Coordination via Delay: Theory and Experiment*

Ye Jin† Zhen Zhou‡ Adam Brandenburger§

Abstract

This paper studies the effect of introducing an option of delay in coordination games, that is, of allowing players to wait and then choose between the risk-dominant and payoff-dominant actions. The delay option enables forward-induction reasoning to operate, whereby a player's waiting and not choosing the risk-dominant action right away signals an intention to choose the payoff-dominant action later. If players have ϵ-social preferences – they help other players if at no cost to themselves – then iterated weak dominance yields a unique outcome where everyone waits and then chooses the payoff-dominant action if everyone else waited. Thus, efficient coordination results. Experimental evidence from a binary-action minimum-effort game confirms that adding a delay option can significantly increase the occurrence of efficient outcomes.

Key Words: *Coordination, Forward Induction, Iterated Weak Dominance*

JEL Classification Numbers: *C73, C92, D83*

*We thank Ala Avoyan, Deepal Basak, Giacomo Bonanno, Jordi Brandts, Stefan Bucher, Emiliano Catonini, Archishman Chakraborty, Vered Kurtz, Barry Nalebuff, Joao Ramos, Satoru Takahashi, John Wooders (discussant), and seminar and conference participants at NUS, NYU Shanghai, UC Davis, ESWC2020 and WESSI2020 for helpful comments and suggestions. We thank Shuhuai Zhang and Dandan Zhao for outstanding research assistance, and the National Science Foundation of China (project no. 71703101), the NYU Stern School of Business, NYU Shanghai, and J.P. Valles for financial support. We are grateful to the Smith Lab at Shanghai Jiao Tong University for hosting the lab experiment.

†NYU Shanghai, Shanghai, China 200122. Email: wendyjin@nyu.edu.

‡PBC School of Finance, Tsinghua University, Beijing, China 100083. Email: zhouzh@pbcsf.tsinghua.edu.cn.

§Stern School of Business, Tandon School of Engineering, NYU Shanghai, New York University, New York, NY 10012, U.S.A.. Email: adam.brandenburger@nyu.edu.

1 Introduction

Coordination games are models of the challenge of coordination among economic (or other) players. The coordination challenge consists of two parts. First, there is the challenge of achieving a Nash equilibrium of the game. Consider a simple coordination game as shown in Table 1. In this 2×2

		Player 2	
		A	B
Player 1	A	$4, 4$	$0, x$
	B	$x, 0$	x, x

Table 1: A Simple Coordination Game

game with $x \in (2, 4)$, there are two pure-strategy Nash equilibria – one in which all players choose the safe but inefficient action B and one in which all players choose the risky but efficient action A. However, miscoordination may occur if one player chooses A, while the other chooses B

The (pure-strategy) Nash equilibria can be Pareto-ranked, so that the second challenge is attaining an efficient equilibrium. In this simple example, the payoff-dominant equilibrium (A, A) differs from the risk-dominant equilibrium (B, B) in the sense of Harsanyi and Selten (1988). Carlsson and van Damme (1993) show that the risk-dominant equilibrium (B, B) is uniquely selected if we relax common knowledge of the payoffs (for example, common knowledge of the parameter x). The selection of the risk-dominant equilibrium is supported by ample experimental evidence (see, for example, Cooper et al. (1992*a*) and Van Huyck, Battalio and Beil (1990)).

How can the players overcome the challenge of coordination? In reality, coordination games are often played dynamically, and the option of "wait and see" is often available. For instance, in the bank-run game, which is a prime example of a coordination game, each investor might be able to wait and then make their final withdrawal decision conditional on the information they observe.[1] This paper explores a class of dynamic games that allow each player to exercise such an option to delay choice of the efficient action. Furthermore, the paper shows that the addition of the delay option can help players overcome miscoordination and also achieve the Pareto-dominant outcome.

The delay option, if exercised, enables a player to observe the past history of play by other players. However, more than observability, this paper highlights the idea that exercising the delay option and not taking the inefficient action early enables a player to signal their intention of taking the efficient choice in future play – signaling this to other players who also exercise that option. We show, both theoretically and experimentally, that signaling through exercising the delay option, i.e.,

[1] For other examples of applications of coordination games, see, among other examples, new technology adoption (Farrell and Saloner, 1985; Katz and Shapiro, 1986), team production (Bryant, 1983), search (Diamond, 1982), currency attacks (Morris and Shin, 1998), and debt crises (Corsetti, Guimaraes and Roubini, 2006).

adopting the strategy of waiting and then taking the efficient action if all others wait, can work effectively to overcome the challenge of coordination.

The main result is proved for a multiple-player coordination game in which the efficient outcome is achieved if all players choose action A. For now, we will rely on the simple 2×2 game above to illustrate the intuition. The game unfolds in two periods: $t = 0, 1$ when the option of delay is available. At $t = 0$, each player can choose between the irreversible choice B and "wait." A player who chooses to wait observes whether or not the other player chooses B at $t = 0$, and then makes their final choice between A and B at $t = 1$. There is no cost associated with the delay option. A player who does not choose B at $t = 0$ should, in some sense, be signaling that they intend to choose A at $t = 1$. That is, there is a forward-induction flavor to choosing "wait." Intuitively, if a player intends to play B and secure the safe payoff x, then they can do so right away, rather than waiting and doing so later, which does not result in any extra benefit.[2] Next, we describe our analysis in more detail.

First, observe that if a player waits and then receives the "B" message (i.e., the other player chooses B at $t = 0$), they will optimally choose B at $t = 1$. Formally, any strategy that involves choosing A after the "B" message is weakly dominated. Next, consider the situation in which a player receives the message "no-B." The game then enters a simultaneous-move subgame in which both players make final choices between A and B. A player might decide to choose B in this subgame if they believe that the other player will do so. But then this player chooses B after either message. Compare this with the strategy of choosing B at $t = 0$. The two strategies yield the same payoff to our player for each strategy profile of the other player. But choosing B after the message "no-B" hurts the opponent if they choose A after observing "no-B." To capture this idea, we assume that the players in our game have ϵ-social preferences. By this, we mean that each player has a utility function that is given by their original payoff function plus an infinitesimal weight on the other players' payoff functions. In other words, a player will not gratuitously hurt another player, where, by "gratuitously" we mean that one makes a choice that hurts others without helping oneself.

Formally, we have just argued that, with ϵ-social preferences, the strategy of waiting and then playing B, regardless of the message received, is weakly dominated – in the second round of elimination – by playing B immediately (at $t = 0$). Once this strategy is eliminated, the strategy of playing B immediately is weakly dominated by waiting and then playing B after the "B" message and A after the "no-B" message. In effect, by choosing "wait," a player signals their intention to follow the message and, in particular, to choose the efficient action A after the "no-B" message. This is the sole strategy that survives iterated weak dominance, and the result is that each player chooses

[2] The simple idea that any player, if choosing to wait, tends to take different actions based on different observed histories is also in Chamley and Gale (1994) and Gul and Lundholm (1995), who study the delay option in a (non-strategic) social learning setup.

"wait" at $t = 0$, collectively generating the "no-B" message, and then takes action A. In that way, the Pareto-dominant outcome is achieved, even when it is risk-dominated.

There are two key components to this analysis. The first is forward induction. Introduced by Kohlberg and Mertens (1986), forward induction has been formalized in two different ways. One way is as an equilibrium refinement (Govindan and Wilson, 2009; Kohlberg and Mertens, 1986). The other way is as iterated weak dominance (Ben-Porath and Dekel, 1992).[3] Cooper et al. (1992*b*) provide early evidence of forward-induction reasoning in coordination games with Pareto-ranked equilibria. They find that granting an outside option with an appropriate payoff to one player significantly improves coordination efficiency. Subsequent papers consider other pre-play moves, such as pre-auctions (Cachon and Camerer, 1996; Van Huyck, Battalio and Beil, 1993) or costly messages (Blume, Kriss and Weber, 2017).[4] While these studies discuss forward induction in terms of the equilibrium-refinement route, we follow the iterated weak dominance route in analyzing what we believe is likely a built-in feature of various real-world situations modeled as coordination – namely, an option to delay the risky choice.

The second key component of our analysis is the inclusion of social preferences. Other-regarding preferences have been identified in various experimental studies (see Fehr and Schmidt (2006) for a survey). We adopt a very weak form of social preference in which there is no trade-off between a player's own payoff and those of other players, since the weight ϵ is infinitesimal. It implies that other players' payoff functions are decisive only when the player is choosing between two equivalent strategies. Here, two strategies are equivalent for a player if, for every strategy profile of the other players, they yield the same payoff to that player. This particular concept should be contrasted with the usual models of altruism, which, in many games, will modify a player's original preferences in more ways than our condition does.

In addition to iterated weak dominance, we analyze the game via symmetric pure-strategy subgame-perfect Nash equilibrium,[5]. This solution concept does not lead to a unique prediction of efficient coordination, even with ϵ-social preferences. To see this, consider our 2×2 example again, and focus on the subgame starting from the "no-B" message. If the other player chooses B at this information set, playing A hurts a player without benefiting the other player. Thus, subgame perfection, together with ϵ-social preferences, does not rule out the Nash equilibrium in which everyone waits and chooses the inefficient action B regardless of the $t = 0$ outcome.

Motivated by this sharp difference in analyses across different solution concepts, we design an experiment by adding the delay option to a binary-action weakest-link game, or minimum-effort game, and examine its effectiveness in overcoming the challenges of coordination. The minimum-effort

[3] Brandenburger, Friedenberg and Keisler (2008) provide an epistemic foundation for iterated weak dominance.

[4] For experimental evidence on forward induction in Battle-of-the-Sexes games and entry games, see Cooper et al. (1993), Brandts and Holt (1995), Huck and Müller (2005), and Brandts, Cabrales and Charness (2007). Brandts, Cabrales and Charness (2007) adopt the iterated weak dominance formalization of forward induction.

[5] Throughout the paper, we focus on symmetric Nash equilibria in pure strategies.

game, first examined experimentally in Van Huyck, Battalio and Beil (1990), is a well-studied example of coordination with multiple Pareto-ranked equilibria. A large literature has documented the difficulty in promoting the efficient outcome. Our experimental evidence confirms that, while the efficient outcome is hard to achieve in the static version of the game, with the presence of the delay option, waiting can be a meaningful signal for future play of the risky but efficient action A. We adopted the strategy method to elicit the subjects' full plans of play, and we find that, conditional on waiting, over 90 percent of the subjects made the efficient choice A at $t = 1$ after receiving the "No-B" message. Consistent with our theory, on average, 75 to 85 percent of the subjects in each round adopted the unique strategy surviving iterated weak dominance – namely, waiting and then choosing A after the "no-B" message and B after the "B" message. Because of the prevalence of this particular strategy, notably, around 60 percent of the groups achieved efficient coordination. In contrast, in the static treatment without the delay option, the efficiency rate was only 14 percent.[6] In addition, the experimental evidence indicates that the delay option significantly reduces the incidence of miscoordination and improves the individual subjects' payoffs.

