On the empirical relevance of correlated equilibrium

Daniel Friedman [a,b], Jean Paul Rabanal [c], Olga A. Rud [c], Shuchen Zhao [d,*]

[a] *Economics Departments, University of California Santa Cruz, United States of America*
[b] *University of Essex, United Kingdom of Great Britain and Northern Ireland*
[c] *Department of Economics and Finance, University of Stavanger, Norway*
[d] *Institute for Advanced Economic Research, Dongbei University of Finance and Economics, 217 Jianshan Street, Dalian, 116025, China*

Abstract

Absent coordinating signals from an exogenous benevolent agent, can an efficient correlated equilibrium emerge? Theoretical work in adaptive dynamics suggests a positive answer, which we test in a laboratory experiment. In the well-known Chicken game, we observe time average play that is close to the asymmetric pure Nash equilibrium in some treatments, and in other treatments we observe collusive play. In a game resembling rock-paper-scissors or matching pennies, we observe time average play close to a correlated equilibrium that is more efficient than the unique Nash equilibrium. Estimates and simulations of adaptive dynamics capture much of the observed heterogeneity across player pairs as well as dynamic regularities.

JEL classification: C72; C73; C92

Keywords: Correlated equilibrium; Laboratory experiment; Adaptive dynamics

1. Introduction

A probability distribution φ over the set of action profiles in a normal form game is a *correlated equilibrium* (CE) if no player has an incentive to deviate from any action j in her component

* Corresponding author.
E-mail addresses: dan@ucsc.edu (D. Friedman), jeanpaulrab@gmail.com (J.P. Rabanal), olga.rud@gmail.com (O.A. Rud), fenix.zhaoshuchen@dufe.edu.cn (S. Zhao).

of the support of φ. As an equilibrium concept, CE has several advantages over Nash equilibrium (NE). First, as pointed out by Aumann (1974), the rationality assumptions are much more attractive – a common prior and common knowledge of rationality suffice for CE but do not suffice for NE. Second, CE is more general in that it does not require statistical independence, while φ must be a player-by-player product distribution in mixed NE. Third, the set of all CE is convex and much easier to compute than the set of all NE, which may have disconnected components. Perhaps for these reasons, CE is the standard game theoretical equilibrium concept for many computer scientists.

Nevertheless, correlated equilibrium is of only minor interest to most economists and other social scientists. We believe that two obstacles block more widespread interest. First, the standard interpretation of CE is that the distribution φ is implemented via an exogenous benevolent agent who recommends a particular action to each player. Such agents, such as a traffic cop who cannot stop to write tickets, exist only in very special circumstances. Second, there are very few well known examples of games with distinctive CE, e.g., that achieve greater efficiency and fairness than any NE.

The present paper attempts to surmount those obstacles by means of a laboratory experiment with no exogenous agents. We investigate two games (one better known than the other) with efficient and fair CE. Our project addresses general questions such as: in relevant games, does the overall average joint distribution approximate a product distribution? If so, is it close to a NE distribution? If not, is it close to an efficient and fair CE distribution? Does the joint distribution display dynamic regularities? If so, are they consistent with relevant adaptive models?

There are several levels at which an equilibrium concept such as CE may be empirically relevant. At the aggregate level, it might accurately predict the distribution of play averaged over time and over player populations. At a finer level, it might predict time average play in a particular ongoing game. Finally, it might predict cross-sectional average play in a population playing one-shot games. Much of the theoretical work with CE has focused on the last level, as have most previous laboratory experiments. It is at that level that the traffic cop interpretation is most cogent, but even there it is not required. The definition of CE requires only that, whenever player i happens (for any reason) to choose a strategy j that is part of a CE, her beliefs do not discourage her from playing j, and those beliefs are not contradicted by her personal experience. We will see shortly that such beliefs can arise from certain sorts of adaptive processes, and that their long run behavior may converge to CE. Such processes will be central to our experiment.

We study two symmetric bimatrix games. The first is the well known Chicken game (CH), which has two pure NE as well as a mixed strategy NE, and an attractive collusive strategy profile. Our second game is a variant of Matching Pennies proposed by Moulin and Vial (1978) (MV) with a mixed strategy NE, and no pure strategy NE nor an attractive collusive profile. The two games are quite different, but each has a "target" CE that is more efficient and fair than any of the NE.

Our experiment features long sequences of repeated matches for player pairs but, consistent with our research questions, we are mainly concerned with time average play and how it compares to the stage game equilibrium. Fortunately, those research questions are best addressed in low information environments that are not conducive to well-known repeated game strategies such as grim trigger. In our low information environment, players do not know the payoffs of other players, but they always know either their own historical average payoff for each action or, alternatively, the counterfactual payoff that they would have earned if they had always made a different choice. For robustness, we also examine a (relatively) high information treatment where players know other players' previous actions and payoffs.

The results show that CE has empirical relevance, but within limits. With more conducive treatments (low information, counterfactual regret), time average behavior in both games is roughly consistent with CE on the aggregate level and at the player pair level. Behavior is also consistent with NE in the Chicken game but not in the MV game. With less conducive treatments, we see widespread collusion in the Chicken game, while in the MV game behavior is still roughly consistent with the target CE.

We work with adaptive dynamics models to estimate how players respond to "regret" given counterfactual or, alternatively, historical average payoff information. Using a logit model, we find that players respond strongly to positive regret where they could have earned a higher payoff from an alternative action. When we allow for inertia, data shows heterogeneity in subject responses to negative regret.

In the following section we present an overview of the literature, including a handful of experiments studying CE, followed by the theoretical foundations for our experiment. They derive largely from Foster and Vohra (1998), and Hart and Mas-Colell (2000, 2001) who propose regret-based adaptive dynamics that promote convergence of the time average distribution φ to the CE set. The intuition is that adaptive behavior leads players away from "regrettable" actions that strongly violate the inequalities that define CE. We also run simulations that shed new light on the behavior produced by such models. Results from our experiments are then presented, first at the aggregate level, then at the player-pair level, and finally at the level of period-by-period individual player response. A concluding discussion briefly summarizes our findings, and points to new avenues for future research. Supplementary appendices offer further data analysis and simulation results.

2. Previous literature

Early literature on correlated equilibrium, e.g., Aumann (1974, 1987) and Brandenburger and Dekel (1987), focused on epistemic foundations. These papers show that common knowledge of rationality (together with a common prior) suffices to achieve CE. By contrast, Aumann and Brandenburger (1995) show that to ensure NE a profile of conjectures must arise from a common prior, and players must have mutual knowledge of the payoff functions and of rationality, and common knowledge of the conjectures.

Forges and Peck (1995) illustrate how the concept of CE covers sunspot equilibrium in exchange markets. Moulin and Vial (1978) propose a new equilibrium concept called coarse correlated equilibria (CCE): the set of probability distributions over action profiles that can be supported if each player either commits to playing according to the recommendation of a device, or else plays freely with no access to the device recommendation. They note that the equilibrium sets satisfy NE $\subset$ CE $\subset$ CCE.

The theoretical literature on convergent dynamics is especially relevant. Fudenberg and Levine (1999) propose a smooth fictitious play procedure which guarantees almost sure convergence to the set of correlated ϵ-equilibria. Foster and Vohra (1997) introduce "calibrated" strategies with the property that time average counterfactual regret (defined in the next section) converges to zero irrespective of what strategies the other players adopt. The authors show that time average play will converge to CE if all players adopt calibrated strategies. As explained in the next section, Hart and Mas-Colell (2000, 2001) introduce specific regret-responsive strategies that, if followed by all players, guarantee convergence to the set of CE. Metzger (2018) models evolutionary dynamics given exogenous coordinating signals, and finds that they can lead to CE outcomes that are not NE. Arifovic et al. (2019) implement evolutionary learning simulations and obtain outcomes similar to those seen in the experiments mentioned below of Duffy and Feltovich (2010) and Duffy et al. (2017).

It may be worth noting that some learning rules which incorporate stochastic choice can also converge to a pure NE in two player games, using either sampling and regret computations (Foster and Young, 2006), or trial and error (Young, 2009). In the absence of a pure NE, Pradelski and Young (2012) propose that the long run behavior under a log linear learning rule depends on the sum of payoff across all players, and the gain from a unilateral deviation by some player. In a low information public goods experiment, Nax et al. (2016) find that inertia, reversion and reinforced adjustment are key features of learning dynamics.

Thus far, empirical work on CE has been sparse and has focused mainly on laboratory experiments with exogenous coordinating signals. Chicken (or hawk-dove) games with private recommendations have been studied by Cason and Sharma (2007) and Duffy and Feltovich (2010). The former study concludes that recommendations are followed when subjects play automated counterparties who always follow recommendations, while the latter study finds that players follow recommendations when they implement a CE that is more efficient than relevant NE. Another game that has been studied in the laboratory is Battle of the Sexes (BoS). Duffy et al. (2017) find that direct messages improve coordination on a CE, relative to indirect messages; Anbarcı et al. (2018) find that subjects are more likely to follow recommendations when payoffs are more symmetric; and Bone et al. (2013) find that subjects follow recommendations from a public device and coordinate more in a symmetric BoS than in the game of Chicken. Georgalos et al. (2020) consider a game proposed by Moulin and Vial (1978) that has a pure NE and a Pareto superior CCE. They find that a small and declining fraction of subjects choose the correlation device for that CCE; in key treatments most players prefer free choice, and thus their play eventually approximates the pure NE.

In an experiment without an external device but allowing for preplay communication, Moreno and Wooders (1998) study a constant sum, three-player matching pennies game, and find that average play is a noisy version of a non-NE target CE. Palfrey and Pogorelskiy (2019) find support for CE in a voter turnout game with communication within parties. Furthermore, without an external correlation device, Cason et al. (2020) show that correlation of beliefs in some prisoner's dilemma games is less frequent than in hawk-dove (a variant of chicken) games and some coordination games. In addition, correlated equilibrium can also be achieved in repeated play with turn taking strategies (e.g., Zhao, 2021).

The present empirical paper seems to be the first to focus on whether convergence to CE can arise from a regret-based adaptive process that requires no signal nor external devices nor communication among players.

