *mixed* otherwise.

Let $V_k$ denote an agent's *continuation value*, the expected total payoff from the beginning of a round in state $k$ onward when all agents follow $\sigma$. Let $Q_k(X)$ and $Q_k(Y)$ denote the expected payoff from choosing $X$ or $Y$ in the current round and following $\sigma$ thereafter. Choosing $Y$ guarantees that the round is not all-$X$, so the next state is 0, and the payoff is overwritten to $-d$ if the shock arrives:

$$Q_k(Y) = \delta\big(u_Y + V_0\big) + (1-\delta)(-d).$$

An agent who chooses $X$ obtains an all-$X$ round only if all other $n-1$ members also choose $X$, which occurs with probability $q_k = \sigma_k^{n-1}$. If $k < \bar{K} - 1$, even an all-$X$ round leaves the group unprotected, whereas if $k \geq \bar{K} - 1$, an all-$X$ round reaches (or keeps) the protected state, so the shock no longer triggers the loss:

$$Q_k(X) = \begin{cases} q_k\Big(\delta\big(1 + V_{k+1}\big) + (1-\delta)(-d)\Big) + (1-q_k)\Big(\delta V_0 + (1-\delta)(-d)\Big) & \text{if } k < \bar{K} - 1, \\ q_k\Big(\delta\big(1 + V_{\bar{K}}\big) + (1-\delta)\Big) + (1-q_k)\Big(\delta V_0 + (1-\delta)(-d)\Big) & \text{if } k \geq \bar{K} - 1. \end{cases}$$

A symmetric stationary MPE is a collection $\{V_k, \sigma_k\}_{k=0}^{\bar{K}}$ such that, for each state $k$,

$$V_k = \max\{Q_k(X), Q_k(Y)\}, \qquad \sigma_k = \begin{cases} 1 & \text{if } Q_k(X) > Q_k(Y), \\ 0 & \text{if } Q_k(X) < Q_k(Y), \\ \in (0,1) & \text{only if } Q_k(X) = Q_k(Y). \end{cases}$$

That is, agents choose $X$ for sure when $X$ is strictly better, $Y$ for sure when $Y$ is strictly

better, and mix only when they are indifferent.[3] The game has many MPEs. I focus on *monotone* MPEs, in which the probability of choosing $X$ does not decrease as the group makes progress.

*Definition* 1 (Monotone Markov strategy). A symmetric stationary Markov strategy $\sigma$ is *monotone* if $\sigma_0 \leq \sigma_1 \leq \cdots \leq \sigma_{\bar{K}}$.

Throughout the paper, *monotone* means weakly increasing in $k$. A higher state means that the group needs fewer additional all-$X$ rounds to become protected, so it is difficult to justify behaviorally a strategy that chooses $X$ at a low state but abandons $X$ at a higher one. Section 2.2.3 briefly discusses the non-monotone MPEs that this restriction excludes.

For brevity, I write a strategy as a string that lists $\sigma_k$ in order of $k = 0, 1, \ldots, \bar{K}$, where 0 denotes $\sigma_k = 0$ (choose $Y$ for sure), 1 denotes $\sigma_k = 1$ (choose $X$ for sure), and $m$ denotes interior mixing, $\sigma_k \in (0, 1)$. For example, with $\bar{K} = 2$, $(m11)$ prescribes mixing at $k = 0$ and choosing $X$ for sure at $k = 1$ and $k = 2$. Any MPE with $\sigma_0 = 0$ never leaves the initial state: all members choose $Y$ at $k = 0$, so the state never rises to $k = 1$. I call such an MPE *off-path*. By contrast, I call an MPE *on-path* if $\sigma_0 > 0$, so that equilibrium play reaches higher states with positive probability.

### 2.2.1 Monotone Pure MPEs

In a symmetric pure strategy, the other members either all choose $Y$ or all choose $X$ in a given state. If they all choose $Y$, choosing $X$ cannot produce an all-$X$ round, so $Y$ is a best response. If they all choose $X$, choosing $X$ is a best response, because it yields a higher current payoff ($1 > u_Y$) and advances (or keeps) the state instead of resetting it. Hence, every symmetric pure Markov strategy is an MPE.[4] Among them, the monotone ones take a simple threshold form.

*Proposition* 1 (Monotone pure MPEs). The symmetric monotone pure MPEs are exactly the threshold strategies

$$a^{(r)}(k) = \begin{cases} Y, & k < r, \\ X, & k \geq r, \end{cases} \qquad r \in \{0, 1, \ldots, \bar{K} + 1\}.$$

Hence, there are exactly $\bar{K} + 2$ symmetric monotone pure MPEs.

[3] The condition is "only if" rather than "if" because indifference is necessary but not sufficient for mixing. When $Q_k(X) = Q_k(Y)$, every $\sigma_k \in [0, 1]$ is a best response, including the pure choices $\sigma_k = 0$ and $\sigma_k = 1$. Thus, an agent who mixes must be indifferent, but an indifferent agent need not mix.

[4] See Appendix A1.4 for the formal statement and proof, as well as the proof of Proposition 1 below.

Only $r = 0$, under which all members choose $X$ in every state, is on-path. For every $r \geq 1$, members choose $Y$ at $k = 0$, so play never leaves $k = 0$ and is identical to that of the all-$Y$ MPE ($r = \bar{K}+1$), the natural coordination-failure benchmark. Thus, along the path of play, monotone pure MPEs predict either full coordination on $X$ or permanent coordination on $Y$, and neither says how coordination should vary with the state, the safe payoff, or group size. The mixed MPEs below predict how play varies with the state, and the dynamic basin in Section 2.3 also predicts how it varies with the safe payoff and group size.

### 2.2.2 Monotone Mixed MPEs

In a mixed MPE, $\sigma_k \in (0, 1)$ in at least one state, which requires $Q_k(X) = Q_k(Y)$ in that state. To find all monotone mixed MPEs, I first list every candidate pattern of 0, $m$, and 1 that never moves "down" in the order $0 \to m \to 1$ as $k$ increases (e.g., $(0m1)$ but not $(m01)$); I call these *regime-monotone* candidates. I then solve the indifference conditions of each candidate in closed form and keep only the candidates whose solved probabilities satisfy $\sigma_0 \leq \sigma_1 \leq \cdots \leq \sigma_{\bar{K}}$. The second step is necessary because a candidate such as $(mm)$ can have $\sigma_0 > \sigma_1$ once the two mixing probabilities are solved.[5]

*Proposition* 2 (On-path monotone mixed MPEs). For $\bar{K} \in \{1, 2, 3\}$, there is exactly one on-path monotone mixed MPE: $(m1)$ for $\bar{K} = 1$, $(m11)$ for $\bar{K} = 2$, and $(m111)$ for $\bar{K} = 3$. In each case, members mix at $k = 0$ and choose $X$ for sure in every state $k \geq 1$.

Table 2 summarizes the monotone pure and mixed MPEs for $\bar{K} = 1, 2, 3$, and Figure 1 reports the closed-form solutions of the three on-path monotone mixed MPEs. In each of them, strategic uncertainty is concentrated at the initial state. Once the group has achieved one all-$X$ round, all members choose $X$ for sure, so the group reaches the protected state and remains there. The mixed MPE $(m11)$ with $\bar{K} = 2$ serves as the mixed-MPE benchmark for the experiment.

The mixing probability $\sigma_0$ admits two interpretations. Under the *equilibrium interpretation*, $\sigma_0$ is the probability with which each member chooses $X$ at $k = 0$. Under the *best-response interpretation*, $\sigma_0$ is a *belief cutoff*: if a member believes that the others choose $X$ for sure at $k \geq 1$ and choose $X$ at $k = 0$ with probability above $\sigma_0$, then $X$ is a strict best response at $k = 0$. As with mixed equilibria in coordination games generally, the equilibrium interpretation implies counterintuitive comparative statics ($\sigma_0$ rises with $u_Y$ and $n$), because each member's mixing probability is pinned down by the requirement that the *other* members be indifferent. I adopt the best-response interpretation, under which $1-\sigma_0$ is the largest probability of $Y$ per other member that a member can tolerate while $X$ remains optimal.

[5] Appendix A1 reports the full procedure and the solutions to every candidate for $\bar{K} = 1, 2, 3$.

Table 2: Symmetric Monotone MPEs for $\bar{K} = 1, 2, 3$

| | Pure | | Mixed | |
|---|---|---|---|---|
| $\bar{K}$ | All monotone | On-path | All monotone | On-path |
| 1 | (00), (01), (11) | (11) | $(0m)$, $(m1)$ | $(m1)$ |
| 2 | (000), (001), (011), (111) | (111) | $(00m)$, $(0mm)$, $(0m1)$, $(m11)$ | $(m11)$ |
| 3 | (0000), (0001), (0011), (0111), (1111) | (1111) | $(000m)$, $(00mm)$, $(00m1)$, $(0m11)$, $(m111)$ | $(m111)$ |

*Note:* Each profile lists $\sigma_k$ for $k = 0, 1, \ldots, \bar{K}$. 0 and 1 denote $\sigma_k = 0$ and $\sigma_k = 1$; $m$ denotes interior mixing, $\sigma_k \in (0, 1)$. An MPE is on-path if $\sigma_0 > 0$, so that play can leave the initial state $k = 0$. Appendix Table A1.1 also lists the mixed candidates that fail monotonicity.

For example, in $(m11)$ with $\delta = 0.8$ and $d = 1$, $\sigma_0 = 0.02$ when $n = 2$ and $u_Y = 0.15$, whereas $\sigma_0 = 0.62$ when $n = 4$ and $u_Y = 0.85$. Section 2.3 shows that this cutoff coincides with the dynamic basin at the initial state.

### 2.2.3 Non-monotone MPEs

For completeness, the game also has MPEs that are not monotone. Because every symmetric pure Markov strategy is an MPE, there are $2^{\bar{K}+1}$ pure MPEs, of which only the $\bar{K} + 2$ in Proposition 1 are monotone. For example, with $\bar{K} = 2$, (100), under which members choose $X$ at $k = 0$ but switch to $Y$ as soon as the group makes progress, is a pure MPE. Some mixed MPEs are also non-monotone. For example, with $\bar{K} = 1$, the mixed MPE $(mm)$ has $\sigma_0 > \sigma_1$.[6] Such strategies abandon $X$ precisely as the group approaches protection. In this paper, however, I focus only on monotone MPEs.

## 2.3 History-dependent Strategies and Dynamic Basin of Attraction of Coordination

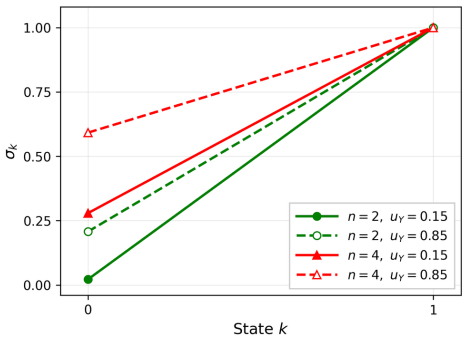
The game has multiple equilibria, including full coordination on $X$ and permanent coordination on $Y$. Which one agents select depends on *strategic uncertainty*, that is, uncertainty about whether the other members will do their part. In the experimental literature on indefinitely repeated games, the existence of a cooperative equilibrium alone provides limited guidance for equilibrium selection, whereas measures of robustness to strategic uncertainty provide sharper behavioral predictions (Blonski et al., 2011; Dal Bó and Fréchette, 2011, 2018). This subsection develops such a measure for the present game, in which the robustness of coordination varies with the state.

[6] See Appendix A1 for the solution of $(mm)$ and the other non-monotone mixed candidates.

Figure 1: Solutions for the On-path Monotone Mixed MPEs $(m1)$, $(m11)$, and $(m111)$

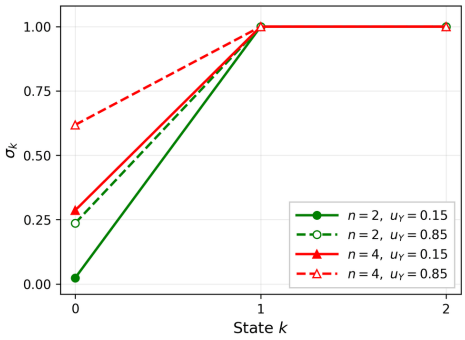
$[\bar{K} = 1$**: MPE** $(m1)]$

$$V_0 = \frac{\delta}{1-\delta} u_Y - d, \qquad V_1 = \frac{1}{1-\delta},$$
$$\sigma_0 = \left( \frac{\delta(1-\delta)u_Y}{1 - \delta^2 u_Y + (1-\delta)d} \right)^{\frac{1}{n-1}}, \qquad \sigma_1 = 1.$$



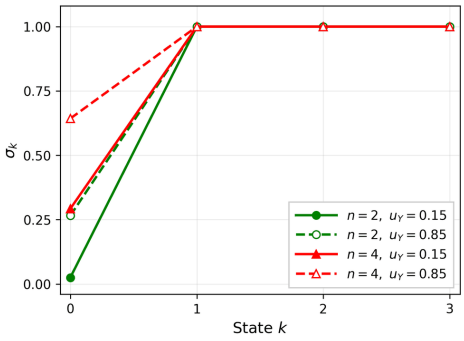

$[\bar{K} = 2$**: MPE** $(m11)]$

$$V_0 = \frac{\delta}{1-\delta} u_Y - d, \qquad V_1 = V_2 = \frac{1}{1-\delta},$$
$$\sigma_0 = \left( \frac{(1-\delta)u_Y}{(2-\delta) - \delta u_Y + (1-\delta)d} \right)^{\frac{1}{n-1}}, \quad \sigma_1 = \sigma_2 = 1.$$




$[\bar{K} = 3$**: MPE** $(m111)]$

$$V_0 = \frac{\delta}{1-\delta} u_Y - d, \qquad V_1 = \frac{\delta(2-\delta)}{1-\delta} - (1-\delta)d,$$
$$V_2 = V_3 = \frac{1}{1-\delta}, \qquad \sigma_1 = \sigma_2 = \sigma_3 = 1,$$
$$\sigma_0 = \left( \frac{(1-\delta)u_Y}{1 + \delta(1-\delta)(1+d) - \delta u_Y} \right)^{\frac{1}{n-1}}.$$




*Notes:* The plots illustrate $\sigma_k$ by state for $\delta = 0.8$ and $d = 1$; green for $n = 2$, red for $n = 4$, solid lines for $u_Y = 0.15$, and dashed lines for $u_Y = 0.85$. Derivations are in Appendix A1; additional parameter comparisons for $(m11)$ are in Appendix A2.

**Static basin of attraction.** I begin with the one-shot benchmark of Dal Bó et al. (2021). In a two-player one-shot stag hunt, the size of the *basin of attraction* of the risky action is the largest probability that the other player chooses the safe action such that the risky action remains a best response. With the payoffs in panel (a) of Table 1, if agent $i$ believes that agent $j$ chooses $Y$ with probability $q$, then $X$ yields $1 - q$ and $Y$ yields $u_Y$. Hence, $X$ is a best response whenever $q \leq 1 - u_Y$, and the size of the basin of attraction of $X$ is

$$\mathcal{B} = 1 - u_Y.$$

I refer to $\mathcal{B}$ as the *static basin*. Action $X$ is *risk dominant* when $\mathcal{B} \geq 0.5$, that is, when $X$ is a best response to the belief that the other player is equally likely to choose either action. A larger $u_Y$ shrinks $\mathcal{B}$, so $X$ is optimal for a smaller range of beliefs and coordination on $X$ becomes strategically riskier.

**Dynamic basin of attraction.** A dynamic analogue must specify which strategies an agent is uncertain about. Comparing unconditional All-$X$ (choose $X$ in every round) with All-$Y$ (choose $Y$ in every round) is problematic in a dynamic game: once an All-$X$ agent observes a $Y$, continuing to choose $X$ is no longer optimal, yet All-$X$ cannot respond. Following Dal Bó and Fréchette (2011), I instead compare All-$Y$ with *Grim Trigger* (GT), under which an agent chooses $X$ as long as every previous round has been all-$X$ and switches permanently to $Y$ after the first round that is not. GT keeps coordination on $X$ possible while responding optimally to a coordination failure. GT is an SPE, and in this game it induces the same play as Tit-for-Tat and state-dependent variants of both rules, both on the equilibrium path and after any unilateral deviation.[7]

Suppose that agent $i$ believes that each other member, independently, is a GT type with probability $p$ and an All-$Y$ type with probability $1-p$. Following Boczoń et al. (2025), this measures strategic uncertainty *per opponent*. Let $m \equiv \max\{\bar{K} - k, 1\}$ denote the number of consecutive all-$X$ rounds the group still needs to become protected from state $k$, truncated at one because a single all-$X$ round secures protection from both $k = \bar{K} - 1$ and $k = \bar{K}$. Agent $i$ prefers GT to All-$Y$ at a cooperative history in state $k$, that is, a history with no coordination failure so far, if and only if $p \geq p_k^\star$,[8] where

$$p_k^\star \equiv \left( \frac{u_Y}{\frac{1-\delta u_Y}{1-\delta} + (1+d)\delta^{m-2}} \right)^{\frac{1}{n-1}}.$$

*Definition* 2 (Dynamic basin of attraction). The *dynamic basin of attraction* of $X$ at state $k$ (hereafter, the *dynamic basin*) is

$$\mathcal{DB}_k \equiv 1 - p_k^\star,$$

the largest probability that each other member is an All-$Y$ type such that Grim Trigger remains optimal at state $k$.

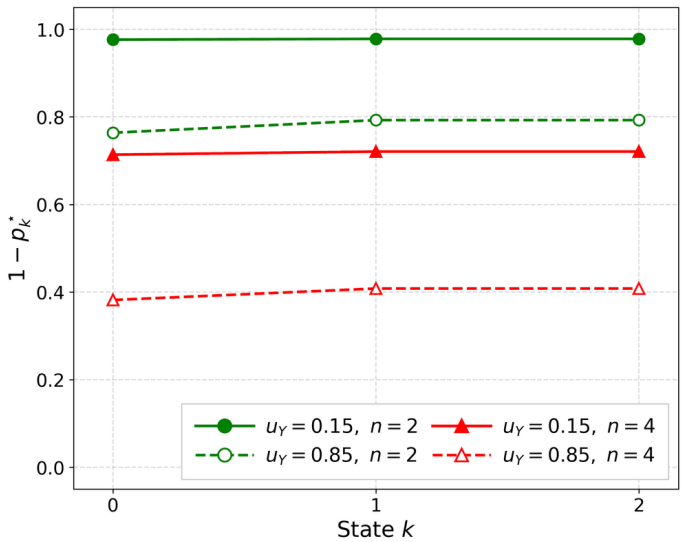
[7] See Appendix A3.1 for the definitions of these rules, the proof that they induce the same play, and the proof that they are SPEs.

[8] Appendix A3.2 derives this cutoff from the continuation values of the two strategies.

A larger dynamic basin means that coordination on $X$ remains optimal under more pessimistic beliefs about the other members. Three comparative statics follow directly. First, $\mathcal{DB}_k$ decreases in $u_Y$: a more attractive safe action makes $Y$ optimal under a wider range of beliefs. Second, $\mathcal{DB}_k$ decreases in $n$: GT pays off only if all $n-1$ other members are GT types, which is less likely in a larger group. Third, $\mathcal{DB}_k$ strictly increases in $k$ for $k \leq \bar{K}-1$ and is constant between $k=\bar{K}-1$ and $k=\bar{K}$, because from both states a single all-$X$ round secures protection. Thus, the dynamic basin expands as the group makes progress, but the expansion stops once protection is one round away.

At the initial state, the dynamic basin coincides with the belief cutoff from the monotone mixed MPE in Section 2.2.2: for $\bar{K}=1,2,3$, $p_0^\star$ equals $\sigma_0$ in $(m1)$, $(m11)$, and $(m111)$, so $\mathcal{DB}_0 = 1-\sigma_0$. In both constructions, a failure at $k=0$ leads to the All-$Y$ continuation value and success leads to the cooperative path, so the indifference conditions are identical.[9] The Markov and Grim-Trigger approaches therefore yield the same measure of strategic uncertainty at the initial state. Figure 2 illustrates the experimental parameterization, $\bar{K}=2$, $\delta=0.8$, $d=1$, $n\in\{2,4\}$, and $u_Y\in\{0.15,0.85\}$. With $n=2$ and $u_Y=0.15$, the dynamic basin is close to one at all states, so $X$ remains optimal under a wide range of beliefs about the other members. With $n=4$ and $u_Y=0.85$ (treatment $4h$ in Section 3), the dynamic basin is below 0.5 at all states, so choosing $X$ requires substantially more optimistic beliefs. This is therefore the most strategically difficult of the four treatment conditions. In every condition, the dynamic basin expands from $k=0$ to $k=1$ and is flat from $k=1$ to $k=2$.[10]

Figure 2: Dynamic Basin of Attraction for Grim Trigger against All-$Y$




[9] Appendix A3.3 shows the algebra for $\bar{K}=1,2,3$.
[10] Additional parameter comparisons are reported in Appendix A2.

# 3 Experimental Design

## 3.1 Treatments and Parameters

The experiment varies two treatment dimensions: the normalized payoff of the safe action, $u_Y \in \{0.15, 0.85\}$, and group size, $n \in \{2, 4\}$. Group size varies between sessions, whereas each subject experiences both values of $u_Y$, with the order counterbalanced across sessions. The remaining game parameters are fixed at $\bar{K} = 2$, $d = 1$, and $\delta = 0.8$ across all treatment conditions. Payoffs are implemented in experimental points by multiplying the normalized stage payoffs in Table 1 by 100. Thus, $u_Y = 0.15$ and $u_Y = 0.85$ correspond to 15 and 85 experimental points, respectively. The stage payoffs used in the laboratory are shown in Table 3.

| | | *Group outcome* | |
|---|---|---|---|
| | | all-$X$ | not all-$X$ |
| *You* | $X$ | 100 | 0 |
| | $Y$ | | 15 |

*(a) When $u_Y$ is low*

| | | *Group outcome* | |
|---|---|---|---|
| | | all-$X$ | not all-$X$ |
| *You* | $X$ | 100 | 0 |
| | $Y$ | | 85 |

*(b) When $u_Y$ is high*

Table 3: Stage Payoffs in Experimental Points

The parameter values are chosen to generate meaningful variation in strategic uncertainty while keeping the dynamic structure parsimonious. I set $\bar{K} = 2$ because with $\bar{K} = 1$ there is no intermediate unprotected state: one successful all-$X$ round immediately reaches protection. Moreover, for the monotone mixed-MPE benchmark, increasing the threshold from $\bar{K} = 2$ to $\bar{K} = 3$ mainly extends the sequence of pure-$X$ states following the initial mixing state rather than introducing a qualitatively different pattern. Thus, $\bar{K} = 2$ captures the main transition of interest in a more parsimonious experimental design. I set $\delta = 0.8$, implying an expected match length of five rounds. This is long enough, on average, for a group to reach the threshold, experience a coordination failure and reset, and subsequently reach the threshold again, creating within-match variation useful for distinguishing state from history dependence.

I use $n \in \{2, 4\}$ to generate a substantial change in coordination difficulty while keeping group formation manageable in the laboratory. Although four is not a large group in absolute terms, it doubles the two-player benchmark and creates a stronger theoretical contrast than $n = 3$. In particular, the individual belief threshold depends nonlinearly on group size through the power $1/(n-1)$, so increasing group size from two to four produces a larger change in the per-opponent belief cutoff than increasing it from two to three. I choose

$u_Y \in \{0.15, 0.85\}$ to generate a sharper contrast in the attractiveness of the safe action than the more moderate parameterization $u_Y \in \{0.25, 0.75\}$. Finally, I set $d = 1$ so that the terminal loss equals the normalized payoff from successful coordination in magnitude; together with the 100-point endowment that each subject receives at the beginning of every match, this prevents negative match earnings. Appendix A2 reports the corresponding theoretical comparisons under alternative values of $\bar{K}$, $u_Y$, $\delta$, and $d$.

Part 1 is the main game of the experiment and consists of 10 matches. Subjects are randomly rematched before each match. Group size $n$ is fixed throughout a session, while each value of $u_Y$ is used for five consecutive matches, with one payoff condition applying to Matches 1-5 and the other to Matches 6-10. The four Part 1 treatment conditions are denoted $2l$, $2h$, $4l$, and $4h$, where the number denotes group size and $l$ and $h$ denote the low and high safe-payoff conditions, respectively. Accordingly, the four treatment orders are $2l + 2h$, $2h + 2l$, $4l + 4h$, and $4h + 4l$.

Each of the four treatment orders was conducted once under each of the three predrawn termination sequences[11], giving 12 treatment sessions. Table 4 reports the realized number of subjects in each session. The treatment sample contains 164 subjects: 68 in the $n = 2$ sessions and 96 in the $n = 4$ sessions. The planned sample size was 168; two of the $n = 2$ sessions were conducted with 10 rather than 12 subjects due to actual attendance.

After completing the treatment sessions, I conducted two additional baseline sessions as a robustness benchmark. The baseline sessions use the same stage-game payoffs and random-termination procedure but omit the public state counter and the last-round payoff penalty. Both baseline sessions have group size $n = 4$ and use Sequence 2, with the two payoff orders counterbalanced. The baseline sample contains 32 subjects, bringing the total number of subjects across all non-pilot sessions to 196.

## 3.2 Implementation

The experiment was approved by the Purdue University IRB (STUDY2026-00000142) and conducted at the Vernon Smith Experimental Economics Laboratory (VSEEL) at Purdue University in August and September 2026. The experiment was implemented in oTree 6.0.14 (Chen et al., 2016). Pilot sessions conducted in April 2026 were used to test the software, instructions, and parameterization; pilot data are not included in the analysis.

Subjects were Purdue University undergraduate students from a variety of disciplines and were recruited by email using ORSEE (Greiner, 2015). Each subject participated in only one session and provided electronic informed consent before the experiment began. Given

[11] Appendix C2 lists the three predrawn termination sequences under $\delta = 0.8$.

Table 4: Treatment Orders and Experimental Sessions

| Treatment order | Safe payoff $u_Y$ (points) | | Sequence | | | # Sessions | # Subjects |
|---|---|---|---|---|---|---|---|
| | Matches 1-5 | Matches 6-10 | 1 | 2 | 3 | | |
| *Treatments* | | | | | | | |
| $2l + 2h$ | 15 | 85 | 12 | 12 | 10 | 3 | 34 |
| $2h + 2l$ | 85 | 15 | 10 | 12 | 12 | 3 | 34 |
| $4l + 4h$ | 15 | 85 | 16 | 16 | 16 | 3 | 48 |
| $4h + 4l$ | 85 | 15 | 16 | 16 | 16 | 3 | 48 |
| Treatment total | | | 54 | 56 | 54 | 12 | 164 |
| *Baseline* | | | | | | | |
| $4l + 4h$ | 15 | 85 | - | 16 | - | 1 | 16 |
| $4h + 4l$ | 85 | 15 | - | 16 | - | 1 | 16 |
| Total | | | 54 | 88 | 54 | 14 | 196 |

*Notes:* Entries under Sequences 1-3 report the number of subjects in each session. Each treatment was conducted once under each of the three sequences. The baseline sessions use the same stage-game payoffs but omit the public counter and the last-round penalty. For the treatments, $\bar{K} = 2$, $d = 1$, and $\delta = 0.8$ are fixed across treatment conditions.

the importance of common information in coordination games (Chen et al., 2021), subjects were informed on the Overview page that all participants in a session received the same instructions and that deception would not be used.

Each session consisted of three parts. Part 1 contained the 10 matches of the dynamic stag-hunt game described above. Before beginning Part 1, subjects received the instructions and completed a paid six-question comprehension quiz designed to reinforce and assess their understanding of the game. The payoff table for Matches 1-5 differed from that for Matches 6-10 only in the safe-option payoff $u_Y$. Subjects were informed at the outset that Part 1 contained 10 matches, but the change in the safe action payoff was revealed only after they completed Match 5.

The following procedure applies to the 12 treatment sessions. The two baseline sessions followed the same procedures, but without the public state counter and the last-round payoff penalty. The public state was displayed to subjects as a public counter. Let $k_t$ denote the updated state at the end of round $t$, with $k_0 = 0$. At the beginning of round $t$, subject $i$ observes the group-level inherited state $k_{t-1}$, displayed on the interface as the current public counter. Subject $i$ then uses a slider to assign a probability to the action $X$, denoted by $p_i(X) \in \{0, \ldots, 100\}$, individually and simultaneously. To identify whether subjects use mixed strategies at each state, I use an elicitation method similar to Romero and Rosokha (2023): Alongside the slider, a $10 \times 10$ grid containing 100 numbered cells visualizes the

chosen probabilities: $p_i(X)$ cells are displayed in yellow for $X$, and the remaining $100-p_i(X)$ cells are displayed in blue for $Y$. After subject $i$ submits the decision $p_i(X)$, the computer draws an integer uniformly from $\{1,\ldots,100\}$. If the draw falls in a yellow cell, action $X$ is implemented for the subject in that round; otherwise, action $Y$ is implemented. The realized action is shown to the participant. After all group members have submitted their decisions and actions have been realized, subjects observe the group outcome (a list of each member's action), their own round payoff, and the updated state $k_t$ at the end of a round. An all-$X$ outcome increases the state by one until the threshold $\bar{K}=2$ is reached, whereas a non-all-$X$ outcome resets the state to zero. Once the threshold has been reached, another all-$X$ outcome maintains the group at the protected state.

Each match used random termination with continuation probability $\delta=0.8$. The continuation draws were generated in advance as independent draws from the discrete uniform distribution on $\{1,\ldots,100\}$. At the end of each round, the corresponding draw was displayed to subjects through a spinner with an 80% white continuation region and a 20% gray stopping region. A draw at or below 80 continued the match, while a draw above 80 ended it. Subjects knew the continuation probability but did not know the predrawn sequence or, when making their decisions, which round would ultimately be the last. Appendix C2 reports the three predrawn continuation sequences.

In the treatment sessions, if the state at the end of the last round satisfied $k_T<\bar{K}$, every group member's payoff for round $T$ was overridden to $-100$ experimental points. If $k_T=\bar{K}$, the payoff for the last round was determined by the stage-payoff table as usual. This last-round payoff rule was absent in the baseline sessions.

Part 2 elicits subjects' risk preferences using a multiple-price-list task adapted from Cason et al. (2025). Part 3 measures subjects' social preferences using the Social Value Orientation (SVO) task introduced by Murphy et al. (2011). Subjects receive incentive payments for both Parts 2 and 3. After completing Part 3, subjects answer a short survey collecting demographic information and brief descriptions of the strategies they used in Part 1.

Total earnings from the comprehension quiz and all three parts were displayed only on the final page of the experiment and were paid in cash. The average payment was \$17.15 per subject, with an interquartile range of [\$14.96, \$19.86]. Across the 14 non-pilot sessions, 196 subjects participated: 164 in the 12 main treatment sessions and 32 in the two baseline sessions. Due to actual attendance, the treatment sessions with $n=2$ included 10-12 subjects, while all $n=4$ sessions included 16 subjects. Sessions lasted approximately 45 minutes, including instructions and payment distribution. See Appendix C1 for the experimental instructions, comprehension quiz, game interfaces, and survey.

## 3.3 Hypotheses

For the experimental parameters $\bar{K} = 2$, $\delta = 0.8$, and $d = 1$, the theoretical framework generates three comparative static predictions that guide the experiment. The dynamic basin expands as the public state $k$ increases, is larger when the safe-option payoff $u_Y$ is lower, and is larger in smaller groups. At the same time, unanimous $X$ remains socially efficient in every state and treatment. Thus, the experimental variation changes the strategic difficulty of coordination while holding the efficient action fixed.

Based on the equilibrium analysis and the dynamic basin of attraction derived in Section 2, I formulate three main hypotheses corresponding to those specified in the preregistration. The first concerns how behavior varies with the current public state, the second distinguishes state-dependent from history-dependent strategies, and the third concerns comparative statics in group size and the payoff from the safe action. Note that throughout the empirical analysis, $k$ denotes the state observed when the subject makes the current-round decision, corresponding to $k_{t-1}$ in the time-indexed game. The study was preregistered on AsPredicted (#308239) prior to data collection.

*Hypothesis* 1 (Monotonicity in state). The probability assigned to $X$ is weakly increasing in the current public state $k$, and the increase from $k = 0$ to $k = 1$ is greater than the increase from $k = 1$ to $k = 2$.

Both the state-based MPE analysis and the history-based Grim-Trigger benchmark imply that coordination on $X$ becomes more attractive as the group moves closer to the threshold. In the state-based framework, a higher $k$ increases the continuation value associated with successful coordination. In the dynamic-basin framework, a higher $k$ expands the range of beliefs under which continued coordination on $X$ is preferred. Thus, although the two frameworks differ in whether behavior depends only on the current state or also on the history leading to that state, both predict that the probability assigned to $X$ should weakly increase with $k$.

The stronger prediction that the increase from $k = 0$ to $k = 1$ exceeds the increase from $k = 1$ to $k = 2$ follows from the shape of the dynamic basin: with $\bar{K} = 2$, the basin expands from $k = 0$ to $k = 1$ but not from $k = 1$ to $k = 2$ (Section 2.3, Figure 2). I therefore examine not only whether $p(X)$ increases across states, but also whether

$$p(X \mid k = 1) - p(X \mid k = 0) > p(X \mid k = 2) - p(X \mid k = 1).$$

*Hypothesis* 2 (State- versus history-dependent strategies). Holding the current public state $k$ fixed, the probability assigned to $X$ does not systematically depend on the history leading to that state. In particular, behavior at $k = 0$ following a reset should be similar to behavior

at the initial $k = 0$ state of a match.

The current public state $k$ is a sufficient statistic for the objective continuation problem. Conditional on the same value of $k$, the continuation probability, payoff structure, and distance from the threshold are the same regardless of the history through which that state was reached. In particular, after a coordination failure resets the state to $k = 0$, the objective incentives to begin coordinating toward the threshold are the same as at the initial $k = 0$ state of a match.

Moreover, unlike in a repeated prisoner's dilemma, coordination on $X$ does not require a history-dependent punishment to deter a profitable unilateral deviation: when all other group members coordinate on $X$, choosing $X$ is itself a best response. The current state is also publicly displayed at every decision, making it a salient basis for behavior. These features make state-dependent strategies a natural behavioral benchmark.

Previous coordination experiments show that past play and historical precedents can shape subsequent equilibrium selection (Van Huyck et al., 1991; Rankin et al., 2000; Schmidt et al., 2003). In this paper, the public state allows history dependence to be examined while holding the current payoff-relevant state fixed. History dependence nevertheless remains plausible. A coordination failure may cause subjects to update their beliefs about the willingness of their current group members to coordinate, even though the objective continuation problem has returned to the same state. The distinction is therefore an empirical one rather than a mechanical consequence of the game. Hypothesis 2 tests whether behavior is summarized by the current state or retains systematic dependence on the path through which that state was reached.

*Hypothesis* 3a (Comparative statics in group size). For a given payoff from the safe action $Y$, the probability assigned to $X$ is higher in smaller groups.

Successful coordination requires confidence that all other $n - 1$ group members will coordinate on $X$. Under the dynamic-basin framework, this becomes more difficult as group size increases because each subject must hold sufficiently optimistic beliefs about a larger number of opponents. Equivalently, the per-opponent dynamic basin shrinks with $n$. This predicted group-size effect is also related to experimental evidence of greater coordination difficulty in larger-group environments (Van Huyck et al., 1990; Weber, 2006). The theoretical comparative statics hold state by state. Experimentally, Hypothesis 3a implies two pairwise comparisons:

$$p(X \mid 2l) > p(X \mid 4l) \qquad \text{and} \qquad p(X \mid 2h) > p(X \mid 4h).$$

*Hypothesis* 3b (Comparative statics in safe-option payoff). For a given group size, the probability assigned to $X$ is higher when the payoff from choosing $Y$ is lower.

In the dynamic-basin analysis, a higher payoff from the safe action $Y$ makes persistent play of $Y$ more attractive and reduces the range of beliefs under which coordination on $X$ is preferred. Hence, the dynamic basin is larger when $u_Y$ is lower. Experimentally, Hypothesis 3b implies two pairwise comparisons:

$$p(X \mid 2l) > p(X \mid 2h) \qquad \text{and} \qquad p(X \mid 4l) > p(X \mid 4h).$$

Together, Hypotheses 3a and 3b provide four pairwise comparisons that vary one experimentally manipulated parameter at a time.

As an additional implication of Hypothesis 1, I examine whether coordination stabilizes once the group reaches the threshold $k = \bar{K}$. At the threshold, successful coordination protects the group from the terminal overwrite loss, so deviations to $Y$ should become relatively rare. I treat this protection-stability comparison as a secondary analysis rather than as a separate main hypothesis.

The primary outcome is subject $i$'s stated probability of choosing $X$. For the empirical analysis, I divide the elicited 0-100 probability by 100 and denote the resulting measure by $p_{isgmr}(X) \in [0, 1]$, where $s$ indexes sessions, $g$ groups, $m$ matches, and $r$ rounds. The realized individual action and group outcome are treated as secondary outcomes because they additionally reflect the random draws used to implement subjects' stated probabilities. The empirical analysis follows the order of the hypotheses, examining state dependence, state versus history dependence, and the comparative statics in group size and the payoff from the safe action. The corresponding regression specifications are introduced together with the results in Section 4. For round-level regressions, inference is based on wild cluster bootstrap-$t$ tests clustered at the session level with 5,000 bootstrap replications. The group-size comparisons use session-level randomization inference because group size is assigned at the session level. Tests of the directional hypotheses are one-sided at the 5% significance level, while the test of state versus history dependence is two-sided.

# 4 Results

Before turning to the main results, Table 5 summarizes subject characteristics and elicited preferences by group-size treatment. Participants were randomly assigned to sessions with $n = 2$ or $n = 4$. The table reports treatment-group means and $t$-statistics for differences in means or proportions as descriptive measures of realized balance across the two conditions.

The reported characteristics are broadly similar across group-size treatments. Attendance at an English-speaking high school is included as a proxy for prior exposure to an English-speaking environment and, more broadly, cultural background. Social preference is coded as a binary indicator equal to 1 for Prosocial and 0 for Individualistic. Only one of the 164 treatment participants is classified as Altruistic, and that subject is included in the Prosocial category.

Table 5: Balance of Subject Characteristics and Elicited Preferences

| Subject characteristic | $n = 2$ | | $n = 4$ | | $t$-statistic |
|---|---|---|---|---|---|
| | Mean | (SE) | Mean | (SE) | |
| Gender (1 = Male) | 0.53 | (0.06) | 0.49 | (0.05) | 0.50 |
| Major (1 = Engineering/Science) | 0.57 | (0.06) | 0.57 | (0.05) | 0.01 |
| High school language (1 = English) | 0.85 | (0.04) | 0.88 | (0.03) | -0.41 |
| Risk-averse (1 = Yes) | 0.44 | (0.06) | 0.44 | (0.05) | 0.05 |
| Social Preference (1 = Prosocial) | 0.50 | (0.06) | 0.61 | (0.05) | -1.46 |
| Sessions | 6 | | 6 | | |
| Subjects | 68 | | 96 | | |

*Notes:* Standard errors are shown in parentheses. All characteristics are binary variables. The $t$-statistics for differences in means or proportions are reported as $n = 2$ minus $n = 4$.

## 4.1 Monotonicity in state

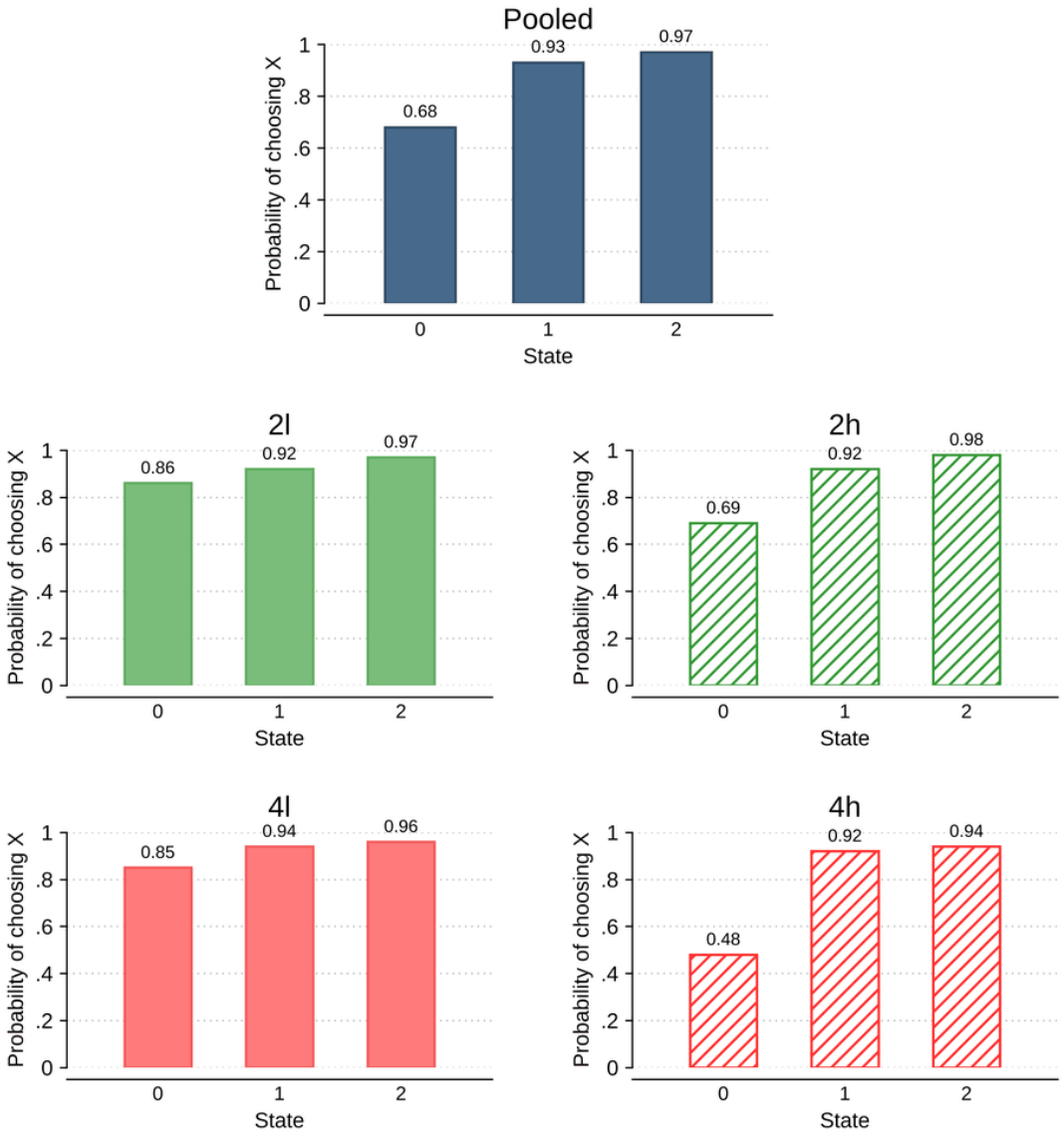
Hypothesis 1 concerns the monotonicity of the probability assigned to $X$ in state $k$. To formally test Hypothesis 1, I estimate

$$p_{isgmr}(X) = \alpha_i + \mu_{sgm} + \lambda_r + \beta_1 1\{k = 1\} + \beta_2 1\{k = 2\} + \varepsilon_{isgmr}, \tag{1}$$

where $\alpha_i$ denotes subject fixed effects, $\mu_{sgm}$ group-by-match fixed effects, and $\lambda_r$ round-within-match fixed effects. The omitted category is state $k = 0$. Thus, $\beta_1$ measures the difference between $k = 1$ and $k = 0$, $\beta_2$ measures the difference between $k = 2$ and $k = 0$, while $\beta_2 - \beta_1$ measures the difference between $k = 2$ and $k = 1$. Hypothesis 1 predicts that both $\beta_1$ and $\beta_2$ are weakly positive and that the increase from $k = 0$ to $k = 1$ exceeds that from $k = 1$ to $k = 2$, corresponding to $2\beta_1 - \beta_2 > 0$. I also estimate a specification replacing the state indicators with the linear state variable $k$. Because the public state $k$ is generated by previous play rather than experimentally assigned, these estimates are interpreted as conditional associations rather than causal effects of $k$.