Efficient equilibrium in dynamic coordination games has been studied in the literature. However, the dynamic setup we consider, in terms of the asynchronicity of moves and the irreversibility and observability of actions, is different. Unlike Lagunoff and Matsui (1997) and Dutta (2012), who study the repeated play of coordination game under infinite and finite horizons, respectively, players in our model play the 2-period game once. Throughout the paper, we consider only the cases in which actions are either fully reversible or fully irreversible, while Calcagno et al. (2014) investigate the preparation stage, in which actions are partially reversible, since they can be revised stochastically with a deadline.[7]

Our dynamic setup is also different from the case in which players move sequentially to coordinate. When the moves are asynchronous in a pre-specified order, because of the observability of past actions, backward induction would predict the efficient outcome (see Farrell and Saloner (1985)). Next suppose the order is endogenous and the efficient choice (instead of the inferior one in our benchmark setting)[8] is irreversible in the earlier period. Even then, the players, knowing that their early actions can be observed and expecting others to follow, may take the lead by committing to the efficient action earlier and, thereby, arrive at the efficient outcome.[9]

[6] The efficiency rate of a round in a certain treatment is defined as the percentage of groups in which every group member's realized choice is the efficient action A.

[7] In a recent study, Avoyan and Ramos (2019) apply the stochastic revision mechanism to the minimum-effort game and show, theoretically and experimentally, that this mechanism of partial commitment can help to promote efficient coordination on high effort making.

[8] We also investigate the case in which the efficient action is the only irreversible one (see Section 2.2) and experimentally test the efficacy of such a delay option in promoting efficient coordination (see Section 4.3).

[9] Farrell and Saloner (1985) discuss this case in a multiple-period setting, but they restrict their attention to an endogenously chosen order of asynchronous moves; that is, they exclude the possibility that players can take the efficient action simultaneously at any period.

This paper highlights a mechanism enabled by the option of delay, which is different from backward induction in a dynamic setting with asynchronous moving. Exercising the delay option signals the players' intention to play the efficient action in the future and, at the same time, enables the player in question to observe other players' delay choices and to infer their intentions for future play. Efficient coordination can be achieved through everyone's waiting and then taking the efficient action simultaneously.

Furthermore, when the delay option is available, and if both actions are reversible, then, for each player, taking a certain action in the first period can be thought of as a way of sending a non-binding message expressing their intention for future play. It is appealing that pre-play communication can lead to the efficient outcome in common-interest games.[10] There is ample experimental evidence confirming the effectiveness of costless communication in coordination games.[11] To investigate the difference between signaling intention by waiting – our mechanism – and communicating intention via "cheap talk," we compare these two mechanisms in our experiment by examining the difference between the $t = 0$ choice (and the resulting messages). We find that coordination via delay is more effective than coordination via pre-play communication.

The remainder of the paper is organized as follows. Section 2 presents the static benchmark model and the theory of the dynamic protocol. The experimental design and procedure are discussed in Section 3. Section 4 reports the experimental results, and Section 5 concludes. Some of the proofs are relegated to the Appendix.

2 Theory

We start with a static binary-action coordination game. There are $N \geq 2$ players, indexed by $i \in \mathcal{N} = \{1, 2, \ldots, N\}$. Each player i simultaneously makes a decision d_i from $D_i := \{A, B\}$. For any action/strategy profile $d \in \prod D_i$, the monetary payoff of player i is denoted by $\pi_i(d)$. For each player i, the payoff $\pi_i(\cdot)$ satisfies the conditions: (a) if a player chooses $d_i = B$, then $\pi_i(d_i = B; d_{-i}) = b$,

[10] Regarding the credibility of non-binding messages in one-way communication, see Farrell (1988) for the notion of self-committing and Aumann (1990) for the notion of self-signaling. For follow-up works on coordination games, see Baliga and Morris (2002), Sobel (2017), and Lo (2020). It is worth noting that, in the case of $N = 2$ (with one sender and one receiver), the message "A" is self-committing in one-way communication (i.e., the sender will take A at $t = 1$ if they expect the receiver to trust the message), and it is self-signaling with the assumed social preferences (i.e., the sender wants the receiver to trust the message if and only if the announcement is credible and they would choose A at $t = 1$). However, in our game, there are multiple players, and each of them can "send messages" to all the other players. In this case, since we need to consider a profile of messages, the notions of self-committing and self-signaling are not well defined (see the discussion in Blume and Ortmann (2007)).

[11] Cooper et al. (1992*b*) document the efficacy of two-way communication, and find that one-way messages are much less effective. By contrast, Charness (2000) finds that one-way communication is powerful in facilitating efficient coordination. Blume and Ortmann (2007) demonstrate the effectiveness of multi-way communication in minimum-effort and median-effort games.

regardless of the other players' choice d_{-i}; (b) if a player chooses $d_i = A$, then, $\pi_i(d_i = A; d_{-i}) = c < b$ for any d_{-i} that involves $d_j = B$ for some player $j \neq i$; and (c) $\pi_i(d_i = A; (d_j = A)_{j \neq i}) = a > b$.

In words, B is a safe choice, which yields a payoff of b, regardless of other players' choices, while A is a risky choice, which yields a high payoff a only if coordination is successful – that is, if all players choose A. Otherwise, if any other player chooses B, choosing A yields a low payoff of c.

We further assume that each player i has the following utility function:

$$u_i(d) = \pi_i(d) + \epsilon \sum_{j \neq i} \omega_{ij} \pi_j(d), \tag{1}$$

where ϵ is positive and infinitesimal, and the w_{ij} are strictly positive numbers. We say that player i has *ϵ-social preferences*. Under this assumption, player i's preference is lexicographic - there is no tradeoff between i's own monetary payoff and that of any other player j.

Proposition 1 *In this static binary-action coordination game, the pure-strategy Nash equilibria are $d_i = A$ for all i and $d_i = B$ for all i.*

Clearly, when players are selfish ($\epsilon = 0$), there are two equilibria in this simple static coordination game. Proposition 1 confirms that the set of equilibria is the same with ϵ-social preferences.

Coordinating on the risky choice A yields the highest payoff for all players, regardless of whether or not players have the social preferences as defined in (1). Formally, *efficient coordination* is achieved if and only if $d_i = A$ for all $i \in \mathcal{N}$. As shown in Proposition 1, efficient coordination is not guaranteed, since coordinating on the safe action B constitutes another equilibrium. *Miscoordination* occurs if players fail to coordinate on a certain equilibrium – that is, if there are some player(s) who choose A, while some other(s) choose B. Miscoordination incurs a significant loss to the players who choose A.

Next, we add a dynamic structure to the static coordination game, which enables each player to exercise a delay option so that they can choose between A and B at a later date. We will investigate how this delay option changes the outcome.

2.1 Dynamic Structure with Irreversible Choice of B

There are two periods, $t = 0, 1$. At $t = 0$, each player chooses between B and "wait." The choice of B is irreversible. That is, if a player chooses B at $t = 0$, they cannot make any further changes. However, the players who wait at $t = 0$ get to choose between A and B at $t = 1$. There is no cost associated with waiting. By waiting, players can observe a binary message m that indicates whether a player chose B earlier, at $t = 0$ ($m = 1$), or not ($m = 0$). The realized message is $m = 0$ if and only if all players choose to wait at $t = 0$.

We denote the set of pure strategies as

$$\mathcal{S} \equiv \{B, WBB, WBA, WAB, WAA\}.$$

The strategy of not waiting and taking B at $t = 0$ is denoted as B. Any strategy involving waiting at $t = 0$ is a plan contingent on the message m. We denote such a strategy as "W, $d_i(m = 1)$, and $d_i(m = 0)$," respectively, where $d_i(m)$ is defined as the action chosen conditional on m at $t = 1$. For example, if a player takes strategy WAB, they will wait at $t = 0$ and then choose A after observing $m = 1$; otherwise, they will choose B.

For any strategy profile of other players $s_{-i} = (s_j)_{j \in \mathcal{N} \setminus \{i\}}$, if player i chooses to wait, then the total number of B choices at $t = 0$ can be written as

$$n(s_{-i}) = \big|\{j \in \mathcal{N} \setminus \{i\} | s_j = B\}\big|,$$

and, accordingly, the binary message that player i will receive after waiting is

$$m = \mathbb{1}\{n(s_{-i}) \geq 1\}.$$

$m = 0$ and 1 correspond to the "no-B" and "B" messages, respectively.

Proposition 2 *For any $s_i \in \mathcal{S}$, the strategy profile $(s_i)_{i=1}^N$ constitutes a pure-strategy Nash equilibrium. The subgame-perfect equilibria are $(s_i = B)_{i=1}^N, (s_i = WBA)_{i=1}^N$, and $(s_i = WBB)_{i=1}^N$.*

It is easy to see that choosing A in the subgame following the message $m = 1$ (the "B" message) cannot be part of an equilibrium in this subgame. Still, as Proposition 2 demonstrates, subgame perfection does not yield a unique outcome or imply efficient coordination. In the subgame-perfect equilibria in which all players choose $s_i = B$, or in which they all choose $s_i = WBB$, each player ends up choosing B, and, therefore, efficiency does not result.

In the following theorem, we formalize forward induction as iterated simultaneous maximal deletion of weakly dominated strategies, which we henceforth simply call iterated weak dominance. The theorem shows that this procedure yields a unique strategy profile, which achieves efficient coordination.

Theorem 1 *The unique strategy profile that survives iterated weak dominance is $(s_i = WBA)_{i=1}^N$. Under this strategy profile, efficient coordination is achieved.*

The argument involves three rounds of elimination. Here, for the purpose of illustration, we use the payoff matrix of a two-player example (see Table 2) to illustrate the elimination process. We give the main argument for each step of elimination and relegate the complete proof to the Appendix.

		Player 2				
		B	WBB	WBA	WAB	WAA
	B	b, b	b, b	b, b	b, c	b, c
	WBB	b, b	b, b	b, c	b, b	b, c
Player 1	WBA	b, b	c, b	a, a	c, b	a, a
	WAB	c, b	b, b	b, c	b, b	b, c
	WAA	c, b	c, b	a, a	c, b	a, a

Table 2: 2-Player Payoff Matrix

First Round (Eliminate WAB and WAA) WAB is weakly dominated by WBB. To see this, note that after the message $m = 0$, these two strategies yield equivalent outcomes. When $m = 1$, WAB involves choosing A and yields a private payoff $\pi_i = c$, while WBB yields a private payoff $\pi_i = b > c$. The same argument can be used to show that WAA is weakly dominated by WBA.[12], [13]

Second Round (Eliminate WBB) After the first round of elimination, the remaining pure strategies are B, WBB, and WBA. Regardless of what other players choose, the realized choice under both strategies B and WBB is B. Thus, these two strategies yield the same payoff $\pi_i = b$ to any player i.

Both strategies induce the same payoff to player $j \neq i$ in all but one case, in which all players $j \neq i$ choose to wait at $t = 0$, and at least some $j \neq i$ choose the strategy WBA. In this case, if player i chooses B, a player j who chooses WBA gets payoff $\pi_j = b$ from playing B after seeing $m = 1$. However, player j's payoff is reduced to c if player i chooses WBB because, in this case, player j's realized choice is A, following $m = 0$. Therefore, under the assumption of ϵ-social preferences, B weakly dominates WBB.

Third Round (Eliminate B) Two strategies remain after the second round: B and WBA. If at least one player $j \neq i$ chooses B at $t = 0$, these two strategies yield the same payoff to player i and to all other players. However, if all $j \neq i$ choose WBA, then WBA yields a strictly higher payoff to i. Thus, B is weakly dominated by WBA.

[12] This round of elimination holds for both $\epsilon = 0$ and $\epsilon > 0$. In fact, both WAB (resp. WAA) and WBB (resp. WBA) yield the same payoff to each of the other players, regardless of any possible strategies they take.

