Coordination with preferences over the coalition size[*]

Philippos Louis[†] Orestis Troumpounis[‡] Nikolas Tsakas[§]
Dimitrios Xefteris[¶]

September 2, 2021

Abstract

We study a coordination game where N players simultaneously and independently decide whether to take a certain action or not. Players' payoffs depend only on how many players take each action (i.e., the coalition size) and there is incomplete information on players' types: "Dominant" types have a dominant strategy and care about "enough" people taking an action. "Non-dominant" types do not have a dominant strategy due to non-monotone preferences over the coalition size: their payoffs are maximized when "enough" but "not too many" people take an action. We focus on the behavior of "non-dominant" types and show how the frequency of taking each action and (mis)coordination outcomes depend on the distribution of types and types' preference heterogeneity. Our experimental results are (mostly) in line with our theoretical predictions: The frequency of coordination failure is not only increasing in the preference heterogeneity —as predicted by the theory, but is also increasing in the share of "non-dominant" types.

Keywords: coordination; anti-coordination; laboratory experiment; protest voting.
JEL classification: D72

[*]A previous version was circulated as "Protest voting in the laboratory". Financial support by University of Cyprus Starting Grant 8037E-32806 is gratefully acknowledged. We are grateful to seminar and conference participants at various (mostly online) events.

[†]Department of Economics, University of Cyprus, PO Box 20537, 1678 Nicosia, Cyprus, e-mail: louis.philippos@ucy.ac.cy

[‡]University of Padova, via Del Santo 33, 35137, Italy & University of Lancaster, The Management School, UK, e-mail: orestis.troumpounis@unipd.it

[§]Department of Economics, University of Cyprus, PO Box 20537, 1678 Nicosia, Cyprus, e-mail: tsakas.nikolaos@ucy.ac.cy, Corresponding author.

[¶]Department of Economics, University of Cyprus, PO Box 20537, 1678 Nicosia, Cyprus, e-mail: xefteris.dimitrios@ucy.ac.cy

1 Introduction

Imagine you just received an invitation to a picnic party. Each person invited is asked to contribute either food or beverages, with an equal dollar value, and all food and beverages are shared among invitees. What should you bring along? A picnic with no beverages is going to be dry. Not enough food and folks will be hungry. Achieving the perfect mix of the two requires coordination that is not trivial in the absence of communication and can be further complicated if preferences are private information.

This strategic situation poses a coordination problem with three core elements: First, individual players only care about the distribution of choices, i.e., *the size of the coalition*: how much food and beverages are there at the picnic, but not who brought what. This feature is also prevalent in other coordination games, as for instance in standard voting games. Second, preferences over the size of the coalition may not be monotone,[1] which is a common feature of anti-coordination games. Third, there is incomplete information about other players' preferences, which is often the case in several real-world scenarios.

Examples of such situations are found beyond social life. They can arise in teams working on tasks in which individual effort by any team member may have undesired externalities. Or in cases where teams are working towards multiple goals simultaneously. In politics, a similar coordination problem arises in what is known as protest voting.[2] In this case voters are in favor of an incumbent, but may consider voting against her, hoping that she only gets reelected marginally, in order to express dissatisfaction about the way she handled certain issues.

Let us now further illustrate the characteristics of the setting we consider by returning to our simple picnic example. In this case, there are three possible outcomes: *a*) a picnic with plenty of beverages but not enough food, *b*) a picnic with plenty of food but not enough beverages, or *c*) a successful picnic offering the right mixture of food and beverages. Suppose invitees can be of two types. 'Appetite' types enjoy the picnic as

[1] In section 1.1 we further discuss the relationship of our work to the existing literature on coordination.

[2] See Kselman and Niou (2011); Myatt (2017).

long as enough food is provided. They are indifferent between outcomes b and c, and strictly prefer both to a. 'Balance' types also appreciate when enough food is available, but they strictly prefer a picnic that also offers enough beverages. They strictly prefer outcome c, followed by b and consider a the worst. Assuming incomplete information and random draws of types from a commonly known distribution, what should each invitee contribute to the party? Or, more generally, if in such a coordination problem types are private information, how are individuals' choices determined by what they know about others' preferences? This is the first question this paper addresses.

Individual choices will of course determine the success of the picnic party. Outcome c is clearly optimal due to invitees' *efficient coordination*: the "right" mixture of beverages and food is available and both types are satisfied. Outcome a at the other extreme is a *coordination failure*: not enough food is provided, letting both types unsatisfied. Outcome b instead is an outcome of *partial coordination*: while invitees successfully coordinate so that enough food is provided, they fail to coordinate so that enough beverages are also available. 'Appetite' types are satisfied, but 'balance' types are not. What determines the frequency of *efficient coordination*, *partial coordination*, or *coordination failure*? This is the second question this paper addresses.

The picnic party example captures the key elements of the coordination problem at hand: a number of individuals simultaneously and independently decide whether to take a certain action or not and there is incomplete information on individual types. *Dominant types* have monotone preferences over the number of individuals taking the action (i.e., the coalition size) and hence for them taking the action is a dominant strategy (for the 'appetite' types in the example, contributing food is a dominant strategy). *Non-dominant types* instead have non-monotone preferences over the coalition size and therefore no dominant strategy (the 'balance' types in our example).[3] In other words, dominant types care about the coordination outcome over one threshold: "enough" people taking the action. Non-dominant types instead care about the coordination outcome over two

[3]We borrow the definition of types from Baliga and Sjöström (2004); Jelnov et al. (2020).

thresholds: "enough" but not "too many" people take the action. An important feature of this setting is that while payoffs depend on one's type and the number of individuals taking the action, individual payoffs are the same for two individuals of the same type taking different actions.

While for dominant types taking the action seems the obvious choice, non-dominant types (may) need to randomize between taking the action or not. Overall, efficient coordination cannot be guaranteed. What affects the behavior of non-dominant types and the frequency of coordination outcomes? Our analysis focuses on two key parameters of the above-presented setting: i) the distribution of individuals across types, and ii) the preference heterogeneity across types.

Returning to our simple example, consider the case where 'balance' types are rare. It is evident then that a 'balance' type should contribute to the party beverages and the equilibrium is in pure strategies. As these types become more frequent, the (type-symmetric) equilibrium is obtained in mixed strategies. In this mixed strategy equilibrium, 'balance' types contribute beverages less often as their (expected) share increases. Interestingly, in equilibrium the frequency of each outcome is not affected by the (expected) share of non-dominant types. For example, an increase in the (expected) share of 'balance' types is fully offset by the decrease in the probability that each of them contributes beverages to the picnic, leaving the frequency of each outcome unaffected.

However, the prospects of (un)successful coordination reacts to the degree of preference heterogeneity across types. Types are heterogeneous due to the additional utility 'balance' types assign to outcome c compared to outcome b, with 'appetite' types indifferent between outcomes b and c. When this utility difference goes to zero, 'balance' and 'appetite' types are quite "similar": both types care about avoiding a party with not enough food. As this utility difference is getting larger, the type heterogeneity increases: while 'appetite' types still only care about the availability of enough food, 'balance' types have stronger incentives to guarantee the availability of enough beverages. Therefore, an increase in preference heterogeneity makes 'balance' types contribute beverages more often, resulting in more parties failing due to the availability of too many beverages and not enough food, and in fewer parties with too much food and not enough beverages. More formally, our results show that the frequency non-dominant types take the action is decreasing in preference heterogeneity. This leads to an increase in the frequency of coordination failures and a decrease in the frequency of partial coordination. The effect of preference heterogeneity on the frequency of efficient coordination instead, is inverse-U shaped.

But is this equilibrium analysis pertinent to settings of empirical interest? To give a first answer, we take the above theoretical predictions to the laboratory and consider the simplest possible experimental setup to test them: Two subjects are randomly assigned either the dominant or non-dominant type with a given probability and choose simultaneously and independently whether to take the action or not. *Efficient coordination* is achieved when one subject takes the action and the other does not. Both subjects taking the action results to *partial coordination*, while none of the two subjects taking the action results to *coordination failure*. In this setting, we use a 2x2 factorial design where we vary the two main parameters of our theoretical setting as far as comparative statics are concerned: i) the probability subjects are assigned each type (i.e., the distribution of types), and ii) the utility differential for non-dominant types (i.e., types' preference heterogeneity).

Overall, we observe subjects "over-taking" the action in three out of four treatments, compared to the equilibrium benchmarks. Nevertheless, the comparative statics regarding individual behavior across treatments are very much in line with the theoretical predictions. Regarding coordination outcomes, in treatments with frequent non-dominant types, perhaps surprisingly, subjects achieve higher than anticipated levels of efficient coordination. We provide evidence that indicates that subjects employ what we call *role playing*: instead of mixing, a subject adopts a specific role of either always or never taking the action when assigned a non-dominant type. As we explain in the theory section, this type of behavior, which can explain the high levels of efficient coordination, would also appear if subjects were playing an asymmetric Bayes-Nash equilibrium.

1.1 Related Literature

Theoretical and subsequently experimental research on strategic coordination remains active for the past three decades (see Cooper et al. 1990; Van Huyck et al. 1990, 1991 and Ochs 1990 for some early experimental work; Devetag and Ortmann 2007, Weidenholzer 2010 and Camerer 2011 for more recent surveys of the literature). Two opposing building blocks of these models are present in coordination and anti-coordination games (Chierchia et al. 2018). In the classical examples of coordination individuals have incentives to match their actions (e.g., pure coordination games, stag-hunt games, standard voting games, and many others, see for instance Mehta et al. 1994; Battalio et al. 2001). In the classical example of anti-coordination instead individuals have incentives to mismatch their actions (e.g., entry games, see for instance Selten and Güth 1982; Camerer and Lovallo 1999).

Our setup shares features with some coordination games in which players care about the distribution of choices, but do not necessarily care about the identity of the players who chose each action. This feature is particularly prevalent in standard voting games —it does not matter who voted for whom but the election winner. It also shares elements with anti-coordination games, as a distinctive feature of our model is that preferences over the size of the coalition need not be monotone.

A common class of coordination problems with preferences over the coalition size – and irrespective of the specific individual choices– includes work on political economy and voting in binary elections: voters obtain their type-specific payoff depending on how many individuals voted for each candidate regardless of whether they voted for one candidate or another. In classical models of strategic voting in multi-candidate elections preferences are monotone and there is need for coordination among subjects belonging to the majority, to avoid a coordination failure where the Condorcet loser wins the election (Palfrey 1989; Forsythe et al. 1993; Myerson and Weber 1993; Fey 1997; Bouton and Castanheira 2012; Bouton et al. 2017). Instead, non-monotone preferences –as in our setting– appear in the model of protest voting proposed by Myatt (2017).[4] In the simplest version of such model, all voters agree that the incumbent is preferred over her opponent. While dominant types just rank the two candidates, non-dominant types' most preferred outcome is a "successful protest": enough votes for the incumbent so that she wins the election, but not too many votes for her so that voters signal their dissatisfaction. Given that in equilibrium non-dominant types (may) randomize between the two alternatives, coordination failures resulting to electoral accidents where the incumbent loses cannot be ruled out. Our experimental analysis directly applies to this setting and presents one of the first experiments on protest voting. Our results from the laboratory would suggest that non-dominant types protest more often as they become more fanatic or less popular.