Figure 3 shows the mean $p(X)$ by state in the pooled sample and across the four treat-

Figure 3: Probability Assigned to $X$ by State




ments.[12] In every treatment, $p(X)$ is lowest at $k = 0$ and close to one at $k = 1$ and $k = 2$. Table 6 relates these differences at $k = 0$ to the theoretical belief cutoffs. Under the experimental parameterization, the $(m11)$ mixing probability $\sigma_0$ coincides with the Grim-Trigger cutoff $p_0^\star$, providing a common state-0 benchmark for the two theoretical approaches.[13] The treatment ordering of observed $p(X)$ at $k = 0$ largely follows the ordering of the dynamic basin: $X$ is chosen most often in $2l$ and least often in $4h$. The main departure is that $4l$ exhibits more $X$ than $2h$, although its basin is slightly smaller (0.71 versus 0.76). The

[12]Figure B2.1 in Appendix B2 shows the corresponding pattern for realized all-$X$ coordination.

[13]See Section 2.2.2 for the $(m11)$ mixed-MPE benchmark and Section 2.3 for the Grim-Trigger cutoff.

comparisons in Figure 3 and Table 6 are descriptive. Because group-matches with higher underlying propensities to choose $X$ are more likely to reach higher states, the fixed-effects regressions below isolate within-group-match state variation and can yield smaller, or even differently signed, state gradients.

Table 6: Theoretical Cutoffs and Observed Behavior at $k = 0$

| Treatment | Belief cutoff $\sigma_0 = p_0^\star$ | Dynamic basin $\mathcal{DB}_0 = 1 - \sigma_0$ | Observed mean $p(X)$ at $k = 0$ |
|---|---|---|---|
| $2l$ | 0.023 | 0.977 | 0.86 |
| $2h$ | 0.236 | 0.764 | 0.69 |
| $4l$ | 0.286 | 0.714 | 0.85 |
| $4h$ | 0.618 | 0.382 | 0.48 |

*Notes:* Theoretical values use $\bar{K} = 2$, $\delta = 0.8$, and $d = 1$. The cutoff $\sigma_0 = p_0^\star$ is the minimum probability assigned to $X$ by each other group member at $k = 0$ for $X$ to be a best response. Observed values are mean elicited probabilities assigned to $X$ at $k = 0$, pooling initial and post-reset observations as in Figure 3.

Table 7: Monotonicity in the Public State

| Contrast | Estimate | Clustered SE | One-sided $p$ | 95% CI |
|---|---|---|---|---|
| *Panel A: Categorical state specification* | | | | |
| $k = 1$ minus $k = 0$ | 0.064 | 0.014 | 0.002 | [ 0.032, 0.095] |
| $k = 2$ minus $k = 0$ | 0.097 | 0.033 | 0.004 | [ 0.027, 0.170] |
| $k = 2$ minus $k = 1$ | 0.033 | 0.022 | 0.073 | [-0.013, 0.084] |
| $(k = 1 - k = 0) - (k = 2 - k = 1)$ | 0.030 | 0.018 | 0.056 | [-0.010, 0.066] |
| *Panel B: Linear state specification* | | | | |
| One-unit increase in $k$ | 0.049 | 0.016 | 0.006 | [ 0.012, 0.085] |
| Observations | | 7,812 | | |
| Subjects | | 164 | | |
| Sessions | | 12 | | |

*Notes:* All specifications include subject, group-by-match, and round-within-match fixed effects. Analytic weights give each of the 12 sessions equal total weight. Standard errors and wild-cluster-bootstrap-$t$ tests are clustered by session; bootstrap inference uses 5,000 replications. Reported $p$-values are one-sided in the direction predicted by Hypothesis 1; confidence intervals are two-sided 95% wild-bootstrap intervals. Alternative specifications replacing subject fixed effects with observed participant controls are reported in Appendix B1, Table B1.1.

Table 7 reports the tests of Hypothesis 1. The probability assigned to $X$ increases by 0.064 from $k = 0$ to $k = 1$ ($p = 0.002$), and by 0.033 from $k = 1$ to $k = 2$ ($p = 0.073$). The overall difference between $k = 0$ and $k = 2$ is 0.097 ($p = 0.004$). The estimated difference between the two adjacent increments is 0.030 in the predicted direction, but falls just short

of statistical significance at the 5% level ($p = 0.056$). Consistent with the overall monotonic pattern, the linear-state specification also yields a positive and statistically significant coefficient of 0.049 ($p = 0.006$).

Table 8: Monotonicity in the Public State by Treatment

| | $2l$ | $2h$ | $4l$ | $4h$ |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *Panel A: Categorical state specification* | | | | |
| $k = 1$ minus $k = 0$ | -0.005 | 0.066** | 0.024** | 0.099*** |
| | (0.036) | (0.014) | (0.014) | (0.033) |
| $k = 2$ minus $k = 0$ | -0.055 | 0.065** | 0.044 | 0.168** |
| | (0.050) | (0.032) | (0.037) | (0.076) |
| $k = 2$ minus $k = 1$ | -0.049 | -0.001 | 0.019 | 0.069 |
| | (0.017) | (0.022) | (0.028) | (0.057) |
| Difference between increments | 0.044* | 0.067** | 0.005 | 0.030 |
| | (0.026) | (0.018) | (0.025) | (0.055) |
| *Panel B: Linear state specification* | | | | |
| One-unit increase in $k$ | -0.026 | 0.034** | 0.022 | 0.084** |
| | (0.025) | (0.016) | (0.018) | (0.037) |
| Observations | 1,620 | 1,616 | 2,288 | 2,288 |
| Subjects | 68 | 68 | 96 | 96 |
| Sessions | 6 | 6 | 6 | 6 |

*Notes:* All specifications include subject, group-by-match, and round-within-match fixed effects. Analytic weights give each of the six sessions represented in a treatment column equal total weight. Session-clustered standard errors are in parentheses. Stars use one-sided wild-cluster-bootstrap-$t$ tests with 5,000 replications in the direction predicted by Hypothesis 1: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Columns denote group size and the low or high safe-option payoff.

Table 8 reports the same specifications separately by treatment. The state relationship varies substantially across treatment conditions. The increase from $k = 0$ to $k = 1$ is positive and statistically significant in 2h, 4l, and 4h, but is close to zero in 2l. In contrast, the increase from $k = 1$ to $k = 2$ is not statistically significant in any treatment. The difference between the two increments is positive in all four treatments, although it is statistically significant at the 5% level only in 2h and at the 10% level in 2l. Consistent with this heterogeneity, the linear-state coefficient is positive and statistically significant in 2h and 4h, but not in 2l or 4l. Overall, the monotonic state relationship is strongest in the high-$u_Y$ treatments.

The observed state pattern is also qualitatively consistent with the ($m11$) mixed-MPE benchmark characterized in Section 2. For the experimental threshold $\bar{K} = 2$, this equilibrium prescribes interior mixing at $k = 0$ and pure $X$ at $k = 1$ and $k = 2$. Consistent with this qualitative pattern, the probability assigned to $X$ is substantially lower at $k = 0$ and rises sharply once the state reaches $k = 1$, with relatively little additional increase from

$k = 1$ to $k = 2$. The observed probabilities at the higher states are nevertheless below one, so the aggregate state pattern should be interpreted as resembling, rather than exactly implementing, the $(m11)$ equilibrium.

As a secondary analysis, Table B1.1 in Appendix B1 examines heterogeneity by elicited risk and social preferences. Risk aversion is not significantly associated with $p(X)$, while prosociality is positively associated with $p(X)$, although the latter relationship is only weakly statistically significant and is sensitive to the inclusion of additional participant controls.

**Result 1:** Consistent with Hypothesis 1, the probability assigned to $X$ increases with the public state $k$ in the pooled sample. The increase from $k = 0$ to $k = 1$ is larger than that from $k = 1$ to $k = 2$ in magnitude, but the difference is not statistically significant at the 5% level. The strength of the state relationship varies across treatments.

## 4.2 State-dependent and history-dependent strategies

Hypothesis 2 distinguishes state-dependent from history-dependent behavior. Under a state-dependent strategy, subjects condition their choices only on the current public state $k$, so two observations with the same value of $k$ should generate similar behavior regardless of the path by which that state was reached. By contrast, a history-dependent strategy allows behavior to depend on previous play even after conditioning on the current state.

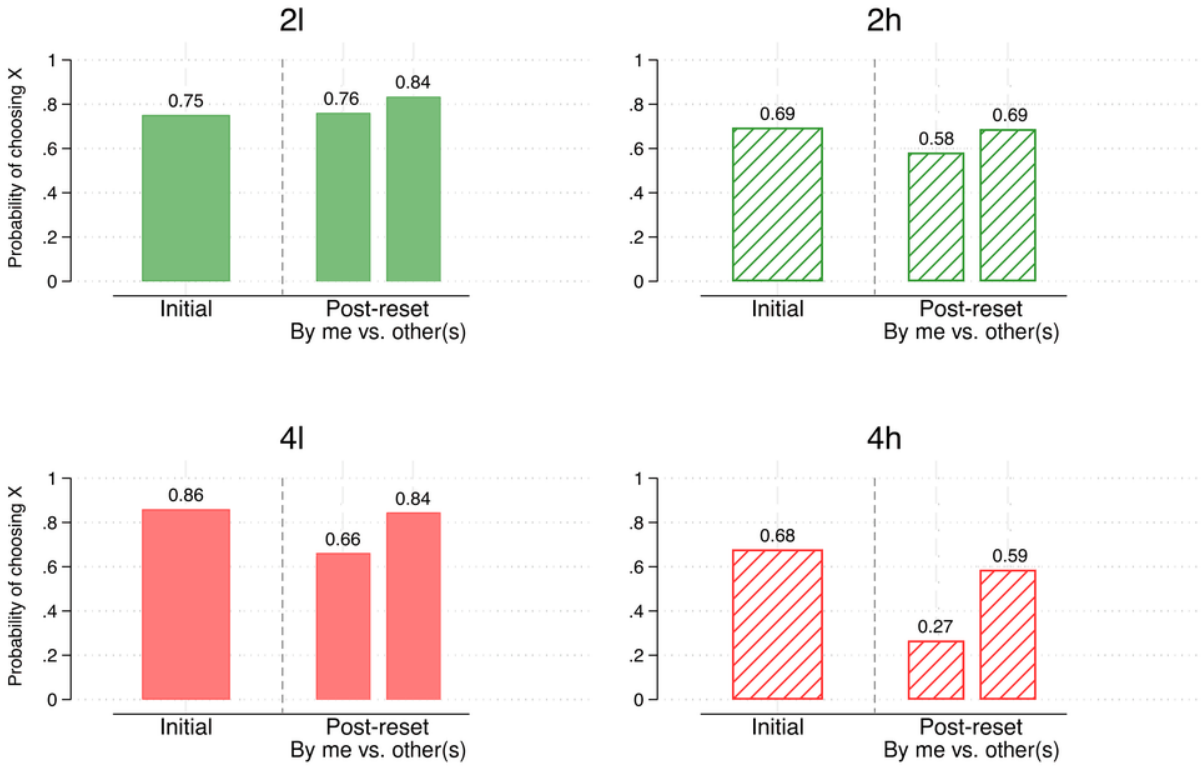
To test Hypothesis 2, I focus on observations with $k = 0$ because this state can arise in two distinct ways: at the beginning of a match, i.e., Round 1, when no previous history has yet accumulated, and after a reset following a failure to coordinate on all-$X$. Comparing these two cases therefore provides a direct test of whether behavior depends on history beyond the information contained in the current state. I estimate

$$p_{isgmr}(X) = \alpha_i + \mu_{sgm} + \theta Reset0_{sgmr} + \varepsilon_{isgmr}, \qquad k = 0, \tag{2}$$

where $Reset0_{sgmr} = 1$ for a $k = 0$ observation that follows a reset caused by other group members, that is, a failed round in which subject $i$ chose $X$ but at least one other member chose $Y$, and $Reset0_{sgmr} = 0$ for the initial $k = 0$ observation. Post-reset observations that follow the subject's own $Y$ are excluded: because the realized action is drawn from the elicited probability, subjects whose own $Y$ caused the reset are selected on a lower propensity to choose $X$, which would confound history dependence with persistence in individual behavior. Under state-dependent behavior, the same state should induce the same behavior regardless of the preceding history, implying $\theta = 0$. Hypothesis 2 is therefore evaluated using a two-sided test of $\theta = 0$. Note that round-within-match fixed effects are not included because

the initial $k = 0$ observation necessarily occurs in Round 1, whereas post-reset observations occur only in later rounds.

Figure 4: Mean $p(X)$ at Initial and Post-Reset $k = 0$, by Source of Reset




*Notes:* Post-reset observations are split by whether the subject chose $Y$ in the failed round (by me) or the subject chose $X$ and at least one other member chose $Y$ (by other(s)). The sample is restricted to group-by-match cells containing both the initial Round-1 $k = 0$ observation and at least one post-reset $k = 0$ observation.

Figure 4 provides a descriptive comparison of initial and post-reset behavior among group-by-match cells containing both the initial Round-1 $k = 0$ observation and at least one post-reset $k = 0$ observation. This restriction is necessary because some matches end after only one round and others never return to $k = 0$ after Round 1, and it ensures that the initial and post-reset means are calculated over the same set of group-by-match cells. Post-reset observations are further split by whether the reset was caused by the subject's own $Y$ or by other group members. After a reset caused by the subject's own $Y$, mean $p(X)$ is lower than at the initial state in $2h$, $4l$, and especially $4h$ (0.27 versus 0.68), and roughly unchanged in $2l$. These observations are selected on a low propensity to choose $X$, so they partly reflect persistence in individual behavior rather than a response to history. After a reset caused by others, mean $p(X)$ is close to its initial level in $2h$ and $4l$, higher in $2l$ (0.84 versus 0.75),

and lower in $4h$ (0.59 versus 0.68). These raw means are in turn selected toward subjects with a high propensity to choose $X$, because the subject chose $X$ in the failed round. The regression analysis therefore compares initial behavior with behavior after resets caused by others, controlling for subject fixed effects.

Table 9 reports the corresponding fixed-effect estimates using the full $k = 0$ estimation sample rather than the restricted sample in Figure 4. In the pooled specification, the post-reset coefficient is $-0.050$ and is statistically significant at the 5% level, so I reject equality between initial and post-reset behavior: after a coordination failure caused by others, subjects assign a 5.0 percentage-point lower probability to $X$ than at the start of the match, even though the payoff-relevant state is the same. The treatment-specific estimates show that this decline is concentrated in the $4h$ treatment, where post-reset $p(X)$ is 20.1 percentage points lower ($p < 0.01$). The estimates in $2l$ and $4l$ are small, the estimate in $2h$ ($-0.089$) is imprecise, and none of the three is statistically significant. The positive raw difference in $2l$ in Figure 4 does not survive the inclusion of subject fixed effects, consistent with the selection described above. Moreover, because a subject is more likely to have chosen $X$ in the failed round when her propensity to choose $X$ is high, any within-subject persistence in this propensity biases the post-reset estimates upward, so the estimated declines are, if anything, conservative. Because each treatment-specific estimate is based on six sessions, these results should be interpreted cautiously.

Table 9: Initial versus Post-Reset Behavior at $k = 0$: Resets Caused by Others

| | Probability assigned to $X$ | | | | |
|---|---|---|---|---|---|
| | Pooled | $2l$ | $2h$ | $4l$ | $4h$ |
| | (1) | (2) | (3) | (4) | (5) |
| Post-reset $k = 0$, by others | -0.050** | -0.019 | -0.089 | -0.016 | -0.201*** |
| | (0.022) | (0.028) | (0.077) | (0.013) | (0.031) |
| Observations | 2,723 | 443 | 467 | 830 | 983 |
| Subjects | 164 | 68 | 68 | 96 | 96 |
| Sessions | 12 | 6 | 6 | 6 | 6 |
| Subject fixed effects | Yes | Yes | Yes | Yes | Yes |
| Group-by-match fixed effects | Yes | Yes | Yes | Yes | Yes |

*Notes:* The sample contains initial $k = 0$ observations and post-reset $k = 0$ observations for which the participant chose $X$ in the preceding round, so one or more other group members caused the reset. Initial $k = 0$ is the omitted category. All specifications include subject and group-by-match fixed effects. Analytic weights give each eligible session–treatment cell equal total weight. Session-clustered standard errors are in parentheses. Stars use two-sided wild-cluster-bootstrap-$t$ tests with 5,000 replications: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The pattern is consistent with the size of the dynamic basin. The largest and only significant decline occurs in $4h$, the treatment with the smallest basin at $k = 0$ (Table 6),

whereas the estimates in the three treatments with larger basins are small or imprecise. One interpretation is that a coordination failure caused by others leads subjects to revise downward their beliefs about their current partners. When the dynamic basin is small, as in $4h$, choosing $X$ already requires optimistic beliefs, so even a modest downward revision can make $Y$ the better response; when the basin is larger, the same revision is less likely to overturn $X$ as a best response. With only four treatment conditions, this pattern is descriptive rather than a formal test. Even after a failure, behavior rarely resembles the permanent switch to $Y$ prescribed by Grim Trigger: in $4h$, where the post-reset decline is largest, the mean probability assigned to $X$ after a reset caused by others remains 0.59 (Figure 4).[14]

**Result 2:** Behavior is not fully summarized by the current state. After a reset caused by others, the probability assigned to $X$ is significantly lower than at the initial $k = 0$ state, with the decline concentrated in $4h$, the treatment with the smallest dynamic basin. Post-reset behavior rarely resembles the permanent switch to $Y$ prescribed by Grim Trigger.

## 4.3 Comparative statics in $u_Y$ and $n$

I next examine the comparative statics in group size $n$ and the payoff from the safe action $u_Y$. The experimental design varies group size between sessions and $u_Y$ within subjects, so the two treatment dimensions are analyzed separately. For each comparison, I report three samples. The first pools all rounds and therefore all states. Because the state is generated by play, the distribution of states can differ across treatments, so this comparison mixes the effect of the treatment at a given state with differences in how often each state is visited. Since the theoretical comparative statics hold state by state, the second sample holds the state fixed at $k = 0$, including both initial and post-reset observations. The third sample uses only the first round of each match, which is always at $k = 0$ and precedes any history within the match.

Hypothesis 3a predicts a higher probability of choosing $X$ in smaller groups, i.e., when group size $n = 2$. Because group size is assigned at the session level, I conduct the primary comparison at the session level, separately for each value of $u_Y$. For each subject, I first calculate the mean probability assigned to $X$ under a given payoff condition and then average

[14] I classify a participant-match as a strict Grim Trigger case if the participant assigns $p(X) = 1$ in every round through the first non-all-$X$ round and $p(X) = 0$ in every remaining round. Among the 826 participant-matches with a failure before the final round, the only ones in which such a switch can be observed, 46 (5.6%) are strict Grim Trigger cases: 0 in $2l$, 3 in $2h$, 4 in $4l$, and 39 in $4h$ (11.3% of informative failures in that treatment). These cases come from 30 of 164 participants, about 1.5 matches per participant on average, so they reflect match-level patterns rather than stable Grim Trigger types.

these subject-level means within each session. Let $\bar{p}_{su}$ denote this session-level mean for payoff condition $u \in \{l, h\}$. I estimate

$$\bar{p}_{su} = \alpha_u + \beta_{n,u} 1\{n_s = 2\} + \varepsilon_{su}, \qquad u \in \{l, h\}, \tag{3}$$

where $\beta_{n,l}$ compares treatments $2l$ and $4l$, while $\beta_{n,h}$ compares treatments $2h$ and $4h$. Hypothesis 3a predicts $\beta_{n,u} > 0$. Inference is based on session-level randomization tests.

Table 10: Effect of Group Size on the Probability Assigned to $X$

| | Low $u_Y$ | | High $u_Y$ | |
|---|---|---|---|---|
| | Unadjusted | Adjusted | Unadjusted | Adjusted |
| | (1) | (2) | (3) | (4) |
| *Panel A: All rounds* | | | | |
| $n = 2$ minus $n = 4$ | 0.012 | 0.012 | 0.240** | 0.240** |
| | (0.036) | (0.023) | (0.086) | (0.085) |
| *Panel B: $k = 0$ only* | | | | |
| $n = 2$ minus $n = 4$ | 0.010 | 0.010 | 0.247** | 0.247** |
| | (0.037) | (0.026) | (0.089) | (0.090) |
| *Panel C: Round 1 only* | | | | |
| $n = 2$ minus $n = 4$ | −0.023 | −0.023 | 0.099 | 0.099 |
| | (0.041) | (0.022) | (0.078) | (0.081) |
| High-first | No | Yes | No | Yes |
| Sequence fixed effects | No | Yes | No | Yes |
| Sessions | 12 | 12 | 12 | 12 |

*Notes:* The outcome is the session mean of subject-level mean $p(X)$ under each safe-option payoff. Panel B includes both initial and post-reset $k = 0$ rounds. Standard errors are in parentheses. Stars use one-sided exact session-level randomization tests based on the 64 assignments obtained by swapping the $n = 2$ and $n = 4$ labels within each of the six sequence-by-order pairs: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 10 reports the results for the three samples. Columns (2) and (4) additionally control for payoff order (High-first) and predetermined termination-sequence fixed effects. Coefficients are reported as $n = 2$ minus $n = 4$, so positive estimates support Hypothesis 3a. When $u_Y$ is low, the estimated difference between $n = 2$ and $n = 4$ is small and not statistically significant in any sample, ranging from $-2.3$ to 1.2 percentage points. In contrast, when $u_Y$ is high, participants in groups of two assign a 24.0 percentage-point higher probability to $X$ than participants in groups of four over all rounds (Panel A) and a 24.7 percentage-point higher probability at $k = 0$ (Panel B). Both differences are statistically significant at the 5% level and unchanged after adding the controls. In Round 1 (Panel C), the difference falls to 9.9 percentage points and is not statistically significant, although its

standard error is similar to those in Panels A and B. Because the $k = 0$ sample combines Round 1 with post-reset rounds, the larger gap at $k = 0$ comes from rounds that follow a coordination failure, consistent with the post-reset decline in the $4h$ treatment (Section 4.2).

**Result 3a:** The group-size comparative static in Hypothesis 3a is supported when the safe payoff is high, but not when it is low. Under high $u_Y$, smaller groups assign substantially higher probabilities to $X$ over all rounds and at $k = 0$, but not significantly so in Round 1. Under low $u_Y$, there is no detectable group-size difference in any sample.

Hypothesis 3b predicts that lowering the payoff from the safe action increases the probability assigned to $X$. Since each subject experiences both payoff conditions, I exploit within-subject variation and estimate the following specification separately for $n = 2$ and $n = 4$:

$$p_{isgmr}(X) = \alpha_i + \lambda_m + \beta_{Y,n} LowY_{sm} + \varepsilon_{isgmr}, \qquad n \in \{2, 4\}, \tag{4}$$

where $LowY_{sm}$ equals one when $u_Y = 0.15$, $\alpha_i$ denotes subject fixed effects, and $\lambda_m$ denotes match-number fixed effects. Thus, $\beta_{Y,n}$ measures the within-subject effect of moving from the high to the low safe-option payoff. Hypothesis 3b predicts $\beta_{Y,n} > 0$.

Table 11 reports the estimates separately for $n = 2$ and $n = 4$ in the same three samples. Positive coefficients indicate a higher probability assigned to $X$ when $u_Y = 0.15$ than when $u_Y = 0.85$. Columns (1), (2), (4), and (5) exploit within-subject variation in $u_Y$. Columns (1) and (4) include absolute match-number fixed effects, where the absolute match number is the match number in Part 1 (1-10). Columns (2) and (5) instead include relative match-number fixed effects, where the relative match number is the position within each five-match payoff block (1-5), so Matches 6-10 are recoded as 1-5 after the safe-option payoff changes. As a robustness check for potential order or carryover effects, Columns (3) and (6) provide a between-session comparison using only Matches 1-5, before participants have been exposed to the alternative payoff condition.

Columns (1) and (4) report the primary within-subject estimates. Over all rounds (Panel A), lowering $u_Y$ increases the probability assigned to $X$ by 0.100 in groups of two and by 0.327 in groups of four, with both effects statistically significant at the 1% level. The effects are larger at $k = 0$ (Panel B), at 0.146 and 0.354, consistent with behavior being close to pure $X$ at higher states under both payoffs, which dilutes the all-rounds comparison. They are smaller in the first round of each match (Panel C), at 0.061 and 0.183, but remain significant at the 1% level. As with group size, the payoff effect therefore appears largest in rounds that follow a coordination failure. With relative match-number fixed effects in Columns (2)

Table 11: Effect of the Safe-Option Payoff on the Probability Assigned to $X$

| | $n = 2$ | | | $n = 4$ | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Panel A: All rounds* | | | | | | |
| Low $u_Y$ minus High $u_Y$ | 0.100*** | 0.100* | 0.151 | 0.327*** | 0.327*** | 0.211 |
| | (0.028) | (0.064) | (0.037) | (0.036) | (0.053) | (0.085) |
| Observations | 3,236 | 3,236 | 1,580 | 4,576 | 4,576 | 2,240 |
| *Panel B: $k = 0$ only* | | | | | | |
| Low $u_Y$ minus High $u_Y$ | 0.146*** | 0.150* | 0.262 | 0.354*** | 0.367*** | 0.249 |
| | (0.045) | (0.080) | (0.051) | (0.027) | (0.050) | (0.048) |
| Observations | 1,268 | 1,268 | 744 | 2,732 | 2,732 | 1,308 |
| *Panel C: Round 1 only* | | | | | | |
| Low $u_Y$ minus High $u_Y$ | 0.061*** | 0.061 | 0.099 | 0.183*** | 0.183*** | 0.094 |
| | (0.018) | (0.078) | (0.015) | (0.037) | (0.041) | (0.065) |
| Observations | 680 | 680 | 340 | 960 | 960 | 480 |
| Subject fixed effects | Yes | Yes | | Yes | Yes | |
| Absolute match-number FE | Yes | | Yes | Yes | | Yes |
| Relative match-number FE | | Yes | | | Yes | |
| First-half sample | | | Yes | | | Yes |
| Subjects | 68 | 68 | 68 | 96 | 96 | 96 |
| Sessions | 6 | 6 | 6 | 6 | 6 | 6 |

*Notes:* Panel B includes both initial and post-reset $k = 0$ rounds. Session-clustered standard errors are in parentheses. Stars use one-sided wild-cluster-bootstrap-$t$ tests in columns (1), (2), (4), and (5), and one-sided exact session-level permutation tests in columns (3) and (6). First-half columns use equal-session weights and Matches 1–5, before exposure to the alternative payoff: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

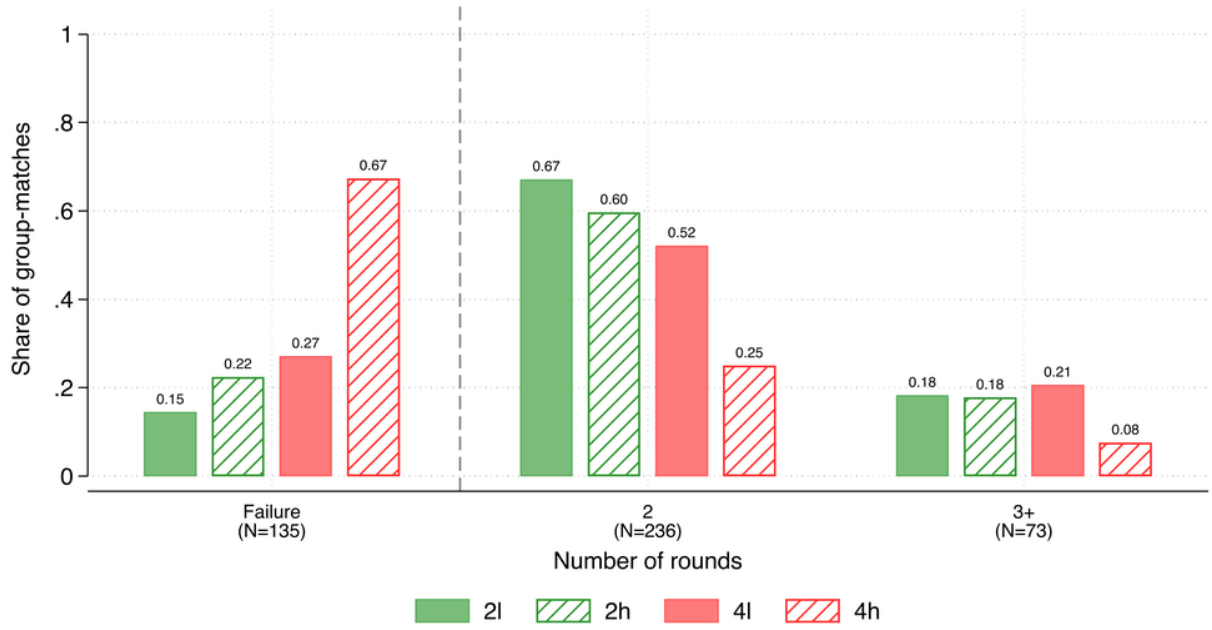
and (5), the point estimates are similar; the estimate for $n = 4$ remains significant at the 1% level in every sample, whereas the estimate for $n = 2$ is significant only at the 10% level in Panels A and B and is not significant in Panel C. In the first-half sample in Columns (3) and (6), the estimated effects remain positive in every panel, ranging from 0.094 to 0.262, but none is statistically significant under the exact session-level permutation tests.

**Result 3b:** Consistent with Hypothesis 3b, reducing the safe payoff significantly increases the probability assigned to $X$ for both group sizes in the primary within-subject specifications, in all three samples. The effect is larger in groups of four and smaller in Round 1 than at $k = 0$. The corresponding first-half estimates are positive but not statistically significant.

## 4.4 Persistence of Coordination

As a secondary implication of Hypothesis 1, I examine whether coordination remains stable once a group reaches the threshold $k = 2$. Because unanimous $X$ is socially efficient at every state, reaching the threshold corresponds to sustaining the efficient coordination path for two consecutive rounds. I therefore examine how quickly groups reach the threshold and whether all-$X$ coordination persists thereafter.

Figure 5: Number of Rounds to First Reach the Threshold

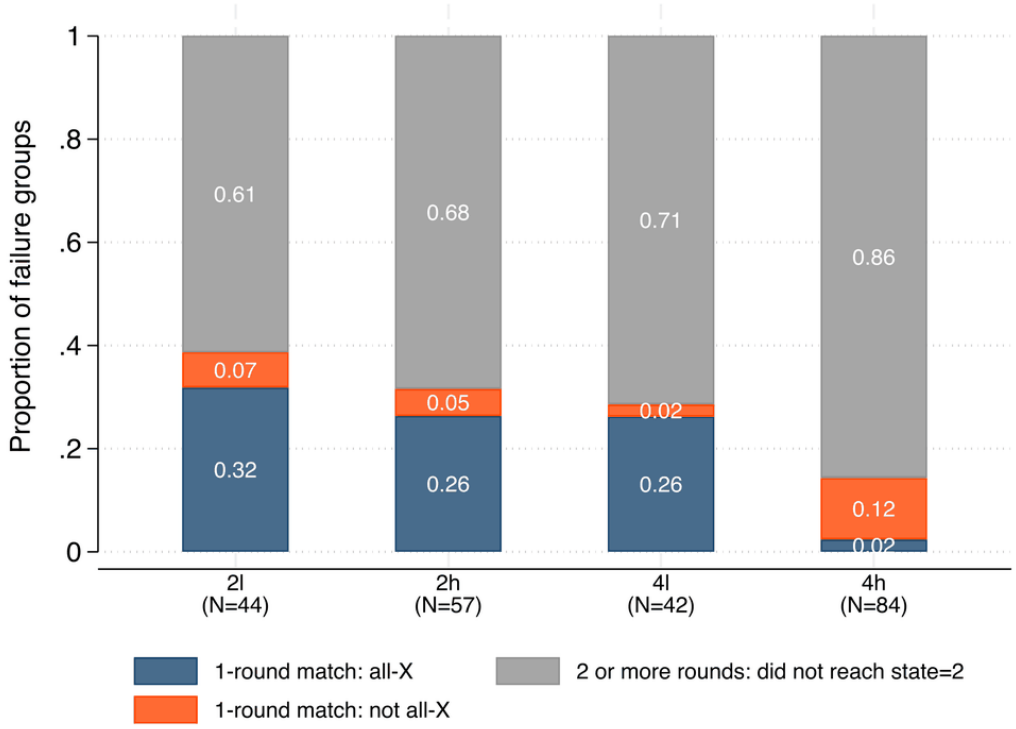



*Notes:* The sample includes only matches lasting at least three rounds.

Figure 5 shows how quickly group-matches reach the threshold, $k = 2$. To separate coordination failure from early random termination, the figure includes only matches that last at least three rounds, in which a group can still reach $k = 2$ after one failed round. Among group-matches that reach the threshold, coordination typically emerges rapidly: in all four treatments, more than 70% reach $k = 2$ in the minimum possible two rounds. The rate of failure to reach the threshold, however, varies sharply across treatments: it is 15%, 22%, and 27% in the $2l$, $2h$, and $4l$ treatments, but 67% in the $4h$ treatment, which has the smallest dynamic basin in Figure 2. Failure rates thus rise as the dynamic basin at $k = 0$ shrinks. Because the sample excludes matches that end too early, the high failure rate in $4h$ cannot be attributed to random termination.

Figure 6 decomposes the groups that do not reach $k = 2$, where some are mechanical because a match that ends after one round cannot reach the threshold in the full sample.

Figure 6: Composition of Groups That Do Not Reach the Threshold



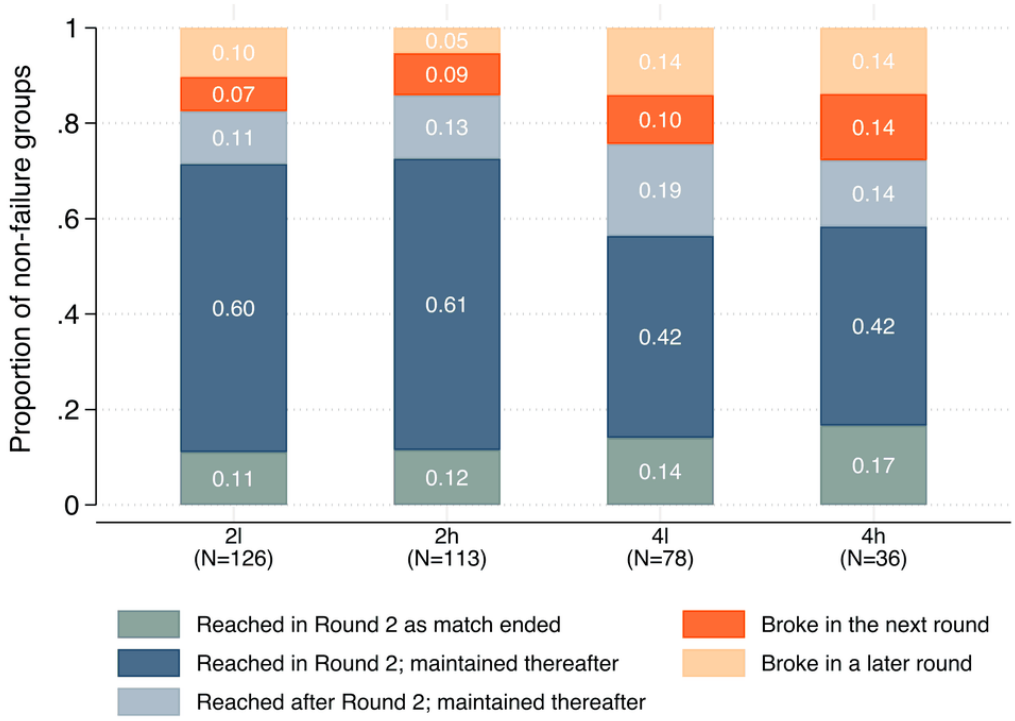

However, in every treatment, the majority of non-attainment occurs in matches lasting at least two rounds. This pattern is particularly pronounced in the $4h$ treatment, where 86% of the group-matches that do not reach $k = 2$ last for at least two rounds. Thus, the relatively high rate of non-attainment in this treatment cannot be explained solely by early random termination.[15]

Figure 7 examines behavior after a group reaches the threshold. Most threshold-reaching group-matches maintain all-$X$ coordination for the remainder of the match: 71%, 74%, 61%, and 56% in the 2l, 2h, 4l, and 4h treatments, respectively. An additional 11-17% reach $k = 2$ only in the final round, so subsequent persistence cannot be observed. Conditional on the match continuing for at least one additional round after the threshold is first reached, approximately 67-84% maintain coordination through the end of the match. The $4h$ treatment also has the lowest rate of maintaining all-$X$ coordination through the remainder of the match.[16]

**Result 4:** Consistent with the secondary implication of Hypothesis 1, coordination is highly persistent once groups reach the threshold. Successful groups typically reach $k = 2$

[15]The corresponding decomposition by experimental half is reported in Appendix B1, Figure B1.1.
[16]The corresponding decomposition by experimental half is reported in Appendix B1, Figure B1.2.

Figure 7: Outcomes after Reaching the Threshold




in the minimum possible two rounds, and most observable post-threshold paths maintain coordination through the end of the match. Failure to reach the threshold is most frequent in the 4h treatment.

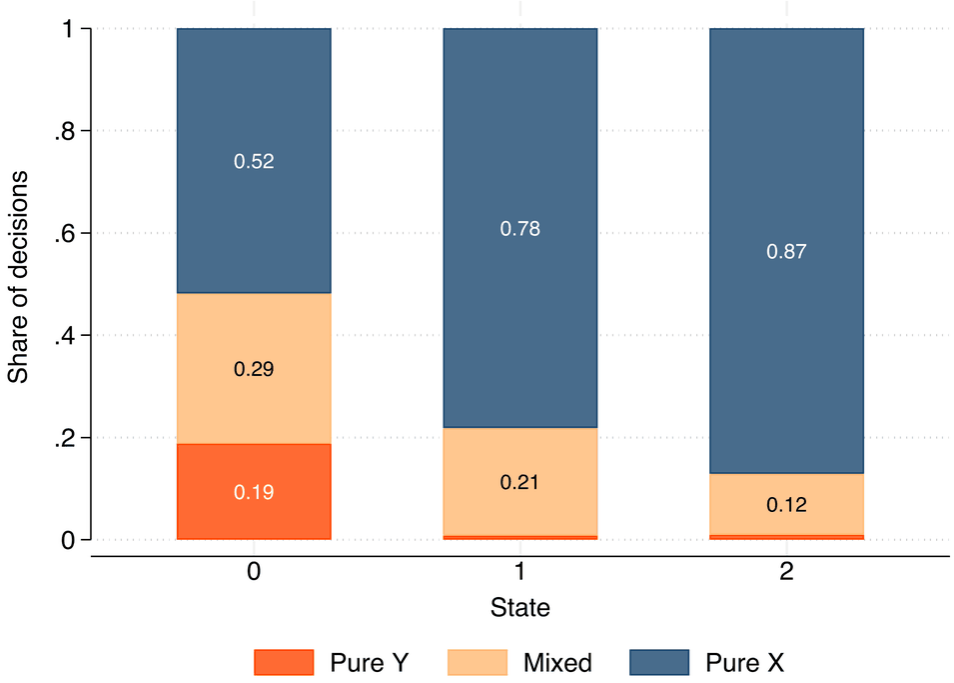
## 4.5 Pure and Mixed Decisions

A feature of the experimental design is that subjects report a probability of choosing $X$, rather than making a binary choice between actions $X$ and $Y$. This elicitation can distinguish pure decisions, $p(X) = 0$ or $p(X) = 1$, from mixed behavior, $0 < p(X) < 1$. Such behavior would not be directly observable from binary choices alone, since an observed action does not reveal whether it was chosen with certainty or generated from an interior choice probability.

Figure 8 shows the shares of pure $Y$, mixed, and pure $X$ decisions at each state. Mixing is most common at $k = 0$, where about 30% of decisions involve interior probabilities. As the state increases, the share of mixed decisions falls to 12% at $k = 2$, while pure $X$ rises from 52% to 87% and pure $Y$ becomes rare. Thus, higher states are associated with a shift from mixed choices toward pure $X$.

The state comparison does not reveal whether mixed behavior also changes as subjects gain experience with the game. I therefore next examine how the prevalence of pure and

Figure 8: Pure and Mixed Decisions by State




mixed choices evolves across matches, both in the pooled sample and separately across treatments. Figure 9 shows how pure and mixed choice behavior evolves over the experiment. Mixed choices are common initially, accounting for about one-third of decisions in Match 1, but their share declines over subsequent matches. Over the same period, the share of pure $X$ choices increases substantially, whereas the share of pure $Y$ choices remains relatively small. Thus, the decline in mixing is primarily associated with convergence toward pure $X$ rather than pure $Y$.

Figure 10 separates the evolution of mixed behavior by treatment. The decline in mixing is not concentrated in a single treatment: mixed choices generally become less frequent over matches. However, both their initial prevalence and the magnitude of the decline vary across treatments. The treatment-specific patterns therefore suggest a common experience-related decline in mixing with substantial heterogeneity across experimental conditions.