3. Theoretical considerations

A probability distribution φ over the set $S = S^1 \times \ldots \times S^N$ of action profiles in an N player normal form game Γ is a *correlated equilibrium* (CE) if, for every player $i \in N$ and every two actions $j, k \in S^i$, we have

$$\sum_{s \in S: s^i = j} \varphi(s)[u^i(k, s^{-i}) - u^i(s)] \leq 0. \qquad (1)$$

That is, no player has an incentive to deviate from any action j in her component of support φ. Nash equilibrium (NE) is the special case where φ is a product distribution, i.e., each player's realization in the mix is independent of other players' realizations. By contrast, in (1) the

Table 1
Chicken (CH) game.

	L	R
U	100, 100	600, 200
D	200, 600	500, 500

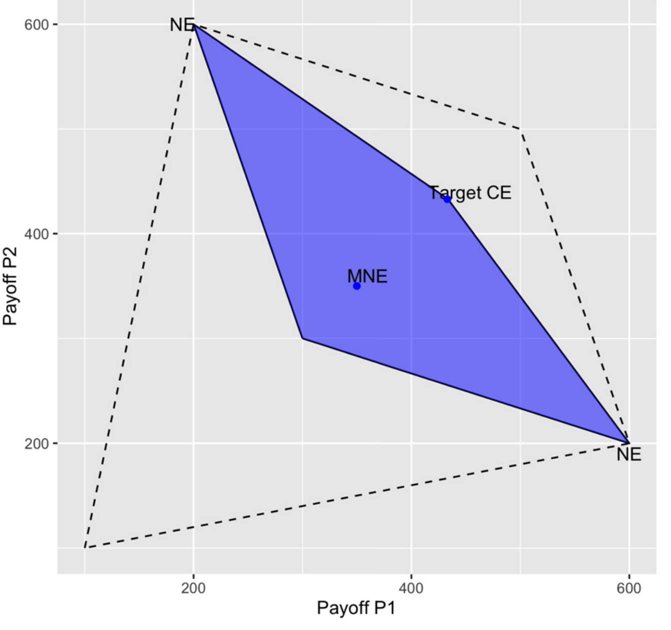


Fig. 1. **Payoff space for Chicken game (CH).** The blue shaded region represents payoff vectors for all CE profiles, while payoffs for all feasible profiles are bounded by the dashed lines. MNE marks the mixed NE payoff vector $(350, 350)$. It is Pareto dominated by the target CE payoff $(\frac{1300}{3}, \frac{1300}{3})$, which is dominated by the collusion payoff (500,500). NE denotes the asymmetric pure NE payoffs, e.g., (200,600), whose sum is less than that of the target CE. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

realizations can be correlated. The set of CE consists of all joint distributions that satisfy all the linear inequalities of equation (1). That set is convex and can be computed using standard linear programming packages; see Appendix A for an example.

3.1. Example games

Table 1 shows a 2×2 bimatrix for the well-known Chicken Game (CH) with three possible NE: two asymmetric pure NE $(U, R), (D, L)$ and the symmetric mixed NE $(0.5, 0.5) = \frac{1}{4}$(U, L) + $\frac{1}{4}$(D, L) + $\frac{1}{4}$(U, R) + $\frac{1}{4}$(D,R). It also has a CE, $\frac{1}{3}$(D, L) + $\frac{1}{3}$(U, R) + $\frac{1}{3}$(D,R), that we will refer to as our *target CE*. The target CE, with expected payoff vector $(\frac{1300}{3}, \frac{1300}{3})$, has higher efficiency (i.e., higher payoff sum) and is at least as fair (i.e., smaller payoff difference) as any of the NE. Note that pure collusion (D,R) is the most efficient and fair profile, but it is not a CE (nor, a fortiori, a NE).

The set of mixed profiles for CH is a three dimensional subset of four-dimensional space. It may be more useful to consider the two-dimensional payoff space shown in Fig. 1, where the

Table 2
Moulin and Vial (1978) (MV) game.

	L	C	R
T	0, 0	100, 200	200, 100
M	200, 100	0, 0	100, 200
B	100, 200	200, 100	0, 0

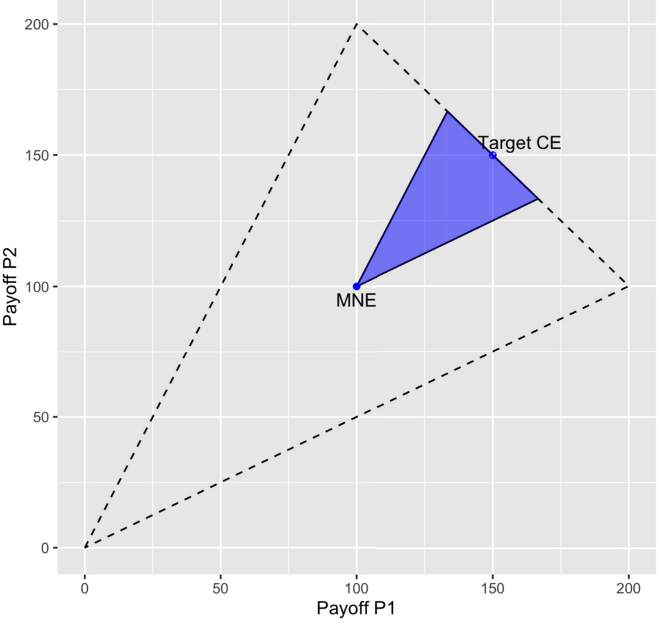


Fig. 2. **Payoff space for MV game.** The blue region represents CE payoff vectors, while the dashed boundary encloses all feasible payoff vectors. The MNE payoff (100,100) is Pareto dominated by the target CE payoff (150,150).

payoff vectors associated with CE lie in the blue quadrilateral, which covers $\frac{4}{15} \approx 27\%$ of the feasible payoff region.[1]

Table 2 presents a 3x3 game based on Moulin and Vial (1978). This game, denoted MV below, resembles the rock-paper-scissors game in that it has a best response cycle.[2] It is also reminiscent of matching pennies in that the off-diagonal profiles are constant sum, and therefore provide little scope for collusion. The MV game has a unique NE: the symmetric uniform independent mixture $(1/3, 1/3, 1/3) \times (1/3, 1/3, 1/3)$. This mixed NE is Pareto dominated by the *target CE*, which assigns a probability of zero to the main diagonal profiles and a probability of $\frac{1}{6}$ to each off-diagonal profile. Fig. 2 shows the feasible payoff space, and the CE region which covers $1/3^2 \approx 11\%$ of the feasible space.

[1] The ratio of the areas is the fraction of the main diagonal in the blue region. The main diagonal has length $500\sqrt{2}$ and its blue segment has length $(\frac{1300}{3} - 300)\sqrt{2}$.

[2] For later reference, note that the best response cycle in MV tours all six off-diagonal profiles in a particular order: ...→(T,C)→(B,C)→(B,L)→(M,L)→(M,R)→(T,R)→(T,C)→...

3.2. Regret

We now present adaptive dynamic models, specifying how players compare current action j to alternative actions k and how they use these comparisons to choose next period's action. The models closely follow Hart and Mas-Colell (2000), and require only that players know the history of realized profiles so far and their own payoff function. They do not require the knowledge of other players' payoffs, nor an exogenous agent to provide coordinating signals. Instead, correlation comes from a common history of play.

To formalize, suppose that the normal form stage game $\Gamma = (S, u)$ is played repeatedly in discrete time $t = 1, 2, \ldots, T \leq \infty$. Let $s_t^i \in S^i$ denote the (realized) choice of player i at time t, and let $u^i(s_t)$ denote the payoff of player i in period t. For any two distinct actions $j \neq k \in S^i$ for any given player i at any time $t < T$, suppose that player i had replaced action j, every time $\tau \leq t$ that it was played so far, by action k, with no other changes in the profiles. Then player i's payoff at time τ becomes $u^i(k, s_\tau^{-i})$ if $s_\tau^i = j$. **Unnormalized counterfactual (UC) regret** is the resulting difference in i's per period payoff so far,

$$\hat{r}_t^i(j, k) = \frac{1}{t} \sum_{\tau \leq t : s_\tau^i = j} \left[u^i(k, s_\tau^{-i}) - u^i(s_\tau) \right]. \tag{2}$$

The intuition is that a player may wish that they had done something different to the extent that it would have generated a higher payoff in the past, assuming no impact on other players' choices.

For the purpose of our experiment, it is useful to define **counterfactual (C) regret**, a normalization of (2) which looks only at periods where the counterfactual is relevant,

$$r_t^i(j, k) = \frac{\sum_{\tau \leq t : s_\tau^i = j} u^i(k, s_\tau^{-i})}{|\tau \leq t : s_\tau^i = j|} - \frac{\sum_{\tau \leq t : s_\tau^i = j} u^i(s_\tau)}{|\tau \leq t : s_\tau^i = j|} \equiv m_t^i(j, k) - M_t^i(j). \tag{3}$$

The last expression summarizes counterfactual regret as the signed difference between $m_t^i(j, k)$, which is the per-period mean counterfactual payoff of playing k instead of j whenever j was actually played, and $M_t^i(j)$, the mean actual payoff so far from playing j. Such normalization (averaging payoffs only for periods when j is played) is useful because, for any strategy j that is rarely played, the UC regret (2) will automatically become small as t gets large, regardless of whether the alternatives to j have relatively high or low payoffs.

Additionally, we work with **average (A) regret**, defined as the difference in per period actual payoffs between an alternative action and the current action,

$$R_t^i(j, k) = \frac{\sum_{\tau \leq t : s^i = k} u^i(s_\tau)}{|\tau \leq t : s^i = k|} - \frac{\sum_{\tau \leq t : s^i = j} u^i(s_\tau)}{|\tau \leq t : s^i = j|} \equiv M_t^i(k) - M_t^i(j). \tag{4}$$

Average regret is of interest due to its informational economy: players can use it even when they do not know their own payoff function nor other players' actions, but do know the history of their own actions and realized payoffs.

How do the players respond to regret? According to Hart and Mas-Colell (2000), the action of player i in period $t + 1$ is chosen via a linear probability model using an appropriate version of regret. Applied to UC regret (2), the **HM response** rule states that given current action j, a player will choose action k next period with probability

$$p_{t+1}^i(k) = \frac{1}{\mu} \hat{r}_t^i(j, k)_+, \quad k \neq j \tag{5}$$

$$p_{t+1}^i(j) = 1 - \sum_{k \in S^i : k \neq j} p_{t+1}^i(k); \text{ where}$$
$$\hat{r}_t^i(j,k)_+ = \max\{\hat{r}_t^i(j,k), 0\},$$

and the parameter $\mu > 0$ is large enough to ensure that $p_{t+1}^i(j)$ is positive, i.e., that the player will always continue to play strategy j with a positive probability. A larger μ indicates greater inertia. In continuous time, the degree of inertia can be tied to a sampling interval $\Delta t < 1$ by changing the scaling factor from $\frac{1}{\mu}$ to $\frac{\Delta t}{\mu}$.