The presence of incentives towards anti-coordination reveals some similarities with entry games. However, in standard entry games payoffs depend both on the coalition size and on the choice of the individual, which is not the case in our setup. An additional distinction of our paper with respect to the respective experimental literature is that, with some exceptions (e.g., Heinemann et al. 2004; Cabrales et al. 2007; Kaplan and Ruffle 2012), entry game experiments typically feature complete information (e.g., Sundali et al. 1995; Rapoport et al. 1998; Erev and Rapoport 1998; Zwick and Rapoport 2002; Duffy and Hopkins 2005). In our game, players are uncertain about their opponent types. Nevertheless, our results about individual behavior exhibit some parallelism with the findings of this literature. Nash predictions do a good job explaining behavior in the experiment, but whether subjects converge to the unique symmetric mixed strategy equilibrium or an asymmetric pure strategy equilibrium depends on the particular treatment parameters used. Hence, our results contribute to the experimental literature of strategic coordination games with incomplete information.

[4]The coordination game we present is in fact inspired by Myatt (2017). Nevertheless, the two models are not nested: Myatt (2017) considers richer environments regarding types and information (i.e., aggregate uncertainty), but focuses on bell-shaped distributions, which make it hard to disentangle between the two key elements of our model: the distribution across types and preference heterogeneity. One can view our experimental results as (mostly) corroborating the main insights of Myatt's (2017) formal analysis. For a broader review of the literature on protest voting see Alvarez et al. (2018).

In what follows we first derive our hypotheses by studying a simple model of coordination with non-monotone preferences (Section 2), we then present our experimental design (Section 3) and results (Section 4), and, finally, we conclude (Section 5).

2 Formal arguments

Consider a society of N individuals with $N \geq 2$. Among them, n individuals, where $2 \leq n \leq N$, have to make a payoff-relevant binary choice $a \in \{0, 1\}$, henceforth referred to as take an action ($a = 1$) or not ($a = 0$). Individuals are of two types: "dominant" or "non-dominant", denoted $\{d, nd\}$. The utility of individual i depends on the size m of the coalition of individuals whose choice is payoff relevant and take the action, where $0 \leq m \leq n$. Dominant types only care about "enough" individuals taking the action. Non-dominant types also care about "enough" individuals taking the action, but they obtain an even higher utility if "enough" *but not* "too many" individuals take the action. Formally,

$$U_i^d(m) = \begin{cases} 0 & \text{if } m < t_1 \\ 1 & \text{if } m \geq t_1 \end{cases} \qquad U_i^{nd}(m) = \begin{cases} 0 & \text{if } m < t_1 \\ 1 + h & \text{if } t_1 \leq m \leq t_2 \\ 1 & \text{if } m > t_2 \end{cases}$$

where thresholds t_1 and t_2, with $1 \leq t_1 \leq t_2 \leq n - 1$, quantify the notion of "enough" and "too many" individuals taking the action.[5] The two thresholds determine three possible coordination outcomes: (i) *coordination failure*, i.e., $m < t_1$, (ii) *efficient coordination*, i.e., $t_1 \leq m \leq t_2$, and (iii) *partial coordination*, i.e., $m > t_2$.[6] Parameter $h > 0$ denotes the preference heterogeneity across types. The above preferences are summarized in Figure 1.

[5] The case of $t_2 = n$ would be rather uninteresting, as that would mean that non-dominant types would also have a unique threshold at t_1 and would just receive a higher utility than dominant types if this was surpassed. Thus, they would also have no incentive to not take the action.

[6] Of course, the payoffs in the partial and efficient coordination outcomes can coincide for certain realization of types, a situation pertinent in our experimental setting. For consistency, we use this terminology throughout the paper.

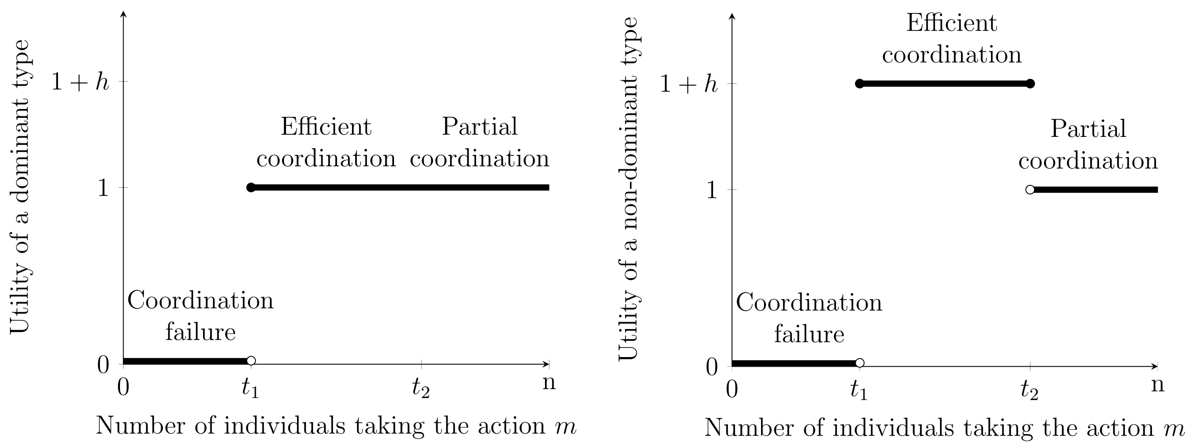


Figure 1: Preferences for dominant and non-dominant types over the coalition size m and the three possible coordination outcomes.

The timing is as follows: Nature draws uniformly at random the n individuals out of the N that will be relevant, i.e., whose actions will be relevant for payoffs. Then, nature also draws individuals' types with *i.i.d.* draws from a binomial distribution with parameter $q \in (0, 1)$, which is common information and represents the probability that a random individual is assigned the non-dominant type. Each individual observes her own type and whether or not she is among the n individuals whose actions will be payoff relevant, but does not observe any of this information for other individuals. Subsequently, individuals choose simultaneously their (mixed) strategies, where the strategy of player i is the probability p_i with which she takes the action.

At this point, we should stress that the modeling choice of considering a subset of $n \leq N$ individuals chosen at random to make payoff-relevant choices is a generalization of the more natural and standard environment where all N individuals do so. The reason we do that is because it allows us to get some additional insights on the adoption of asymmetric equilibrium strategy profiles, i.e., equilibria in which individuals of the same type choose different strategies. It also permits us to bring the theoretical model closer to our experimental setting, were we employed a random matching in groups of four with two individuals playing together.[7] The main theoretical results we describe below also

[7]This assumption is actually realistic in some particular instances, for example elections with sor-

hold for the case where $n = N$.

Given that strategic decisions are made simultaneously, the natural solution approach is to seek for Bayesian Nash Equilibria (BNE). More specifically, we consider the agent-normal form version of the game, in which each individual at each information set is considered as a distinct player. An information set here is a pair of type and relevance –dominant/non-dominant, and relevant/non-relevant. Nevertheless, for simplicity and given that the choices of non-relevant players do not affect anyone's payoffs, without loss of generality, we assume that they always choose to take the action for sure. Thus, with a slight abuse of terminology, in our subsequent analysis we condition individual behavior only on type and describe equilibria accordingly.

Throughout the analysis, we treat N, n, t_1 and t_2 as given, and focus on the comparative statics of the distribution of types, q, and types' preference heterogeneity, h, on individual behavior and coordination outcomes. Let us define $\mathcal{HQ} := \mathbb{R}_+ \times (0, 1)$, which is the set of admissible pairs (h, q).

Non-Responsive equilibria

The game we described admits a large number of equilibria in which players' actions and the outcome do not depend on the draw of types. These include, for example, the cases where $t_1 > 1$ and $t_2 < n - 1$ and either all or none of the players take the action. It also includes equilibria in which specific players take the action and the rest do not.

We find these equilibria to be less interesting in our setup, since in many cases these equilibrium strategy profiles require the use of dominated strategies by some players, namely dominant types required to not take the action in equilibrium. Thus, in the current setup it is unlikely for such behavior to emerge in a decentralized manner.

For the remainder of the paper we focus on equilibria that are *responsive*: the equilibrium outcome is not constant in the players' type distribution. That is, given an

tition, dating back in the Athenian democracy where individuals were randomly chosen to participate in the election (see for example Saran and Tumennasan 2013, 2019 for theoretical research on elections with sortition).

equilibrium strategy profile, there is at least one possible draw of players' types that results in a different outcome. Thus, we do not require that every player changes their strategy according to their type, but some do and are able to affect the outcome.

Type-symmetric equilibria

We start our analysis focusing on symmetric equilibria. *Symmetry* requires that all players of the same type employ the same strategy. It rules out equilibria in which different players behave differently when they are assigned the same type.

The following proposition characterizes the unique BNE that is both symmetric and responsive in our setting.

Proposition 1. *Let*

$$\tilde{q}(h) = 1 - \frac{1}{1 + \left[\frac{h}{1+h}\frac{(t_1-1)!(n-t_1)!}{t_2!(n-t_2-1)!}\right]^{\frac{1}{t_2-t_1+1}}}$$

For each pair $(h,q) \in \mathcal{HQ}$, there exists a unique BNE that is both symmetric and responsive:

1. *When $q \leq \tilde{q}(h)$, the equilibrium is in pure strategies: each individual chooses to take the action when assigned a dominant type and not to take the action when assigned a non-dominant type.*

2. *When $q > \tilde{q}(h)$, the equilibrium is in mixed strategies: each individual chooses to take the action with probability equal to 1 when assigned a dominant type and to take the action with probability equal to $p(h,q) = \frac{q-\tilde{q}(h)}{q}$ when assigned a non-dominant type.*

The threshold value $\tilde{q}(h)$ determines whether non-dominant types follow a pure strategy and never take the action or they randomize between the two actions. If the probability with which individuals are assigned a non-dominant type is relatively low, i.e., if $q \leq \tilde{q}(h)$, then most likely there are not going to be enough non-dominant types to give rise to a coordination failure, and hence non-dominant types never take the action. On the contrary, if $q > \tilde{q}(h)$, non-dominant types follow a mixed strategy where they take the action with some positive probability $p(h, q)$. This probability is computed so that non-dominant types are indifferent between taking the action or not. Given $p(h, q)$, we can also compute the probability with which a random payoff-relevant individual, henceforth called a *random individual*, takes the action in equilibrium, which we shall denote by $r(h, q)$.[8]

When individuals are assigned a non-dominant type with low probability, i.e., when $q < \tilde{q}(h)$, the comparative statics are obvious. In this region, both types follow pure strategies, with dominant types taking the action and non-dominant types not taking the action. Thus, the probability that a random individual takes the action equals the probability with which an individual is assigned a dominant type, i.e., $r(h, q) = 1 - q$, and is hence decreasing in q. The arguably more interesting comparative statics appear in the region where individuals are assigned a non-dominant type with high enough probability, i.e., when $q > \tilde{q}(h)$. In these cases, the probability that a random individual takes the action is $r(h, q) = (1 - q) + q \times p(h, q)$. These comparative statics results summarized in the following proposition pave the ground for our main empirical hypotheses; the ones that we subsequently test in the laboratory.