Table 12 examines learning in pure and mixed behavior over repeated play. I use match number within each five-match treatment block as a measure of learning; the variable runs from 1 to 5 and resets after $u_Y$ changes following Match 5. Columns (1)-(3) use an indicator for a mixed decision as the dependent variable, while columns (4)-(6) use an indicator for a pure $X$ decision. Across specifications, mixed behavior declines with learning. In the preferred specification in column (3), an additional match within the block is associated

Figure 9: Proportion of Decisions by Match

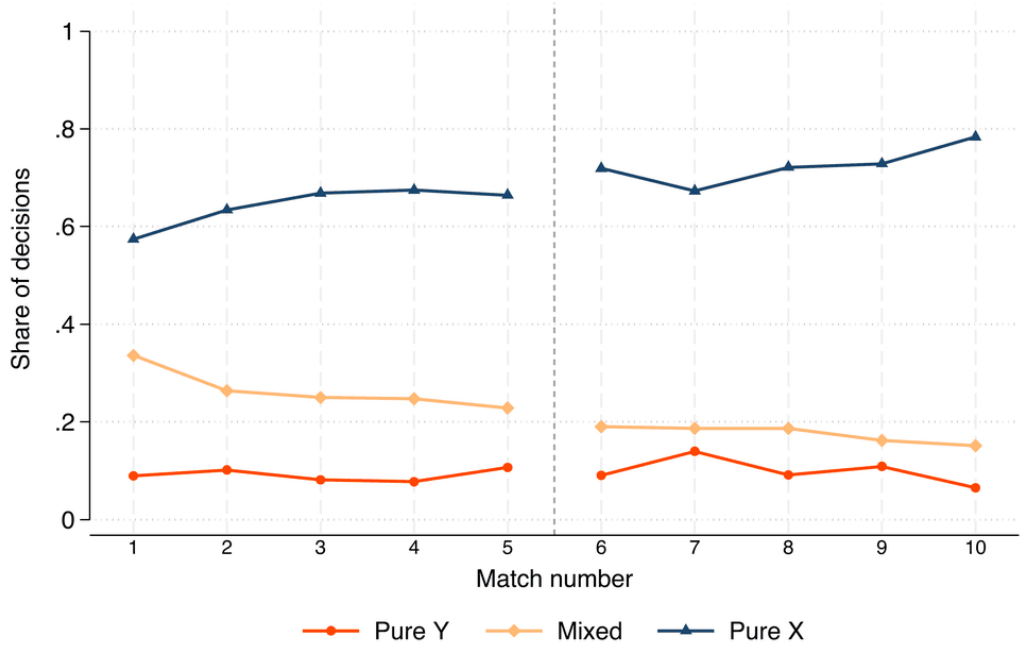
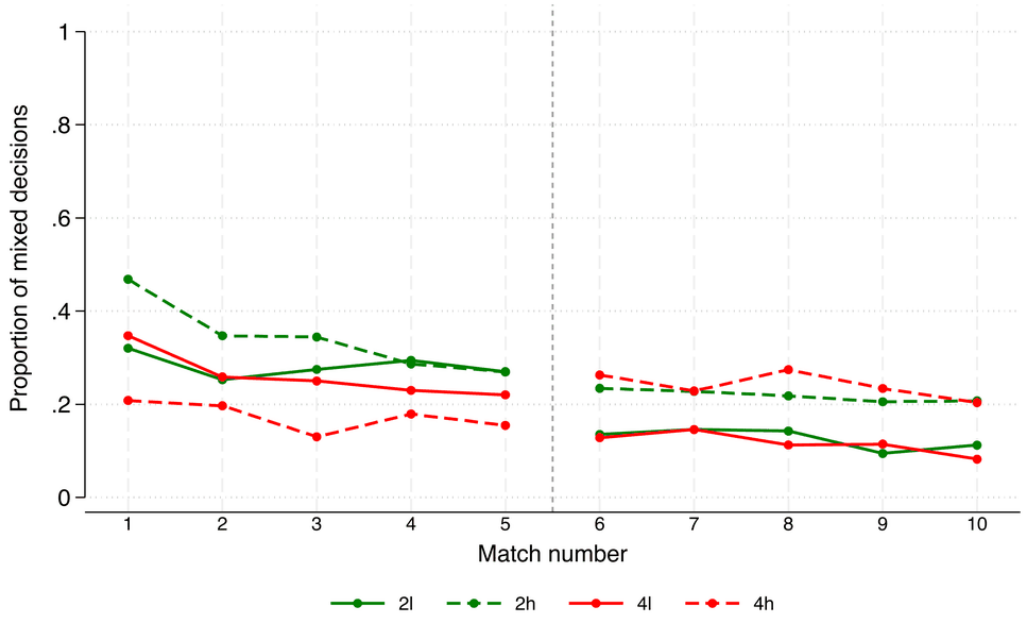



Figure 10: Proportion of Mixed Decisions by Match and Treatment




with a 0.9 percentage-point reduction in the probability of mixing ($p < 0.05$). Mixing also declines substantially with the public state: relative to $k = 0$, the probability of mixing is 5.4 percentage points lower at $k = 1$ and 8.1 percentage points lower at $k = 2$.

The corresponding estimates for pure $X$ behavior show a different pattern. Conditional

Table 12: Learning and Mixed Decisions

| | Mixed decision | | | Pure X decision | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Match within block | -0.007* | -0.006* | -0.009** | 0.005 | 0.004 | 0.005 |
| | (0.004) | (0.003) | (0.004) | (0.009) | (0.005) | (0.005) |
| State = 1 | | -0.067*** | -0.054*** | | 0.306*** | 0.258*** |
| | | (0.027) | (0.019) | | (0.034) | (0.025) |
| State = 2 | | -0.106*** | -0.081*** | | 0.426*** | 0.360*** |
| | | (0.034) | (0.019) | | (0.049) | (0.040) |
| Low $u_Y$ | | | -0.031 | | | 0.145*** |
| | | | (0.025) | | | (0.033) |
| Second half | | | -0.083*** | | | 0.056 |
| | | | (0.026) | | | (0.035) |
| Subject FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Round-within-match FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Equal-session weights | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 7,812 | 7,812 | 7,812 | 7,812 | 7,812 | 7,812 |
| $R^2$ | 0.620 | 0.628 | 0.638 | 0.442 | 0.542 | 0.567 |

*Notes:* Session-clustered standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Analytic weights give every session equal total weight. Stars use two-sided wild-cluster-bootstrap-$t$ tests with 5,000 replications. All specifications include subject and round-within-match fixed effects. State = 0, High $u_Y$, and the first half (Matches 1-5) are the omitted categories.

on the public state and treatment controls, the within-block learning trend is small and statistically insignificant. In contrast, the state effects are large: relative to $k = 0$, the probability of choosing pure $X$ is 25.8 percentage points higher at $k = 1$ and 36.0 percentage points higher at $k = 2$. Thus, learning is associated with a declining propensity to mix, whereas the shift toward pure $X$ is primarily associated with reaching higher public states.

This composition of choices provides a more direct comparison with the $(m11)$ benchmark than the mean-$p(X)$ results alone. The $(m11)$ equilibrium predicts interior mixing at $k = 0$ followed by pure $X$ at $k = 1$ and $k = 2$. The data display the same qualitative movement: interior probabilities are most prevalent at $k = 0$, while pure-$X$ choices become increasingly dominant at higher states. The correspondence is not exact, however, because mixing remains present at $k = 1$ and $k = 2$. Thus, the evidence is consistent with movement toward the $(m11)$ state-contingent pattern, but does not indicate exact play of that equilibrium.

**Result 5:** Mixed behavior declines with both the public state and learning over repeated play. Higher states are associated with a shift toward pure $X$, while the learning effect primarily appears as a reduction in mixing.

## 4.6 Baseline comparison

A natural question is what would happen in a standard repeated stag-hunt setting without the public state or the last-round penalty rule. To provide a descriptive benchmark, I conducted two additional $n = 4$ baseline sessions with the same stage-game payoffs and random-termination procedure, but without the public counter or the last-round payoff penalty.[17] The treatment and baseline samples are similar in observable characteristics, although the small number of baseline sessions limits the informativeness of formal balance comparisons.[18]

Table 13 reports the corresponding descriptive outcome comparisons. When $u_Y = 0.15$, both conditions exhibit high rates of $X$ choice and coordination. Mean $p(X)$ is 0.906 in the treatment and 0.967 in the baseline, while the frequency of all-$X$ coordination is 0.719 and 0.854, respectively. Thus, when the safe payoff is low, coordination on $X$ is already close to the upper bound in the absence of the treatment mechanism. The comparison is markedly different when $u_Y = 0.85$. Mean $p(X)$ is 0.577 in the treatment compared with 0.367 in the baseline, a difference of 0.210. Similarly, pure $X$ choices are 18.7 percentage points more frequent and all-$X$ coordination is 16.4 percentage points more frequent in the treatment. In contrast, the frequency of mixed choices is nearly identical across the two conditions. These suggest that the treatment-baseline difference under high $u_Y$ primarily reflects greater willingness to choose $X$ rather than a change in the prevalence of mixed behavior.

Table 13: Descriptive Comparison of $n = 4$ Treatment and Baseline

| | Low $u_Y$ | | | High $u_Y$ | | |
|---|---|---|---|---|---|---|
| Outcome | Treatment | Baseline | Difference | Treatment | Baseline | Difference |
| Mean $p(X)$ | 0.906 | 0.967 | -0.061 | 0.577 | 0.367 | 0.210 |
| Pure $X$: $p(X) = 1$ | 0.780 | 0.829 | -0.049 | 0.475 | 0.288 | 0.187 |
| Mixed: $0 < p(X) < 1$ | 0.188 | 0.168 | 0.020 | 0.210 | 0.203 | 0.007 |
| all-$X$ coordination | 0.719 | 0.854 | -0.135 | 0.300 | 0.136 | 0.164 |
| Sessions | 6 | 2 | | 6 | 2 | |

*Notes:* All entries are descriptive. Outcomes are first averaged within an independent session and then averaged across sessions, so each session receives equal weight. Pure and mixed choices are measured at the subject-round level. All-$X$ coordination is measured once per group-round. Difference is treatment minus baseline.

How do these treatment-baseline differences evolve within and across matches? Figure 11 plots mean $p(X)$ by round within a match. When $u_Y = 0.15$, $X$ choice remains high throughout the match in both conditions, consistent with the ceiling-level behavior reported

[17] Appendix C1 includes the Experimental Instructions and Interfaces for both $n = 2$ and $n = 4$ baseline sessions, but only $n = 4$ baseline sessions were conducted due to budget constraints.

[18] See Table B1.3 in Appendix B1.

in Table 13. When $u_Y = 0.85$, however, the two conditions diverge sharply after the first round. In the baseline, $p(X)$ falls rapidly and remains around 0.3 in later rounds, whereas the decline is substantially smaller in the treatment. The aggregate difference between treatment and baseline under high $u_Y$ therefore appears to emerge primarily through different within-match dynamics.

Figure 11: Comparison of $p(X)$ by Round between Treatment and Baseline Sessions

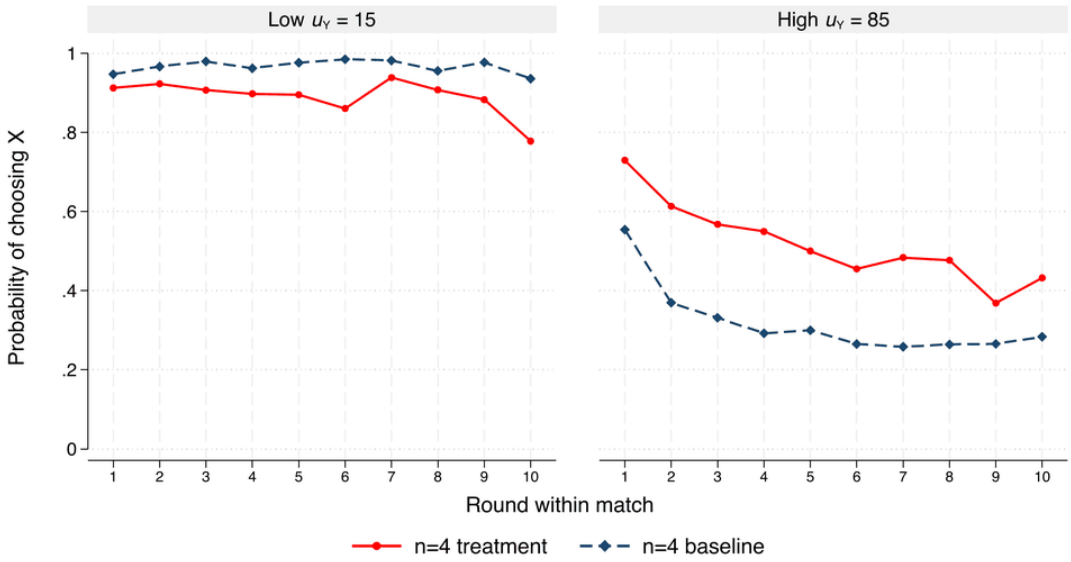



Figure 12: Comparison of $p(X)$ by Match between Treatment and Baseline Sessions

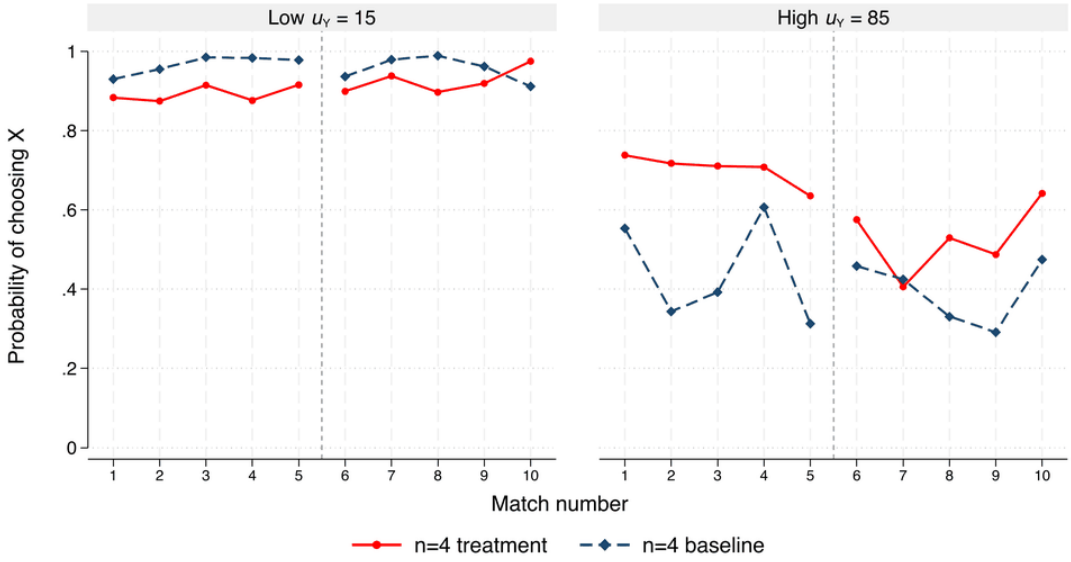
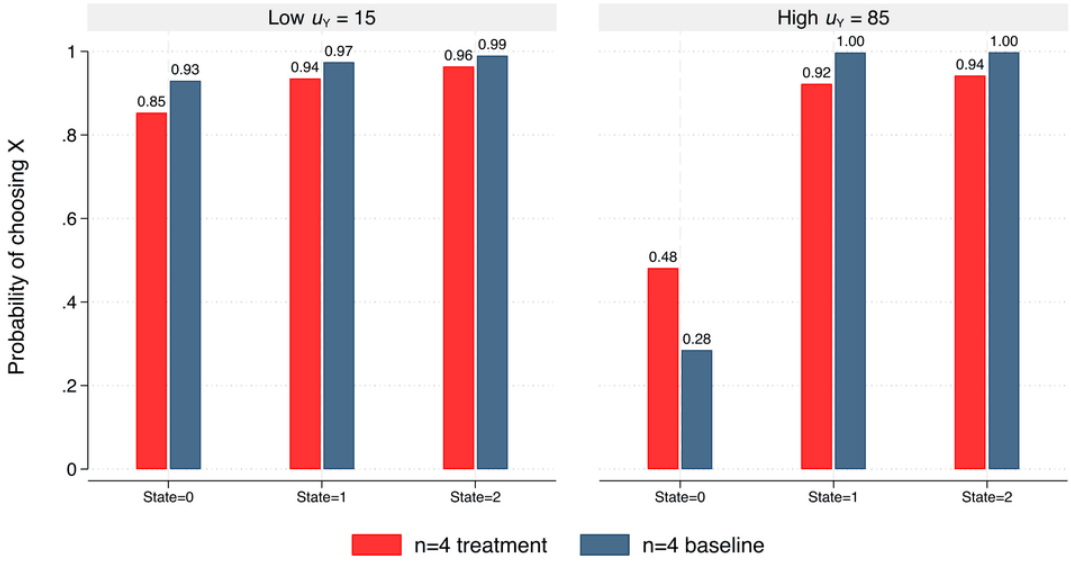



Figure 12 instead plots behavior across match numbers. The across-match pattern is less

systematic. Under low $u_Y$, both conditions remain close to the upper bound throughout the experiment. Under high $u_Y$, $p(X)$ is generally higher in the treatment than in the baseline, although both series fluctuate across matches. Compared with the pronounced pattern within matches, there is little evidence of a smooth treatment-baseline divergence over repeated matches.

Figure 13: Comparison of $p(X)$ by State between Treatment and Baseline Sessions




The preceding figures describe how the treatment-baseline difference evolves. To examine more directly where this difference arises in the coordination process, Figure 13 compares behavior by recent coordination history. For the baseline sessions, I construct an "as-if" state using the same transition rule as in the treatment: each match begins at $k = 0$, the state increases following an all-$X$ round up to $k = 2$, and otherwise resets to zero. This constructed state was not observed by baseline subjects and is used only to summarize recent coordination history. When $u_Y = 0.15$, $p(X)$ is high in both conditions at every state. When $u_Y = 0.85$, however, the treatment-baseline difference is concentrated at $k = 0$: the mean $p(X)$ is 0.48 in the treatment compared with 0.28 in the baseline. At $k = 1$ and $k = 2$, $p(X)$ is above 0.9 in both conditions. Thus, the descriptive difference under high $u_Y$ appears primarily when groups have not yet established recent all-$X$ coordination; conditional on having recently coordinated on all-$X$, behavior is highly persistent in both conditions.

**Result 6:** Relative to the baseline, the treatment is associated with greater $X$ choice when the safe payoff is high, with the difference concentrated at $k = 0$. When the safe payoff is low, $X$ choice is near the upper bound in both conditions.

## 4.7 Self-Reported Strategies

The post-experiment survey asked participants about the information they used when making decisions, the consistency of their strategies across matches, and whether experience in earlier matches affected their later strategies.

Table 14: Self-Reported Strategies in the Post-Experiment Survey

| Response | Treatment ($N = 164$) | Baseline ($N = 32$) |
|---|---|---|
| *Within-match decisions based on the current counter and/or previous group outcomes* | | |
| Current counter only | 34 (20.7%) | — |
| Group outcomes in previous rounds only | 62 (37.8%) | — |
| Both current counter and previous outcomes | 68 (41.5%) | — |
| *Within-match decisions based on previous group outcomes* | | |
| Yes | — | 24 (75.0%) |
| No | — | 8 (25.0%) |
| *Strategy consistency within five-match blocks* | | |
| Same in both blocks | 98 (59.8%) | 14 (43.8%) |
| Same in Matches 1-5 only | 25 (15.2%) | 11 (34.4%) |
| Same in Matches 6-10 only | 24 (14.6%) | 6 (18.8%) |
| Different in both blocks | 17 (10.4%) | 1 (3.1%) |
| *Strategies in Matches 6-10 influenced by Matches 1-5* | | |
| Yes | 97 (59.1%) | 16 (50.0%) |
| No | 67 (40.9%) | 16 (50.0%) |

*Notes:* Entries are participant counts, with column percentages in parentheses. The first question differs across conditions because baseline sessions did not display the public counter.

As shown in Table 14, participants generally reported paying attention to past group outcomes rather than relying only on the current public state. In the treatment sessions, 37.8% reported using previous group outcomes only and another 41.5% reported using both previous outcomes and the current counter; only 20.7% reported relying on the counter alone. Similarly, 75.0% of baseline participants reported that their within-match decisions depended on previous group outcomes. Strategies also exhibit some reported persistence across repeated matches. In the treatment sessions, 59.8% of participants reported using the same strategy across matches within both five-match blocks, while 59.1% reported that their strategies in Matches 6-10 were influenced by their experience in Matches 1-5. The corresponding share for the latter question is 50.0% in the baseline sessions. Taken together, the survey responses suggest that subjects consciously attended to both the current decision environment and their prior coordination experience, complementing the behavioral evidence

on state dependence, history dependence, and learning reported above.

# 5 Discussion and Conclusion

This paper shows that coordination in a dynamic stag hunt depends systematically on accumulated progress. As groups move toward the protection threshold, subjects assign higher probabilities to the risky action and coordination becomes more persistent. The effects of the safe-option payoff and, in strategically difficult environments, group size are also consistent with the comparative statics of the dynamic basin. At the same time, lower coordination after failures caused by others and incomplete convergence to pure $X$ decisions show that the public state does not fully summarize behavior.

The first main finding is that the public state is strongly associated with coordination behavior (Results 1 and 4). In the pooled sample, the probability assigned to $X$ increases as the group moves toward the threshold, with the largest increase occurring between $k = 0$ and $k = 1$. The stronger prediction that the first state increment exceeds the second receives suggestive support in the pooled analysis, although the difference is not statistically significant at the 5% level. The treatment-specific estimates also reveal substantial heterogeneity. Once groups reach the threshold, however, coordination is highly persistent: most observable post-threshold paths remain coordinated through the end of the match. Together, these patterns are consistent with the central idea of the dynamic-basin framework that coordination becomes more robust as accumulated progress increases. Because the public state is generated by previous group behavior rather than randomly assigned, the estimated state gradients should be interpreted as conditional behavioral associations rather than causal effects of the state itself.

The second main finding concerns state- versus history-dependent behavior (Result 2). Comparing the initial $k = 0$ state with $k = 0$ after a reset caused by other group members, subjects assign a significantly lower probability to $X$ after the failure, even though the objective continuation problem has returned to the same state. The decline is concentrated in the $4h$ treatment, where the dynamic basin is smallest, and is small or imprecisely estimated in the other three treatments. Thus, the current public state provides an important summary of the strategic environment, but it does not fully describe behavior. One possible interpretation is that a coordination failure leads subjects to revise downward their beliefs about their current partners, and that such a revision matters most when coordination already requires optimistic beliefs. The initial-versus-reset comparison should nevertheless be interpreted cautiously because the initial observation occurs in the first round of a match whereas a post-reset observation necessarily occurs later in the match.

The third main finding concerns the comparative statics of strategic uncertainty (Results 3a and 3b). Lowering the safe-option payoff substantially increases the probability assigned to $X$ for both group sizes. The effect of group size is more heterogeneous: when the safe-option payoff is high, groups of two assign substantially higher probabilities to $X$ than groups of four, whereas little difference is observed when the safe-option payoff is low and coordination is already near the upper bound, leaving little scope for group size to matter. These conclusions hold when the comparisons are restricted to $k = 0$. In the first round of each match, however, both treatment differences are smaller and the group-size difference is no longer significant, suggesting that the treatment differences widen after coordination failures. More generally, the comparative statics of strategic uncertainty are most visible when the safe alternative is attractive enough that coordination is not already near the upper bound.

The fourth main finding comes from the probability-elicitation design, which reveals behavior that would be hidden by binary action data (Result 5). Interior choice probabilities are common, particularly at the initial state, demonstrating that observed heterogeneity in actions cannot be attributed exclusively to heterogeneity in pure decisions. Mixing becomes substantially less frequent at higher public states, while pure-$X$ choices become increasingly prevalent. Mixing also declines with experience over repeated matches, whereas the corresponding within-block trend in pure-$X$ behavior is small after conditioning on the public state. These patterns suggest two distinct forms of convergence: movement toward higher states is associated primarily with a shift toward pure $X$, while repeated experience reduces subjects' use of interior probabilities. Because the model admits multiple pure and mixed MPEs, the existence of interior choices alone does not identify selection of a particular mixed equilibrium; the comparison with the $(m11)$ benchmark instead provides a way to assess how closely observed state-contingent behavior resembles the mixed equilibrium predicted by the model.

The baseline sessions provide an additional descriptive benchmark for the role of the dynamic environment (Result 6). When the safe-option payoff is low, both treatment and baseline sessions exhibit high levels of $X$ choice, consistent with a ceiling effect. When the safe-option payoff is high, by contrast, the treatment sessions exhibit greater $X$ choice and all-$X$ coordination, and the difference develops primarily over rounds within a match rather than gradually across matches. This pattern is consistent with the dynamic incentives becoming most important when coordination is otherwise difficult to sustain. The baseline comparison should not, however, be interpreted as isolating a single mechanism: the baseline jointly removes the public counter and the last-round payoff penalty, and only two baseline sessions were conducted. I therefore treat these results as a robustness benchmark rather

than as a separate causal treatment comparison.

The experiment is deliberately stylized. The motivating examples, including disease elimination, the adoption of new practices, and minority-culture preservation, are applications in which collective success may require sustained participation and accumulated progress, and the same strategic structure applies more broadly whenever coordinated activity must be maintained long enough to become resilient to an outside disruption. Real applications involve richer institutions, heterogeneous preferences, communication, and longer time horizons than those represented in the laboratory. The contribution of the experiment is therefore not to reproduce a particular real-world setting, but to isolate a dynamic coordination problem in which current incentives depend on accumulated collective progress.

Overall, the results show that equilibrium selection in a dynamic stag hunt cannot be summarized by the underlying one-shot payoff matrix alone. Coordination behavior varies systematically with accumulated progress and with the attractiveness of the safe option, and, in strategically difficult environments, with the number of agents whose actions must be coordinated and with whether the group has just failed. Individual mixing is also quantitatively important at early states but gives way increasingly to pure coordination as groups approach protection. These findings suggest that a state-dependent dynamic basin of attraction provides a useful framework for organizing strategic uncertainty in coordination problems where successful collective action today changes the environment in which coordination takes place tomorrow.

# References


Agranov, M., Frechette, G., Palfrey, T., and Vespa, E. (2016). Static and dynamic underinvestment: An experimental investigation. *Journal of Public Economics*, 143:125–141.

Al-Ubaydli, O., Jones, G., and Weel, J. (2013). Patience, cognitive skill, and coordination in the repeated stag hunt. *Journal of Neuroscience, Psychology, and Economics*, 6(2):71.

Battaglini, M., Nunnari, S., and Palfrey, T. R. (2012). Legislative bargaining and the dynamics of public investment. *American Political Science Review*, 106(2):407–429.

Battaglini, M., Nunnari, S., and Palfrey, T. R. (2016). The dynamic free rider problem: A laboratory study. *American Economic Journal: Microeconomics*, 8(4):268–308.

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

Bisin, A. and Verdier, T. (2000). "Beyond the Melting Pot": Cultural transmission, marriage, and the evolution of ethnic and religious traits. *The Quarterly Journal of Economics*, 115(3):955–988.

Blonski, M., Ockenfels, P., and Spagnolo, G. (2011). Equilibrium selection in the repeated prisoner's dilemma: Axiomatic approach and experimental evidence. *American Economic Journal: Microeconomics*, 3(3):164–192.

Boczoń, M., Vespa, E., Weidman, T., and Wilson, A. J. (2025). Testing models of strategic uncertainty: Equilibrium selection in repeated games. *Journal of the European Economic Association*, 23(2):784–814.

Brocas, I. and Carrillo, J. D. (2025). Dynamic coordination in efficient and fair outcomes: A developmental perspective. *Economic Theory*, 80(3):827–861.

Cason, T. N., Hudja, S., and Rosokha, Y. (2025). Innovation and information free-riding: An experimental investigation of collective experimentation. *Available at SSRN 5253468*.

Chen, D. L., Schonger, M., and Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Chen, R., Chen, Y., and Riyanto, Y. E. (2021). Best practices in replication: A case study of common information in coordination games. *Experimental Economics*, 24(1):2–30.

Church, J. and King, I. (1993). Bilingualism and network externalities. *Canadian Journal of Economics*, pages 337–345.

Cooper, R. W., DeJong, D. V., Forsythe, R., and Ross, T. W. (1990). Selection criteria in coordination games: Some experimental results. *The American Economic Review*, 80(1):218–233.

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

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

Dal Bó, P., Fréchette, G. R., and Kim, J. (2021). The determinants of efficient behavior in coordination games. *Games and Economic Behavior*, 130:352–368.

Dubois, D., Willinger, M., and 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. and Fehr, D. (2018). Equilibrium selection in similar repeated games: Experimental evidence on the role of precedents. *Experimental Economics*, 21(3):573–600.

Dutta, P. K. (1995). A folk theorem for stochastic games. *Journal of Economic Theory*, 66(1):1–32.

Ellison, G. (1993). Learning, local interaction, and coordination. *Econometrica: Journal of the Econometric Society*, pages 1047–1071.

Gerlagh, R. and van der Heijden, E. (2024). Going green: Framing effects in a dynamic coordination game. *Journal of Behavioral and Experimental Economics*, 108:102148.

Gill, D., Prowse, V., and Reddinger, J. L. (2026). Dynamic investment and coordination in teams: Theory and experimental evidence. Working paper.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association*, 1(1):114–125.

Harsanyi, J. C. and Selten, R. (1988). A general theory of equilibrium selection in games. *MIT Press Books*, 1.

Jagau, S. (2024). To catch a stag: Identifying payoff-and risk-dominance effects in coordination games. *Games and Economic Behavior*, 147:429–448.

Kandori, M., Mailath, G. J., and Rob, R. (1993). Learning, mutation, and long run equilibria in games. *Econometrica: Journal of the Econometric Society*, pages 29–56.

Kendall, R. (2022). Decomposing coordination failure in stag hunt games. *Experimental Economics*, 25(4):1109–1145.

Lazear, E. P. (1999). Culture and language. *Journal of political Economy*, 107(S6):S95–S126.

Maskin, E. and Tirole, J. (1988a). A theory of dynamic oligopoly, I: Overview and quantity competition with large fixed costs. *Econometrica: Journal of the Econometric Society*, pages 549–569.

Maskin, E. and Tirole, J. (1988b). A theory of dynamic oligopoly, II: Price competition, kinked demand curves, and edgeworth cycles. *Econometrica: Journal of the Econometric Society*, pages 571–599.

Maskin, E. and Tirole, J. (2001). Markov perfect equilibrium: I. Observable actions. *Journal of Economic Theory*, 100(2):191–219.

Murphy, R. O., Ackermann, K. A., and Handgraaf, M. J. (2011). Measuring social value orientation. *Judgment and Decision making*, 6(8):771–781.

Rankin, F. W., Van Huyck, J. B., and Battalio, R. C. (2000). Strategic similarity and emergent conventions: Evidence from similar stag hunt games. *Games and Economic Behavior*, 32(2):315–337.

Romero, J. and Rosokha, Y. (2023). Mixed strategies in the indefinitely repeated prisoner's dilemma. *Econometrica*, 91(6):2295–2331.

Rosokha, Y., Lyu, X., Tverskoi, D., and Gavrilets, S. (2024). The political hare and the stag hunt. *Available at SSRN 4891324*.

Rosokha, Y., Lyu, X., Tverskoi, D., and Gavrilets, S. (2025). Cooperation under the shadow of political inequality. *Journal of Economic Dynamics and Control*, 172:104988.

Rosokha, Y., Lyu, X., Tverskoi, D., and Gavrilets, S. (2026). Evolution of cooperation in the indefinitely repeated collective action with a contest for power: Y. rosokha et al. *Economic Theory*, 81(1):453–483.

Rosokha, Y. and Wei, C. (2024). Cooperation in queueing systems. *Management Science*, 70(11):7597–7616.

Salz, T. and Vespa, E. (2020). Estimating dynamic games of oligopolistic competition: an experimental investigation. *The RAND Journal of Economics*, 51(2):447–469.

Schmidt, D., Shupp, R., Walker, J. M., and 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.

Silva, R. (2024). Coordination in stag hunt games. *Journal of Behavioral and Experimental Economics*, 113:102290.

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

Van Huyck, J. B., Battalio, R. C., and Beil, R. O. (1991). Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games. *The Quarterly Journal of Economics*, 106(3):885–910.

Vespa, E. (2020). An experimental investigation of cooperation in the dynamic common pool game. *International Economic Review*, 61(1):417–440.

Vespa, E. and Wilson, A. J. (2016). Experimenting with equilibrium selection in dynamic games. Working paper.

Vespa, E. and Wilson, A. J. (2019). Experimenting with the transition rule in dynamic games. *Quantitative Economics*, 10(4):1825–1849.

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

Young, H. P. (1993). The evolution of conventions. *Econometrica: Journal of the Econometric Society*, pages 57–84.

# Appendix A1: Solutions for Regime-monotone Mixed MPEs when $\bar{K} = 1, 2, 3$ and Selected Proofs

Let $\sigma_k \in [0, 1]$ denote the probability that a representative agent chooses $X$ in state $k$. Under a symmetric strategy profile, the probability that all other $n-1$ agents choose $X$ is $q_k \equiv \sigma_k^{n-1}$. Thus, $\sigma_k = q_k^{1/(n-1)}$. The analysis below applies to the group sizes used in the experiment, $n \in \{2, 4\}$, and more generally to any $n \geq 2$. Let $\delta \in (0, 1)$ denote the continuation probability. Recall the stage payoffs:

$$u_{XX} = 1, \qquad u_{XY} = 0, \qquad u_Y \in (0, 1).$$

Let $V_k$ denote the continuation value at the beginning of a round in state $k$, and let $Q_k(X)$ and $Q_k(Y)$ denote the corresponding action values. Choosing $Y$ ensures that the current outcome is not all-$X$, so the next state is 0 and the overwrite rule applies if the shock arrives. Hence, for every $k$,

$$Q_k(Y) = \delta\big(u_Y + V_0\big) + (1-\delta)(-d). \tag{5}$$

If the agent chooses $X$, then with probability $q_k$ all other $n-1$ agents also choose $X$. If an all-$X$ outcome does not yet reach the threshold, i.e. $k \leq \bar{K} - 2$, then

$$Q_k(X) = q_k\Big[\delta\big(1 + V_{k+1}\big) + (1-\delta)(-d)\Big] + (1-q_k)\Big[\delta V_0 + (1-\delta)(-d)\Big]. \tag{6}$$

At $k = \bar{K} - 1$, an all-$X$ outcome reaches the protected state $\bar{K}$, so

$$Q_{\bar{K}-1}(X) = q_{\bar{K}-1}\Big[\delta\big(1 + V_{\bar{K}}\big) + (1-\delta)\Big] + (1-q_{\bar{K}-1})\Big[\delta V_0 + (1-\delta)(-d)\Big]. \tag{7}$$

Finally, at the protected state $k = \bar{K}$,

$$Q_{\bar{K}}(X) = q_{\bar{K}}\Big[\delta\big(1 + V_{\bar{K}}\big) + (1-\delta)\Big] + (1-q_{\bar{K}})\Big[\delta V_0 + (1-\delta)(-d)\Big]. \tag{8}$$

In any state in which agents mix, $\sigma_k \in (0, 1)$ and hence $q_k \in (0, 1)$. The indifference condition is

$$Q_k(X) = Q_k(Y) = V_k.$$

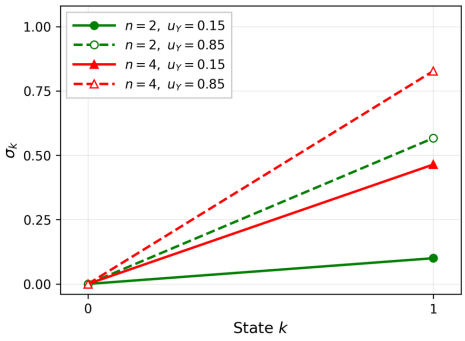
Since $q_k = \sigma_k^{n-1}$, solving for $q_k$ is sufficient; the corresponding individual mixing probability is $\sigma_k = q_k^{1/(n-1)}$. Below is the complete illustration of all regime-monotone mixed MPEs for $\bar{K} = 1, 2, 3$. Note that the point values in the corresponding figures take the parameterization of $\delta = 0.8, d = 1$, and $n = 2$ (in green), $n = 4$ (in red), $u_Y = 0.15$ (in solid lines), $u_Y = 0.85$

(in dashed lines). Some candidates have blue tags to explain why they fail to be the final solution for the regime-monotone mixed MPE. Selected proofs of candidates without a blue tag are shown right after. Section A1.4 provides the proofs for pure MPEs, and Section A1.5 counts the regime patterns.

## A1.1 Solutions when $\bar{K} = 1$

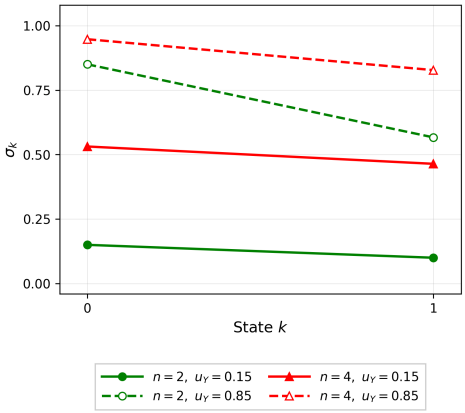
**[Candidate** $(0m)$**]: off-path**

$$V_0 = \frac{\delta}{1-\delta} u_Y - d, \qquad V_1 = V_0,$$
$$\sigma_0 = 0, \qquad \sigma_1 = \left(\frac{\delta u_Y}{1 + (1-\delta)d}\right)^{\frac{1}{n-1}}.$$



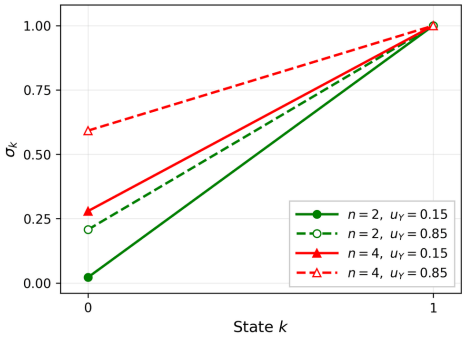

**[Candidate** $(mm)$**]: not monotone**

$$V_0 = \frac{\delta}{1-\delta} u_Y - d, \qquad V_1 = V_0$$
$$\sigma_0 = u_Y^{\frac{1}{n-1}}, \qquad \sigma_1 = \left(\frac{\delta u_Y}{1 + (1-\delta)d}\right)^{\frac{1}{n-1}}.$$




**[Candidate** $(m1)$**]**

$$V_0 = \frac{\delta}{1-\delta} u_Y - d, \qquad V_1 = \frac{1}{1-\delta},$$
$$\sigma_0 = \left(\frac{\delta(1-\delta)u_Y}{1 - \delta^2 u_Y + (1-\delta)d}\right)^{\frac{1}{n-1}}, \qquad \sigma_1 = 1.$$




*Proof.* Consider

$$\sigma_0 \in (0,1), \qquad \sigma_1 = 1, \qquad \text{or equivalently,} \qquad q_0 \in (0,1), \qquad q_1 = 1.$$

Since $Y$ is played with positive probability at $k = 0$,

$$V_0 = Q_0(Y)$$

Using (5),

$$V_0 = \delta\big(u_Y + V_0\big) + (1-\delta)(-d),$$

which gives

$$V_0 = \frac{\delta}{1-\delta} u_Y - d. \tag{9}$$

At $k = 1$, the candidate prescribes $X$ with probability one. The protected state is therefore absorbing along the equilibrium path:

$$V_1 = \delta\big(1 + V_1\big) + (1-\delta).$$

Hence

$$V_1 = \frac{1}{1-\delta}. \tag{10}$$

Mixing at $k = 0$ requires $Q_0(X) = Q_0(Y) = V_0$. Since an all-$X$ outcome reaches protection immediately,

$$V_0 = q_0\Big[\delta\big(1 + V_1\big) + (1-\delta)\Big] + (1-q_0)\Big[\delta V_0 + (1-\delta)(-d)\Big].$$

Substituting (9) and (10) and rearranging yields

$$q_0 = \frac{\delta(1-\delta)u_Y}{1 - \delta^2 u_Y + (1-\delta)d}. \tag{11}$$

Therefore,

$$\sigma_0 = \left[\frac{\delta(1-\delta)u_Y}{1 - \delta^2 u_Y + (1-\delta)d}\right]^{1/(n-1)}. \tag{12}$$

Writing $q_0 = N/D$, we have

$$D - N = 1 - \delta u_Y + (1-\delta)d > 1 - u_Y > 0.$$

Thus $D > N > 0$, implying $q_0 \in (0,1)$ and hence $\sigma_0 = q_0^{1/(n-1)} \in (0,1)$. It remains to verify that $X$ is optimal at $k = 1$. Under the candidate,

$$Q_1(X) = V_1, \qquad Q_1(Y) = V_0.$$

Thus $X$ is optimal if $V_1 \geq V_0$, or

$$\frac{1}{1-\delta} \geq \frac{\delta}{1-\delta} u_Y - d,$$

which is equivalent to

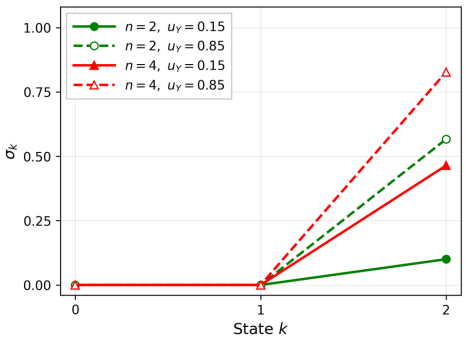
$$1 \geq \delta u_Y - (1-\delta)d.$$

This holds because $1 > u_Y > \delta u_Y$ and $d > 0$. □

## A1.2 Solutions when $\bar{K} = 2$

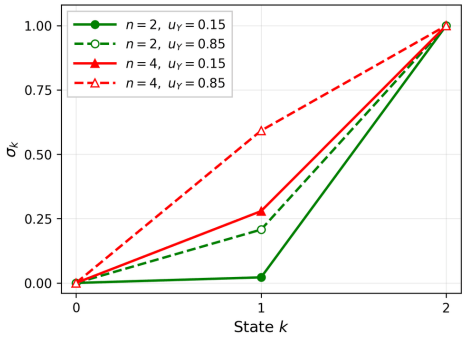
**[Candidate $(00m)$]: off-path**

$$V_0 = V_1 = V_2 = \frac{\delta}{1-\delta} u_Y - d,$$
$$\sigma_0 = 0, \qquad \sigma_1 = 0, \qquad \sigma_2 = \left(\frac{\delta u_Y}{1 + (1-\delta)d}\right)^{\frac{1}{n-1}}.$$



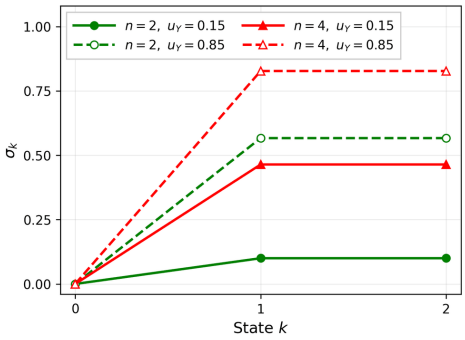

**[Candidate $(0m1)$]: off-path**

$$V_0 = V_1 = \frac{\delta}{1-\delta} u_Y - d, \qquad V_2 = \frac{1}{1-\delta},$$
$$\sigma_0 = 0, \qquad \sigma_1 = \left(\frac{\delta(1-\delta)u_Y}{1 - \delta^2 u_Y + (1-\delta)d}\right)^{\frac{1}{n-1}}, \qquad \sigma_2 = 1.$$



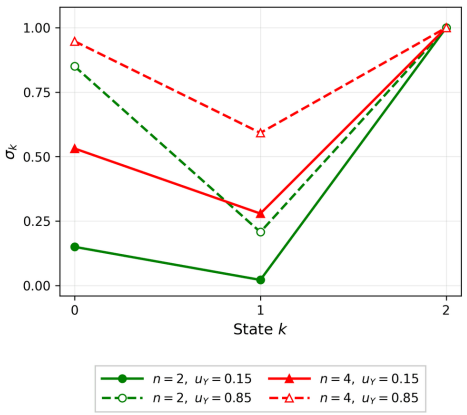

**[Candidate $(0mm)$]: off-path**

$$V_0 = V_1 = V_2 = \frac{\delta}{1-\delta} u_Y - d,$$
$$\sigma_0 = 0, \qquad \sigma_1 = \sigma_2 = \left(\frac{\delta u_Y}{1 + (1-\delta)d}\right)^{\frac{1}{n-1}}.$$



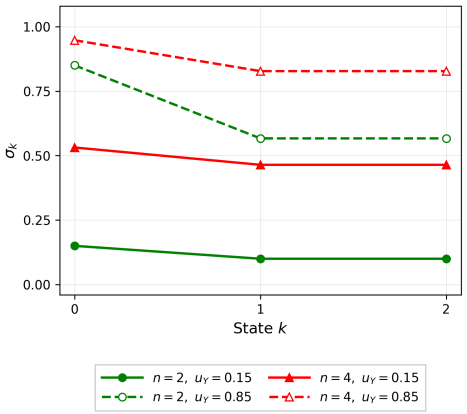
**[Candidate $(mm1)$]: not monotone**

$$V_0 = V_1 = \frac{\delta}{1-\delta} u_Y - d, \qquad V_2 = \frac{1}{1-\delta},$$
$$\sigma_0 = u_Y^{\frac{1}{n-1}},$$
$$\sigma_1 = \left( \frac{\delta(1-\delta)u_Y}{1-\delta^2 u_Y + (1-\delta)d} \right)^{\frac{1}{n-1}}, \qquad \sigma_2 = 1.$$