Note that the HM response rule in (5) is applied to $\hat{r}_+$, i.e., after truncating UC regret below at zero. That truncation ensures that choice probabilities will be non-negative, but it also has a substantive interpretation: insensitivity to the extent of negative regret. Whether $\hat{r}(j,k)$ is very negative or only slightly negative doesn't matter; the player is not at all tempted to switch from current action j to any alternative action k with negative regret. Of course, the HM response rule can similarly be applied to C or A regret, again after truncating at zero.

An alternative response rule, widely used in empirical literature, is for choices to follow the logit model. Applied to C regret, the **logit response** rule with parameter $\beta > 0$ sets

$$p_{t+1}^i(k) = \frac{e^{\beta r_t^i(j,k)}}{\sum_{\ell \in S^i} e^{\beta r_t^i(j,\ell)}} = \frac{e^{\beta m_t^i(j,k)}}{\sum_{\ell \in S^i} e^{\beta m_t^i(j,\ell)}}, \quad \text{for all } k \in S^i. \tag{6}$$

The last expression uses equation (3) to cancel the common factor $e^{-\beta M_t^i(j)}$ in the numerator and the denominator. The parameter $\beta > 0$ measures the intensity of response to regret. The rule can also be applied to UC or A regret.

Compared to (5), equation (6) assigns no special inertia to the current action, even though that might matter in practice. Therefore we also consider the **inertial logit response** rule with parameters $\beta > 0$ and $\Delta \in (0, 1]$, defined as

$$p_{t+1}^i(k) = \frac{e^{\beta r_t^i(j,k)} \Delta}{\sum_{\ell \in S^i} e^{\beta r_t^i(j,\ell)}}, \quad k \neq j \tag{7}$$
$$p_{t+1}^i(j) = 1 - \sum_{k \in S^i : k \neq j} p_{t+1}^i(k),$$

where j denotes the action of player i at time t. When Δ is small, the player is more likely to stay with the current strategy j. Of course, when the player does switch, the relative probabilities of different alternative actions are governed by logit choice independent of Δ.

3.3. Hart and Mas-Colell convergence results

Hart and Mas-Colell (2000) Main Theorem: Given any finite normal form game $\Gamma = (S, u)$, suppose that in the infinitely repeated game with stage game Γ, every player follows the HM response rule (5) applied to UC regret (2). Then the time average profile $z_t = \frac{1}{t}\sum_{\tau \leq t} s_\tau$ converges almost surely to the set CE of correlated equilibria as $t \to \infty$.

Recall that counterfactual regret (either C or UC) implicitly assumes that each player i knows the complete profile history and her own payoff function. A follow-up paper, Hart and Mas-Colell (2001), considers average (A) regret, which requires players to know only their own actions and own payoffs so far. The authors prove convergence to CE under A regret, but only after substantial modifications to the response rule, e.g., including diminishing trembles.

These conclusions are impressively broad in that they apply to all finite normal form games. However, the conclusions concern convergence to the CE set, and not necessarily to some point within the set. They offer few hints about when a regret-based adaptive process will converge to a particular CE of interest. Nor do they tell us much about what will happen with more easily implemented forms of regret or with standard response rules such as logit.[3] We investigate such matters via simulations in the next section, and thereby seek to generate sharper testable hypotheses for our experiment.

4. Simulations

We conduct simulations for the CH game and for the MV game introduced in Section 3.1, employing three types of regret (UC, C, and A) and three response rules (HM, logit, and inertial logit). Baseline parameter values are $(\mu, \beta, \Delta) = (600, 1, 0.8)$, where μ and $\Delta > 0$ capture the level of inertia respectively for HM and for inertial logit response. The restriction $\mu > \max_{i,j,k}(m^i - 1)|u^i(k, s^{-i}) - u^i(j, s^{-i})|$, where m^i is player i's number of actions, ensures that the current action is played next period with positive probability. The restriction $\Delta \in (0, 1]$ ensures greater inertia in inertial logit than in plain logit. Baseline parameters easily satisfy these restrictions, e.g., that $\mu > 500$ in the CH game and $\mu > 400$ in the MV game. The precision parameter for logit response is only restricted to $\beta > 0$, which again is easily met at baseline. The qualitative behavior described below is robust to a substantial range around baseline parameter values. For example, adjustment speed varies but long run behavior seems unaffected by varying $\mu \in [600, 2000]$ and $\Delta \in [0.5, 1]$. On the other hand, logit dynamics for $\beta < 0.1$ are so noisy that behavior approximates the MNE.

For each of the 3×3 combinations of regret type and response rule, we run 500 simulation trials, with each trial running for 500 periods. To initialize regret, we draw iid actions for the first 50 periods[4] and then, for the remaining 450 periods, we follow the specified combination of dynamics. The analysis below omits the first 50 periods in each trial.

4.1. CH games

Table 3 reports the empirical aggregate joint action distributions for CH game simulations. For UC and C regret, most of the mass is in the pure NE profiles (U,R) and (D,L). There is little difference between the UC and C data, suggesting that it is safe to normalize. There is not much difference across the three response rules either, suggesting some robustness to different dynamic processes. However, there is a remarkable difference between A regret and the counterfactual regrets. Average regret supports a much higher frequency of collusion, exceeding the 0.33 predicted by the target CE.

Next, we consider the less aggregated pair level data. Fig. 3 shows the time average payoff for each simulation trial in the CH game. The left and center panels confirm that there is little difference between UC and C regret, and that convergence is primarily to the two pure NEs. On the other hand, we do see substantial differences across the response rules in that MNE seems

[3] Nor do the authors offer much intuition on the reasons for convergence. Note that replacing the true population distribution φ in the CE-defining inequalities (1) by a particular sort of historical estimate, e.g. (2), will yield expressions that say the corresponding sort of regret is non-positive. So our rough intuition is that convergence to CE arises in the long run when all players favor regret-free actions.

[4] The dynamics are robust to some other forms of random initialization, e.g., with 100 or more iid random profiles.

Table 3
Profile distributions in simulations of the CH game.

HM response, Eq. (5)

	UC		*C*		*A*	
	L	R	L	R	L	R
U	0.018	0.497	0.011	0.467	0.073	0.245
D	0.467	0.018	0.509	0.012	0.273	0.408

logit response, Eq. (6)

	UC		*C*		*A*	
	L	R	L	R	L	R
U	0.128	0.323	0.138	0.393	0.034	0.219
D	0.421	0.128	0.331	0.138	0.212	0.535

inertia logit response, Eq. (7)

	UC		*C*		*A*	
	L	R	L	R	L	R
U	0.026	0.506	0.032	0.450	0.032	0.209
D	0.443	0.025	0.485	0.032	0.178	0.580

Notes: Regret varieties are unnormalized counterfactual (UC), counterfactual (C) and Average (A). Simulations use baseline parameters $(\mu, \beta, \Delta) = (600, 1, 0.8)$.

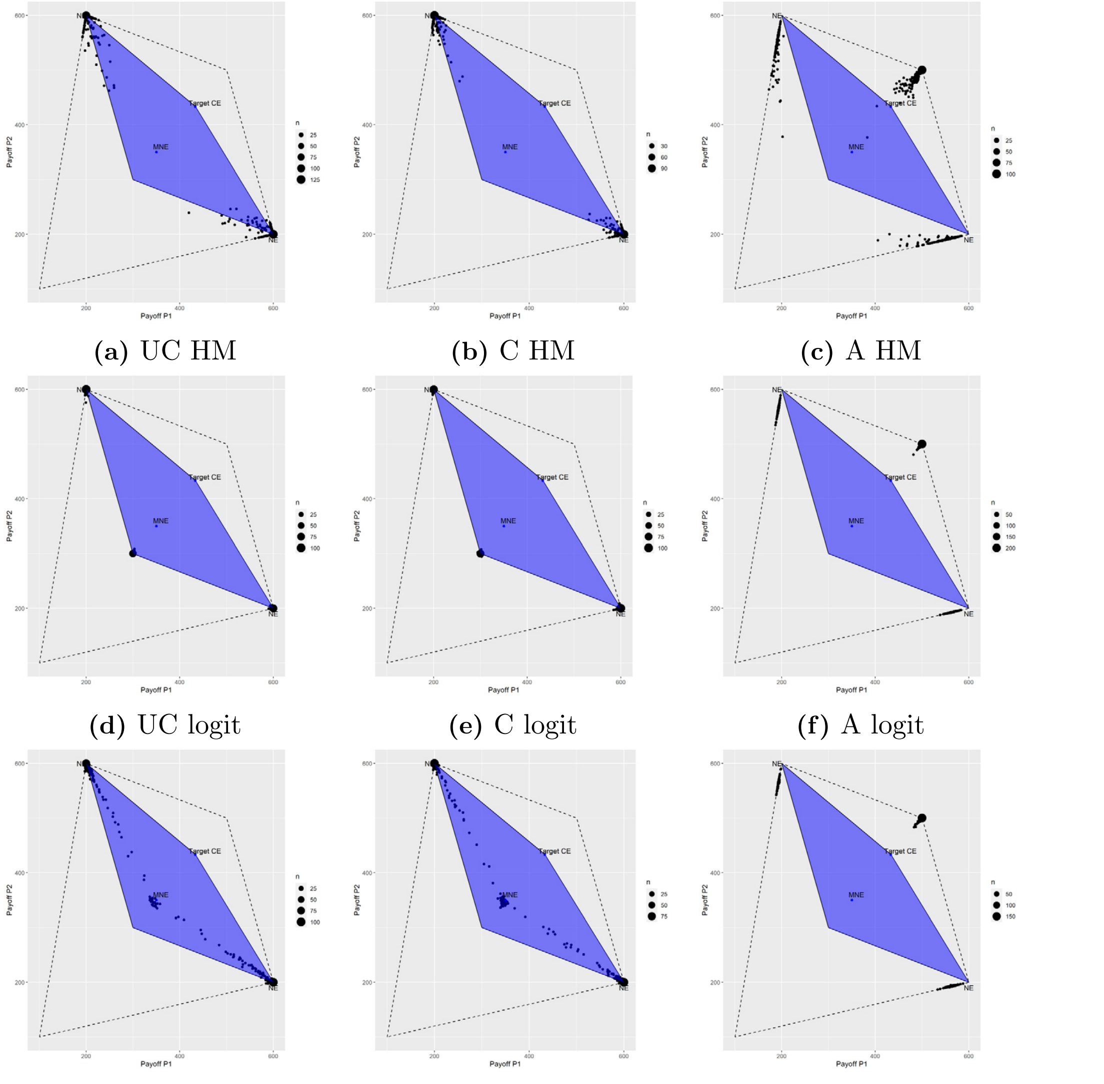
more stable under inertial logit. Under plain logit, a nontrivial fraction of trials appear to converge to the symmetric inefficient CE. However, as noted in Appendix D, that fraction shrinks when we look at the later periods of the simulations.

