Overcoming Inefficient Lock-in in Coordination Games with Sophisticated and Myopic Players

Aidas Masiliūnas

Global Asia Institute, National University of Singapore,
10 Lower Kent Ridge Road, Singapore 119076

Abstract

Path-dependence in coordination games may cause lock-in on inefficient outcomes, such as inferior technologies (Arthur, 1989) or inefficient economic institutions (North, 1990). To calculate the conditions under which lock-in can be overcome, we develop a solution concept that makes ex-ante predictions about the adaptation process following lock-in in a critical mass game. We assume that some players are myopic, forming beliefs according to weighted fictitious play, while others are sophisticated, anticipating the learning process of the myopic players. We propose a solution concept based on a Nash equilibrium of the strategies chosen by sophisticated players. Our model predicts that no players would switch from the efficient to the inefficient action, but deviations in the other direction are possible. Three types of equilibria may exist: in the first type lock-in is sustained, while in the other two types lock-in is overcome. We determine the existence conditions for each of these equilibria and show that the equilibria in which lock-in is overcome are more likely and the transition is faster when sophisticated players have a longer planning horizon, or when the history of inefficient coordination is shorter.

Keywords: game theory, learning, lock-in, farsightedness, coordination
JEL classification: C73, D83

Email address: aidas.masiliunas@gmail.com (Aidas Masiliūnas)

1. Introduction

Research in natural and social sciences identified the presence of alternative stable states in many important settings.[1] The main problem in these situations is the emergence of an inefficient state,[2] or "inefficient lock-in", from which no individual has incentives to deviate. Lock-in has been identified as the primary cause of inefficient social customs (Akerlof, 1980), inefficient economic and political institutions (North, 1990) and inefficient technologies (Shapiro and Varian, 1999; Cowan, 1990; David, 1985). For example, employees might tolerate misconduct, overwork or use inefficient procedures and technologies as long as sufficiently many others engage in such behavior. Lock-in might be prevented, but policy makers often deal with situations in which lock-in has already occurred, and it is important to know how it can be overcome. This paper studies whether lock-in can be endogenously overcome in a population of sophisticated and myopic agents and identifies how the likelihood and the speed of such transitions depend on the parameters of the game.

No previous literature has addressed the question of whether the presence of multiple sophisticated players can help overcome inefficient lock-in. Game theory can make predictions about whether lock-in will occur (Harsanyi and Selten, 1988, Kandori et al., 1993, Young, 1993), but not whether it can be overcome, because the predictions of standard solution concepts are invariant to the history of play. The question of how to anticipate and avoid or facilitate transitions between states has been studied in other disciplines,[3] often using models with adaptive agents from complexity science (Gao et al., 2016, Battiston et al., 2016). However, these models cannot predict whether lock-in will persist if agents are strategic and farsighted, as they are in human societies.

We are able to model and study lock-in through the interaction of two types of players, "myopic" and "sophisticated". Myopic players use adaptive learning: they form beliefs about the actions of other players based on observed history and choose the myopic best-response. Inefficient lock-in is modeled through the beliefs of myopic players, who experienced a history of inefficient coordination. Belief-based learning by myopic players creates incentives for sophisticated players to use "strategic teaching", that is to deviate from the inefficient state to induce a future deviation by myopic players. Sophisticated players anticipate how their actions will affect subsequent beliefs and actions of myopic players, and our solution concept requires the choice path of sophisticated players to be optimal given the choices of all other players. A combination of game-theoretic reasoning and adaptive learning generates predictions about the occurrence and speed of transitions between stable states, in contrast to models with only strategic or only adaptive players,

[1] Examples include ecosystems such as forests (Hirota et al., 2011), coral reefs (Nyström et al., 2000) and lakes (Scheffer et al., 1993); mood (van de Leemput et al., 2014); unemployment (Cooper and John, 1988; Ball and Romer, 1991); political institutions (North, 1990; Greif and Laitin, 2004); bank runs (Diamond and Dybvig, 1983); revolts (Kuran, 1989); technological standards (David, 1985; Arthur, 1989).

[2] We use the term "state" rather than "equilibrium" when referring to the Nash equilibrium of a stage game to avoid confusion with the repeated game equilibrium.

[3] Scheffer et al. (2009), Scheffer et al. (2012), Folke et al. (2004).

which cannot make such predictions.[4] We show that three types of outcomes are possible on the equilibrium path: sophisticated players might deviate from the inefficient state immediately, after a delay, or they might never deviate. No player ever deviates from the efficient state, therefore inefficient lock-in is overcome in the first two cases, but not in the third. We specify the existence conditions and the speed of transition for each equilibrium type, obtaining testable predictions about the behavioral patterns[5] and about the comparative statics.[6]

Models with adaptive and sophisticated players have been used in the literature to address issues other than lock-in. Camerer et al. (2002) and Chong et al. (2006) use a sophisticated experience-weighted attraction (EWA) model in which some players are adaptive (as in Camerer and Ho, 1999), while others are sophisticated, anticipate the behavior of adaptive players and use strategic teaching. Sophisticated EWA can be used to explain why sophisticated borrowers repay loans to adaptive lenders in early periods to secure loans in later periods of a repeated trust game. A simplified version of sophisticated EWA is used by Brandts et al. (2016) to explain why players choose to lower the effort cost of other group members in a minimum effort game following coordination failure. The learning model in Brandts et al. (2016) assumes that unsophisticated types best-respond to the distribution of actions observed in the previous round, while sophisticated types expect all others to be unsophisticated, therefore efficient coordination in future rounds can be facilitated by lowering the effort cost of other group members. Hyndman et al. (2009) explain behavior in two-player coordination games using a model that combines adaptive players, who follow weighted fictitious play, with farsighted players, who anticipate the learning process and maximize the discounted sum of expected payoffs.[7]

Our model rests on the concept of strategic teaching, just as the sophisticated EWA and other models discussed above, but we use a different approach for a different purpose and therefore make three important contributions to the literature. First, we use strategic teaching to refine rather than alter the predictions of standard solution concepts. Strategic teaching has been previously used to explain deviations from the Nash equilibrium (e.g. repaying loans in early periods, as in Camerer et al., 2002, Chong et al., 2006, or helping others at a cost to oneself, Brandts et al., 2016). In the setting we study, standard solution concepts generate vacuous predictions, therefore we use strategic teaching to refine the paths of play that can be supported in equilibrium. As a result, we develop the first game-

[4]Solely adaptive models (see Fudenberg and Levine, 1998, Camerer, 2003) predict no deviations from the inefficient state once it has been implemented. At the other extreme, Nash equilibrium and its refinements are invariant to the history of play, therefore there will always be some repeated game equilibria in which lock-in is overcome, and others in which lock-in persists.

[5]For example, only efficiency-enhancing transitions are predicted to occur; sophisticated players will deviate from the inefficient state earlier than myopic players.

[6]For example, a longer history of inefficient coordination and a shorter planning horizon should increase the likelihood that lock-in will persist, while a larger number of sophisticated players should speed up the transition to the efficient state.

[7]Terracol and Vaksmann (2009) and Hyndman et al. (2012) use similar models to study strategic teaching in different games.

theoretic solution concept that can make predictions about the likelihood and timing of transitions between stable states. Second, we address the problem of strategic uncertainty, devising a method to apply models of strategic teaching to games with multiple strategic agents. Previously, the scope of applications was limited because the existing models bypassed the problem of strategic uncertainty by assuming that sophisticated players expect all opponents to be myopic.[8] We show that the issue of strategic uncertainty can be resolved using standard game-theoretic tools, resulting in a model that is both more flexible and more realistic, and therefore could potentially fit data better. Third, the models in the previous literature can explain experimental data only ex-post, while we make testable ex-ante predictions. The characterization of the solution concept allows us to study its properties, such as the speed of convergence, and the conditions under which each equilibrium type exists, generating novel testable predictions. In the previous literature, ex-ante predictions could be obtained only using simulations with specific parameter values (e.g. Chong et al., 2006).

Compared to the previous models of strategic teaching, we make different assumptions about the behavior of the two types of players, because stronger assumptions are needed to obtain a tractable model. First, we characterize the sophisticated player equilibrium only for critical mass games. Our intention is not to devise a solution concept that explains behavioral regularities in different games (although it could be characterized in other games), but rather a model that can explain how and when inefficient lock-in can be overcome. Second, we assume that myopic players update beliefs using weighted fictitious play (just as in Hyndman et al., 2009, and Ellison, 1997) instead of EWA (Camerer et al., 2002). We use a belief-based updating rule because under the assumptions made here, sophisticated players cannot directly influence the payoffs of the myopic players; therefore, if myopic players used a payoff-based learning rule (such as reinforcement learning, Erev and Roth, 1998), strategic teaching would have no effect. It has also been found that belief-based models can explain behavior in experiments with coordination games,[9] and they typically fit better than other models (Ho and Weigelt, 1996, Battalio et al., 1998). Weighted fictitious play is sufficiently rich to accurately explain actions and beliefs in critical mass game experiments (Masiliūnas, 2017), but it is also sufficiently simple, allowing us to set up a tractable model of strategic teaching. Other updating rules might be more attractive from a descriptive viewpoint (such as EWA, Camerer and Ho, 1999) or a normative viewpoint (standard Bayesian updating using a Dirichlet prior), but they could not be used to characterize equilibria with the methods used in this study.

Other papers study the interaction between farsighted and myopic players using different methods and in different games, and are therefore less related to our study than the strategic teaching literature. Ellison (1997) models a population of adaptive players,

[8] Sometimes this assumption was made explicitly (e.g. Brandts et al., 2016), other times only two-player games were considered (e.g. Camerer et al., 2002, Hyndman et al., 2009)

[9] Examples include pure coordination games (Van Huyck et al., 1997); coordination games with Pareto-ranked equilibria (Sefton, 1999; Cheung and Friedman, 1997); critical mass games (Masiliūnas, 2017); order statistic games (Crawford, 1995).

who learn according to fictitious play and are repeatedly paired to play a binary choice coordination game. Adding one rational player to the population of adaptive players can change the outcome from coordination on the inefficient equilibrium to coordination on the efficient one, as long as the number of players is fixed and the rational player is sufficiently patient. Acemoglu and Jackson (2014) develop an overlapping generations model that shows how a social norm of low cooperation can be overturned by a single forward-looking player. Schipper (2019) uses an optimal control model with two players and shows how a strategic player can control an adaptive player in repeated games with strategic substitutes or strategic complements. Mengel (2014) studies adaptive players who are also forward-looking and finds that in two-player coordination games the efficient equilibrium may be stochastically stable, in contrast to the case with only adaptive players. Models studied in these papers do not deal with strategic uncertainty, and therefore could not be applied to games with multiple sophisticated players.

The rest of the paper is organized as follows. Section 2 provides a general definition of our solution concept, the "sophisticated player equilibrium". The entire section 3 constructs the characterization of this solution concept in a critical mass game, which is defined in section 3.1. Section 3.2 specifies the behavior of myopic players, conditional on the observed history. Section 3.3 refines the behavior of sophisticated players, under the assumptions that are discussed in section 3.4. Section 3.5 shows that the efficient state is absorbing, and is implemented when myopic players deviate from the inefficient state. Section 3.6 specifies this switching time of myopic players, as well as the payoffs received by sophisticated players. Section 3.7 characterizes the three types of equilibria. Section 3.8 illustrates the theoretical results using a numerical example. Section 3.9 shows how the speed of transition and the types of equilibria that exist respond to changes in game parameters. Section 4 concludes.

2. Sophisticated Player Equilibrium

Consider N players, indexed by $i \in \mathcal{N} \equiv \{1, 2, \ldots, N\}$, who play a repeated game of duration T in continuous time. The duration of the game could be determined by length of the interaction, or by the length of the planning horizon of sophisticated players.[10] At each point in time $t \in [0, T]$ players choose an action from a stage game action space $\{A, B\}$. The two actions can be though of as two competing technologies, with coordination on A being more efficient than on B.[11] We model inefficient lock-in by assuming that all players were playing action B prior to time 0. The duration of this history is $T_h \in (0, \infty)$.

[10]We will typically refer to T as the length of the planning horizon, because in many situations the actual duration of the interaction is unknown or uncertain. Player's behavior therefore depends on the subjective belief about the duration of interaction, or the planning horizon that the player can take into account. Even when the game duration is known, as in laboratory experiments, players often do not plan ahead for the entire duration of the game (e.g. see Johnson et al., 2002, Mantovani, 2016).

[11]For the lack of a better term, we will sometimes refer to action A as an "efficient" and to B as an "inefficient" action. These labels refer to the efficiency of the state that would be implemented if the action was chosen by all players.

Table 1: Main variables that describe the choices of sophisticated and myopic players.

Variable	**Sophisticated**	**Myopic**
Number of players	S	$N-S$
Characteristic member	$s \in \mathcal{S}$	$m \in (\mathcal{N} \setminus \mathcal{S})$
Choice plan/response function $t \to \{A, B\}$	$c_s \in \mathcal{C}$	$r_m(\mathbf{c}_S)$
Vector of mappings for all players of this type	$\mathbf{c}_S \equiv \{c_s \| s \in S\} \in \mathcal{C}^S$	$\mathbf{r}_M(\mathbf{c}_S) \equiv \{r_m \| m \in (N \setminus S)\}$
Action of player s/m at time t	$c_s(t) \in \{A, B\}$	$r_m(t, \mathbf{c}_S) \in \{A, B\}$
Action profile of all players of this type in t	$\mathbf{c}_S(t) \in \{A, B\}^S$	$\mathbf{r}_M(t, \mathbf{c}_S) \in \{A, B\}^{N-S}$
Choice plan of all sophisticated players besides s	$\mathbf{c}_{-s} \in \mathcal{C}^{S-1}$	–
Action profile of all sophisticated players besides s	$\mathbf{c}_{-s}(t) \in \{A, B\}^{S-1}$	–
Action profile of all other players	$\{\mathbf{c}_{-s}(t), \mathbf{r}_M(t, \mathbf{c}_S)\}$	$\mathbf{r}_{-m}(t, \mathbf{c}_S) \in \{A, B\}^{N-1}$
Number of all other players who choose A in t	$\alpha(\{\mathbf{c}_{-s}(t), \mathbf{r}_M(t, \mathbf{c}_S)\})$	$\alpha(\mathbf{r}_{-m}(t, \mathbf{c}_S))$
Belief of myopic player m in t	–	$x_m(t) \in [0, 1]$
Subjective expected payoff of playing action a in t	–	$SEP(a, x_m(t))$
Realized payoff in t	$\pi[c_s(t), \{\mathbf{c}_{-s}(t), \mathbf{r}_M(t, \mathbf{c}_S)\}]$	$\pi[\mathbf{r}_m(t, \mathbf{c}_S), \mathbf{r}_{-m}(t, \mathbf{c}_S)]$

We study the population composed of S sophisticated and $N - S$ myopic players. Sophisticated players are indexed by $s \in \mathcal{S}$ and myopic players by $m \in (\mathcal{N} \setminus \mathcal{S})$. The two types of players follow different choice rules, respectively denoted by r_m and c_s, which prescribe an action for each moment in time. We will refer to r_m as a *response function* of a myopic player m and to c_s as a *choice plan* of a sophisticated player s. Denote the action of player m at time t by $r_m(t)$ and the action of player s by $c_s(t)$. Denote the combination of choice plans for all sophisticated players except s by $\mathbf{c}_{-s}$ and the combination of choice plans of all sophisticated players, or a choice plan profile, by $\mathbf{c}_S$, with $\mathbf{c}_S(t)$ denoting the action profile of sophisticated players at time t. Denote the combination of response functions of all myopic players by $\mathbf{r}_M(\mathbf{c}_S)$, with $\mathbf{r}_M(t, \mathbf{c}_S)$ denoting the action profile of myopic players at time t.[12] Denote the action profile of all players besides m by $\mathbf{r}_{-m}(t, \mathbf{c}_S)$. Define a function that counts the number of players who choose A in an action profile $\mathbf{c}'$ by $\alpha(\mathbf{c}') = |c \in \mathbf{c}' : c = A|$. The response functions of myopic players are determined by the history of play, while the choice plans of sophisticated players must be optimal given the choices of all other players. The payoff function π maps player's action and the vector of the actions of all other players into a payoff. The payoff at time t is $\pi[\mathbf{r}_m(t, \mathbf{c}_S), \mathbf{r}_{-m}(t, \mathbf{c}_S)]$ for a myopic player and $\pi[c_s(t), \{\mathbf{c}_{-s}(t), \mathbf{r}_M(t, \mathbf{c}_S)\}]$ for a sophisticated player. Table 1 displays the notation of choice plans, response functions and the spaces of these functions and variables.[13]

Belief of a myopic player is a probability assigned to the event that a randomly chosen other group member chooses action A. Denote the belief of a myopic player m at time t by $x_m(t)$. Belief formation is assumed to follow a one parameter weighted fictitious

[12] All myopic players share the same response function and observed history, therefore their choices will be identical at any point in time.

[13] We denote a vector containing element a repeated k times by $\{(a)_{\times k}\}$, and the set of vectors of length j containing all possible combinations of elements A and B by $\{A, B\}^j$.

play model,[14] originally proposed by Cheung and Friedman (1997) and adjusted to fit the N-person games studied here.[15] Beliefs are formed according to the following rule:

$$x_m(t) = \frac{\int_{k=0}^{t} \gamma^k \frac{\alpha(\mathbf{r}_{-m}(t-k,\mathbf{c}_S))}{N-1}\,\mathrm{d}k}{\int_{k=0}^{t+T_h} \gamma^k\,\mathrm{d}k} \tag{1}$$

The integral in the numerator measures the weighted frequency of action A choices by all other players. When $k = 0$, the integrand in the numerator reduces to $\frac{\alpha(\mathbf{r}_{-m}(t,\mathbf{c}_S))}{N-1}$, the relative frequency of action A played at time t. As k increases, observed relative frequencies in more distant past are multiplied by the discount factor γ. The summation runs backwards to time 0. Prior to time 0, only action B has been played, therefore this history does not increase the frequency of observed action A. The denominator normalizes the expression to the range between 0 and 1, thus the belief can be interpreted as a weighted average of the relative frequency of action A.

The γ parameter measures the rate at which old observations are forgotten. We assume that $\gamma \in (0, 1)$, where values close to 1 indicate that all past observations receive similar weights, while values close to 0 indicate that only the most recent experience is taken into account. Weighted fictitious play therefore generalizes two popular belief learning models: fictitious play, in which beliefs are proportional to the empirical frequency of the actions played over all past periods ($\gamma = 1$), and Cournot best-response, in which beliefs are proportional to the empirical frequency in the last round ($\gamma = 0$).

Subjective expected payoff is the payoff that a myopic player expects to receive by playing action $a \in \{A, B\}$, conditional on the player's expectations about the behavior of other group members, measured by the subjective belief $x_m(t)$. Subjective expected payoff is equal to the realized payoff if beliefs are correct, but this will generally not be the case in the short run because beliefs are formed from the observed history of play. To calculate the subjective expected payoff, player's beliefs are used to assign a probability to each action

[14] Fictitious play corresponds to Bayesian updating of the probability that any group member will choose A, using a Dirichlet prior and assuming that the choice of each group member was independently drawn from the distribution about which players are learning.

[15] In the original weighted fictitious play, beliefs about the likelihood that the opponent will choose a given action are formed by counting the empirical frequency of this action played in the past. We extend the model that was initially specified for two-player games to N-person games by assuming that a joint distribution of choices is used to form beliefs about the actions of group members, but players do not distinguish between the identities of others. Beliefs are therefore homogeneous (Rapoport, 1985; Rapoport and Eshed-Levy, 1989): a single belief is formed about the probability that any other player will choose A. There also are other ways how weighted fictitious play could be extended to N-person games. One way could be to assume that players form beliefs about the joint distribution of the actions of all others and update it using observed aggregate feedback: for example, Crawford (1995) assumes that players form beliefs and observe feedback about an order statistic of all the choices. Another way is to assume that separate beliefs are formed about every other player j based on the empirical distribution of j's choices (e.g. Monderer and Shapley, 1996). We combine the two approaches by assuming that players use the joint distribution of choices to form beliefs about the action of any opponent, but do not distinguish between their identities.

profile of other group members, and payoffs at all possible action profiles are summed up using these probabilities as weights:

$$SEP(a, x_m(t)) = \sum_{\mathbf{c}' \in \{A,B\}^{N-1}} Pr[\mathbf{r}_{-m}(t, \mathbf{c}_S) = \mathbf{c}' | x_m(t)] \pi(a, \mathbf{c}') =$$
$$= \sum_{\mathbf{c}' \in \{A,B\}^{N-1}} x_m(t)^{(\alpha(\mathbf{c}'))} (1 - x_m(t))^{(N-1-\alpha(\mathbf{c}'))} \pi(a, \mathbf{c}') \qquad (2)$$

Response function $r_m(t, \mathbf{c}_S)$ prescribes an action for a myopic player m at any point in time $t \in [0, T]$, conditional on the profile of choice plans chosen by sophisticated players, $\mathbf{c}_S$. We assume that myopic players choose the action that maximizes the flow of subjective expected payoff and ties are broken in favour of action A:

$$r_m(t, \mathbf{c}_S) = \begin{cases} A & \text{if } SEP(A, x_m(t)) \geq SEP(B, x_m(t)) \\ B & \text{otherwise} \end{cases} \qquad (3)$$

Choice plans of sophisticated players ($\mathbf{c}_S$) are explicitly included in the response function to make it transparent that myopic player actions can be affected by sophisticated players. Note that the response function depends only on the current round payoffs and beliefs, which are determined by observed history, therefore it is possible to anticipate the behavior of myopic players at any history. Since all myopic players observe the same history and form beliefs using the same weighted fictitious play rule, they will have the same beliefs and therefore take the same actions.