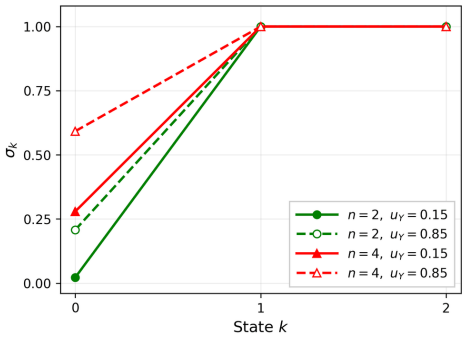

**[Candidate $(mmm)$]: not monotone**

$$V_0 = V_1 = V_2 = \frac{\delta}{1-\delta} u_Y - d,$$
$$\sigma_0 = u_Y^{\frac{1}{n-1}}, \qquad \sigma_1 = \sigma_2 = \left( \frac{\delta u_Y}{1 + (1-\delta)d} \right)^{\frac{1}{n-1}}.$$




**[Candidate $(m11)$]**

$$V_0 = \frac{\delta}{1-\delta} u_Y - d, \qquad V_1 = V_2 = \frac{1}{1-\delta},$$
$$\sigma_0 = \left( \frac{(1-\delta)u_Y}{(2-\delta) - \delta u_Y + (1-\delta)d} \right)^{\frac{1}{n-1}}, \quad \sigma_1 = \sigma_2 = 1.$$




*Proof.* Consider

$$\sigma_0 \in (0,1), \qquad \sigma_1 = \sigma_2 = 1, \qquad \text{or equivalently,} \qquad q_0 \in (0,1), \qquad q_1 = q_2 = 1.$$

At the protected state,

$$V_2 = \delta\big(1 + V_2\big) + (1-\delta),$$

so

$$V_2 = \frac{1}{1-\delta}. \tag{13}$$

At $k = 1$, choosing $X$ reaches protection in the current round:

$$V_1 = \delta\big(1 + V_2\big) + (1 - \delta) = \frac{1}{1 - \delta}.$$

Thus,

$$V_1 = V_2 = \frac{1}{1 - \delta}. \tag{14}$$

Since agents mix at $k = 0$,

$$V_0 = Q_0(Y),$$

and therefore, as in (9),

$$V_0 = \frac{\delta}{1 - \delta} u_Y - d. \tag{15}$$

At $k = 0$, an all-$X$ outcome moves the state to $1 < \bar{K}$, so the overwrite rule still applies if a shock arrives. The indifference condition is

$$q_0(1 + V_1) + (1 - q_0)V_0 = u_Y + V_0.$$

Hence

$$q_0 = \frac{u_Y}{1 + V_1 - V_0}.$$

Substituting (14) and (15),

$$q_0 = \frac{(1 - \delta)u_Y}{(2 - \delta) - \delta u_Y + (1 - \delta)d}. \tag{16}$$

Therefore,

$$\sigma_0 = q_0^{1/(n-1)}. \tag{17}$$

Write $q_0$ as $q_0 = N/D$, then

$$D - N = (2 - \delta) - u_Y + (1 - \delta)d > (2 - \delta)u_Y - u_Y > 0.$$

Thus $q_0 \in (0, 1)$ and therefore $\sigma_0 \in (0, 1)$. For $k = 1, 2$, $Q_k(Y) = V_0$ while

$$Q_1(X) = V_1, \qquad Q_2(X) = V_2.$$

Since

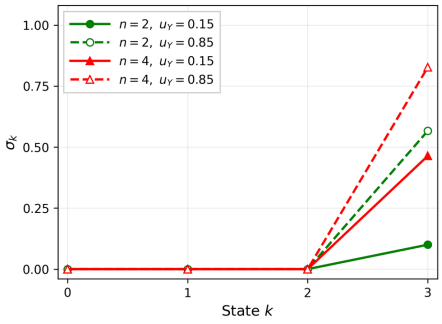
$$V_1 = V_2 = \frac{1}{1 - \delta} > \frac{\delta}{1 - \delta} u_Y - d = V_0,$$

choosing $X$ is optimal in both states. □

## A1.3 Solutions when $\bar{K} = 3$

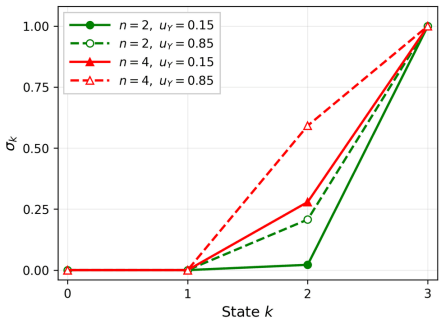
**[Candidate $(000m)$]: off-path**

$$V_0 = V_1 = V_2 = V_3 = \frac{\delta}{1-\delta} u_Y - d,$$
$$\sigma_0 = \sigma_1 = \sigma_2 = 0, \qquad \sigma_3 = \left( \frac{\delta u_Y}{1 + (1-\delta)d} \right)^{\frac{1}{n-1}}.$$



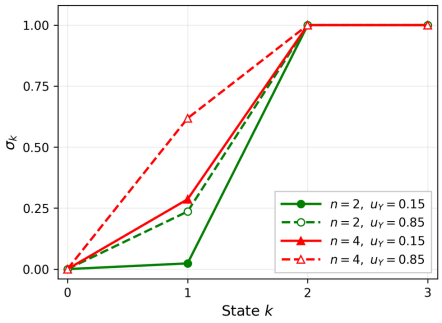

**[Candidate $(00m1)$]: off-path**

$$V_0 = V_1 = V_2 = \frac{\delta}{1-\delta} u_Y - d, \qquad V_3 = \frac{1}{1-\delta},$$
$$\sigma_0 = \sigma_1 = 0, \qquad \sigma_2 = \left( \frac{\delta(1-\delta) u_Y}{1 - \delta^2 u_Y + (1-\delta)d} \right)^{\frac{1}{n-1}},$$
$$\sigma_3 = 1.$$



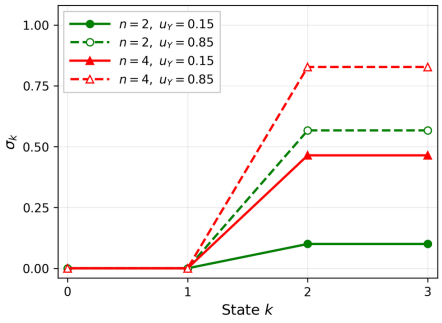

**[Candidate $(0m11)$]: off-path**

$$V_0 = V_1 = \frac{\delta}{1-\delta} u_Y - d, \qquad V_2 = V_3 = \frac{1}{1-\delta},$$
$$\sigma_0 = 0, \qquad \sigma_1 = \left( \frac{(1-\delta) u_Y}{(2 - \delta - \delta u_Y) + (1-\delta)d} \right)^{\frac{1}{n-1}},$$
$$\sigma_2 = \sigma_3 = 1.$$



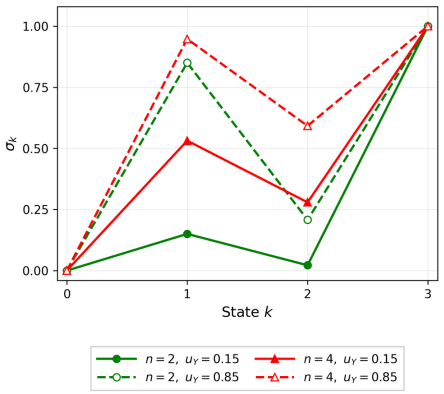

**[Candidate $(00mm)$]: off-path**

$$V_0 = V_1 = V_2 = V_3 = \frac{\delta}{1-\delta} u_Y - d,$$
$$\sigma_0 = \sigma_1 = 0, \qquad \sigma_2 = \sigma_3 = \left( \frac{\delta u_Y}{1 + (1-\delta)d} \right)^{\frac{1}{n-1}}.$$



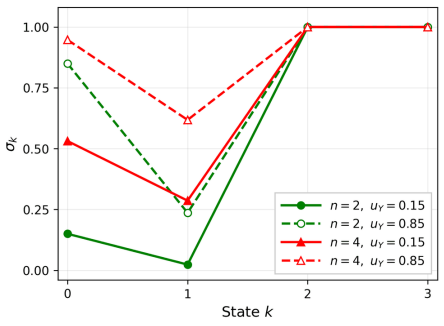
**[Candidate $(0mm1)$]: off-path, not monotone**

$$V_0 = V_1 = V_2 = \frac{\delta}{1-\delta} u_Y - d, \qquad V_3 = \frac{1}{1-\delta},$$
$$\sigma_0 = 0, \qquad \sigma_1 = u_Y^{\frac{1}{n-1}},$$
$$\sigma_2 = \left( \frac{\delta(1-\delta)u_Y}{1-\delta^2 u_Y + (1-\delta)d} \right)^{\frac{1}{n-1}} \qquad \sigma_3 = 1.$$



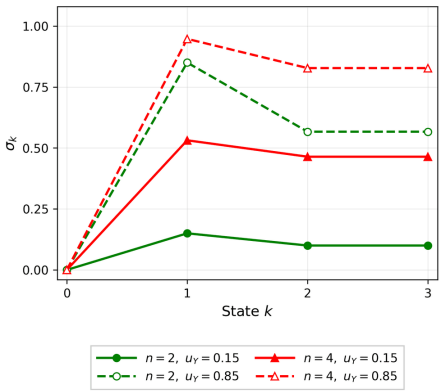

**[Candidate $(mm11)$]: not monotone**

$$V_0 = V_1 = \frac{\delta}{1-\delta} u_Y - d, \qquad V_2 = V_3 = \frac{1}{1-\delta},$$
$$\sigma_0 = u_Y^{\frac{1}{n-1}},$$
$$\sigma_1 = \left( \frac{(1-\delta)u_Y}{(2-\delta-\delta u_Y) + (1-\delta)d} \right)^{\frac{1}{n-1}},$$
$$\sigma_2 = \sigma_3 = 1.$$



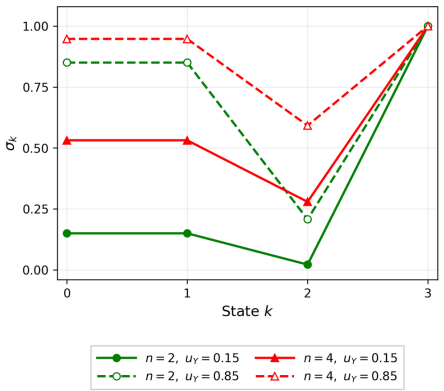

**[Candidate $(0mmm)$]: off-path, not monotone**

$$V_0 = V_1 = V_2 = V_3 = \frac{\delta}{1-\delta} u_Y - d,$$
$$\sigma_0 = 0, \qquad \sigma_1 = u_Y^{\frac{1}{n-1}},$$
$$\sigma_2 = \sigma_3 = \left( \frac{\delta u_Y}{1 + (1-\delta)d} \right)^{\frac{1}{n-1}}.$$



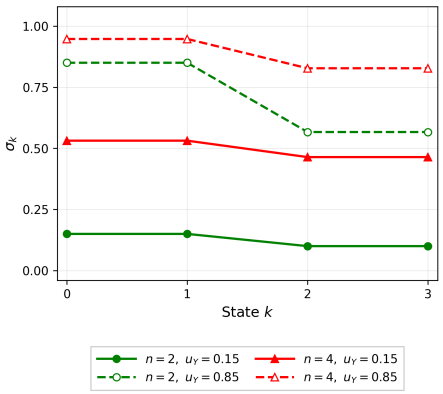

**[Candidate $(mmm1)$]: not monotone**

$$V_0 = V_1 = V_2 = \frac{\delta}{1-\delta} u_Y - d, \qquad V_3 = \frac{1}{1-\delta},$$
$$\sigma_0 = \sigma_1 = u_Y^{\frac{1}{n-1}},$$
$$\sigma_2 = \left( \frac{\delta(1-\delta)u_Y}{1-\delta^2 u_Y + (1-\delta)d} \right)^{\frac{1}{n-1}}, \qquad \sigma_3 = 1.$$



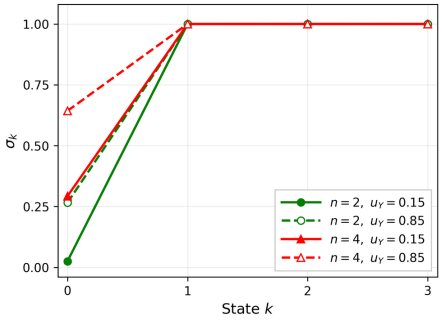
**[Candidate** $(mmmm)$**]: not monotone**

$$V_0 = V_1 = V_2 = V_3 = \frac{\delta}{1-\delta} u_Y - d,$$
$$\sigma_0 = \sigma_1 = u_Y^{\frac{1}{n-1}},$$
$$\sigma_2 = \sigma_3 = \left(\frac{\delta u_Y}{1+(1-\delta)d}\right)^{\frac{1}{n-1}}.$$




**[Candidate** $(m111)$**]**

$$V_0 = \frac{\delta}{1-\delta} u_Y - d, \qquad V_1 = \frac{\delta(2-\delta)}{1-\delta} - (1-\delta)d,$$
$$V_2 = V_3 = \frac{1}{1-\delta},$$
$$\sigma_0 = \left(\frac{(1-\delta)u_Y}{1+\delta(1-\delta)(1+d)-\delta u_Y}\right)^{\frac{1}{n-1}},$$
$$\sigma_1 = \sigma_2 = \sigma_3 = 1.$$




*Proof.* Consider

$$\sigma_0 \in (0,1), \qquad \sigma_1 = \sigma_2 = \sigma_3 = 1, \qquad \text{or equivalently,} \qquad q_0 \in (0,1), \qquad q_1 = q_2 = q_3 = 1.$$

At $k = 3$,

$$V_3 = \delta(1 + V_3) + (1-\delta),$$

so

$$V_3 = \frac{1}{1-\delta}. \tag{18}$$

At $k = 2$, choosing $X$ reaches protection in the current round, implying

$$V_2 = \frac{1}{1-\delta} = V_3. \tag{19}$$

At $k = 1$, an all-$X$ outcome moves the state to $2 < \bar{K}$, so the overwrite rule still applies if the shock arrives:

$$V_1 = \delta(1 + V_2) + (1-\delta)(-d).$$

Substituting (19),

$$V_1 = \frac{\delta(2-\delta)}{1-\delta} - (1-\delta)d. \tag{20}$$

Since agents mix at $k = 0$,

$$V_0 = \frac{\delta}{1-\delta} u_Y - d. \tag{21}$$

At $k = 0$, the indifference condition gives

$$q_0(1 + V_1) + (1 - q_0)V_0 = u_Y + V_0.$$

Therefore,

$$q_0 = \frac{u_Y}{1 + V_1 - V_0}.$$

Substituting (20) and (21),

$$q_0 = \frac{(1-\delta)u_Y}{(1 + \delta - \delta^2) - \delta u_Y + \delta(1-\delta)d}. \tag{22}$$

Hence

$$\sigma_0 = q_0^{1/(n-1)}. \tag{23}$$

Write $q_0$ as $q_0 = N/D$, then

$$D - N = (1 + \delta - \delta^2) - u_Y + \delta(1-\delta)d > (1 + \delta - \delta^2)u_Y - u_Y > 0.$$

Thus $q_0 \in (0, 1)$ and $\sigma_0 \in (0, 1)$. For $k = 2, 3$, $Q_k(X) = u_{XX}/(1-\delta) > V_0 = Q_k(Y)$. At $k = 1$,

$$Q_1(X) = V_1, \qquad Q_1(Y) = V_0,$$

and

$$V_1 - V_0 = \frac{\delta}{1-\delta}\Big[(2 - \delta) - u_Y\Big] + \delta d > 0.$$

Thus $X$ is optimal at every $k \geq 1$. □

## A1.4 Proofs for Pure MPEs

*Lemma* 1 (Best response at a given state). Fix a state $k$.

1. If at least one other member chooses $Y$ at state $k$, then $Y$ is a dominant best response.
2. If all other members choose $X$ at state $k$, then $X$ is a best response.

*Proof.* If at least one other member chooses $Y$, choosing $X$ cannot produce a group outcome of all-$X$, so the current payoff from $X$ is $u_{XY} = 0$, and the value of the next state $k'$ is 0. Choosing $Y$ yields a payoff $u_Y \in (0, 1)$ and also induces $k' = 0$. Hence $Y$ strictly dominates

$X$. If all other members choose $X$, choosing $X$ yields current payoff $u_{XX} = 1$ and moves the value of the next state to $k' = \min\{k+1, \bar{K}\}$, which is (weakly) closer to the threshold $\bar{K}$; while deviating to $Y$ yields current payoff $u_Y \in (0,1)$ and resets the value of the next state to $k' = 0$. Therefore, $X$ is a best response. □

*Proposition* 3 (All pure symmetric MPEs). Every symmetric pure Markov strategy profile $a(\cdot)$ with $a(k) \in \{X, Y\}$ for all $k$ is a Markov Perfect Equilibrium. Hence there are exactly $2^{\bar{K}+1}$ pure symmetric MPEs.

*Proof.* Take any symmetric pure Markov strategy $a(\cdot)$. At any state $k$, if $a(k) = Y$, then all other members also choose $Y$ at that state, so by Lemma 1, $Y$ is a best response. If $a(k) = X$, then all other members also choose $X$, so by Lemma 1, $X$ is a best response. Therefore, the prescribed action is optimal at every state, and $a(\cdot)$ is a pure symmetric MPE. Since at each of the $\bar{K}+1$ states the action can be chosen independently as either $X$ or $Y$, the total number of pure symmetric MPEs is $2^{\bar{K}+1}$. □

*Proof of Proposition 1.* Let $a(\cdot)$ be a monotone pure Markov strategy. If $a(k) \equiv Y$ for all $k$, then this is the case $r = \bar{K}+1$. Otherwise, let $r \equiv \min\{k : a(k) = X\}$. By monotonicity, $a(k) = Y$ for all $k < r$ and $a(k) = X$ for all $k \geq r$, so $a(\cdot) = a^{(r)}(\cdot)$ for some $r \in \{0, 1, \ldots, \bar{K}\}$. The case $r = 0$ gives $a(k) \equiv X$. Conversely, each threshold strategy $a^{(r)}(\cdot)$ is a symmetric pure MPE by Proposition 3, since every symmetric pure Markov strategy is a pure MPE. Hence, the monotone pure symmetric MPEs are exactly those listed above. □

## A1.5 Counting Regime Patterns

*Proposition* 4 (Counting mixed-MPE regime patterns). Consider symmetric, stationary Markov profiles on the state space $\{0, 1, \ldots, \bar{K}\}$. In any stationary MPE, each state $k$ must fall into exactly one of three regimes:

$$\sigma_k = \begin{cases} 1, & \text{if } Q_k(X) > Q_k(Y), \\ 0, & \text{if } Q_k(X) < Q_k(Y), \\ \in (0,1), & \text{only if } Q_k(X) = Q_k(Y). \end{cases}$$

Hence, the number of distinct regime patterns is $3^{\bar{K}+1}$, the number of pure (no-mixing) patterns is $2^{\bar{K}+1}$, and the number of mixed *candidate* regime patterns is $3^{\bar{K}+1} - 2^{\bar{K}+1}$.

*Definition* 3 (Regime-monotone mixed Markov profile). A stationary mixed Markov profile is regime-monotone if its regime pattern is non-decreasing in the order $0 < m < 1$; that is, whenever $k < \bar{K}$,

$$\sigma_k = 1 \;\Rightarrow\; \sigma_{k+1} = 1, \qquad \sigma_k = m \;\Rightarrow\; \sigma_{k+1} \in \{m, 1\}.$$

Equivalently, $\sigma_k$ can move only "up" in the order $0 \to m \to 1$, never down.

*Proposition* 5 (Counting regime-monotone mixed candidates). Fix $\bar{K} \geq 1$. Then the number of regime-monotone mixed candidate patterns equals

$$\binom{\bar{K}+3}{2} - (\bar{K}+2) = \frac{(\bar{K}+1)(\bar{K}+2)}{2}.$$

*Proof.* A regime-monotone pattern over $\{0, m, 1\}$ on the state space $\{0, 1, \ldots, \bar{K}\}$ must be of the form

$$\underbrace{0 \cdots 0}_{z}\ \underbrace{m \cdots m}_{r}\ \underbrace{1 \cdots 1}_{o}, \qquad z, r, o \geq 0, \quad z + r + o = \bar{K} + 1.$$

Thus, such patterns are in one-to-one correspondence with triples $(z, r, o)$ of nonnegative integers summing to $\bar{K}+1$. By stars-and-bars there are $\binom{\bar{K}+3}{2}$ such triples. Among them, those with $r = 0$ are pure 0/1 threshold patterns, and there are $\bar{K}+2$ of these. Subtracting yields $\frac{(\bar{K}+1)(\bar{K}+2)}{2}$. □

Table A1.1: Regime-monotone and Monotone Mixed MPEs for $\bar{K} = 1, 2, 3$

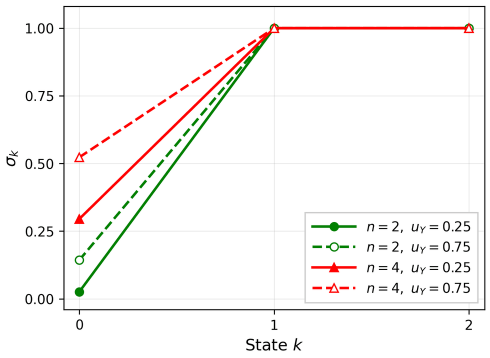
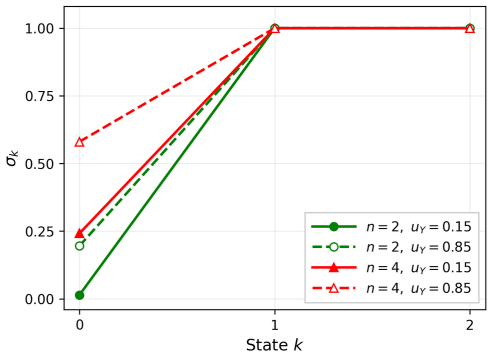
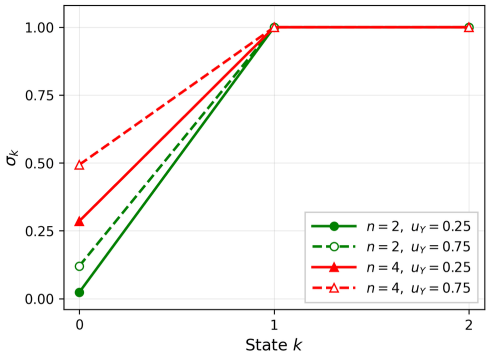
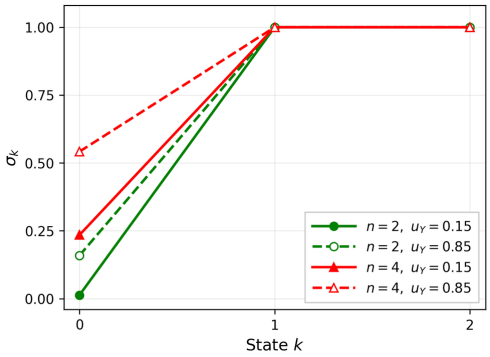
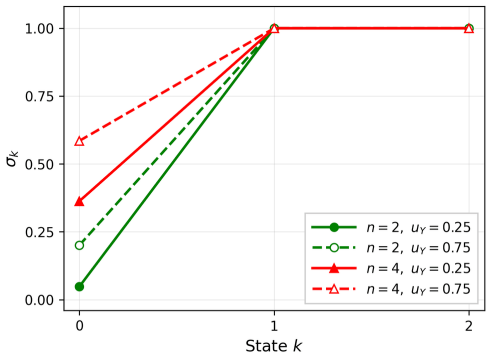
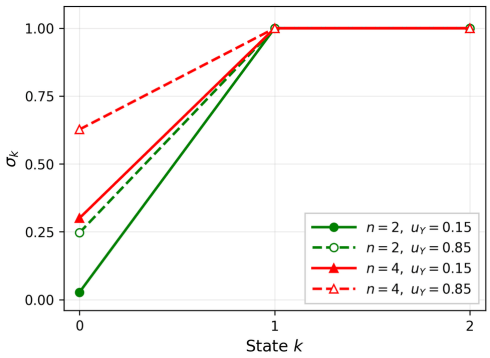
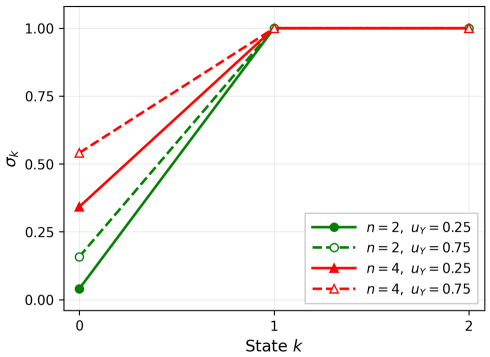
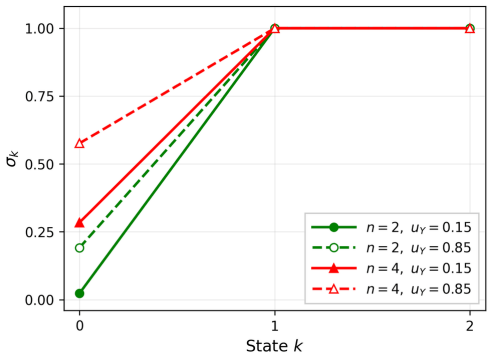
| | $\bar{K}=1$ | | $\bar{K}=2$ | | $\bar{K}=3$ | |
|---|---|---|---|---|---|---|
| | $N$ | $\{\sigma_k\}$ profiles | $N$ | $\{\sigma_k\}$ profiles | $N$ | $\{\sigma_k\}$ profiles |
| Regime-monotone | 3 | $(0m)$, $(mm)$, $(m1)$ | 6 | $(00m)$, $(0mm)$, $(0m1)$, $(mmm)$, $(mm1)$, $(m11)$ | 10 | $(000m)$, $(00mm)$, $(00m1)$, $(0mmm)$, $(0mm1)$, $(0m11)$, $(mmmm)$, $(mmm1)$, $(mm11)$, $(m111)$ |
| Monotone | 2 | $(0m)$, $(m1)$ | 4 | $(00m)$, $(0mm)$, $(0m1)$, $(m11)$ | 5 | $(000m)$, $(00mm)$, $(00m1)$, $(0m11)$, $(m111)$ |
| Monotone (on-path) | 1 | $(m1)$ | 1 | $(m11)$ | 1 | $(m111)$ |

Note: The listed profiles are nested across rows: the monotone mixed MPEs form a subset of the regime-monotone candidates, and the on-path monotone mixed MPEs form a subset of the monotone mixed MPEs. $m$ denotes interior mixing $\sigma_k \in (0, 1)$; 0 and 1 denote $\sigma_k = 0$ and $\sigma_k = 1$.

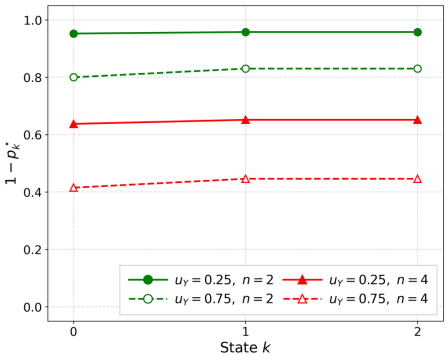
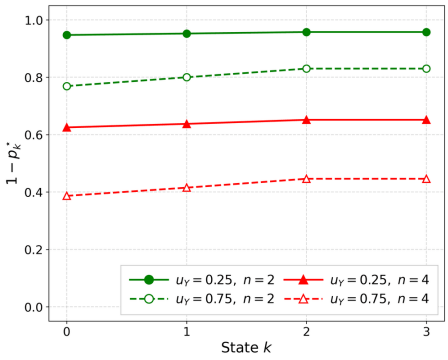
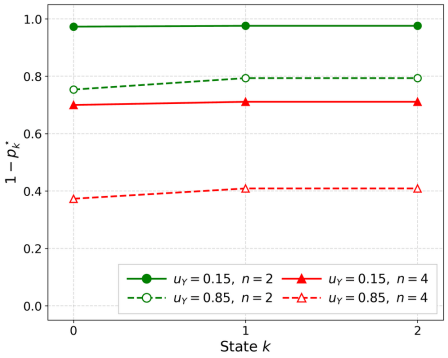
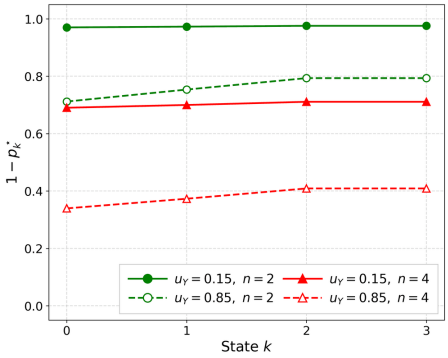
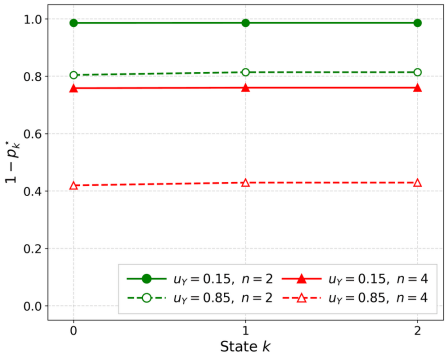
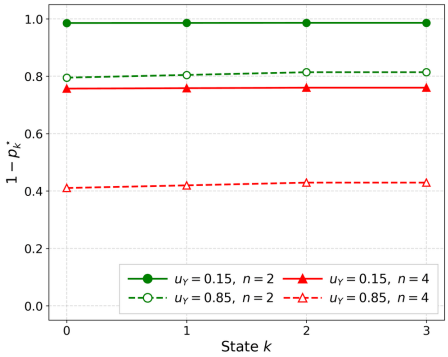
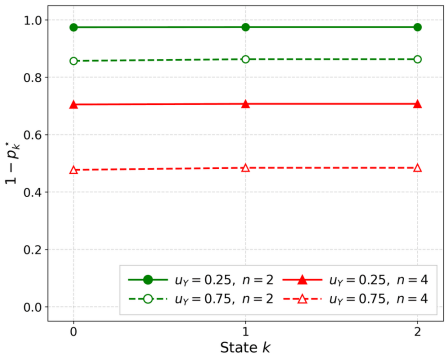
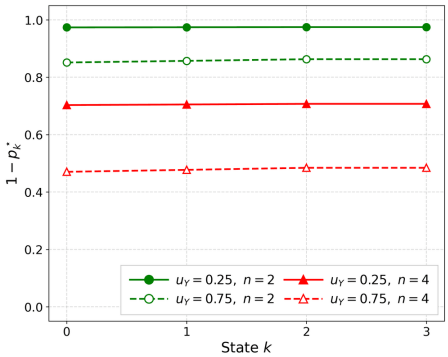
Definition 3 restricts only the potential shape of the profile. It does not guarantee that the solved mixing probabilities satisfy $\sigma_0 \leq \sigma_1 \leq \cdots \leq \sigma_{\bar{K}}$. Table A1.1 lists all regime-monotone candidates and the monotone mixed MPEs for $\bar{K} = 1, 2, 3$.

# Appendix A2: Parameter Comparisons

## A2.1 Parameter Comparison for the $(m11)$ Mixed MPE, $\bar{K} = 2$

| *Parameters* | $u_Y \in \{0.25, 0.75\}$ | $u_Y \in \{0.15, 0.85\}$ |
|---|---|---|
| $\delta = 0.9$, $d = 1$ |  |  |
| $\delta = 0.9$, $d = 2$ |  |  |
| $\delta = 0.75$, $d = 1$ |  |  |
| $\delta = 0.75$, $d = 2$ |  |  |

## A2.2 Parameter Comparisons for the Dynamic Basin of Attraction of Grim Trigger against All-Y

| *Parameters* | $\bar{K} = 2$ | $\bar{K} = 3$ |
|---|---|---|
| $\delta = 0.75$<br>$u_Y \in \{0.25, 0.75\}$ |  |  |
| $\delta = 0.75$<br>$u_Y \in \{0.15, 0.85\}$ |  |  |
| $\delta = 0.9$<br>$u_Y \in \{0.15, 0.85\}$ |  |  |
| $\delta = 0.9$<br>$u_Y \in \{0.25, 0.75\}$ |  |  |

# Appendix A3: History-dependent Strategies and Derivation of the Dynamic Basin

## A3.1 Equivalent History-dependent Rules and Subgame Perfection

Markov Perfect Equilibria are state-based strategies that depend only on the current state, regardless of the history. By contrast, purely history-dependent strategies condition only on past play, while hybrid strategies condition on both the public state and the realized history. Consider the following four candidate rules. Two are purely history-dependent: (i) Grim Trigger independent of the state, denoted $GT$, which starts with $X$ and switches permanently to $Y$ after the first group failure, and (ii) Tit-for-Tat independent of the state, denoted $TFT$, which starts with $X$ and in each later round mimics the previous group outcome, choosing $X$ after all-$X$ and $Y$ otherwise. The other two are hybrid state-and-history-dependent rules: (iii) a state-dependent Grim Trigger, denoted $GT(k)$, that uses the same trigger rule but allows a state-based override only at $k = \bar{K}$, and (iv) a state-dependent Tit-for-Tat, denoted $TFT(k)$, that follows the same group-outcome rule with the same $k = \bar{K}$ override.

Under the public monitoring and reset rule in the present game, these four labels are distinct descriptions of the same induced strategy. After any group failure, the state next period is $k = 0$. Hence the override at $k = \bar{K}$ never binds off the cooperative path. Moreover, once $TFT$ prescribes $Y$ after the first group failure, all members choose $Y$, so the public outcome remains not all-$X$ forever and $TFT$ also stays at $Y$ forever. Therefore, along the path of play and after any unilateral deviation from it, all four rules prescribe the same actions; they can differ only after histories in which every member simultaneously returns to $X$ following a failure.

These four rules are also all Subgame Perfect Equilibria. After any history at which the rule prescribes $Y$ for all members, choosing $X$ cannot produce all-$X$ and does not change continuation play, so $Y$ is a strict best response. Along a cooperative public history at state $k$, let $m \equiv \max\{\bar{K} - k, 1\}$ (truncated at one because an all-$X$ outcome at $k = \bar{K} - 1$ already reaches protection). If all other members follow any one of the four rules above, then deviating to $Y$ immediately triggers the continuation value

$$V_k^{AY} = \frac{\delta}{1-\delta} u_Y - d.$$

By contrast, if agent $i$ also follows the rule and all others continue with $X$, the cooperative continuation value is

$$W_m = \frac{\delta}{1-\delta} - d + (1+d)\delta^{m-1}, \qquad m \geq 1.$$

Therefore,

$$W_m - V_k^{AY} = \frac{\delta(1-u_Y)}{1-\delta} + (1+d)\delta^{m-1} > 0,$$

so no profitable one-shot deviation exists on the cooperative path either. Hence all four candidate rules are SPE. Because these four rules are behaviorally identical in this environment, I use $GT$ as shorthand for this common induced strategy.

## A3.2 Derivation of the Dynamic Basin

As in Section 2.3, agent $i$ believes that each other member $j \neq i$ is a *Grim Trigger* type with probability $p$ and an All-$Y$ type with probability $1-p$. Agent $i$'s continuation value as an All-$Y$ type is

$$V_k^{AY} = V_0^{AY} = \delta\big(u_Y + V_0^{AY}\big) + (1-\delta)(-d).$$

Hence,

$$V_k^{AY} = \frac{\delta}{1-\delta} u_Y - d.$$

Agent $i$'s continuation value as a *Grim Trigger* type has two components. First, if all other $n-1$ members are also *Grim Trigger*, which occurs with probability $p^{n-1}$, then all members coordinate on $X$ until the threshold is reached. Denote $m \equiv \max\{\bar{K} - k, 1\}$ as the distance from the current state $k$ to the threshold $\bar{K}$. The continuation value of the first part is therefore $W_m$, the same cooperative value as in Section A3.1. It satisfies

$$W_0 = W_1 = \frac{1}{1-\delta},$$

and for $2 \leq m \leq \bar{K}$,

$$W_m = \delta\big(1 + W_{m-1}\big) + (1-\delta)(-d).$$

Solving this recursion gives

$$W_m = \frac{\delta}{1-\delta} - d + (1+d)\delta^{m-1}.$$

Second, if at least one other member is an All-$Y$ type, which occurs with probability $1-p^{n-1}$, then coordination on $X$ fails immediately. In that case, agent $i$'s current-stage payoff from choosing $X$ is 0, and from the next period onward all Grim Trigger types switch to $Y$. Hence,

the continuation value of the second part is

$$\delta V_k^{AY} + (1-\delta)(-d) = \frac{\delta^2}{1-\delta} u_Y - d.$$

Therefore,

$$V_k^{GT} = p^{n-1} W_m + \left(1 - p^{n-1}\right) \left( \frac{\delta^2}{1-\delta} u_Y - d \right).$$

Solving $V_k^{GT} \geq V_k^{AY}$ gives

$$p^{n-1} \geq \frac{\delta u_Y}{\frac{\delta(1-\delta u_Y)}{1-\delta} + (1+d)\delta^{m-1}}.$$

Dividing the numerator and denominator by $\delta$,

$$p \geq p_k^\star \equiv \left( \frac{u_Y}{\frac{1-\delta u_Y}{1-\delta} + (1+d)\delta^{m-2}} \right)^{\frac{1}{n-1}}.$$

## A3.3 Equivalence with the Mixed-MPE Cutoff at $k = 0$

At $k = 0$, the distance to protection is $m = \max\{\bar{K}, 1\} = \bar{K}$. Multiplying the numerator and denominator of $(p_0^\star)^{n-1}$ by $(1-\delta)$, and for $\bar{K} = 1$ also by $\delta$,

$$\begin{aligned}
\bar{K} = 1: \quad (p_0^\star)^{n-1} &= \frac{\delta(1-\delta)u_Y}{\delta(1-\delta u_Y) + (1+d)(1-\delta)} = \frac{\delta(1-\delta)u_Y}{1 - \delta^2 u_Y + (1-\delta)d}, \\
\bar{K} = 2: \quad (p_0^\star)^{n-1} &= \frac{(1-\delta)u_Y}{1 - \delta u_Y + (1+d)(1-\delta)} = \frac{(1-\delta)u_Y}{(2-\delta) - \delta u_Y + (1-\delta)d}, \\
\bar{K} = 3: \quad (p_0^\star)^{n-1} &= \frac{(1-\delta)u_Y}{1 - \delta u_Y + (1+d)\delta(1-\delta)} = \frac{(1-\delta)u_Y}{1 + \delta(1-\delta)(1+d) - \delta u_Y}.
\end{aligned}$$

These equal $\sigma_0^{n-1}$ in the $(m1)$, $(m11)$, and $(m111)$ MPEs (Figure 1), so $p_0^\star = \sigma_0$ and $\mathcal{DB}_0 = 1 - \sigma_0$.

# Appendix B1: Additional Empirical Results for p($X$)

Table B1.1: Monotonicity in State: Alternative Specifications with Participant Controls

| | State increments | | | Linear state | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| State = 1 minus State = 0 | 0.063*** | 0.063*** | 0.063*** | | | |
| | (0.015) | (0.015) | (0.015) | | | |
| State = 2 minus State = 1 | 0.032* | 0.032* | 0.032* | | | |
| | (0.022) | (0.022) | (0.022) | | | |
| Difference between increments | 0.031* | 0.031* | 0.031* | | | |
| | (0.018) | (0.018) | (0.018) | | | |
| State | | | | 0.049*** | 0.049*** | 0.049*** |
| | | | | (0.016) | (0.016) | (0.016) |
| Risk-averse | | -0.021 | -0.022 | | -0.021 | -0.022 |
| | | (0.013) | (0.014) | | (0.013) | (0.014) |
| Prosocial | | 0.037* | 0.035 | | 0.038* | 0.036 |
| | | (0.018) | (0.021) | | (0.019) | (0.021) |
| Gender (1=Male) | | | -0.008 | | | -0.008 |
| | | | (0.020) | | | (0.020) |
| Major (1=Engineering/Science) | | | 0.019 | | | 0.020 |
| | | | (0.020) | | | (0.019) |
| High school language (1=English) | | | 0.022 | | | 0.022 |
| | | | (0.022) | | | (0.021) |
| Group-by-match fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Round-within-match fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 7,812 | 7,812 | 7,812 | 7,812 | 7,812 | 7,812 |
| $R^2$ | 0.567 | 0.569 | 0.570 | 0.566 | 0.569 | 0.570 |

*Notes:* *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Session-clustered standard errors are in parentheses. Equal weight per session. Inference uses 5,000 wild-cluster-bootstrap-$t$ replications.

As a secondary analysis, Table B1.1 reports alternative specifications that include elicited risk and social preferences and other participant characteristics. Columns (1)-(3) parameterize the state variable in terms of adjacent increments. The estimated increase in the probability assigned to $X$ from $k = 0$ to $k = 1$ is 0.063, while the additional increase from $k = 1$ to $k = 2$ is 0.032. Thus, the point estimates are monotonic in the state and indicate a larger increase from $k = 0$ to $k = 1$ than from $k = 1$ to $k = 2$. The difference between the two increments is 0.031. These estimates are essentially unchanged as participant controls are added across columns (1)-(3). Columns (4)-(6) provide a complementary specification in which the state enters linearly. The estimated coefficient on $k$ is 0.049, implying that a one-unit increase in the state is associated with an approximately 5 percentage-point increase in the probability assigned to $X$. The estimate is stable across specifications and statistically significant at the 1% level. Risk aversion is not significantly associated with $p(X)$, while

prosociality is positively associated with $p(X)$, although the latter relationship is only weakly statistically significant and is sensitive to the inclusion of additional participant controls.

Table B1.2: Initial versus Post-Reset Behavior Caused by Others at $k = 0$

| | Probability assigned to $X$ | | | | |
|---|---|---|---|---|---|
| | Pooled | 2l | 2h | 4l | 4h |
| | (1) | (2) | (3) | (4) | (5) |
| Post-reset $k = 0$, by others | 0.003 | 0.085*** | 0.011 | 0.026 | -0.088*** |
| | (0.022) | (0.023) | (0.064) | (0.021) | (0.025) |
| Risk-averse | -0.043* | -0.009 | -0.126* | -0.048 | 0.018 |
| | (0.024) | (0.065) | (0.061) | (0.030) | (0.063) |
| Prosocial | 0.055* | 0.098 | -0.066** | 0.051** | 0.127 |
| | (0.028) | (0.076) | (0.035) | (0.025) | (0.067) |
| Male | 0.004 | 0.020 | 0.023 | -0.036 | -0.009 |
| | (0.032) | (0.057) | (0.068) | (0.025) | (0.074) |
| Major | 0.021 | 0.056 | 0.034 | 0.055*** | -0.052 |
| (1=Engineering/Science) | (0.030) | (0.077) | (0.094) | (0.013) | (0.061) |
| High school language | 0.036 | -0.006 | 0.039 | 0.102 | 0.004 |
| (1=English) | (0.034) | (0.068) | (0.068) | (0.046) | (0.074) |
| Group-by-match fixed effects | Yes | Yes | Yes | Yes | Yes |
| Observations | 2,723 | 443 | 467 | 830 | 983 |
| Subjects | 164 | 68 | 68 | 96 | 96 |
| Sessions | 12 | 6 | 6 | 6 | 6 |
| Groups | 580 | 170 | 170 | 120 | 120 |
| $R^2$ | 0.382 | 0.399 | 0.518 | 0.238 | 0.274 |

*Notes:* The sample contains initial $k = 0$ observations and post-reset $k = 0$ observations for which the participant chose $X$ in the preceding round, so one or more other group members caused the reset. Initial $k = 0$ is the omitted category. Treatment columns are estimated separately. Subject and round-within-match fixed effects are omitted. Analytic weights give each eligible session–treatment cell equal total weight. SEs are clustered by session. Groups are distinct group-by-match cells. Stars use two-sided wild-cluster-bootstrap-$t$ tests with 5,000 replications: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Treatment-specific inference uses six sessions and should be interpreted cautiously.