[13] The strategy WBB is not weakly dominated by B in the first round of elimination. To see this, consider the case in which all other players choose WAB. In this case, compared with B, WBB generates strictly higher payoffs to the other players. As such, ϵ-social preference cannot eliminate WBB before the second round of elimination.

Coordination Outcome Since all players choose the strategy WBA, the realized message is $m = 0$, and, thus, the realized choice is $d_i = A$ for all $i \in \mathcal{N}$. Therefore, efficient coordination is achieved.

2.2 Discussion

We consider a simple binary-action coordination game with $N \geq 2$ players. By incorporating a delay option into the static game, we create a dynamic variant in which the safe but inefficient choice B is the only irreversible action. The players who exercise the delay option can observe a binary message about whether or not all players have taken the delay option. Somewhat surprisingly, there is a unique strategy WBA that survives iterated weak dominance in the resulting dynamic game. Under this strategy, a player, by giving up the safe but inefficient choice and exercising the delay option at $t = 0$, signals their intention to play the risky but efficient choice A (conditional on observing that all other players chose to wait).

Under this unique strategy profile, efficient coordination is achieved. This result is built on forward-induction reasoning, which becomes possible only when players have ϵ-social preferences. Next, we discuss how our result depends on the extensive form that governs the play, i.e., on the observability of the history of play and the (ir)reversiblity of the actions.

Observability of Past Actions

In our benchmark model, players who choose to wait can observe only a binary message regarding the history. This is a deliberate assumption meant to capture the difficulty of observing the precise history of play in a multiple-player setting. But a delay option per se does not rule out cases in which players can observe more information about the past history.

Here, we consider an environment in which any player i who exercises the delay option can observe the exact number of irreversible choices that occurred at $t = 0$. We denote this number by $n(s_{-i})$ and say that this scenario exhibits *finer information*.[14] With finer information, the strategy of waiting at $t = 0$ and then choosing A at $t = 1$ if and only if $n = 0$ remains the unique strategy profile that survives iterated weak dominance. To reduce the notational burden, in the finer-information setting, we continue to write "WBA" for this strategy.

Proposition 3 *With finer information, the unique strategy profile that survives iterated weak dominance is* $(s_i = WBA)_{i=1}^N$. *Under this strategy profile, efficient coordination is achieved.*

Note that, efficient coordination cannot be achieved as long as someone chooses B at $t = 0$, that is, $n(s_{-i}) \geq 1$, irrespective of the exact number of B choices. After any information set $n(s_{-i}) \geq 1$, the

[14] Note that "finer information" here is different from perfect information, since the identities of the players who choose B and who choose to wait at $t = 0$ remain unknown.

best response for any player i who waited is to play B at $t = 1$. Therefore, as Proposition 3 states, providing finer information by partitioning the information set of $m = 1$ ("B" message) in the binary message setting does not change the unique strategy players take or the coordination outcome. Our mechanism is robust to finer information because the intention to coordinate efficiently is signaled via an information set that is a singleton ($n = 0$), which is exactly the same as the information set $m = 0$ ("no B" message) in the binary message setting.

(Ir)reversibility Structure

We have argued that a delay option can resolve the coordination problem if the inefficient choice B is the only binding choice at $t = 0$. What if both actions A and B are reversible, or, the efficient choice A, instead of B, is the only irreversible choice at $t = 0$? In this subsection, we discuss the essentiality of the reversibility structure to our result.

Neither Choice is Irreversible We first consider the case in which neither A nor B is irreversible at $t = 0$. More precisely, players choose between A and B at $t = 0$, but the first-period choice is not binding. At $t = 1$, they first observe the number of A and B choices at $t = 0$, and then make a final choice between A and B. Since a player's payoff depends only on their action at $t = 1$, their choice at $t = 0$ is payoff-irrelevant.

In fact, we can interpret the play at $t = 0$ as costless pre-play communication and the play at $t = 1$ as the actual coordination game. It is easy to check that neither subgame-perfect equilibrium nor iterated weak dominance generates a unique prediction, even with the assumed ϵ-social preferences. For instance, there is a subgame-perfect equilibrium in which all players choose A at $t = 0$, and then all switch to B at $t = 1$, regardless of the information they observe. In another equilibrium, everyone chooses B at $t = 0$ and makes no change at $t = 1$, regardless of the information observed. These strategy profiles also survive iterated weak dominance and lead to inefficient outcomes.

Action A is Irreversible Now consider the case in which the efficient choice A is the only irreversible choice at $t = 0$. More precisely, players first choose between "wait" and A at $t = 0$. After observing the number of wait and A choices at $t = 0$, players who waited choose between A and B at $t = 1$. This setting has an irreversibility structure that is the opposite of our benchmark setup. The following proposition shows that, iterated weak dominance and subgame-perfect equilibrium predict the efficient outcome only in the case of $N = 2$. For $N \geq 3$ players, theory fails to yield (uniquely) the coordination outcome.

Proposition 4 *In a coordination game, in which A is the only irreversible choice at $t = 0$:*

1. *if $N = 2$, both iterated weak dominance and subgame-perfect equilibrium yield efficient strategy profiles;*

2. *if $N \geq 3$, both iterated weak dominance and subgame-perfect equilibrium allow inefficient strategy profiles.*

As shown in the proof of Proposition 4 in the Appendix, for $N \geq 3$ players, neither iterated weak dominance nor subgame-perfect equilibrium guarantees efficiency. For example, the strategy profile where all players choose wait at $t = 0$ and choose A at $t = 1$ only if all other players have chosen A at $t = 0$ survives iterated weak dominance, and it constitutes a subgame-perfect equilibrium.[15] Under this strategy profile, each player's final choice is B and, thus, efficient outcome is not achieved.

One might think that, in this setting, efficient coordination is easier to achieve because, now, any player can take the lead by committing to the binding choice A early, at $t = 0$. Intuitively, these early A choices would encourage other players to follow. However, Proposition 4 says that this intuition is false. Intuitively, the only message that ensures that a player will follow and choose A at $t = 1$ is that all other players have chosen A at $t = 0$. When the game is played by $N \geq 3$ players, it is possible that all players who waited would choose B at $t = 1$ if more than one player waited at $t = 0$, since they anticipate that other players will do the same. Therefore, efficient coordination cannot be guaranteed.

2.3 Summary

Our main results show that the option to delay facilitates efficient coordination by allowing players to signal their intentions. The above discussion demonstrates, at a theoretical level, that signaling intentions by waiting is different from signaling intentions by non-binding communication (when both actions are reversible) and also different from signaling intentions by taking the efficient action early (when the efficient action A is irreversible). Therefore, although the mechanism we emphasize is robust to the availability of finer information, it relies crucially on the irreversibility of the inefficient choice.

For completeness, we extend the model further to show that the delay mechanism can work in a more general coordination game, in which the successful coordination does not require all players to choose the efficient choice A. We also discuss the case in which both actions are irreversible choices, as well as the case in which delay is costly.[16] Since these extensions are not essential to our theoretical analysis and experimental tests, we relegate them to the online appendix.[17]

[15]This strategy is, in fact, the most frequently used one in our experiment.

[16]In a more general class of coordination games, or when both A and B are irreversible in our benchmark setup, our result is sensitive to the information available to the players who exercise the delay option. Specifically, our result holds in the binary-information setting but does not hold in the finer information-setting. Moreover, although the mechanism of signaling intentions can still work with a costly delay option, this mechanism cannot ensure efficient coordination.

[17]The online appendix can be found at www.zhenzhoueconomics.com/research.

3 Experimental Design

Our theory demonstrates that the dynamic structure with an irreversible B choice admits a unique prediction of efficient coordination via iterated dominance. However, the inferior outcome still qualifies as a subgame-perfect equilibrium, even with the assumption of ϵ-social preferences. Therefore, we design an experiment to test the efficacy of this delay structure and check whether participants' choices are consistent with our theoretical prediction based on iterated dominance.

The alternative irreversibility structures do not admit unique predictions via iterated dominance. Thus, it is an open question whether these structures also facilitate coordination. In the experiment, we compare the efficacy of the three irreversibility structures (A-irreversible, B-irreversible, neither-irreversible) in promoting efficiency under finer information.

The binary-action coordination game we consider can be interpreted as a weakest-link game, or minimum-effort game, with high effort level A and low effort level B, as the group coordination is determined by the lowest choice in the group. In our main treatments, we follow the design of the minimum-effort games literature, which started with Van Huyck, Battalio and Beil (1990),[18] so as to make our experimental findings comparable to those of the existing studies.

Following the standard protocol in the minimum-effort game literature, subjects played a game for 15 rounds in fixed four-person groups for all static and dynamic treatments. Throughout a session, subjects' strategies were elicited using the strategy method in the dynamic games. The parameters chosen were $a = 55$, $b = 45$, $c = 5$, and $N = 4$. A more detailed description of the experimental implementation will be given in Section 3.3.

3.1 Main Treatments

The main treatments compare the coordination efficiency in static games and the dynamic games with the irreversible B choice. To examine the efficacy of the delay option, we are interested in whether efficient coordination can be achieved in the dynamic treatment and, in particular, whether efficiency is achieved through the unique iteratedly undominated strategy WBA.

Static Game ("St-b")

The "St-b" treatment is the static version of the binary-action coordination game. At the end of each round, subjects were informed of whether the efficient outcome is achieved. Feedback about only the coordination outcome was the standard protocol in the minimum-effort literature; that is, subjects

[18] A large subsequent literature investigates various mechanisms to promote efficient coordination. In particular, Van Huyck, Battalio and Beil (1993) and Cachon and Camerer (1996) study the effect of pre-game auctions on coordination in the minimum-effort games, which involves forward-induction reasoning. For other examples, see Weber (2006), Brandts and Cooper (2006), Devetag and Ortmann (2007), and Chen and Chen (2011). See Devetag and Ortmann (2007) for a survey.

observe only the minimum effort chosen in the previous rounds. It differs from the finer-information setting, which discloses the number of players choosing A or B. The binary feedback is referred to as "binary information" ("b" for short) in the static treatment.

Dynamic Game with Irreversible B Action ("BI-b")

The main treatment, "BI-b," follows the dynamic structure proposed in Section 2.1 (see Proposition 2 and Theorem 1), where B is the only irreversible action ("BI" for short), and players could observe the binary message m at the end of $t = 0$. Each round of the game consists of two periods. At $t = 0$, each player chooses between B and the wait option. If a subject chooses to wait at $t = 0$, they receive a binary message ($m = 0, 1$). $m = 0$ ($m = 1$) corresponds to "nobody (someone) in the group has chosen B in $t = 0$." After that, the player can make the final choice of A or B.

In the dynamic treatments, the term "binary information" ("b" for short) is characterized by two information settings. First, subjects received binary feedback at the end of each round (same as in the static stage game). Second, binary information also refers to the binary message available to the players who chose to wait at $t = 0$ - that is, whether someone chose B at $t = 0$. Therefore, either between two periods, or at the end of a round, subjects were always informed of the binary outcome - that is, whether or not B was chosen, but not of the exact numbers of the irreversible choices.

3.2 Additional Treatments

Finer-Information Treatments ("St-f" and "BI-f")

In addition to the main treatments, in which subjects received binary feedback only about the coordination outcome, we also tested the finer-information versions of these two treatments: "St-f" (static, finer information) and "BI-f" (B-irreversible, finer information). In all finer-information treatments ("f" for short), subjects were informed of how many group members chose A or B at the end of each round. In the dynamic treatments, under finer information, subjects who waited also received information about the number of irreversible choices at $t = 0$.