Sophisticated players anticipate the learning process of myopic players and plan for the entire game in advance, choosing the choice plan for the interval $[0, T]$. Variable T measures the perceived duration of the game, which can be limited by the objective duration of the game, or by the ability of players to plan ahead.

Choice plan c_s prescribes an action for a sophisticated player s at any point in time $t \in [0, T]$. Denote the set of all choice plans by $\mathcal{C}$. The choice plan is assumed to be an open-loop strategy, which depends only on time and not on observed history.[16] Sophisticated players face no strategic uncertainty about the actions of myopic players, but they face uncertainty about the actions of the other sophisticated players. Payoffs associated with a choice plan c_s depend on the vector of choice plans of the other sophisticated players, $\mathbf{c}_{-s}$, and on the response function of myopic players, $\mathbf{r}_M(t, \mathbf{c}_S)$, which also depends on the choice plans of all sophisticated players.

[16] Traditionally, a strategy is assumed to be closed-loop, prescribing an action for each information set. If strategies are defined this way, then almost any sequence of actions could be supported on the equilibrium path in coordination games such as the one analyzed here. This occurs because all players could use strategies that prescribe playing the efficient equilibrium only if a particular sequence of actions has been observed previously. Note that sophisticated players are assumed to know about the decision-making process of the myopic players, and they therefore can anticipate beforehand what these choices will be for each combination of sophisticated player strategies.

Since the sophisticated players can perfectly anticipate the actions of myopic players, the game can be reduced to a static game between sophisticated players. Nash equilibrium is the standard solution concept in static games, and we follow the logic of Nash equilibrium by requiring equilibrium choice plans to be mutual best-responses.

Definition 1. *A combination of choice plans* $\mathbf{c}_S^* = \{(c_s^*)_{\times S}\}$ *is a symmetric* ***sophisticated player equilibrium*** *if for each player* $s \in S$*,* c_s^* *satisfies:*

$$\int_0^T \pi[c_s^*(t), \{\mathbf{c}_{-s}^*(t), \mathbf{r}_M(t, \{c_s^*, \mathbf{c}_{-s}^*\})\}] \, \mathrm{d}t \geq \int_0^T \pi[c_s(t), \{\mathbf{c}_{-s}^*(t), \mathbf{r}_M(t, \{c_s, \mathbf{c}_{-s}^*\})\}] \, \mathrm{d}t \tag{4}$$
$$\forall c_s \in \mathcal{C}, \quad \mathbf{c}_{-s}^* = \{(c_s^*)_{\times(S-1)}\}, \; r_m(t, \mathbf{c}_S) \text{ is defined in } (3)$$

If there were no myopic players, equation (4) would reduce to the standard Nash equilibrium. If all players were myopic, equation (4) would not be needed because the choices of all players would be calculated using weighted fictitious play. We will look at an intermediate case where both myopic and sophisticated players are present.

The assumption of two distinct types of players, as well as the specific assumptions about their behavior, is motivated by previous experimental evidence. Evidence for the dichotomy between myopic and sophisticated types is found in a turnaround game, where a sharp distinction is observed between "leaders", who initiate the transition, and "laggards", who follow leaders (Brandts and Cooper, 2006, Brandts et al., 2007); in a planning task, where participants either plan completely, or do not plan at all (Hey and Knoll, 2007); and in a critical mass game, where farsighted players are much more likely to deviate from the inefficient convention and to spend more time thinking ahead (Masiliūnas, 2017). Evidence for the existence of sophisticated players, who engage in strategic teaching, has been observed in repeated two player games with two Pareto-ranked Nash equilibria (Hyndman et al., 2009), three non Pareto-ranked equilibria (Terracol and Vaksmann, 2009) and a unique equilibrium (Hyndman et al., 2012). Additional evidence for strategic teaching in coordination games comes from increased rates of efficient coordination when the shadow of the future is increased (Berninghaus and Ehrhart, 1998), or when the cost to disclose one's action is decreased (Masiliūnas, 2017). Strategic teaching has also been proposed as a likely explanation for deviations from the inefficient state in studies that observe such deviations (Brandts and Cooper, 2006, Devetag, 2003, Friedman, 1996).

3. Sophisticated Player Equilibria in a Critical Mass Game

This section will characterize the symmetric sophisticated player equilibria for a repeated N-person critical mass coordination game. We are particularly interested in whether there are equilibria in which lock-in is overcome, and if so, how long this process takes.

The construction of the sophisticated player equilibrium involves several steps. Proposition 1 shows that myopic players will choose the efficient action if their beliefs exceed a certain threshold. Furthermore, in the sophisticated player equilibrium myopic players will switch from an inefficient to the efficient action at most once. The single switch and the assumption that there are sufficiently many myopic players means that the efficient state is absorbing, therefore the switching time is the only information needed for sophisticated players to calculate their payoffs. Proposition 3 shows exactly how the switching time of myopic players can be calculated if beliefs were formed using weighted fictitious play. The switching time depends on the strategies taken by sophisticated players, which could prescribe many switches from one action to the other. The task of specifying the switching time is therefore greatly simplified by Proposition 2, which shows that only the sophisticated player strategies prescribing at most one switch from the inefficient to the efficient action survive the elimination of strictly dominated strategies, allowing a strategy to be identified by the switching time.

The ability to anticipate the speed of a transition allows sophisticated players to calculate how their payoffs depend on their own strategies and on the strategies chosen by the other sophisticated players. The mapping from strategies to payoffs specified in Corollary 1 is used to identify strategy profiles in which all sophisticated players are best-responding to each other. Three types of symmetric equilibria are possible: sophisticated players may play the efficient action right away, they may switch to the efficient action later or they may never switch. In the first two cases myopic players eventually start playing the efficient action, while in the third case all players always choose the inefficient action. Which types of equilibria exist and how long a transition to the efficient state takes depends on the game parameters, as specified in Propositions 4, 5 and 6. Finally, Corollaries 2, 3 and 4 show how these existence conditions depend on the history of inefficient coordination, length of the planning horizon of sophisticated players and the player composition.

3.1. Payoffs in a Critical Mass Game

Recall that we defined a sophisticated player equilibrium for a class of games with n players and an action space $\{A, B\}$. A special case in this game class is a critical mass game, in which payoffs depend on player's action, denoted by a, and on the total number of other group members who chose action A, denoted by $\alpha(\mathbf{c}')$, where $\mathbf{c}'$ is the action profile of all other players:[17]

$$\pi(a, \mathbf{c}') = \begin{cases} H & \text{if } \alpha(\mathbf{c}') \geq \theta \text{ and } a = A \\ 0 & \text{if } \alpha(\mathbf{c}') < \theta \text{ and } a = A \\ M & \text{if } \alpha(\mathbf{c}') \geq \theta \text{ and } a = B \\ L & \text{if } \alpha(\mathbf{c}') < \theta \text{ and } a = B \end{cases} \tag{5}$$

We require $H > M$ and $L > 0$ for the game to belong to the class of coordination games. The coordination requirement is determined by an exogenous threshold θ: payoff from choosing A exceeds the payoff from B if at least θ other group members choose A. The stage game contains two stable states in pure strategies: in one all players choose A, in the other all players choose B. We assume that states are Pareto-ranked by requiring

[17] The payoff flow of myopic and sophisticated players, calculated using the payoff function from equation (5) and the choice variables of both player types, is displayed in table 1.

that $H > L$. Finally, we assume that $M \geq L$, so that players who choose B also prefer a situation in which the threshold is exceeded.

Assumption 1: H > M ≥ L > 0.

In addition, we assume that there are at least 2 sophisticated players so that the equilibrium could be defined using equation (4). We also assume that the number of myopic players is sufficiently large to implement the efficient state, and the number of sophisticated players is small enough so that sophisticated players on their own could not implement the efficient state.

Assumption 2: $2 \leq S < \theta \leq N - S$.

The importance of these two assumptions will be discussed in section 3.4.

3.2. Response Function of Myopic Players

Myopic players form beliefs about the actions of other players and choose the action that maximizes immediate payoffs. This subsection specifies this response function $r_m(t, \mathbf{c}_S)$, which prescribes an action for player m at time t when sophisticated players are using choice plans $\mathbf{c}_S$.

Proposition 1. *Suppose that in a game with payoffs defined by* (5) *at time t myopic player m holds beliefs $x_m(t)$. Then the response function defined in* (3) *is equivalent to:*

$$r_m(t, \mathbf{c}_S) = \begin{cases} A & \text{if } x_m(t) \geq I^{-1}_{\frac{L}{L+H-M}}(\theta, N-\theta) \\ B & \text{otherwise} \end{cases} \tag{6}$$

where I^{-1} is the inverse of an incomplete regularized beta function.

Proof: see Appendix B.1.

Proposition 1 states that a myopic player chooses A instead of B if the belief exceeds $I^{-1}_{\frac{L}{L+H-M}}(\theta, N-\theta)$, a threshold value that depends only on the game parameters. For brevity, we will refer to this threshold value by I^{-1}. The properties of inverse regularized beta functions imply that I^{-1} is increasing in L, M and θ, but decreasing in H and N.

Since the number of myopic players exceeds θ (Assumption 2) and they choose the same action, the efficient state is implemented when beliefs exceed the threshold. The next subsection shows that once exceeded, the threshold remains exceeded, therefore the efficient state is absorbing. To anticipate the state, it is thus sufficient to know the first time when beliefs reach the threshold value.

3.3. Undominated Choice Plans of Sophisticated Players

Although sophisticated players could use choice plans that prescribe many switches from one action to the other, we will show that undominated choice plans must prescribe at most one switch from B to A and no switches from A to B. The strategy space of the sophisticated players can therefore be restricted to a set of real numbers that denote a switching time from A to B.

Definition 2. *Denote by U_s (for "undominated") the set of choice plan profiles in which no sophisticated player is choosing strictly dominated choice plans:*

$$U_s = \{\mathbf{c}_S \in \mathcal{C}^S | \nexists c'_s \in \mathcal{C}, \nexists c''_s \in \mathbf{c}_S : \int_0^T \pi[c'_s(t), \{\mathbf{c}_{-s}(t), \mathbf{r}_M(t, \{c'_s, \mathbf{c}_{-s}\})\}] \, \mathrm{d}t > \int_0^T \pi[c''_s(t), \{\mathbf{c}_{-s}(t), \mathbf{r}_M(t, \{c''_s, \mathbf{c}_{-s}\})\}] \, \mathrm{d}t, \forall \mathbf{c}_{-s} \in \mathcal{C}^{S-1}\}$$

A choice plan profile is *dominated* if it is not in set U_s, that is if in this choice plan profile at least one sophisticated player is choosing a dominated choice plan.

We will show that the set of undominated choice plans cannot contain any strategies that prescribe a switch from A to B. The proof requires two additional lemmas.

Lemma 1. *If two choice plans of the sophisticated player prescribe the same action at time t, the payoff flow is higher for the choice plan that induced higher beliefs of myopic players:*

$$\pi[c'_s(t), \{\mathbf{c}_{-s}(t), \mathbf{r}_M(t, \{c'_s, \mathbf{c}_{-s}\})\}] \geq \pi[c''_s(t), \{\mathbf{c}_{-s}(t), \mathbf{r}_M(t, \{c''_s, \mathbf{c}_{-s}\})\}] \quad \text{if} \quad x'(t) \geq x''(t) \quad \text{and} \quad c'_s(t) = c''_s(t)$$

where $x'(t)$ is the belief held by myopic players if the sophisticated player uses choice plan c'_s and $x''(t)$ is the belief if the sophisticated player uses choice plan c''_s.

Proof: see Appendix B.7.

Lemma 1 shows that sophisticated players can only benefit from myopic players assigning a higher probability to others choosing A. The proof rests on a finding that higher beliefs can only increase the number of myopic players choosing A, which can only increase the payoffs of sophisticated players (from Assumption 1).

Definition 3. *Denote by AB_M the set of choice plan profiles for sophisticated players with which myopic players switch from A to B:*

$$AB_M = \{\mathbf{c}_S \in \mathcal{C}^S | \exists t_1, t_2 \in [0, T] : \quad t_1 < t_2, \quad r_m(t_1, \mathbf{c}_S) = A, \quad r_m(t_2, \mathbf{c}_S) = B\}$$

Lemma 2. *All choice plan profiles for sophisticated players with which myopic players switch from A to B are strictly dominated:*

$$AB_M \cap U_s = \emptyset$$

Proof: see Appendix B.7.

Lemma 2 shows that if sophisticated players choose undominated choice plans, myopic players would never switch from A to B. If a switch from A to B was to occur, all myopic players would switch at the same time, because they share the observed history. For the switch to occur, the belief must fall below the threshold value, therefore at least one sophisticated player must be playing B right before the switch – otherwise beliefs would remain above the threshold. But the choice plan of this sophisticated player would be dominated, because the player could earn more by instead choosing A before the switch (playing A generates a payoff flow of H because $N - S \geq \theta$, from Assumption 2, and cannot decrease future payoffs, from Lemma 1).

Definition 4. *Denote by AB_S the set of choice plan profiles for sophisticated players with which at least one sophisticated player switches from A to B:*

$$AB_S = \{\mathbf{c}_S \in \mathcal{C}^S | \exists c_s \in \mathbf{c}_S, t_1, t_2 \in [0, T]: \quad t_1 < t_2, \quad c_s(t_1) = A, \quad c_s(t_2) = B\}$$

Proposition 2. *Choice plan profiles for sophisticated players that prescribe a switch from A to B for at least one sophisticated player are dominated:*

$$AB_S \cap U_s = \emptyset$$

Proof: see Appendix B.2.

The logic of the proof is as follows: suppose that at some point in time the sophisticated player switches from A to B. Such a switch could not be optimal if myopic players switched from B to A at the same time or earlier, because then the sophisticated player could increase payoffs by never switching to B. If myopic players never switch to A, the sophisticated player would be better off always playing B. If myopic players switch after the sophisticated player, payoffs could be strictly increased by teaching less at the start of the game and teaching more later. Doing so would not reduce the payoffs prior to the switch, but would strictly decrease the switching time of the myopic players, because weighted fictitious play puts more weight on recent experience. Consequently, sophisticated players would never switch from A to B because they would be better off delaying strategic teaching to just before the predicted switch of myopic players.

3.4. Assumptions

Proposition 2 is the key step towards the characterization of the sophisticated player equilibrium, reducing the space of undominated sophisticated player strategies to only those that prescribe a single switch from B to A. We will shortly review the importance of Assumptions 1 and 2, which are necessary for the proof of Proposition 2.

Assumption 1: $H > M \geq L > 0$. Restrictions $H > M$ and $L > 0$ are needed to set up the critical mass game. The key assumption is $M \geq L$, implying that players who choose B either prefer the number of A players to exceed the threshold, or are indifferent to it.[18] This assumption is necessary to create incentives for strategic teaching. Lemma 1 would not hold without Assumption 1, because then higher myopic player beliefs could increase the number of players choosing A and therefore decrease the payoffs of the sophisticated players who choose B. Without Lemma 1, we could not prove Lemma 2 because it would not be possible to construct a choice plan that would dominate a choice plan with which myopic players switch from A to B. Choice plans constructed following the procedure in the proof of Lemma 2 increase the beliefs of myopic players, and thus might reduce the payoffs of sophisticated players who use such choice plans. If such choice plans cannot be constructed, it cannot be claimed that choice plans that induce myopic players to switch from A to B

[18]Critical mass games are commonly used to study riots and political revolts (Granovetter, 1978; Edmond, 2013; Kiss et al., 2017; Masiliūnas, 2017), where this assumption usually holds, as the citizens who do not take part in the revolt are better off when the regime is overthrown.

are dominated. Identical problems would arise in the proof of Proposition 2, since the proof relies on constructing choice plans that dominate choice plans with a switch from A to B. The dominance may not hold because the newly constructed choice plans increase the beliefs of myopic players and could therefore decrease the payoffs of sophisticated players.

Assumption 2: $2 \leq S < \theta \leq N - S$. The key part of this assumption is that sophisticated players cannot on their own achieve a transition from one state to the other: state is inefficient if all myopic players choose B (because $S < \theta$), and efficient if they choose A (because $N - S \geq \theta$). Sophisticated players can therefore affect the state only indirectly, through strategic teaching. If $S \geq \theta$, sophisticated players could directly implement the efficient state and our solution concept would reduce to the standard Nash equilibrium. If $N - S < \theta$, sophisticated players could directly implement the transition from the efficient to the inefficient state, and such a transition could be supported in equilibrium. This would not only increase the number of equilibria, but also permit sophisticated players to change actions multiple times (Proposition 2 would no longer hold), therefore their strategies could not be described by a single switching time. This can no longer occur when the number of myopic players exceeds the threshold, because then the transition to the inefficient state would require myopic players to change their actions from A to B, which takes some time and sophisticated players could increase payoffs by playing A over this period.

Overall, the failure of either Assumption 1 or Assumption 2 would inhibit the current approach of reducing the choice plans to those with a single switching point, and therefore the equilibria could not be characterized with the methods used here.

3.5. Single-switching Choice Plans and Response Functions

Lemma 2 and Proposition 2 show that if sophisticated players do not choose dominated choice plans, both myopic and sophisticated players will switch from B to A at most once, thus the equilibrium choice plans and response functions can be described by a scalar indicating the switching time.

Each choice plan of the sophisticated players must have the following structure on the equilibrium path:

$$c_s(t) = \begin{cases} B & \text{if } t \in [0, y_s) \\ A & \text{if } t \in [y_s, T] \end{cases} \qquad \forall \mathbf{c}_S \in U_s$$

Sophisticated players therefore only choose how long to delay strategic teaching, a variable defined as $y_s \in [0, T]$, which will be called a **strategy** of player s. If $y_s = T$, player s never uses strategic teaching.

In principle, to characterize all sophisticated player equilibria we would need to calculate the sum of payoffs from the entire game for all possible strategy profiles. However, since we only characterize symmetric equilibria, it is sufficient to specify the payoffs in such equilibria and in possible unilateral deviations from it. The sum of payoffs from the entire game will be denoted by $\Pi(y, \bar{y})$, where the player's strategy is $y_s = y$ and the strategies of the other sophisticated players are $\mathbf{y}_{-s} = \{(\bar{y})_{\times(S-1)}\}$. This payoff depends on the strategies of the sophisticated players, but also on the **switching period** of the myopic players, which is a function of the strategies of sophisticated players.

The switching period of myopic players will be denoted by $\hat{t}(y,\bar{y}) \in (0,\infty]$, defined as the time when the myopic players start playing A.[19] Lemma 2 shows that on the equilibrium path myopic players must switch from B to A at most once, therefore the response function must have the following structure:

$$r_m(t,\mathbf{c}_S) = \begin{cases} B & \text{if } t \in [0,\hat{t}(y,\bar{y})) \\ A & \text{if } t \in [\hat{t}(y,\bar{y}),T] \end{cases} \qquad \forall \mathbf{c}_S \in U_s$$

The payoff of the sophisticated players can be expressed as a function of the strategies of the sophisticated players:

$$\Pi(y,\bar{y}) = \int_0^T \pi(\hat{c}(t,y),\{(\hat{c}(t,\bar{y}))_{\times(S-1)},(\hat{c}(t,\hat{t}(y,\bar{y})))_{\times(N-S)}\})\,\mathrm{d}t \tag{7}$$

where $\hat{c}(t,k)$ stands for the choice of either a myopic player with a switching period at time k, or a sophisticated player whose choice plan prescribes a switch from B to A at time k:

$$\hat{c}(t,k) = \begin{cases} B & \text{if } t \in [0,k) \\ A & \text{if } t \in [k,T] \end{cases}$$

The next subsection will specify how the switching period of the myopic players and the payoffs of the sophisticated players depend on the game parameters.

3.6. Switching Period of Myopic Players and Payoffs of Sophisticated Players

The switching period of the myopic players is important because it determines the time of transition to the efficient state, and therefore the payoffs of the sophisticated players. The switching period $\hat{t}(y,\bar{y})$ is defined as a function of the strategy of one sophisticated player (y) and all other $(S-1)$ sophisticated players ($\bar{y}$). The switching period is calculated by finding the time at which the beliefs of myopic players reach the threshold value I^{-1}. The specification consists of four cases, depending on the sequence of events: the switching period is calculated by $\hat{t}_1(y,\bar{y})$ if myopic players switch to A after all sophisticated players; by $\hat{t}_2(y)$ if the switch occurs after one but before the other $(S-1)$ sophisticated players; by $\hat{t}_3(\bar{y})$ if the switch occurs after $(S-1)$ but before the last sophisticated player and it is equal to infinity if myopic players never switch.

[19] All myopic players will switch to A at the same time, because they experienced the same history and use the same weighted fictitious play rule to form beliefs. Note that $\hat{t}(y,\bar{y}) > 0$ because equation (1) implies that $x_m(0) = 0$, thus myopic players choose B at time 0. If $\hat{t}(y,\bar{y}) > T$, myopic players will never be observed switching to A.