Table B1.2 compares initial $k = 0$ behavior with behavior following a reset caused by others, replacing subject fixed effects with participant characteristics. A post-reset observation is included only when the participant chose $X$ in the preceding round. Without subject fixed effects, this sample over-represents participants with a high propensity to choose $X$, because such participants are more likely to have chosen $X$ in the failed round (Section 4.2). The post-reset coefficients are therefore likely biased upward relative to the within-subject estimates in Table 9. Consistent with this, the pooled coefficient is small and insignificant (0.003), and the estimate in $2l$ is positive (0.085, $p < 0.01$). The decline in $4h$ nevertheless remains significant ($-0.088$, $p < 0.01$), so the main qualitative finding, a post-failure decline

concentrated in $4h$, holds even in a specification biased against finding a decline. Treatment-specific results should be interpreted cautiously because each relies on six sessions.

Figure B1.1: Composition of Groups That Do Not Reach the Threshold ($k = 2$)

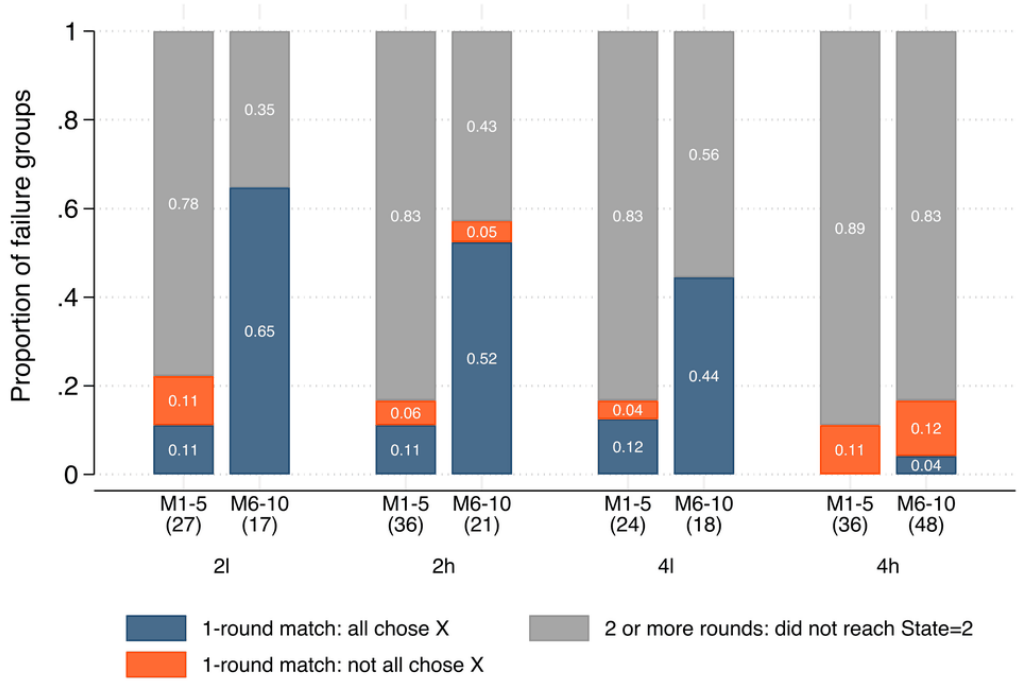



Table B1.3: Balance of Subject Characteristics: $n = 4$ Treatment and Baseline

| Subject Characteristic | Treatment | | Baseline | | $t$-statistic |
|---|---|---|---|---|---|
| | Mean | (SE) | Mean | (SE) | |
| Gender (1=Male) | 0.49 | (0.03) | 0.63 | (0.06) | -1.96 |
| Major (1=Engineering/Science) | 0.57 | (0.06) | 0.41 | (0.16) | 0.99 |
| High school language (1=English) | 0.88 | (0.05) | 0.97 | (0.03) | -1.63 |
| Risk averse (1=Yes) | 0.44 | (0.04) | 0.41 | (0.09) | 0.31 |
| Social Preference (1=Prosocial) | 0.61 | (0.03) | 0.53 | (0.03) | 1.81 |
| Sessions | 6 | | 2 | | |
| Subjects | 96 | | 32 | | |

*Notes:* Each characteristic is first averaged within an independent session. Standard errors are standard errors of the session means. The final column reports the unequal-variance $t$-statistic for treatment minus baseline as a descriptive balance diagnostic.

Figure B1.2: Outcomes among Groups That Reached the Threshold ($k = 2$)

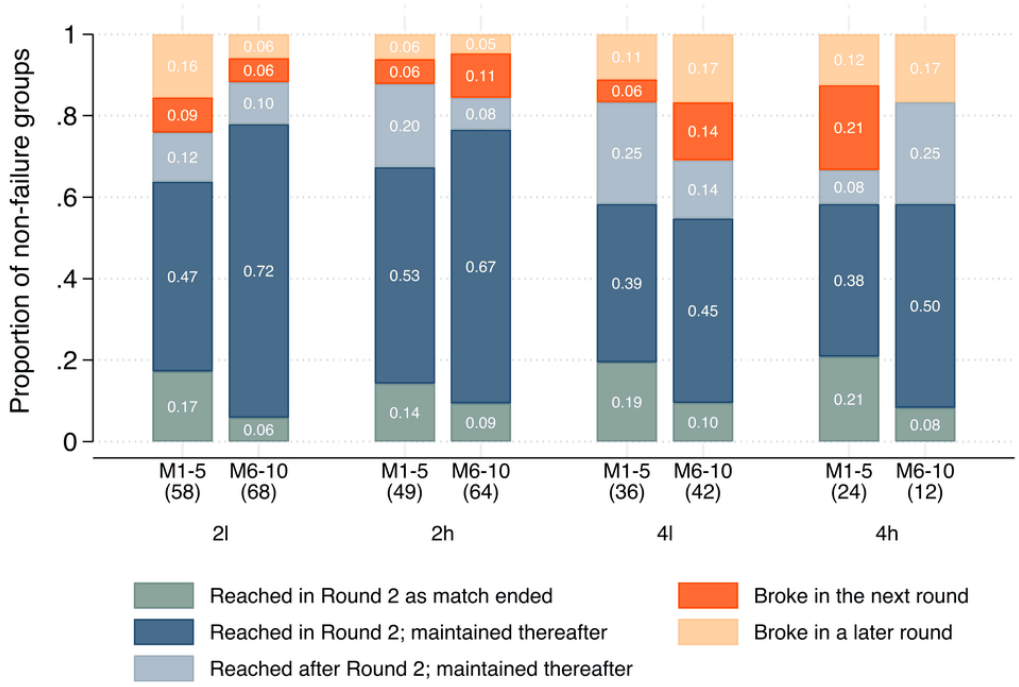
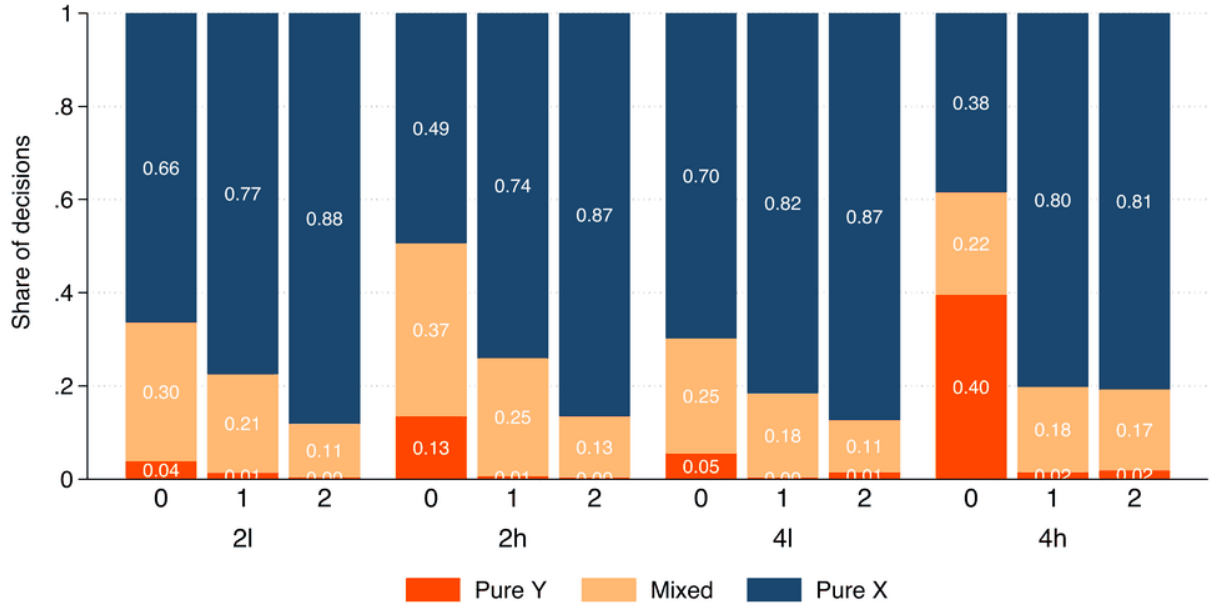



Figure B1.3: Decisions by State and Treatment



# Appendix B2: Additional Empirical Results for All-$X$

This appendix reports corresponding analyses using realized all-$X$ coordination as a secondary group-level outcome.

Figure B2.1: Proportion of Group-Rounds with All-$X$ Coordination by State

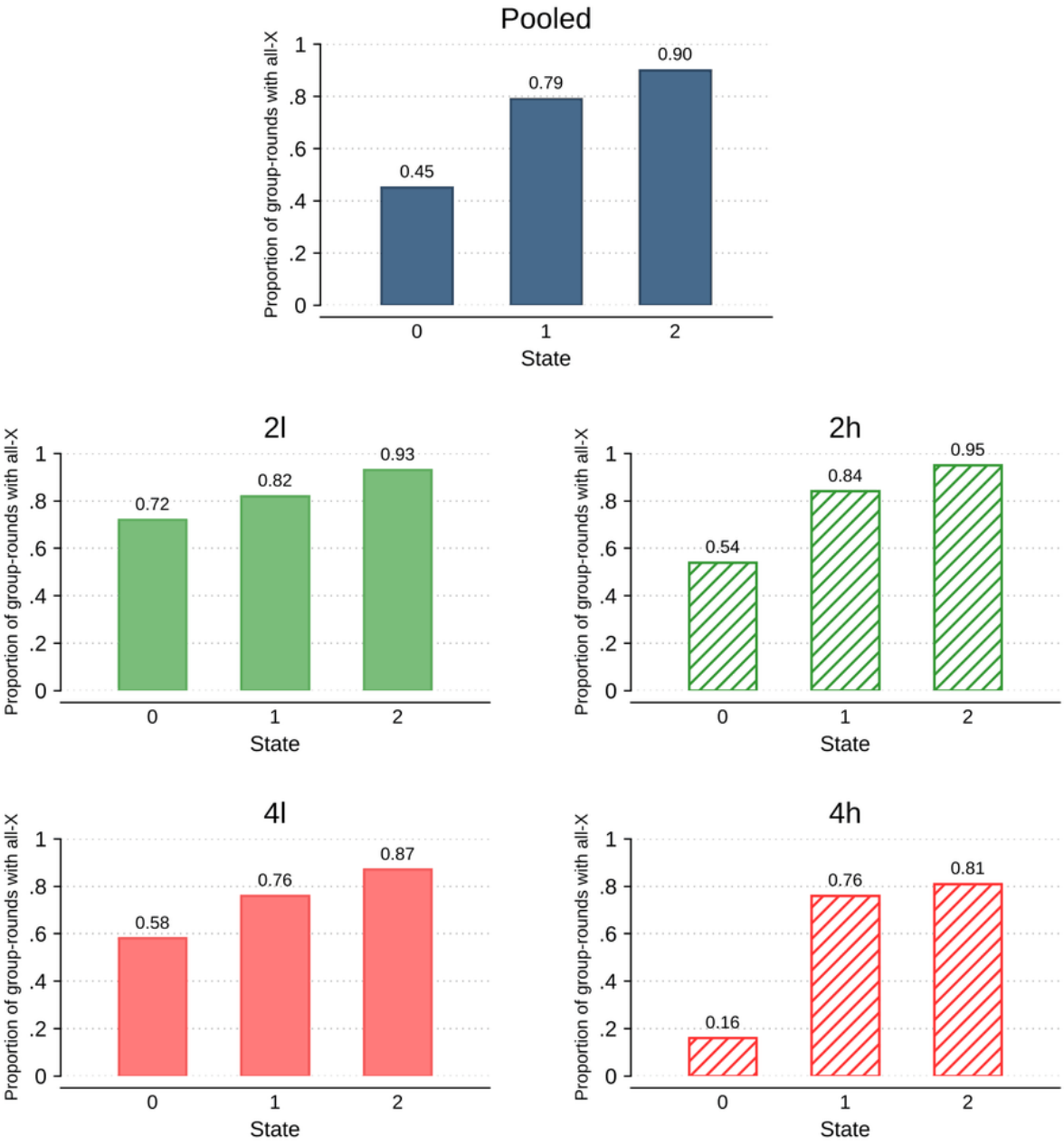
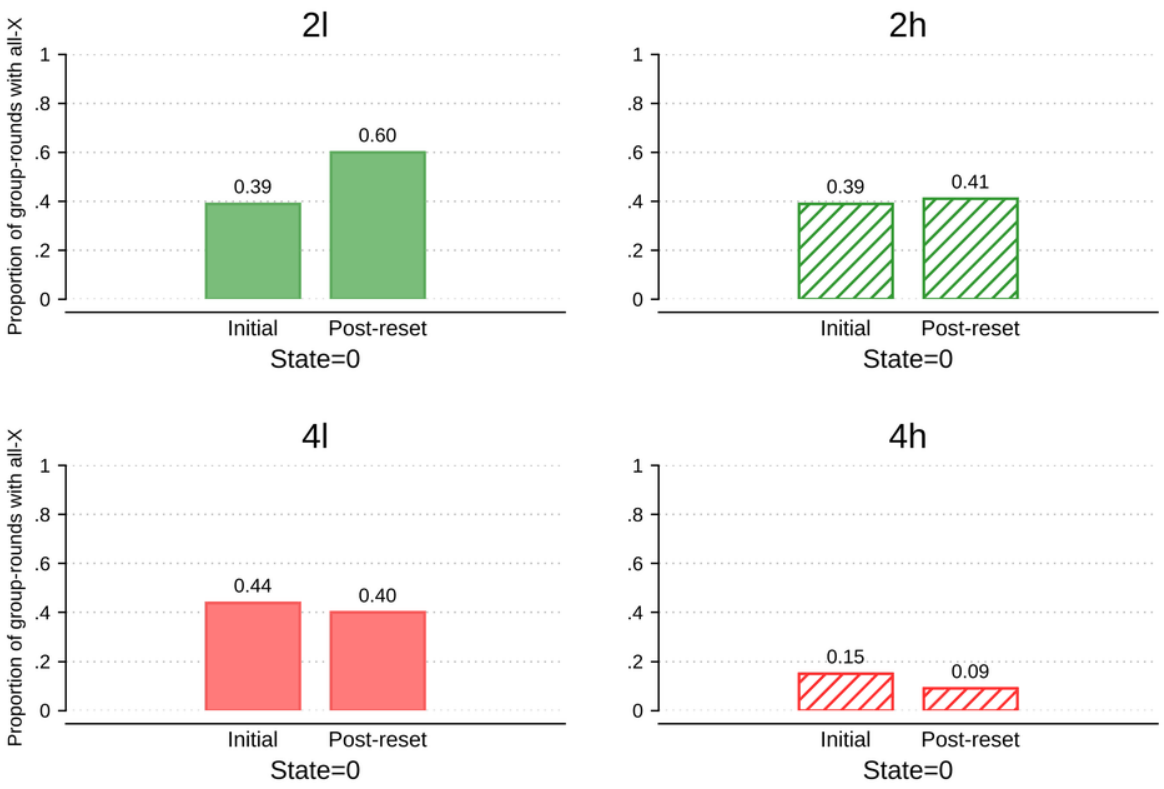


Figure B2.2: Proportion of Group-Rounds with All-$X$ Coordination at Initial and Post-Reset $k = 0$ by Treatment




*Notes:* The sample is restricted to group-by-match cells containing both the initial Round-1 $k = 0$ observation and at least one post-reset $k = 0$ observation.

Table B2.1: Group Size and Realized All-$X$ Coordination

| | Low $u_Y$ | | High $u_Y$ | |
|---|---|---|---|---|
| | Unadjusted (1) | Adjusted (2) | Unadjusted (3) | Adjusted (4) |
| $n = 2$ minus $n = 4$ | 0.109 | 0.109 | 0.415** | 0.415** |
| | (0.082) | (0.060) | (0.112) | (0.101) |
| High-first | No | Yes | No | Yes |
| Sequence fixed effects | No | Yes | No | Yes |
| Sessions | 12 | 12 | 12 | 12 |

*Notes:* The binary outcome equals one when all group members choose X. One observation is retained per group-round, then averaged within each session and payoff condition. Standard errors are in parentheses. Adjusted columns include High-first and sequence fixed effects. Stars use one-sided exact randomization tests (64 assignments). *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Positive estimates are in the direction predicted by Hypothesis 3a.

Table B2.2: Safe-Option Payoff and Realized All-$X$ Coordination

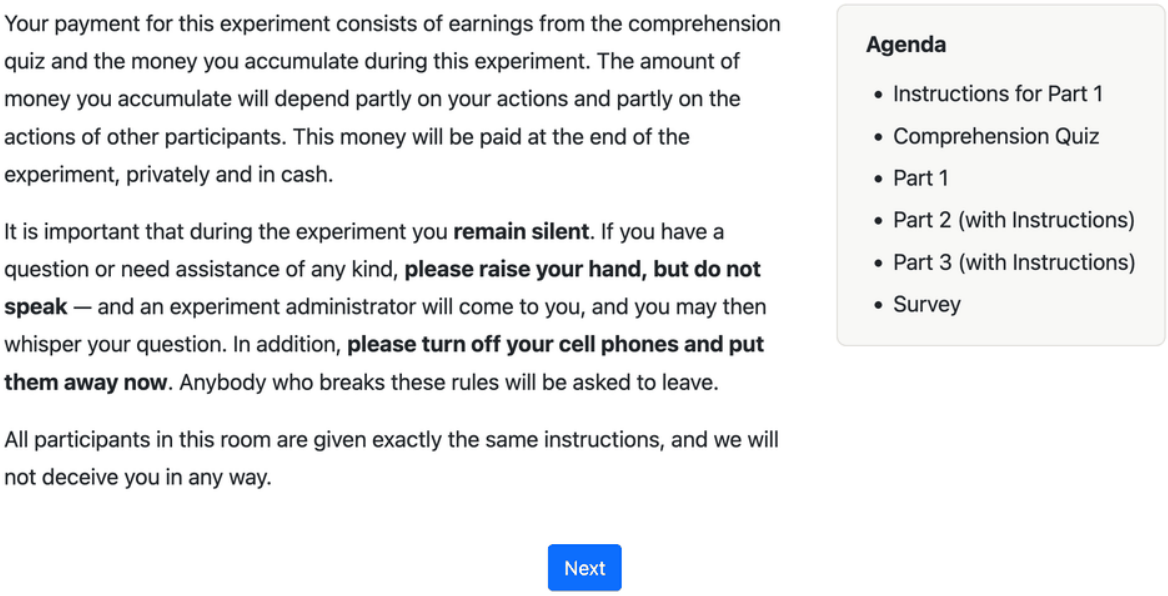
| | $n = 2$ | | | $n = 4$ | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Low $u_Y$ minus High $u_Y$ | 0.112** | 0.112 | 0.169 | 0.417*** | 0.417*** | 0.293 |
| | (0.040) | (0.091) | (0.041) | (0.071) | (0.073) | (0.180) |
| Session fixed effects | Yes | Yes | | Yes | Yes | |
| Absolute match-number FE | Yes | | Yes | Yes | | Yes |
| Relative match-number FE | | Yes | | | Yes | |
| First-half sample | No | No | Yes | No | No | Yes |
| Group-rounds | 1,618 | 1,618 | 790 | 1,144 | 1,144 | 560 |
| Sessions | 6 | 6 | 6 | 6 | 6 | 6 |
| $R^2$ | 0.103 | 0.048 | 0.044 | 0.286 | 0.283 | 0.090 |

*Notes:* The binary outcome equals one when all group members choose X; one observation is retained per group-round. Full-sample columns use equal-session analytic weights, session fixed effects, and session-clustered standard errors. Stars use one-sided wild-cluster-bootstrap-$t$ tests in columns (1), (2), (4), and (5), and one-sided exact session-level permutation tests in columns (3) and (6). First-half columns use equal-session weights and behavior before exposure to the alternative payoff. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

# Appendix C1: Experimental Instructions and Interfaces

*The decisions, random draws, player tags, outcomes, and payment amounts shown in the screenshots are illustrative examples generated for interface documentation.*

## C1.1 Experiment Overview



## C1.2 Instructions for Part 1 ($n = 2$ treatment sessions)

Please read the instructions carefully. A comprehension quiz follows. **Each correct answer earns you $0.50.**

### [How a Match Works]

**Part 1 consists of 10 matches.** At the beginning of each match, you will be randomly grouped with 1 other participant in this room. You will then play a number of rounds with that 1 participant. This is what we call a "match". The two of you form a "group". Once a match ends, you will be randomly grouped with 1 other participant for a new match. You will not be able to identify who you've interacted with in previous or future matches.

The length of a match, that is, the number of rounds in a match, is randomly determined as follows: **After each round, there is an 80% probability that the match will continue for at least another round.** Specifically, after each round, whether the match continues for another round will be determined by a spinner with 80% white area and 20% grey area. If the arrow stops in the white area, the match will continue for at least another round; Otherwise, it will end. For example, if you are in round 2, the probability that there will be a third round is 80%, and if you are in round 5, the probability that there will be a sixth round is also 80%. At any point in a match, the probability that the match will continue is 80%.



✓ The match continues to the next round.

(arrow in white area)



✕ The match ends at the end of this round.

(arrow in grey area)

### [Payment for Part 1]

All payoffs in Part 1 are in units of experimental points and will be converted to US Dollars at the rate of 400 Experimental Points = 1 US Dollar. You will have an endowment of 100 points at the beginning of each match. The total earnings of a match are the sum of the endowment and the payoffs received in each round of that match. **Your payment for Part 1 is the sum of your total earnings across all 10 matches.**

## [Actions and Payoffs]

Each round, you and 1 other matched participant will each choose X or Y individually and simultaneously. Your round payoff is determined by your action and the group outcome, as shown in the payoff table. The group outcome shows each member's action in the order of Players 1 and 2 (P1 P2). The group player tag is randomly assigned to you at the beginning of each match and remains the same within that match.

Note: The payoffs here are for illustration only. You will be informed of the real payoffs at the beginning of each match.

| Your action | Group outcome (P1 P2) | Your payoff |
|---|---|---|
| X | X X | 100 |
| | At least one action is Y | 0 |
| Y | At least one action is Y | 50 |

**Rather than directly choosing between X and Y, you assign probabilities to the two actions, and the computer decides which action to take based on these probabilities**. You can think of this as allocating an area between the two actions. Below are 100 grids for you to allocate between Yellow (X) and Blue (Y) using a slider. Drag the slider to choose how many grids to allocate to each action. After you confirm your allocation, the computer will randomly draw an integer number between 1 and 100.

- If the number is in the Yellow area, your action is X.
- If the number is in the Blue area, your action is Y.

Let's try it. **Click anywhere on the slider to activate it.** Then, drag the slider thumb to your preferred position and click the Confirm Decision button to see the result. This is a practice, so you may drag the slider and try as many times as you like. **In real matches, however, you CANNOT change the probabilities once you click the Confirm Decision button.** Click Next when you are ready to move on.




### [The Public Counter]

**The public counter tracks the number of consecutive rounds in which the group outcome is X X among the two most recent rounds.** You will see the public counter at the beginning of each round before you drag the slider. The public counter is initially set to 0 and shown at the beginning of Round 1, since there is no previous round. At the end of each round, the public counter is updated based on the group outcome and carried over to the beginning of the next round:

- If group outcome is X X, the counter **increases by 1, up to a maximum of 2**.
- If at least one action in the group outcome is Y, the counter **resets to 0**.




### [Payoff in the Last Round of a Match]

As stated before, your round payoffs are determined by the payoff table. **However, your payoff in the last round of a match may be affected by the public counter.** At the end of the last round:

- If the counter **equals 2**, your payoff in the last round is determined by the payoff table.
- If the counter **is below 2**, everyone's payoff in the last round **is overridden to -100**, regardless of what is shown in the payoff table.

Note that you don't know whether a round is the last round of that match when you assign probabilities to X and Y in that round.

## C1.3 Instructions for Part 1 ($n = 4$ treatment sessions)

Please read the instructions carefully. A comprehension quiz follows. **Each correct answer earns you $0.50.**

**[How a Match Works]**

**Part 1 consists of 10 matches.** At the beginning of each match, you will be randomly grouped with 3 other participants in this room. You will then play a number of rounds with those 3 participants. This is what we call a "match". The four of you form a "group". Once a match ends, you will be randomly grouped with 3 other participants for a new match. You will not be able to identify who you've interacted with in previous or future matches.

The length of a match, that is, the number of rounds in a match, is randomly determined as follows: **After each round, there is an 80% probability that the match will continue for at least another round.** Specifically, after each round, whether the match continues for another round will be determined by a spinner with 80% white area and 20% grey area. If the arrow stops in the white area, the match will continue for at least another round; Otherwise, it will end. For example, if you are in round 2, the probability that there will be a third round is 80%, and if you are in round 5, the probability that there will be a sixth round is also 80%. At any point in a match, the probability that the match will continue is 80%.



✓ The match continues to the next round.

(arrow in white area)



✕ The match ends at the end of this round.

(arrow in grey area)

**[Payment for Part 1]**

All payoffs in Part 1 are in units of experimental points and will be converted to US Dollars at the rate of 400 Experimental Points = 1 US Dollar. You will have an endowment of 100 points at the beginning of each match. The total earnings of a match are the sum of the endowment and the payoffs received in each round of that match. **Your payment for Part 1 is the sum of your total earnings across all 10 matches.**

## [Actions and Payoffs]

Each round, you and 3 other matched participants will each choose X or Y individually and simultaneously. Your round payoff is determined by your action and the group outcome, as shown in the payoff table. The group outcome shows each member's action in the order of Players 1, 2, 3, and 4 (P1 P2 P3 P4). The group player tag is randomly assigned to you at the beginning of each match and remains the same within that match.

Note: The payoffs here are for illustration only. You will be informed of the real payoffs at the beginning of each match.

| Your action | Group outcome (P1 P2 P3 P4) | Your payoff |
|---|---|---|
| X | X X X X | 100 |
| | At least one action is Y | 0 |
| Y | At least one action is Y | 50 |

**Rather than directly choosing between X and Y, you assign probabilities to the two actions, and the computer decides which action to take based on these probabilities**. You can think of this as allocating an area between the two actions. Below are 100 grids for you to allocate between Yellow (X) and Blue (Y) using a slider. Drag the slider to choose how many grids to allocate to each action. After you confirm your allocation, the computer will randomly draw an integer number between 1 and 100.

- If the number is in the Yellow area, your action is X.
- If the number is in the Blue area, your action is Y.

Let's try it. **Click anywhere on the slider to activate it.** Then, drag the slider thumb to your preferred position and click the Confirm Decision button to see the result. This is a practice, so you may drag the slider and try as many times as you like. **In real matches, however, you CANNOT change the probabilities once you click the Confirm Decision button.** Click Next when you are ready to move on.




### [The Public Counter]

**The public counter tracks the number of consecutive rounds in which the group outcome is X X X X among the two most recent rounds.** You will see the public counter at the beginning of each round before you drag the slider. The public counter is initially set to 0 and shown at the beginning of Round 1, since there is no previous round. At the end of each round, the public counter is updated based on the group outcome and carried over to the beginning of the next round:

- If group outcome is X X X X, the counter **increases by 1, up to a maximum of 2**.
- If at least one action in the group outcome is Y, the counter **resets to 0**.




### [Payoff in the Last Round of a Match]

As stated before, your round payoffs are determined by the payoff table. **However, your payoff in the last round of a match may be affected by the public counter.** At the end of the last round:

- If the counter **equals 2**, your payoff in the last round is determined by the payoff table.
- If the counter **is below 2**, everyone's payoff in the last round **is overridden to -100**, regardless of what is shown in the payoff table.

Note that you don't know whether a round is the last round of that match when you assign probabilities to X and Y in that round.

## C1.4 Instructions for Part 1 ($n = 2$ baseline sessions)

Please read the instructions carefully. A comprehension quiz follows. **Each correct answer earns you $0.50.**

**[How a Match Works]**

**Part 1 consists of 10 matches.** At the beginning of each match, you will be randomly grouped with 1 other participant in this room. You will then play a number of rounds with that 1 participant. This is what we call a "match". The two of you form a "group". Once a match ends, you will be randomly grouped with 1 other participant for a new match. You will not be able to identify who you've interacted with in previous or future matches.

The length of a match, that is, the number of rounds in a match, is randomly determined as follows: **After each round, there is an 80% probability that the match will continue for at least another round.** Specifically, after each round, whether the match continues for another round will be determined by a spinner with 80% white area and 20% grey area. If the arrow stops in the white area, the match will continue for at least another round; Otherwise, it will end. For example, if you are in round 2, the probability that there will be a third round is 80%, and if you are in round 5, the probability that there will be a sixth round is also 80%. At any point in a match, the probability that the match will continue is 80%.



✓ The match continues to the next round.

(arrow in white area)



✕ The match ends at the end of this round.

(arrow in grey area)

**[Payment for Part 1]**

All payoffs in Part 1 are in units of experimental points and will be converted to US Dollars at the rate of 400 Experimental Points = 1 US Dollar. You will have an endowment of 100 points at the beginning of each match. The total earnings of a match are the sum of the endowment and the payoffs received in each round of that match. **Your payment for Part 1 is the sum of your total earnings across all 10 matches.**

## [Actions and Payoffs]

Each round, you and 1 other matched participant will each choose X or Y individually and simultaneously. Your round payoff is determined by your action and the group outcome, as shown in the payoff table. The group outcome shows each member's action in the order of Players 1 and 2 (P1 P2). The group player tag is randomly assigned to you at the beginning of each match and remains the same within that match.

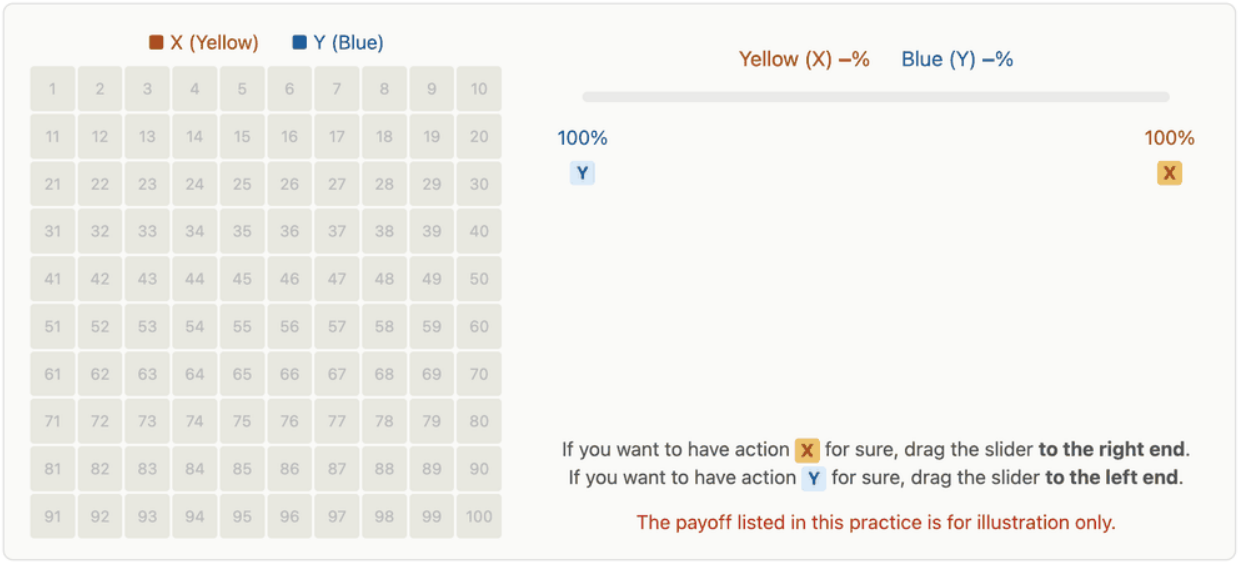
Note: The payoffs here are for illustration only. You will be informed of the real payoffs at the beginning of each match.

| Your action | Group outcome (P1 P2) | Your payoff |
|---|---|---|
| X | X X | 100 |
| | At least one action is Y | 0 |
| Y | At least one action is Y | 50 |

**Rather than directly choosing between X and Y, you assign probabilities to the two actions, and the computer decides which action to take based on these probabilities**. You can think of this as allocating an area between the two actions. Below are 100 grids for you to allocate between Yellow (X) and Blue (Y) using a slider. Drag the slider to choose how many grids to allocate to each action. After you confirm your allocation, the computer will randomly draw an integer number between 1 and 100.

- If the number is in the Yellow area, your action is X.
- If the number is in the Blue area, your action is Y.

Let's try it. **Click anywhere on the slider to activate it.** Then, drag the slider thumb to your preferred position and click the Confirm Decision button to see the result. This is a practice, so you may drag the slider and try as many times as you like. **In real matches, however, you CANNOT change the probabilities once you click the Confirm Decision button.** Click Next when you are ready to move on.


X (Yellow)
Y (Blue)
1 2 3 4 5 6 7 8 9 10
11 12 13 14 15 16 17 18 19 20
21 22 23 24 25 26 27 28 29 30
31 32 33 34 35 36 37 38 39 40
41 42 43 44 45 46 47 48 49 50
51 52 53 54 55 56 57 58 59 60
61 62 63 64 65 66 67 68 69 70
71 72 73 74 75 76 77 78 79 80
81 82 83 84 85 86 87 88 89 90
91 92 93 94 95 96 97 98 99 100
Yellow (X) –%
Blue (Y) –%
100%
Y
100%
X
If you want to have action X for sure, drag the slider to the right end.
If you want to have action Y for sure, drag the slider to the left end.
The payoff listed in this practice is for illustration only.

## C1.5 Instructions for Part 1 ($n = 4$ baseline sessions)

Please read the instructions carefully. A comprehension quiz follows. **Each correct answer earns you $0.50.**

**[How a Match Works]**

**Part 1 consists of 10 matches.** At the beginning of each match, you will be randomly grouped with 3 other participants in this room. You will then play a number of rounds with those 3 participants. This is what we call a "match". The four of you form a "group". Once a match ends, you will be randomly grouped with 3 other participants for a new match. You will not be able to identify who you've interacted with in previous or future matches.

The length of a match, that is, the number of rounds in a match, is randomly determined as follows: **After each round, there is an 80% probability that the match will continue for at least another round.** Specifically, after each round, whether the match continues for another round will be determined by a spinner with 80% white area and 20% grey area. If the arrow stops in the white area, the match will continue for at least another round; Otherwise, it will end. For example, if you are in round 2, the probability that there will be a third round is 80%, and if you are in round 5, the probability that there will be a sixth round is also 80%. At any point in a match, the probability that the match will continue is 80%.



✓ The match continues to the next round.

(arrow in white area)



✕ The match ends at the end of this round.

(arrow in grey area)

**[Payment for Part 1]**

All payoffs in Part 1 are in units of experimental points and will be converted to US Dollars at the rate of 400 Experimental Points = 1 US Dollar. You will have an endowment of 100 points at the beginning of each match. The total earnings of a match are the sum of the endowment and the payoffs received in each round of that match. **Your payment for Part 1 is the sum of your total earnings across all 10 matches.**

## [Actions and Payoffs]

Each round, you and 3 other matched participants will each choose X or Y individually and simultaneously. Your round payoff is determined by your action and the group outcome, as shown in the payoff table. The group outcome shows each member's action in the order of Players 1, 2, 3, and 4 (P1 P2 P3 P4). The group player tag is randomly assigned to you at the beginning of each match and remains the same within that match.

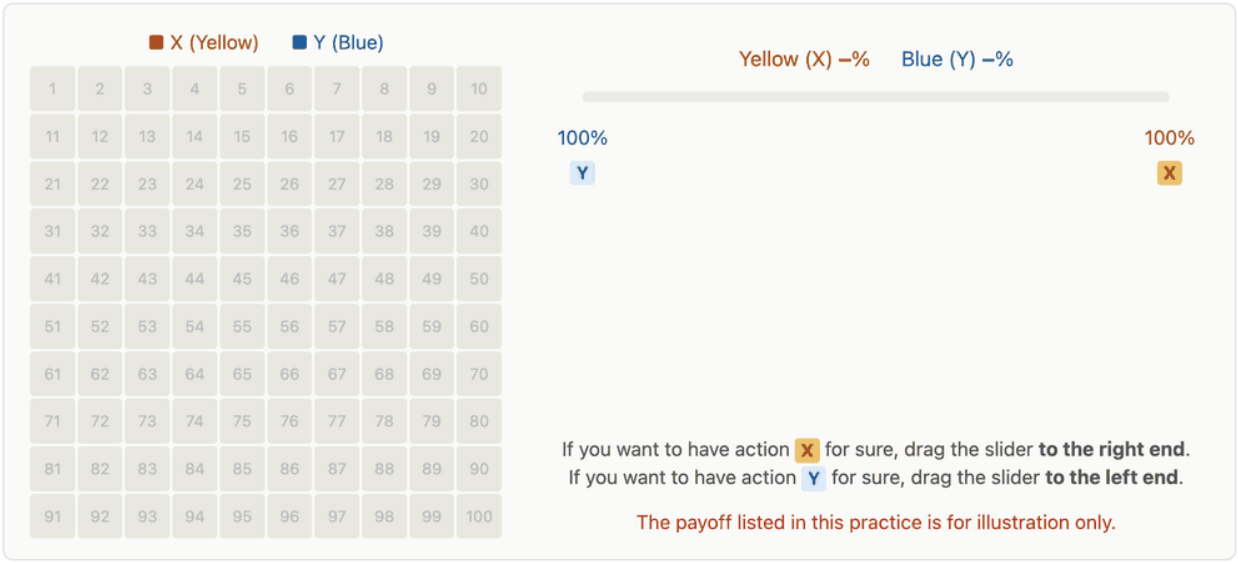
Note: The payoffs here are for illustration only. You will be informed of the real payoffs at the beginning of each match.

| Your action | Group outcome (P1 P2 P3 P4) | Your payoff |
|---|---|---|
| X | X X X X | 100 |
| | At least one action is Y | 0 |
| Y | At least one action is Y | 50 |

**Rather than directly choosing between X and Y, you assign probabilities to the two actions, and the computer decides which action to take based on these probabilities**. You can think of this as allocating an area between the two actions. Below are 100 grids for you to allocate between Yellow (X) and Blue (Y) using a slider. Drag the slider to choose how many grids to allocate to each action. After you confirm your allocation, the computer will randomly draw an integer number between 1 and 100.

- If the number is in the Yellow area, your action is X.
- If the number is in the Blue area, your action is Y.

Let's try it. **Click anywhere on the slider to activate it.** Then, drag the slider thumb to your preferred position and click the Confirm Decision button to see the result. This is a practice, so you may drag the slider and try as many times as you like. **In real matches, however, you CANNOT change the probabilities once you click the Confirm Decision button.** Click Next when you are ready to move on.


X (Yellow)
Y (Blue)
1 2 3 4 5 6 7 8 9 10
11 12 13 14 15 16 17 18 19 20
21 22 23 24 25 26 27 28 29 30
31 32 33 34 35 36 37 38 39 40
41 42 43 44 45 46 47 48 49 50
51 52 53 54 55 56 57 58 59 60
61 62 63 64 65 66 67 68 69 70
71 72 73 74 75 76 77 78 79 80
81 82 83 84 85 86 87 88 89 90
91 92 93 94 95 96 97 98 99 100
Yellow (X) –%
Blue (Y) –%
100%
Y
100%
X
If you want to have action X for sure, drag the slider to the right end.
If you want to have action Y for sure, drag the slider to the left end.
The payoff listed in this practice is for illustration only.