More precisely, in the "BI-f" treatment, if a subject decided to wait at $t = 0$, they would face four possible situations: everybody waited, or 1, 2, or 3 group members chose B. Therefore, the subject's strategy would be to wait or not to wait at $t = 0$, and, if they waited, a full plan on these four contingencies.

We added these two finer-information treatments for two reasons. First, when subjects played the coordination game repeatedly in our lab experiment, they were able to learn about the coordination outcome and the strategies adopted by other subjects in the earlier rounds. The learning effect was not captured by our theory and might have affected the coordination outcome in the later rounds. In particular, the cross-round learning could, presumably, have affected the coordination outcome of the static and dynamic games in different ways.[19] Since our theory predicts that the same efficient outcome could be generated in both the binary ("BI-b") and the finer-information ("BI-f") settings, treatments with the finer-information setting ("St-f" and "BI-f"), which enable subjects to learn more about their opponents' strategies in both static and dynamic games, serve as a robustness check.[20]

Moreover, examining the benchmark treatment with finer information ("BI-f") allows for a fair comparison across different irreversibility structures. The alternative irreversibility structures are theoretically studied and experimentally tested based on the finer-information setting.[21]

Alternative Irreversibility Treatments ("NI-f" and "AI-f")

We also tested the two alternative irreversibility structures discussed in Section 2. We first considered the "NI-f" (neither action being irreversible, finer information) treatment, in which both the choices of A and B at $t = 0$ were reversible. At $t = 0$, subjects chose between A and B. There was no wait option in the first period. Then, at $t = 1$, upon observing the distribution of the choices from $t = 0$, they could freely switch to the other choice at no cost. Under this dynamic setting, a player could still express their intention to play A or B, but in a non-binding way. Thus, testing the "NI-f" protocol allowed us to compare the efficacy of the main treatment with the non-binding dynamic structure.[22]

In the "AI-f" (A-irreversible, finer information) treatment, only the A choice was binding at $t = 0$. In $t = 0$, subjects chose between A and the wait option. Then, in $t = 1$, those who chose the wait option could decide between A and B after observing the number of A choices at $t = 0$. This delay structure allowed a player to credibly signal their intention to choose the efficient action A. However,

[19] For example, in treatment "BI-b," subjects may have an incentive to explore others' strategy adopted in the earlier rounds by waiting, while that cannot happen in the "St-b" treatment. According to the survey after the experiment, under binary information, some subjects also played strategically to receive a bit extra information in the dynamic game treatment. To see this, consider the case in which a subject who waited learned that someone chose B at $t = 0$, and then that subject also chose B at $t = 1$. The subject could then infer that at least two subjects had B as their final choices in the group. However, in the static game with limited feedback, if a subject chose B, there was no way for them to find out whether someone else also chose B in that round. Since the information about the history of play in the earlier rounds might have affected efficient coordination, the finer-information setting ensured that no extra information was revealed in the dynamic treatment.

[20] In addition, there is a minor concern that relates to the framing effect. In the "BI-b" treatment, subjects may have felt tempted to choose differently for the $m = 0$ and $m = 1$ messages, thereby inducing more choices of WBA and WAB than of WBB and WAA. The finer-information treatments with four contingencies should have helped mitigate the framing effect.

[21] For alternative irreversibility structures, it is reasonable to focus our analysis on the finer-information setting. For example, when both actions are reversible, it is natural to allow subjects to observe the number of A and B choices at $t = 0$, as in Blume and Ortmann (2007).

[22] It is worth noting that an alternative design of the "neither irreversible" treatment is to allow the agents to choose between B and "wait" at $t = 0$ (instead of B and A). Then, after observing the "message" of $t = 0$ choices, each player chooses between A and B at $t = 1$. Theoretically, this design is equivalent to the "NI-f" treatment used in the experiment, except for the labeling of first-period actions. Therefore, the alternative design does not differ from our "NI-f" treatment if the face value of messages do not matter. By contrast, if the face value of the message matters, we expect the effect of first-period communication in the alternative setting to be weaker since the players, now, lack the message to express their intention to choose A.

as discussed in Section 2.2, there is no unique prediction of efficient coordination by SPNE or weak dominance for a four-person group ($N = 4$). The "AI-f" treatment further helps us understand the predictive power of iterated weak dominance and how different irreversible structures affect efficient coordination.

3.3 Experimental Procedure

The experiment was implemented by a web-based program in the Smith Lab at Shanghai Jiao Tong University in 2019. A total of 356 undergraduate and graduate students participated in 18 sessions. At the beginning of each session, each subject arriving at the lab was randomly assigned a seat number. The program then randomly put them into groups of four that were fixed throughout the sessions.

We adopted a between-subject design. In each session, subjects played the game from one treatment for 15 rounds with their groupmates. The choices were labeled "1" and "2" instead of "A" and "B." There was no time limit for making the choices.

In the static treatment, subjects simply submitted their choices of "1" or "2" in each round. In the dynamic treatments, subjects' complete strategies were elicited using the strategy method. For example, on the choice page of our main treatment ("BI-b"), subjects were first asked to choose between "1" and "Wait." If their choice was "Wait," then two additional choices would appear, asking them to choose an action for each of the two possible realizations of the message, $m = 0, 1$. Subjects were made aware that only one of the choices would be realized, based on the outcome in the first period. Instead, if any subject's first-period choice was "1," then there would be a notice telling them that they did not need to make any choice for the second period. However, the subject still needed to click a "confirm" button for each possible realization of the binary message to finish this round. With these two "confirm" buttons, the total number of clicks would be the same whether a subject chose to wait or not to wait at $t = 0$. Thus, subjects would not be able to infer others' choices from the number of clicks.

In the finer-information treatments, after choosing "Wait" (or either of the two actions in the "NI-f" treatment), the four possible outcomes from the first period would appear, and the subject needed to choose an action for each of the four contingencies.[23] If a subject chose not to wait, then they needed to click on the four "confirm" buttons.

At the beginning of the experiment, the instructions were first read aloud in the lab. Then, the subjects completed a short comprehension test before the 15-round play of the experiment. After all participants finished the experiment, we gave them unincentivized and anonymous questionnaires about their decision rules. Participants had not been informed about the questionnaires beforehand.

[23] In all treatments with finer information, there were $N = 4$ contingencies that could arrive at the end of $t = 0$ - specifically, all four possible numbers $(0, 1, 2, 3)$ of the irreversible choices made by the other three group members in "BI-f" and "AI-f" treatments. The same held true for the number of B choices in the "NI-f" treatment.

At the end of each session, subjects were paid based on their cumulative payoffs from all rounds. Each session took about 45 minutes, and the average earnings were 55 RMB (or 8.5 USD), including a participation fee of 5 RMB. The numbers of subjects in each session and treatment are summarized in Table 3.

Treatment	# Sessions	# 4-player Groups
"St-b" (static, binary info)	5	21
"BI-b" (B irreversible, binary info)	5	21
"St-f" (static, finer info)	2	11
"BI-f" (B irreversible, finer info)	2	12
"NI-f" (neither irreversible, finer info)	2	12
"AI-f" (A irreversible, finer info)	2	12

Table 3: Experimental Design

4 Experimental Results

4.1 The Effectiveness of the B-irreversible Structure

We first compare the efficiency rate, defined as the percentage of groups that achieved efficient coordination, of the static treatment and the main dynamic treatment, followed by a decomposition of the strategies adopted in the dynamic games.

Result 1 (Group-level Efficiency) *The efficiency rates are significantly higher in "BI-b" than in "St-b," and the higher efficiency rates in "BI-b" can be sustained over time.*

The left panel of Figure 1 plots the frequencies of efficient outcome in the "St-b" treatment and the main treatment, "BI-b." Efficient coordination was hard to achieve in the static games, which replicates the findings in the minimum-effort game literature. Only 14 percent of the groups managed to coordinate on the efficient A choice. This number stayed constant throughout the 15 rounds, which implies that if a group starts with the inferior outcome, efficient coordination in the later rounds is impossible.

In sharp contrast, the average efficiency rate over the 15 rounds was significantly higher – over 60 percent – in our main treatment, "BI-b." It started at 52 percent, and then gradually increased to 67 percent in the later rounds. This high rate of efficient coordination was sustained, except for a drop (to 52 percent) in the last round. Statistical tests comparing the group efficiency rates and individual earnings between the two treatments can be found in Appendix B.

It should be noted that the distinction of efficiency rates is present in the first round, absent of any learning effect. Although the fixed group protocol could induce concerns for reputation, if such concerns are similar in the static and dynamic games, we can still conclude that the sharp increase in the efficiency rate is caused by our dynamic structure with the delay option.

(a) Group: Efficiency Rate (b) Individual: Proportion of Realized A

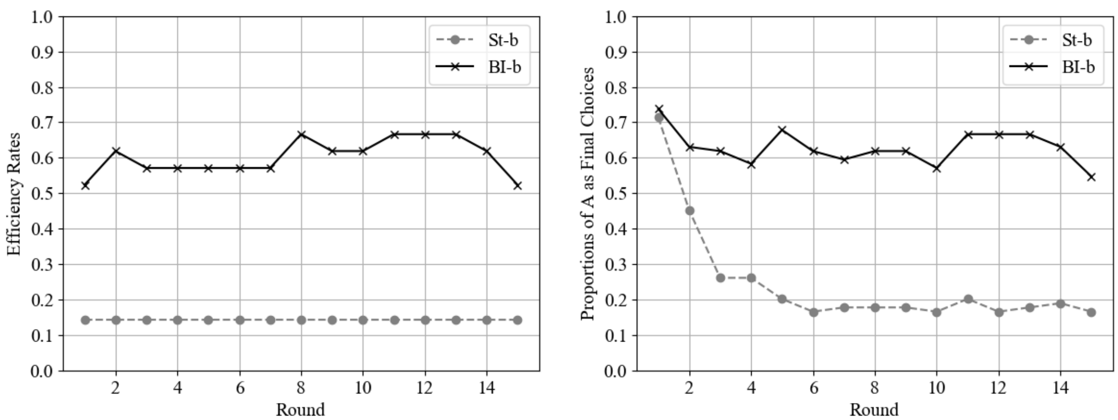

Figure 1: Group Efficiency and Individual Choices: BI-b v.s. St-b Treatments

Note: The left panel presents the percentage of groups that achieved efficient coordination in each round of play in both the "St-b" and "BI-b" treatments. The right panel presents the proportion of subjects who had A as the realized choice in each round of play in both the "St-b" and "BI-b" treatments.

Realized Choices and Individual Payoffs We examine individuals' final realized choices between the two treatments in the right panel of Figure 1. In "St-b," although over 70 percent of the subjects showed their willingness to cooperate and started with A in the first period, the proportion of A choices dropped rapidly, to below 20 percent within six rounds, and it never recovered. The dramatic drop was caused by the high miscoordination rates in the static treatment, as depicted in Figure 2. A large proportion of subjects who initially chose A experienced coordination failure due to at least one B choice in the group. As a result, they switched to B expeditiously in three or four rounds.

In our main treatment ("BI-b"), the proportion of subjects whose realized choice was A started at 73 percent, very close to that in "St-B." However, unlike in the static treatment, the proportion of realized A choices in the "BI-b" treatment fluctuated over time, without a distinct decline. That explains why the efficiency rate in the "BI-b" treatment maintained a high level.