With Average regret, a large fraction of simulation trials indeed converge to the collusion profile, while smaller fractions converge to the two pure NEs. This confirms that the choice between A and C regret is consequential.

4.2. MV games

Table 4 reports the joint action distributions over all 500 trials in the MV game simulations. These population level data approximate the target CE fairly well for all nine combinations of regret varieties and response rules. On the other hand, Fig. 4 shows that player pairs in those simulations often have asymmetric payoffs, which is inconsistent with the target CE. Moreover, the time average payoff vector is often outside the CE region for A regret, while it usually near the efficient frontier of CE for UC and C regret.

To better understand dynamic patterns in the MV simulations, we calculate S2, which we define as the sum of the two largest profile fractions. The value of S2 is 0.22 at MNE and is 0.33 at the target CE, but approaches 1.0 when player pairs alternate between two profiles or when they never switch. Fig. 5 plots the cumulative distribution (CDF) across all trials, and shows that S2 is nearly uniformly distributed between 0.6 and 0.8 under logit or inertial logit dynamics using UC or C regret. The range shifts out to 0.7 - 0.95 with HM response. For all response rules using A regret, S2 is above 0.9. This suggests that, although correlated equilibrium might be a good prediction of what happens across a population of trials, it doesn't fully capture what happens



(a) UC HM **(b)** C HM **(c)** A HM

(d) UC logit **(e)** C logit **(f)** A logit

(g) UC inertial logit **(h)** C inertial logit **(i)** A inertial logit

Fig. 3. **Pair level payoff space data for CH game simulations.** Response rules HM, logit and inertial logit follow equations (5), (6) and (7), respectively. Regret varieties are unnormalized counterfactual (UC), counterfactual (C) and Average (A). Baseline parameters are $(\mu, \beta, \Delta) = (600, 1, 0.8)$.

over the span of 450 periods in a single trial. It seems that play in a single simulation trial is much less varied than in the target CE or MNE, and it often is confined to just one or two profiles.

We also ran a few much longer trials with 20,000 periods. For a typical long trial with HM response using C regret, Fig. 6 shows the action profiles (upper panel) and regret value of the better of the two alternative actions (lower panel). Dynamics are quite similar for long simulations using UC regret and other response rules with parameters near the baseline. The upper panel shows that from about period 2000 to 4500, Row (blue lines) plays M and Column (red

Table 4
Profile distributions in simulations of the MV game.

HM response, Eq. (5)

	UC			*C*			*A*		
	L	C	R	L	C	R	L	C	R
T	0.003	0.178	0.180	0.003	0.157	0.167	0.008	0.152	0.166
M	0.146	0.004	0.168	0.173	0.002	0.172	0.140	0.007	0.180
D	0.150	0.167	0.003	0.166	0.158	0.003	0.196	0.145	0.007

logit response, Eq. (6)

	UC			*C*			*A*		
	L	C	R	L	C	R	L	C	R
T	0.020	0.153	0.152	0.014	0.163	0.163	0.002	0.161	0.168
M	0.154	0.020	0.155	0.161	0.015	0.162	0.159	0.002	0.173
D	0.164	0.162	0.020	0.156	0.152	0.014	0.161	0.172	0.002

inertia logit response, Eq. (7)

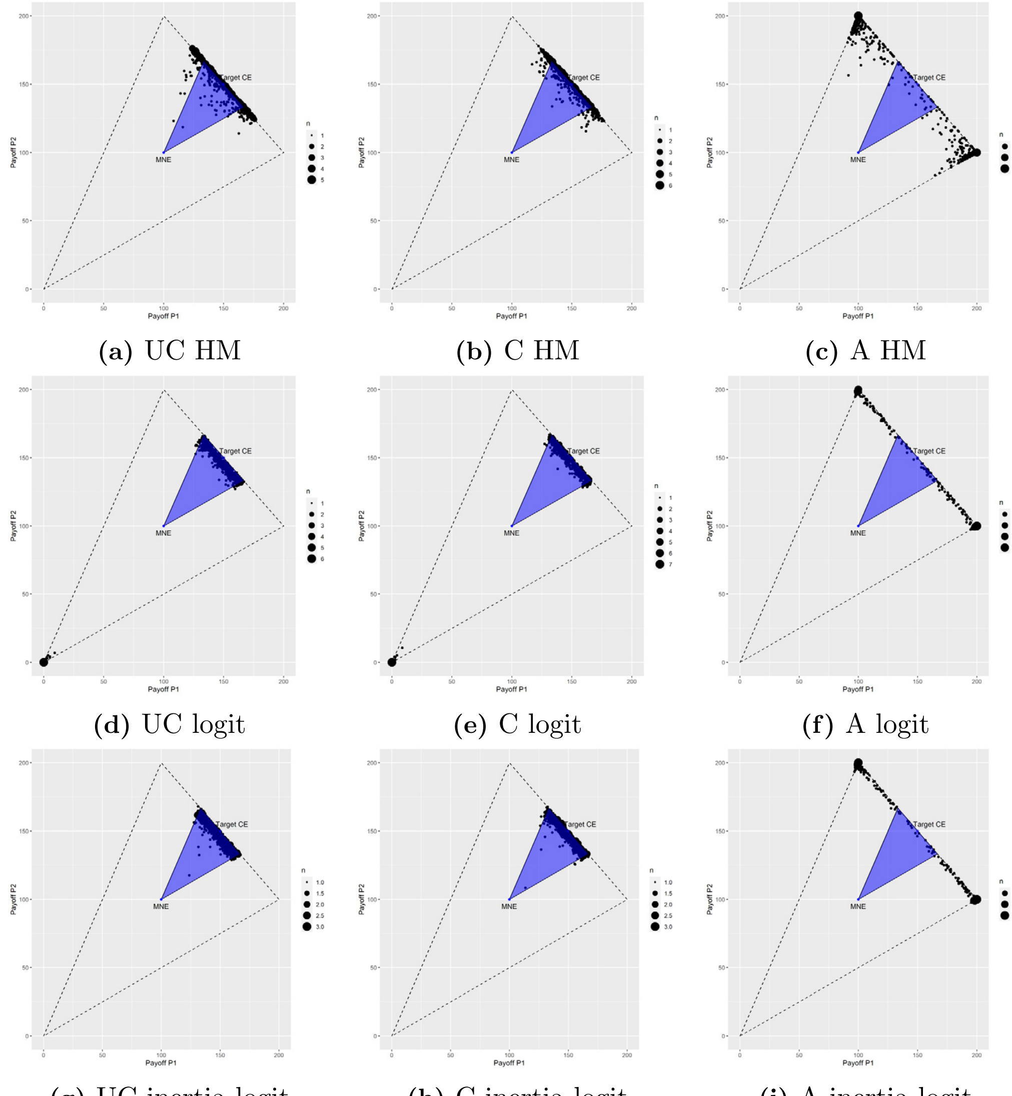
	UC			*C*			*A*		
	L	C	R	L	C	R	L	C	R
T	0.006	0.165	0.171	0.005	0.160	0.166	0.002	0.167	0.159
M	0.157	0.006	0.161	0.164	0.005	0.162	0.158	0.002	0.153
D	0.160	0.168	0.006	0.169	0.164	0.005	0.188	0.168	0.002

Notes: Regret varieties are unnormalized counterfactual (UC), counterfactual (C) and Average (A). Simulations use baseline parameters $(\mu, \beta, \Delta) = (600, 1, 0.8)$.

lines) plays R. According to the lower panel, during this time Column's regret remains near -100, while Row's regret gradually rises and becomes positive around period 4500. At that point, Row switches action several times but settles on T, her best response to R in the bimatrix stage game. Then her regret rapidly falls to -100, so she has no incentive to switch. However, over the next several thousand periods, Column's regret rises gradually and becomes positive around period 10000. After switching a few times, he settles on action C, his best response to T. Then Row's regret increases and by the end of the simulation is not far below zero. Given more time, she would evidently switch to B, her best response to C.

These long MV simulations suggest that with counterfactual regret, player pairs will trace out the entire six-profile best response cycle. However, the time between switches gets longer and longer because when there is a longer history, it takes more periods in the current profile to outweigh previous evidence. Thus it seems that, in the very long run, play in a single trial will converge to the CE set, and the target CE is the average of its limit points. However, convergence is glacial, and a simple time average of an initial segment will skew towards the more recent profiles visited.

Long trials using A regret, on the other hand, do not appear to exhibit best response cycles. As illustrated in Fig. 7, typical simulation trials eventually fail to produce positive regret, thus locking players into one of the off-diagonal profiles, and all six are equally likely a priori. That reconciles the aggregate results, which approximate the target CE, with the player pair payoff vectors, which are often so asymmetric that they lie outside the CE region. Appendix D provides more examples with logit response.



(a) UC HM (b) C HM (c) A HM

(d) UC logit (e) C logit (f) A logit

(g) UC inertia logit (h) C inertia logit (i) A inertia logit

Fig. 4. **Pair level payoff space data for MV game simulations.** Response rules HM, logit and inertia logit follow equations (5), (6) and (7), respectively. Regret varieties are unnormalized counterfactual (UC), counterfactual (C) and Average (A). Baseline parameters are $(\mu, \beta, \Delta) = (600, 1, 0.8)$.

5. Experimental design

Do the simulations and theory give insight into strategic interaction among human players? To find out, we ran a full factorial laboratory experiment with three treatment variables. The first is the game that subjects play — either standard Chicken (CH) as shown in Table 1 or Moulin and Vial (1978) (MV) as shown in Table 2. Recall that the NE sets and collusion possibilities differ sharply between these two bimatrix games.