## C1.6 Comprehension Quiz

*To avoid repetition, all screenshots in this part show interfaces where participants select incorrect answers; correct answers are marked in green. Some quiz questions apply only to particular sessions, as indicated by the figure titles. For example, "$n = 2$ sessions" refers to both $n = 2$ treatment and baseline sessions, while "baseline sessions" refers to both $n = 2$ and $n = 4$ baseline sessions. Figures without such labels apply to all sessions.*

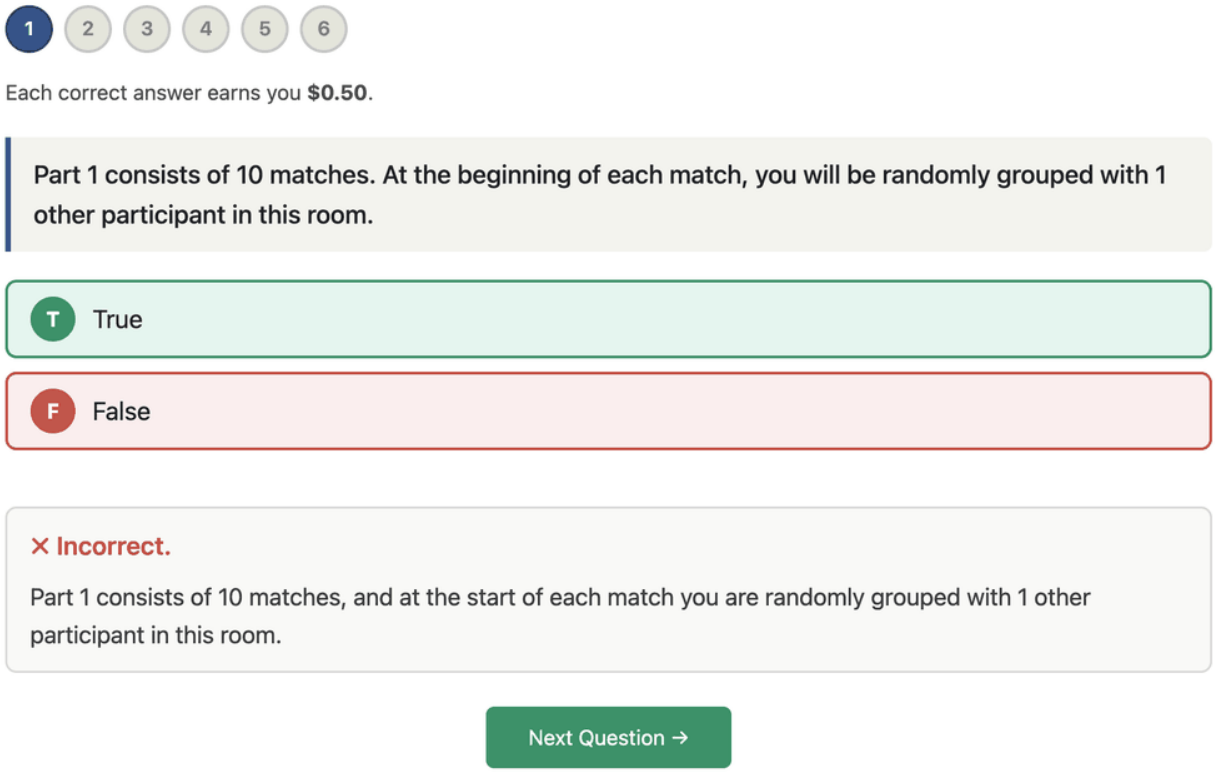



Figure C1.6.1: Question 1 ($n = 2$ sessions)

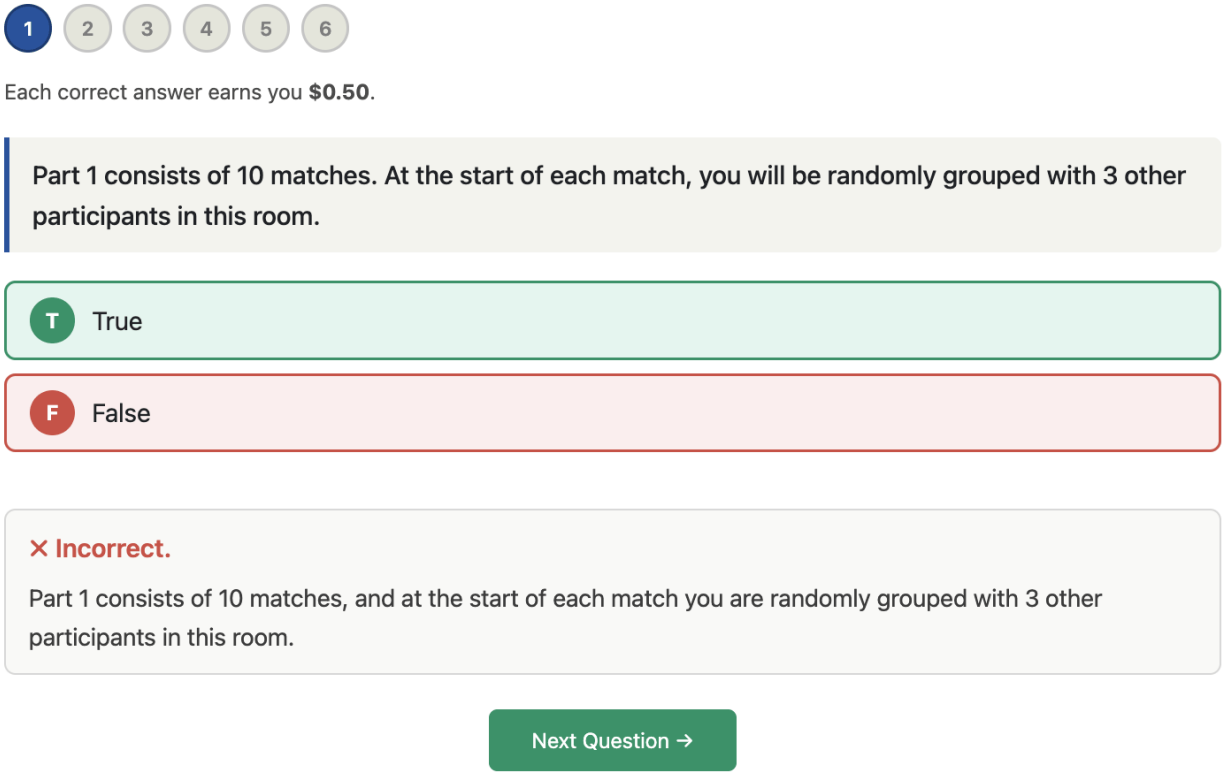



Figure C1.6.2: Question 1 ($n = 4$ sessions)

1 2 3 4 5 6

Each correct answer earns you **$0.50**.

At the end of each round, the computer spins an arrow on a 20% grey-80% white pie. Suppose in this round, the arrow stops in the white area. Then, the match ends at the end of this round.



T True

F False

✕ Incorrect.

If the arrow stops in the white area, the match will continue for at least another round; Otherwise, it will end. In this round the arrow stops in the white area, so the match does NOT end at the end of this round.

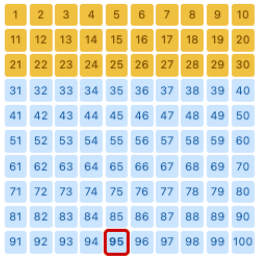
Next Question →

Figure C1.6.3: Question 2

1 2 3 4 5 6

Each correct answer earns you **$0.50**.

**Suppose this is NOT the last round of a match. You allocate the 100 grids as 1-30 for Yellow (X) and 31-100 for Blue (Y). The integer drawn by the computer is 95. Which of the following options is correct?**




Integer drawn: **95**

| Your action | Group outcome (P1 P2) | Your payoff |
|---|---|---|
| X | X X | **100** |
| | At least one action is Y | **0** |
| Y | At least one action is Y | **50** |

A Your action is Y, and your round payoff can be 100, 0, or 50.

B Your action is Y, and your round payoff is 50.

C Your action is X, and your round payoff is 100.

D Your action is X, and your round payoff is 0.

**✕ Incorrect.**

The integer drawn by the computer is 95, which is in the Blue area (grids 31-100), so your action is Y. At least one action in the group outcome is Y, so your round payoff is 50.

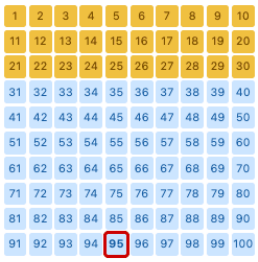
Next Question →

Figure C1.6.4: Question 3 ($n = 2$ sessions)

1 2 3 4 5 6

Each correct answer earns you **$0.50**.

Suppose this is NOT the last round of a match. You allocate the 100 grids as 1-30 for Yellow ( X ) and 31-100 for Blue ( Y ). The integer drawn by the computer is 95. Which of the following options is correct?




Integer drawn: **95**

| Your action | Group outcome (P1 P2 P3 P4) | Your payoff |
|---|---|---|
| X | X X X X | **100** |
| | At least one action is Y | **0** |
| Y | At least one action is Y | **50** |

A Your action is Y, and your round payoff can be 100, 0, or 50.

B Your action is Y, and your round payoff is 50.

C Your action is X, and your round payoff is 100.

D Your action is X, and your round payoff is 0.

✕ **Incorrect.**

The integer drawn by the computer is 95, which is in the Blue area (grids 31-100), so your action is Y. At least one action in the group outcome is Y, so your round payoff is 50.

Next Question →

Figure C1.6.5: Question 3 ($n = 4$ sessions)

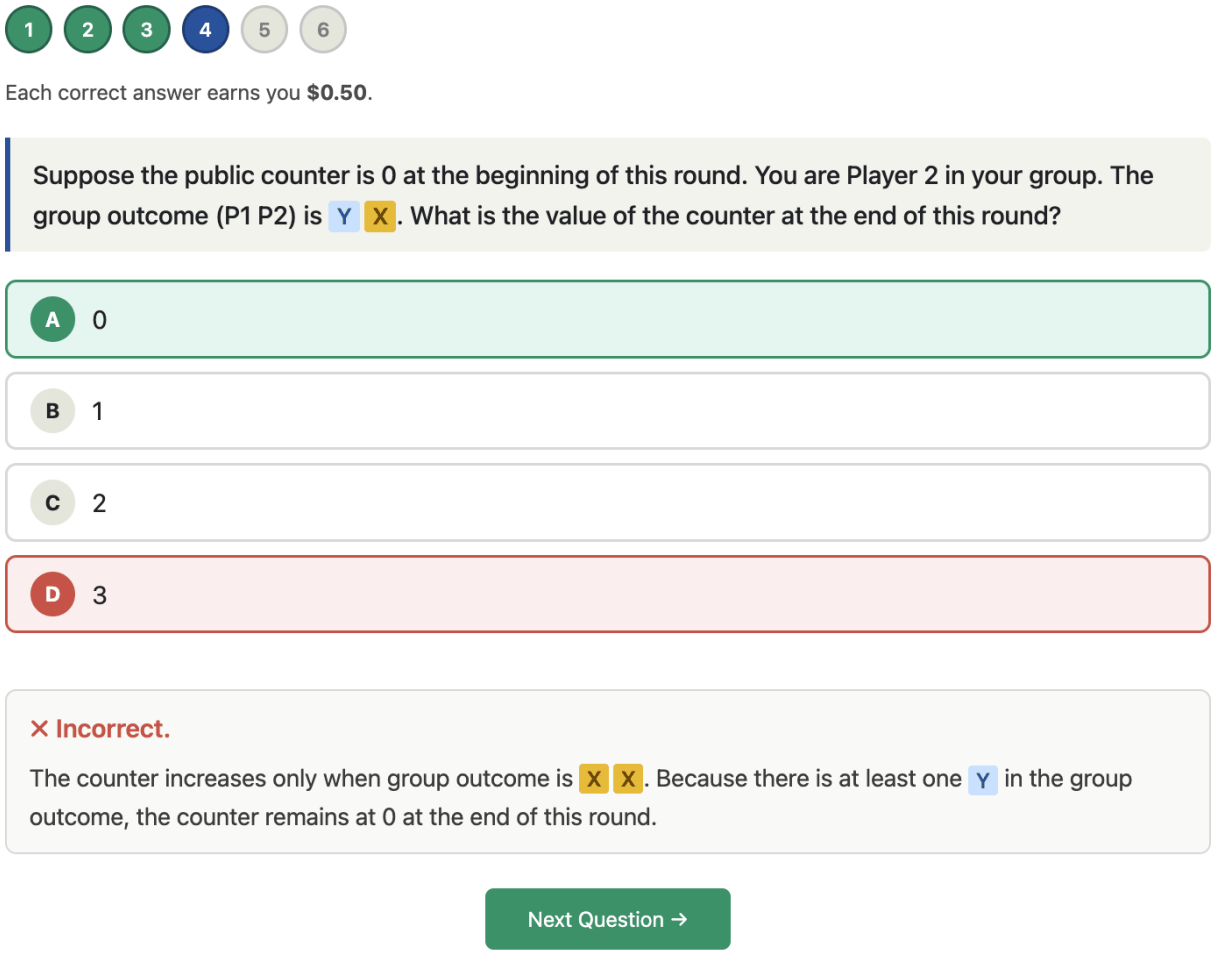



Figure C1.6.6: Question 4 ($n = 2$ treatment sessions)

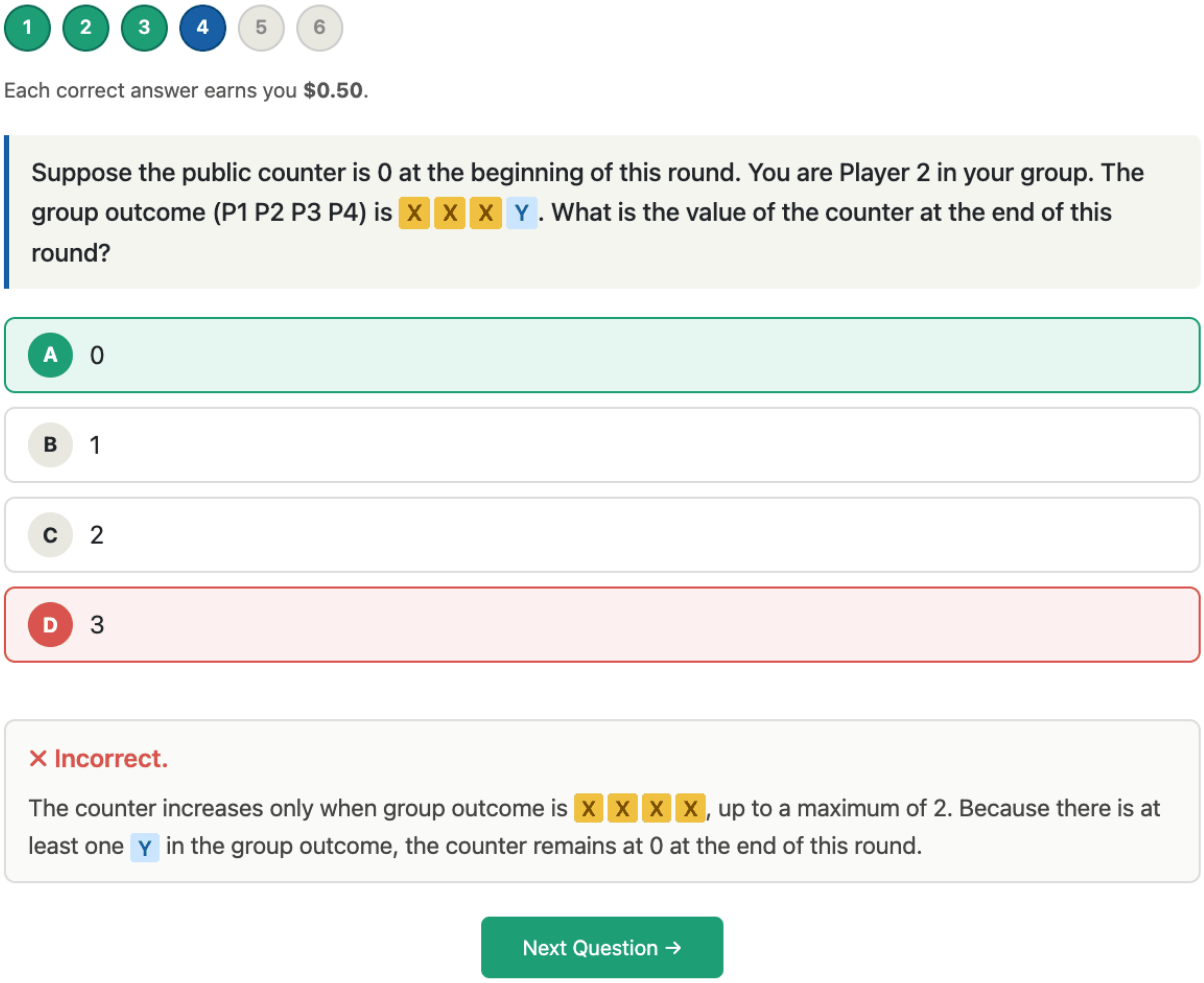



Figure C1.6.7: Question 4 ($n = 4$ treatment sessions)

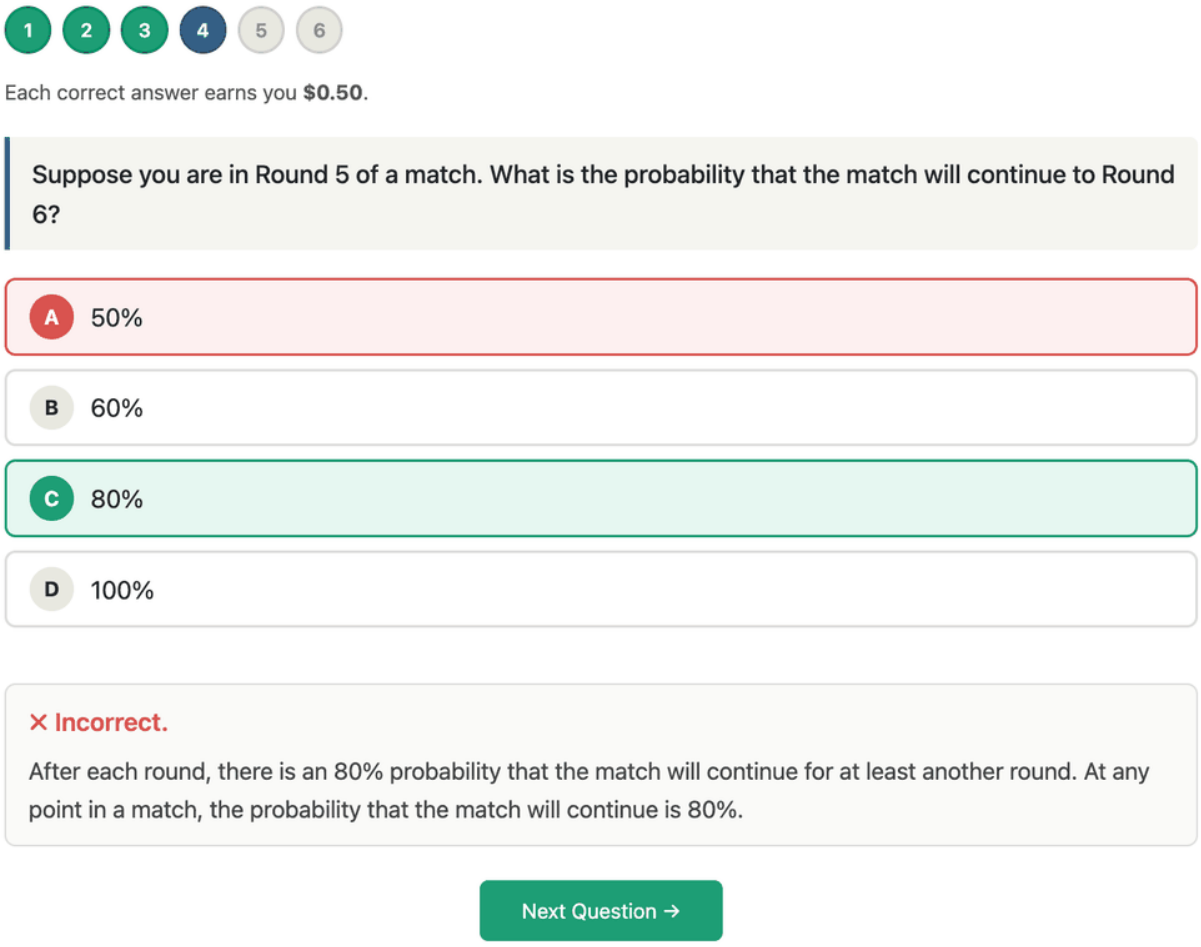



Figure C1.6.8: Question 4 (baseline sessions)

1 2 3 4 5 6

Each correct answer earns you **$0.50**.

| Your action | Group outcome (P1 P2) | Your payoff |
|---|---|---|
| X | X X | 100 |
| | At least one action is Y | 0 |
| Y | At least one action is Y | 50 |



Ends

Suppose the public counter is 0 at the beginning of this round, and the group outcome (P1 P2) is X X. You later find out that the spinner for this round stops in the grey area. What's your payoff in this round?

A 100

B 50

C 0

D −100

✕ Incorrect.

The arrow stops in the grey area, so this round turns out to be **the last round of this match**. The group outcome is X X, so the counter increases from 0 to 1 at the end of this round. The counter is below 2 at the end of the last round, so everyone's payoff of the last round **is overridden to −100**, regardless of what is shown in the payoff table.

Next Question →

Figure C1.6.9: Question 5 ($n = 2$ treatment sessions)

1 2 3 4 5 6

Each correct answer earns you **$0.50**.

| Your action | Group outcome (P1 P2 P3 P4) | Your payoff |
|---|---|---|
| X | X X X X | 100 |
| | At least one action is Y | 0 |
| Y | At least one action is Y | 50 |



**Ends**

Suppose the public counter is 0 at the beginning of this round, and the group outcome (P1 P2 P3 P4) is X X X X. You later find out that the spinner for this round stops in the grey area. What's your payoff in this round?

A 100

B 50

C 0

D –100

✕ **Incorrect.**

The arrow stops in the grey area, so this round turns out to be **the last round of this match**. The group outcome is X X X X, so the counter increases from 0 to 1 at the end of this round. The counter is below 2 at the end of the last round, so everyone's payoff of the last round **is overridden to –100**, regardless of what is shown in the payoff table.

Next Question →

Figure C1.6.10: Question 5 ($n = 4$ treatment sessions)

1 2 3 4 5 6

Each correct answer earns you **$0.50**.

| Your action | Group outcome (P1 P2) | Your payoff |
|---|---|---|
| X | X X | 100 |
| | At least one action is Y | 0 |
| Y | At least one action is Y | 50 |



**Ends**

Suppose you are Player 1 (P1) and in this round, the group outcome (P1 P2) is X X. You later find out that in this round, the arrow stops in the grey area. What's your payoff in this round?

A 100

B 50

C 0

D We don't know for sure.

✕ Incorrect.

Your round payoff is determined by your action and the group outcome, as shown in the payoff table. Your action is X, and the group outcome is X X, so your round payoff is 100. The arrow stops in the grey area, so the match ends at the end of this round.

Next Question →

Figure C1.6.11: Question 5 ($n = 2$ baseline sessions)

1 2 3 4 5 6

Each correct answer earns you **$0.50**.

| Your action | Group outcome (P1 P2 P3 P4) | Your payoff |
|---|---|---|
| X | X X X X | 100 |
| | At least one action is Y | 0 |
| Y | At least one action is Y | 50 |



**Ends**

**Suppose you are Player 1 (P1) and in this round, the group outcome (P1 P2 P3 P4) is X X X X. You later find out that in this round, the arrow stops in the grey area. What's your payoff in this round?**

A 100

B 50

C 0

D We don't know for sure.

**✕ Incorrect.**

Your round payoff is determined by your action and the group outcome, as shown in the payoff table. Your action is X, and the group outcome is X X X X, so your round payoff is 100. The arrow stops in the grey area, so the match ends at the end of this round.

Next Question →

Figure C1.6.12: Question 5 ($n = 4$ baseline sessions)

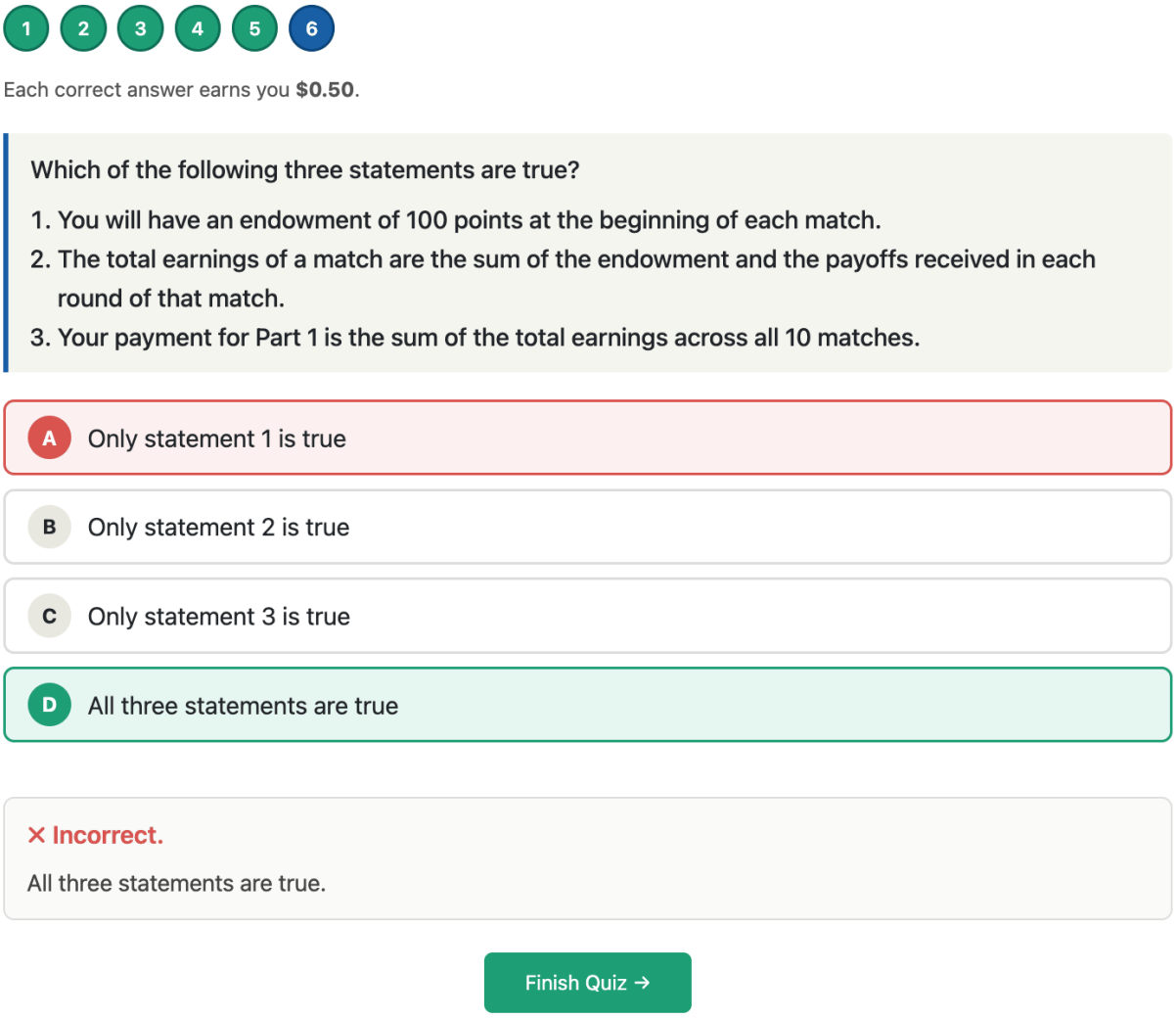



Figure C1.6.13: Question 6

## C1.7 Part 1: 10 Matches ($n = 2$ sessions)

*To reduce repetition, all screenshots in this part show the Part 1 interfaces in page order, for the $n = 2$ sessions only. Explanations are provided in figure titles or between figures.*

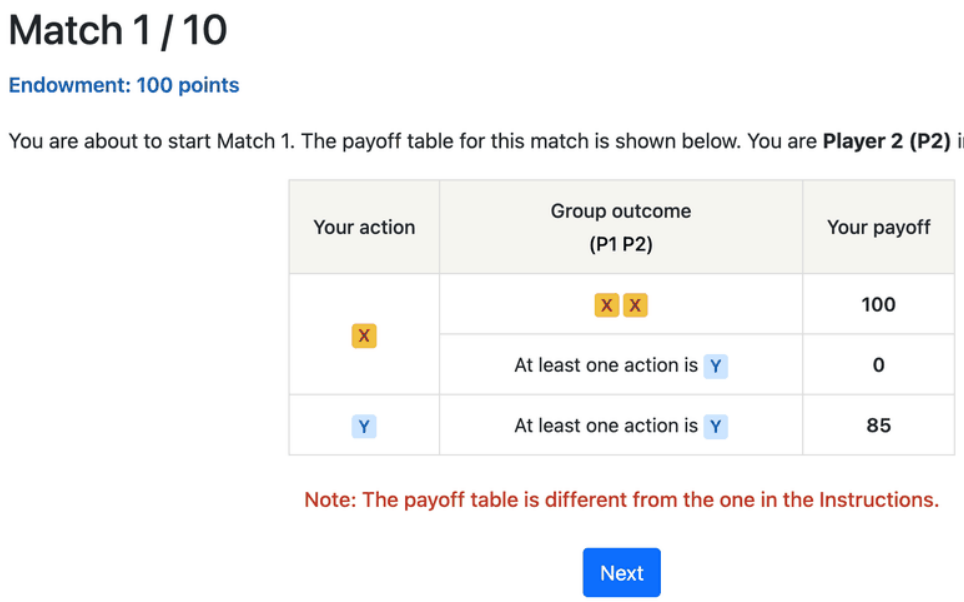



Figure C1.7.1: Welcome page for Matches 1-5 ($n = 2$ sessions)

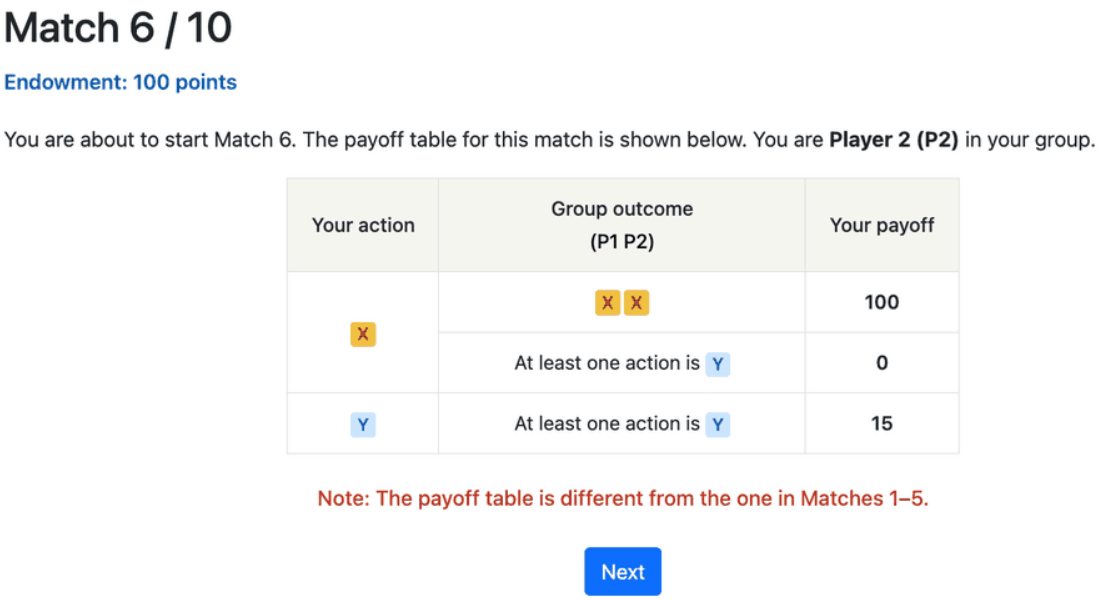



Figure C1.7.2: Welcome page for Matches 6-10 ($n = 2$ sessions)

*The order of the welcome pages shown here ($u_Y = 85$ for Matches 1-5, $u_Y = 15$ for Matches 6-10) is for the $2h{+}2l$ treatment and baseline sessions. For the $2l{+}2h$ treatment and baseline*

*sessions, the two payoff tables will reverse the order in display ($u_Y = 15$ for Matches 1-5, $u_Y = 85$ for Matches 6-10).*

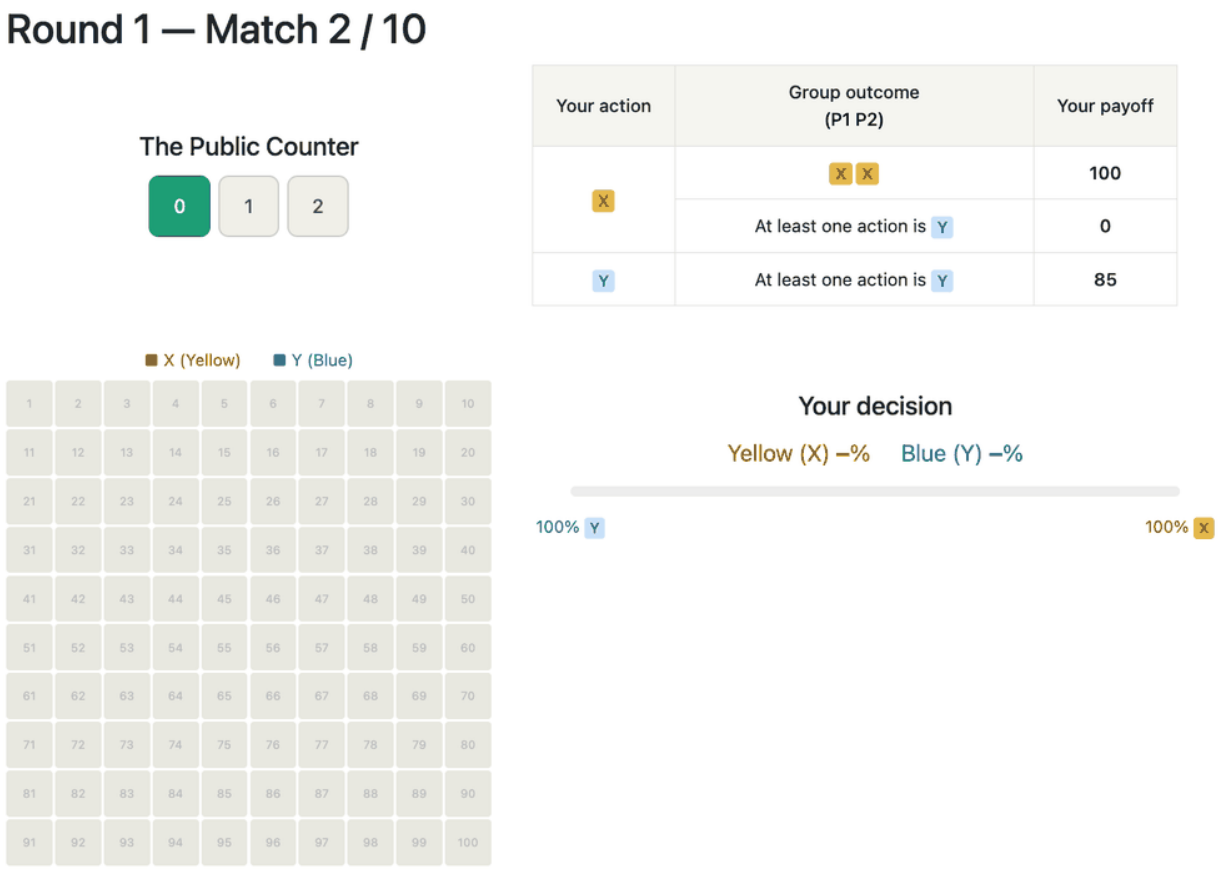



Figure C1.7.3: Decision page when Counter=0 ($n = 2$ treatment sessions)

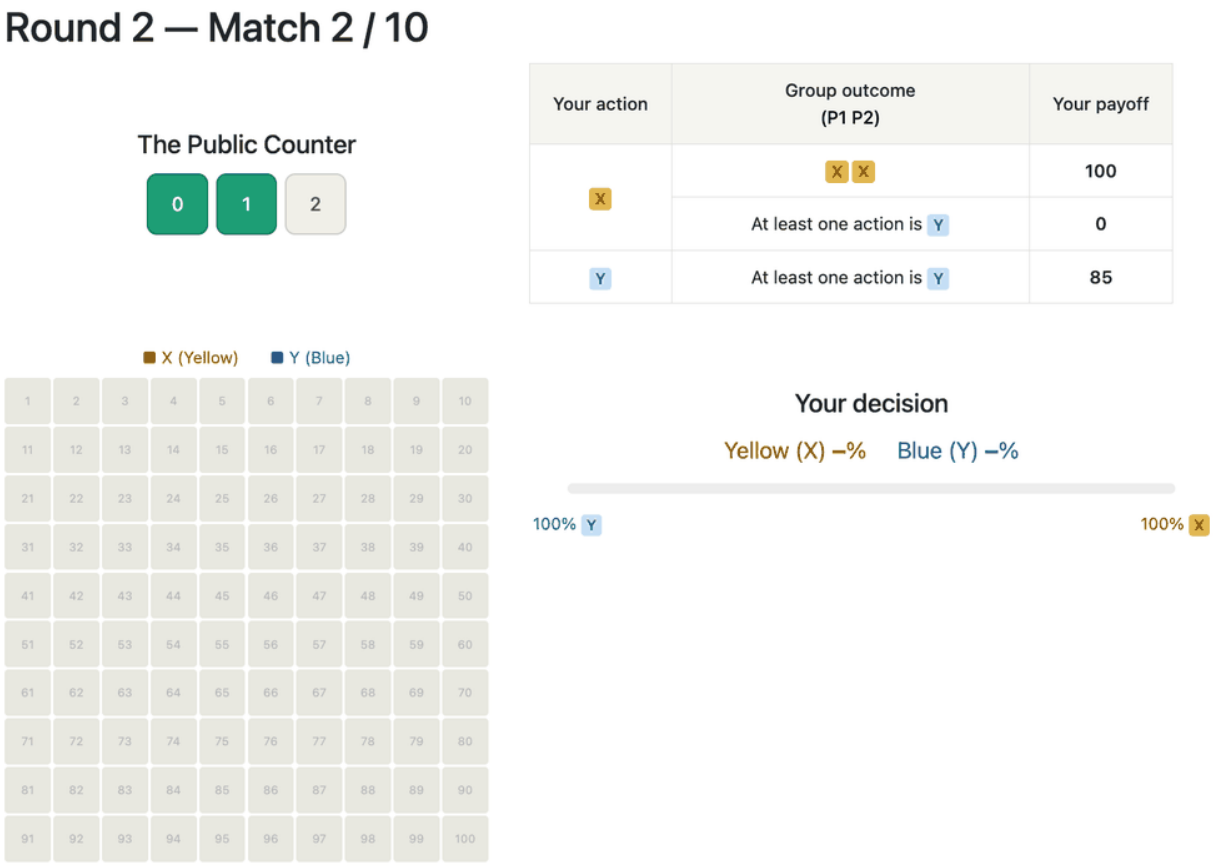



Figure C1.7.4: Decision page when Counter=1 ($n = 2$ treatment sessions)

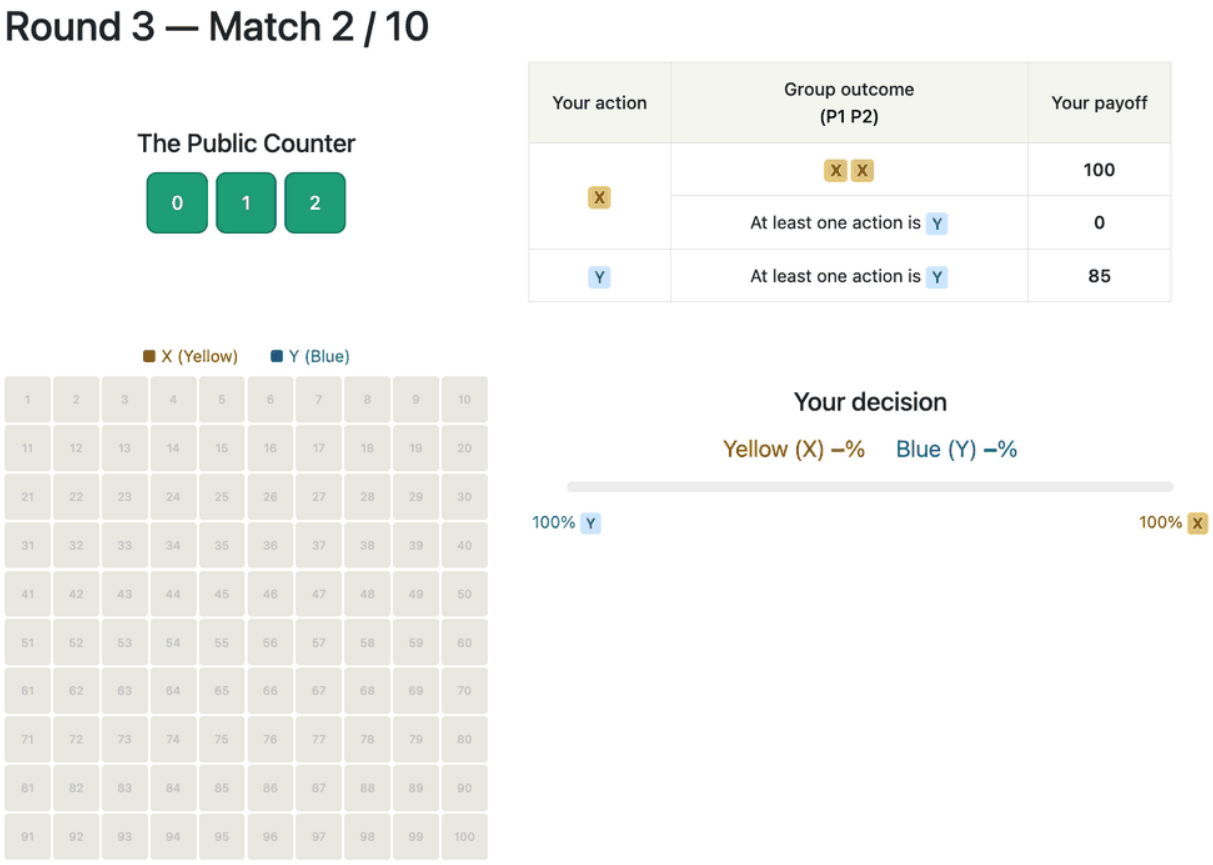



Figure C1.7.5: Decision page when Counter=2 ($n = 2$ treatment sessions)

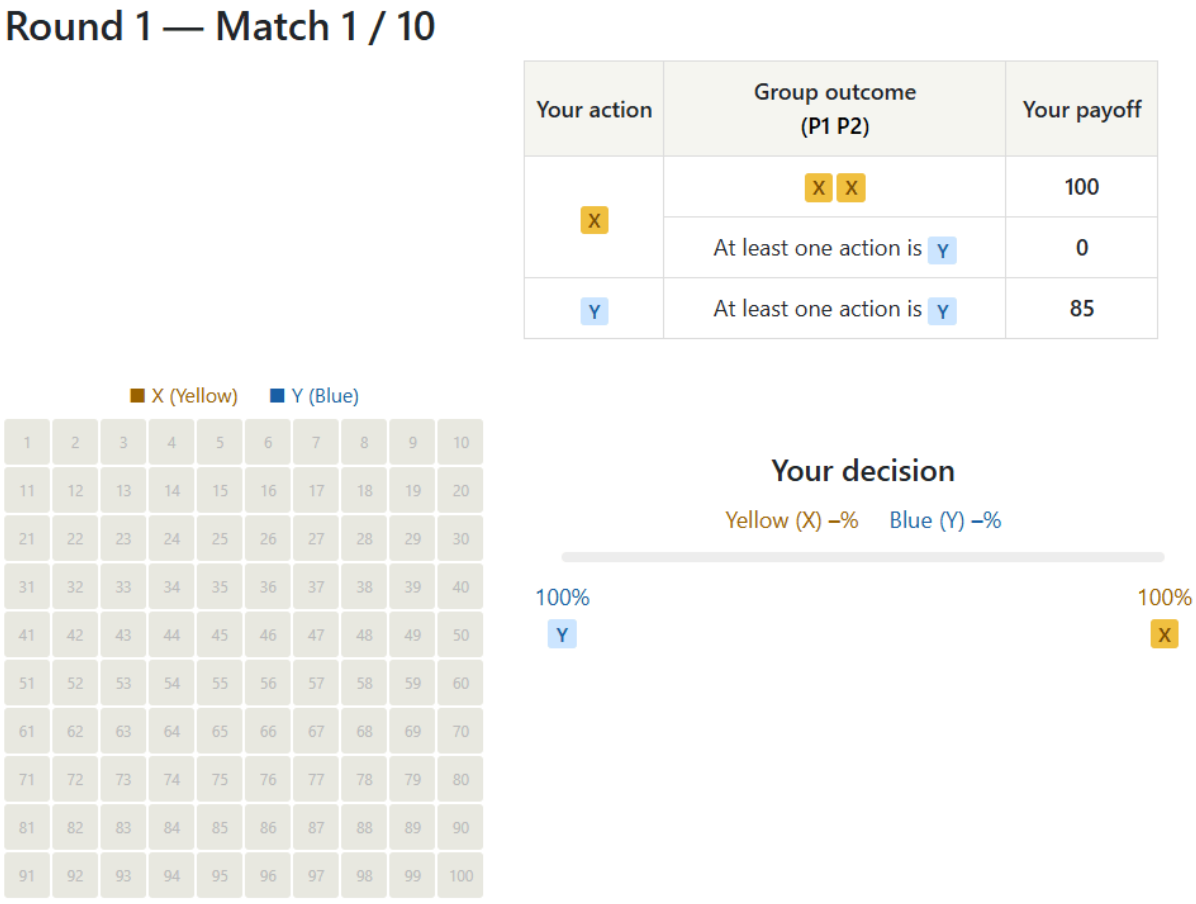



Figure C1.7.6: Decision page ($n = 2$ baseline sessions)