The higher incidence of efficient coordination in the dynamic setting led to significantly higher individual payoffs. The average payoff in the static games was 42.4 per round, while it was 49.6 per round in the dynamic games. Given the fact that the highest possible payoff (when achieving efficient coordination) was 55, and the B choice secured a payoff of 45, the dynamic structure with the irreversible B choice significantly recovered the efficiency loss in the static games. T-tests show that the differences in the group total payoffs between the two treatments were statistically significant at the five-percent level in all but the 4th, 5th, and 15th rounds (see Appendix B).

Miscoordination Apart from promoting the Pareto-dominant equilibrium play, we also found the delay option to significantly reduce the rate of miscoordination, or the percentage of groups in which members ended up with different realized choices.

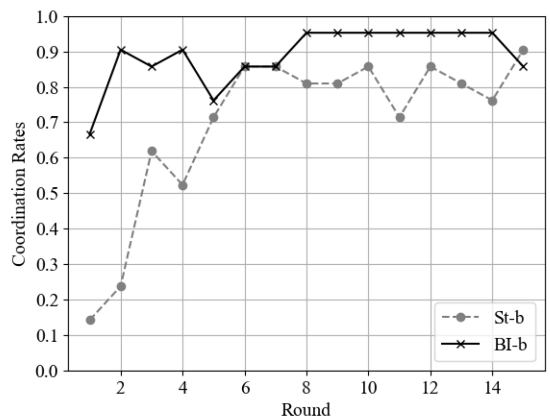

Figure 2: Group Coordination Rates

Note: The figure presents the proportion of groups in which all members had the same realization choice in each round of play in the "St-b" and "BI-b" treatments. The coordination rate is equal to 1 minus the rate of miscoordination.

Figure 2 shows, interestingly, that although the proportions of realized A choices were very similar in the two treatments in the first round, the mis-coordinate rate surpassed 85 percent in the static treatment, while it was about 35 percent when the delay option was available. Such a significant gap in the coordination rates also existed in the later rounds, despite the fact that many of the groups in the static treatment gradually converged to the inferior equilibrium.

Result 2 (Adoption of Strategies) *In "BI-b," more than 70 percent of the subjects took the unique iteratedly undominated strategy WBA.*

The strategy method allowed us to examine a decomposition of the strategies adopted in the "BI-b" treatment. Apart from the group-level coordination outcome, we are interested in whether the strategy WBA, which, according to our theory, is the unique iteratedly undominated strategy, is adopted frequently in our experiment. Figure 3 plots the distribution of strategies B, WBB, WBA and the dominated strategies (WAB and WAA) adopted by subjects as the experiment went on.

Consistent with the theoretical prediction built on iterated weak dominance, most subjects adopted the strategy WBA. The proportion of WBA choices was 85 percent in the first round, suggesting a high willingness to wait and cooperate if all group members made the same positive gesture.[24] However, if any group member adopted the strategy WBB, subjects who chose WBA would end up with a payoff of 5 instead of a secured payoff of 45 under strategies B or WBB. In our experiments, subjects could identify whether a WBB choice in the group was the (pure) reason for the inferior coordination outcome via the feedback after each round. Therefore, the proportion of WBA dropped slightly over time and ended up being 75 percent in the last round.[25]

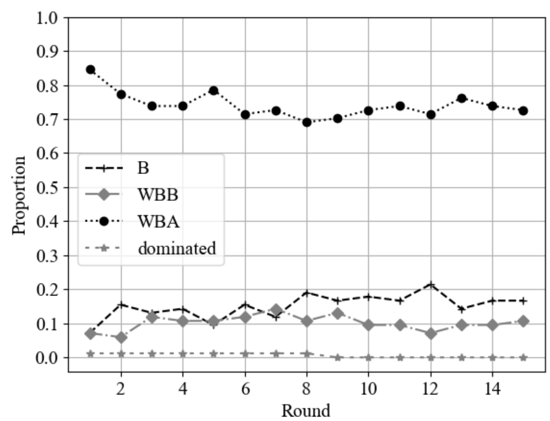

Figure 3: Decomposition of Strategies in the "BI-b" Treatment

Note: The figure presents the frequency of different strategies chosen by subjects in each round of play in the "BI-b" treatment. "Dominated strategy" denotes strategies WAB and WAA, which are strictly dominated strategies in the first round of elimination.

Aside from those who chose WBA, 15 percent of the subjects chose B and 10 percent chose WBB, on average, over time. In particular, there was an increase in the proportion of the subjects who chose B, from 7 percent in the first round to over 15 percent in the later rounds, possibly due to the inferior outcome in the previous rounds.[26]

In the online appendix, we report subjects' choice dynamics in the "BI-b" treatment. When a group failed to achieve the efficient outcome in the previous round, the subjects who initially chose

[24]There is a gap between the proportion of the WBA strategy (85% in the first round) and the proportion of having A as the realized action (73% in the first round), which is due to the presence of some B choices in the first period. In this case, the realized action for the WBA strategy would be B.

[25]An analysis of individual choices (see part C in the online appendix for details) and the questionnaires reveal that some participants who started with WBA switched to B or WBB because they were hurt by their groupmates who chose WBB.

[26]It is not surprising to have a small fraction of subjects choose strategies B and WBB, as this can be understood through the lens of our model. Choosing B over WBB might suggest that the subject had a social preference to not hurt others, but did not believe in the social preferences of their group mates. For the 10 percent of subjects choosing WBB, one plausible explanation would be selfish or spiteful social preferences.

WBA were much more likely to switch to B than to WBB, which suggests that, overall, WBA choosers preferred B to WBB. More specifically, if some B choices led to the inferior outcome, 10.3% of the WBA choosers switched to B, and no one switched to WBB. In contrast, when nobody in the group chose B in the previous round, and coordination failure was caused purely by some "harmful" WBB choices, 29% of the WBA choosers switched to B, and 6% switched to WBB (possibly an act of retaliation). This evidence is consistent with players having other-regarding social preferences.

We collected subjects' self-reported motives for choosing a certain strategy from the survey after the experiment. The survey was unincentivized and anonymous and served only as a piece of anecdotal evidence. The results from the subjects who self-reported having chosen WBB showed that most of them, indeed, had negative social preferences, but that some of them were also confused and failed to realize that adopting WBB could have hurt others (compared with strategy B).[27]

4.2 Finer Information

Result 3 (Group-level Efficiency) *The efficiency rates were significantly higher in the dynamic treatment with finer information (i.e., "BI-f") than those in the static treatment with finer information (i.e., "St-f"), and the higher efficiency rates in "BI-f" could be sustained over time.*

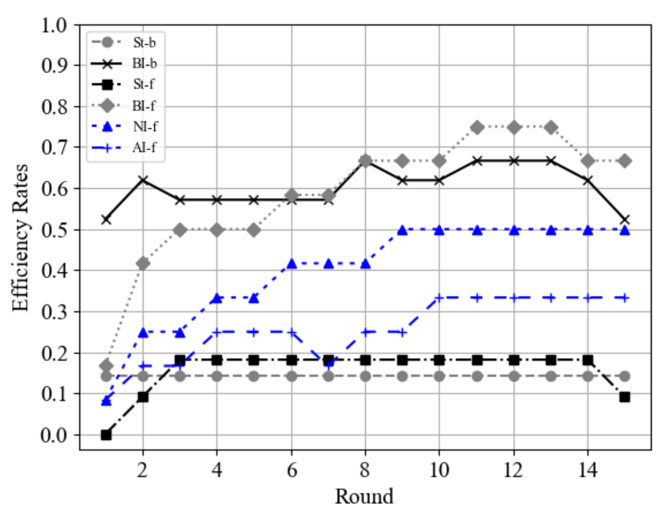

Figure 4: Efficiency Rates: All Treatments

Figure 4 presents the efficiency rates in all treatments in the study. Chi-squared tests of the pairwise comparisons of group efficiency rates are presented in Appendix B. We first discuss the

[27]Some participants explicitly mentioned that they chose WBB to hurt other subjects or to retaliate against the teammate who chose WBB, which suggests negative reciprocity and the opposite of social preference.

results from the finer-information versions of the main treatments - namely, "St-f" and "BI-f". Overall, we still observe a significantly large difference between the group-level efficiencies in the static and dynamic treatments with finer-information, which serves as a robustness check of the main treatment.

A higher proportion of subjects started with B choices in the first round of the finer-information treatment "BI-f" than that of "BI-b."[28] One explanation could be that some subjects were confused by the four contingencies and, therefore, chose the most conservative strategies at the beginning of the experiment. In spite of that, the rate of efficient coordination rapidly converged somewhere above the binary information treatments, though the differences were, in general, not significant.

Result 4 (Adoption of Strategies) *Over 60 percent of the subjects adopted the strategy equivalent to WBA in the "BI-f" treatment.*

We next investigate the strategies chosen in the "BI-f" treatment. As depicted in Figure 5, in the earlier rounds, about 60 to 70 percent of the subjects chose WBA.[29] This proportion is lower than that in "BI-b" (around 75 percent), but it rapidly caught up in the later rounds. This finding demonstrates that our main experimental finding is robust to the introduction of finer information, which is consistent with Proposition 3. It also indicates that the high adoption rate of the WBA strategy was unlikely to have been driven by the framing effect of the binary information.

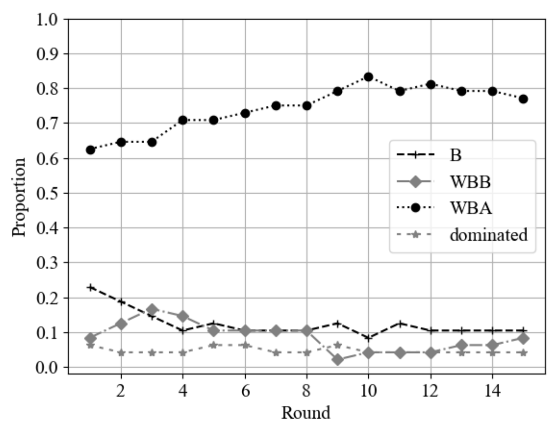

Figure 5: Decomposition of Strategies in the "BI-f" Treatment

[28]The pattern of the increasing efficiency in the first couple of rounds is also present in Avoyan and Ramos (2019) and Brandts and Cooper (2006), in which finer information is available. In the treatments with finer information, subjects could observe the exact number of A and B choices in the previous rounds, which might have facilitated coordination. For example, if a subject realized that they are the only one playing B in one round, they might have wanted to switch from B to A in the next round.

[29]Recall that we refer to the following strategy as WBA in the "BI-f" treatment: "wait in the first period; choose A if $n = 0$, and choose B if $n \geq 1$," where n is the number of B choices in $t = 0$. Any strategy involving choosing A after observing $n \geq 1$ is called a dominated strategy.

Apart from the WBA strategy, about 10 percent of the subjects used some dominated strategies at least once in the first 5 rounds, whereas this proportion was merely 1 percent with binary information. In addition, they did not learn to avoid using these dominated strategies until the end of the experiment.[30]

4.3 Alternative Irreversibility Structures

Result 5 (Group-level Efficiency) *The efficiency rates in both the "NI-f" and "AI-f" treatments were significantly higher than those in "St-f" but significantly lower than those in "BI-f."*

In Figure 4, we also plot the efficiency rates in the games in which neither action was irreversible ("NI-f") or only A was irreversible ("AI-f").