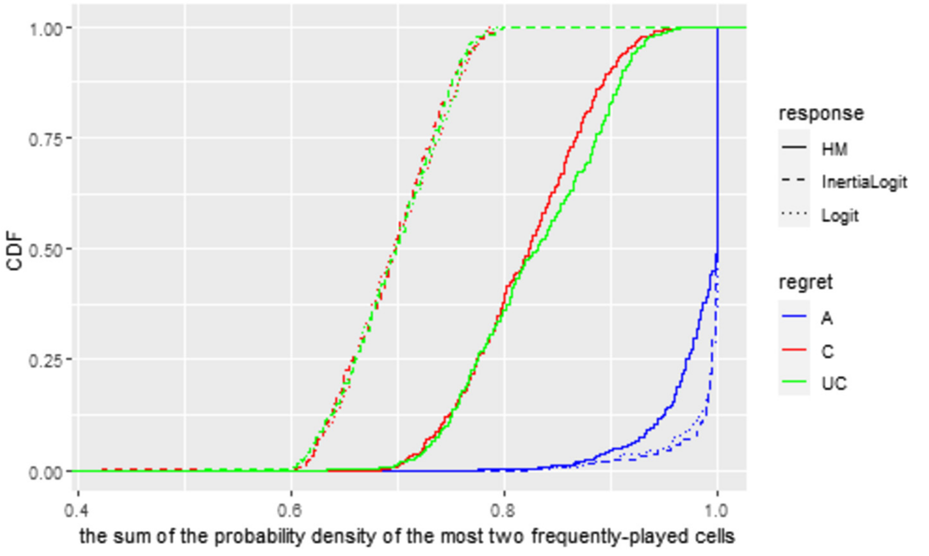


Fig. 5. CDFs for the sum of the top two profile frequencies.

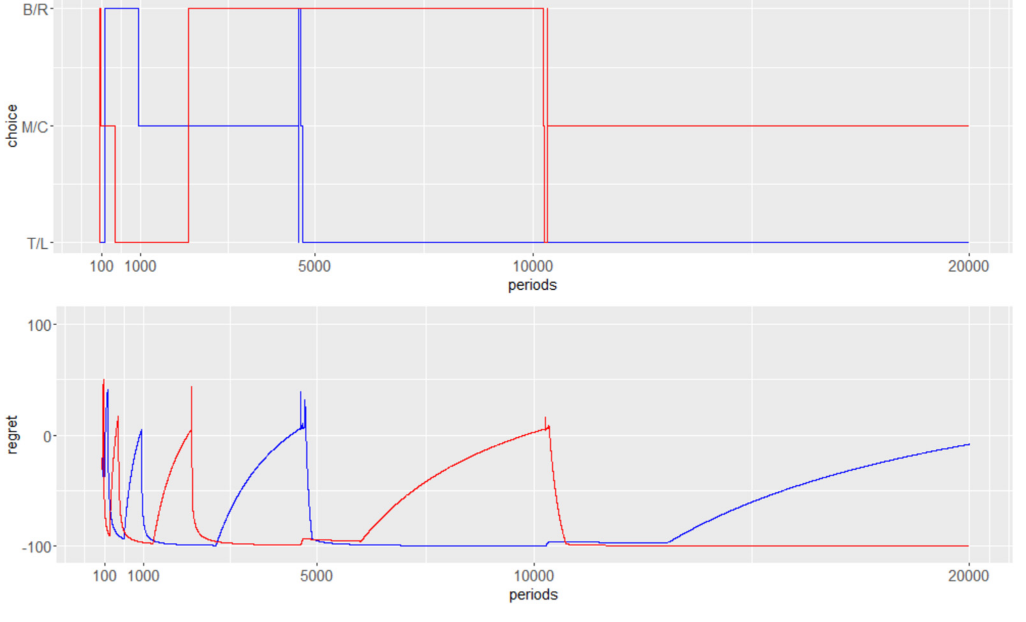


Fig. 6. **A long MV trial** with HM response and C regret. Blue (red) lines refer to row (column) players' action and regret.

The second treatment variable is the level of information provided by the user interface. In low information treatment (L), subjects see very little information besides what is required to verify the regret display. High information treatment (H) includes information on payoff functions and opponent actions, and thus brings us closer to laboratory environments used in most previous bi-

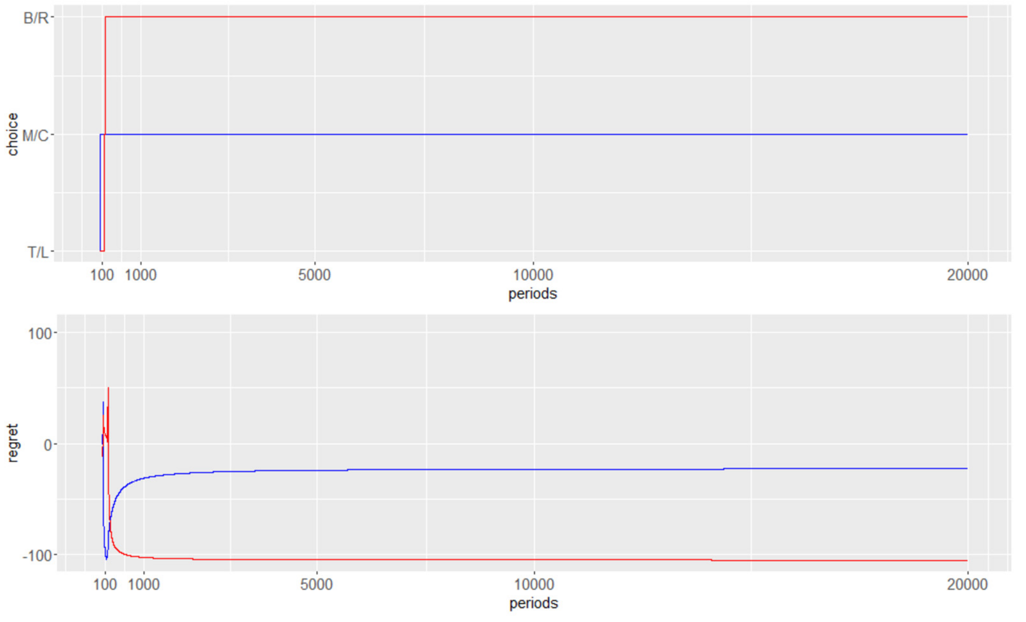


Fig. 7. **A long MV trial** with HM response and A regret. Blue (red) lines refer to row (column) players' action and regret.

Fig. 8. **User interface for low information (L) Chicken (CH) game.** The horizontal orange bars next to the radio buttons A and B (with text 380 and 350) show the current regret components. The upper graph shows the payoffs earned each period so far, while the lower graph shows the corresponding actions selected. The green bar in top right corner shows time remaining in the current period.

matrix game experiments. The point is to check the robustness of behavior to additional feedback beyond the regret information.

The third treatment variable is the type of regret subjects see in the user interface. The components of regret are displayed as orange bars next to each radio button, as shown in Fig. 8. In

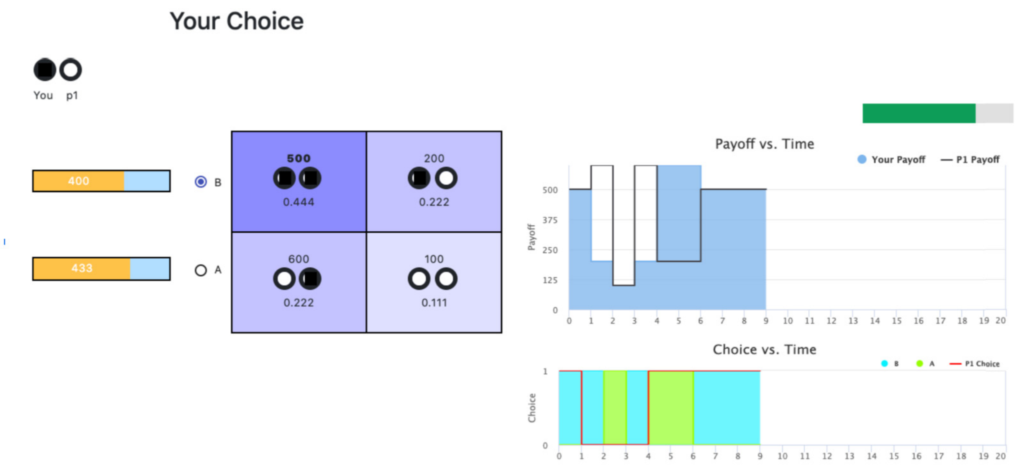


Fig. 9. **User interface for high information (H) Chicken (CH) game.** Horizontal bars next to radio buttons still represent current regret components (here 433 and 400). Each cell of the 2 × 2 matrix reports own payoff (top number), own previous choice (first filled or empty circle), counterparty's previous choice (second circle), and the frequency of past play (bottom number, and shading intensity). The upper graph shows own (blue area) and counterparty payoffs (black line) so far. The bottom graph displays own choices (color coded), and counterparty (binary coded red line) so far.

Counterfactual (C) treatments, these bars display the counterfactual per period payoffs $m_t^i(j,k)$ defined in equation (3). (We use normalized regret C because with unnormalized regret UC, the difference between orange bars would often shrink towards invisibility in later periods.) In Average (A) treatments, the orange bar for each action k shows its actual historical per-period payoff $M_t^i(k) = m_t^i(k,k)$.

Fig. 8 exemplifies the user interface for CH×L treatments. Besides the orange regret bars just described, the features include the green bar in the upper right corner. It shows the time remaining until the player's action (here, either A or B, entered by clicking a radio button) becomes final for that period. There are two time graphs which show current and past payoffs, and current and past actions, for each period so far.

Fig. 9 shows an analogous screen for the high information treatments CH×H. Beyond the information conveyed in L treatments, subjects in H treatments can observe (i) current and past actions, and payoffs, of the other player (P1 in the Figure), (ii) their own payoff matrix, and (iii) the frequency of past profiles. The latter is shown by the intensity of purple shading across profile cells — the more intense the purple color, the more frequently this cell has been played. Lastly, the filled circles indicate the player's own current choice and last period's choice by P1. User interfaces for the MV games are analogous.

5.1. Procedures

Treatment variables are held constant within each session and varied across sessions. Each session consists of 2 practice supergames with 20 periods each, and 8 salient supergames with 50 periods each.[5] In L treatments, each period within each supergame lasts 4 seconds. In H

[5] The simulations and the experiment operate at very different time scales because they have different objectives. The purpose of the long-trial simulations is not to mimic what human subjects might do, but rather to elucidate the relevant

Table 5
Session information.