Proposition 2 (Comparative Statics)**.** *For each pair $(h, q) \in \mathcal{HQ}$, in the unique BNE that is both symmetric and responsive, the probability $p(h, q)$ with which non-dominant types takes the action and the probability $r(h, q)$ with which a random individual takes the action satisfy the following:*

- *If (h, q) is such that $q > \tilde{q}(h)$:*

  1. $p(h, q)$ is strictly increasing in q and strictly decreasing in h,

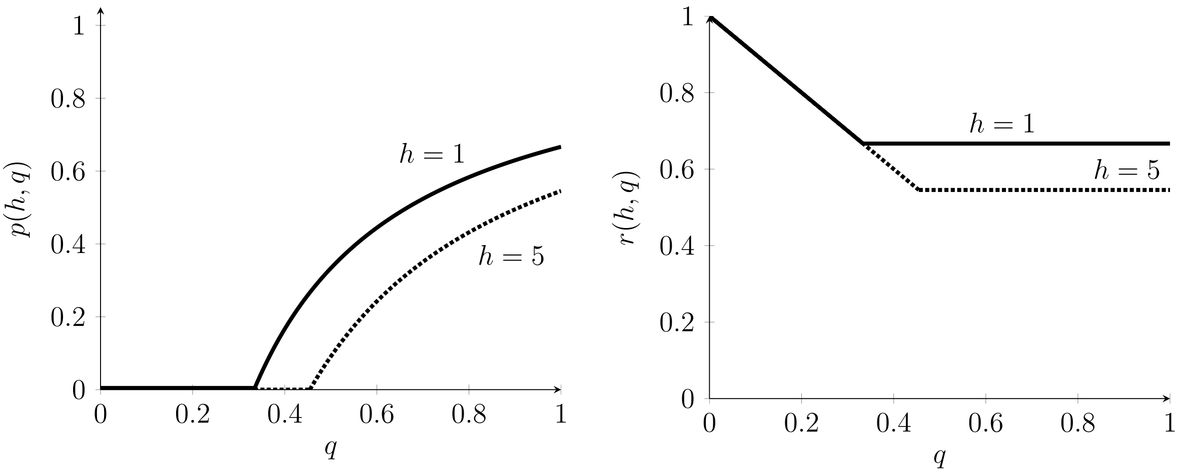
  2. $r(h, q) = (1 - q) + q \times p(h, q)$ does not vary in q and is decreasing in h.

- *If (h, q) is such that $q < \tilde{q}(h)$:*

[8] Under the assumption of symmetry, this probability is the same for all individuals and it is calculated conditional on the individual being payoff relevant, but before the realization of types.

1. $p(h, q) = 0$*,*

2. $r(h, q) = 1 - q$ *is strictly decreasing in* q *and does not vary with* h*.*



(a) Probability that an individual who is assigned a non-dominant type takes the action.

(b) Probability that a random individual takes the action.

Figure 2: An example where $n = 2$, $t_1 = t_2 = 1$. Solid line for $h = 1$, dashed line for $h = 5$. Value of $\tilde{q}(h)$ is 1/3 for $h = 1$ and 5/11 for $h = 5$.

Given some level of type heterogeneity h, for any value of $q > \tilde{q}(h)$ the equilibrium probability with which non-dominant types take the action is increasing in q. As non-dominant types are assigned more frequently, they have incentives to take the action more often to avoid a coordination failure. The exact shape of $p(h, q)$ for an example with $n = 2$ is depicted in Figure 2a. Figure 2b instead illustrates the probability with which a random individual (without conditioning on type) takes the action. As we have already mentioned, for $q < \tilde{q}(h)$ this probability is linearly decreasing in q, whereas, more interestingly, for $q > \tilde{q}(h)$ it does not depend on q. The mixed strategy played by non-dominant types ensures exactly that. More specifically, in expectation, non-dominant types choose how often to take the action so that the "right" share of individuals is taking the action. This "right" share ends up being independent of q and is, in fact, equal to $1 - \tilde{q}(h)$. Hence, a change in the distribution of types (i.e., in q) triggers an appropriate downward adjustment in the individual frequency with which non-dominant types take the action. The two effects counterbalance each other and as a result they leave the likelihood of each coordination outcome unaffected.

Given a fixed probability q, as long as (h, q) is such that $q > \tilde{q}(h)$, changes in type heterogeneity h also affect the probability that a randomly chosen individual takes the action and therefore the likelihood of each coordination outcome. As the type heterogeneity increases the utility difference for a non-dominant type between the efficient and partial coordination increases, providing incentives to non-dominant types to take the action less often (see Figure 2a). Given that q is fixed, an increase in h –as long as it does not result to $q < \tilde{q}(h)$– results to a decrease in the probability that a random individual is taking the action. This increases the frequency of coordination failures and decreases the frequency of partial coordination. The relationship between this probability and the frequency of efficient coordination instead is non-monotonic. The left panel of Figure 3 illustrates the relationship between the probability a random individual takes the action and the frequency of coordination outcomes for our example with $n = 2$ and $t = 1$.

On the right panel of Figure 3, we represent the same frequencies of coordination outcomes on the simplex. The dashed line on the simplex represents the expected frequencies of the three outcomes assuming that a randomly chosen individual takes the action with *any* probability in the $[0, 1]$ interval. If this probability is one, individuals partially coordinate with certainty and we are at the northern corner of the simplex. As the probability to take the action decreases, we move along the dashed line till coordination fails surely when randomly selected individuals never take the action. Both panels of Figure 3 also illustrate the expected frequencies of the three outcomes in our two-individuals example with $h = 1$ and $h = 5$ when non-dominant types play the mixed strategy characterized in Proposition 1 and hence the probability of a random voter taking the action is $6/11$ and $2/3$ respectively.

Type-asymmetric equilibria in pure strategies

Next, we extend our analysis by relaxing the assumption of *symmetry*, but maintaining the requirement of equilibria to be *responsive*. Out of all possible type-asymmetric equilibria,

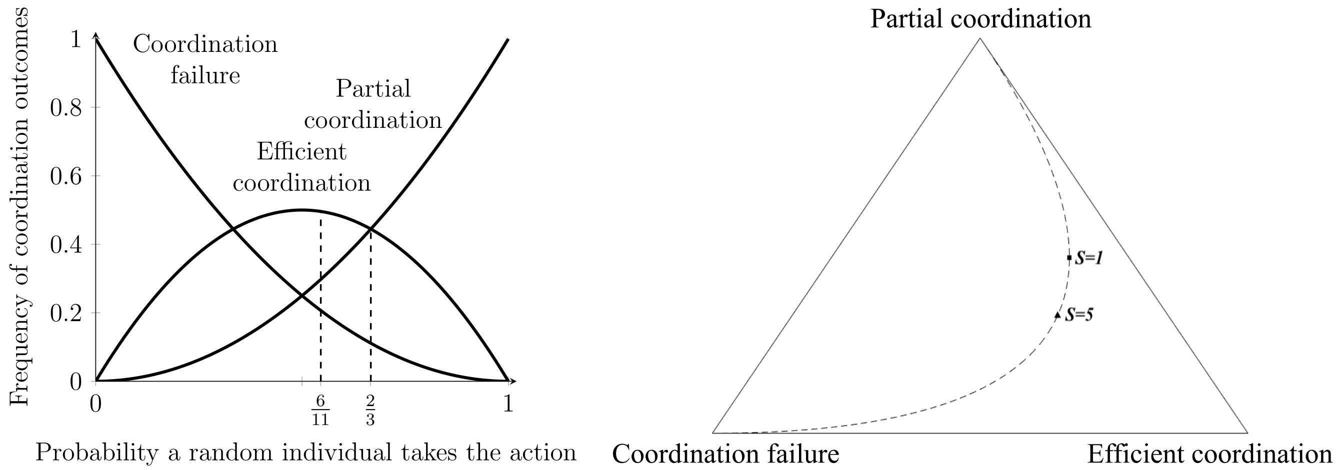


Figure 3: The left panel presents the frequencies of coordination outcomes in the example where $n = 2$, $t_1 = t_2 = 1$ assuming that a randomly selected individual takes the action with a given probability $[0, 1]$. For $h = 1$ the probability a random individual takes the action is 2/3. For $h = 5$ this probability is 6/11. The right panel contains a representation of such frequencies on the simplex. Both representations highlight that this specific increase in h —accompanied by a decrease in the probability a random individual takes the action— increases the frequency of coordination failures and efficient coordination and decreases the frequency of partial coordination.

we restrict our attention to type-asymmetric BNE in pure and undominated strategies.

As we shall see, asymmetries allow different players to assume different *"roles"* in the game, with some of them conditioning their choices on their type and others always simply taking the action irrespective of their type. Such *role-playing* behavior is interesting in its own right, even beyond the scope of equilibrium, especially in light of our experimental findings.

The characteristics of asymmetric responsive BNE in pure and undominated strategies that we focus on are the following: (1) each individual chooses to take the action when assigned a dominant type, (2) K out of the N individuals choose to take the action also when assigned a non-dominant type, whereas the rest $N - K$ individuals do not take the action when assigned a non-dominant type, and (3) $1 \leq K \leq N - n + t_2$. Condition (3) guarantees that the equilibrium will be responsive. If in some equilibrium most of the players, though not necessarily all, were to choose to take the action even when assigned a non-dominant type (i.e., $K \geq N - n + t_2 + 1$), then any draw of players' types would lead to the same outcome, thus would not be responsive. Such equilibria may exist here, but condition (3) rules them out. Henceforth, we refer to such a BNE as a *role-playing equilibrium.*

Proposition 3. *Let $\tilde{q}(h)$ be the same as defined in Proposition 1. Then, for a pair of parameters $(h, q) \in \mathcal{HQ}$:*

1. *When $q \leq \tilde{q}(h)$, the game has no role-playing equilibrium.*

2. *When $q > \tilde{q}(h)$, the game has at least one role-playing equilibrium and in any role-playing equilibrium it holds that $K < N - n + t_1$.*

Proposition 3 provides useful intuition. First, it strengthens further the appeal of the symmetric pure strategy equilibrium for parameter values below the threshold $q \leq \tilde{q}(h)$. Second, for values above the threshold, not only guarantees the existence of a role-playing equilibrium, but also provides a partial characterization. To observe that, note that the threshold of K below which role-playing equilibria are proven to exist is lower than the threshold that induces responsiveness. This means that role-playing equilibria will typically involve a substantial share of players assuming each of the two roles.

Additionally, the proof of this result contains an explicit expression of the conditions on expected utilities that need to be satisfied in an equilibrium, which are quite intuitive. Namely, increasing (decreasing) the number of non-dominant types who take the action by one has two opposite effects: it increases (decreases) the likelihood of passing from a coordination failure to efficient coordination, but it also increases (decreases) the likelihood of passing from efficient coordination to partial coordination. An equilibrium is such that either an increase or a decrease in the number of non-dominant types who take the action makes the respective negative effect prevail over the positive one.

The explicit equilibrium conditions –despite their complexity– can also allow us to point down all such equilibria for any specific choice of parameter values. This is helpful, because it could allow us to obtain a better understanding of the number and properties of such equilibria for any configuration of interest. The following corollary characterizes the role-playing equilibria for each of our four experimental treatments.