"NI-f" Treatment In the "NI-f" treatment, the efficiency rate was also raised significantly, compared to that of the static game. The opportunity to communicate one's intentions using the non-binding first-period actions did help promote an coordination outcome. This is consistent with the findings in the literature on multi-sided costless pre-play communication in common-interest coordination games (Blume and Ortmann, 2007; Cooper et al., 1992*a*). Nevertheless, the efficiency rates in the "NI-f" treatment were significantly lower than those in "BI-f." Chi-squared tests confirm that, compared with "NI-f," the "BI-f" protocol generated significantly higher efficiency rates in all rounds of play.[31] On average, "BI-f" raised the efficiency rates by 18 percentage points, which is a 40-percent increase compared with "NI-f."

Next, we make a detailed comparison between the two mechanisms: signaling intention by waiting and expressing intention via costless pre-play communication. Based on our theory, the choice of "wait" signals the intention to play A at $t = 1$ (if all others choose to wait). From this perspective, the wait choice in "BI-f" or "BI-b" is comparable to the non-binding A choice in the "NI-f" treatment.

We first examine the trustworthiness of the two comparable messages - that is, the "no-B" message (i.e., all others chose to wait) in "BI-b" and "BI-f" and the "all-A" message (i.e., all others chose the non-binding action A at $t = 0$) in "NI-f." As the right panel of Figure 6 shows, among the subjects who chose to wait in "BI-b" and "BI-f," the proportion of A choices following the "no-B" message is comparable to the A choice following the "all-A" message among subjects who chose A at $t = 0$ in "NI-f."

[30] The reason for adopting dominated strategies might be that the strategy method, which requires subjects to decide from a full menu of all four possible contingencies after waiting at $t = 0$, is somewhat challenging in its implementation with finer information.

[31] Given that the subjects could observe all information about the $t = 0$ choices, we believe that "BI-f" is a better comparison for "NI-f" than "BI-b." But if we compare "NI-f" with "BI-b" (our main treatment), the differences in the group efficiency rates are still significant in all but the last round. See Appendix B for details.

(a) $t = 0$ choices: Wait or A

(b) H choice after "no-B" and "all-A"

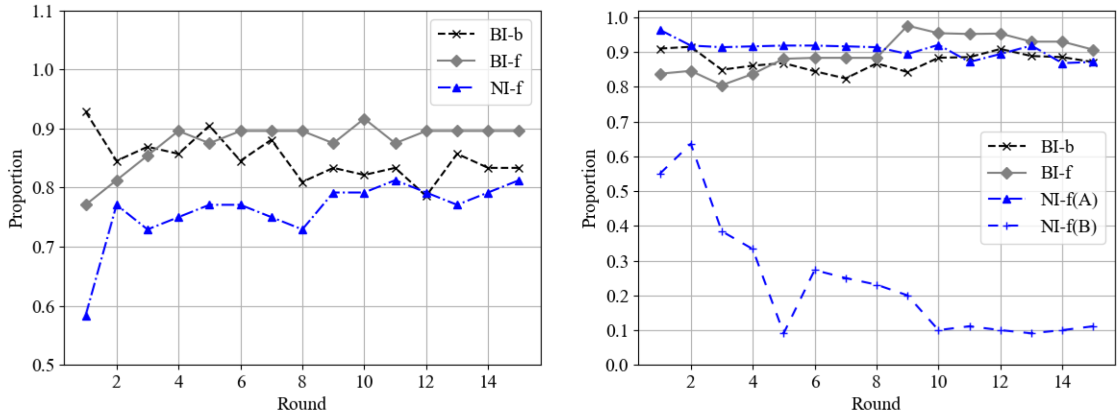

Figure 6: Comparison of the "no-B" and "all-A" messages

Note: The panel on the left shows the frequency of the "wait" choice at $t = 0$ in the "BI-b" and "BI-f" treatments, and the frequency of the "A" choice at $t = 0$ in the "NI-f" treatment. The panel on the right shows, among the subjects who waited, the proportion of those who chose A after the "no-B" message (in the "BI-b" and "BI-f" treatments), and, among the subjects who chose A or B in $t = 0$, the proportion of those who chose A after the "all-A" message (in the "NI-f" treatment).

Recall that the "no-B" message simply means that no one has taken the binding choice B, or "nothing has happened yet." In fact, all messages in the "NI" setting have that meaning since all actions are reversible. But the "all-A" message, by its face value, says that "nothing has happened yet, but all subjects intend to choose A." If the face value of messages can, indeed, affect players' beliefs and their subsequent moves (not in a strategic sense but in a linguistic sense), it is striking that the "no-B" message could be as effective as the "all-A" message. The trustworthiness of the "no-B" message is implied by our theory. In fact, a subject would trust this message if they believe that other subjects who chose to wait would also have trust in it.

Given that the effectiveness of "no-B" (for subjects who chose "wait") and that of "all-A" (for subjects who chose the non-binding A) were comparable in $t = 1$, efficient rates should have been similar in these two irreversibility structures, if a similar number of subjects chose to wait or chose A at $t = 0$. This is, however, not the case, and the lower efficiency rates were observed in "NI-f" for two reasons. First, in the "NI-f" treatment, not only A, but also the non-binding choice B, was allowed at $t = 0$. As shown in the right panel of Figure 6, among those who chose B in $t = 0$, most would continue to take B at $t = 1$, even after the "all-A" message. Taking them into consideration, the "all-A" message is no longer as effective as the "no-B" message. Second, as can be seen from the left panel of Figure 6, the frequency of "wait" choices in "BI-b" and "BI-f" is much higher than the "A" choices in "NI-f" at $t = 0$, which means that the "no-B" message was generated more frequently. This is, again, consistent with the prediction of iterated weak dominance.

"AI-f" Treatment When A is the only binding choice at $t = 1$, as discussed in Section 2.2, neither SPNE nor weak dominance can provide a clear prediction about efficient coordination for any multiple-player group ($N \geq 3$). Although choosing A early may serve as a signal to induce the players who have waited to follow, it is, indeed, a very risky choice because all the subjects who waited still faced a coordination problem at $t = 1$.

The experimental evidence shows that the "AI-f" protocol could promote efficient coordination (compared with the static benchmark with finer information), but the efficiency rates were, on average, 33 percentage points lower than those in the main treatment. On average, only 15 out of the 48 subjects chose A in $t = 0$ in each round. Out of the subjects who chose to wait, only 80 percent would choose A if all other groupmates had already committed to A, and only 33 percent would choose A if two of the other three groupmates had committed to A.

4.4 Summary

Overall, the experimental evidence confirms our theoretical prediction that efficiency is promoted in the dynamic in which the inefficient action B is the only binding choice in the first period. The evidence is also consistent with the mechanism of signaling intention by waiting, as over 90 percent of the subjects who chose to wait at $t = 0$ were willing to take the risky but efficient choice A after the "no-B" message. In fact, the majority of the subjects adopted the unique iteratedly undominated strategy "WBA." The additional experimental tests demonstrate that the mechanism of signaling intention via delay is different from sending non-binding messages to signal intention, and that it is more effective in promoting efficient coordination.

5 Conclusion

This paper highlights a distinctive function of a delay option in strategic interactions that can be modeled as coordination. The option enables forward-induction reasoning to operate, and, in this way, each player, by delaying their choice, can signal their intention to take the risky and efficient action. We show that this mechanism of signaling intentions via delay can work to achieve the efficient outcome. This idea is formalized via iterated weak dominance, when players have ϵ-social preferences. We also provide experimental evidence to support our theoretical analysis regarding use of the strategy that survives iterated weak dominance and the resulting coordination outcome. We further demonstrate the advantage of our mechanism in promoting efficient coordination over pre-play non-binding communication.

The unique strategy surviving iterated weak dominance – waiting and then taking the efficient action if and only if none of the other players took the inefficient action earlier – can be interpreted as "no first use (of the inefficient action)." Obviously, if everyone commits to such a strategy, that can lead to the efficient outcome, and this way of achieving efficiency becomes possible only if each player is granted the option of delay. We believe that this simple idea should be applicable to more complex coordination settings,[32] as well as to other strategic interactions beyond pure coordination. We leave this to future work.

References

Aumann, Robert. 1990. "Nash equilibria are not self-enforcing." *Economic decision making: Games, econometrics and optimisation*, 201–206.

Avoyan, Ala, and Joao Ramos. 2019. "A Road to Efficiency Through Communication and Commitment." *Available at SSRN 2777644*.

Baliga, Sandeep, and Stephen Morris. 2002. "Co-ordination, spillovers, and cheap talk." *Journal of Economic Theory*, 105(2): 450–468.

Basak, Deepal, and Zhen Zhou. 2020. "Timely Persuasion." *Working Paper*.

Ben-Porath, Elchanan, and Eddie Dekel. 1992. "Signaling future actions and the potential for sacrifice." *Journal of Economic Theory*, 57(1): 36–51.

Blume, Andreas, and Andreas Ortmann. 2007. "The effects of costless pre-play communication: Experimental evidence from games with Pareto-ranked equilibria." *Journal of Economic theory*, 132(1): 274–290.

Blume, Andreas, Peter H Kriss, and Roberto A Weber. 2017. "Pre-play communication with forgone costly messages: experimental evidence on forward induction." *Experimental Economics*, 20(2): 368–395.

Brandenburger, Adam, Amanda Friedenberg, and H Jerome Keisler. 2008. "Admissibility in Games." *Econometrica*, 76(2): 307–352.

Brandts, Jordi, and Charles A Holt. 1995. "Limitations of dominance and forward induction: Experimental evidence." *Economics Letters*, 49(4): 391–395.

Brandts, Jordi, and David J Cooper. 2006. "A change would do you good.... An experimental study on how to overcome coordination failure in organizations." *American Economic Review*, 96(3): 669–693.

[32]For example, Basak and Zhou (2020) apply a similar idea to design information disclosure policy in a dynamic regime change game with irreversible attacks under an incomplete information environment.

Brandts, Jordi, Antonio Cabrales, and Gary Charness. 2007. "Forward induction and entry deterrence: an experiment." *Economic Theory*, 33(1): 183–209.

Bryant, John. 1983. "A simple rational expectations Keynes-type model." *The Quarterly Journal of Economics*, 98(3): 525–528.

Cachon, Gerard P, and Colin F Camerer. 1996. "Loss-avoidance and forward induction in experimental coordination games." *The Quarterly Journal of Economics*, 111(1): 165–194.

Calcagno, Riccardo, Yuichiro Kamada, Stefano Lovo, and Takuo Sugaya. 2014. "Asynchronicity and coordination in common and opposing interest games." *Theoretical Economics*, 9(2): 409–434.

Carlsson, Hans, and Eric van Damme. 1993. "Global Games and Equilibrium Selection." *Econometrica*, 61(5): 989–1018.

Chamley, Christophe, and Douglas Gale. 1994. "Information revelation and strategic delay in a model of investment." *Econometrica*, 62(5): 1065–1085.

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

Chen, Roy, and Yan Chen. 2011. "The potential of social identity for equilibrium selection." *American Economic Review*, 101(6): 2562–89.

Cooper, Russell, Douglas V De Jong, Robert Forsythe, and Thomas W Ross. 1992*a*. "Communication in Coordination Games." *Quarterly Journal of Economics*, 107: 739–771.