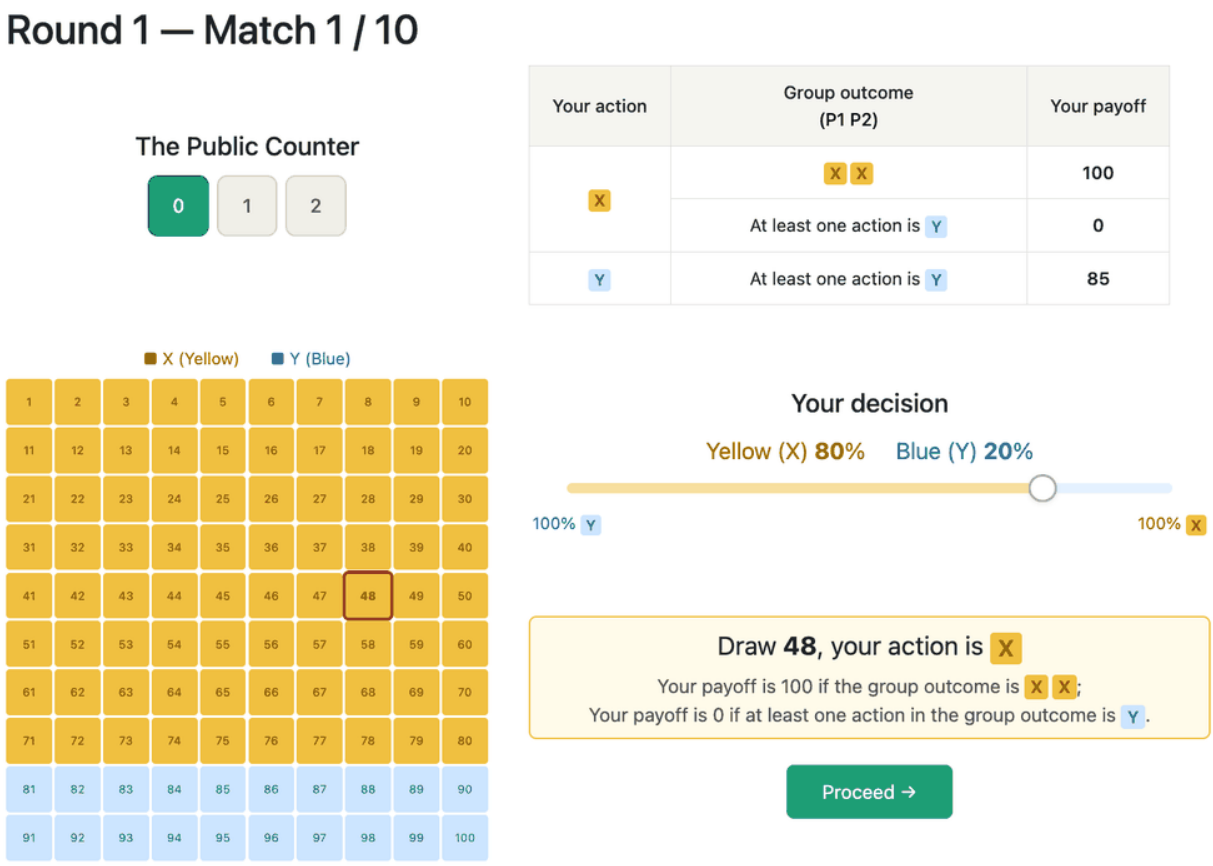



Figure C1.7.7: RevealAction page when Action X is realized ($n = 2$ treatment sessions)

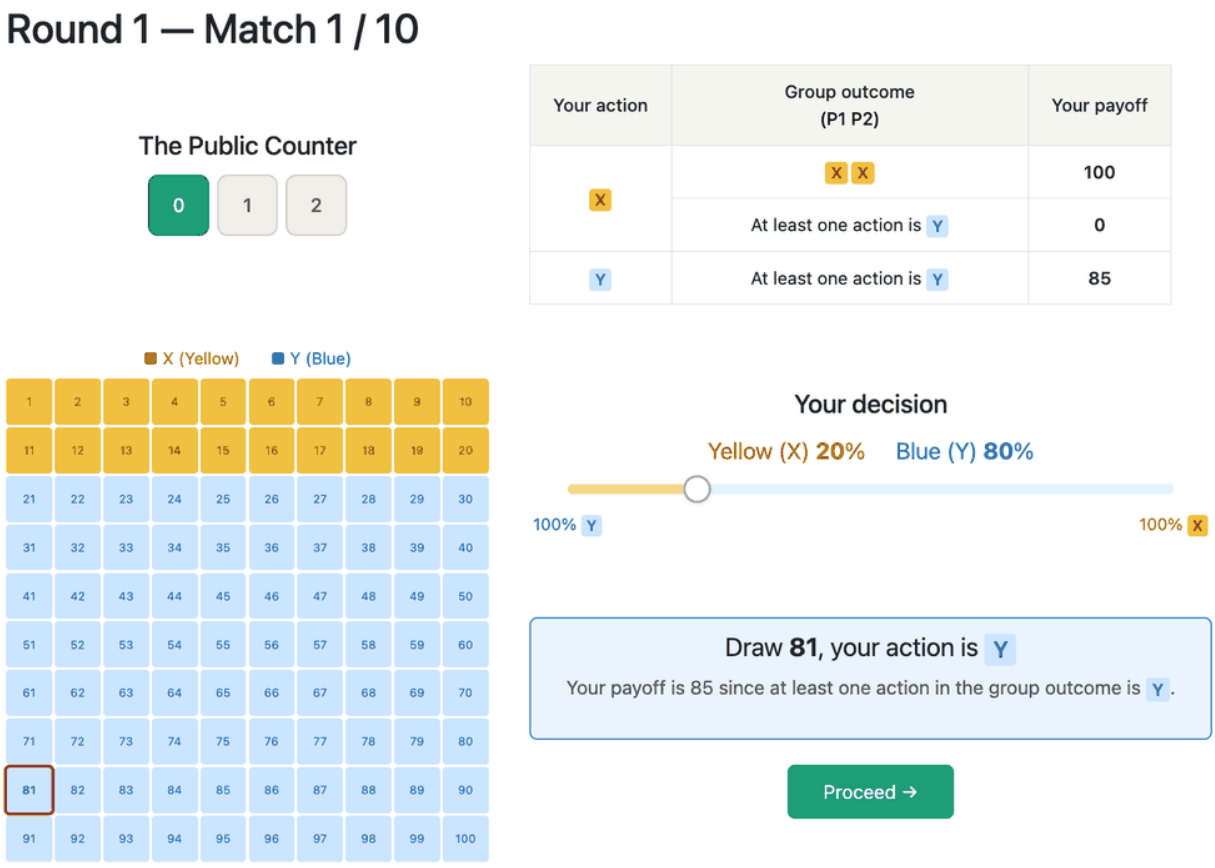



Figure C1.7.8: RevealAction page when Action Y is realized ($n = 2$ treatment sessions)

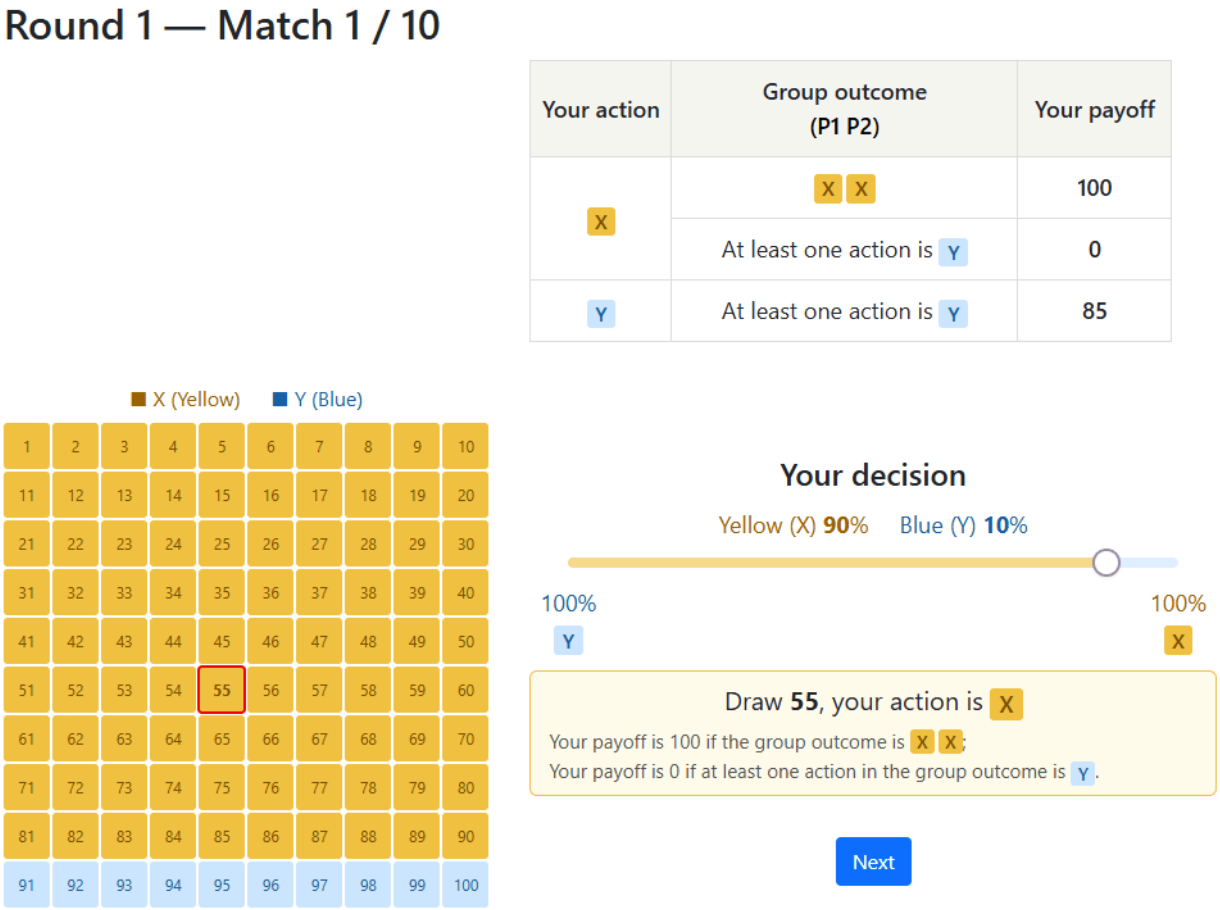



Figure C1.7.9: RevealAction page when Action X is realized ($n = 2$ baseline sessions)

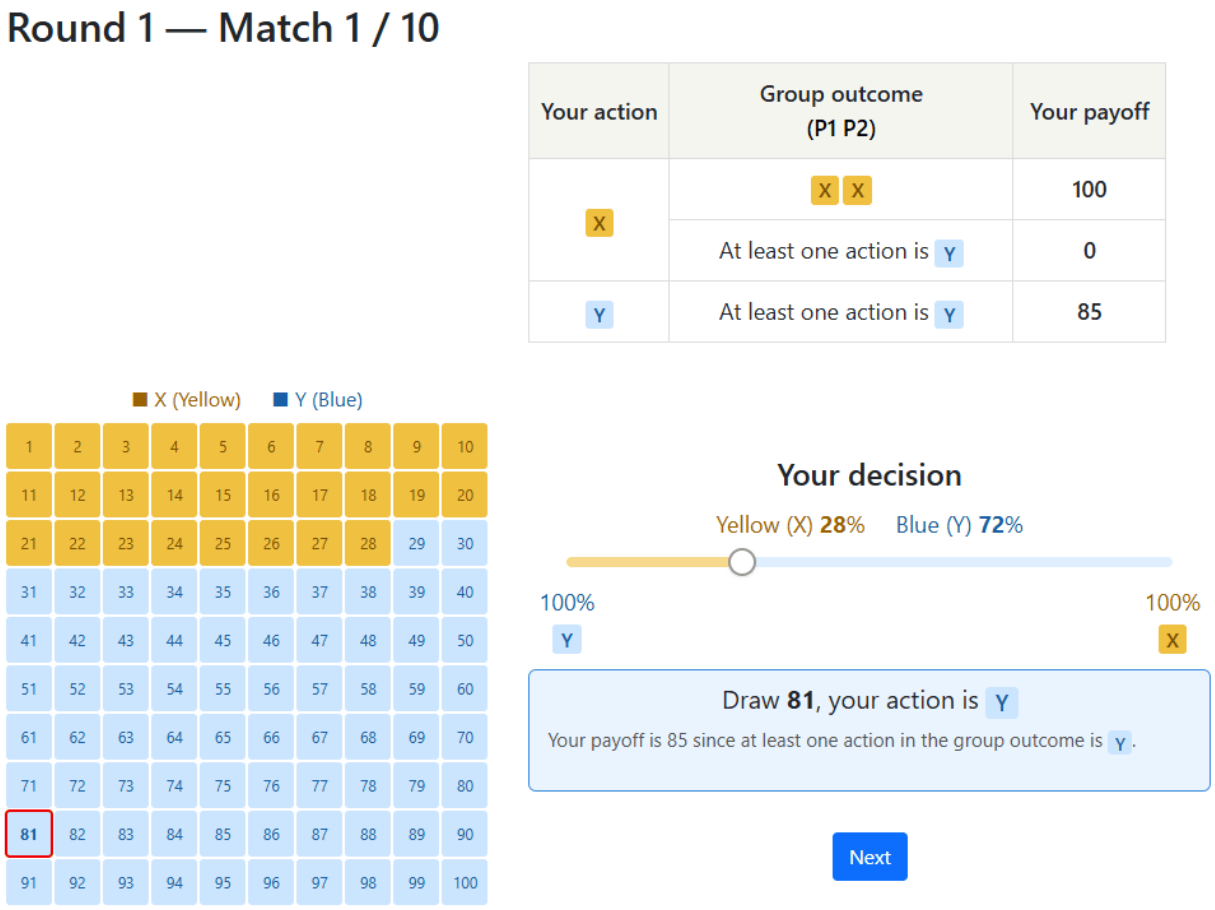



Figure C1.7.10: RevealAction page when Action Y is realized ($n = 2$ baseline sessions)

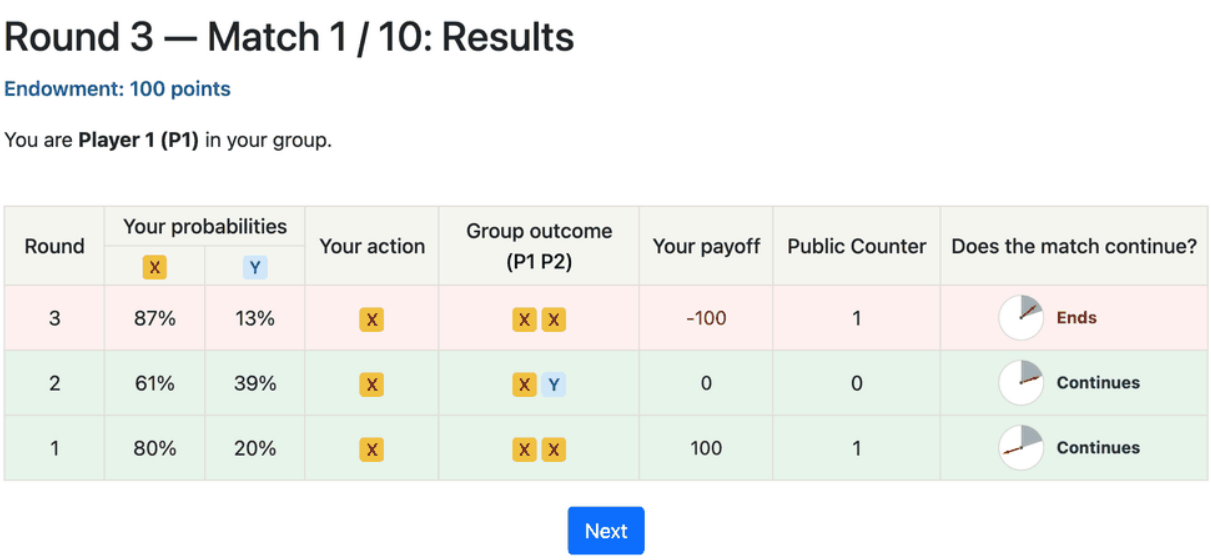



Figure C1.7.11: Result page with a penalty in the last round ($n = 2$ treatment sessions)

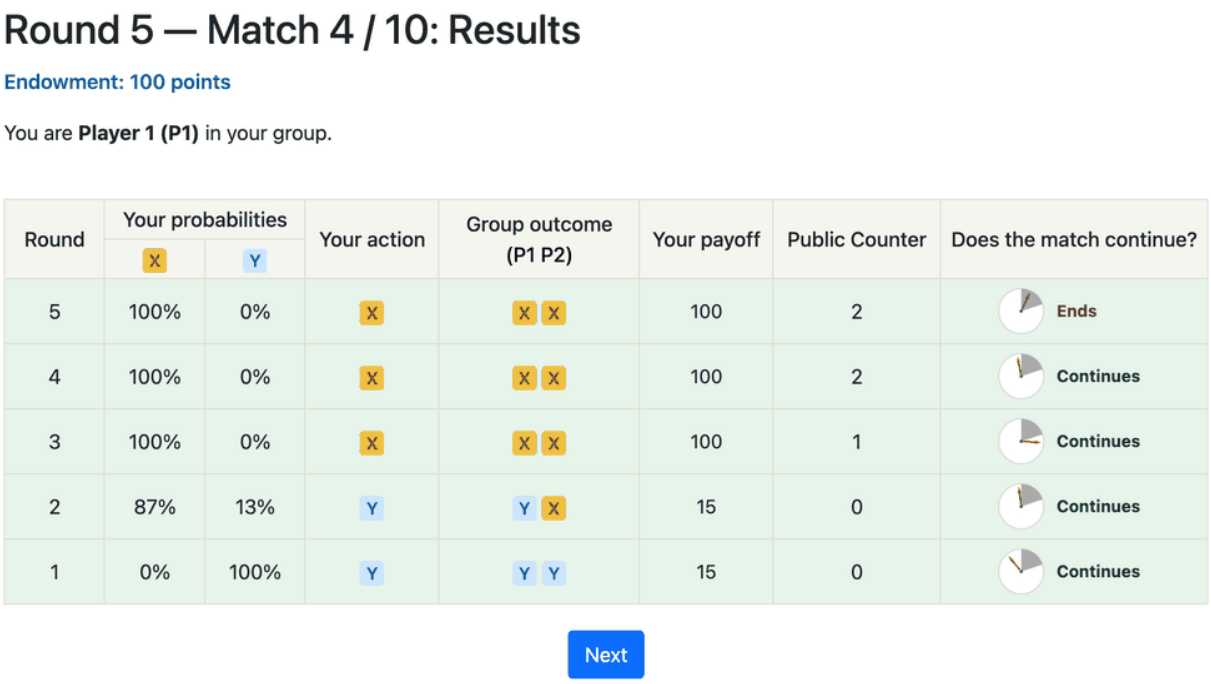



Figure C1.7.12: Result page with no penalty in the last round ($n = 2$ treatment sessions)

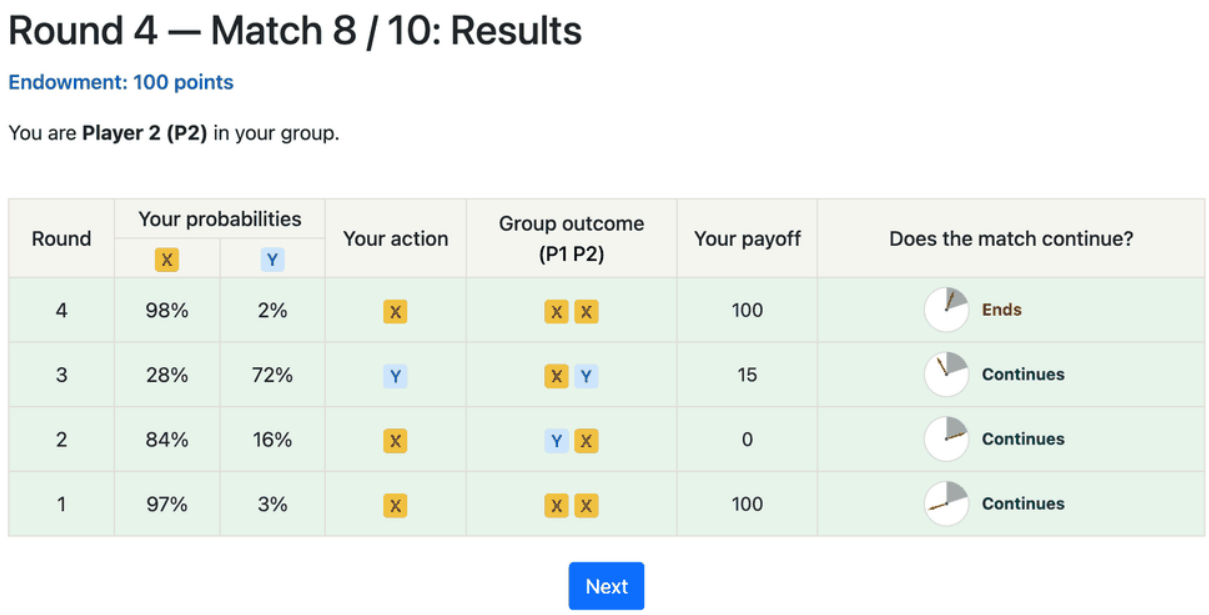



Figure C1.7.13: Result page ($n = 2$ baseline sessions)

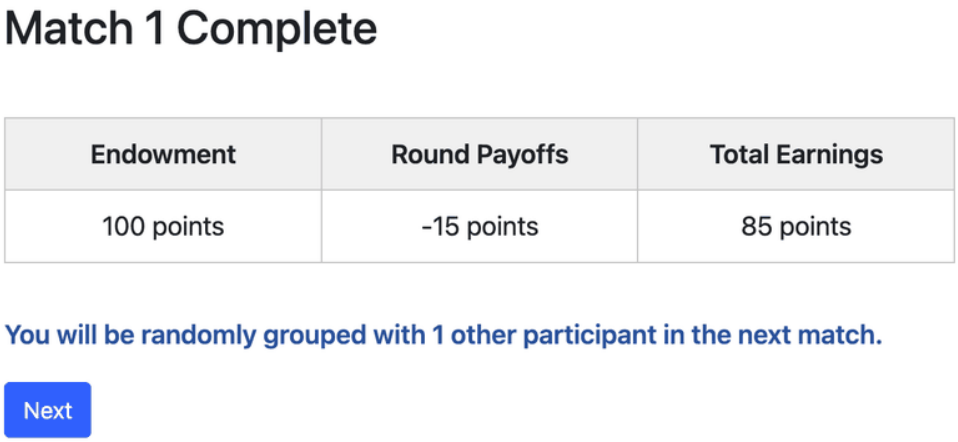



Figure C1.7.14: Match end page

## C1.8 Part 1: 10 Matches ($n = 4$ sessions)

*To reduce repetition, all screenshots in this part show the Part 1 interfaces in page order, for the $n = 4$ sessions only. Explanations are provided in figure titles or between figures.*

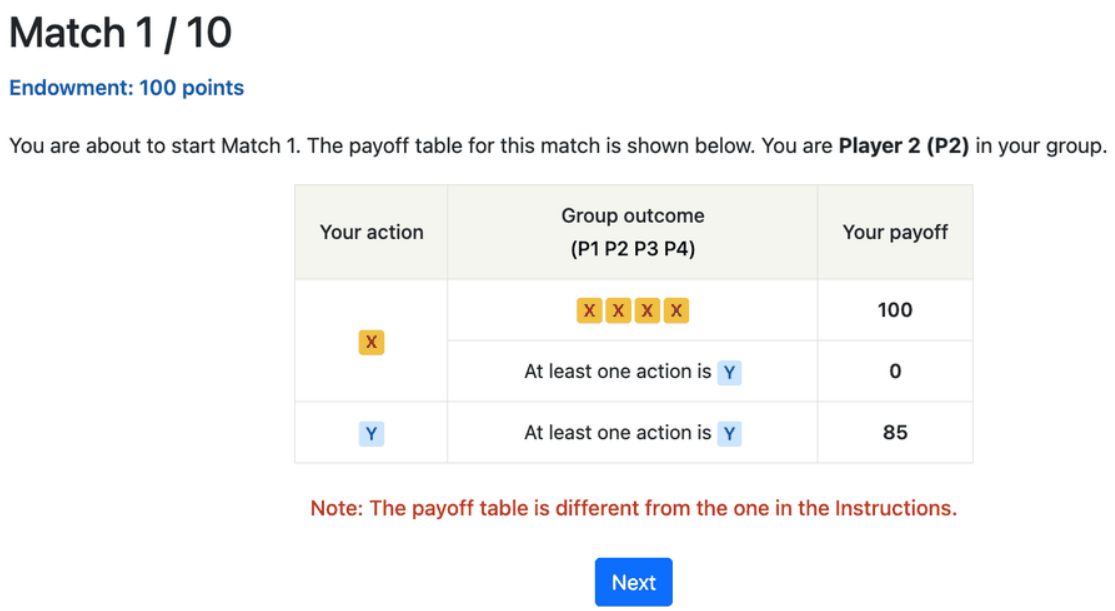



Figure C1.8.1: Welcome page for Matches 1-5 ($n = 4$ sessions)

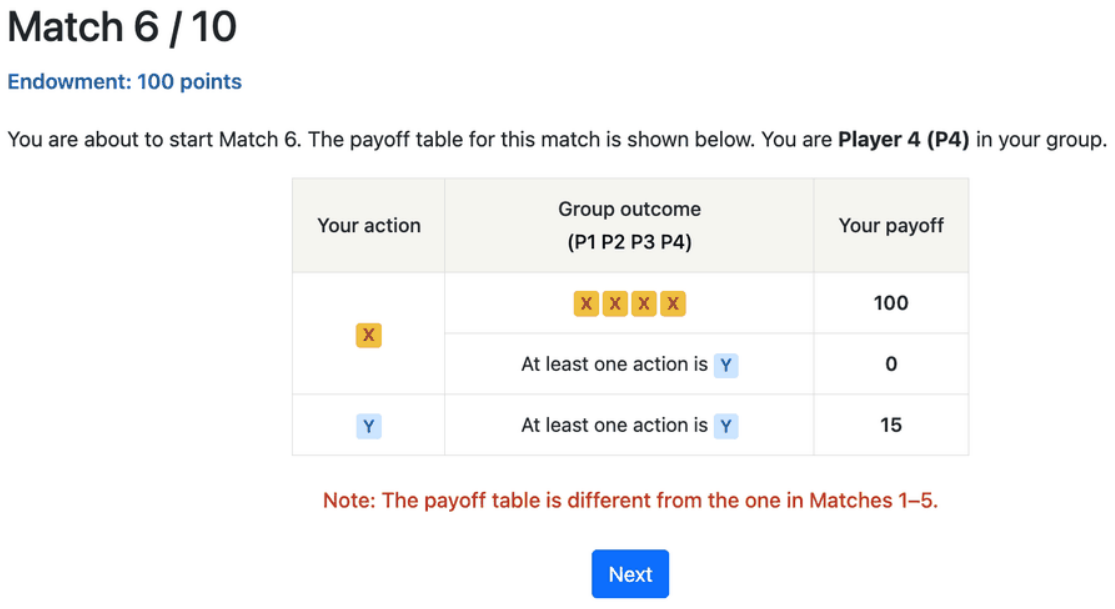



Figure C1.8.2: Welcome page for Matches 6-10 ($n = 4$ sessions)

*The order of the welcome pages shown here ($u_Y = 85$ for Matches 1-5, $u_Y = 15$ for Matches*

*6-10) is for the 4h+4l treatment and baseline sessions. For the 4l+4h treatment and baseline sessions, the two payoff tables will reverse the order in display ($u_Y = 15$ for Matches 1-5, $u_Y = 85$ for Matches 6-10).*

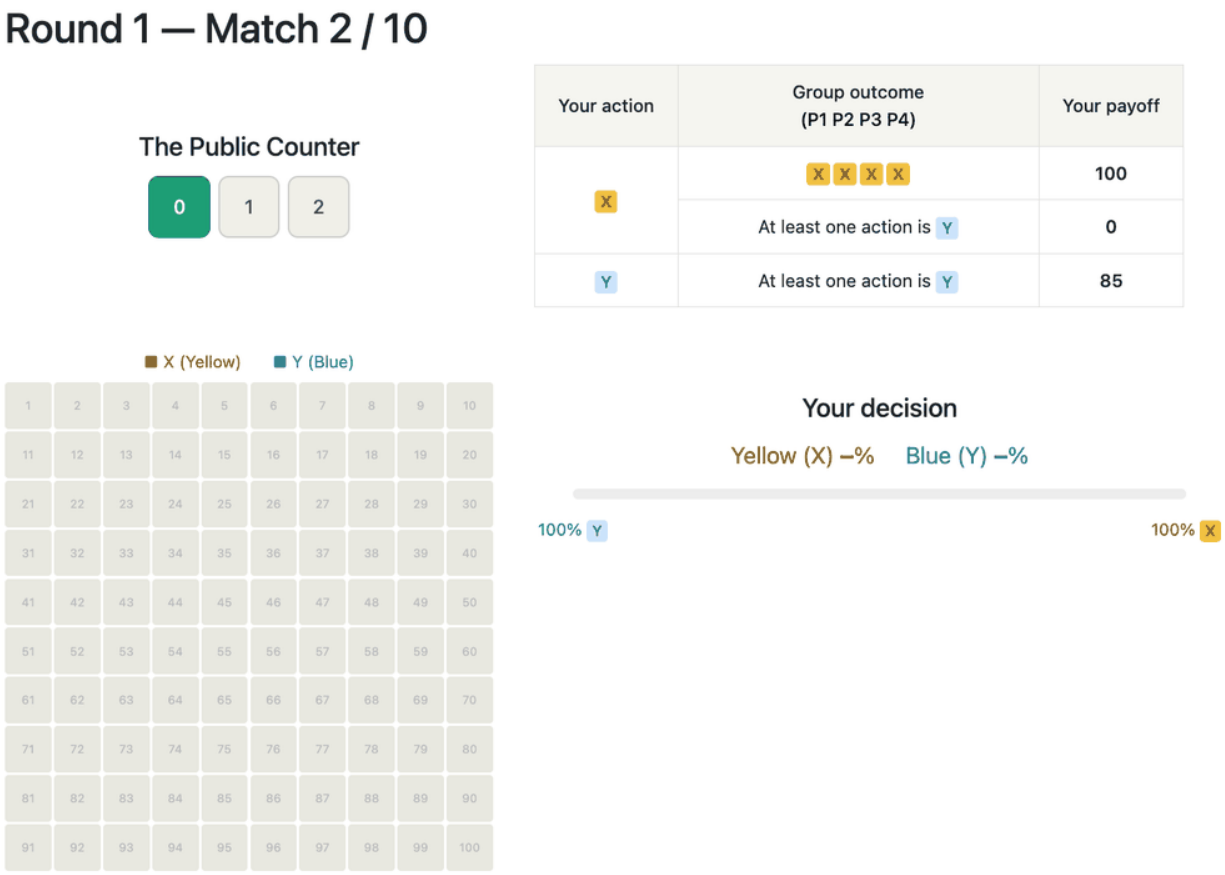



Figure C1.8.3: Decision page when Counter=0 ($n = 4$ treatment sessions)

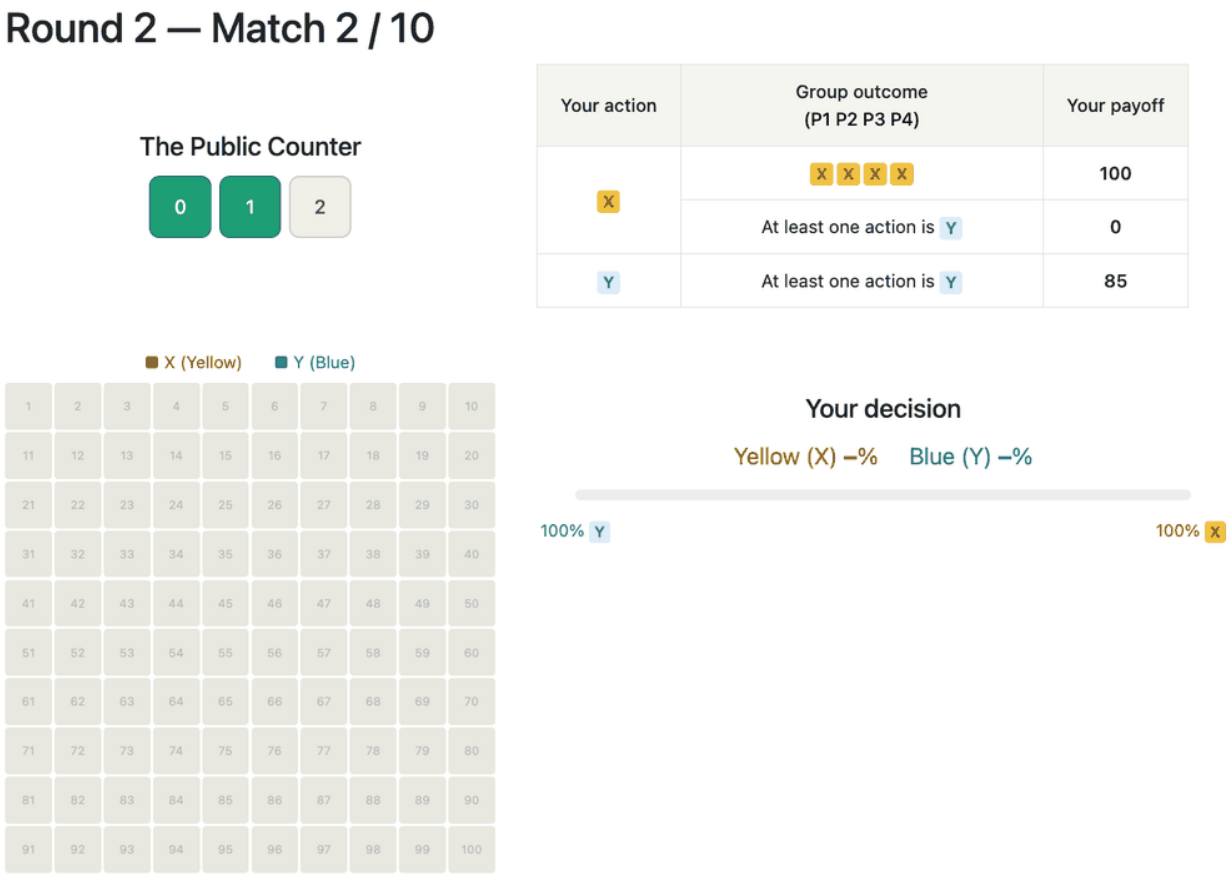



Figure C1.8.4: Decision page when Counter=1 ($n = 4$ treatment sessions)

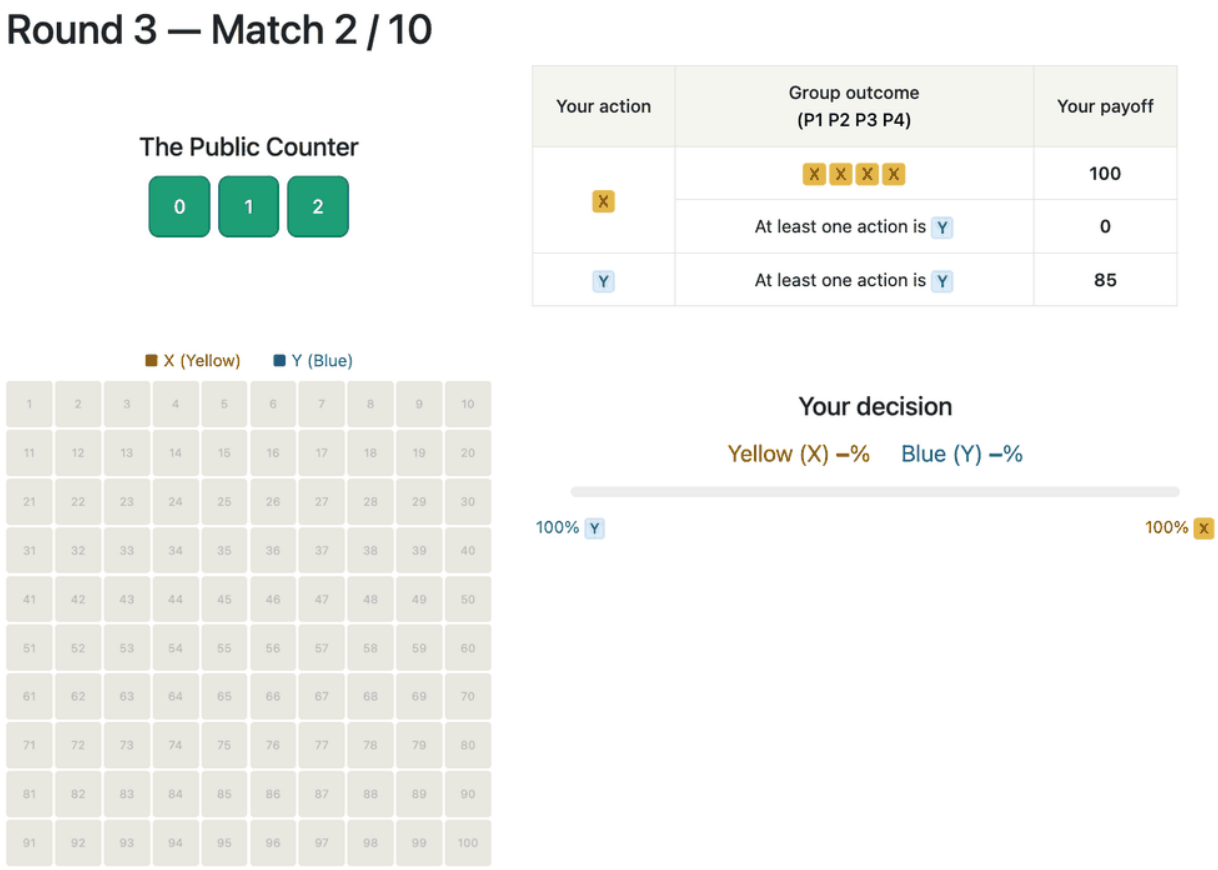



Figure C1.8.5: Decision page when Counter=2 ($n = 4$ treatment sessions)

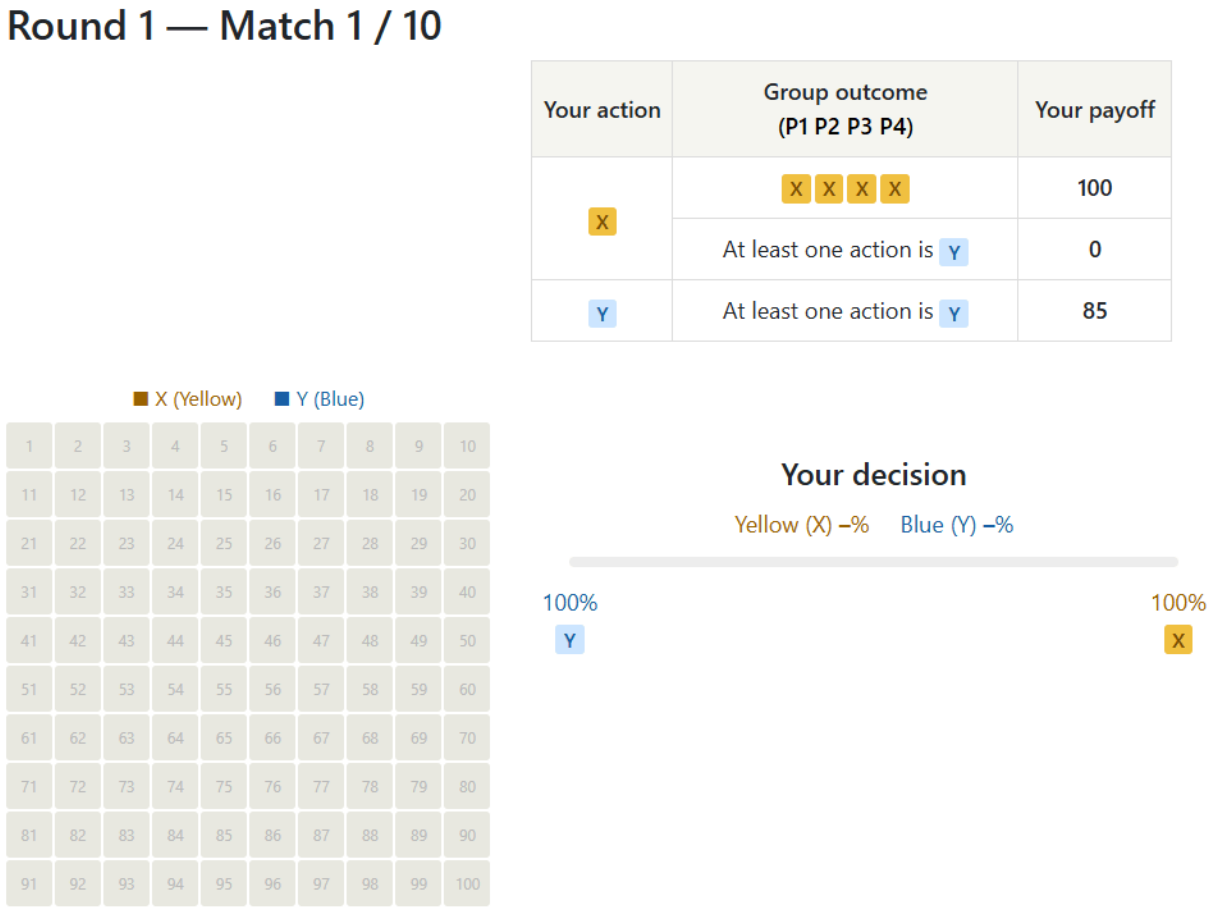



Figure C1.8.6: Decision page ($n = 4$ baseline sessions)