Corollary 1. *Let $N = 4$, $n = 2$ and $t_1 = t_2 = 1$. Then,*

1. *for $(h, q) = (1, 0.5)$, there are two role-playing equilibria, with $K = 1$ and $K = 2$.*
2. *for $(h, q) = (1, 0.8)$, there is one role-playing equilibrium, with $K = 2$.*
3. *for $(h, q) = (5, 0.5)$, there is one role-playing equilibrium, with $K = 1$.*
4. *for $(h, q) = (5, 0.8)$, there is one role-playing equilibrium, with $K = 2$.*

3 The Experiment

3.1 Design

The experiment took place at the Laboratory for Experimental Economics at the University of Cyprus. A total of 128 subjects were recruited in 8 sessions, with 16 subjects in each session.[9] The experiment consisted of 100 rounds, prior to which there were 2 practice rounds that aimed at helping the subjects familiarize with the environment. Average total payment was approximately 17.4 euros, including 5 euros as a participation fee, and the experiment lasted about 90 minutes.[10]

The experiment employed the simplest version of the coordination game described in Section 2 with two subjects simultaneously deciding whether to take the action or not. Efficient coordination was achieved when the two subjects mismatched their actions with matching actions instead resulting to either partial coordination or coordination failure.[11] We used a 2×2 factorial design with two sessions per treatment. The two treatment variables were: (i) the types heterogeneity, H, and (ii) the distribution of types, q. Heterogeneity refers to the payoff (in tokens) that non-dominant types enjoy in

[9]Recruitment was done via ORSEE (Greiner et al. 2004).
[10]The experiment was designed and run on z-Tree (Fischbacher 2007).
[11]This would correspond to threshold values $t_1 = t_2 = 1$ and $n = 2$.

the event of an efficient coordination, and may take one of two values in each treatment: $H \in \{300, 700\}$. The probability q of a subject being assigned the non-dominant type was also selected from two possible values in each treatment: $q \in \{0.5, 0.8\}$. A summary of the experimental design is shown in Table 1.

Treatment	Heterogeneity (H)	Distribution (q)	Subjects	Sessions	Subgroups
$(300, 0.5)$	300	0.5	32	2	8
$(300, 0.8)$	300	0.8	32	2	8
$(700, 0.5)$	700	0.5	32	2	8
$(700, 0.8)$	700	0.8	32	2	8
		Total:	128	8	32

Table 1: The four treatments.

The two subjects of each pair were asked to either take the action or not. When no subject took the action, coordination failed and both players received 100 tokens. When both subjects took the action, partial coordination arose and both players received 200 tokens. When the two subjects mismatched their actions: dominant types received 200 tokens and non-dominant types received H tokens. The value of H varied in different treatments. Table 2 summarizes the possible outcomes and payoffs. In the experiment we used a neutral frame.[12]

Note that the payoffs in tokens are a monotone transformation of those described in Section 2. Namely, for an outcome that would yield utility U to a subject, the subject would receive $100(1+U)$ tokens. In the current setup, this transformation does not affect the theoretical predictions. Heterogeneity $H = 300$ ($H = 700$) corresponds to $h = 1$ ($h = 5$) in the theoretical analysis.

In each round, subjects were randomly matched in pairs and played a one–shot game with their assigned pair. We used a stranger matching protocol to avoid any repeated game effects. To maintain a sufficient number of independent observations, pairs were

[12]The two available actions are presented as “choosing **X** or **O**”. Again, see the instructions in the Appendix for details.

actions: outcome:	**00** Coordination failure	**10** or **01** Efficient coordination	**11** Partial coordination
Non-dominant	100	$S \in \{300, 700\}$	200
Dominant	100	200	200

Table 2: Round payoffs in tokens, for each type, depending on outcomes.

drawn from subgroups of four subjects (i.e., $N = 4$ and $n = 2$). Subjects were told that the matching was random and that in each round it was more likely to not be matched with the same subject as in the previous round, while the existence of these subgroups was not detailed to subjects.[13] This design choice is not behaviorally neutral and we come back to that in Section 4.3.

Types were assigned randomly and independently for each subject in each round. Subjects knew their own type but not the type of their partner. They also knew the distribution of types. The process was represented to them as draws (with replacement) from a single urn with red and blue balls. In the treatments where $q = 0.5$ the urn contained five balls of each color, whereas in treatments where $q = 0.8$ the urn contained two red balls and eight blue balls. Drawing a blue ball meant that a subject was a non-dominant type in that round. Both subjects knew the composition of the urn. They could see the ball that was drawn for them, but not the one drawn for their partner. The composition of the urn remained the same throughout each session.

At the end of each round, subjects were informed about the choice of their partner and their payoff from that round. Final earnings were determined by the sum of the subject's payoffs in 10 randomly selected rounds out of the 100.[14] The conversion rate used in the experiment was 1 euro for every 250 tokens.

[13]See instructions in the Appendix.

[14]Subjects were paid based on a random subset of rounds to avoid any wealth effects. Given the large number of rounds in the experiment, we did not pay for a single randomly picked round to keep monetary incentives salient, as some subjects may underweight the very small likelihood of a particular round being picked (Hertwig et al. 2004).

3.2 Testable Hypotheses

Similar to our theoretical setting, dominant types have a straightforward payoff-maximizing option (i.e., to take the action). Hypothesis 0 summarizes this behavior. Hypotheses 1 and 2 instead summarize non-dominant types' behavior as predicted by Proposition 2. Recall that as summarized in Figure 2a, for a given q, non-dominant types take the action less often as the heterogeneity increases. Similarly, for a given level of heterogeneity, non-dominant types take the action more often as q increases. The exact values for the mixed strategy equilibrium probabilities that a non-dominant type takes the action in our experimental setup are given in Table 3 in the following section.[15]

Hypothesis 0. *Dominant types take the action always in all treatments.*

Hypothesis 1. *For a given probability* $q \in \{0.5, 0.8\}$*, non-dominant types take the action less often in the high heterogeneity treatment* $(700, q)$ *than in the low heterogeneity treatment* $(300, q)$*.*

Hypothesis 2. *For a given heterogeneity* $H \in \{300, 700\}$*, non-dominant types take the action more often in the high probability treatment* $(H, 0.8)$ *than in the low probability treatment* $(H, 0.5)$*.*

Hypotheses 3 and 4 and Table 4 summarize the predictions of our model regarding the likelihood of outcomes across treatments (see Figure 3 for the corresponding theoretical predictions). Recall that there are three possible outcomes from a pair's choice: i) *Coordination failure:* None of the two subjects takes the action and each individual gets the lowest possible payoff, irrespective of her type, ii) *Efficient coordination:* One subject takes the action and the other not, and any subject that is a non-dominant type gets the maximum payoff, and iii) *Partial coordination:* Both subjects take the action and hence the intermediate payoff, regardless of their type.

[15] Those values can be directly obtained from Proposition 1 by substituting the relevant parameters of our experimental treatments (i.e., $n = 2$, $t_1 = t_2 = 1$, and $q \in \{0.5, 0.8\}$, $h \in \{1, 5\}$). Actually, those values are the ones that are used in our example of Figures 2 and 3.

Hypothesis 3 fixes the probability q and permits the heterogeneity H to vary. One of our main comparative statics shows that coordination failures occur more often as the heterogeneity increases. This increase in coordination failures goes hand in hand with an increase in the frequency of instances of efficient coordination and a decrease in instances of partial coordination.

Hypothesis 3. *For a given probability* $q \in \{0.5, 0.8\}$*, coordination failures and efficient coordination occur more often in the high heterogeneity treatment* $(700, q)$ *than in the low heterogeneity treatment* $(300, q)$*. Partial coordination occurs more often in the low heterogeneity treatments* $(300, q)$ *than in the high heterogeneity treatments* $(700, q)$*.*

Hypothesis 4 instead fixes the heterogeneity and varies the probability with which nature assigns non-dominant types. As we have shown, changes in q do not affect the frequency of different coordination outcomes. This is because the mixed strategy equilibrium strategies adjust to changes in heterogeneity (Hypothesis 2), letting the frequency of outcomes invariant to changes in heterogeneity.

Hypothesis 4. *For a given heterogeneity* $H \in \{300, 700\}$*, coordination failures, partial coordination, and efficient coordination occur equally often in the high and low probability treatments* $(H, 0.8)$ *and* $(H, 0.5)$*.*

4 Results

We first present results regarding individual behavior and then we argue how such behavior translates to coordination outcomes.

4.1 Individual behavior (Hypotheses 0-2)

Table 3 shows the symmetric (mixed Bayesian Nash) equilibrium predictions for each type in each treatment, as well as the corresponding mean frequency of individuals taking the action observed in the data. Focusing on the lower part of the table, it is reassuring to note that in line with **Hypothesis 0**, dominant types seem to realize that taking the action is payoff maximizing and almost always follow this strategy. Thus, we can now focus on the behavior of non-dominant types.

Figure 4 summarizes the behavior of non-dominant types across treatments, compares that with the Nash predictions, and summarizes relevant tests for differences across treatments. In general we find significant treatment effects as subjects respond to the changes in the treatment variables. When comparing behavior to that predicted by theory, we find that in three out of four treatments, non-dominant types take the action more often than the equilibrium prediction. The exception is the treatment $(300, 0.8)$, where the frequency observed in the data is not statistically different from a BNE. For all other treatments, we have statistically significant over-taking of the action (see Figure 4).

		$H = 300$		$H = 700$	
		$q = 0.5$	$q = 0.8$	$q = 0.5$	$q = 0.8$
Non-dominant	Nash	0.333	0.583	0.091	0.432
	Data	0.409	0.56	0.308	0.483
Dominant	Nash	1	1	1	1
	Data	0.988	0.971	0.959	0.976

Table 3: Equilibrium and actual frequency of taking the action.

Regarding **Hypotheses 1 and 2**, although the frequency of taking the action is mostly different from the equilibrium predictions, comparative statics move in the predicted directions. Performing Wilcoxon rank sum tests comparing the data in two treatments aggregated at the subgroup level (8 observations per treatment) yields low p-values for all treatment effects on the average frequency of taking the action. First, for a given probability q, taking the action among non-dominant types is decreasing in the hetero-

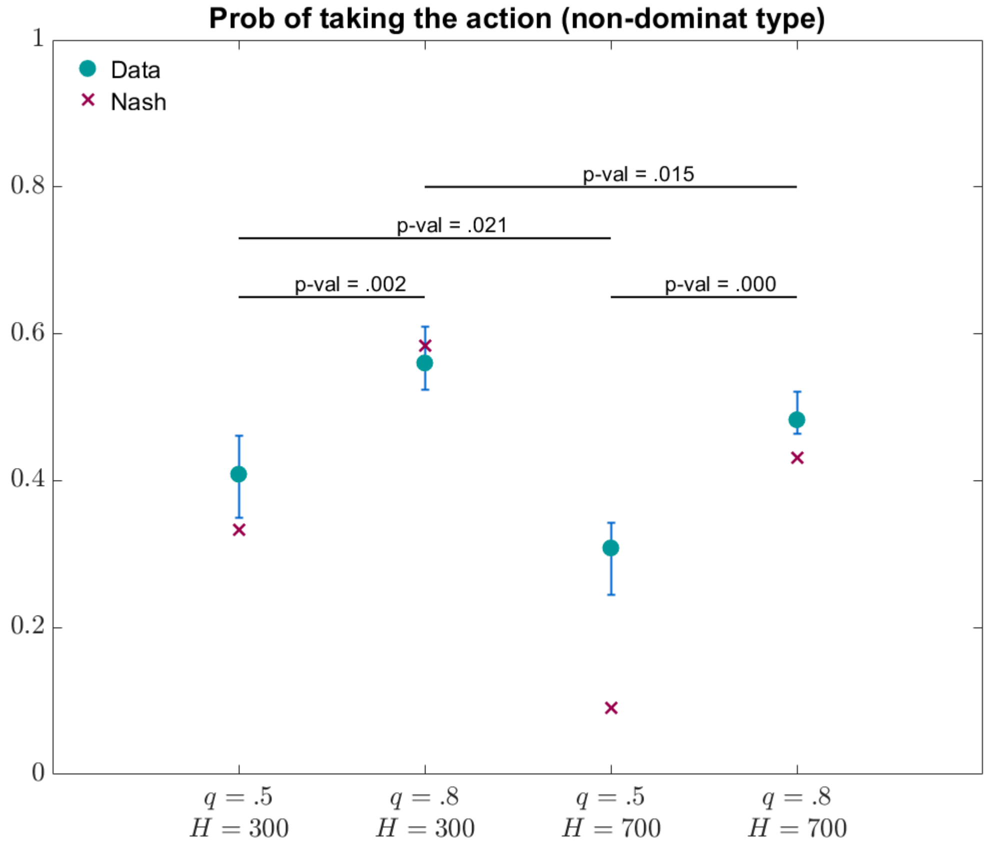


Figure 4: The cyan dots indicate the average frequency of non-dominant types taking the action in each treatment. Error bars indicate 95% confidence intervals for the mean, constructed using 5000 bootstrap samples from the data clustered at the subgroup level. The red X's mark the Nash prediction for each treatment. The p-values refer to the result of a Wilcoxon rank sum test comparing the data in two treatments, indicated by the horizontal bars, aggregated at the subgroup level (8 observations per treatment)

geneity H (p-values <0.01). Second, for a given heterogeneity H, taking the action among non-dominant types is increasing in the probability q (p-values <0.05).