Proposition 3. *If one sophisticated player uses strategy $y_s = y$ and the other $(S-1)$ players use strategies $\mathbf{y}_{-s} = \{(\bar{y})_{\times(S-1)}\}$, the switching period of myopic players is:*

$$\hat{t}(y,\bar{y}) = \begin{cases} \hat{t}_1(y,\bar{y}) & \text{if } \max\{y,\bar{y}\} < \hat{t}_1(y,\bar{y}) \quad \text{and} \quad \frac{S}{N-1} > I^{-1} \\ \hat{t}_2(y) & \text{if } y < \hat{t}_2(y) \leq \bar{y} \quad \text{and} \quad \frac{1}{N-1} > I^{-1} \\ \hat{t}_3(\bar{y}) & \text{if } \bar{y} < \hat{t}_3(\bar{y}) \leq y \quad \text{and} \quad \frac{S-1}{N-1} > I^{-1} \\ \infty & \text{otherwise} \end{cases} \tag{8}$$

such that

$$\hat{t}_1(y,\bar{y}) = \frac{\log(\frac{S}{N-1} - I^{-1}) - \log(\gamma^{-\bar{y}}\frac{S-1}{N-1} + \gamma^{-y}\frac{1}{N-1} - \gamma^{T_h}I^{-1})}{\log(\gamma)} \tag{9}$$

$$\hat{t}_2(y) = \frac{\log(\frac{1}{N-1} - I^{-1}) - \log(\gamma^{-y}\frac{1}{N-1} - \gamma^{T_h}I^{-1})}{\log(\gamma)} \tag{10}$$

$$\hat{t}_3(\bar{y}) = \frac{\log(\frac{S-1}{N-1} - I^{-1}) - \log(\gamma^{-\bar{y}}\frac{S-1}{N-1} - \gamma^{T_h}I^{-1})}{\log(\gamma)} \tag{11}$$

Proof: see Appendix B.3.

It is never possible that more than one condition of (8) is satisfied because $\hat{t}_1(y,\bar{y}) \leq \hat{t}_2(y)$ and $\hat{t}_1(y,\bar{y}) \leq \hat{t}_3(y)$ (see Lemma 10 in Appendix B).

The specification of the switching period derived in Proposition 3 is necessary to calculate the payoffs of sophisticated players from the entire game. We specify these payoffs in a corollary of Proposition 3.

Corollary 1. *If a sophisticated player s uses strategy $y_s = y$ and the other sophisticated players use strategies $\mathbf{y}_{-s} = \{(\bar{y})_{\times(S-1)}\}$, the total payoff received by player s over period $[0,T]$ is:*

$$\Pi(y,\bar{y}) = \begin{cases} \Pi_1 = yL + (T - \hat{t}_1(y,\bar{y}))H & \text{if } \hat{t}_1(y,\bar{y}) \leq T,\ \hat{t}_2(y) \geq \bar{y},\ \hat{t}_3(\bar{y}) \geq y \quad \text{(12a)} \\ \Pi_2 = yL + (T - \hat{t}_2(y))H & \text{if } \hat{t}_2(y) < \bar{y} \quad \text{(12b)} \\ \Pi_3 = \hat{t}_3(\bar{y})L + (y - \hat{t}_3(\bar{y}))M + (T-y)H & \text{if } \hat{t}_3(\bar{y}) < y \quad \text{(12c)} \\ \Pi_4 = yL & \text{if } \hat{t}_1(y,\bar{y}) > T \quad \text{(12d)} \end{cases}$$

where $\hat{t}_1(y,\bar{y})$, $\hat{t}_2(y)$ and $\hat{t}_3(\bar{y})$ are specified in Proposition 3.

Proof: see Appendix B.5.

3.7. Characterization of Symmetric Sophisticated Player Equilibria

The payoff specification from Corollary 1 transforms the repeated game into a static game between the sophisticated players. We make predictions in this static game using the standard solution concept for static games – Nash equilibrium – which requires mutual best responses for each player. We do so by combining the definition of a symmetric sophisticated player equilibrium in equation (4) with the definition of the payoff function from the entire game in equation (7). A sophisticated player equilibrium is a combination of *strategies* (y^*, y^*) that satisfies:

$$\Pi(y^*, y^*) \geq \Pi(y, y^*), \quad \forall y \in [0, T] \tag{13}$$

We look at the existence of three types of equilibria: interior solutions with $y^* \in (0, T)$, a corner solution with $y^* = 0$ and a corner solution with $y^* = T$. Equilibria are labeled according to the time at which sophisticated players start teaching: "immediate teaching" equilibrium ($y^* = 0$), "delayed teaching" equilibrium ($y^* \in (0, T)$) and "no teaching" equilibrium ($y^* = T$). Propositions 4, 5 and 6 specify the conditions under which each type of equilibrium exists.

3.7.1. "Immediate teaching" equilibrium

In the "immediate teaching" equilibrium all sophisticated players switch to A at the start of the game, so that equilibrium strategies are $y^* = \bar{y}^* = 0$.

Proposition 4. *A combination of strategies* $(0, 0)$ *is a symmetric sophisticated player equilibrium ("immediate teaching" equilibrium) if and only if conditions* ***I1*** *and* ***I2*** *are satisfied:*

$$\frac{S - H/L}{N - 1} \leq \gamma^{T_h} I^{-1}, \quad \textbf{\textit{(I1)}}$$
$$\hat{t}_1(0, 0) \leq T(1 - L/H), \quad \textbf{\textit{(I2)}}$$

Proof: see Appendix B.4.

Conditions **I1** and **I2** are needed to prevent incentives to deviate from the "immediate teaching" equilibrium. **I1** ensures that the derivative at $y = \bar{y} = 0$ is negative, therefore delayed teaching would decrease the payoff. Condition **I2** ensures that players have no incentives to never teach. Compared to immediate teaching, no teaching increases the payoff flow from 0 to L in the period between 0 and $\hat{t}_1(0, 0)$, but decreases it from H to L in the period between $\hat{t}_1(0, 0)$ and T. Condition **I2** ensures that the costs of not teaching exceed the benefits.

3.7.2. "Delayed Teaching" Equilibrium

In the "delayed teaching" equilibrium all sophisticated players switch to A at time $y^* \in (0, T)$.

Proposition 5. *A combination of strategies* (y^*, y^*) *with* $y^* \in (0, T)$ *is a symmetric sophisticated player equilibrium ("delayed teaching" equilibrium) if and only if conditions* ***D1****,* ***D2****,* ***D3*** *and* ***D4*** *are satisfied:*

$$\hat{t}_1(y^*, y^*) < T, \quad \textbf{\textit{(D1)}}$$
$$y^* > 0, \quad \textbf{\textit{(D2)}}$$
$$\hat{t}_1(y^*, y^*) - y^* L/H \leq \hat{t}_2(0), \quad \textbf{\textit{(D3)}}$$
$$\hat{t}_1(y^*, y^*) - y^* L/H \leq T(1 - L/H), \quad \textbf{\textit{(D4)}}$$

where equilibrium strategies are calculated by

$$y^* = \frac{\log\left(\frac{S - H/L}{I^{-1}(N-1)}\right)}{\log(\gamma)} - T_h$$

Proof: see Appendix B.4.

If conditions **D1** to **D4** are satisfied, no sophisticated player will find it profitable to deviate from the "delayed teaching" equilibrium. In equilibrium, players cannot have incentives to deviate to neighboring strategies, therefore the first derivative of the payoff function at y^* should be equal to zero. This condition is used to determine the value of y^*. The "delayed teaching" equilibrium does not exist if the first derivative is non-zero for all $t \in (0, T)$. Conditions **D1** and **D2** ensure that this does not happen, and **D1** additionally ensures that the switching period of myopic players occurs before the end of the game. Condition **D3** ensures that players have no incentives to start teaching immediately. Immediate teaching accelerates the transition, increasing the payoff flow from 0 to H for the period between $\hat{t}_2(0)$ and $\hat{t}_1(y^*, y^*)$, but strategic teaching needs to be used for an additional duration y^*, with a total cost of $y^* L$. Condition **D3** requires the costs of the deviation to immediate teaching to exceed the benefits. Condition **D4** ensures that players have no incentives to never teach. By not teaching, a sophisticated player benefits from a payoff flow of L instead of 0 in the period between y^* and $\hat{t}_1(y^*, y^*)$, but the transition to the efficient state would never occur, reducing the payoff to L instead of H in the period between $\hat{t}_1(y^*, y^*)$ and T. Condition **D4** ensures that the costs of the deviation to no teaching exceed the benefits.

3.7.3. "No Teaching" Equilibrium

In the third type of a symmetric sophisticated player equilibrium all sophisticated players choose B for the entire duration of the game, that is $y^* = \bar{y}^* = T$.

Proposition 6. *A combination of strategies (T, T) is a symmetric sophisticated player equilibrium ("no teaching" equilibrium) if and only if condition **N1** is satisfied:*

$$\hat{t}_2(0) \geq T(1 - L/H) \qquad \textbf{\textit{(N1)}}$$

Proof: see Appendix B.4.

The proof of Proposition 6 shows that the most profitable deviation from the "no teaching" equilibrium is to start teaching immediately. Condition **N1** ensures that the costs of such deviation exceed the benefits. Immediate teaching increases the payoff flow by from L to H for the period between $\hat{t}_2(0)$ and T, but decreases it from L to 0 in the period between 0 and $\hat{t}_2(0)$.

3.8. Overview of a Numerical Example

To keep the results general, we have placed as few restrictions on the parameter values as possible, but the intuition of the proofs and the importance of the results can be better understood by going through a specific example. A detailed example that illustrates findings from Propositions 1, 3, 4, 5, 6 and Corollaries 1, 2 is described in Appendix A. In the example, we chose the simplest setup that can illustrate how the existence of multiple equilibrium types depends on parameter values. We chose the smallest player composition that satisfies Assumption 2 (with 2 sophisticated, 3 myopic players and $\theta = 3$), a payoff structure that satisfies Assumption 1 ($H = 1$, $M = L = 0.05$) and a reasonable parameter value for weighted fictitious play ($\gamma = 0.2$).[20] This subsection provides an overview of the example, illustrating how the theoretical results are linked to develop predictions about the behavioral patterns in a specific game.

First, the example graphically illustrates how the beliefs of myopic players respond to the strategies of two sophisticated players. At the time when beliefs reach a threshold value, myopic players switch to action A (from Proposition 1). Using Proposition 3, we calculate and graphically illustrate how the switching period responds to the strategy of a single sophisticated player (holding the strategy of the other player constant). The switching period determines the speed of transition to the efficient state, and therefore the sophisticated player's payoffs, calculated using Corollary 1. In the example, the relationship between the payoff and the strategy of a sophisticated player illustrates the trade-off that players face: delayed teaching delays the transition, but shortens the duration of costly teaching. Using iso-profit lines and a profit graph, we show how the sophisticated player would find the best response to the other player's strategy. This allows us to verify which of the three equilibrium types exist. In our example, only the "immediate teaching" equilibrium exists, because neither player has any incentives to delay teaching if the other player does not delay. Figure A.4 plots the types of equilibria that exist for a set of parameter combinations, utilizing the conditions derived in Propositions 4, 5 and 6. Findings from these propositions allow the existence to be evaluated directly, without the need to calculate beliefs, switching period and payoffs. A visual representation in Figure A.4 simplifies the interpretation of the effect that parameter values have on the existence of equilibria: for example, as the planning horizon of the sophisticated players goes up, the system shifts from a unique "no teaching" equilibrium to multiple equilibria to a unique "immediate teaching" equilibrium. Corollary 2 shows that this result holds more generally, and Corollaries 3 and 4 show similar results for the number of sophisticated players and the history of inefficient coordination.

3.9. Summary and Comparative Statics

Table 2 summarizes the findings from Propositions 4, 5 and 6, listing the conditions that need to be satisfied for each type of equilibrium to exist and the speed of transition

[20] When weighted fictitious play is estimated using experimental data, the median estimated values of γ range from 0.1 to 0.5, depending on the game type and model specification (Cheung and Friedman, 1997)

Table 2: Summary of the types of symmetric sophisticated player equilibria that may exist, speed of transition to the efficient state and the conditions that need to be satisfied for each type of equilibrium to exist.

	Immediate Teaching	**Delayed Teaching**	**No Teaching**
Equilibrium strategy	$y^* = 0$	$y^* = \frac{\log\left(\frac{S-H/L}{I^{-1}(N-1)}\right)}{\log(\gamma)} - T_h$	$y^* = T$
Speed of transition	$\hat{t}(y^*, y^*) = \hat{t}_1(0, 0)$	$\hat{t}(y^*, y^*) = \hat{t}_1(y^*, y^*)$	$\hat{t}(y^*, y^*) > T$
Equilibrium payoffs	$\Pi(0, 0) = (T - \hat{t}_1(0, 0))H$	$\Pi(y^*, y^*) = y^*L + (T - \hat{t}_1(y^*, y^*))H$	$\Pi(T, T) = TL$
		Existence conditions	
No deviation to neighbouring strategies	**I1:** $\frac{S-H/L}{N-1} \leq \gamma^{T_h} I^{-1}$	**D2:** $y^* > 0$	–
No deviation to $y = 0$	–	**D3:** $\hat{t}_1(y^*, y^*) - y^*L/H \leq \hat{t}_2(0)$	**N1:** $T(1 - L/H) \leq \hat{t}_2(0)$
No deviation to $y = T$	**I2:** $\hat{t}_1(0, 0) \leq T(1 - L/H)$	**D1:** $\hat{t}_1(y^*, y^*) < T$ **D4:** $\hat{t}_1(y^*, y^*) - y^*L/H \leq T(1 - L/H)$	–

if the equilibrium exists. Note that the uniqueness and the existence of equilibria cannot be guaranteed without additional restrictions on the parameter values. Multiplicity of equilibria is a result of strategic complementarity between the strategies of sophisticated players. Equilibria may fail to exist because we characterize only the symmetric equilibria in pure strategies. Characterization of asymmetric and mixed strategy equilibria would require a different approach than the one used here, and the existence problem would be solved at the expense of increased multiplicity.

From the policy perspective, it is important to know how a change in a parameter value affects the types of equilibria that exist and the speed of transition to the efficient state. In practice, some parameters cannot be measured and known in advance, resulting in uncertainty about the existence of equilibria. Uncertainty depends on the range of parameter values that could satisfy the existence conditions: an equilibrium is "more likely" to exist if a wide range of parameter values could satisfy all conditions, and "less likely" if these conditions are met only with very specific values. To measure how each parameter affects the existence of equilibria, we look at each parameter separately and identify whether a change in its value would increase or decrease the set of other parameter values with which existence conditions are satisfied. We say that an existence condition from Table 2 is satisfied for a smaller set of other parameter values under parameter value p than value p' if the set of other parameter values with which the condition is satisfied when the value is p is a strict subset of the other parameter values with which the condition is satisfied when the value is p'. If a parameter affects all existence conditions of a certain equilibrium type in the same direction, we can identify the overall effect on the existence of this equilibrium type.

In terms of parameters, we consider the length of the planning horizon of sophisticated players (T), the number of sophisticated players (S) and the strength of initial lock-in (T_h). In terms of outcomes, we evaluate which types of equilibria will exist and how long the transition takes if an equilibrium exists.

3.9.1. Planning Horizon of the Sophisticated Players

The first parameter of interest is T, the length of the planning horizon for sophisticated players.

Corollary 2. *If sophisticated players have a longer planning horizon, then:*

1. *The speed of transition in any equilibrium is not affected.*
2. *"Immediate teaching" equilibrium exists for a larger set of other parameter values.*
3. *"Delayed teaching" equilibrium exists for a larger set of other parameter values.*
4. *"No teaching" equilibrium exists for a smaller set of other parameter values.*

Proof: see Appendix B.6.

The first part of the proof holds because the speed of transition in the "immediate teaching" and "delayed teaching" equilibria does not depend on the planning horizon of the sophisticated players. It does, however, affect the existence conditions because strategic teaching becomes more attractive when a longer planning horizon increases the length of time over which the benefits of teaching are received. Consequently, a longer planning horizon makes immediate teaching and delayed teaching more attractive compared to not teaching at all (hence conditions **I2** and **D4** are easier to satisfy). The no-teaching equilibrium, on the other hand, exists for a smaller set of parameter values, because players have more incentives to use strategic teaching (condition **N1** is harder to satisfy).

3.9.2. Player Composition

Instead of measuring the degree of sophistication by the length of the planning horizon, we can measure it by the fraction of the sophisticated players in the group. We therefore investigate the effect of increasing the number of sophisticated players (S) and at the same time decreasing the number of myopic players to keep the total number of players constant.

Corollary 3. *If myopic players are replaced by sophisticated players, then:*

1. *Transition is faster in the "delayed teaching" and "immediate teaching" equilibria.*
2. *The effect on the existence of an "immediate teaching" equilibrium or "delayed teaching" equilibrium is ambiguous:*
   (a) *there are more incentives to deviate to neighboring choice plans*
   (b) *there are less incentives to never choose A.*
3. *There is no change in the existence conditions of the "no teaching" equilibrium.*

Proof: see Appendix B.6.

When myopic players are gradually replaced by sophisticated players, the transition in "immediate teaching" and "delayed teaching" equilibria is faster, because myopic players observe a larger share of other players playing action A. Faster transition increases payoffs in these equilibria, making the option to never teach less attractive (conditions **I2**, **D1** and **D4** would be easier to satisfy). On the other hand, a larger fraction of sophisticated players makes the transition period less sensitive to the strategy of a single sophisticated player, increasing incentives to slightly delay teaching. Such incentives may lead to a break-down of "immediate teaching" and "delayed teaching" equilibria (conditions **I1** and **D2** would be harder to satisfy). The fraction of sophisticated players does not affect the existence of the "no teaching" equilibrium, which would fail to exist only if each sophisticated player had incentives to teach immediately, even in the absence of teaching by other sophisticated players.

3.9.3. Length of the History of Inefficient Coordination

The third factor that we look at is the strength of the initial lock-in to an inefficient state, measured by the length of history of inefficient coordination, T_h.

Corollary 4. *If the history of inefficient coordination is longer, then:*

1. *Transition is slower in the "immediate teaching" equilibrium but faster in a "delayed teaching" equilibrium.*
2. *"Immediate teaching" equilibrium exists for a smaller set of other parameter values*

3. *The effect on the existence of a "delayed teaching" equilibrium is ambiguous*
4. *"No teaching" equilibrium exists for a larger set of other parameter values.*

Proof: see Appendix B.6.

A longer history of inefficient coordination increases the strength of inefficient lock-in, lowering the beliefs of myopic payers and delaying the transition in the "immediate teaching" equilibrium. A similar reduction in beliefs occurs in the "delayed teaching" equilibrium, but, in addition, the equilibrium strategies change too, since sophisticated players need to start teaching earlier to offset a stronger lock-in. Overall, Lemma 12 shows that the effect on the equilibrium strategy exceeds the effect on beliefs, therefore a longer history of inefficient coordination speeds up the transition in the "delayed teaching" equilibrium, in contrast to the "immediate teaching" equilibrium.

Lower beliefs of myopic players increase the cost of strategic teaching, which now has to be used longer to achieve the same result. Consequently, there are more incentives to never teach (condition **I2** is harder to satisfy) or to delay teaching (condition **I1** is harder to satisfy), reducing the set of parameter values for which the "immediate teaching" equilibrium exists. The effect on the existence of the "delayed teaching" equilibrium is ambiguous. A faster transition increases equilibrium payoffs, therefore players have less incentives to stop teaching (conditions **D1** and **D4** are easier to satisfy) or to teach immediately (condition **D3** is easier to satisfy), but there are more incentives to slightly delay teaching because teaching starts later in equilibrium (condition **D2** is harder to satisfy). A longer history of inefficient coordination increases the set of parameter values for which the "no teaching" equilibrium exists, because lower beliefs make strategic teaching more expensive, therefore a unilateral deviation to immediate teaching becomes less attractive.

4. Conclusion

To make predictions in a critical mass coordination game following lock-in we propose a new solution concept based on a Nash equilibrium between sophisticated players who anticipate the learning process of the myopic players. Myopic players make choices based on observed history of play, while sophisticated players have correct beliefs about the actions of all other players, plan ahead and choose actions that maximize the sum of payoff flows. In equilibrium, players may deviate from an inefficient state, but none deviate from the efficient one, therefore there is a unique point in time at which play transitions from the inefficient to the efficient state. For sophisticated players, choice plans that prescribe a switch from efficient to the inefficient action are dominated, therefore a switch to the efficient action occurs at most once. We calculate how such choice plans of sophisticated players affect the switching period of myopic players, and how the latter affects the payoffs of sophisticated players. This mapping from sophisticated player choice plans to payoffs is then used to determine the combinations of choice plans that are mutual best responses.

We show that three types of symmetric equilibria may exist in repeated critical mass games. In the first two types we observe "strategic teaching", that is sophisticated players deviate from the inefficient state to induce the myopic players to deviate in the future, attaining efficient coordination. In the "immediate teaching" equilibrium, sophisticated players deviate from the inefficient state at the start of the game, and in the "delayed teaching" equilibrium sophisticated players initially stay in the inefficient state, but deviate later. In a "no teaching" equilibrium, sophisticated players never deviate from the inefficient state, and neither do the myopic players. Inefficient lock-in is therefore overcome in the first two types of equilibria, but not in the third.

For each equilibrium type, we specify the parameter combinations with which an equilibrium of that type exists, as well as the speed of transition to the efficient state if it does occur. We then show how the likelihood of existence and the speed of transition respond to changes in parameter values. A longer planning horizon of sophisticated players increases the likelihood of the "immediate teaching" and the "delayed teaching" equilibria, and decreases the likelihood of the "no teaching" equilibrium. A longer history of inefficient coordination reduces the likelihood and decreases the speed of transitions in the "immediate teaching" equilibrium. The effect of player composition is ambiguous: on one hand, a larger number of sophisticated players leads to a faster transition and higher profits in the "immediate teaching" and "delayed teaching" equilibria, reducing incentives to completely stop teaching. On the other hand, it reduces the impact that each sophisticated player has on the speed of transition, increasing incentives to delay teaching and leading to a potential breakdown of strategic teaching.