Game	Information	Regret	#subjects	#sessions
CH	L	Average	44	4
CH	H	Average	22	2
MV	L	Average	50	4
MV	H	Average	24	2
CH	L	Counterfactual	44	4
CH	H	Counterfactual	26	2
MV	L	Counterfactual	42	4
MV	H	Counterfactual	22	2

treatments, periods are 8 seconds for the first 4 supergames, and then shorten to 6 seconds for the last 4 supergames. The additional time in H treatments allows subjects to absorb the greater amount of information in the display. Occasionally a player does not click any radio button during a period, in which case their action from the previous period is carried over. Each supergame starts with a random choice of actions, and we change the order of actions in the payoff matrices every other supergame to encourage subjects to remain attentive. The matching of players was random (fixed) between (within) supergames.

Table 5 summarizes the sessions by treatment. We completed 4 low information sessions and 2 high information sessions for each type of game and each type of regret. There were 10 to 14 participants in each session, with a total of 274 subjects in the 24 online sessions. Subjects were drawn from the MonLEE (Monash University) subject pool. The experimenter read the instructions aloud, and answered questions via private chat. Participants received all points earned in the 8 salient supergames, which were converted to Australian Dollars (AUD) at the rate of 0.63 per 100 points in CH sessions, and 1.89 per 100 points in MV sessions. On average, subjects received 19 AUD (treatment averages ranged from 17.25 to 19.92) on top of the 5 AUD show-up fee. All payments were made via bank transfer. Sessions in low information treatments lasted less than one hour, and lasted up to 1.5 hours in high information treatments.

5.2. Hypotheses

Hypothesis 1. *(a) In later periods, the overall observed (row, column) joint distribution of actions will be well approximated by a CE distribution. (b) That approximation will be better under counterfactual regret than under average regret.*

Hypothesis 1 focuses on aggregate time average profile frequencies. H1a will be rejected if players' action profiles systematically deviate from the CE distributions. As for H1b, better convergence to CE with counterfactual than with average regret is suggested by the theoretical results of Hart and Mas-Colell (2000) and Hart and Mas-Colell (2001). It is also strongly suggested by our CH simulation results: Table 3 shows aggregate frequencies consistent with the pure Nash equilibria under C regret, but under A regret the frequencies are much closer to the collusive outcome, which is not part of the CE set. On the other hand, the MV simulations reported in Table 4 show no major differences in the aggregate behavior for different types of regret; they in-

adaptive models and better understand their implications. The purpose of many laboratory experiments, including ours, is to test the predictive power of relevant models on time scales that work for human subjects.

dicate profiles closer to the target CE than to the MNE, as simulated players avoid the inefficient diagonal profiles.

Theory and simulations have little to say about the role of information, but we conjecture that High information treatments in the CH game will make regret less salient and may lead to aggregate behavior that is closer to the collusive profile.

Hypothesis 2. *(a) Time average payoff vectors for player pairs will disproportionately lie in or near the CE region. (b) The fraction of player pairs with payoff vectors in the CE region will be higher under counterfactual regret than under average regret.*

Hypothesis 2 focuses on long-run behavior of individual player pairs. The theoretical results of Hart and Mas-Colell and others predict that player pairs' time average play will converge to the CE set, but not necessarily to any particular point in the CE set, and H2a is worded accordingly.[6] On the other hand, convergence to a Pareto efficient and "fair" (equal payoffs for the two players) CE would be especially interesting in our symmetric games. Therefore, we will also look for convergence to the target CE payoff vector. H2b is a disaggregated version of H1b that applies to player pair payoffs.

The CH game simulations in Fig. 3 suggest that most pairs will converge towards one of the two pure Nash equilibria under C regret, and will collude more often under A regret. In the MV game, the simulations in Fig. 4 suggest that time average payoff vectors will be closer to the target CE under C regret than under A regret. Again we conjecture that low information will make counterfactual regret especially salient and so will facilitate convergence to the CE region.

Hypothesis 3. *Subjects are more likely switch to an alternative action when its positive regret is larger.*

Hypothesis 3 focuses on individual players' period-by-period behavior. The hypothesis is the basic premise of all adaptive dynamics models we consider. Logit models are the natural specification for testing Hypothesis 3, but we also use linear probability model as a robustness check. Lastly, we will also test for the response asymmetry between positive and negative regret assumed in the HM response rule.

6. Results

Except when otherwise noted, the data reported below include only the last 60% of observations (viz., the last 30 of 50 periods) from each supergame. We exclude early periods because regret is initially very noisy, and because hypotheses H1 and H2 focus on long-run behavior. As noted in Appendix E, results are generally similar with all observations and with the last 20 periods but noisier when all observations are included.

6.1. Aggregate behavior

Table 6 presents aggregate time average profile frequencies by treatment across CH games. The left panel summarizes the frequency of play in low information treatments, for Counterfactual (C) as well as for Average (A) regret, while the right panel does the same for high information

[6] Any sort of weighted average of limit points in the CE set will also lie in the CE set, since that set is convex.

Table 6
Time average frequency in CH games.

	Low information				High information			
	C		**A**		**C**		**A**	
	L	R	L	R	L	R	L	R
U	0.07	0.40	0.12	0.22	0.17	0.12	0.23	0.18
D	0.45	0.07	0.27	0.39	0.14	0.58	0.16	0.44

Table 7
Time average frequency in MV games.

	Low information						High information					
	C			**A**			**C**			**A**		
	L	C	R	L	C	R	L	C	R	L	C	R
T	0.04	0.14	0.12	0.07	0.14	0.13	0.04	0.13	0.10	0.06	0.15	0.12
M	0.13	0.05	0.12	0.13	0.07	0.13	0.15	0.06	0.17	0.11	0.08	0.13
D	0.18	0.17	0.06	0.12	0.15	0.07	0.15	0.15	0.06	0.14	0.12	0.08

treatments. Recall that the target CE puts a frequency of zero in the (U,L) cell and 0.33 in the other cells. The table shows that this target CE is best approximated with A regret and Low information: the lowest frequency of play here is indeed the (U,L) cell (0.12), while the collusion cell (D,R) has frequency 0.39, modestly higher than the target 0.33, but a bit closer to it than in any of the simulations. With C regret, subjects mostly play the two pure NE cells (U,R) and (D,L) with similar frequencies (0.40 and 0.45), and put little weight on the main diagonal (0.07 each on (U,L) and (D,R)). This CE distribution is very similar to those in C and UC simulations. In the high information treatment, collusive behavior dominates for both A (with frequency 0.44) and C (0.58) regret, which is not consistent with CE.

Table 7 presents aggregate frequencies for MV sessions; all treatments yield similar outcomes. Roughly consistent with the target CE for this game, and inconsistent with the unique NE, every diagonal cell has lower frequency than any of the six off-diagonal cells. This is especially true in the C treatments.

Result 1. *Consistent with Hypothesis 1a, aggregate time average action profiles in later periods of MV games are well approximated by a CE distribution. Consistent with H1b, the approximation is better under counterfactual regret than under average regret. In CH games with low information, aggregate profiles resemble the target CE with A regret, and resemble a roughly equal combination of the two pure NE with C regret.*

Appendix B offers a more quantitative test confirming Hypotheses 1a and 1b, and further investigates the correlation between players' actions. The test is related to the likelihood ratio test of independence featured in Moreno and Wooders (1998) but is less direct due to the lack of independent observations within each player pair, and across pairs in a session.

6.2. Player pair average behavior

Fig. 10 shows pair-level results relevant to Hypothesis 2 from Chicken (CH) games. In L×C sessions, most pairs generate average payoff vectors at or close to one of the two pure NE. In the

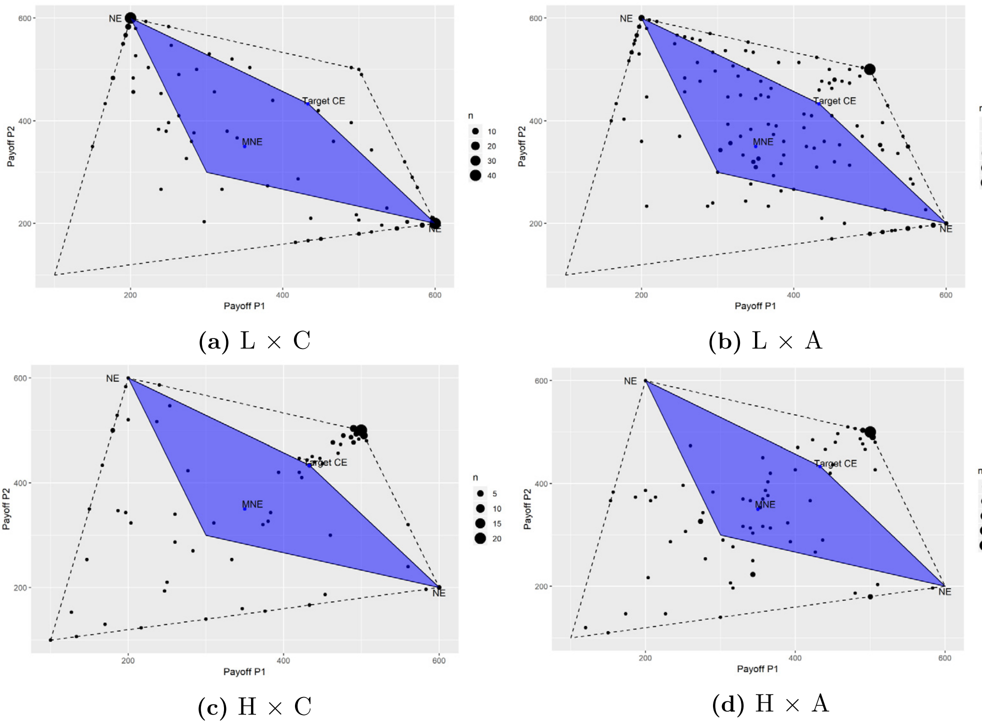


Fig. 10. **Payoffs for player pairs in CH games.** The four panels separate data by combinations of L vs H information and C vs A regret. The CE region is shaded blue.