Cooper, Russell, Douglas V De Jong, Robert Forsythe, and Thomas W Ross. 1992*b*. "Forward induction in coordination games." *Economics Letters*, 40(2): 167–172.

Cooper, Russell, Douglas V DeJong, Robert Forsythe, and Thomas W Ross. 1993. "Forward induction in the battle-of-the-sexes games." *The American Economic Review*, 1303–1316.

Corsetti, Giancarlo, Bernardo Guimaraes, and Nouriel Roubini. 2006. "International lending of last resort and moral hazard: A model of IMF's catalytic finance." *Journal of Monetary Economics*, 53(3): 441–471.

Devetag, Giovanna, and Andreas Ortmann. 2007. "When and why? A critical survey on coordination failure in the laboratory." *Experimental economics*, 10(3): 331–344.

Diamond, Peter A. 1982. "Aggregate demand management in search equilibrium." *Journal of political Economy*, 90(5): 881–894.

Dutta, Prajit K. 2012. "Coordination need not be a problem." *Games and Economic Behavior*, 76(2): 519–534.

Farrell, Joseph. 1988. "Communication, coordination and Nash equilibrium." *Economic Letters*, , (27): 209–214.

Farrell, Joseph, and Garth Saloner. 1985. "Standardization, compatibility, and innovation." *the RAND Journal of Economics*, 70–83.

Fehr, Ernst, and Klaus M Schmidt. 2006. "The economics of fairness, reciprocity and altruism–experimental evidence and new theories." *Handbook of the economics of giving, altruism and reciprocity*, 1: 615–691.

Govindan, Srihari, and Robert Wilson. 2009. "On forward induction." *Econometrica*, 77(1): 1–28.

Gul, Faruk, and Russell Lundholm. 1995. "Endogenous timing and the clustering of agents' decisions." *Journal of Political Economy*, 103(5): 1039–1066.

Harsanyi, John C, and Reinhard Selten. 1988. "A general theory of equilibrium selection in games." *MIT Press Books*, 1.

Huck, Steffen, and Wieland Müller. 2005. "Burning money and (pseudo) first-mover advantages: an experimental study on forward induction." *Games and Economic Behavior*, 51(1): 109–127.

Katz, Michael L, and Carl Shapiro. 1986. "Technology adoption in the presence of network externalities." *Journal of political economy*, 94(4): 822–841.

Kohlberg, Elon, and Jean-Francois Mertens. 1986. "On the Strategic Stability of Equilibria." *Econometrica*, 54(5): 1003–1037.

Lagunoff, Roger, and Akihiko Matsui. 1997. "Asynchronous choice in repeated coordination games." *Econometrica*, 65(6): 1467–1477.

Lo, Melody. 2020. "Language and coordination games." *Economic Theory*, 1–44.

Morris, Stephen, and Hyun Song Shin. 1998. "Unique equilibrium in a model of self-fulfilling currency attacks." *American Economic Review*, 587–597.

Sobel, Joel. 2017. "A note on pre-play communication." *Games and Economic Behavior*, 102: 477–486.

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

Van Huyck, John B, Raymond C Battalio, and Richard O Beil. 1993. "Asset markets as an equilibrium selection mechanism: Coordination failure, game form auctions, and tacit communication." *Games and Economic Behavior*, 5(3): 485–504.

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

Appendix A Proofs

Proof of Proposition 1 Consider player $i \in \mathcal{N}$. When all other players take $d_j = B$, their monetary payoffs will be $\pi_j = b$ (for all $j \in \mathcal{N} \setminus \{i\}$), which is independent of player i's choice. For player i, choosing B yields $\pi_i = b$, while choosing A yields $\pi_i = c$. Since $b > c$, $u_i(d_i = B, (d_j = B)_{j \in \mathcal{N} \setminus \{i\}}) > u_i(d_i = A, (d_j = B)_{j \in \mathcal{N} \setminus \{i\}})$, and, thus, $\{d_i = B\}_{i=1}^N$ is a Nash equilibrium.

Similarly, when all others are taking $d_j = A$, choosing $d_i = A$ yields the same monetary payoff a for player i and all other players, while choosing $d_i = B$ yields $\pi_i = b$ for player i and $\pi_j = c$ for all $j \in \mathcal{N} \setminus \{i\}$. Hence, $u_i(d_i = A, (d_j = A)_{j \in \mathcal{N} \setminus \{i\}}) > u_i(d_i = B, (e_j = A)_{j \in \mathcal{N} \setminus \{i\}})$, and, thus, $\{d_i = A\}_{i=1}^N$ is a Nash equilibrium. □

Proof of Proposition 2 First, consider the case in which all other players take $s_j = B$. In this case, $m = 1$, and, $\pi_j = b$ for all $j \in \mathcal{N} \setminus \{i\}$ independent of s_i. For player i, the private payoff from choosing $s_i = B$ is b, while deviating to other strategies never strictly increases this private payoff but possibly strictly decreases it to c (for example, deviating to WAB). Hence, $(s_i = B)_{i=1}^N$ is a Nash equilibrium.

Next, consider the case in which all other players choose WAA. In this case, WAA is a best response because (1) choosing WAA yields $\pi_i = a$ and $\pi_j = a$ for all $j \in \mathcal{N} \setminus \{i\}$; (2) deviating to B yields $\pi_i = b < a$ (and $\pi_j = c < a$); (3) deviating to WBB or WAB yields $\pi_i = b < a$ (and $\pi_j = c < a$); and (4) deviating to WBA yields the same π_i and π_j as choosing WAA. Hence, $(s_i = WAA)_{i=1}^N$ is a Nash equilibrium. Following similar arguments, we can show that $(s_i = WBA)_{i=1}^N$ is a Nash equilibrium.

Let us further consider the case in which all other players choose WAB. Given that, WAB is a best response because (1) choosing WAB yields $\pi_i = b$ and $\pi_j = b$ for all $j \in \mathcal{N} \setminus \{i\}$; (2) deviating to B yields $\pi_i = b$ and $\pi_j = c < b$; (3) deviating to WAA or WBA yields $\pi_i = c < b$ (and $\pi_j = b$); and (4) deviating to WBB yields the same π_i and π_j as choosing WAB. Hence, $(s_i = WAB)_{i=1}^N$ is a Nash equilibrium. Following similar arguments, we can show that $(s_i = WBB)_{i=1}^N$ is a Nash equilibrium.

Lastly, choosing A on the information set $m = 1$ is not subgame perfect. That is because, in the subgame starting at $t = 1$ following $m = 1$ – i.e., after someone has already taken B at $t = 0$ deviating from A to B increases one's payoff from c to b (without changing others' payoffs). □

Proof of Theorem 1

First round of elimination Consider any player i and any strategy profile $s_{-i} = (s_j)_{j \in \mathcal{N} \setminus \{i\}}$. We want to show that WAB (WAA) is weakly dominated by WBB (WBA). Consider two mutually exclusive and collectively exhaustive cases. In the first case, the other players adopt the strategy profile $s_{-i} = (s_j)_{j\in\mathcal{N}\setminus\{i\}}$, which satisfies $\left|\{j \in \mathcal{N}\setminus\{i\}|s_j = B\}\right| \geq 1$; that is, some other players choose B at $t = 0$ (or $m = 1$, regardless of s_i). Given s_{-i}, $\pi_i(s_i = WBB, s_{-i}) = b > \pi_i(s_i = WAB, s_{-i}) = c$, whereas $\pi_j(s_i = WBB, s_{-i}) = \pi_j(s_i = WAB, s_{-i})$ for any s_j and any $j \in \mathcal{N} \setminus \{i\}$. In the other case, $s_{-i} = (s_j)_{j\in\mathcal{N}\setminus\{i\}}$ satisfies $\left|\{j \in \mathcal{N} \setminus \{i\}|s_j = B\}\right| = 0$, meaning that $m = 0$, regardless of s_i. Given this s_{-i}, $\pi_i(s_i = WBB, s_{-i}) = \pi_i(s_i = WAB, s_{-i}) = b$ and $\pi_j(s_i = WBB, s_{-i}) = \pi_j(s_i = WAB, s_{-i})$ for any s_j and any $j \in \mathcal{N} \setminus \{i\}$. Hence, based on the utility function u_i defined in (1), WAB is weakly dominated by WBB. The same argument can be applied to show that WAA is weakly dominated by WBA.

To see why WBA, WBB and B cannot be eliminated in this round, consider any mixed strategy that might dominate any of these three strategies. If such a mixed strategy exists, and if it assigns positive probabilities to WAA and/or WAB, then we can reassign those probabilities to WBA and/or WBB, respectively, to generate another strategy that still satisfies the weak dominance relationship (since WAA and WAB are weakly dominated by WBA and WBB, respectively). So, we need only consider the mixtures of WBA, WBB, and B.

First, note that any mixed strategy consisting of B and WBB cannot dominate WBA because WBA is the best response to $s_{-i} = (s_j = WBA)_{j\in\mathcal{N}\setminus\{i\}}$.

Now, suppose that B can be weakly dominated by $s^0 = p_0 \cdot WBB \oplus (1 - p_0) \cdot WBA$ for some $p_0 \in [0, 1]$. Consider the case in which the other players' strategy profile is $s_{-i} = (s_j = WBB)_{j\in\mathcal{N}\setminus\{i\}}$. Then, $\pi_i(s_i = B, s_{-i}) = b$, $\pi_i(s_i = s^0, s_{-i}) = p_0 b+(1-p_0)c$, while $\pi_j(s_i = B, s_{-i}) = \pi_j(s_i = s^0, s_{-i}) = b$. So, weak dominance requires $p_0 = 1$, which means that B can be dominated only by the pure strategy WBB. Next, fix $p_0 = 1$ in s^0 and consider the other case in which $s'_{-i} = (s_j = WBA)_{j\in\mathcal{N}\setminus\{i\}}$. Then, $\pi_i(s_i = B, s'_{-i}) = \pi_i(s_i = s^0, s'_{-i}) = b$, while $\pi_j(s_i = B, s'_{-i}) = b > \pi_j(s_i = s^0, s'_{-i}) = c$, which means that B is preferred to WBB in this case. Therefore, no such $p_0 \in [0, 1]$ exists, and B cannot be weakly dominated by any possible mixed strategy.

To see that WBB cannot be weakly dominated either, suppose that a mixed strategy $s^1 = p_1 \cdot B \oplus (1 - p_1) \cdot WBA$ for some $p_1 \in [0, 1]$ weakly dominates WBB. Consider the case in which the other players' strategy profile is $s_{-i} = (s_j = WBB)_{j\in\mathcal{N}\setminus\{i\}}$. Then, $\pi_i(s_i = WBB, s_{-i}) = b$, $\pi_i(s_i = s^1, s_{-i}) = p_1 b + (1 - p_1)c$, while $\pi_j(s_i = WBB, s_{-i}) = \pi_j(s_i = s^1, s_{-i}) = b$. So, weak dominance requires $p_1 = 1$, which means that WBB could only be dominated by the pure strategy B. Next, consider the other case, in which $s'_{-i} = (s_j = WAB)_{j\in\mathcal{N}\setminus\{i\}}$. Then, $\pi_i(s_i = WBB, s'_{-i}) = \pi_i(s_i = B, s'_{-i}) = b$, while $\pi_j(s_i = WBB, s'_{-i}) = b > \pi_j(s_i = B, s'_{-i}) = c$, which means that WBB is preferred to B in this case. Therefore, no such $p_1 \in [0, 1]$ exists, and WBB cannot be weakly dominated by any possible mixed strategy.