The problem that motivated this paper was the lack of a suitable theoretical model that could be used to make predictions in a game with a history of lock-in. A small change in the assumptions – instead of assuming all players to be farsighted we assume that some players are learning from history – leads to a large difference in theoretical predictions, with the existence of at most three types of equilibria in the repeated game. The methods demonstrated in this study pave the way for the development of a successful positive model of human behavior, which should combine adaptive learning with strategic behavior.

An important subsequent question is whether the predictions of the sophisticated player equilibrium are supported by empirical data. Masiliūnas (2017) ran an experiment with the setup used in this study (a critical mass game following lock-in to an inefficient state) and found support for some of the predictions: farsighted players deviate from the inefficient state more often than the less farsighted ones, and transitions occur from the inefficient to the efficient state, but not the other way around. However, on an individual level, some players start using strategic teaching, but later stop it if other group members fail to follow. Such behavior is not possible in our theoretical model, because strategic players are assumed to be perfectly aware of the composition and future behavior of all other players. The explanatory power of the sophisticated player equilibrium might therefore be increased by relaxing the perfect information assumption and instead assuming that beliefs about the composition of types and their learning rate are updated based on observed feedback. Data in Masiliūnas (2017) shows evidence for the "no teaching" equilibrium (groups that never overcome lock-in) and for the "immediate teaching" equilibrium (groups in which strategic teaching starts right away, and lock-in is soon overcome), but equilibria in which teaching is delayed are not observed. One reason for the failure to observe such equilibria is that they require coordination between the sophisticated players, which is difficult without a coordination device and without a possibility to learn across supergames.

Findings from this paper help to understand some of the previous experimental results; for example, it has been found that a longer planning horizon increases the rates of efficient coordination, either when the planning horizon is a personal characteristic, as in Masiliūnas (2017), or when it is an experimental parameter, as in Berninghaus and Ehrhart (1998). Other results obtained here have not been tested, and while some findings are intuitive, others are unexpected and would be interesting to investigate in the future. For example, we show that a higher number of sophisticated players might introduce incentives to free-ride, and if confirmed, this finding would suggest that it might be beneficial to limit the number of farsighted players when a few of them are sufficient to overcome lock-in.

Acknowledgment

I am grateful to two anonymous referees, associate editor, Friederike Mengel and Philipp J. Reiss for helpful comments and suggestions.

References

Acemoglu, D. and Jackson, M. O. (2014). History, expectations, and leadership in the evolution of social norms. *The Review of Economic Studies*, 82(2):423–456.

Akerlof, G. A. (1980). A theory of social custom, of which unemployment may be one consequence. *The Quarterly Journal of Economics*, 94(4):749–775.

Arthur, W. B. (1989). Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal*, 99(394):116–131.

Ball, L. and Romer, D. (1991). Sticky prices as coordination failure. *The American Economic Review*, 81(3):539–552.

Battalio, R. C., Samuelson, L., Van Huyck, J. B., et al. (1998). *Risk dominance, payoff dominance and probabilistic choice learning*. Social Systems Research Institute, University of Wisconsin.

Battiston, S., Farmer, J. D., Flache, A., Garlaschelli, D., Haldane, A. G., Heesterbeek, H., Hommes, C., Jaeger, C., May, R., and Scheffer, M. (2016). Complexity theory and financial regulation. *Science*, 351(6275):818–819.

Berninghaus, S. K. and Ehrhart, K.-M. (1998). Time horizon and equilibrium selection in tacit coordination games: Experimental results. *Journal of Economic Behavior & Organization*, 37(2):231–248.

Brandts, J. and Cooper, D. J. (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.

Brandts, J., Cooper, D. J., and Fatas, E. (2007). Leadership and overcoming coordination failure with asymmetric costs. *Experimental Economics*, 10(3):269–284.

Brandts, J., Cooper, D. J., Fatas, E., Qi, S., et al. (2016). Stand by me—experiments on help and commitment in coordination games. *Management Science*, 62(10):2916–2936.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

Camerer, C. and Ho, T.-H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica*, 67(4):827–874.

Camerer, C. F., Ho, T.-H., and Chong, J.-K. (2002). Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. *Journal of Economic Theory*, 104(1):137–188.

Cheung, Y.-W. and Friedman, D. (1997). Individual learning in normal form games: Some laboratory results. *Games and Economic Behavior*, 19(1):46–76.

Chong, J.-K., Camerer, C. F., and Ho, T. H. (2006). A learning-based model of repeated games with incomplete information. *Games and Economic Behavior*, 55(2):340–371.

Cooper, R. and John, A. (1988). Coordinating coordination failures in keynesian models. *The Quarterly Journal of Economics*, 103(3):441–463.

Cowan, R. (1990). Nuclear power reactors: a study in technological lock-in. *The Journal of Economic History*, 50(03):541–567.

Crawford, V. P. (1995). Adaptive dynamics in coordination games. *Econometrica*, 63(1):103–143.

David, P. (1985). Clio and the economics of QWERTY. *American Economic Review*, 75(2):332–37.

Devetag, G. (2003). Coordination and information in critical mass games: an experimental study. *Experimental Economics*, 6(1):53–73.

Diamond, D. W. and Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, 91(3):401–419.

Edmond, C. (2013). Information manipulation, coordination, and regime change. *Review of Economic Studies*, 80(4):1422–1458.

Ellison, G. (1997). Learning from personal experience: One rational guy and the justification of myopia. *Games and Economic Behavior*, 19(2):180–210.

Erev, I. and Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, 88(4):848–881.

Folke, C., Carpenter, S., Walker, B., Scheffer, M., Elmqvist, T., Gunderson, L., and Holling, C. (2004). Regime shifts, resilience, and biodiversity in ecosystem management. *Annual Review of Ecology, Evolution, and Systematics*, 35:557–581.

Friedman, D. (1996). Equilibrium in evolutionary games: Some experimental results. *The Economic Journal*, 106(434):1–25.

Fudenberg, D. and Levine, D. K. (1998). *The theory of learning in games*, volume 2. MIT press.

Gao, J., Barzel, B., and Barabási, A.-L. (2016). Universal resilience patterns in complex networks. *Nature*, 530:307–312.

Granovetter, M. (1978). Threshold models of collective behavior. *American Journal of Sociology*, 83(6):1420–1443.

Greif, A. and Laitin, D. D. (2004). A theory of endogenous institutional change. *American Political Science Review*, 98(4):633–652.

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

Hey, J. D. and Knoll, J. A. (2007). How far ahead do people plan? *Economics Letters*, 96(1):8–13.

Hirota, M., Holmgren, M., Van Nes, E. H., and Scheffer, M. (2011). Global resilience of tropical forest and savanna to critical transitions. *Science*, 334(6053):232–235.

Ho, T.-H. and Weigelt, K. (1996). Task complexity, equilibrium selection, and learning: An experimental study. *Management Science*, 42(5):659–679.

Hyndman, K., Ozbay, E. Y., Schotter, A., and Ehrblatt, W. Z. (2012). Convergence: an experimental study of teaching and learning in repeated games. *Journal of the European Economic Association*, 10(3):573–604.

Hyndman, K., Terracol, A., and Vaksmann, J. (2009). Learning and sophistication in coordination games. *Experimental Economics*, 12(4):450–472.

Johnson, E. J., Camerer, C., Sen, S., and Rymon, T. (2002). Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of Economic Theory*, 104(1):16–47.

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

Kiss, H. J., Rodríguez-Lara, I., and Rosa-García, A. (2017). Overthrowing the dictator: a game-theoretic approach to revolutions and media. *Social Choice and Welfare*, 49(2):329–355.

Kuran, T. (1989). Sparks and prairie fires: A theory of unanticipated political revolution. *Public Choice*, 61(1):41–74.

Mantovani, M. (2016). Limited foresight in sequential games: an experiment. Working paper.

Masiliūnas, A. (2017). Overcoming coordination failure in a critical mass game: strategic motives and action disclosure. *Journal of Economic Behavior & Organization*, 139:214–251.

Mengel, F. (2014). Learning by (limited) forward looking players. *Journal of Economic Behavior & Organization*, 108:59–77.

Monderer, D. and Shapley, L. S. (1996). Fictitious play property for games with identical interests. *Journal of Economic Theory*, 68(1):258–265.

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge university press.

Nyström, M., Folke, C., and Moberg, F. (2000). Coral reef disturbance and resilience in a human-dominated environment. *Trends in Ecology & Evolution*, 15(10):413–417.

Rapoport, A. (1985). Provision of public goods and the MCS experimental paradigm. *American Political Science Review*, 79(01):148–155.

Rapoport, A. and Eshed-Levy, D. (1989). Provision of step-level public goods: Effects of greed and fear of being gypped. *Organizational Behavior and Human Decision Processes*, 44(3):325–344.

Scheffer, M., Bascompte, J., Brock, W., Brovkin, V., Carpenter, S., Dakos, V., Held, H., van Nes, E., Rietkerk, M., and Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature*, 461:53–59.

Scheffer, M., Carpenter, S. R., Lenton, T. M., Bascompte, J., Brock, W., Dakos, V., Van de Koppel, J., Van de Leemput, I. A., Levin, S. A., Van Nes, E. H., et al. (2012). Anticipating critical transitions. *Science*, 338(6105):344–348.

Scheffer, M., Hosper, S., Meijer, M., Moss, B., and Jeppesen, E. (1993). Alternative equilibria in shallow lakes. *Trends in Ecology & Evolution*, 8(8):275–279.

Schipper, B. C. (2019). Dynamic exploitation of myopic best response. *Dynamic Games and Applications*.

Sefton, M. (1999). A model of behavior in coordination game experiments. *Experimental Economics*, 2(2):151–164.

Shapiro, C. and Varian, H. R. (1999). *Information rules: a strategic guide to the network economy*. Harvard Business School Press, Boston, MA.

Terracol, A. and Vaksmann, J. (2009). Dumbing down rational players: Learning and teaching in an experimental game. *Journal of Economic Behavior & Organization*, 70(1):54–71.

van de Leemput, I. A., Wichers, M., Cramer, A. O., Borsboom, D., Tuerlinckx, F., Kuppens, P., van Nes, E. H., Viechtbauer, W., Giltay, E. J., Aggen, S. H., et al. (2014). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*, 111(1):87–92.

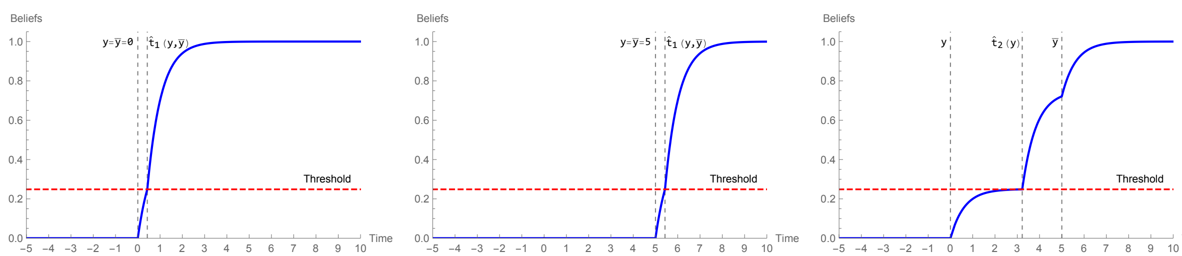
Van Huyck, J. B., Battalio, R. C., and Rankin, F. W. (1997). On the origin of convention: Evidence from coordination games. *The Economic Journal*, 107(442):576–596.

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

Online Appendix for "Overcoming Inefficient Lock-in in Coordination Games with Sophisticated and Myopic Players"

Appendix A. Numerical Example

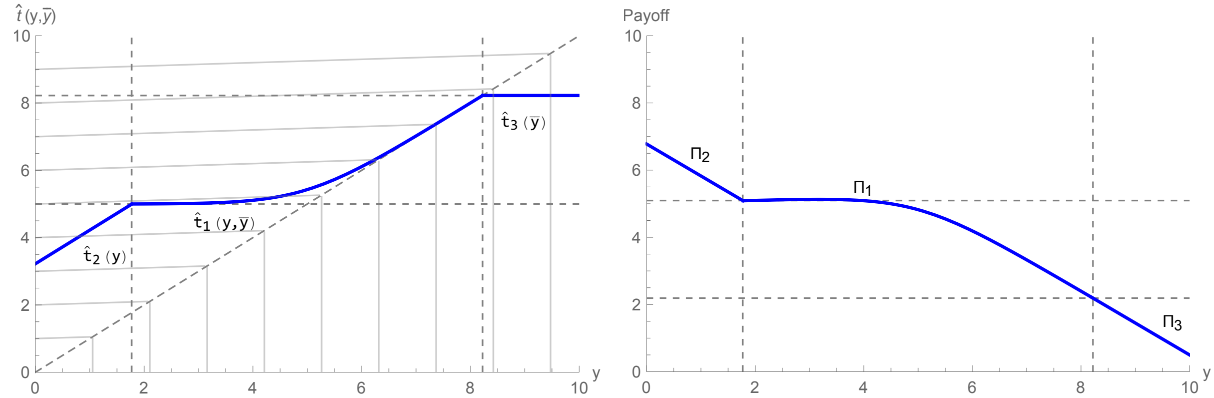
We assume 2 sophisticated and 3 myopic players ($n = 5$ and $s = 2$), the smallest group composition that satisfies Assumption 2. We assume that $\theta = 3$, the smallest value that satisfies Assumption 2 for the chosen group composition. Without the loss of generality, we normalize the payoff in the efficient state to $H = 1$. We also assume that $M = L = 0.05$, $\gamma = 0.2$, $T_h = 5$ and $T = 10$. We illustrate the beliefs and switching period of myopic players and the payoffs of a sophisticated player who chooses strategy y if the other sophisticated player chooses $\bar{y}$.



(a) $y = 0$, $\bar{y} = 0$ (b) $y = 5$, $\bar{y} = 5$ (c) $y = 0$, $\bar{y} = 5$

Figure A.1: Myopic player beliefs, assuming $n = 5$, $s = 2$, $\theta = 3$, $H = 1$, $M = L = 0.05$, $\gamma = 0.2$, $T_h = 5$.

Proposition 1 indicates that in a game with the chosen parameter values myopic players will choose A if their beliefs exceed the threshold value $I^{-1}_{0.05}(\theta, N - \theta) = 0.249$. Since the threshold value is below $1/(N-1) = 0.25$, a transition to the efficient state may be possible because of strategic teaching by a single sophisticated player (over time, beliefs approach 0.25 if one out of four other group members always plays A). Panel (a) of Figure A.1 illustrates the evolution of beliefs, calculated using equation (1), if $y = \bar{y} = 0$. Beliefs reach the threshold value at time $\hat{t}_1(0, 0) = 0.43$, and all players choose A from this point onward, therefore beliefs start to go up at a faster pace. Similarly, panel (b) shows that myopic players switch to A at time $\hat{t}_1(5, 5) = 5.43$ if sophisticated players delay the start of strategic teaching to time $y = \bar{y} = 5$. If sophisticated players use different strategies, a transition to the efficient state can occur before the second player starts teaching. This case is illustrated in panel (c), assuming $y = 0$ and $\bar{y} = 5$. Myopic player beliefs grow slowly at first, when only one sophisticated player is teaching, faster starting from time $\hat{t}_2(0) = 3.22$, when all myopic players switch to A, and even faster after time $\bar{y} = 5$, when the second sophisticated player switches to A. However, the strategy profile illustrated in panel (c) could not be supported in equilibrium because the second sophisticated player could increase earnings by switching to A earlier. In a symmetric equilibrium, both players must start teaching at the start of the game (as in panel (a)), in the middle of the game (panel (b)) or never, a case in which myopic player beliefs remain 0 throughout the game.



(a) Switching time of myopic players

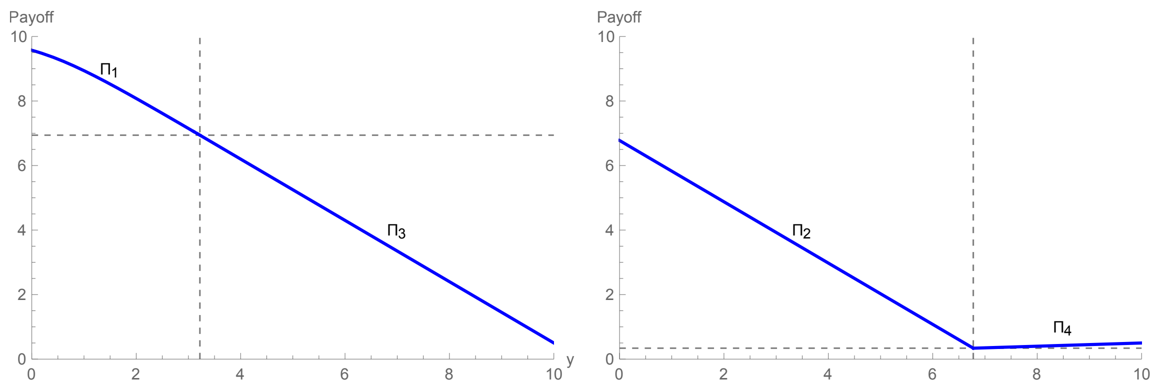
(b) Sophisticated player payoffs

Figure A.2: Switching time and payoffs of a sophisticated player if the other sophisticated player chooses $\bar{y} = 5$, assuming $n = 5$, $s = 2$, $\theta = 3$, $H = 1$, $M = L = 0.05$, $\gamma = 0.2$, $T_h = 5$, $T = 10$.

Panel (a) in Figure A.2 plots the switching time of myopic players (from Proposition 3) for all y, assuming that the other sophisticated player chooses $\bar{y} = 5$. If $y = 0$, myopic players switch at $\hat{t}_2(0) = 3.22$, a case that was illustrated in panel (c) of Figure A.1. The longer the delay, the later is the switching period, and the relationship between y and $\hat{t}_2(y)$ is initially nearly linear. The slope decreases when $\hat{t}_2(y)$ reaches $\bar{y} = 5$, at the point $y = 1.77$. If y exceeds this value, myopic players switch after both sophisticated players, and the switching period is calculated by $\hat{t}_1(y, \bar{y})$. Since beliefs jump up when both sophisticated players are observed teaching, myopic players would switch to A soon after $\bar{y}$, and a higher y does little to delay the transition. The slope increases when y approaches $\bar{y} = 5$ (the case when $y = \bar{y} = 5$ was illustrated in panel (b) of Figure A.1). The slope increases because when y exceeds $\bar{y}$, the transition period occurs soon after y, when both sophisticated players start teaching. If y is sufficiently large, myopic players switch prior to y, after observing strategic teaching by the other sophisticated player since time $\bar{y} = 5$. The switch occurs at $\hat{t}_3(5) = 8.22$. Any increase of y beyond $\hat{t}_3(5)$ does not affect the switching period of myopic players.

Information about the switching period allows sophisticated players to calculate the payoffs from the entire game, as specified in Corollary 1. In panel (a) of Figure A.2, payoffs are illustrated using iso-profit lines, which consist of two parts. The lines are vertical if $\hat{t}(y, \bar{y}) < y$, because payoffs depend only on y and a shorter delay increases the payoff flow from M to H. Payoffs are not affected by $\hat{t}$ because we assume that $L = M$. The lines have a positive slope if $\hat{t}(y, \bar{y}) \geq y$, because the payoff flow is decreasing both in $\hat{t}$ (from H to 0) and in y (from L to 0). The slope of the iso-profit line is therefore determined by the L/H ratio. In the example shown in panel (a), the highest (i.e. closest to the origin) iso-profit line is reached by setting $y = 0$, therefore the "delayed teaching" equilibrium with $y = \bar{y} = 5$ does not exist.

The payoff-maximizing action is easier identified by plotting the payoff function $\Pi(y, \bar{y})$ from Corollary 1 (panel (b) of Figure A.2). If $y = 0$, the payoff flow is 0 until myopic players switch to A at time $\hat{t}_2(0) = 3.22$, and then 1 from $\hat{t}_2(0)$ until the end of the game, therefore $\Pi(0, 5) = 6.78$. A higher y increases the transition period in a nearly linear way, decreasing the payoff that is calculated by $\Pi_2(y, \bar{y})$. Lemma 3 shows that this pattern would hold in all games in which a transition is possible by strategic teaching of a single sophisticated player. If $y > 1.77$, the payoff starts being calculated by $\Pi_1(y, \bar{y})$ and is increasing in y, because $\hat{t}$ is initially insensitive to y. Once the sensitivity goes up, the payoffs start to decrease. The decrease continues at the same rate even when y passes $\hat{t}_3(\bar{y}) = 8.22$ and the payoffs start being calculated by $\Pi_3(y, \bar{y})$. Any further increase in y delays the time at which the sophisticated player starts to earn H and thus decreases payoffs.



(a) $\bar{y} = 0$ (b) $\bar{y} = 10$

Figure A.3: Sophisticated player payoffs, assuming $n = 5$, $s = 2$, $\theta = 3$, $H = 1$, $M = L = 0.05$, $\gamma = 0.2$, $T_h = 5$, $T = 10$.