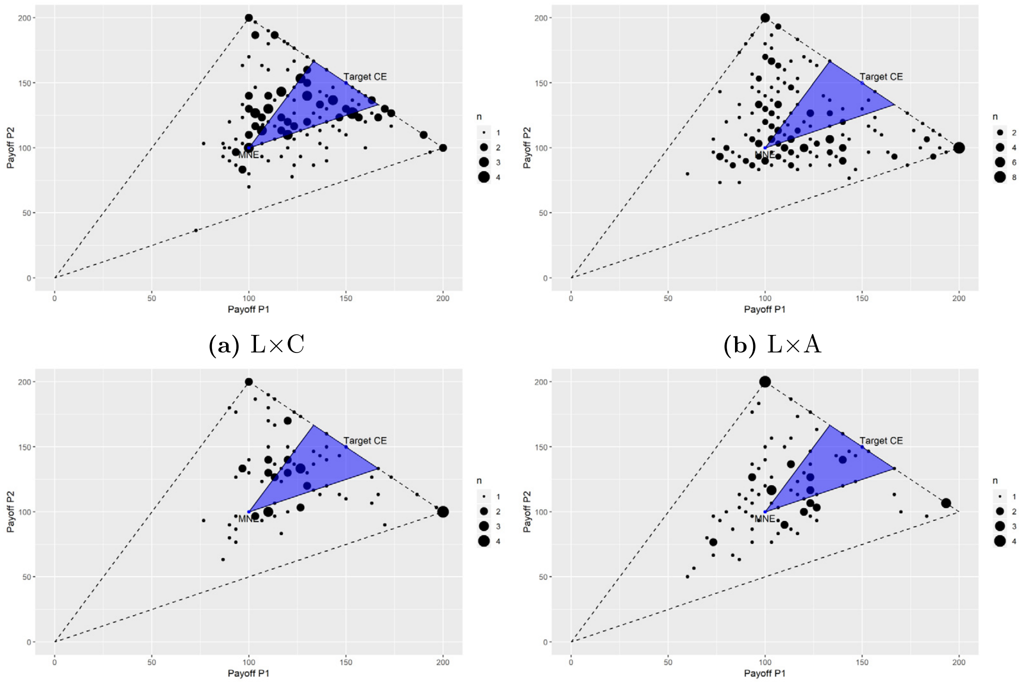
other treatments, many pairs collude to obtain payoffs near (500, 500) while payoff vectors for other pairs are widely scattered. These figures are broadly consistent with the simulation data.

Fig. 10 also suggests that high information destabilizes CE (more specifically, NE) in our counterfactual regret treatment for the CH game. With low information, most pairs converge to one of the two pure NE, but with high information most pairs seem to seize the opportunity to collude on a fairer and more efficient outcome.

Fig. 11 similarly summarizes the pair payoffs for the four MV treatments. The time average payoffs are mostly in or near the CE zone, especially in L×C sessions. Payoff vectors tend to be closer to the target CE (and are more likely to Pareto dominate the unique NE) in Counterfactual than in corresponding Average regret treatments.

Result 2. *Consistent with Hypothesis 2a, player pairs in CH games with Low information are disproportionately likely to earn average payoffs near or in the CE region. The same is true for all treatments in MV games. Consistent with Hypothesis 2b, for each game the fractions of pair payoffs in the CE region are largest in L×C treatments.*

Table 8 offers a more quantitative test of Hypothesis 2. The first column presents for each treatment the observed fractions of pair payoffs that are in the CE region. We perform a proportional test with null hypothesis that the observed fraction of pairs is no greater than the uniform random success rate of 0.27 for the CH game and 0.11 for the MV game; recall from Section 3.1 that these rates represent the fractions of feasible payoff space that lie in the CE region. For the



(a) L×C (b) L×A

(c) H×C (d) H×A

Fig. 11. **Payoffs for player pairs in MV games.** The four panels separate data by combinations of L vs H information and C vs A regret. The CE region is shaded blue.

Table 8
Fraction of pairs in CE region; distance to target CE.

	Fraction in CE	Distance Target CE
CH game	0.27	
L×C	0.57 (0.002)	253.3
L×A	0.36 (0.005)	157.1
H×C	0.17 (0.989)	150.1
H×A	0.28 (0.438)	157.0
MV game	0.11	
L×C	0.43 (0.000)	45.4
L×A	0.20 (0.000)	58.5
H×C	0.35 (0.000)	48.4
H×A	0.29 (0.000)	58.9

Notes: The entries 0.27 for CH and 0.11 for MV are the fractions of feasible payoff space occupied by CE payoffs; other entries in that column are fractions of observed pair payoffs in the CE zone. (P-values are shown in parentheses for proportional tests with errors clustered at the session level, for the null hypothesis that the observed fraction is no greater than 0.27[0.11] in CH[MV] games.) Last column reports mean Euclidean distance of observed payoff vector from target CE payoff vector.

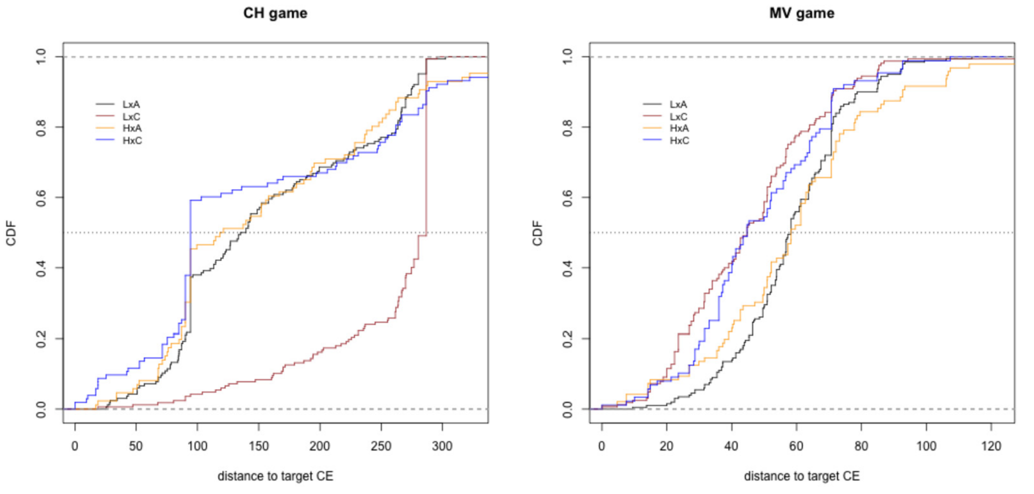


Fig. 12. CDF of the distance to target CE (pair data).

high information treatments in the CH game, we fail to reject that the observed fraction is equal or lower to a random draw. However, as claimed in Result 2, we strongly reject that null hypothesis in both low information treatments. In particular, consistent with Hypothesis 2b, the observed fraction 0.57 of CE-consistent play in the CH×L×C treatment is far larger than the 0.27 random benchmark. Of course, from Fig. 10 we know that much of that CE play arises from the two pure NE.

For all MV game treatments, we strongly reject that the observed fraction is no greater that the random proportion 0.11; all p-values are less than 0.001. Again, the largest fraction of payoff pairs in the CE region (0.43) is for counterfactual regret and low information. On the other hand, under Average regret the fraction of pairs in the CE region is larger in the high information (0.29) than in the low information environment (0.20).[7]

Result 2*. *In the CH game, the distance from the target CE payoff vector is driven mainly by collusion in high information treatments, and by pure NE play in the L×C treatment. In the MV game, the distance is smaller under counterfactual regret, with no significant differences between high and low information environments.*

This result is foreshadowed by the mean distances reported in last column of Table 8, but it relies mainly on the finer grained analysis displayed in Fig. 12. We calculate the Euclidean distance from the target CE payoff for each player pair, and build the cumulative distribution functions (CDFs) shown in Fig. 12. The left panel shows that in the CH game, a large fraction of pairs in the L×C treatment has distance of 287, which is the distance from either pure NE payoff vector to the target CE. In high information treatments, play is closer to the target CE mainly due to the mass of play near the collusion profile, which has distance of 95 from the target CE.

[7] Table E.5 in Appendix E shows the fraction of pair payoff vectors in the MV game that lie in the disk centered on MNE, using the same area as for the CE region. We confirm that for C regret (but not for A regret) the CE region has more pairs than the disk centered on MNE.

Table 9
Maximum Likelihood Estimation.

Treatment	Obs	logit	inertial logit		inertial truncated logit		
		β	Δ	β	Δ	β_1	β_2
CH game							
L×C	10,020	2.70	0.45	2.95	0.61	0.04	3.00
		(0.000)	(0.000)	(0.000)	(0.000)	(0.496)	(0.000)
L×A	9,870	0.77	0.48	0.44	0.81	0.06	0.88
		(0.000)	(0.000)	(0.000)	(0.000)	(0.740)	(0.000)
H×C	6,104	−1.70	0.28	−1.71	0.59	−2.97	4.77
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
H×A	5,116	0.60	0.57	0.65	0.76	−0.29	1.18
		(0.000)	(0.000)	(0.000)	(0.000)	(0.036)	(0.000)
MV game							
L×C	9,737	1.85	0.98	1.91	1.00	1.59	0.61
		(0.000)	(0.000)	(0.000)	(0.076)	(0.000)	(0.000)
L×A	11,896	0.93	0.97	0.88	0.97	0.75	0.04
		(0.000)	(0.000)	(0.000)	(0.602)	(0.000)	(0.000)
H×C	5,228	0.96	0.97	1.21	1.02	0.82	0.78
		(0.000)	(0.000)	(0.000)	(0.254)	(0.000)	(0.000)
H×A	5,760	1.09	0.96	1.13	0.98	1.03	0.19
		(0.000)	(0.000)	(0.000)	(0.202)	(0.000)	(0.231)

Notes: Logit specification refers to equation (6); inertia logit to equation (7); and inertia truncated logit uses a multiplicative dummy when the regret is negative. In parentheses are p-values given error clustering at the subject level for the null hypotheses that βs are zero and that $\Delta = 1$.

The right panel of Fig. 12 shows that distances from target CE in the MV game are generally greater (almost first order stochastic domination) under average than under counterfactual regret. Table E.6 in Appendix E confirms this conclusion, which reports linear regressions of distance to target CE on treatment dummies. Session level clustered error p-values are 0.018 and 0.001, confirming greater distance under A than C regret, for low and high information treatments, respectively. Appendix C offers further investigation into the MV game dynamics.