Second round of elimination The remaining strategies are B, WBB and WBA. For player i, given any s_{-i}, $\pi_i(s_i = WBB, s_{-i}) = \pi_i(s_i = B, s_{-i}) = b$. If $n(s_{-i}) \geq 1$, then $m = 1$ regardless of s_i, and, therefore, $\pi_j(s_i = B, s_{-i}) = \pi_j(s_i = WBB, s_{-i})$ for all $j \in \mathcal{N} \setminus \{i\}$. This means that player i is indifferent between B and WBB. The indifference also holds when $s_j = WBB$ for all j. However, if, among other players, no one chooses B and some players choose WBA – i.e., $n(s_{-i}) = 0$ and $\left|\{j \in \mathcal{N} \setminus \{i\} | s_j = WBA\}\right| \geq 1$ – then $\pi_{j'}(s_i = B, s_{-i}) = b > \pi_{j'}(s_i = WBB, s_{-i}) = c$ for all $j' \in \{j \in \mathcal{N} \setminus \{i\} | s_j = WBA\}$, and $\pi_j(s_i = B, s_{-i}) = \pi_j(s_i = WBB, s_{-i}) = b$ for all $j \in \{j \in \mathcal{N} \setminus \{i\} | s_j = WBB\}$. Hence, under the $\epsilon-$social preference assumption, WBB is weakly dominated by B.

No other strategies can be eliminated in this round. WBA is the unique best response if all others take WBA. When all others choose WBB, compared with strategy WBA, choosing B yields a strictly higher payoff to player i but the same payoffs to other players. Therefore, B cannot be dominated by WBA. Since we have already shown that B weakly dominates WBB, B cannot be eliminated in this round.

Third round of elimination The remaining strategies are B and WBA. Again, consider two mutually exclusive and collectively exhaustive cases regarding s_{-i}. First, suppose that s_{-i} satisfies $\left|\{j \in \mathcal{N} \setminus \{i\} | s_j = B\}\right| \geq 1$, which means $m = 1$, regardless of s_i. Then, player i is indifferent between B and WBA. Second, suppose that s_{-i} satisfies that $\left|\{j \in \mathcal{N} \setminus \{i\} | s_j = B\}\right| = 0$ – i.e., all other players choose WBA; then, $\pi_i(s_i = WBA, s_{-i}) = a > \pi_i(s_i = B, s_{-i}) = b$, and $\pi_j(s_i = WBA, s_{-i}) = a > \pi_j(s_i = B, s_{-i}) = b$ for all $j \in \mathcal{N} \setminus \{i\}$. Hence, B is weakly dominated by WBA. $\square$

Proof of Proposition 3 As in the proof of Theorem 1, in the first round of elimination, we can eliminate any strategies that involve waiting and taking A following any message that indicates $n(s_{-i}) \geq 1$; that is, someone else has already chosen B at $t = 0$. Then, the proofs of second-round and third-round elimination follow immediately from that of Theorem 1. $\square$

Proof of Proposition 4 For the $N = 2$ case, waiting and then taking B after observing that the other player chose A at $t = 0$ is dominated by waiting and then taking A based on this history. Given that, choosing A at $t = 0$ weakly dominates waiting and then choosing A after observing that the other player chose A, and choosing B (or A) after observing that the other player chose to wait. The symmetric subgame-perfect equilibria are (1) A at $t = 0$ and (2) wait and always choose A. It is worth mentioning that the strategy "waiting and choosing A if the other player chooses A; otherwise, choosing B" cannot constitute a symmetric equilibrium, as each player would profit from deviating to choosing A at $t = 0$.

In this proof for groups with $N \geq 3$, we consider a simple case with $N = 3$, and we will find all possible symmetric strategy profiles that are consistent with the iterated elimination of weakly dominated strategies. The result can be easily generalized to cases with $N > 3$.

Iterated Weak Dominance In this three-player case, we can write the strategy as A, $WBBB$, $WBBA$, $WBAB$, $WABB$, $WAAB$, $WABA$, $WBAA$ and $WAAA$. The strategy of taking A at $t = 0$ is denoted as A. For any strategy profile of other players s_{-i}, let $n^A(s_{-i}) := \big|j \in \mathcal{N}\backslash\{i\}|s_j = A\big|$ denote the number of the irreversible A choices in $t = 0$. Then, we denote any player i's strategy associated with waiting at $t = 0$ as follows. "W" stands for waiting at $t = 0$. The first letter after "W" is for the case in which no one chose A at $t = 0$ ($n^A = 0$), and the second (third) letter is for the choice of action when $n^A = 1$ ($n^A = 2$).

In $t = 1$, it is strictly better to choose A after observing that $n^A = 2$. Therefore, we can eliminate $WBBB$ (by $WBBA$), $WBAB$ (by $WBAA$), $WAAB$ (by $WAAA$) and $WABB$ (by $WABA$). The remaining ones are A, $WBBA$, $WABA$, $WBAA$ and $WAAA$.

We will show that none of the other strategies can be eliminated in this round. Consider the case in which one player chooses $WBBA$, and the other chooses a mixed strategy $p \cdot A \oplus (1-p)WBBA$ with $p \in (0, \frac{b-c}{a-c})$. As can be seen from the following table, in this case, $WBBA$ and $WBBB$ are the only two pure strategies that serve as the best responses because they both generate the highest (expected) payoff π_i, and they both generate the same payoff to other players (as n^A obtains a value of 0 or 1, but these two strategies differ only when $n^A = 2$.)

Strategy	Payoff π_i
A	$pa + (1-p)c$
$WBBA$ (or $WBBB$)	b
$WBAA$ (or $WBAB$)	$pc + (1-p)b$
$WABA$ (or $WABB$)	$pb + (1-p)c$
$WAAA$ (or $WAAB$)	c

Therefore, $WBBA$ can be weakly dominated only by a mixture of $WBBA$ and $WBBB$, which is not possible, since $WBBA$ weakly dominates $WBBB$. In this way, we show that $WBBA$ cannot be dominated by any other strategies.

Similarly, $WABA$ and $WABB$ are the only best-responding pure strategies for one player choosing $WABA$ and the other player choosing $p \cdot A \oplus (1-p) \cdot WABA$ with $p \in (0, \frac{a-c}{2a-b-c})$. Moreover, $WBAA$ and $WBAB$ are the only best-responding pure strategies for one player choosing the mixed strategy $p \cdot A \oplus (1-p) \cdot WBBA$ with $p \in (0, 1)$ and the other choosing $WBAA$. Lastly, $WAAA$ and $WAAB$ are the only best-responding pure strategies for one player choosing $WAAA$ and the other choosing a mixed strategy $p \cdot A \oplus (1-p) \cdot WABA$ with $p \in (0, 1)$. Following the above logic, we can show that $WABA$, $WBAA$, and $WAAA$ cannot be dominated. In addition, A is the unique best response to all other players choosing $WBAA$.

Therefore, in the first round of elimination, we can eliminate any strategy that involve choosing B after seeing all others choosing A ($n^A = N - 1$) at $t = 0$. However, the strategy of not waiting (i.e., A), and strategies that involves waiting and then choosing either B or A after any $n^A < N - 1$ (i.e., $WBBA$, $WABA$, $WBAA$ and $WAAA$), cannot be eliminated.

After eliminating $WBBB$, $WBAB$, $WABB$ and $WAAB$, by repeating the same arguments for why other strategies cannot be eliminated in the first round, we can show that each strategy that survives the first round of elimination is, in fact, a unique best response to some strategies taken by other players. Hence, none of them can be eliminated further.

To summarize, strategy profiles consistent with iterated weak dominance are (1) all players choose A at $t = 0$; and (2) all players wait and choose A or B when $n^A < N - 1$ but choose A when $n^A = N - 1$.

Subgame-Perfect Nash Equilibria The subgame-perfect equilibria take the following forms. All players choose A at $t = 0$. In all other cases, all players choose "wait" at $t = 0$, choose A when $n^A = N - 1$, and choose A or B if $n^A = 2, ..., N - 2$. There are multiple possibilities for $m^A = 0, 1$. In one case, all players also choose A following $n^A = 0, 1$. In another case, all players choose B following $n^A = 0, 1$. In the third case, all players choose A following $n^A = 0$ and choose B following $n^A = 1$.

It is easy to check that any of the following strategies – A, $WAAA$, $WBBA$ and $WABA$ – can constitute a pure-strategy symmetric subgame-perfect equilibrium. To see why each player choosing the strategy $WBAA$ is not such an equilibrium, let us consider the case in which the other two players choose $WBAA$. In that case, a player would choose A and receive a monetary payoff a rather than the strategy $WBAA$ and receiving a monetary payoff of b. $\square$

Appendix B Statistical Tests

Rounds	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
(1) Chi-squared tests: rates of group efficient coordination in St-b and BI-b															
χ	99.556	155.556	126.000	126.000	126.000	126.000	126.000	188.222	155.556	155.556	188.222	188.222	188.222	155.556	99.556
p	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
(2) T tests: group total earnings in St-b and BI-b															
t	3.686	5.777	2.525	4.151	0.750	1.397	2.351	4.183	3.833	3.721	4.421	4.091	4.183	3.960	0.662
p	0.001	0.000	0.016	0.000	0.458	0.170	0.024	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.512
(3) Chi-squared tests: rates of group efficient coordination in St-b and St-f															
χ	7.333	0.970	0.545	0.545	0.545	0.545	0.545	0.545	0.545	0.545	0.545	0.545	0.545	0.545	0.970
p	0.007	0.325	0.460	0.460	0.460	0.460	0.460	0.460	0.460	0.460	0.460	0.460	0.460	0.460	0.325
(4) Chi-squared tests: rates of group efficient coordination in BI-b and BI-f															
χ	24.545	8.337	1.000	1.000	1.000	0.028	0.028	0.000	0.462	0.462	1.500	1.500	1.500	0.462	3.927
p	0.000	0.004	0.317	0.317	0.317	0.868	0.868	1.000	0.497	0.497	0.221	0.221	0.221	0.497	0.048
(5) Chi-squared tests: rates of group efficient coordination in BI-f and St-f															
χ	8.800	19.210	17.818	17.818	17.818	29.184	29.184	46.545	46.545	46.545	75.758	75.758	75.758	46.545	65.636
p	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
(6) Chi-squared tests: rates of group efficient coordination in NI-f and BI-f															
χ	213.351	61.016	46.286	21.429	21.429	8.278	8.278	21.600	4.762	4.762	9.333	9.333	9.333	4.762	0.190
p	0.000	0.000	0.000	0.000	0.000	0.004	0.004	0.000	0.029	0.029	0.002	0.002	0.002	0.029	0.663
(7) Chi-squared tests: rates of group efficient coordination in AI-f and BI-f															
χ	4.364	7.111	16.000	6.000	6.000	5.486	5.486	12.343	5.333	5.333	12.000	12.000	12.000	5.333	5.333
p	0.037	0.008	0.000	0.014	0.014	0.019	0.019	0.000	0.021	0.021	0.001	0.001	0.001	0.021	0.021

Table 4: Statistical tests

Note: (1),(3), (4), (5), (6), and (7) report the Chi-squared of the percentages of groups achieving efficient coordination between two treatments (as noted) in each round. The null hypothesis is that the distributions are the same in the two treatments. (2) reports the T-tests of group total earnings in St-b and BI. The null hypothesis is that the group earnings are the same in the two treatments.