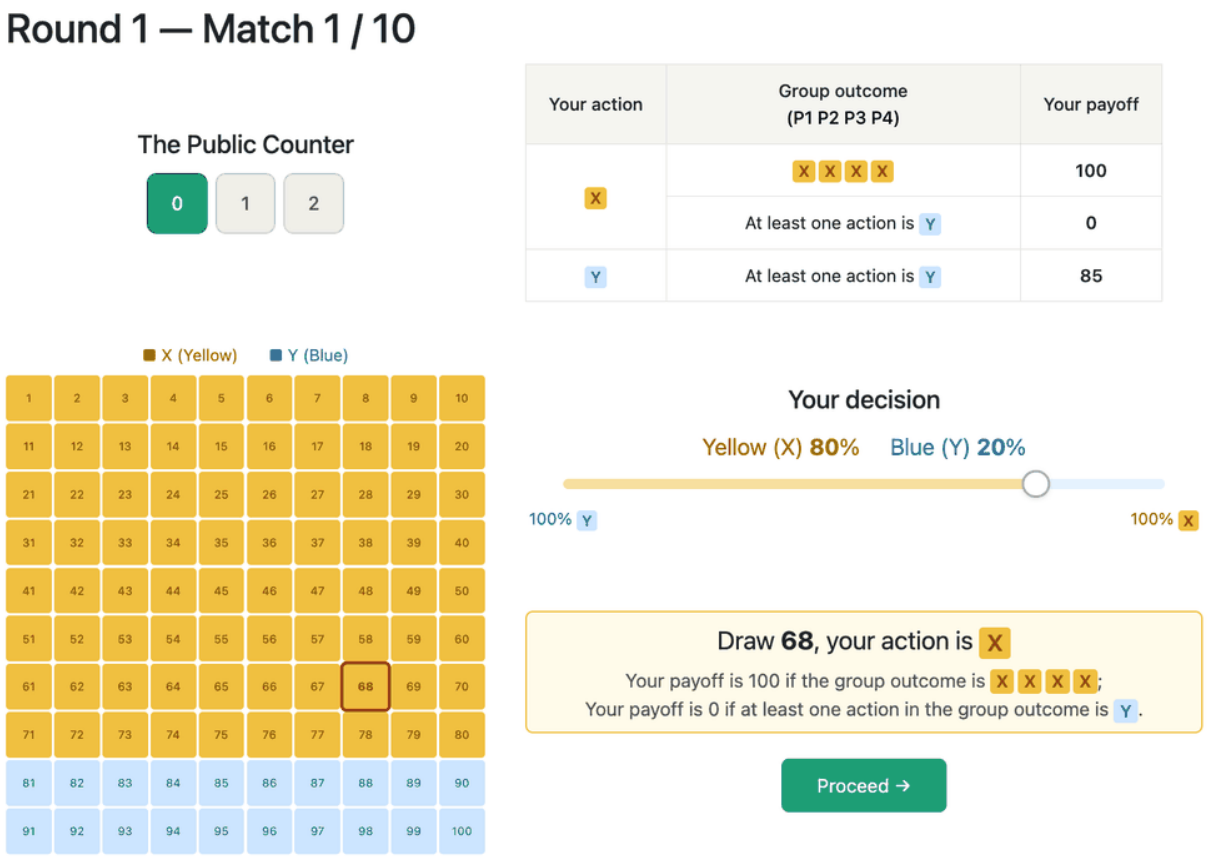



Figure C1.8.7: RevealAction page when Action X is realized ($n = 4$ treatment sessions)

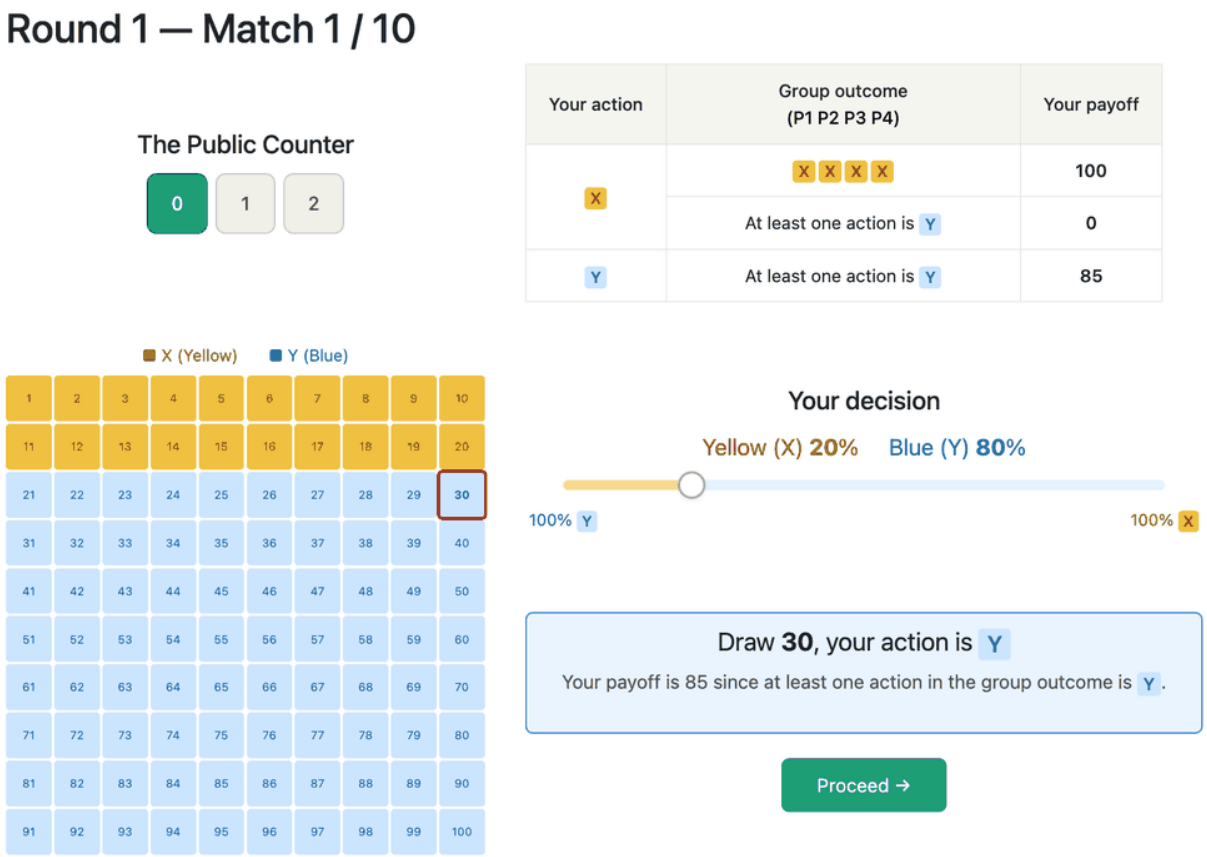



Figure C1.8.8: RevealAction page when Action Y is realized ($n = 4$ treatment sessions)

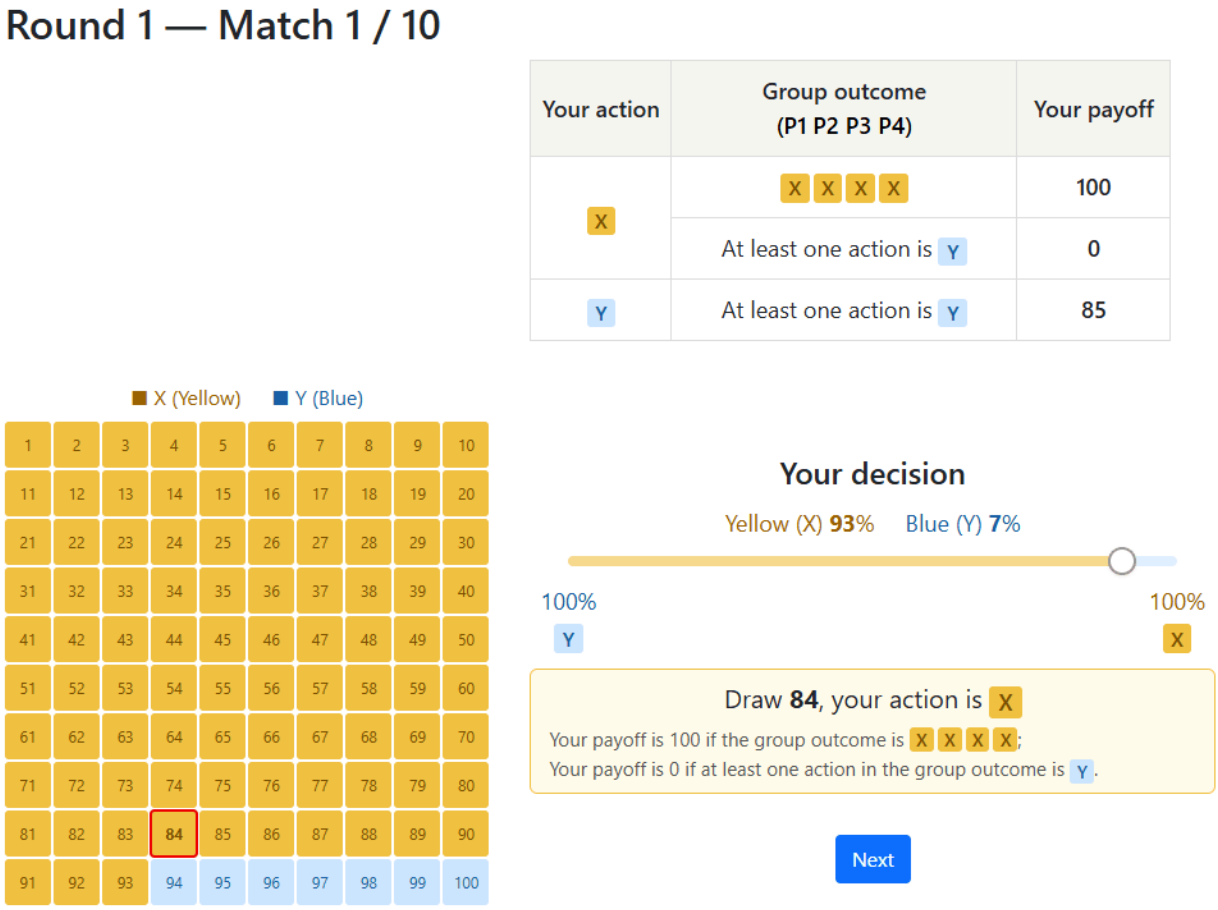



Figure C1.8.9: RevealAction page when Action X is realized ($n = 4$ baseline sessions)

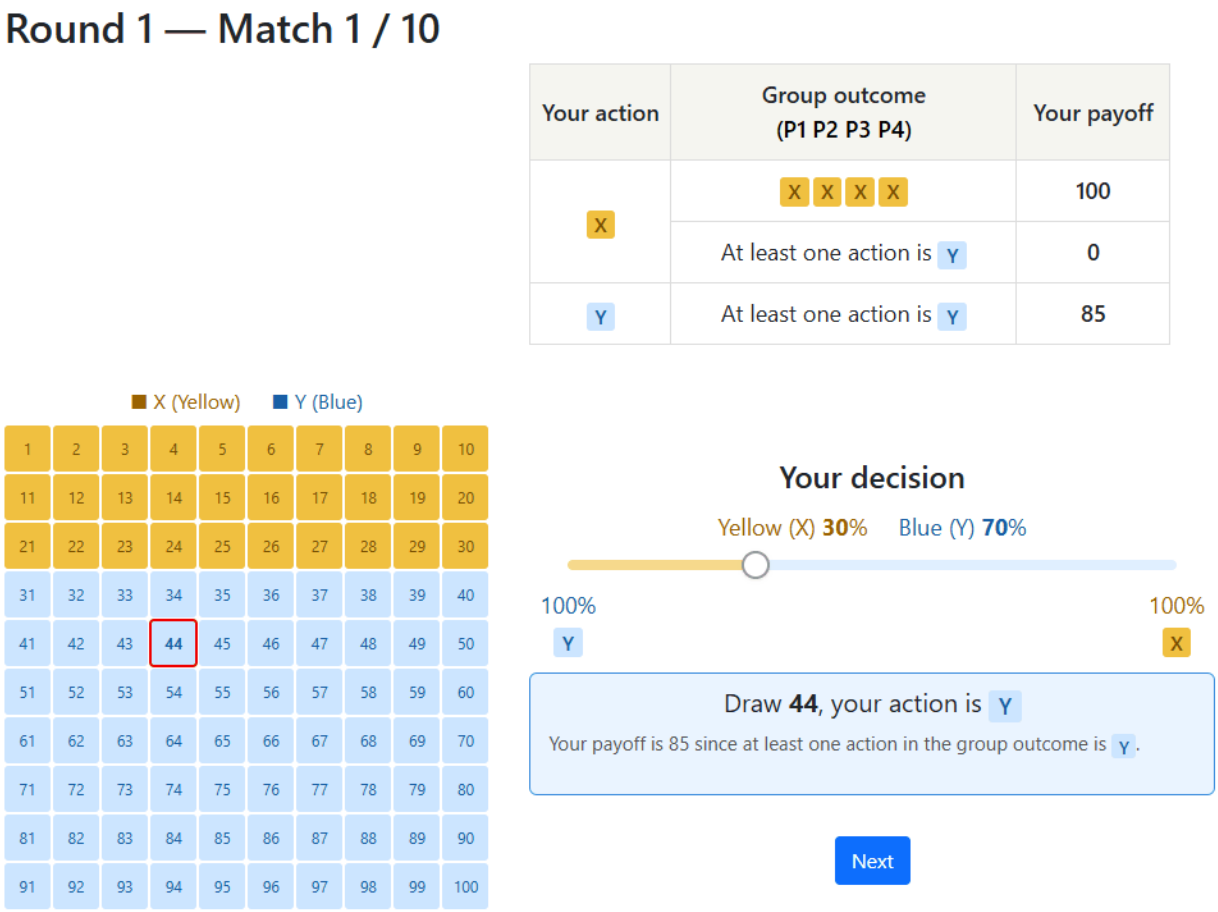



Figure C1.8.10: RevealAction page when Action Y is realized ($n = 4$ baseline sessions)

## Round 3 — Match 1 / 10: Results

**Endowment: 100 points**

You are **Player 2 (P2)** in your group.

| Round | Your probabilities | | Your action | Group outcome (P1 P2 P3 P4) | Your payoff | Public Counter | Does the match continue? |
|---|---|---|---|---|---|---|---|
| | X | Y | | | | | |
| 3 | 100% | 0% | X | X X X Y | -100 | 0 | Ends |
| 2 | 100% | 0% | X | X X X X | 100 | 1 | Continues |
| 1 | 88% | 12% | X | X X Y Y | 0 | 0 | Continues |

Next

Figure C1.8.11: Result page with a penalty in the last round ($n = 4$ treatment sessions)

## Round 5 — Match 2 / 10: Results

**Endowment: 100 points**

You are **Player 3 (P3)** in your group.

| Round | Your probabilities | | Your action | Group outcome (P1 P2 P3 P4) | Your payoff | Public Counter | Does the match continue? |
|---|---|---|---|---|---|---|---|
| | X | Y | | | | | |
| 5 | 100% | 0% | X | X X X X | 100 | 2 | Ends |
| 4 | 100% | 0% | X | X X X X | 100 | 2 | Continues |
| 3 | 100% | 0% | X | X X X X | 100 | 1 | Continues |
| 2 | 100% | 0% | X | X Y X X | 0 | 0 | Continues |
| 1 | 100% | 0% | X | X X X Y | 0 | 0 | Continues |

Next

Figure C1.8.12: Result page with no penalty in the last round ($n = 4$ treatment sessions)

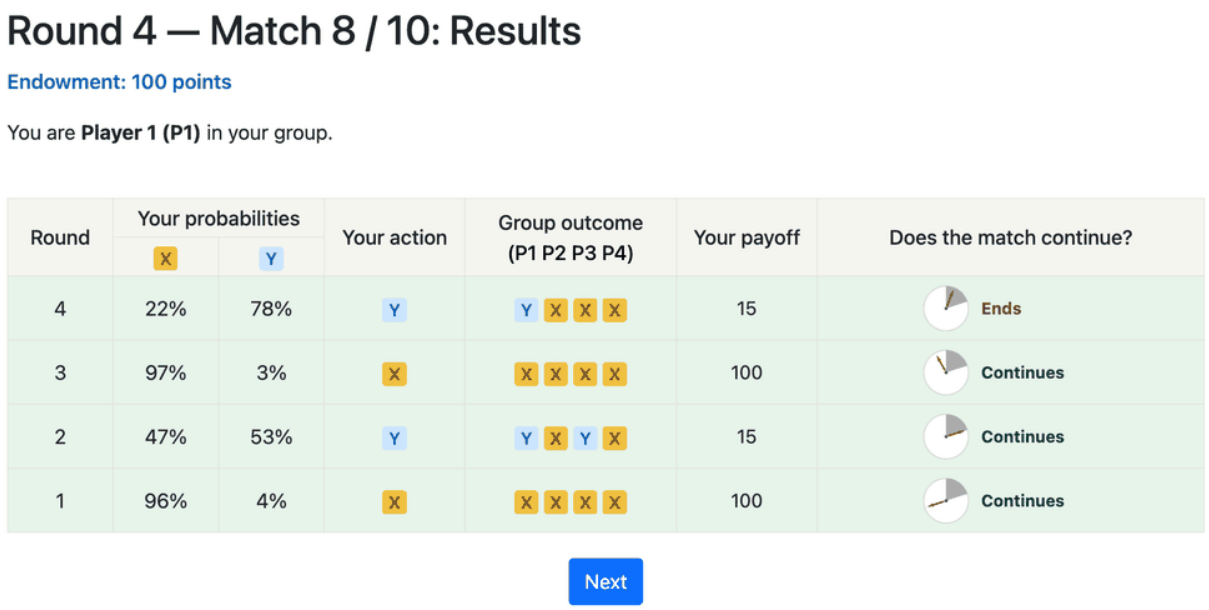



Figure C1.8.13: Result page ($n = 4$ baseline sessions)

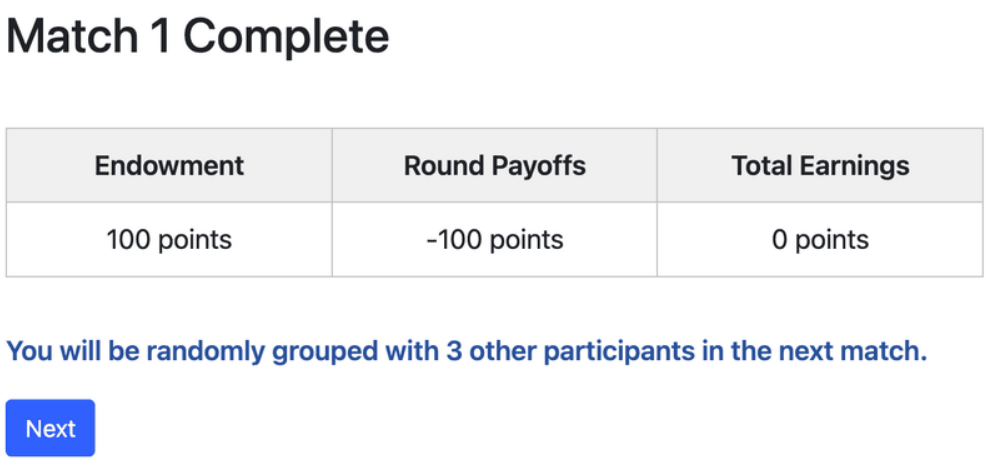



Figure C1.8.14: Match end page

## C1.9 Part 2: Risk Preferences

*Part 2 is identical across all sessions; therefore, only the interfaces are shown, without additional figure titles.*

# Part 2: Instructions

In Part 2, you will be asked to make a series of decisions. Each decision is a choice between two options -- **Option A** and **Option B**. The decisions will be organized into a table like the one below. Notice that for practice there are a total of 5 decisions (one for each rows of the table). You should think of each row as a separate decision you need to make. At this time, please make your choice for each decision row. Note that your compensation will not depend on the practice task.

| Decision | Option A | | Option B | Your Choice |
|---|---|---|---|---|
| #1 | **$5.00** with 50% chance; **$0.00** with 50% chance | ○ ○ | **$1.00** | __ |
| #2 | **$5.00** with 50% chance; **$0.00** with 50% chance | ○ ○ | **$2.00** | __ |
| #3 | **$5.00** with 50% chance; **$0.00** with 50% chance | ○ ○ | **$3.00** | __ |
| #4 | **$5.00** with 50% chance; **$0.00** with 50% chance | ○ ○ | **$4.00** | __ |
| #5 | **$5.00** with 50% chance; **$0.00** with 50% chance | ○ ○ | **$5.00** | __ |

# Part 2: Instructions

For Part 2, you will make 10 decisions. At the end of the experiment, one of these decisions will be chosen at random to be carried out. As an example, we will demonstrate how compensation works with the practice decisions that you just did.

The procedure involves two steps:

- First, one of the decisions will be chosen at random. For example, in the practice task you had 5 decisions, so one will be chosen randomly by the computer (with equal probability).
- Second, depending on the chosen decision in step one, a random draw might be necessary. For example, if you chose Option A over Option B, then a random draw will be made to determine the outcome of the technology. For example, in the practice task, the technology in Option A had 50% of $5.00 and 50% of $0.00.

To demonstrate the compensation procedure, we will perform the two steps using the practice task.

| Decision | Option A | Option B | Your Choice |
|---|---|---|---|
| #1 | **$5.00** with 50% chance; **$0.00** with 50% chance | $1.00 | A |
| #2 | **$5.00** with 50% chance; **$0.00** with 50% chance | $2.00 | A |
| #3 | $5.00 with 50% chance; $0.00 with 50% chance | $3.00 | B |
| #4 | $5.00 with 50% chance; $0.00 with 50% chance | $4.00 | B |
| #5 | $5.00 with 50% chance; $0.00 with 50% chance | $5.00 | B |

1. The randomly chosen decision is Decision #2

2. The random realization of Option A was $0.00.

**Therefore, your payoff is $0.00.**

To gain further understanding of the compensation procedure, you may try again.

Try Again

In the actual experiment, this procedure will be carried out **only once**. If you are ready to begin, please click **Next**.

Next

# Part 2: Decisions

Please make a choice for **each of the 10 decisions**.

| Decision | Option A | | Option B | Your Choice |
|---|---|---|---|---|
| #1 | **$5.00** with 50% chance; **$0.00** with 50% chance | ○ ○ | **$0.50** | __ |
| #2 | **$5.00** with 50% chance; **$0.00** with 50% chance | ○ ○ | **$1.00** | __ |
| #3 | **$5.00** with 50% chance; **$0.00** with 50% chance | ○ ○ | **$1.50** | __ |
| #4 | **$5.00** with 50% chance; **$0.00** with 50% chance | ○ ○ | **$2.00** | __ |
| #5 | **$5.00** with 50% chance; **$0.00** with 50% chance | ○ ○ | **$2.50** | __ |
| #6 | **$5.00** with 50% chance; **$0.00** with 50% chance | ○ ○ | **$3.00** | __ |
| #7 | **$5.00** with 50% chance; **$0.00** with 50% chance | ○ ○ | **$3.50** | __ |
| #8 | **$5.00** with 50% chance; **$0.00** with 50% chance | ○ ○ | **$4.00** | __ |
| #9 | **$5.00** with 50% chance; **$0.00** with 50% chance | ○ ○ | **$4.50** | __ |
| #10 | **$5.00** with 50% chance; **$0.00** with 50% chance | ○ ○ | **$5.00** | __ |

## C1.10 Part 3: Social Value Orientations

*Part 3 is identical across all sessions; therefore, only the interfaces are shown, without additional figure titles.*

### Part 3 : Instructions

In this part of the study, you will be randomly paired with another person, whom we will refer to as the **other**. You will not know who the other person is, nor will the other person be informed about your identity. You will make a series of choices among several alternative allocations of points. These points will be converted into Dollars at a rate of **50 points = $1.00**. Note that this rate is different from the one in Part 1.

You will be making a series of decisions about allocating points between you and this other person. For each of the questions, please indicate the distribution you prefer most by selecting the corresponding button in the middle row. You can only make one choice for each question. There are no right or wrong answers, this is all about personal preference.

---

In the example below, a person chose the allocation giving 50 points to herself, and 40 points to the unknown other person. In terms of US dollars, this yields an allocation of 50 / 50 = $1.00 for the person making the choice and 40 / 50 = $0.80 for the unknown other.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| You Receive | 30 | 35 | 40 | 45 | 50 | 55 | 60 | 65 | 70 |
| | ○ | ○ | ○ | ○ | ◉ | ○ | ○ | ○ | ○ |
| Other Receives | 80 | 70 | 60 | 50 | 40 | 30 | 20 | 10 | 0 |
| | | | | | | | | | Submit |

---

As you can see, your choices influence both the number of points you receive, as well as the number of points the other person receives.

After you have made all your choices, one of the allocation choices will be randomly selected by the computer. **For this choice, the computer will randomly assign one person from your group (you or the other) the role of "Receiver" and the other the role of the "Sender". The allocation choice made by the Sender will be enforced.** This allocation will be paid to both the Sender and the Receiver.

If you have any questions, please raise your hand.

Next

### Part 3 : Decision 1 of 6

Choose a distribution between yourself and the other individual

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| You Receive | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 | 85 |
| | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Other Receives | 85 | 76 | 68 | 59 | 50 | 41 | 33 | 24 | 15 |
| | | | | | | | | | Submit |

## Part 3 : Decision 2 of 6

Choose a distribution between yourself and the other individual

| You Receive | 85 | 87 | 89 | 91 | 93 | 94 | 96 | 98 | 100 |
|---|---|---|---|---|---|---|---|---|---|
| | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Other Receives | 15 | 19 | 24 | 28 | 33 | 37 | 41 | 46 | 50 |

Submit

## Part 3 : Decision 3 of 6

Choose a distribution between yourself and the other individual

| You Receive | 50 | 54 | 59 | 63 | 68 | 72 | 76 | 81 | 85 |
|---|---|---|---|---|---|---|---|---|---|
| | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Other Receives | 100 | 98 | 96 | 94 | 93 | 91 | 89 | 87 | 85 |

Submit

## Part 3 : Decision 4 of 6

Choose a distribution between yourself and the other individual

| You Receive | 50 | 54 | 59 | 63 | 68 | 72 | 76 | 81 | 85 |
|---|---|---|---|---|---|---|---|---|---|
| | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Other Receives | 100 | 89 | 79 | 68 | 58 | 47 | 36 | 26 | 15 |

Submit

## Part 3 : Decision 5 of 6

Choose a distribution between yourself and the other individual

| You Receive | 100 | 94 | 88 | 81 | 75 | 69 | 63 | 56 | 50 |
|---|---|---|---|---|---|---|---|---|---|
| | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Other Receives | 50 | 56 | 63 | 69 | 75 | 81 | 88 | 94 | 100 |

Submit

## Part 3 : Decision 6 of 6

Choose a distribution between yourself and the other individual

| You Receive | 100 | 98 | 96 | 94 | 93 | 91 | 89 | 87 | 85 |
|---|---|---|---|---|---|---|---|---|---|
| | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Other Receives | 50 | 54 | 59 | 63 | 68 | 72 | 76 | 81 | 85 |

Submit

## C1.11 Survey

*Questions 1 to 4 and 7 are identical across sessions. Question 5 differs between treatment and baseline sessions, while Question 6 differs between the $n = 2$ and $n = 4$ sessions.*

Please complete this short survey while we prepare your payment information.

1. What's your gender?
   - Female
   - Male
   - Other
   - Prefer not to say

2. Did you complete high school in an English-speaking country or region?
   - Yes
   - No
   - Prefer not to say

3. Which of the following best describes your major?
   - Agriculture
   - Business and Economics
   - Education
   - Engineering
   - Health and Human Sciences
   - Liberal Arts and Humanities
   - Other Social Sciences
   - Science and Mathematics
   - Others

4. Using 1-2 sentences, briefly describe your strategies of assigning probabilities to X and Y.

   Please describe your strategy...

5. *(For treatment sessions)* Within a match, did your decision of assigning probabilities to X and Y in a round depend on the current counter, on group outcomes in previous rounds, or both?

- The current counter only
- Group outcomes in previous rounds only
- Both the current counter and group outcomes in previous rounds

5. *(For baseline sessions)* Within a match, did your decision of assigning probabilities to X and Y in a round depend on the group outcomes in previous rounds?

- Yes
- No

6. Within Matches 1-5 and 6-10, was your strategy different between the initial matches and the later matches?

$n = 2$ *sessions*

Payoff table for Matches 1-5

| Your action | Group outcome (P1 P2) | Your payoff |
|---|---|---|
| X | X X | 100 |
| | At least one action is Y | 0 |
| Y | At least one action is Y | 85 |

Payoff table for Matches 6-10

| Your action | Group outcome (P1 P2) | Your payoff |
|---|---|---|
| X | X X | 100 |
| | At least one action is Y | 0 |
| Y | At least one action is Y | 15 |

$n = 4$ *sessions*

Payoff table for Matches 1-5

| Your action | Group outcome (P1 P2 P3 P4) | Your payoff |
|---|---|---|
| X | X X X X | 100 |
| | At least one action is Y | 0 |
| Y | At least one action is Y | 85 |

Payoff table for Matches 6-10

| Your action | Group outcome (P1 P2 P3 P4) | Your payoff |
|---|---|---|
| X | X X X X | 100 |
| | At least one action is Y | 0 |
| Y | At least one action is Y | 15 |

- Same within Matches 1-5; Same within Matches 6-10
- Different within Matches 1-5; Same within Matches 6-10
- Same within Matches 1-5; Different within Matches 6-10
- Different within Matches 1-5; Different within Matches 6-10

*The payoff tables shown above correspond to the high-first treatment and baseline sessions ($2h + 2l$ and $4h + 4l$). In the low-first treatment and baseline sessions ($2l + 2h$ and $4l + 4h$), the two payoff tables are displayed in reverse order.*

7. Were the strategies you use for Matches 6-10 influenced by the ones you use for Matches 1-5?

- Yes
- No

## C1.12 Final Payment Page

### Payment Summary

All 10 matches from Part 1 are paid. Exchange rate: **400 points = $1.00**.

| Match | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Points | 355 | 780 | 270 | 525 | 695 | 220 | 220 | 160 | 115 | 130 | **3470** |

| | |
|---|---|
| Comprehension quiz (0/6 correct) | $0.00 |
| Part 1 (3470 points) | $8.68 |
| Part 2 | $5.00 |
| Part 3 | $1.00 |
| **Final payment** | **$14.68** |

Please remain seated. The experimenter will come to you shortly.

# Appendix C2: Continuation Draws and Power Analysis

## C2.1 Continuation Draws

Table C2.1: Continuation Draws and Realized Match Lengths

| Match | Draws (Arrow landing positions) | Rounds |
|---|---|---|
| | **Sequence #1** | **Total: 48** |
| 1 | [51, 1, 95] | 3 |
| 2 | [20, 64, 74, 23, 20, 28, 57, 87] | 8 |
| 3 | [27, 98] | 2 |
| 4 | [70, 79, 7, 78, 88] | 5 |
| 5 | [37, 69, 41, 39, 60, 43, 85] | 7 |
| 6 | [25, 10, 48, 30, 62, 41, 2, 93] | 8 |
| 7 | [9, 77, 60, 7, 25, 31, 8, 91] | 8 |
| 8 | [51, 1, 73, 87] | 4 |
| 9 | [95] | 1 |
| 10 | [14, 90] | 2 |
| | **Sequence #2** | **Total: 45** |
| 1 | [49, 67, 88] | 3 |
| 2 | [48, 40, 17, 21, 93] | 5 |
| 3 | [25, 89] | 2 |
| 4 | [95] | 1 |
| 5 | [67, 8, 79, 11, 60, 33, 43, 5, 90] | 9 |
| 6 | [70, 42, 83] | 3 |
| 7 | [72, 7, 49, 23, 89] | 5 |
| 8 | [78, 74, 10, 58, 55, 96] | 6 |
| 9 | [14, 39, 62, 27, 75, 13, 48, 71, 20, 92] | 10 |
| 10 | [94] | 1 |
| | **Sequence #3** | **Total: 50** |
| 1 | [68, 14, 85] | 3 |
| 2 | [74, 19, 19, 1, 35, 44, 30, 10, 64, 97] | 10 |
| 3 | [78, 3, 46, 93] | 4 |
| 4 | [6, 74, 17, 44, 14, 89] | 6 |
| 5 | [66, 81] | 2 |
| 6 | [50, 4, 91] | 3 |
| 7 | [80, 6, 22, 87] | 4 |
| 8 | [46, 69, 90] | 3 |
| 9 | [18, 43, 66, 12, 68, 82] | 6 |
| 10 | [55, 55, 50, 25, 28, 33, 44, 24, 84] | 9 |

Table C2.1 reports the predrawn continuation draws for each round of the 10 matches in Part 1. Each draw is an integer from the discrete uniform distribution on $\{1, \ldots, 100\}$. A

draw at or below 80 indicates that the match continues to another round, whereas a draw above 80 terminates the match, corresponding to a continuation probability of $\delta = 0.8$. The continuation draw is revealed at the end of every round, so the match ends immediately following the first draw above 80. Accordingly, the final entry in each row of Table C2.1 is the terminating draw, and the number of draws in that row equals the realized length of the match. For example, in Match 2 of Sequence #1, the first seven draws are at or below 80 and the eighth draw is 87; the match therefore ends after Round 8.

The random draws are presented to subjects using a spinner with an 80% white continuation region and a 20% gray stopping region. A draw at or below 80 corresponds to the arrow landing in the white region, while a draw above 80 corresponds to the arrow landing in the gray region. Subjects observe the spinner realization after each round and learn at that point whether the match continues or ends. The three sequences contain 48, 45, and 50 rounds in total, respectively, across their 10 matches.

Each of the three sequences was used in four treatment sessions, once for each treatment order: $2l + 2h$, $2h + 2l$, $4l + 4h$, and $4h + 4l$. Thus, the three sequences generated the 12 treatment sessions. The $n = 2$ sessions contained 10-12 subjects, while each $n = 4$ session contained 16 subjects, for a total treatment sample of 164 subjects.

## C2.2 Power Analysis

I conducted an ex ante power analysis to assess the precision of alternative sample sizes for the main experiment. Because the group-size treatment is assigned at the session level, the comparison between $n = 2$ and $n = 4$ is the most demanding of the treatment comparisons and therefore serves as the power-driving test. The outcome used for this analytical power calculation is the subject-level mean probability assigned to action $X$, denoted by $\bar{p}_i(X) \in [0, 1]$.

The experiment is implemented in balanced blocks based on the predetermined sequences of match length.[19] For each sequence, I conduct four sessions corresponding to the four treatment combinations: $n = 2$ with $u_Y = 0.15$ first, $n = 2$ with $u_Y = 0.85$ first, $n = 4$ with $u_Y = 0.15$ first, and $n = 4$ with $u_Y = 0.85$ first. Each $n = 2$ session contains 12 subjects, and each $n = 4$ session contains 16 subjects. Thus, each complete block contains four sessions and $12 + 12 + 16 + 16 = 56$ subjects. Let $S$ denote the number of complete sequence blocks. Since each block contains two $n = 2$ sessions and two $n = 4$ sessions, there are $2S$ sessions in each group-size treatment. Table C2.2 summarizes the assumptions.

Given the experimental budget constraint, I report minimum detectable effects (MDEs)

[19] See Appendix C2 for predrawn sequences.

Table C2.2: Power Analysis: Assumptions

| Quantity | Value | Notes |
|---|---|---|
| Primary outcome | $\bar{p}_i(X)$ | Subject-level mean probability assigned to $X$ |
| Subject-level SD, $\sigma$ | 0.25 | Baseline assumption |
| | 0.20-0.30 | Alternatives |
| Session-level ICC, $\rho$ | 0.05 | Baseline assumption |
| | 0.10 | Conservative case |
| Subjects per $n = 2$ session | 12 | |
| Subjects per $n = 4$ session | 16 | |
| Sessions per sequence block | 4 | 2 sessions for each group-size treatment |
| Subjects per sequence block | 56 | |
| Significance level $\alpha$ | 0.05 | One-sided |
| Target power | 0.80 | |
| treatment comparison | | $n = 2$ vs. $n = 4$ |

rather than assuming a particular treatment-effect magnitude. The calculations use 80% power and a one-sided significance level of $\alpha = 0.05$, consistent with the directional prediction that coordination on $X$ is higher in smaller groups. I use $\sigma = 0.25$ as the baseline standard deviation of the subject-level mean probability assigned to $X$ and report sensitivity to $\sigma \in \{0.20, 0.25, 0.30\}$. To account for dependence among subjects within the same session, I consider a session-level intraclass correlation of $\rho = 0.05$ as the baseline case and $\rho = 0.10$ as a more conservative case. For this analytical approximation, I apply the same value of $\rho$ to the two group-size treatments.

Assume that the subject-level mean probability assigned to $X$ has variance

$$\mathrm{Var}(\bar{p}_i(X)) = \sigma^2.$$

Let $\rho$ denote the intraclass correlation (ICC) between the subject-level mean probabilities of any two subjects $i$ and $j$ in the same session. Thus,

$$\mathrm{Corr}(\bar{p}_i(X), \bar{p}_j(X)) = \rho,$$

which implies

$$\mathrm{Cov}(\bar{p}_i(X), \bar{p}_j(X)) = \rho\sigma^2.$$

The temporary groups formed within a session are not treated as independent clusters because group membership changes across matches. Dependence induced by repeated interaction and rematching is instead summarized by the session-level ICC $\rho$. For a session $s$

containing $m$ subjects, define the session-level mean as

$$\bar{p}_s(X) = \frac{1}{m}\sum_{i=1}^{m}\bar{p}_i(X).$$

The variance of the session mean is therefore

$$\begin{aligned}
\mathrm{Var}(\bar{p}_s(X)) &= \mathrm{Var}\left(\frac{1}{m}\sum_{i=1}^{m}\bar{p}_i(X)\right) \\
&= \frac{1}{m^2}\left[\sum_{i=1}^{m}\mathrm{Var}(\bar{p}_i(X)) + \sum_{i\neq j}\mathrm{Cov}(\bar{p}_i(X),\bar{p}_j(X))\right] \\
&= \frac{1}{m^2}\left[m\sigma^2 + m(m-1)\rho\sigma^2\right] \\
&= \frac{\sigma^2}{m}\left[1+(m-1)\rho\right] \\
&= \sigma^2\left[\rho + \frac{1-\rho}{m}\right].
\end{aligned}$$

Here, $\rho$ captures the component of subject-level variance associated with variation shared by subjects within the same session, while $1-\rho$ captures subject-specific variation, which is reduced by a factor of $m$ when averaged across subjects.

In the $n = 2$ treatment, each session contains 12 subjects. Hence, the variance of an $n = 2$ session mean is

$$V_2 = \sigma^2\left[\rho + \frac{1-\rho}{12}\right].$$

In the $n = 4$ treatment, each session contains 16 subjects. Hence, the variance of an $n = 4$ session mean is

$$V_4 = \sigma^2\left[\rho + \frac{1-\rho}{16}\right].$$

With $S$ complete sequence blocks, there are $2S$ sessions in each group-size treatment. The variance of the mean outcome in each treatment is therefore

$$\mathrm{Var}(\bar{p}_{n=2}(X)) = \frac{V_2}{2S}, \qquad \mathrm{Var}(\bar{p}_{n=4}(X)) = \frac{V_4}{2S}.$$

Since the $n = 2$ and $n = 4$ sessions are independent, the variance of the difference in treatment means is

$$\mathrm{Var}\left(\bar{p}_{n=2}(X) - \bar{p}_{n=4}(X)\right) = \frac{V_2}{2S} + \frac{V_4}{2S}.$$

The standard error of the group-size treatment difference is therefore

$$se = \sqrt{\frac{V_2}{2S} + \frac{V_4}{2S}}.$$

Because the two group-size treatments contain different numbers of subjects per session and therefore have different session-mean variances, I use the Welch-Satterthwaite approximation for the degrees of freedom:

$$\nu = \frac{\left(\frac{V_2}{2S} + \frac{V_4}{2S}\right)^2}{\frac{(V_2/(2S))^2}{2S-1} + \frac{(V_4/(2S))^2}{2S-1}}.$$

Let $\Delta^{\min}$ denote the minimum detectable group-size effect. Its corresponding noncentrality parameter is

$$\lambda = \frac{\Delta^{\min}}{se}.$$

For a one-sided test with significance level $\alpha = 0.05$, $\Delta^{\min}$ is chosen such that

$$\Pr\left(T_{\nu,\lambda} > t_{0.95,\nu}\right) = 0.80,$$

where $T_{\nu,\lambda}$ denotes a noncentral $t$ random variable with $\nu$ degrees of freedom and noncentrality parameter $\lambda$. For example, under the baseline assumptions $\sigma = 0.25$ and $\rho = 0.05$, with three complete sequence blocks ($S = 3$),

$$V_2 = 0.25^2\left[0.05 + \frac{0.95}{12}\right] = 0.008073,$$
$$V_4 = 0.25^2\left[0.05 + \frac{0.95}{16}\right] = 0.006836.$$

Since $S = 3$ implies six sessions in each group-size treatment,

$$se = \sqrt{\frac{0.008073}{6} + \frac{0.006836}{6}} = 0.04985.$$

The Welch-Satterthwaite degrees of freedom are

$$\nu = \frac{(0.008073/6 + 0.006836/6)^2}{\frac{(0.008073/6)^2}{5} + \frac{(0.006836/6)^2}{5}} \approx 9.93.$$

For a one-sided 5% significance level,

$$t_{0.95,9.93} \approx 1.814.$$

Solving

$$\Pr\left(T_{9.93,\lambda} > 1.814\right) = 0.80$$

gives

$$\lambda \approx 2.675.$$

Therefore,

$$\Delta^{\min} = \lambda \times se = 2.675 \times 0.04985 \approx 0.1333 = 13.3 \text{ percentage points.}$$

Thus, under the baseline variance and ICC assumptions, three sequence blocks, corresponding to 12 sessions and 168 subjects, provide 80% power to detect a group-size difference of approximately 13.3 percentage points, as shown in Table C2.3.

Table C2.3: Minimum Detectable Group-Size Effects and Candidate Sample Sizes

| # Sequence Sequence ($S$) | # Sessions | # Subjects | Minimum Detectable Effects | |
|---|---|---|---|---|
| | | | $\rho = 0.05$ | $\rho = 0.10$ |
| 1 | 4 | 56 | 34.5 *pp* | 40.6 *pp* |
| 2 | 8 | 112 | 17.2 *pp* | 20.3 *pp* |
| 3 | 12 | 168 | 13.3 *pp* | 15.7 *pp* |
| 4 | 16 | 224 | 11.3 *pp* | 13.3 *pp* |

Note: *pp* stands for percentage point. Calculations use $\sigma = 0.25$ and 80% power.

Because the standard deviation of the subject-level mean probability is not known prior to the main data collection, I additionally report sensitivity of the MDEs to $\sigma \in \{0.20, 0.25, 0.30\}$ for the two most relevant candidate sample sizes. Table C2.4 shows that the gain from a fourth sequence block persists across the range of variance and ICC assumptions.

Table C2.4: Sensitivity of Minimum Detectable Effects to the Assumed Standard Deviation

| $\sigma$ | 3 Sequences: 168 Subjects | | 4 Sequences: 224 Subjects | |
|---|---|---|---|---|
| | $\rho = 0.05$ | $\rho = 0.10$ | $\rho = 0.05$ | $\rho = 0.10$ |
| 0.20 | 10.6 *pp* | 12.6 *pp* | 9.0 *pp* | 10.6 *pp* |
| 0.25 | 13.3 *pp* | 15.7 *pp* | 11.3 *pp* | 13.3 *pp* |
| 0.30 | 16.0 *pp* | 18.8 *pp* | 13.6 *pp* | 16.0 *pp* |

Note: MDEs are calculated at 80% power with a one-sided significance level of 5%.

Based on the pilot sessions conducted in April, the estimated average payment is around \$22 per subject, implying an expected total participant payment of approximately \$3,700. With the available experimental budget in the range of \$3,000 - \$4,000, I therefore target three complete sequence blocks. This sample size also provides reasonable statistical preci-

sion for the power-driving group-size comparison. Under the baseline assumptions $\sigma = 0.25$ and $\rho = 0.05$, the three-block design has 80% power to detect a group-size effect of approximately 13.3 percentage points. Under the more conservative ICC assumption $\rho = 0.10$, the corresponding minimum detectable effect is approximately 15.7 percentage points. The sensitivity analysis shows that the resulting MDEs remain within a comparable range under alternative assumptions about the subject-level standard deviation.