As discussed in Section 4.4, there is some evidence of learning, as individual subjects behave differently in earlier compared to later rounds. Nevertheless, on an aggregate level there is no difference in the frequency of taking the action across rounds in any of the four treatments.

4.2 Coordination outcomes (Hypotheses 3 & 4)

Table 4 shows the outcome distribution in the experimental treatments and the corresponding Nash equilibrium distribution. Figure 5 summarizes those outcome distributions on the simplex. It also reports the p-values for a bootstrapped Hotelling test for compositional data (Tsagris et al. 2017) comparing treatments.[16]

The prediction of **Hypothesis 3** is supported in our data. As we see in Table 4, for a given probability q, we observe more instances of efficient coordination and more coordination failures as the heterogeneity increases, while the frequency of instances of partial coordination decreases. As we see in Figure 5, the corresponding p-values are 0.057 for the high probability treatment (i.e., $q = 0.8$), and 0.053 for the low probability treatment (i.e., $q = 0.5$). In line with our theoretical prediction: For a given q, as types' heterogeneity H increases, non-dominant types realize the potential gain from an efficient coordination and take the action less often (Hypothesis 1), therefore increasing the frequency of coordination failures and of instances of efficient coordination.

Contrary to what we find for the previous hypothsesis, there is little evidence in support of **Hypothesis 4**. Recall that according to theory, for a given H, changes in q should not affect the frequency of the three outcomes. Table 4 instead indicates that for either level of heterogeneity, as the non-dominant types are assigned more frequently, the frequency of outcomes varies. Our data shows that for a given H there are more coordination failures, more instances of efficient coordination, and fewer instances of partial coordination in the high probability treatments than in the low probability treatments. These results are also visualized in Figure 5. For a given level of heterogeneity H, an increase in q moves the distribution point south-east, and those movements are statistically significant (p-values are 0.035 for the high heterogeneity treatment (i.e., $H = 700$), and 0.015 for the low heterogeneity treatment (i.e., $H = 300$). While the increase in efficient coordination and coordination failures have opposite effects on welfare, the former

[16]Comparing treatments by each outcome separately and ignoring the compositional nature of the data, i.e., outcome frequencies add-up to one, yields lower p-values. Hence, the test we use is more conservative than other, less appropriate approaches.

		$H = 300$		$H = 700$	
		$q = .5$	$q = .8$	$q = .5$	$q = .8$
Coordination failure	Nash	.111	.111	.207	.207
	Data	.087	.100	.120	.148
Efficient coordination	Nash	.444	.444	.496	.496
	Data	.422	.512	.484	.538
Partial coordination	Nash	.444	.444	.297	.297
	Data	.491	.388	.396	.314

Table 4: Frequencies of the three possible coordination outcomes

is much more pronounced. This leads to an overall higher efficiency when q is higher.

4.3 Asymmetric behavior

Figure 5 depicts expected and actual outcomes' distributions on the simplex. The dashed line coincides with the one on the right panel of Figure 3. It represents the expected frequencies of the three outcomes assuming a *symmetric behavior* by all individuals, i.e., all dominant types taking the action always and all non-dominant types taking the action with *some* given probability in the $[0, 1]$ interval. If non-dominant types always take the action, partial coordination is guaranteed and we are at the northern vertex of the simplex. As the (symmetric) probability of not taking the action increases we move along the dashed line. The ending point of that curve illustrates the frequency of outcomes when non-dominant types never take the action. The illustrated dashed line in Figure 5 is plotted for $q = 0.8$. While varying the level of q does not affect the shape of the curve, it affects its ending point. That is, for lower levels of q this curve would

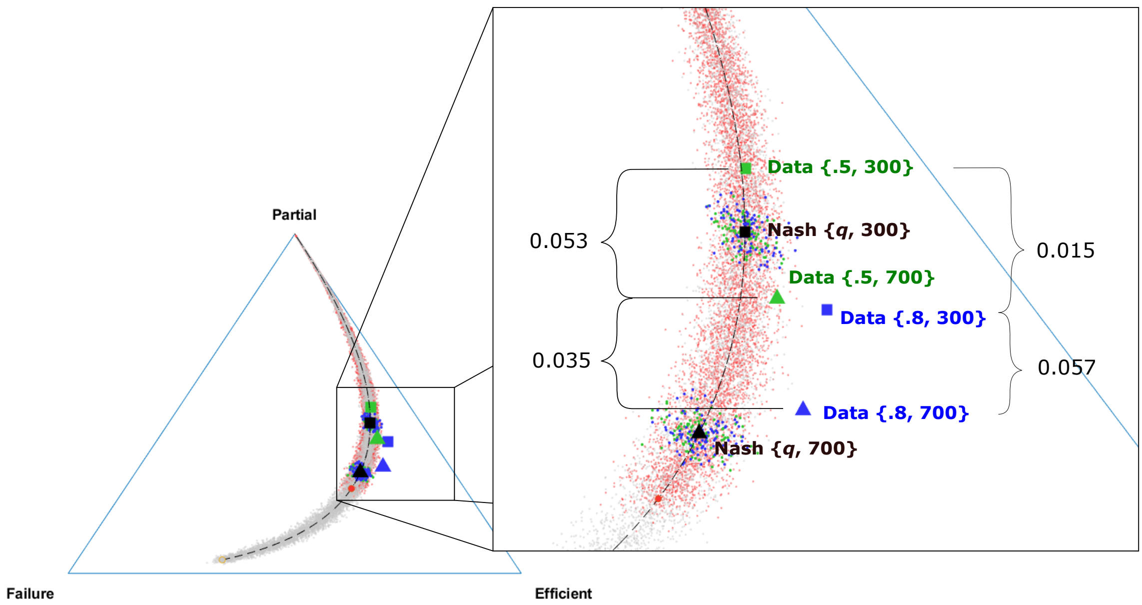


Figure 5: Outcome frequencies for our experimental data across the four treatments and Nash predictions on the simplex. The numbers next to the large braces report p-values for a bootstrapped Hotelling test for compositional data (Tsagris et al. 2017). The dashed line coincides with the one on the right panel of Figure 3. The dot clouds are proxies for confidence intervals and are obtained by simulating behavior by all individuals as detailed in Section 4.3.

be "shorter", and for higher levels of q it would be "longer". The larger red dot on the curve illustrates the ending point for $q = 0.5$. Finally, note that the dashed line passes through the predictions of the symmetric mixed equilibrium of our model.

To obtain a feeling of the distance between the predicted and observed distribution of outcomes we use simulations. In each simulation, a pair is drawn 100 times with $q = 0.5$ (red dots) or $q = 0.8$ (gray dots) and the two individuals play as follows: dominant types always take the action; for non-dominant types we fix a level for the probability of taking the action $p \in \{0, .01, .02, ..., 1\}$ and run the simulation 100 times for each given level. For the simulations where the probability of taking the action equals the one predicted by the symmetric BNE we color the dots green (when $q = 0.5$) and blue (when $q = 0.8$). The colored "clouds" of simulated outcome distributions form pseudo-"confidence intervals" –CI clouds– for the outcome distribution expected in the experiment.

A first observation is that Figure 5 visually confirms what is indicated in Table 4.

Namely, the observed outcomes are far from the corresponding symmetric Nash predictions. Nevertheless, in the treatments in which $q = 0.5$ it cannot be ruled out, based on our simulations, that subjects are using a symmetric strategy, as in both cases the distribution of outcomes is "within the CI cloud" surrounding the dashed line. The same cannot be said for the two treatments where $q = 0.8$. In these treatments the outcome distributions are too far to the right of the dashed line.

The area on the right of the dashed line represents the expected frequencies of coordination outcomes when individuals adopt *asymmetric strategies*. To see why this is the case, consider scenario where half of the non-dominant types take the action with probability $p + \alpha$ and the other half does so with probability $p - \alpha$, for some $\alpha \in [0, p]$. Whatever the value of α, the probability that a random individual takes the action is always the same, namely $(1 - q) + q \times p$. However, it is easily verified that for a fixed p, the probability of efficient coordination is increasing in α, while the probability of the other two outcomes is decreasing.

The above suggests that the observed outcome distributions, especially in the $q = 0.8$ treatments, may be the result of subjects employing asymmetric strategies. In Figure 6, we plot the distribution of the frequencies with which non-dominant types take the action, across treatments. Differences are apparent and in line with those suggested by Figure 5. Observe first the top two panels of Figure 6 that refer to the low probability treatments ($q = 0.5$). The distribution is essentially unimodal, with a great mass of non-dominant types adopting a homogeneous behavior of never taking the action. Instead, in the high probability treatments ($q = 0.8$), shown in the lower two panels, the distribution is bimodal. A mass of non-dominant types always takes, and another mass of them never takes the action; indicating an, arguably, more heterogeneous behavior.