Figure A.2 indicates that there is no "delayed teaching" equilibrium with $y = \bar{y} = 5$ because the best reply to the opponent choosing $\bar{y} = 5$ is $y = 0$. To check whether the "immediate teaching" equilibrium exists, we plot the payoffs for $\bar{y} = 0$ in panel (a) of Figure A.3. Payoffs are decreasing in y because higher y delays the transition, which occurs soon after both players start to teach (if $y < \hat{t}_3(\bar{y}) = 3.22$), or because it shortens the duration of time in which player receives H (if $y \geq \hat{t}_3(\bar{y})$). The optimal response to $\bar{y} = 0$ is $y = 0$, therefore an "immediate teaching" equilibrium exists.

Next, we evaluate whether the "no teaching" equilibrium exists. Panel (b) of Figure A.3 plots the payoff when $\bar{y} = 10$. Payoffs from a small y are determined by $\Pi_2(y)$, and are therefore identical to the payoffs in panel (b) of Figure A.2. A higher y delays the transition and decreases payoffs, up to the point when $\hat{t}_2(y) = 10$, which occurs when $y = 6.77$. This strategy minimizes payoffs, because the player pays the cost of strategic teaching, but receives no benefit from it since the transition occurs at the end of player's planning horizon. Increasing y beyond 6.77 increases payoffs, shortening the duration of wasteful strategic teaching. In this example, the best-response to $\bar{y} = 10$ is $y = 0$, therefore the "no teaching" equilibrium does not exist.

The existence conditions derived in Propositions 4, 5 and 6 allow existence to be eval-

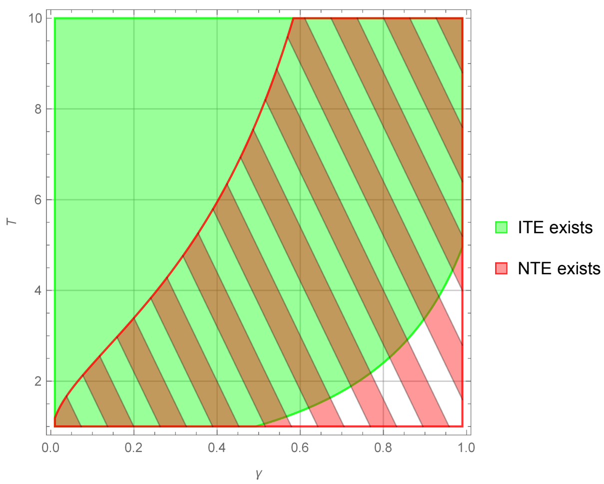


Figure A.4: Combinations of γ and T for which the "immediate teaching" and "no teaching" equilibria exist, assuming $n = 5$, $s = 2$, $\theta = 3$, $H = 1$, $M = L = 0.05$, $T_h = 5$.

uated directly, without the calculation of payoffs. Figure A.4 shows the combination of two variables, γ and T, for which the "immediate teaching" and "no teaching" equilibria exist. These two parameters are interesting to study because precise values of personal characteristics would be difficult to know in practice. With the chosen parameter values, the "delayed teaching" equilibrium does not exist.[21] Figure A.4 shows that with the parameter values assumed in this example ($\gamma = 0.2$, $T = 10$), only the "immediate teaching" equilibrium exists. The "no teaching" equilibrium may start to exist if γ went up, increasing the weight placed on the history of inefficient coordination and delaying transition, or if T went down, decreasing the benefits of strategic teaching. There is a range of parameters for which two types of equilibria co-exist, and a range in which the "no teaching" equilibrium is unique. A higher T increases the set of other parameter values for which the "immediate teaching" equilibrium exists, and decreases the set for the "no teaching" equilibrium; Corollary 2 shows that this result holds more generally, and Corollaries 3 and 4 show similar results for the number of sophisticated players and the history of inefficient coordination.

Appendix B. Proofs

Appendix B.1. Proposition 1

Proposition 1. *Suppose that in a game with payoffs defined by* (5) *at time t myopic player m holds beliefs $x_m(t)$. Then the response function defined in* (3) *is equivalent to:*

[21] "Delayed teaching" equilibrium does not exist because $S - H/L < 0$, therefore the derivative of the profit function with respect to y at any potential equilibrium would be negative (see the calculations in the proof of Lemma 7).

$$r_m(t, \mathbf{c}_S) = \begin{cases} A & \text{if } x_m(t) \geq I^{-1}_{\frac{L}{L+H-M}}(\theta, N-\theta) \\ B & \text{otherwise} \end{cases}$$

where I^{-1} is the inverse of an incomplete regularized beta function.

Proof.

From (3), action A is chosen if the subjective expected payoff of A at time t exceeds the subjective expected payoff of B:

$$r_m(t, \mathbf{c}_S) = A \Leftrightarrow SEP(A, x_m(t)) \geq SEP(B, x_m(t)) \tag{B.1}$$

In critical mass games, player's payoff depends on the chosen action and on whether the number of other group members who choose A exceeds θ. Denote the subjective probability assigned to the latter event by $Pr[\alpha(\mathbf{r}_{-m}(t, \mathbf{c}_S)) \geq \theta | x_m(t)]$. Then subjective expected payoffs from equation (2) can be defined as:

$$\begin{aligned} SEP(A, x_m(t)) &= 0 \times (1 - Pr[\alpha(\mathbf{r}_{-m}(t, \mathbf{c}_S)) \geq \theta | x_m(t)]) + H \times Pr[\alpha(\mathbf{r}_{-m}(t, \mathbf{c}_S)) \geq \theta | x_m(t)] \\ SEP(B, x_m(t)) &= L \times (1 - Pr[\alpha(\mathbf{r}_{-m}(t, \mathbf{c}_S)) \geq \theta | x_m(t)]) + M \times Pr[\alpha(\mathbf{r}_{-m}(t, \mathbf{c}_S)) \geq \theta | x_m(t)] \end{aligned} \tag{B.2}$$

The subjective probability that the threshold will be exceeded is calculated by adding the probabilities assigned to all action profiles of other players in which more than θ players choose A:

$$Pr[\alpha(\mathbf{r}_{-m}(t, \mathbf{c}_S)) \geq \theta | x_m(t)] = \sum_{k=\theta}^{N-1} (x_m(t))^k (1 - x_m(t))^{N-1-k} \binom{N-1}{k} \tag{B.3}$$

Use equations (B.2) and (B.3) to rewrite (B.1) the following way:

$$r_m(t, \mathbf{c}_S) = A \Leftrightarrow \sum_{k=\theta}^{N-1} (x_m(t))^k (1 - x_m(t))^{N-1-k} \binom{N-1}{k} \geq \frac{L}{L+H-M} \tag{B.4}$$

Notation in (B.4) is simplified using the definition of an incomplete regularized beta function:[22]

$$r_m(t, \mathbf{c}_S) = A \Leftrightarrow I_{x_m(t)}(\theta, N-\theta) \geq \frac{L}{L+H-M} \tag{B.5}$$

Taking the inverse of (B.5) and substituting into (3) completes the proof:

$$r_m(t, \mathbf{c}_S) = \begin{cases} A & \text{if } x_m(t) \geq I^{-1}_{\frac{L}{L+H-M}}(\theta, N-\theta) \\ B & \text{otherwise} \end{cases}$$

□

[22] An incomplete regularized beta function is defined as $I_c(a, b) = \sum_{k=a}^{a+b-1} c^k (1-c)^{a+b-1-k} \binom{a+b-1}{k}$. The function is well defined because $\frac{L}{L+H-M} \in (0, 1)$, from Assumption 1.

Appendix B.2. Proposition 2

Proposition 2. *Choice plan profiles for sophisticated players that prescribe a switch from A to B for at least one sophisticated player are dominated:*

$$AB_S \cap U_s = \emptyset$$

Proof.

Take a choice plan profile $\mathbf{c}_S \in AB_S$. We will show that in this profile at least one sophisticated player must be choosing a dominated choice plan.

If $\mathbf{c}_S \in AB_M$, at least one sophisticated player must be choosing a dominated choice plan, from Lemma 2, and the proof would be completed. Alternatively, assume that $\mathbf{c}_S \in \{AB_S \setminus AB_M\}$. By the definition of AB_S, there must be a sophisticated player whose choice plan prescribes a switch from A to B; denote the choice plan of this player by $\widetilde{c}_s$ and denote the switching time prescribed by $\widetilde{c}_s$ by t'. Then there must be some small ε such that $\widetilde{c}_s(t) = A$ if $t \in [t' - \varepsilon, t')$ and $\widetilde{c}_s(t) = B$ if $t \in [t', t' + \varepsilon]$. Since $\mathbf{c}_S \notin AB_M$, myopic players switch from B to A at most once, thus their choices can be described by a number $\hat{t}(\widetilde{c}_s)$ that identifies this switching time: B is chosen in the interval $[0, \hat{t}(\widetilde{c}_s))$ and A is chosen in the interval $[\hat{t}(\widetilde{c}_s), T]$.

First, suppose that $t' \geq \hat{t}(\widetilde{c}_s)$, then myopic players would be choosing A at any time $t \geq t'$. Assumption 2 implies that the threshold will be exceeded at any such point in time, therefore a choice plan $\widetilde{c}_s$ is dominated by a choice plan that prescribes A at each point in time $t \geq \hat{t}(\widetilde{c}_s)$. Next, suppose that $t' < \hat{t}(\widetilde{c}_s)$ and $\hat{t}(\widetilde{c}_s) > T$. Then myopic players will choose B for the entire period that is taken into account by the sophisticated player, thus choice plan $\widetilde{c}_s$ will be dominated by a choice plan that prescribes B for the entire interval.

Alternatively, suppose that $\hat{t}(\widetilde{c}_s) > t'$ and $\hat{t}(\widetilde{c}_s) \leq T$ (see Figure B.5). Choose ε to be sufficiently small to satisfy $\hat{t}(\widetilde{c}_s) > t' + \varepsilon$. Then for any $\widetilde{c}_s$ construct a choice plan c'_s the following way:

$$c'_s(t) = \begin{cases} \widetilde{c}_s(t) & \text{if } t \in [0, t' - \varepsilon) \cup (t' + \varepsilon, T] \\ B & \text{if } t \in [t' - \varepsilon, t'] \\ A & \text{if } t \in (t', t' + \varepsilon] \end{cases}$$

In other words, c'_s is constructed by taking $\widetilde{c}_s$ and swapping choices prescribed in the interval $(t' - \varepsilon, t')$ with choices prescribed in the interval $(t', t' + \varepsilon)$. We will show that $\widetilde{c}_s$ is dominated by c'_s.

The comparison of payoff flows generated by these two choice plans is shown in Figure B.5. In the interval $[0, t + \varepsilon)$ the sum of payoff flows is the same for both choice plans $(\pi_1 + \pi_2 + \pi_3)$. Payoffs are equal because with both choice plans myopic players choose B in this entire interval (both $\hat{t}(c'_s)$ and $\hat{t}(\widetilde{c}_s)$ exceed $t' + \varepsilon$), therefore the participation threshold is never exceeded. Choice plan $\widetilde{c}_s$ prescribes A for the same duration of time as c'_s, therefore the sum of payoffs in the interval $[0, t' + \varepsilon)$ would be the same for both choice plans.

In the interval $[t' + \varepsilon, T]$ the sum of payoffs generated by c'_s is strictly higher than that of $\widetilde{c}_s$. Since $\widetilde{c}_s(t) = c'_s(t)$, $\forall t \in (t' + \varepsilon, T]$, any payoff difference between the two choice plans

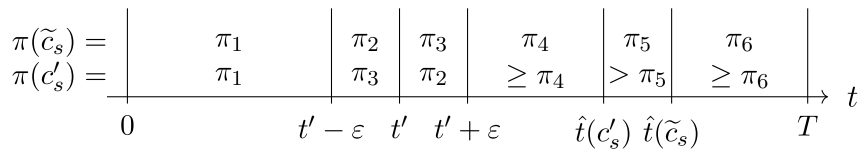


Figure B.5: Payoff flows generated by choice plans $\widetilde{c}_s$ and c'_s for the case $\hat{t}(\widetilde{c}_s) > t'$ and $\hat{t}(\widetilde{c}_s) \leq T$.

in this interval must be due to the choices of myopic players. From equation (1), $x_m(t)$ would be the same under $\widetilde{c}_s(t)$ as under $c_s(t)'$ if γ was equal to 1. But since $\gamma \in (0, 1)$, older observations receive less weight and therefore myopic player beliefs would be strictly higher following c'_s than following $\widetilde{c}_s$ at any time $t \in (t' + \varepsilon, T]$. Then Lemma 1 implies that the payoff flow is always weakly higher for c'_s at any time in the interval $[t' + \varepsilon, T]$. To get strict dominance, note that $\hat{t}(c'_s) < \hat{t}(\widetilde{c}_s)$, for the following reasons. Since $\hat{t}(c'_s) \in (t' + \varepsilon, T]$ and $x_m(t)$ is continuous, the switching period $\hat{t}(c'_s)$ must satisfy $x'_m(\hat{t}(c'_s)) = I^{-1}$. But since $\widetilde{x}_m(t) < x'_m(t)$, $\forall t \in (t' + \varepsilon, T]$, it must also hold that $\widetilde{x}_m(\hat{t}(c'_s)) < x'_m(\hat{t}(c'_s)) = I^{-1}$. Consequently, the intersection of beliefs $\widetilde{x}_m(t)$ and belief threshold I^{-1} must occur strictly later, so that $\hat{t}(c'_s) < \hat{t}(\widetilde{c}_s)$. In the interval $(\hat{t}(c'_s), \hat{t}(\widetilde{c}_s))$ choice plan $\widetilde{c}_s$ provides a flow of payoffs of at most L, while c'_s provides a payoff of H because more than θ players are choosing A.

The comparison of payoff flows associated with choice plans $\widetilde{c}_s$ and c'_s is shown in Figure B.5. The sum of payoff flows generated by c'_s will be strictly higher than the sum of payoff flows generated by $\widetilde{c}_s$, therefore choice plan $\widetilde{c}_s$ that prescribes switching from A to B is strictly dominated by another choice plan c'_s.

$\square$

Appendix B.3. Proposition 3

Proposition 3. *If one sophisticated player uses strategy $y_s = y$ and the other $(S-1)$ players use strategies $\mathbf{y}_{-s} = \{(\bar{y})_{\times(S-1)}\}$, the switching period of myopic players is:*

$$\hat{t}(y, \bar{y}) = \begin{cases} \hat{t}_1(y, \bar{y}) & \text{if } \max\{y, \bar{y}\} < \hat{t}_1(y, \bar{y}) \quad \text{and} \quad \frac{S}{N-1} > I^{-1} \\ \hat{t}_2(y) & \text{if } y < \hat{t}_2(y) \leq \bar{y} \quad \text{and} \quad \frac{1}{N-1} > I^{-1} \\ \hat{t}_3(\bar{y}) & \text{if } \bar{y} < \hat{t}_3(\bar{y}) \leq y \quad \text{and} \quad \frac{S-1}{N-1} > I^{-1} \\ \infty & \text{otherwise} \end{cases}$$

such that

$$\hat{t}_1(y, \bar{y}) = \frac{\log(\frac{S}{N-1} - I^{-1}) - \log(\gamma^{-\bar{y}} \frac{S-1}{N-1} + \gamma^{-y} \frac{1}{N-1} - \gamma^{T_h} I^{-1})}{\log(\gamma)}$$

$$\hat{t}_2(y) = \frac{\log(\frac{1}{N-1} - I^{-1}) - \log(\gamma^{-y} \frac{1}{N-1} - \gamma^{T_h} I^{-1})}{\log(\gamma)}$$

$$\hat{t}_3(\bar{y}) = \frac{\log(\frac{S-1}{N-1} - I^{-1}) - \log(\gamma^{-\bar{y}} \frac{S-1}{N-1} - \gamma^{T_h} I^{-1})}{\log(\gamma)}$$

Proof.

There are three cases to consider. In the first case, $\hat{t}(y,\bar{y}) > \max\{y,\bar{y}\}$, so that myopic players observe no other players choosing A from time 0 to time $\min\{y,\bar{y}\}$, a fraction of $\frac{S}{N-1}$ others choosing A from time $\max\{y,\bar{y}\}$ to $\hat{t}(y,\bar{y})$ and either a fraction of $\frac{1}{N-1}$ others choosing A from time $\bar{y}$ to time y (if $y > \bar{y}$) or a fraction of $\frac{S-1}{N-1}$ others choosing A from time y to $\bar{y}$ (if $\bar{y} > y$). Feedback observed by myopic players in this case is illustrated in Figure B.6.

In the second case, $y < \hat{t}(y,\bar{y}) < \bar{y}$. This will be true only if $\frac{1}{N-1} > I^{-1}$, that is if myopic players would switch to A after observing only one player choosing A. In this case each myopic player will observe no others choosing A from time 0 to y and a fraction of $\frac{1}{N-1}$ others choosing A from time y to $\hat{t}(y,\bar{y})$. See Figure B.7 for a graphical representation.

In the third case, $\bar{y} < \hat{t}(y,\bar{y}) < y$. Then each myopic player will observe no others choosing A from time 0 to $\bar{y}$ and a fraction of $\frac{S-1}{N-1}$ others choosing A from time $\bar{y}$ to $\hat{t}(y,\bar{y})$. See Figure B.8 for a graphical representation.

It is never possible that $\hat{t}(y,\bar{y}) < \min\{y,\bar{y}\}$ because at time $t \in [0,\min\{y,\bar{y}\})$ myopic players observe no others choosing A and therefore always choose B.

Case 1: $\hat{t}(y,\bar{y}) > \max\{y,\bar{y}\}$

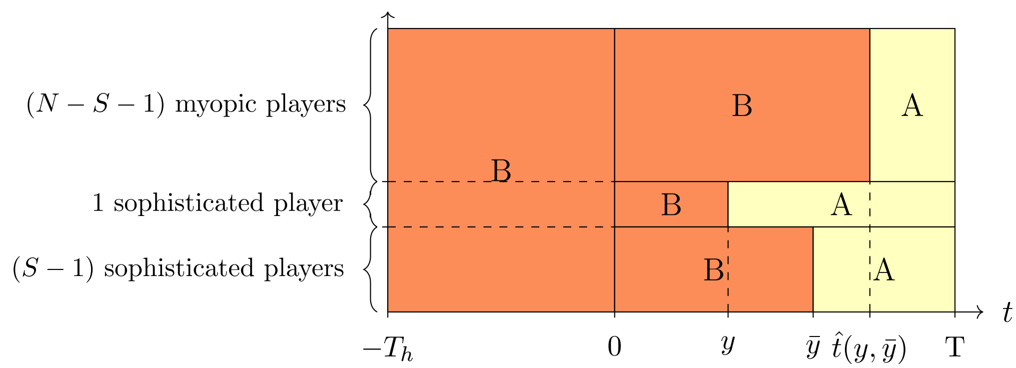


Figure B.6: Illustration of the feedback observed by a single myopic player in the first case, where $\hat{t}(y,\bar{y}) > \max\{y,\bar{y}\}$. In this example $\bar{y} > y$. Vertical axis shows the fraction of other players choosing A or B, horizontal axis shows the passage of time. The first sophisticated player switches from B to A at time y, other $(S-1)$ sophisticated players switch at time $\bar{y}$ and myopic players switch at time $\hat{t}(y,\bar{y})$

Recall that beliefs of myopic players are calculated using weighted fictitious play from equation (1). If sophisticated players are using strategies y and $\bar{y}$, myopic player beliefs at any time $t \in (\max\{y,\bar{y}\},\hat{t}(y,\bar{y})]$ will be calculated using the following rule:

$$x_i(t) = \frac{\int_{k=0}^{t-\bar{y}} \gamma^k (\frac{S-1}{N-1}) \mathrm{d}k + \int_{k=0}^{t-y} \gamma^k (\frac{1}{N-1}) \mathrm{d}k}{\int_{k=0}^{t+T_h} \gamma^k \, \mathrm{d}k} =$$

$$= \frac{(\gamma^{t-\bar{y}} - 1)(\frac{S-1}{N-1}) + (\gamma^{t-y} - 1)(\frac{1}{N-1})}{\gamma^{t+T_h} - 1}$$

The terms in the numerator correspond to the history observed by a myopic player up to time $t \in (\max\{y, \bar{y}\}, \hat{t}(y, \bar{y})]$: $(S-1)$ sophisticated players are observed choosing A for a span of $t - \bar{y}$ and one sophisticated player is observed choosing A for a span of $t - y$. This feedback is illustrated in Figure B.6. The denominator measures the length of the entire history, including the T_h rounds of inefficient coordination.

From Proposition 1, myopic players will choose A at time t if $x_m(t) \geq I^{-1}$:

$$x_m(t) \geq I^{-1} \quad \Leftrightarrow \quad \frac{(\gamma^{t-\bar{y}} - 1)(\frac{S-1}{N-1}) + (\gamma^{t-y} - 1)(\frac{1}{N-1})}{\gamma^{t+T_h} - 1} \geq I^{-1} \quad \Leftrightarrow$$
$$\gamma^{t+T_h}\left(\gamma^{-\bar{y}-T_h}\frac{S-1}{N-1} + \gamma^{-y-T_h}\frac{1}{N-1} - I^{-1}\right) \leq \frac{S}{N-1} - I^{-1} \tag{B.6}$$

If $\frac{S}{N-1} - I^{-1} \leq 0$, equation (B.6) is never satisfied because of the following relationship that contradicts (B.6):

$$\gamma^{t+T_h}\big(\gamma^{-\bar{y}-T_h}\tfrac{S-1}{N-1} + \gamma^{-y-T_h}\tfrac{1}{N-1} - I^{-1}\big) > \gamma^{t+T_h}\big(\tfrac{S}{N-1} - I^{-1}\big) \geq \tfrac{S}{N-1} - I^{-1} \tag{B.7}$$

The first inequality holds because $\gamma^{-\bar{y}-T_h} > 1$ and $\gamma^{-y-T_h} > 1$ and the second inequality holds because $\gamma^{t+T_h} < 1$ and $\frac{S}{N-1} - I^{-1} \leq 0$. But (B.7) contradicts (B.6), therefore if $\frac{S}{N-1} - I^{-1} \leq 0$, equation (B.6) is never satisfied and myopic players would choose B at any time t.