These differences are compatible with the predictions of the asymmetric role-playing equilibria described in Propositions 3 and Corollary 1. Recall from the experimental design that in each round subjects were randomly matched with another subject from a subgroup of four. Therefore, our experimental setup corresponds to our modeling

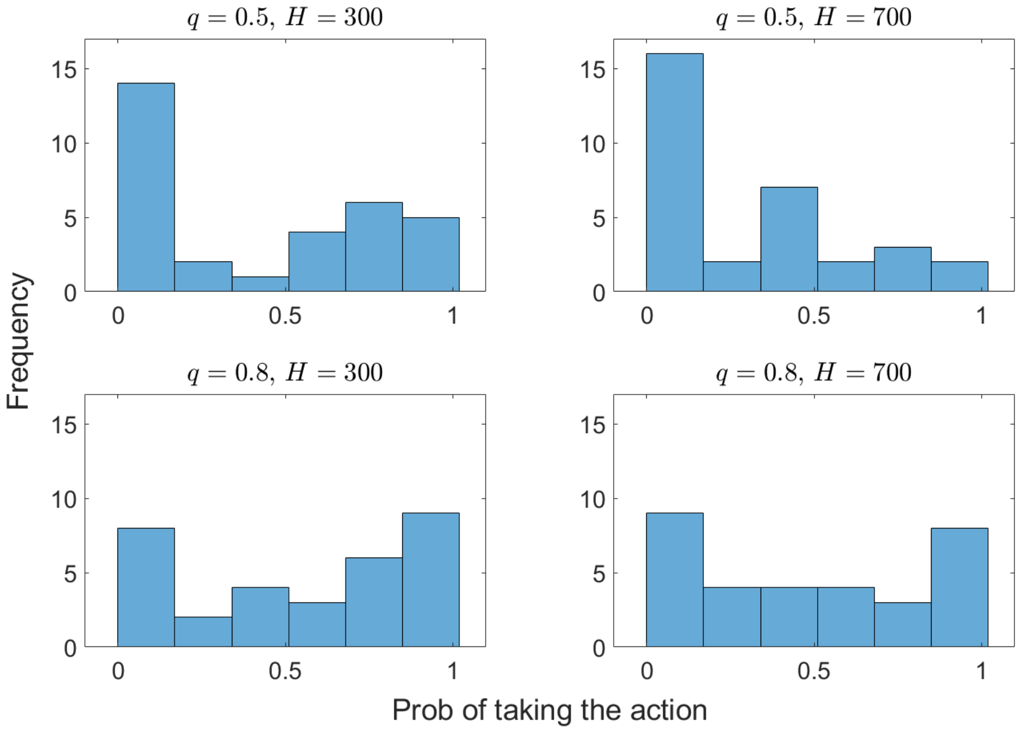


Figure 6: Histograms show the distribution of individual frequencies for non-dominant types taking the action, by treatment. There are 32 individuals in each treatment.

environment with parameter values $N = 4$, $n = 2$ and $t_1 = t_2 = 1$. For these values, Corollary 1 shows that in each of the treatments there is some role-playing equilibrium in which one or two individuals take the action also when assigned a non-dominant type, while the rest do not. More specifically, in the treatments in which non-dominant types are more likely to appear (i.e., $q = 0.8$) the role-playing equilibrium predicts a higher number of individuals who would take the action even when assigned a non-dominant type, in order to decrease the likelihood of a coordination failure. This pattern of behavior is largely observed in Figure 6, which highlights the differences between the unimodal distributions in the low probability treatments, and the bimodal distributions in the high probability treatments.

4.4 Learning

The experiment lasted for 100 rounds, so there was ample opportunity for subjects to adjust their behavior and learn how to play. In this section we take a closer look at possible learning patterns.

At the aggregate level, there seems to be very little change across the experiment. Figure 8 in the appendix shows how the average frequency of taking the action evolved across rounds in each treatment. It is clear that these frequencies remain quite stable. This seems to indicate that subjects do not change their behavior throughout the experiment. Taking a closer look at individual behavior reveals a more nuanced picture.

In Figure 7 we again plot the distribution of the frequencies with which non-dominant types take the action, across treatments, only now we do it separately for every 20 rounds of the experiment. In the first 20 rounds in all treatments, most subjects are mixing to some degree between taking the action or not. Already in rounds 21 to 40, a majority of subjects have switched to some form of role-playing behavior, where they almost always either take the action or not. In later rounds this role-playing behavior becomes even more pronounced.

To summarize, we observe the treatment effects on individual behavior to be present from the start, and all in line with the comparative statics predicted by theory. Still, subjects transition from an initial phase of mostly mixing (rounds 1-20) to adopting role-playing behavior in later rounds. So, while behavior does not get closer to equilibrium, it does move away from mixing and closer to role-playing.

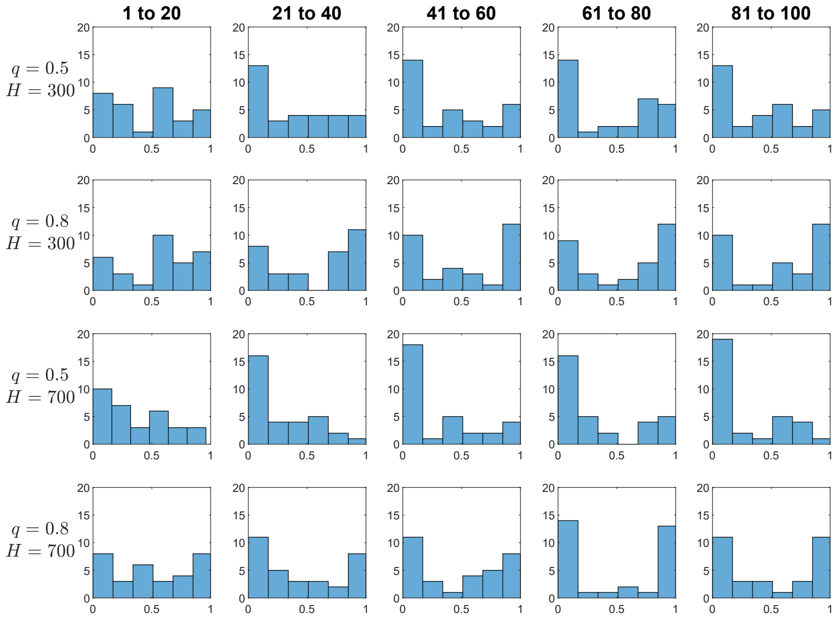


Figure 7: Histograms show the distribution of individual frequencies for non-dominant types taking the action, by treatment, in each 20 round interval. There are 32 individuals in each treatment.

5 Conclusions

We examine a novel coordination problem that combines features of both coordination and anti-coordination games in the presence of incomplete information. The comparative statics obtained by the theoretical analysis do a good job in capturing how individuals react to changes in the game's parameters. Of course, the theoretical predictions do not match the experimental results perfectly. One result that we did not anticipate and is worth highlighting is the fact that subjects in the experiment are able to achieve higher payoffs in the treatments with more uncertainty about their opponents strategy. The behavior of dominant types is almost perfectly predictable, so one might expect things to be easier for a non-dominant type when it is more likely for the opponent to be a dominant type. We find the opposite to be true. This seems to be related to the superiority of role-playing behavior, which subjects did adopt, versus symmetric mixing. The positive effect on payoffs then becomes more salient in the environment with more non-dominant types that role-play.

In the context of protest voting à la Myatt (2017), which provided the first inspiration for this paper, this result issues additional caveats regarding protest campaigns: fringe candidates may gain not only by fanaticising protest voters as suggested by the formal analysis, but also by expanding their popular base. As it was found, such strategy, improves not only the chances of a successful protest, but also of an electoral accident where the alternative preferred by the majority of voters loses against an inferior opponent.

Role playing is interesting beyond the specific treatment effect. The existence of role-playing equilibria that can increase every player's payoff compared to the symmetric mixed strategy equilibrium is in fact a special feature of our game, which only became apparent after conducting the experiment. In the context of teams, such equilibria provide a rationale for specialization that is unrelated to comparative advantage, namely that of a coordination device. For instance, in teams that need to act fast roles are predetermined, not necessarily on the basis of ability, but mainly because it cuts down on time-consuming communication. This is why a combat unit maintains a watch rotation. While this insight is not novel, we are not aware of other simple games that feature such equilibria. Our game might therefore prove useful in theoretical or experimental studies of organizations.

References

Alvarez, R. M., D. R. Kiewiet, and L. Núñez (2018): "A taxonomy of protest voting," *Annual Review of Political Science*, 21, 135–154.

Baliga, S. and T. Sjöström (2004): "Arms races and negotiations," *Review of Economic Studies*, 71, 351–369.

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

Bouton, L. and M. Castanheira (2012): "One person, many votes: Divided majority and information aggregation," *Econometrica*, 80, 43–87.

Bouton, L., M. Castanheira, and A. Llorente-Saguer (2017): "Multicandidate elections: Aggregate uncertainty in the laboratory," *Games and Economic Behavior*, 101, 132–150.

Cabrales, A., R. Nagel, and R. Armenter (2007): "Equilibrium selection through incomplete information in coordination games: An experimental study," *Experimental Economics*, 10, 221–234.

Camerer, C. and D. Lovallo (1999): "Overconfidence and excess entry: An experimental approach," *American Economic Review*, 89, 306–318.

Camerer, C. F. (2011): *Behavioral game theory: Experiments in strategic interaction*, Princeton university press.

Chierchia, G., R. Nagel, and G. Coricelli (2018): "'Betting on nature' or 'betting on others': Anti-coordination induces uniquely high levels of entropy," *Scientific reports*, 8, 1–11.

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

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

Duffy, J. and E. Hopkins (2005): "Learning, information, and sorting in market entry games: theory and evidence," *Games and Economic Behavior*, 51, 31–62.

Erev, I. and A. Rapoport (1998): "Coordination, "magic," and reinforcement learning in a market entry game," *Games and Economic Behavior*, 23, 146–175.

Fey, M. (1997): "Stability and coordination in Duverger's law: A formal model of preelection polls and strategic voting," *American Political Science Review*, 91, 135–147.

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

Forsythe, R., R. B. Myerson, T. A. Rietz, and R. J. Weber (1993): "An experiment on coordination in multi-candidate elections: The importance of polls and election histories," *Social Choice and Welfare*, 10, 223–247.

Greiner, B. et al. (2004): "The online recruitment system orsee 2.0-a guide for the organization of experiments in economics," *University of Cologne, Working paper series in economics*, 10, 63–104.

Heinemann, F., R. Nagel, and P. Ockenfels (2004): "The theory of global games on test: Experimental analysis of coordination games with public and private information," *Econometrica*, 72, 1583–1599.

Hertwig, R., G. Barron, E. U. Weber, and I. Erev (2004): "Decisions from experience and the effect of rare events in risky choice," *Psychological Science*, 15, 534–539.

Jelnov, A., Y. Tauman, and C. Zhao (2020): "Stag Hunt with unknown outside options," *Economic Theory*, 1–33.

Kaplan, T. R. and B. J. Ruffle (2012): "Which way to cooperate," *Economic Journal*, 122, 1042–1068.

Kselman, D. and E. Niou (2011): "Protest voting in plurality elections: a theory of voter signaling," *Public Choice*, 148, 395–418.

Mehta, J., C. Starmer, and R. Sugden (1994): "The nature of salience: An experimental investigation of pure coordination games," *The American Economic Review*, 84, 658–673.

Myatt, D. P. (2017): "A theory of protest voting," *Economic Journal*, 127, 1527–1567.

Myerson, R. and R. Weber (1993): "A theory of voting equilibria," *American Political Science Review*, 87, 102–114.

Ochs, J. (1990): "The coordination problem in decentralized markets: An experiment," *Quarterly Journal of Economics*, 105, 545–559.

Palfrey, T. (1989): "A mathematical proof of Duverger's law," *Models of Strategic Choice in Politics*, 69–92.

Rapoport, A., D. A. Seale, I. Erev, and J. A. Sundali (1998): "Equilibrium play in large group market entry games," *Management Science*, 44, 119–141.

Saran, R. and N. Tumennasan (2013): "Whose opinion counts? Implementation by sortition," *Games and Economic Behavior*, 78, 72–84.

——— (2019): "Eliciting and Aggregating Information By Sortition in Collective Choice," *The Economic Journal*, 129, 1924–1952.

Selten, R. and W. Güth (1982): "Equilibrium point selection in a class of market entry games," in *Games, economic dynamics, and time series analysis*, Springer, 101–116.