Alternatively, if $\frac{S}{N-1} - I^{-1} > 0$, condition (B.6) can be expressed the following way:

$$\gamma^{-t} \geq \frac{\gamma^{-\bar{y}}\frac{S-1}{N-1} + \gamma^{-y}\frac{1}{N-1} - \gamma^{T_h}I^{-1}}{\frac{S}{N-1} - I^{-1}} \tag{B.8}$$

The left-hand side of (B.8) is strictly increasing in t and unbounded for any $\gamma \in (0, 1)$, so (B.8) will be satisfied for some t, although not necessarily with $t \leq T$. Equation (B.8) is not satisfied for $t = 0$ because the RHS of (B.8) is always strictly larger than 1 (RHS is increasing in both y and $\bar{y}$, but RHS > 1 even if $y = \bar{y} = 0$ because $\frac{S}{N-1} - \gamma^{T_h}I^{-1} > \frac{S}{N-1} - I^{-1}$) and $\gamma^{-t} < 1$. Consequently, (B.8) must be satisfied with equality at a unique value of t, which we denote by $\hat{t}_1(y, \bar{y})$, with $\hat{t}_1(y, \bar{y}) \in (0, \infty)$. This value is the first moment in time at which myopic players are indifferent between choosing A and B, thus it is the switching period of myopic players. To get an expression for $\hat{t}_1(y, \bar{y})$, we require (B.8) to be satisfied with equality and rearrange the following way:

$$\hat{t}_1(y, \bar{y}) = \frac{\log(\frac{S}{N-1} - I^{-1}) - \log(\gamma^{-\bar{y}}\frac{S-1}{N-1} + \gamma^{-y}\frac{1}{N-1} - \gamma^{T_h}I^{-1})}{\log(\gamma)} \tag{B.9}$$

Of course, $\hat{t}(y, \bar{y})$ can be calculated using (B.9) only if $\frac{S}{N-1} - I^{-1} > 0$, otherwise myopic players would always play B. The precise characterization of the switching period if case 1 is applicable is as follows:

$$\hat{t}(y, \bar{y}) = \begin{cases} \hat{t}_1(y, \bar{y}) & \text{if } \frac{S}{N-1} - I^{-1} > 0 \\ \infty & \text{otherwise} \end{cases} \tag{B.10}$$

Note that it is not required that $\hat{t}_1(y, \bar{y}) \leq T$, therefore it is possible that the planning horizon of a sophisticated player is too short to take re-coordination into account.

Case 2: $y < \hat{t} < \bar{y}$

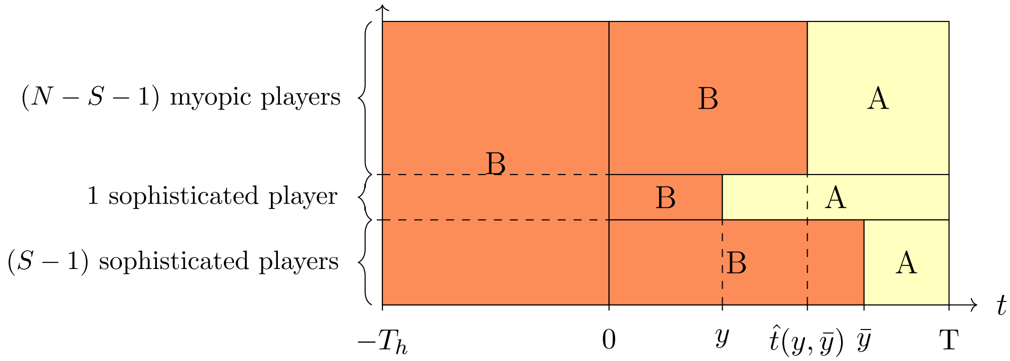


Figure B.7: Illustration of the second case, where $y < \hat{t}(y, \bar{y}) \leq \bar{y}$. The height of the figure shows the fraction of players choosing action A or action B, the width shows the passage of time. The first sophisticated player switches from B to A in period y, other (S-1) sophisticated players switch in period $\bar{y}$ and the myopic players switch in period $\hat{t}$.

The second possibility is that $\hat{t}(y, \bar{y}) \leq \bar{y}$, that is myopic players switch to A earlier than $(S-1)$ sophisticated players. In this case the actual value of $\bar{y}$ will have no influence on the switching period of myopic players, as they will never observe any of the $(S-1)$ sophisticated players choosing A. Therefore the switching period will be a function only of the strategy chosen by a single sophisticated player. At time $t \in (y, \hat{t}]$ beliefs of a myopic player m are $x_m(t)$:

$$x_m(t) = \frac{\int_{k=0}^{t-y} \gamma^k (\frac{1}{N-1})\, \mathrm{d}k}{\int_{k=0}^{t+T_h} \gamma^k \, \mathrm{d}k} = \frac{(\gamma^{t-y} - 1)(\frac{1}{N-1})}{\gamma^{t+T_h} - 1}$$

Player m will choose A in t if:

$$x_m(t) \geq I^{-1} \quad \Leftrightarrow$$

$$\gamma^{t+T_h} \left(\gamma^{-y-T_h} \frac{1}{N-1} - I^{-1} \right) \leq \frac{1}{N-1} - I^{-1} \tag{B.11}$$

If $\frac{1}{N-1} - I^{-1} \leq 0$, equation (B.11) is never satisfied. To see this, notice the following relationship that contradicts (B.11):

$$\gamma^{t+T_h} \left(\gamma^{-y-T_h} \frac{1}{N-1} - I^{-1} \right) > \gamma^{t+T_h} \left(\frac{1}{N-1} - I^{-1} \right) \geq \frac{1}{N-1} - I^{-1}$$

The latter equation holds because $\gamma^{-y-T_h} > 1$, $\gamma^{t+T_h} < 1$ and $\frac{1}{N-1} - I^{-1} \leq 0$.

Alternatively, if $\frac{1}{N-1} - I^{-1} > 0$, (B.11) will be satisfied with equality at time $\hat{t}_2(y) \in (0, \infty)$ that satisfies:

$$\gamma^{-\hat{t}_2(y)} = \frac{\gamma^{-y} \frac{1}{N-1} - \gamma^{T_h} I^{-1}}{\frac{1}{N-1} - I^{-1}} \quad \Leftrightarrow$$

$$\hat{t}_2(y) = \frac{\log(\frac{1}{N-1} - I^{-1}) - \log(\gamma^{-y}\frac{1}{N-1} - \gamma^{T_h} I^{-1})}{\log(\gamma)} \tag{B.12}$$

$\hat{t}(y, \bar{y})$ can be calculated using (B.12) only if $\frac{1}{N-1} - I^{-1} > 0$, otherwise myopic players would never switch from A to B. The switching period if case 2 applies can be expressed as follows:

$$\hat{t}(y, \bar{y}) = \begin{cases} \hat{t}_2(y) & \text{if } \frac{1}{N-1} - I^{-1} > 0 \\ \infty & \text{otherwise} \end{cases} \tag{B.13}$$

Case 3: $\bar{y} < \hat{t} < y$

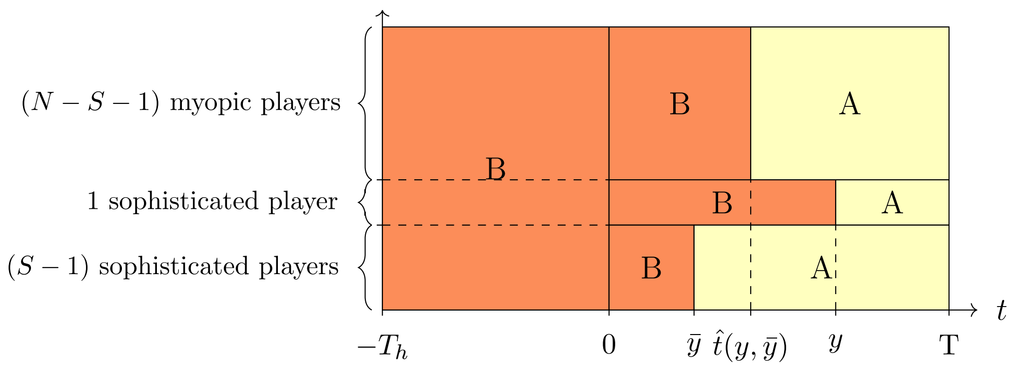


Figure B.8: Illustration of the third case, where $\bar{y} < \hat{t}(y, \bar{y}) \le y$. Height of the figure shows a fraction of players choosing action A or action B, the width shows the passage of time. The first sophisticated player switches from B to A in period y, other $(S-1)$ sophisticated players switch in period $\bar{y}$ and myopic players switch at time $\hat{t}(y, \bar{y})$.

The third possibility is that $\bar{y} < \hat{t}(y, \bar{y}) \le y$, that is at first $(S-1)$ sophisticated players switch to A, then $N-S$ myopic players switch and the last sophisticated player may switch some time after the myopic ones. In this case the switching time is a function only of $\bar{y}$. At time $t \in (\bar{y}, \hat{t}(y, \bar{y})]$ beliefs of a myopic player m are $x_m(t)$:

$$x_m(t) = \frac{\int_{k=0}^{t-\bar{y}} \gamma^k (\frac{S-1}{N-1})\, \mathrm{d}k}{\int_{k=0}^{t+T_h} \gamma^k \, \mathrm{d}k} = \frac{(\gamma^{t-\bar{y}} - 1)(\frac{S-1}{N-1})}{\gamma^{t+T_h} - 1}$$

Player m will choose A in t if:

$$x_m(t) \ge I^{-1} \Leftrightarrow$$

$$\gamma^{t+T_h}\left(\gamma^{-\bar{y}-T_h}\frac{S-1}{N-1} - I^{-1}\right) \le \frac{S-1}{N-1} - I^{-1} \tag{B.14}$$

If $\frac{S-1}{N-1} - I^{-1} \le 0$, condition (B.14) is never satisfied. To see this, notice the following relationship that contradicts (B.14):

$$\gamma^{t+T_h}\left(\gamma^{-\bar{y}-T_h}\frac{S-1}{N-1} - I^{-1}\right) > \gamma^{t+T_h}\left(\frac{S-1}{N-1} - I^{-1}\right) \ge \frac{S-1}{N-1} - I^{-1}$$

The latter conditions holds because $\gamma^{-\bar{y}-T_h} > 1$, $\gamma^{t+T_h} < 1$ and $\frac{S-1}{N-1} - I^{-1} \leq 0$. Therefore if $\frac{S-1}{N-1} - I^{-1} \leq 0$, equation (B.14) is never satisfied and myopic players would choose B at any time t.

Alternatively, if $\frac{S-1}{N-1} - I^{-1} > 0$, (B.14) will be satisfied with equality at time $\hat{t}_3(y) \in (0, \infty)$ that satisfies:

$$\gamma^{-\hat{t}_3(\bar{y})} = \frac{\gamma^{-\bar{y}}\frac{S-1}{N-1} - \gamma^{T_h} I^{-1}}{\frac{S-1}{N-1} - I^{-1}} \Leftrightarrow \tag{B.15}$$

$$\hat{t}_3(\bar{y}) = \frac{\log(\frac{S-1}{N-1} - I^{-1}) - \log(\gamma^{-\bar{y}}\frac{S-1}{N-1} - \gamma^{T_h} I^{-1})}{\log(\gamma)} \tag{B.16}$$

$\hat{t}(y, \bar{y})$ can be calculated using (B.16) only if $\frac{S-1}{N-1} - I^{-1} > 0$. Therefore, the switching period if case 3 applies can be expressed as follows:

$$\hat{t}(y, \bar{y}) = \begin{cases} \hat{t}_3(\bar{y}) & \text{if } \frac{S-1}{N-1} - I^{-1} > 0 \\ \infty & \text{otherwise} \end{cases} \tag{B.17}$$

□

Appendix B.4. Propositions 4, 5 and 6

Proposition 4. *A combination of strategies* $(0, 0)$ *is a symmetric sophisticated player equilibrium ("immediate teaching" equilibrium) if and only if conditions **I1** and **I2** are satisfied:*

$$\frac{S - H/L}{N-1} \leq \gamma^{T_h} I^{-1}, \quad \textbf{(I1)}$$
$$\hat{t}_1(0, 0) \leq T(1 - L/H), \quad \textbf{(I2)}$$

Proof.

$$\begin{array}{lcl}
\Pi(0,0) \geq \Pi(y,0), \forall y \in [0,T] & \xLeftrightarrow{\text{if } \hat{t}_1(0,0) > T} & \Pi_4(0,y^*) \geq \Pi(y,y^*), \forall y \in [0,T] \\
\Updownarrow \text{ if } \hat{t}_1(0,0) \leq T & & \\
\Pi_1(0,0) \geq \Pi(y,0), \forall y \in [0,T] & & \\
\Updownarrow & & \\
\left\{ \begin{array}{l} \Pi_1(0,0) \geq \Pi_1(y,0), \forall y \in [0,y') \\ \Pi_1(0,0) \geq \Pi_4(T,0) \end{array} \right. & \begin{array}{c} \xLeftrightarrow{\text{Lemma 4}} \\ \xLeftrightarrow{\text{Lemma 5}} \end{array} & \begin{array}{l} \textbf{I1} \\ \textbf{I2} \end{array}
\end{array}$$

Figure B.9: Structure of the proof for Proposition 4.

The structure of the proof is shown in Figure B.9. The "immediate teaching" equilibrium exists if (13) is satisfied for $y^* = 0$:

$$\Pi(0,0) \geq \Pi(y,0), \quad \forall y \in [0,T]$$

If $\hat{t}_1(0,0) > T$, condition (12d) is satisfied and equilibrium payoffs are determined by $\Pi(0,0) = \Pi_4(0,0) = 0$, while deviation payoffs are determined by $\Pi(y,0) = yL$. Then an "immediate teaching" equilibrium would not exist because there is a profitable deviation to strategy $y = T$ that provides a payoff of TL. If $\hat{t}_1(0,0) \leq T$, equilibrium payoffs are calculated by $\Pi_1(0,0)$. Condition $\hat{t}_1(0,0) \leq T$ is therefore necessary for the existence of the "immediate teaching" equilibrium. We do not list this condition separately because it is implied by **I2**.

The calculation of payoffs obtained by deviating from the equilibrium depends on the size of the deviation. Payoffs for $y \in [0, \hat{t}_3(0)]$ are calculated by $\Pi_1(y,0)$. Payoffs for $y \in (\hat{t}_3(0), t']$ (where t' solves $\hat{t}_1(t',0) = T$) are dominated by $y = \hat{t}_3(0)$ because increasing y beyond $\hat{t}_3(0)$ does not change the speed of transition but decreases the payoff flow from H to M. If the deviation is larger, that is $y \in [t', T]$, myopic players would never switch to A and deviation profits would be calculated by $\Pi_4(y,0) = yL$. All strategies in this interval would be dominated by strategy $y = T$ that provides a payoff of TL. Overall, two requirements need to be satisfied for an "immediate teaching" equilibrium to exist. First, equilibrium payoffs should be higher than the payoffs from any other $y \in [0, y')$, calculated by $\Pi_1(y,0)$. Lemma 4 specifies the conditions under which this requirement is satisfied. Second, equilibrium payoffs should be higher than the payoff of strategy $y = T$; we derive the conditions for this requirement in Lemma 5. The proofs of these lemmas are in Appendix B.7.

If both **I1** and **I2** hold, equilibrium payoffs are calculated by $\Pi_1(0,0)$ and there are no incentives to deviate to neighbouring strategies or to $y = T$. If one of these conditions was violated, there would be a profitable deviation and the "immediate teaching" equilibrium would not exist.

□

Proposition 5. *A combination of strategies (y^*, y^*) with $y^* \in (0,T)$ is a symmetric sophisticated player equilibrium ("delayed teaching" equilibrium) if and only if conditions **D1**, **D2**, **D3** and **D4** are satisfied:*

$$\hat{t}_1(y^*, y^*) < T, \quad (\mathbf{D1})$$
$$y^* > 0, \quad (\mathbf{D2})$$
$$\hat{t}_1(y^*, y^*) - y^* L/H \leq \hat{t}_2(0), \quad (\mathbf{D3})$$
$$\hat{t}_1(y^*, y^*) - y^* L/H \leq T(1 - L/H), \quad (\mathbf{D4})$$

where equilibrium strategies are calculated by

$$y^* = \frac{\log\left(\frac{S - H/L}{I^{-1}(N-1)}\right)}{\log(\gamma)} - T_h$$

Proof.

The structure of the proof is shown in Figure B.10. Payoffs in a symmetric equilibrium are equal either to $\Pi_1(y^*, y^*)$ or to $\Pi_4(y^*, y^*)$. If condition **D1** holds, condition (12a) will hold as well (Lemma 10), therefore $\Pi(y^*, y^*) = \Pi_1(y^*, y^*)$. If **D1** does not hold, $\Pi(y^*, y^*) = \Pi_4(y^*, y^*) = y^*L$, and a "delayed teaching" equilibrium will not exist because there is a profitable deviation to a strategy $y = T$ that provides a payoff of TL. Condition **D1** is therefore the first necessary existence condition, and we show that it is also jointly sufficient, together with conditions **D2**, **D3** and **D4**. These proofs are shown in additional lemmas. Lemma 6 shows that equilibrium payoffs exceed deviation payoffs if and only if equilibrium payoffs exceed the payoffs of two endpoints, 0 and T, and the payoffs of "neighboring" strategies, calculated by $\Pi_1(y, y^*)$. The proof of Lemma 6 uses a property shown in Lemma 3. Lemmas 7, 8 and 9 show the conditions under which there are no profitable deviations to neighboring strategies (Lemma 7), to $y = 0$ (Lemma 8) and to $y = T$ (Lemma 8). The details of the lemmas and their proofs are in Appendix B.7.

$$
\boxed{
\begin{array}{l}
\Pi(y^*, y^*) \geq \Pi(y, y^*), \forall y \in [0, T] \xLeftrightarrow{\text{if not } \mathbf{D1}} \Pi_4(y^*, y^*) \geq \Pi(y, y^*), \forall y \in [0, T] \\
\qquad\qquad \Updownarrow \text{ if } \mathbf{D1} \\
\Pi_1(y^*, y^*) \geq \Pi(y, y^*), \forall y \in [0, T] \\
\qquad\qquad \Updownarrow \text{ Lemma 6} \\
\left\{
\begin{array}{lcl}
\Pi_1(y^*, y^*) \geq \Pi_1(y, y^*), \forall y \in (y'', y') & \xLeftrightarrow{\text{Lemma 7}} & \mathbf{D2} \\
\Pi_1(y^*, y^*) \geq \Pi_2(0, y^*) & \xLeftrightarrow{\text{Lemma 8}} & \mathbf{D3} \\
\Pi_1(y^*, y^*) \geq \Pi_4(T, y^*) & \xLeftrightarrow{\text{Lemma 9}} & \mathbf{D4}
\end{array}
\right.
\end{array}
}
$$

Figure B.10: Structure of the proof for Proposition 5.

Taken together, Lemmas 7, 8 and 9 prove Proposition 5. Conditions **D1**, **D2**, **D3** and **D4** are jointly sufficient because if all of them are satisfied, there will be no incentives to deviate to any strategy in $[0, T]$. If one of these conditions is violated, the payoff of some strategy will exceed the equilibrium payoff.

□

Proposition 6 *A combination of strategies (T, T) is a symmetric sophisticated player equilibrium ("no teaching" equilibrium) if and only if condition **N1** is satisfied:*

$$\hat{t}_2(0) \geq T(1 - L/H) \qquad (\mathbf{N1})$$

Proof.

If there is a symmetric equilibrium with $y^* = T$, it must hold that:

$$\Pi(T, T) \geq \Pi(y, T), \quad \forall y \in [0, T]$$

The structure of the proof is shown in Figure B.11. Condition (12d) is satisfied, therefore equilibrium payoffs are $\Pi(T, T) = \Pi_4(T, T) = TL$. Deviation payoffs $\Pi(y, T)$ are

$$
\boxed{
\begin{array}{lcl}
\Pi(T,T) \geq \Pi(y,T), \forall y \in [0,T] & & \\
\Updownarrow & & \\
\Pi_4(T,T) \geq \Pi(y,T), \forall y \in [0,T] & & \\
\Updownarrow & & \\
\left\{ \begin{array}{l} \Pi_4(T,T) \geq \Pi_2(0,T) \\ \Pi_4(T,T) \geq \Pi_4(y,T), \forall y \in [y',T] \end{array} \right. & \begin{array}{c} \Longleftrightarrow \\ \Longleftrightarrow \end{array} & \begin{array}{l} \textbf{N1} \\ \text{Always satisfied} \end{array}
\end{array}
}
$$

Figure B.11: Structure of the proof for Proposition 6.

calculated either as $\Pi_4(y,T)$ if $y \in [y',T]$ or as $\Pi_2(y,T)$ if $y \in [0,y')$, where y' solves $\hat{t}_2(y') = T$. In the former case, $\Pi_4(y,T) = yL$, below the payoff of TL provided by strategy $y = T$, therefore the "no teaching" equilibrium exists. In the latter case, deviation payoffs are:

$$\Pi(y,T) = \Pi_2(y,T) = yL + (T - \hat{t}_2(y))H$$

Lemma 3 implies that $\text{argmax}_y(\Pi_2(y,T)) = 0$, that is the most profitable deviation is to strategy $y = 0$. There will be no incentives to deviate to this strategy if the following condition (**N1**) holds:

$$
\begin{aligned}
\Pi_4(T,T) \geq \Pi_2(y,T) \quad &\Leftrightarrow \\
TL \geq (T - \hat{t}_2(0))H \quad &\Leftrightarrow \\
\hat{t}_2(0) \geq T(1 - L/H) &
\end{aligned}
$$

If condition **N1** is satisfied, there will be no incentives to deviate to $y = 0$ and there would be no other profitable deviations, therefore the "no teaching" equilibrium would exist. If **N1** is not satisfied, payoffs could be increased by choosing strategy $y = 0$.