Sundali, J. A., A. Rapoport, and D. A. Seale (1995): "Coordination in market entry games with symmetric players," *Organizational behavior and Human decision processes*, 64, 203–218.

Tsagris, M., S. Preston, and A. T. Wood (2017): "Nonparametric hypothesis testing for equality of means on the simplex," *Journal of Statistical Computation and Simulation*, 87, 406–422.

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

——— (1991): "Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games," *Quarterly Journal of Economics*, 106, 885–910.

Weidenholzer, S. (2010): "Coordination games and local interactions: A survey of the game theoretic literature," *Games*, 1, 551–585.

Zwick, R. and A. Rapoport (2002): "Tacit coordination in a decentralized market entry game with fixed capacity," *Experimental Economics*, 5, 253–272.

A Proofs

Proof of Proposition 1: Let us first look at the behavior of individuals who are assigned a dominant type. The expected utility an arbitrary such individual gets by taking the action is equal to $EU^d(1) = (1) \times \left[\sum_{k=t_1-1}^{n-1} \binom{n-1}{k} r^k (1-r)^{n-1-k} \right]$ and by not taking the action it is equal to $EU^d(0) = (1) \times \left[\sum_{k=t_1}^{n-1} \binom{n-1}{k} r^k (1-r)^{n-1-k} \right]$, where r is the probability that a random individual takes the action in a symmetric strategy profile. Thus, $EU^d(1) \geq EU^d(0)$ for all $1 \leq t_1 \leq n-1$ and for all $r \in [0,1]$. Therefore, whenever an individual is assigned a dominant type, taking the action is a best response to any beliefs and in all but the two cases described below it is actually the unique best response.

Not taking the action when assigned a dominant type could only be a best response when (i) all other individuals are expected to not take the action with certainty irrespective of their type, i.e. $r = 0$, or when (ii) all other individuals are expected to take the action with certainty irrespective of their type, i.e. $r = 1$. In both of these cases, an equilibrium in which the said individual puts a positive probability in not taking the action when being assigned a dominant type would either be non-responsive –case (i)– or asymmetric –case (ii). Namely, in case (i), the unique symmetric equilibrium in which dominant types put positive probability to not taking the action is such that all individuals choose with certainty not to take the action irrespective of their type. This would not be responsive. On the other hand, in case (ii) all other individuals would choose to take the action irrespective of their type, thus a strategy that would put positive probability in not taking the action would violate symmetry.

Hence, let dominant types take the action with certainty. By symmetry, we have that any equilibrium is essentially characterized by a single number, p, which denotes the probability with which an individual takes the action when assigned a non-dominant type. For the equilibrium to also be responsive, this probability, p, should be different than one. This rules out the pure strategy equilibrium where all individuals take the action irrespective of their type.

The probability p is strictly between zero and one only if an individual who is assigned a non-dominant type is indifferent between taking the action or not. If r denotes again the probability that a random individual takes the action, then in equilibrium q, r and p should satisfy that $r = (1-q) + q \times p$, or equivalently $1 - r = q \times (1-p)$. The expected utility obtained by an individual who is assigned a non-dominant type if she takes the action is:

$$EU^{nd}(1) = (1+h) \times \left[\sum_{k=t_1-1}^{t_2-1} \binom{n-1}{k} r^k (1-r)^{n-1-k} \right] + (1) \times \left[\sum_{k=t_2}^{n-1} \binom{n-1}{k} r^k (1-r)^{n-1-k} \right]$$

while her expected utility by not taking the action is

$$EU^{nd}(0) = \begin{cases} (1+h) \times \left[\sum_{k=t_1}^{t_2} \binom{n-1}{k} r^k (1-r)^{n-1-k} \right] + (1) \times \left[\sum_{k=t_2+1}^{n-1} \binom{n-1}{k} r^k (1-r)^{n-1-k} \right] & \text{if } t_2 \leq n-2 \\ (1+h) \times \left[\sum_{k=t_1}^{n-1} \binom{n-1}{k} r^k (1-r)^{n-1-k} \right] & \text{if } t_2 = n-1 \end{cases}$$

Hence, for a non-dominant type to be indifferent between taking the action or not the following equality must hold:

$$EU^{nd}(1) = EU^{nd}(0) \Leftrightarrow (1+h) \binom{n-1}{t_1-1} r^{t_1-1} (1-r)^{n-t_1} = h \binom{n-1}{t_2} r^{t_2} (1-r)^{n-1-t_2}$$

Solving with respect to r we get that

$$r = \frac{1}{1 + \left[\frac{h}{1+h} \frac{(t_1-1)!(n-t_1)!}{t_2!(n-t_2-1)!} \right]^{\frac{1}{t_2-t_1+1}}}$$

And given that $r = (1-q) + q \times p$, we can solve for p to obtain:

$$p(h,q) = \frac{\frac{1}{1+\left[\frac{h}{1+h} \frac{(t_1-1)!(n-t_1)!}{t_2!(n-t_2-1)!} \right]^{\frac{1}{t_2-t_1+1}}} - 1 + q}{q} = \frac{q - \tilde{q}(h)}{q}$$

where $\tilde{q}(h) := 1 - \frac{1}{1+\left[\frac{h}{1+h} \frac{(t_1-1)!(n-t_1)!}{t_2!(n-t_2-1)!} \right]^{\frac{1}{t_2-t_1+1}}}$.

Finally, notice that a mixed-strategy equilibrium exists if and only if $p(h,q) \in (0,1)$. The condition $p(h,q) < 1$ is trivially satisfied because $\tilde{q}(h) > 0$ for all admissible parameter values. On the other hand, the condition $p(h,q) > 0$ requires that $q > \tilde{q}(h)$. Thus, a mixed equilibrium exists whenever $q > \tilde{q}(h)$. When $q < \tilde{q}(h)$ we have that $EU^{nd}(1) < EU^{nd}(0)$ for every $p \in [0,1]$, and, hence, the unique equilibrium that satisfies our properties is the pure one in which non-dominant types do not take the action with certainty. The same holds for $q = \tilde{q}(h)$, as in this case $EU^{nd}(1) = EU^{nd}(0)$ for $p = 0$ and strictly negative for any $p > 0$.

Proof of Proposition 2: We only need to provide a proof for pairs (h,q) for which $q > \tilde{q}(h)$. For these pairs we have that $p(h,q) = \frac{q-\tilde{q}(h)}{q}$, where $\tilde{q}(h)$ is independent of q. Thus, we obtain the following results:

$$\frac{\partial p(h,q)}{\partial q} = \frac{\tilde{q}(h)}{q^2} > 0 \qquad \text{and} \qquad \frac{\partial^2 p(h,q)}{\partial q^2} = -2\frac{\tilde{q}(h)}{q^3} < 0$$

The derivatives of $p(h,q)$ with respect to h are as follows:

$$\frac{\partial p(h,q)}{\partial h} = -\frac{1}{q}\tilde{q}'(h) = -\frac{[1-\tilde{q}(h)]^2}{q}\left(\left[\frac{h}{1+h}\frac{(t_1-1)!(n-t_1)!}{t_2!(n-t_2-1)!}\right]^{\frac{1}{t_2-t_1+1}}\right)' =$$

$$= -\frac{[1-\tilde{q}(h)]^2}{q(t_2-t_1+1)}\left[\frac{(t_1-1)!(n-t_1)!}{t_2!(n-t_2-1)!}\right]^{\frac{1}{t_2-t_1+1}}\left[\frac{h^{\frac{1}{t_2-t_1+1}-1}}{(1+h)^{\frac{1}{t_2-t_1+1}+1}}\right] < 0$$

$$\frac{\partial^2 p(h,q)}{\partial h^2} = -\frac{1}{q(t_2-t_1+1)}\left[\frac{(t_1-1)!(n-t_1)!}{t_2!(n-t_2-1)!}\right]^{\frac{1}{t_2-t_1+1}}\left([1-\tilde{q}(h)]^2\left[\frac{h^{\frac{1}{t_2-t_1+1}-1}}{(1+h)^{\frac{1}{t_2-t_1+1}+1}}\right]\right)' > 0$$

Because $[1-\tilde{q}(h)]^2 > 0$, $\left[\frac{h^{\frac{1}{t_2-t_1+1}-1}}{(1+h)^{\frac{1}{t_2-t_1+1}+1}}\right] > 0$, $([1-\tilde{q}(h)]^2)' = -2[1-\tilde{q}(h)]\tilde{q}'(h) < 0$ and $\left[\frac{h^{\frac{1}{t_2-t_1+1}-1}}{(1+h)^{\frac{1}{t_2-t_1+1}+1}}\right]' = \frac{h^{\frac{1}{t_2-t_1+1}-2}}{(1+h)^{\frac{1}{t_2-t_1+1}+2}}\left[\frac{1}{t_2-t_1+1} - 1 - 2h\right] < 0$ given that $t_2 \geq t_1$.

Moreover, for these pairs it holds that $r(h,q) = 1 - q + q \times p(h,q)$. Thus,

$$\frac{\partial r(h,q)}{\partial q} = \frac{\partial[1-q+q\times p(h,q)]}{\partial q} = -1 + p(h,q) + q\frac{\partial p(h,q)}{\partial q} = -1 + \frac{q-\tilde{q}(h)}{q} + q\frac{\tilde{q}(h)}{q^2} = 0$$

$$\frac{\partial r(h,q)}{\partial h} = \frac{\partial[1-q+q\times p(h,q)]}{\partial h} = q\frac{\partial p(h,q)}{\partial h} < 0 \text{ and } \frac{\partial^2 r(h,q)}{\partial h^2} = q\frac{\partial^2 p(h,q)}{\partial h^2} > 0$$

Proof of Proposition 3: In a role-playing equilibrium each individual takes the action when assigned a dominant type, K of the N individuals take the action also when assigned a non-dominant type, and the remaining $N-K$ individuals do not take the action when assigned a non-dominant type.

For this strategy profile to be an equilibrium, the following two conditions should hold simultaneously: (1) an individual who is assigned a non-dominant type and believes that K of the $N-1$ remaining individuals will take the action when assigned a non-dominant type prefers not to take the action, and (2) an individual who believes that $K-1$ of the $N-1$ remaining individuals will take the action when assigned a non-dominant type prefers to take the action when assigned herself a non-dominant type.

Therefore, let us denote by $\Delta EU^{nd}(K)$ the difference in the expected utility of an arbitrary individual between taking and not taking the action when assigned a non-dominant type and conditional on making a payoff-relevant choice, when she expects K of the remaining individuals to take the action when assigned a non-dominant type and all individuals to take the action when assigned a dominant type. Given this, the equilibrium conditions can be summarized as follows:

$$\Delta EU^{nd}(K) \leq 0 \text{ and } \Delta EU^{nd}(K-1) \geq 0$$

In fact, we can calculate the exact form of $\Delta EU^{nd}(K)$, which is the following:

$$\Delta EU^{nd}(K) = (1+h)\left[\sum_{\lambda=\max\{0,K+n-N\}}^{\min\{K,t_1-1\}} \frac{\binom{K}{\lambda}\binom{N-1-K}{n-1-\lambda}}{\binom{N-1}{n-1}} \binom{n-1-\lambda}{t_1-1-\lambda} q^{n-t_1}(1-q)^{t_1-1-\lambda}\right] - h\left[\sum_{\lambda=\max\{0,K+n-N\}}^{\min\{K,t_2\}} \frac{\binom{K}{\lambda}\binom{N-1-K}{n-1-\lambda}}{\binom{N-1}{n-1}} \binom{n-1-\lambda}{t_2-\lambda} q^{n-1-t_2}(1-q)^{t_2-\lambda}\right]$$

When the maximum value of λ in a sum is smaller than its minimum, the value of the sum to be equal to zero.