□

Appendix B.5. Corollary 1

Corollary 1. *If a sophisticated player s uses strategy $y_s = y$ and the other sophisticated players use strategies $\mathbf{y}_{-s} = \{(\bar{y})_{\times(S-1)}\}$, the total payoff received by player s over period $[0,T]$ is:*

$$
\Pi(y,\bar{y}) = \begin{cases}
\Pi_1 = yL + (T - \hat{t}_1(y,\bar{y}))H & \text{if } \hat{t}_1(y,\bar{y}) \leq T,\ \hat{t}_2(y) \geq \bar{y},\ \hat{t}_3(\bar{y}) \geq y \\
\Pi_2 = yL + (T - \hat{t}_2(y))H & \text{if } \hat{t}_2(y) < \bar{y} \\
\Pi_3 = \hat{t}_3(\bar{y})L + (y - \hat{t}_3(\bar{y}))M + (T - y)H & \text{if } \hat{t}_3(\bar{y}) < y \\
\Pi_4 = yL & \text{if } \hat{t}_1(y,\bar{y}) > T
\end{cases}
$$

where $\hat{t}_1(y,\bar{y})$, $\hat{t}_2(y)$ and $\hat{t}_3(\bar{y})$ are specified in Proposition 3.

Proof.

The payoff function depends on the switching period of myopic players, which is determined by one of the four equations in condition (8). Each possibility is shown in Figure B.12. Consider panel (a), which illustrates a situation where all sophisticated players switch to A first,[23] and myopic players follow later, therefore their switching time is calculated as $\hat{t}_1(y, \bar{y})$. The participation threshold is not exceeded at any time prior to $\hat{t}_1(y, \bar{y})$ and is exceeded afterwards, therefore the payoff flow of a sophisticated player is L prior to time y, 0 between time y and $\hat{t}_1(y, \bar{y})$ and H afterwards. The sum of payoffs in this case would be equal to $\Pi_1(y, \bar{y}) = yL + (T - \hat{t}_1(y, \bar{y}))H$. Panel (a), however, applies only if myopic players switch after all sophisticated ones, that is if $\hat{t}_2(y) \geq \bar{y}$ and $\hat{t}_3(\bar{y}) \geq y$, and if switching occurs prior to time T.

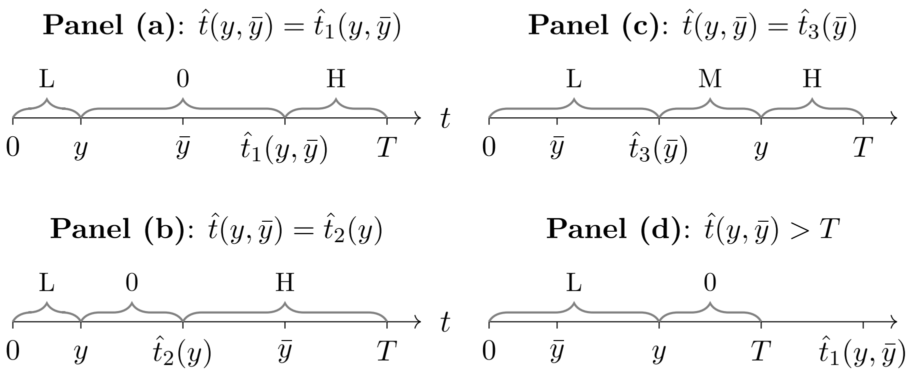


Figure B.12: Stage game payoffs for every possible case. Panel numbering corresponds to equations in (12).

Another possibility is that myopic players switch after observing only one sophisticated player switching to A, a case illustrated in panel (b). Then the sophisticated player will receive a payoff flow equal to L at any time prior to y, a flow of 0 between time y and $\hat{t}_2(y)$ and a flow of H between $\hat{t}_2(y)$ and T. The sum of payoffs in this case would be equal to $\Pi_2(y, \bar{y}) = yL + (T - \hat{t}_2(y))H$. Panel (b) applies only if $\hat{t}_2(y) < \bar{y}$.

In a similar way, $(S - 1)$ sophisticated players may switch first, followed by myopic players and then by a single sophisticated player, illustrated in panel (c). Sophisticated player would receive L until time $\hat{t}_3(\bar{y})$, M from $\hat{t}_3(\bar{y})$ to y and H afterwards. The sum of payoffs would therefore be equal to $\Pi_3(y, \bar{y}) = \hat{t}_3(\bar{y})L + (y - \hat{t}_3(\bar{y}))M + (T - y)H$. Panel (c) applies only if $\hat{t}_3(\bar{y}) < y$.

Finally, myopic players may never switch to A, as illustrated in panel (d). In this case the sophisticated player would receive L until time y, and 0 afterwards, thus the total payoff would be $\Pi_4(y, \bar{y}) = yL$.

□

[23]Panel (a) illustrates the situation with $y < \bar{y}$, but the payoff calculation for $y \geq \bar{y}$ would be equivalent.

Appendix B.6. Corollaries 2, 3 and 4

Corollary 2. *If sophisticated players have a longer planning horizon, then:*

1. *The speed of transition in any equilibrium is not affected.*
2. *"Immediate teaching" equilibrium exists for a larger set of other parameter values.*
3. *"Delayed teaching" equilibrium exists for a larger set of other parameter values.*
4. *"No teaching" equilibrium exists for a smaller set of other parameter values.*

Proof.

Part 1 follows from the definition of the switching period, which depends only on myopic players, and myopic players do not take future payoffs into account. For part 2, note that only condition **I2** depends on the planing horizon, and **I2** is satisfied for a larger set of parameters when T is higher. For part 3, note that conditions **D2** and **D4** depend on the length of the planning horizon, and both are satisfied for a larger set of parameters when T is larger. Part 4 holds because condition **N1** is satisfied for a smaller set of parameters when T is larger.

□

Corollary 3. *If there are more sophisticated players, then:*

1. *Transition is faster in the "delayed teaching" and "immediate teaching" equilibria.*
2. *The effect on the existence of an "immediate teaching" equilibrium or "delayed teaching" equilibrium is ambiguous:*
   (a) *there are more incentives to deviate to neighboring choice plans*
   (b) *there are less incentives to never choose A.*
3. *There is no change in the existence conditions of the "no teaching" equilibrium.*

Proof.

This proof as well as other proofs on comparative statics rely on additional lemmas presented in Appendix B.8. For part 1, see Lemmas 11 and 12. To see part 2 for the "immediate teaching" equilibrium, note that both condition **I1** and condition **I2** depend on player composition. A larger number of sophisticated players leads to **I1** being satisfied for a smaller set of values of other parameters. On the other hand, a larger number of sophisticated players makes condition **I2** satisfied for a larger set of parameters because $\hat{t}_1(0,0)$ is decreasing in S (see Lemma 11). For the "delayed teaching" equilibrium, all four conditions depend on the number of sophisticated players. Incentives to deviate to neighbouring strategies are determined by condition **D2**, which is satisfied for a smaller set of parameters when there are more sophisticated players. To see it, notice that $\frac{\partial y^*}{\partial S} < 0$ (Lemma 12), therefore a higher S decreases y^*, and **D2** is less likely to be satisfied. Incentives to deviate to corner points are determined by conditions **D1**, **D3** and **D4**, all of which are satisfied for a larger set of parameter values when there are more sophisticated players. Conditions **D3** and **D4** are satisfied for a larger set of parameters because $\frac{\partial \hat{t}_1(y^*,y^*)}{\partial S} < \frac{\partial y^*}{\partial S} < \frac{\partial y^*}{\partial S} L/H < 0$ (Lemma 14 and 12). Condition **D1** is also satisfied for a larger set of parameters because $\frac{\partial \hat{t}_1(y^*,y^*)}{\partial S} < 0$, from Lemma 12. Part 3 holds because outcomes in the "no teaching" equilibrium are not affected by the number of sophisticated players.

□

Corollary 4. *If the history of inefficient coordination is longer, then:*

1. *Transition is slower in the "immediate teaching" equilibrium but faster in a "delayed teaching" equilibrium.*
2. *"Immediate teaching" equilibrium exists for a smaller set of other parameter values*
3. *The effect on the existence of a "delayed teaching" equilibrium is ambiguous*
4. *"No teaching" equilibrium exists for a larger set of other parameter values.*

Proof.

Part 1 holds because the derivative of $\hat{t}_1(0,0)$ with respect to T_h is positive while the derivative of $\hat{t}_1(y^*,y^*)$ is negative, as shown in Lemma 11 and Lemma 12. For part 2, parameter T_h affects conditions **I1** and **I2**. An increase in T_h leads to **I1** being satisfied for a smaller set of parameter values, because γ^{T_h} decreases. Condition **I2** is also less likely to be satisfied because of an increase in $\hat{t}_1(0,0)$. For part 3, notice that an increase in T_h satisfies conditions **D1**, **D3** and **D4** for a larger set of parameter values, but satisfies condition **D2** for a smaller set of parameter values. Condition **D1** is satisfied for a larger set of parameter values because $\frac{\partial \hat{t}_1(y^*,y^*)}{\partial T_h} < 0$. Conditions **D3** and **D4** are also satisfied for a larger set of parameter values because $\frac{\partial \hat{t}_1(y^*,y^*)}{\partial T_h} = \frac{\partial y^*}{\partial T_h} < \frac{\partial y^*}{\partial T_h} L/H$ and $\frac{\hat{t}_2(0)}{\partial T_h} > 0$, from Lemma 11, 12 and 13. Condition **D2** is satisfied for a smaller set of parameters because $\frac{\partial y^*}{\partial T_h} < 0$, from Lemma 14. Part 4 holds because $\frac{\partial \hat{t}_2(0)}{\partial T_h} > 0$, from Lemma 13.

□

Appendix B.7. Lemmas 1-10

Lemma 1: *If two choice plans of the sophisticated player prescribe the same action at time t, the payoff flow is higher for the choice plan that induced higher beliefs of myopic players:*

$$\pi[c'_s(t), \{\mathbf{c}_{-s}(t), \mathbf{r}_M(t, \{c'_s, \mathbf{c}_{-s}\})\}] \geq \pi[c''_s(t), \{\mathbf{c}_{-s}(t), \mathbf{r}_M(t, \{c''_s, \mathbf{c}_{-s}\})\}]$$
$$\text{if} \quad x'(t) \geq x''(t) \quad \text{and} \quad c'_s(t) = c''_s(t)$$

where $x'(t)$ is the belief held by myopic players if the sophisticated player uses choice plan c'_s and $x''(t)$ is the belief if the sophisticated player uses choice plan c''_s.

Proof:

Consider two choice plans c'_s and c''_s that prescribe the same action at time t, but different actions prior to time t so that myopic players hold higher beliefs following the history generated by c'_s than by c''_s. Equation (6) in Proposition 1 implies that if myopic players choose A at t following the history generated by c''_s, they must also do so following the history generated by c'_s. Since the choice plans of other strategic players are held constant, a higher tendency to choose A by myopic players increases the total number of other players who choose A at time t. Assumption 1 implies that payoffs are weakly increasing in the number of other players choosing A, therefore the payoff generated by c'_s must be at least as high as the payoff generated by c''_s:

$$\pi[c'_s(t), \{\mathbf{c}_{-s}(t), \mathbf{r}_M(t, \{c'_s, \mathbf{c}_{-s}\})\}] \geq \pi[c''_s(t), \{\mathbf{c}_{-s}(t), \mathbf{r}_M(t, \{c''_s, \mathbf{c}_{-s}\})\}]$$

□

Lemma 2: *All choice plan profiles for sophisticated players with which myopic players switch from A to B are strictly dominated:*

$$AB_M \cap U_s = \emptyset$$

Proof.

Suppose that $\mathbf{c}_S \in AB_M$. Then there are two points in time, t_1 and t_2 (with $t_1 < t_2$), such that myopic players choose A at time t_1 and B at time t_2. Find the first switching period $t_s \in (t_1, t_2]$, such that A is chosen in the interval $[t_1, t_s)$, but B is chosen at time t_s. Since all myopic players share the same history and update beliefs the same way, myopic players will share the same value of t_s and will choose A in the interval $[t_1, t_s)$. If all sophisticated players were also choosing A in the interval $[t_1, t_s)$, the weighted fictitious play rule would imply that $x_m(t_s) \geq x_m(t_1)$ therefore if A was the best-response for a myopic player at time t_1, it will remain a best-response at time t_s, contradicting the definition of t_s. Therefore, a myopic player will choose B at time t_s only if at least one sophisticated player chose B in the interval $[t_1, t_s)$, that is if $c_s(t) = B$ for some $s \in \mathcal{S}$ and $t \in [t_1, t_s)$. Denote the choice plan of this sophisticated player by $\widetilde{c}_s$. We will show that $\widetilde{c}_s$ is dominated by a choice plan c'_s that prescribes A in the entire interval $[t_1, t_s)$ and is otherwise the same as $\widetilde{c}_s$. First, the sum of payoff flows generated by c'_s in the interval $[t_1, t_s)$ is strictly higher than that generated by $\widetilde{c}_s$ because all myopic players are choosing A in this interval, and therefore Assumption 2 implies that the threshold will be exceeded. Second, payoffs generated in the interval $(t_s, T]$ will be equal or higher than those of $\widetilde{c}_s$ because myopic players will hold higher beliefs if c'_s is chosen (due to more A choices being observed) and consequently Lemma 1 implies that higher beliefs will lead to weakly higher payoffs for the sophisticated player at any time $t > t_s$.

□

Lemma 3. $\frac{\partial \hat{t}_2(y)}{\partial y} > 1$.

Proof.

Use the definition of $\hat{t}_2(y)$ from equation (B.12):

$$\hat{t}_2(y) = \frac{\log(\frac{1}{N-1} - I^{-1}) - \log(\gamma^{-y}\frac{1}{N-1} - \gamma^{T_h} I^{-1})}{\log(\gamma)}$$

The partial derivative is calculated as follows:

$$\begin{aligned}\frac{\partial \hat{t}_2(y)}{\partial y} &= \frac{1}{-\log(\gamma)} \times \frac{1}{\gamma^{-y}\frac{1}{N-1} - \gamma^{T_h} I^{-1}} \times \gamma^{-y}\frac{-1}{N-1}\log(\gamma) = \\ &= \frac{\gamma^{-y}\frac{1}{N-1}}{\gamma^{-y}\frac{1}{N-1} - \gamma^{T_h} I^{-1}} > 1\end{aligned}$$

The latter inequality holds because the numerator and the denominator are positive and $\gamma^{T_h I^{-1}} > 0$.

$\square$

Lemma 3 implies that if $\hat{t}_2(0) < T$, it would be optimal for all sophisticated players to choose $y = 0$: increasing y by an amount of ε would increase the payoffs by εL, because of a longer delay, but would simultaneously decrease the payoffs by more than εH because of the longer switching period of myopic players. Lemma 3 therefore shows that when a transition to the efficient state can be achieved by a single sophisticated player, there will be a unique equilibrium in which all sophisticated players immediately use strategic teaching.

Lemma 4. $\Pi_1(0,0) \geq \Pi_1(y,0), \quad \forall y \in [0, y')$ *if and only if condition* ***I1*** *is satisfied:*

$$\frac{S - H/L}{N-1} \leq \gamma^{T_h} I^{-1} \qquad \textbf{\textit{(I1)}}$$

where y' *solves* $\hat{t}_1(y', 0) = T$.

Proof.

Payoffs for any $y \in [0, y')$ are calculated the following way, from equation (12a):

$$\Pi_1(y,0) = yL + (T - \hat{t}_1(y,0))H \tag{B.19}$$

A necessary condition for the payoff to be maximized at $y = 0$ is the non-positive sign of the first derivative of (B.19) with respect to y at $y = 0$. Instead of taking the derivative of the profit function, we first transform it by applying a strictly increasing function $-\gamma^{(\cdot/H)}$, which preserves the sign of the derivative when $\gamma \in (0,1)$. The first derivative of the transformed profit function is non-positive when the following condition holds:

$$\frac{\partial - \gamma^{\Pi_1(y,0)/H}}{\partial y} \leq 0 \quad \Leftrightarrow$$

$$\frac{\log(\gamma)\gamma^T}{I^{-1} - \frac{S}{N-1}} \times \gamma^{-y}\left(\frac{1}{N-1}(L/H - 1) + \frac{S-1}{N-1}L/H - \gamma^{T_h} I^{-1} L/H\right) \leq 0 \quad \Leftrightarrow$$

$$\gamma^{-y}\frac{1}{N-1}(L/H - 1) + \frac{S-1}{N-1}L/H - \gamma^{T_h} I^{-1} L/H \leq 0 \tag{B.20}$$

Inequality (B.20) must hold for $y = 0$:

$$\left.\frac{\partial - \gamma^{\Pi_1(y,0)/H}}{\partial y}\right|_{y=0} \leq 0 \quad \Leftrightarrow$$

$$\frac{1}{N-1}(L/H - 1) + \frac{S-1}{N-1}L/H - \gamma^{T_h} I^{-1} L/H \leq 0 \quad \Leftrightarrow$$

$$L/H\frac{S}{N-1} - \frac{1}{N-1} \leq \gamma^{T_h} I^{-1} L/H \quad \Leftrightarrow$$

$$\frac{S - H/L}{N-1} \leq \gamma^{T_h} I^{-1} \tag{B.21}$$

To obtain the second derivative, differentiate the the left-hand side of (B.20) with respect to y and simplify to get:

$$\frac{\partial^2 - \gamma^{\Pi_1(y,0)/H}}{\partial y^2} = \frac{1}{N-1}(L/H-1)(-1)\log\gamma$$

Note that the second derivative is always negative because $\gamma < 1$ and $H > L$. If condition **I1** is satisfied, the first derivative will be non-positive at point $y = 0$, and it will non-positive for any $y \in (0, t')$. Payoffs would therefore be maximized by choosing $y = 0$. If **I1** does not hold, the first derivative is positive at point $y = 0$ and profits could be increased by choosing $y > 0$.

□

Lemma 5. $\Pi_1(0,0) \geq \Pi_4(T,0)$ *if and only if condition **I2** is satisfied:*

$$\hat{t}_1(0,0) \leq T(1-L/H) \quad \textbf{\textit{(I2)}}$$

Proof.
Deviation payoffs are calculated from (12d): $\Pi_4(T,0) = TL$. There are no incentives to deviate if

$$\Pi_1(0,0) \geq \Pi_4(T,0) \quad \Leftrightarrow \quad \hat{t}_1(0,0) \leq T(1-L/H)$$

□

Lemma 6.

$$\Pi_1(y^*,y^*) \geq \Pi(y,y^*), \forall y \in [0,T] \quad \Leftrightarrow \quad \begin{cases} \Pi_1(y^*,y^*) \geq \Pi_2(0,y^*) \\ \Pi_1(y^*,y^*) \geq \Pi_1(y,y^*), \forall y \in [y'',y'] \\ \Pi_1(y^*,y^*) \geq \Pi_4(T,y^*) \end{cases}$$

If $\hat{t}_3(y^*) \leq T$, then $y' = \hat{t}_3(y^*)$, otherwise y' solves $\hat{t}_1(y',y^*) = T$. If $\hat{t}_2(0) > y^*$, then $y'' = 0$, otherwise y'' solves $\hat{t}_2(y'') = y^*$.

Proof. To specify the deviation payoff, $\Pi(y,y^*)$, we will first look at deviations upwards ($y > y^*$) and then at deviations downwards ($y < y^*$). First, consider a deviation upwards to a strategy $y = y_D > y^*$. The calculation of payoff $\Pi(y_D,y^*)$ depends on the size of the deviation: if y_D is sufficiently small, the payoff is determined by $\Pi(y_D,y^*) = \Pi_1(y_D,y^*)$, but if y is large, myopic players may switch to A prior to y (see an illustration in Figure B.13, panel a), or myopic players may never switch to A (Figure B.13, panel b). The first option is possible only if the myopic players switch to A without ever observing player s choose A, that is if $\hat{t}_3(y^*) < T$. Then the deviation payoffs for a choice plan $y_D \in (\hat{t}_3(y^*), T]$ are calculated by $\Pi_3(y_D,y^*)$. But $\Pi_3(y,y^*)$ is decreasing in y (because $H > M$), thus any strategy in this interval would be strictly dominated by strategy $y = \hat{t}_3(y^*)$. Figure B.13 indicates dominated strategies by an arrow pointing towards the strategy that dominates. Checking for profitable deviations upwards therefore only requires checking for potential deviations in the interval $(y^*, \hat{t}_3(y^*)]$. Also note that $y_D \leq \hat{t}_3(\bar{y})$ together with condition **D1** imply that deviation payoffs for strategies $y_D \in (0, \hat{t}_3(\bar{y}))$ are equal to $\Pi_1(y_D,y^*)$.

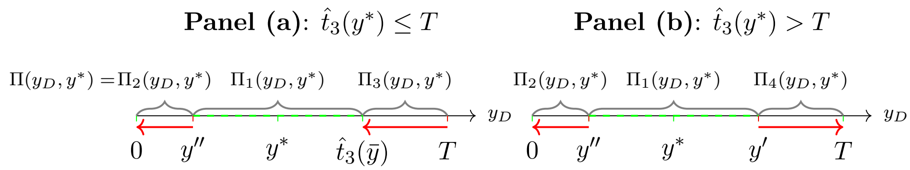


Figure B.13: Calculation of deviation payoffs, $\Pi(y_D, y^*)$ for every possible value of y_D. Green dashed line and green ticks mark undominated strategies. Red arrows mark dominated strategies and the arrow points to the dominant strategy.

The second possibility is that $\hat{t}_3(y^*) \geq T$, so that myopic players do not switch prior to T if they observe only $(S-1)$ sophisticated players switching at y^* (see Figure B.13, panel b). Then because $\hat{t}_1(T, y^*) = \hat{t}_3(y^*) > T$, $\hat{t}_1(y^*, y^*) < T$ (from condition **D1**) and $\hat{t}_1(\cdot, y^*)$ is continuous, there must be a number $y' \in (y^*, T)$ such that $\hat{t}_1(y', y^*) = T$. If $y_D \in (y^*, y']$, (12a) is satisfied and $\Pi(y_D, y^*) = \Pi_1(y_D, y^*)$, because $\hat{t}_1(y, y^*) \leq T$, $\hat{t}_2(y_D) > y_D > y^*$ and $\hat{t}_3(y^*) > T > y^*$. The payoff from any $y_D > y'$ is determined by $\Pi_4(y, y^*) = yL$, and thus all strategies $y_D \in (y', T]$ are dominated by $y_D = T$. Overall, to check for the existence of a "delayed teaching" equilibrium it is sufficient to compare equilibrium payoffs to the payoffs from $y_D \in (y^*, y') \cup T$.

Now consider a possible deviation downwards to $y_D < y^*$. If y_D is only slightly below y^*, the switching period is $\hat{t}_1(y_D, y^*)$ and the deviation payoffs are $\Pi_1(y_D, y^*)$. But if y_D is low enough, myopic players may switch to A prior to y^*, at time $\hat{t}_2(y_D)$. If this does not happen, that is if $\hat{t}_2(0) > y^*$, payoffs from all deviations downwards are calculated by $\Pi_1(y_D, y^*)$. Otherwise, if $\hat{t}_2(0) \leq y^*$, there will be some value y'' that satisfies $\hat{t}_2(y'') = y^*$. For any y below this value, payoffs will be determined by $\Pi_2(y, y^*)$. From Lemma 3, any $y \in (0, y'')$ is dominated by $y = 0$, therefore to check if there are any profitable deviations downwards it is necessary to compare equilibrium payoffs to payoffs from strategies $y_D \in (y'', y^*) \cup 0$.

$\square$

Lemma 7. $\Pi_1(y^*, y^*) \geq \Pi_1(y, y^*), \quad \forall y \in (y'', y')$, *if and only if condition **D2** is satisfied:*

$$\frac{\log(\frac{S-H/L}{I^{-1}(N-1)})}{\log(\gamma)} - T_h > 0, \qquad (\boldsymbol{D2})$$

Proof.

We will calculate the first derivative of the profit function and determine under what conditions the derivative at the equilibrium point is equal to 0 and the second derivative is is non-positive, which ensures that the equilibrium is a local maximum point. We first apply a strictly increasing function $-\gamma^{(\cdot/H)}$ to the profit function an then differentiate the transformed function with respect to y to obtain the following expression:

$$
\begin{aligned}
-\gamma^{\Pi(y,y^*)/H} &= -\gamma^{yL/H+T}\gamma^{-\hat{t}_1(y,y^*)} = \\
&= \frac{1}{I^{-1}-\frac{S}{N-1}}\left(\gamma^{y(L/H-1)+T}\frac{1}{N-1}+\gamma^{yL/H+T-y^*}\frac{S-1}{N-1}-\gamma^{yL/H+T+T_h}I^{-1}\right) = \\
&= \frac{\gamma^{yL/H+T}}{I^{-1}-\frac{S}{N-1}}\left(\gamma^{-y}\frac{1}{N-1}+\gamma^{-y^*}\frac{S-1}{N-1}-\gamma^{T_h}I^{-1}\right) \qquad \text{(B.22)}
\end{aligned}
$$

where $\gamma^{\hat{t}_1(y,y_{-i})}$ has been substituted from (B.8). Differentiate the transformed profit function in (B.22) with respect to y to get

$$
\begin{aligned}
\frac{\partial-\gamma^{\Pi(y,y^*)/H}}{\partial y} &= \frac{\log(\gamma)}{I^{-1}-\frac{S}{N-1}}\times(\gamma^{yL/H+T-y}\frac{1}{N-1}(L/H-1)+ \\
&+\gamma^{yL/H+T-y^*}\frac{S-1}{N-1}L/H-\gamma^{yL/H+T+T_h}I^{-1}L/H) = \\
&= \frac{\log(\gamma)\gamma^{yL/H+T}}{I^{-1}-\frac{S}{N-1}}\left(\gamma^{-y}\frac{1}{N-1}(L/H-1)+\gamma^{-y^*}\frac{S-1}{N-1}L/H-\gamma^{T_h}I^{-1}L/H\right) \\
&\qquad \text{(B.23)}
\end{aligned}
$$

The first derivative is non-negative if:

$$
\begin{aligned}
\frac{\partial-\gamma^{\Pi(y,y^*)/H}}{\partial y} &\geq 0 \quad \Leftrightarrow \\
\gamma^{-y}\frac{L/H-1}{N-1}+\gamma^{-y^*}\frac{S-1}{N-1}L/H-\gamma^{T_h}I^{-1}L/H &\geq 0 \quad \Leftrightarrow \\
\gamma^{-y} &\leq \frac{\gamma^{-y^*}(\frac{S-1}{N-1})-\gamma^{T_h}I^{-1}}{\frac{H/L-1}{N-1}} \qquad \text{(B.24)}
\end{aligned}
$$

The first derivative at point $y=y^*$ is non-negative if:

$$
\begin{aligned}
\left.\frac{\partial-\gamma^{\Pi(y,y^*)/H}}{\partial y}\right|_{y=y^*} &\geq 0 \quad \Leftrightarrow \\
\gamma^{-y^*}\frac{H/L-1}{N-1} &\leq \gamma^{-y^*}\frac{S-1}{N-1}-\gamma^{T_h}I^{-1} \quad \Leftrightarrow \\
\gamma^{T_h+y^*} &\leq \frac{S-H/L}{I^{-1}(N-1)} \quad \Leftrightarrow \\
y^* &\geq \frac{\log\left(\frac{S-H/L}{I^{-1}(N-1)}\right)}{\log(\gamma)}-T_h \qquad \text{(B.25)}
\end{aligned}
$$

The derivative is equal to 0 only if y^* satisfies (B.25) with equality:

$$y^* = \frac{\log\left(\frac{S-H/L}{I^{-1}(N-1)}\right)}{\log(\gamma)} - T_h \tag{B.26}$$

There will be at most one y^* that satisfies (B.26) for any given set of parameters, therefore there can be at most one "delayed teaching" equilibrium in a given game. Equation (B.26) specifies the equilibrium strategy. A necessary condition for the existence of a "delayed teaching" equilibrium is $0 < y^* < T$. But note that condition **D1** from Proposition 5 implies that $y^* < T$ because $y^* < \hat{t}_1(y^*, y^*)$, therefore the only additional condition is that $y^* > 0$.

Condition D2:

$$\frac{\log\left(\frac{S-H/L}{I^{-1}(N-1)}\right)}{\log(\gamma)} - T_h > 0$$

Next, we identify the sign of the second order derivative. The second derivative is obtained by differentiating (B.23) with respect to y:

$$\begin{aligned} \frac{\partial^2 - \gamma^{\Pi(y,y^*)/H}}{\partial y^2} &= \frac{\log(\gamma)^2 \gamma^{yL/H+T}}{I^{-1} - \frac{S}{N-1}} \times (\gamma^{-y}\frac{1}{N-1}(L/H-1)^2 + \\ &+ \gamma^{-y^*}\frac{S-1}{N-1}(L/H)^2 - \gamma^{T_h} I^{-1}(L/H)^2) \end{aligned} \tag{B.27}$$

The second order derivative is negative if:

$$\frac{\partial^2 - \gamma^{\Pi(y,y^*)/H}}{\partial y^2} < 0 \quad \Leftrightarrow$$

$$\gamma^{-y}\frac{1}{N-1}(L/H-1)^2 + \gamma^{-y^*}\frac{S-1}{N-1}(L/H)^2 - \gamma^{T_h} I^{-1}(L/H)^2 > 0 \tag{B.28}$$

If condition **D2** is satisfied, the expression of y^* in (B.26) can be used to rewrite (B.28) as follows:

$$\gamma^{y^*-y} > \frac{L}{L-H} \tag{B.29}$$

Because $L < H$ and $\gamma \in (0,1)$, condition (B.29) is satisfied for all y. The first order condition is therefore both necessary and sufficient for $y = y^*$ to be a local maximum point. Moreover, equation (B.29) states that the second derivative is negative not only at $y = y^*$, but also for any other value of y. Since the first derivative is equal to 0 at point $y = y^*$, and it is decreasing at all y, the payoff function must be increasing at any point $y < y^*$ and decreasing at any point $y > y^*$. Continuity of the profit function therefore implies that $y = y^*$ is not only a local, but also a global maximum in the interval (y'', y) as long as condition **D2** is satisfied.

□

Lemma 8. $\Pi_1(y^*, y^*) \geq \Pi_2(0, y^*)$ *if and only if condition D3 is satisfied:*

$$\hat{t}_1(y^*, y^*) - y^* L/H \;\leq \hat{t}_2(0), \qquad (\boldsymbol{D3})$$

Proof.

Use the profit specification in (8) to get the following expressions for the two profit functions:

$$\Pi_1(y^*, y^*) = yL + (T - \hat{t}_1(y^*, y^*))H$$
$$\Pi_2(0, y^*) = (T - \hat{t}_2(0))H$$

There are no incentives to deviate to $y = T$ if the former expression exceeds the latter:

$$\Pi_1(y^*, y^*) \geq \Pi_2(0, y^*) \quad \Leftrightarrow \quad \hat{t}_1(y^*, y^*) - y^* L/H \;\leq \hat{t}_2(0)$$

□

Lemma 9. $\Pi_1(y^*, y^*) \geq \Pi_4(T, y^*)$ *if and only if condition* ***D4*** *is satisfied:*

$$\hat{t}_1(y^*, y^*) - y^* L/H \leq T(1 - L/H), \qquad (\boldsymbol{D4})$$

Proof.

From (8), deviation payoffs are as follows:

$$\Pi_4(T, y^*) = TL$$

There are no incentives to deviate to $y = T$ if

$$\Pi_1(y^*, y^*) \geq \Pi_4(T, y^*) \quad \Leftrightarrow \quad \hat{t}_1(y^*, y^*) - y^* L/H \leq T(1 - L/H)$$

□

Lemma 10. $\hat{t}_1(y, \bar{y}) \leq \hat{t}_2(y)$ *and* $\hat{t}_1(y, \bar{y}) \leq \hat{t}_3(\bar{y})$

Proof.

Note that $\hat{t}_1(y, \bar{y})$ is increasing both in y and in $\bar{y}$, from equation (B.9). If y is held constant, at any given time t the maximum value of $\hat{t}_1(y, \bar{y})$ will be reached at $\bar{y} = t$. Substituting $\bar{y} = t$ into equation (B.6) reduces it to equation (B.11), thus $\max_{\bar{y}} \hat{t}_1(y, \bar{y}) = \hat{t}_2(y)$. Likewise, setting $y = t$ in equation (B.6) reduces it to equation (B.14), thus $\max_y \hat{t}_1(y, \bar{y}) = \hat{t}_3(\bar{y})$. Therefore $\hat{t}_1(y, \bar{y})$ can never exceed $\hat{t}_2(y)$ or $\hat{t}_3(\bar{y})$.

□

Appendix B.8. Lemmas 11-14 (Comparative Statics)

Appendix B.8.1. Speed of Transition in the Immediate Teaching Equilibrium

The following lemmas show the calculations of comparative statics for $\hat{t}_1(y,y)$ with any y; naturally, they will also hold for the case when $y=0$.

Lemma 11. *Speed of transition to the efficient state in the "immediate teaching" equilibrium depends on the parameter values the following way:*

1. $\frac{\partial \hat{t}_1(y,y)}{\partial y} > 0$
2. $\frac{\partial \hat{t}_1(y,y)}{\partial S} < 0$
3. $\frac{\partial \hat{t}_1(y,y)}{\partial T_h} > 0$

Proof.

Assume that the 'immediate teaching" equilibrium exists, so that $\hat{t}_1(0,0) < T$ and $\frac{S}{N-1} > I^{-1}$. The switching period $\hat{t}_1(y,y)$ is calculated using equation (B.9):

$$\hat{t}_1(y,y) = \frac{1}{-\log(\gamma)}\left[\log\left(\gamma^{-y}\frac{S}{N-1} - \gamma^{T_h}I^{-1}\right) - \log\left(\frac{S}{N-1} - I^{-1}\right)\right] \tag{B.30}$$

1. *Derivative with respect to y:*

$$\frac{\partial \hat{t}_1(y,y)}{\partial y} = \frac{\frac{S}{N-1}}{\frac{S}{N-1} - \gamma^{y+T_h}I^{-1}} \tag{B.31}$$

$\frac{\partial \hat{t}_1(y,y)}{\partial y} > 0$ because $\frac{S}{N-1} > I^{-1}$ and $\gamma \in (0,1)$.

2. *Derivative with respect to S:*

$$\begin{aligned}\frac{\partial \hat{t}_1(y,y)}{\partial S} &= \frac{1}{\log(\gamma)}\frac{1}{N-1}\left(\frac{1}{\frac{S}{N-1} - I^{-1}} - \frac{1}{\frac{S}{N-1} - \gamma^{T_h+y}I^{-1}}\right) = \\ &= \frac{1}{\log(\gamma)}\frac{1}{N-1}\frac{I^{-1}(1-\gamma^{T_h+y})}{(\frac{S}{N-1} - I^{-1})(\frac{S}{N-1} - \gamma^{T_h+y}I^{-1})}\end{aligned} \tag{B.32}$$

$\frac{\partial \hat{t}_1(y,y)}{\partial S} < 0$ because $\frac{S}{N-1} > I^{-1}$ and $\gamma \in (0,1)$.

3. *Derivative with respect to T_h:*

$$\frac{\partial \hat{t}_1(y,y)}{\partial T_h} = -\frac{1}{\log(\gamma)}\frac{1}{\gamma^{-y}\frac{S}{N-1} - \gamma^{T_h}I^{-1}} \times -I^{-1}\gamma^{T_h}\log(\gamma) \tag{B.33}$$

$\frac{\partial \hat{t}_1(y,y)}{\partial T_h} > 0$ because $\frac{S}{N-1} > I^{-1}$ and $\gamma \in (0,1)$.

□

Appendix B.8.2. Speed of Transition in the Delayed Teaching Equilibrium

Lemma 12. *Speed of transition to the efficient state in a "delayed teaching" equilibrium depends on the parameter values the following way:*

1. $\dfrac{\partial \hat{t}_1(y^*, y^*)}{\partial S} < 0$
2. $\dfrac{\partial \hat{t}_1(y^*, y^*)}{\partial T_h} = -1$

Proof.

Assume that a "delayed teaching" equilibrium exists, so that $\hat{t}_1(y^*, y^*) < T$ and $\frac{S}{N-1} > I^{-1}$. In a "delayed teaching" equilibrium, changes in parameter values affect both the equilibrium strategies of sophisticated players and the switching period of myopic players, holding the strategies of sophisticated players constant. To identify the overall effect, we substitute the expression of y^* from equation (B.26) into (B.30) and obtain the following result:

$$\hat{t}_1(y^*, y^*) = \frac{1}{-\log(\gamma)}\left[\log\left(\frac{I^{-1}(N-1)\gamma^{T_h}}{S-H/L}\frac{S}{N-1} - \gamma^{T_h}I^{-1}\right) - \log\left(\frac{S}{N-1} - I^{-1}\right)\right] =$$
$$= \frac{1}{-\log(\gamma)}\left[\log(H/L) + \log(\gamma^{T_h}) + \log(I^{-1}) - \log(S-H/L) - \log\left(\frac{S}{N-1} - I^{-1}\right)\right]$$

1. *Derivative with respect to S:*

$$\frac{\partial \hat{t}_1(y^*, y^*)}{\partial S} = \frac{1}{\log(\gamma)}\left(\frac{1}{S-H/L} + \frac{1}{(\frac{S}{N-1} - I^{-1})(N-1)}\right) \tag{B.34}$$

   $\frac{\partial \hat{t}_1(y^*,y^*)}{\partial S} < 0$ because $\frac{S}{N-1} > I^{-1}$, $\gamma \in (0,1)$ and $S - H/L > 0$ (if a "delayed teaching" equilibrium exists).

2. *Derivative with respect to T_h:*

$$\frac{\partial \hat{t}_1(y^*, y^*)}{\partial T_h} = -1 \tag{B.35}$$

□

Appendix B.8.3. Speed of Transition if One Player is Teaching

We show how the parameters of interest affect $\hat{t}_2(0)$, which measures the transition speed if a single sophisticated player always plays A while all other sophisticated players choose B. This derivative is necessary for Corollary 3 and Corollary 4 because the existence of the "no teaching" equilibrium depends on $\hat{t}_2(0)$.

Lemma 13. *Speed of transition to the efficient state if only one sophisticated player is choosing A depends on the parameter values the following way:*

1. $\dfrac{\partial \hat{t}_2(0)}{\partial S} = 0$
2. $\dfrac{\partial \hat{t}_2(0)}{\partial T_h} > 0$

Proof.

Suppose that $\hat{t}_2(0) < T$, which holds only if $\frac{1}{N-1} > I^{-1}$. Then $\hat{t}_2(0)$ is calculated the following way, from expression (B.12):

$$\hat{t}_2(0) = \frac{\log(\frac{1}{N-1} - I^{-1}) - \log(\frac{1}{N-1} - \gamma^{T_h} I^{-1})}{\log(\gamma)} \tag{B.36}$$

1. *Derivative with respect to S:*

$$\frac{\partial \hat{t}_2(0)}{\partial S} = 0 \tag{B.37}$$

2. *Derivative with respect to T_h:*

$$\frac{\partial \hat{t}_2(0)}{\partial T_h} = \frac{\gamma^{T_h} I^{-1}}{\frac{1}{N-1} - \gamma^{T_h} I^{-1}} \tag{B.38}$$

$\frac{\partial \hat{t}_2(0)}{\partial T_h} > 0$ because $\frac{1}{N-1} > I^{-1}$.

□

Appendix B.8.4. Equilibrium Strategies in the Delayed Teaching Equilibrium

Another variable if interest is the strategy used by sophisticated players in a "delayed teaching" equilibrium, y^*, which has an effect on the existence conditions of the "delayed teaching" equilibrium.

Lemma 14. *The strategies used by sophisticated players in a "delayed teaching" equilibrium depend on parameter values the following way:*

1. $\dfrac{\partial y^*}{\partial S} < 0$
2. $\dfrac{\partial y^*}{\partial T_h} = -1$

   In addition:

3. $\dfrac{\partial y^*}{\partial S} > \dfrac{\partial \hat{t}_1(y^*, y^*)}{\partial S}$

Proof.
Equilibrium strategy is determined by equation (B.26):

$$y^* = \frac{\log\left(\frac{S-H/L}{I^{-1}(N-1)}\right)}{\log(\gamma)} - T_h$$

1. *Derivative with respect to S:*

$$\frac{\partial y^*}{\partial S} = \frac{1}{\log(\gamma)}\left(\frac{1}{S-H/L}\right) \tag{B.39}$$

   $\frac{\partial y^*}{\partial S} < 0$ because $\gamma \in (0,1)$ and $S-H/L > 0$ (because a "delayed teaching" equilibrium exists).

2. *Derivative with respect to T_h:*

$$\frac{\partial y^*}{\partial T_h} = -1 \tag{B.40}$$

3. *Comparison to the derivative of $\hat{t}_1(y^*, y^*)$:*
   Recall the derivative of $\hat{t}_1(y^*, y^*)$ with respect to S from equation (B.34):

$$\frac{\partial \hat{t}_1(y^*, y^*)}{\partial S} = \frac{1}{\log(\gamma)}\left(\frac{1}{S-H/L} + \frac{1}{(\frac{S}{N-1} - I^{-1})(N-1)}\right)$$

   The derivative of y^* calculated in (B.39) is strictly higher (closer to 0) than the derivative of $\hat{t}_1(y^*, y^*)$ because $\log(\gamma) < 0$ and $\frac{1}{(\frac{S}{N-1}-I^{-1})(N-1)} > 0$.

□