The first (resp., second) part of each expression describes the increase (resp., decrease) in the individual's expected utility when by taking the action she increases the number of individuals who take the action from t_1-1 to t_1 (t_2 to t_2+1). The index λ in each sum denotes how many of the individuals who take the action even when assigned a non-domintant type are selected to make a payoff-relevant choice.

Let us first show that a role-playing equilibrium exists if $q \geq \tilde{q}(h)$. By simple substitution, we can get that $\Delta EU^{nd}(0) \geq 0$ if and only if $q \geq \tilde{q}(h)$, with equality holding for $q = \tilde{q}(h)$. Moreover, for any $K \in \{N-n+t_1, \ldots, N-n+t_2\}$ we must have $\Delta EU^{nd}(K) < 0$. The latter holds because for these values of K the first part of the expression is equal to zero, whereas the second part is different than zero and, in fact, negative. Therefore, there exists some $\widetilde{K} \in \{1, \ldots, N-n+t_1\}$ such that $\Delta EU^{nd}(\widetilde{K}-1) \geq 0$ and $\Delta EU^{nd}(\widetilde{K}) \leq 0$. Thus, there is a role-playing equilibrium in which $\widetilde{K}$ individuals take the action when assigned a non-dominant type.

Let us now show that that no role-playing equilibrium exist when $q < \tilde{q}(h)$. To do that, let us slightly reshape $\Delta EU^{nd}(K)$.

We start by defining $F(K,\lambda,t) := \frac{\binom{K}{\lambda}\binom{N-1-K}{n-1-\lambda}}{\binom{N-1}{n-1}}\binom{n-1-\lambda}{t-\lambda}q^{n-1-t}(1-q)^{t-\lambda}$. Given this definition, observe that $F(K,\lambda,t_2) = \left[\frac{(t_1-1-\lambda)!}{(t_2-\lambda)!}\frac{(n-t_1)!}{(n-t_2-1)!}\left(\frac{1-q}{q}\right)^{t_2-t_1+1}\right]F(K,\lambda,t_1-1)$. Then we can rewrite $\Delta EU^{nd}(K)$ as follows:

$$
\begin{aligned}
&\Delta EU^{nd}(K) = \\
&= (1+h)\sum_{\lambda=\max\{0,K+n-N\}}^{\min\{K,t_1-1\}} F(K,\lambda,t_1-1) - h\sum_{\lambda=\max\{0,K+n-N\}}^{\min\{K,t_2\}} F(K,\lambda,t_2) = \\
&= \sum_{\lambda=\max\{0,K+n-N\}}^{\min\{K,t_1-1\}} \left[(1+h)F(K,\lambda,t_1-1) - hF(K,\lambda,t_2)\right] - h\sum_{\lambda=\min\{K,t_1-1\}+1}^{\min\{K,t_2\}} F(K,\lambda,t_2) = \\
&= \sum_{\lambda=\max\{0,K+n-N\}}^{\min\{K,t_1-1\}} h\frac{(n-t_1)!}{(n-t_2-1)!}\left(\frac{1-q}{q}\right)^{t_2-t_1+1} F(K,\lambda,t_1-1)\left[\frac{1+h}{h}\frac{(n-t_2-1)!}{(n-t_1)!}\left(\frac{q}{1-q}\right)^{t_2-t_1+1} - \frac{(t_1-1-\lambda)!}{(t_2-\lambda)!}\right] - \ldots \\
&\qquad \cdots - h\sum_{\lambda=\min\{K,t_1-1\}+1}^{\min\{K,t_2\}} F(K,\lambda,t_2)
\end{aligned}
$$

Observe that, the second part of the expression is different than zero for $K > t_1 - 1$ only, in which cases it is negative. Moreover, note that $\frac{(t_1-1-\lambda)!}{(t_2-\lambda)!}$ is strictly increasing in λ, given that $t_2 \geq t_1$. Thus, given that all the other factors are positive, if

$\left[\frac{1+h}{h}\frac{(n-t_2-1)!}{(n-t_1)!}\left(\frac{q}{1-q}\right)^{t_2-t_1+1} - \frac{(t_1-1-\lambda)!}{(t_2-\lambda)!}\right] < 0$ for $\lambda = 0$, it will also be negative for all $\lambda > 0$, which means that the whole expression will be negative. This factor is negative for $\lambda = 0$ precisely whenever $q < \tilde{q}(h)$, irrespective of K. Hence, if $q < \tilde{q}(h)$ it holds that $\Delta EU^{nd}(K) < 0$ for all relevant K and, thus, the game has no role-playing equilibrium.

Calculations of Corollary 1: First, observe that for $n = 2$ and $t_1 = t_2 = 1$, we get that $\tilde{q}(h) := \frac{1}{1+2h}$. Thus, $\tilde{q}(1) = 1/3$ and $\tilde{q}(5) = 1/11$, which means that in all four cases there is at least one role-playing equilibrium for some $K \in \{1, 2, 3\}$, i.e. in all four cases $\Delta EU^{nd}(0) > 0$ and $\Delta EU^{nd}(3) < 0$. Then,

$$\Delta EU^{nd}(1) = \frac{2}{3}\left[(1+h)q - h(1-q)\right] - \frac{1}{3}h$$
$$\Delta EU^{nd}(2) = \frac{1}{3}\left[(1+h)q - h(1-q)\right] - \frac{2}{3}h$$

Therefore, simple substitutions give us the following equilibria:

For $(h, q) = (1, 0.5)$: $\Delta EU^{nd}(1) = 0$ and $\Delta EU^{nd}(2) = -1/2 < 0 \Rightarrow K \in \{1, 2\}$.

For $(h, q) = (1, 0.8)$: $\Delta EU^{nd}(1) = 3/5 > 0$ and $\Delta EU^{nd}(2) = -1/5 < 0 \Rightarrow K \in \{2\}$.

For $(h, q) = (5, 0.5)$: $\Delta EU^{nd}(1) = -4/3 < 0$ and $\Delta EU^{nd}(2) = -19/6 < 0 \Rightarrow K \in \{1\}$.

For $(h, q) = (5, 0.8)$: $\Delta EU^{nd}(1) = 13/15 > 0$ and $\Delta EU^{nd}(2) = -16/15 < 0 \Rightarrow K \in \{2\}$.

B Outcome simulations

In this section we give some more details on the construction and interpretation of Figure 5 and the underlying simulations.

As we note in the main text, the dashed line represents the expected outcome distribution if all individuals adopt a symmetric, un-correlated strategy (not necessarily a BNE). The expected outcome distribution for the symmetric BNE's will therefore also lie on this line. In Figure 5, these are marked by the black square and triangle for $\{q, 300\}$ and $\{q, 700\}$ respectively.

To move into the area to the left of the dashed line it would be necessary for individuals to coordinate and adopt some type of (positively) correlated strategy. Since there are no means of communication and coordination between subjects, correlated strategies cannot emerge. In fact, we do not observe outcome distributions in that area, so this type of strategies are not discussed any further.

Of course, the points on the dashed line are theoretical predictions. Even if all individuals do adopt the same strategy, the actual distribution of outcomes for a given sample will generally not lie on the line. The larger the size of the sample, the closer to the theoretical prediction we expect the sample distribution to be. The question then is, given the sample size from our experiment, how far from the theoretical prediction can the sample distribution be to not reject it? In other words, what is the "confidence interval"? Calculating exact confidence intervals for the multivariate distribution of outcomes is beyond the scope of the paper. Instead we use Monte Carlo simulations.

In the simulations we use the same sample size as in the experiment: 16 pairs playing 100 rounds. In each run of the simulation we fix the parameter q, which determines whether a player is dominant or not, to either .5 or .8. If a player in a given round is dominant she always takes the action. If she is non-dominant, she takes the action with some probability that remains fixed for all players in all rounds of the run. We repeat each run with the same constellation of parameters 100 times. Each small colored dot in Figure 5 represents the outcome of one run.

First we run the simulation with players adopting one of the two BNE strategies. The green dots creating the "confidence cloud" around the black Nash {p,300} square in the zoomed-in right panel of Figure 5 are the results obtained in the 100 runs of the simulation for $q = .5$ and non-dominants taking the action with the equilibrium probability $p^*(.5, 300) = .333$ (the value for p^* is shown in Table 3). As can be seen in the figure, the sample distribution for $\{q = .5, H = 300\}$, indicated by a green square, lies outside the corresponding "confidence clouds". Hence, the observed data is unlikely to be the result of subjects in the experiment playing the symmetric BNE. The same is true for all treatments and the corresponding confidence clouds (blue corresponds to $q = .8$, squares and triangles correspond to $H = 300$ and $H = 700$ respectively).

Next we look at whether it is possible for subjects to be playing a symmetric strategy, even if it is not a BNE. For that we do 100 runs each time with $q \in \{.5, .8\}$ and $p \in \{0, .01, .02, ..., 1\}$. The red (gray) dots represent the results from runs with $q = .5$ ($q = .8$). In the region of interest, the two confidence clouds (red and gray) essentially coincide. The red and grey clouds also overlap with the green and blue clouds, as the BNE's are a subset of the set of symmetric strategies.

It can be seen in Figure 5 that the sample outcome distribution for the two treatments with $q = .5$ do lie within the red "confidence cloud". Hence, we cannot reject the notion that subjects in these treatments adopt a symmetric strategy. Nonetheless, as we saw, this strategy is probably not the symmetric BNE strategy and is also likely different in the two treatments.

On the other hand, the sample outcome distribution for the two treatments with $q = .8$ do lie outside of the gray "confidence cloud". It is therefore unlikely for them to be the result of some symmetric strategy. As we explain in the main text, outcome distributions to the right of the dashed line can be the result of players adopting asymmetric strategies.

C Learning across rounds

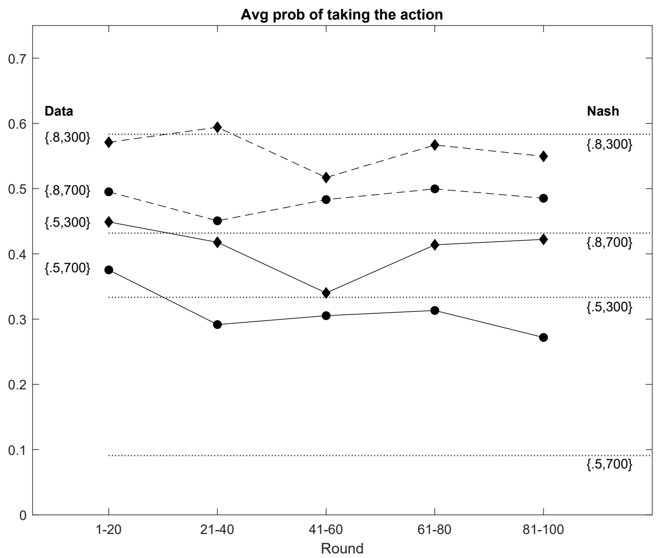


Figure 8: The markers indicate the average frequency of taking the action over 20 rounds in each treatment.As a benchmark, the straight dotted lines indicate the expected frequency for the symmetric BNE's.