6.3. Individual adjustment behavior

Result 3. *Consistent with Hypothesis 3 in all treatments, except CH×H×C where collusive behavior abounds, individual players are more likely to switch to an alternative action the larger its positive regret. There is significant heterogeneity in the response to negative regret. While most subjects respond strongly to negative regret, many subjects are less responsive.*

This result is supported by the regressions reported in Table 9. Under *logit* specification, we estimate equation (6) for all treatments, and find that the value of β is strongly positive (p-value < 0.001) in all cases except the CH game under treatment H×C, likely due to collusive behavior. In specification *inertial logit*, we estimate equation (7), which includes an additional parameter Δ; inertia is present to the degree that its coefficient is less than 1 (but positive). We see very little inertia in the MV games, but see a substantial degree of inertia in CH games. Nevertheless, even in CH games, the β estimates (and significance levels) are not greatly altered by the presence of the additional parameter.

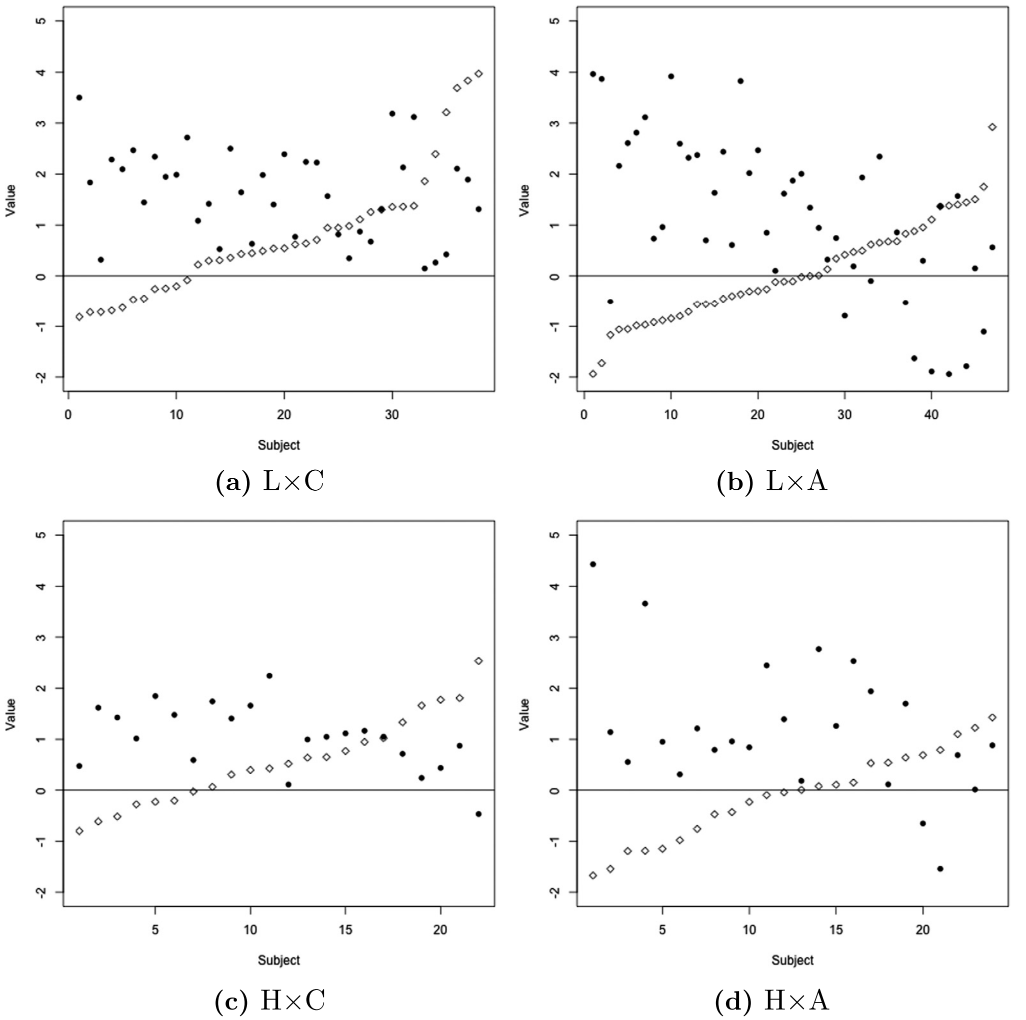


Fig. 13. **Inertial truncated logit betas for MV game (ML estimation)** The four panels separate data by combinations of L vs H information and C vs A regret. β_1 is the solid dot and β_2 the diamond dot. The data is at the subject level, and the plots sort the subjects according to the value estimated of β_2.

For the third specification, *inertial truncated logit*, we replace the expression $\beta r_t^i(j,k)$ in equation (7) by $\beta_1 r_t^i(j,k) + \beta_2 r_t^i(j,k) D[r_t^i(j,k) < 0]$, where the dummy variable D is 1 when regret is negative and otherwise is 0. Thus the response to negative regret is captured by the coefficient sum $\beta_1 + \beta_2$. For all treatments we find a positive or insignificant value of β_2, indicating that overall subjects do respond to negative regret. Further analysis at the subject level in Fig. 13 highlights significant heterogeneity for MV games. While very few subjects have a negative value of β_1, about one quarter to one half of subjects in each treatment have negative β_2. We do not present the individual betas for the CH games due to significant inertia; here the low Δ's (as reported in Table 9) make even the signs of most individual estimated betas ambiguous.

We also estimate μ in HM response following equation (5) using a linear probability regression. See Table E.7 in Appendix E for details. The results support that subjects strongly respond to regret in most treatments; they respond less strongly in the H and A treatments of the CH game due to the high collusion rate.

7. Conclusion

In this paper, we investigate empirically the theoretical possibility, implicit in Foster and Vohra (1998) and Hart and Mas-Colell (2000), that shared experience might enable players to achieve CE even in the absence of exogenous coordinating signals. Our results can be summarized briefly. Correlated equilibrium (CE) indeed has empirical relevance, but within narrower limits than some researchers might have guessed.

- CE does a decent job of predicting time-average behavior in all treatments of the MV game, especially when our human subjects see the more conducive form of regret (counterfactual).
- In the Chicken game featured in most textbook discussions of CE, the predictions of aggregate behavior are less useful. The CE prediction is good in the most conducive treatment (counterfactual regret, low information) but so is the more specific prediction of pure Nash equilibrium. In the less conducive treatments, we see more collusive behavior than allowed in CE.
- These limitations are foreshadowed in our simulations of regret-based adaptive dynamics.
- Fitted models of such dynamics indicate that our human subjects indeed respond systematically to positive regret.
- However, subjects respond quite heterogeneously to negative regret, and not very many of them appear to ignore the magnitude of negative regret.

We do not regard our experiment as definitive, but rather as opening new ways to investigate strategic interaction. The Low (and High, i.e., moderate) information user interfaces and other novel design features can be adapted to study other games with interesting CE, including trimatrix games with some players having two and others having three possible pure actions. Our payoff space analysis makes the number of dimensions manageable, while our simulation approach may provide ways to sharpen results on convergence, e.g., when adaptive dynamics will converge to an interesting non-Nash CE. Simulations can also help pre-explore vast sets of possible games and treatments, and identify especially interesting candidates. Theorists might want to explore adaptive models in the tradition of Hart and Mas-Colell that relax the assumption that players ignore the magnitude of negative regret, and consider allowing greater heterogeneity across players. There is still much to learn about the relevance of correlated equilibrium.

Acknowledgments

We gratefully acknowledge funding for this project from Early Career Grants, Monash Business School. For very helpful comments and feedback we are indebted to Journal Editor Tilman Börgers, two anonymous referees, Aleksandr Alekseev, Tim Cason, John Duffy, Nick Feltovich, Misha Freer, Evan Friedman, Sergiu Hart, Heinrich Nax, Jan Potters, Vjollca Sadiraj, Tridib Sharma, Rakesh Vohra, and participants in numerous seminars and workshops. This paper is dedicated to the memory of Bill Sandholm, whose encouragement and suggestions in 2017 eventually helped us develop the experiment reported herein.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jet.2022.105531.

References

Anbarcı, Nejat, Feltovich, Nick, Gürdal, Mehmet Y., 2018. Payoff inequity reduces the effectiveness of correlated-equilibrium recommendations. Eur. Econ. Rev. 108, 172–190.

Arifovic, Jasmina, Boitnott, Joshua F., Duffy, John, 2019. Learning correlated equilibria: an evolutionary approach. J. Econ. Behav. Organ. 157, 171–190.

Aumann, Robert, Brandenburger, Adam, 1995. Epistemic conditions for Nash equilibrium. Econometrica, 1161–1180.

Aumann, Robert J., 1974. Subjectivity and correlation in randomized strategies. J. Math. Econ. 1 (1), 67–96.

Aumann, Robert J., 1987. Correlated equilibrium as an expression of Bayesian rationality. Econometrica, 1–18.

Bone, John, Droiuvelis, Michalis, Ray, Indrajit, 2013. Coordination in 2x2 games by following recommendations from correlated equilibria. working paper.

Brandenburger, Adam, Dekel, Eddie, 1987. Rationalizability and correlated equilibria. Econometrica, 1391–1402.

Cason, Timothy N., Sharma, Tridib, 2007. Recommended play and correlated equilibria: an experimental study. Econ. Theory 33 (1), 11–27.

Cason, Timothy N., Sharma, Tridib, Vadovič, Radovan, 2020. Correlated beliefs: predicting outcomes in 2×2 games. Games Econ. Behav. 122, 256–276.

Duffy, John, Feltovich, Nick, 2010. Correlated equilibria, good and bad: an experimental study. Int. Econ. Rev. 51 (3), 701–721.

Duffy, John, Lai, Ernest K., Lim, Wooyoung, 2017. Coordination via correlation: an experimental study. Econ. Theory 64 (2), 265–304.

Forges, Francoise, Peck, James, 1995. Correlated equilibrium and sunspot equilibrium. Econ. Theory 5 (1), 33–50.

Foster, Dean, Young, Hobart Peyton, 2006. Regret testing: learning to play Nash equilibrium without knowing you have an opponent. Theor. Econ. 1 (3), 341–367.

Foster, Dean P., Vohra, Rakesh V., 1997. Calibrated learning and correlated equilibrium. Games Econ. Behav. 21 (1–2), 40.

Foster, Dean P., Vohra, Rakesh V., 1998. Asymptotic calibration. Biometrika 85, 379–390.

Fudenberg, Drew, Levine, David K., 1999. Conditional universal consistency. Games Econ. Behav. 29 (1–2), 104–130.

Georgalos, Konstantinos, Ray, Indrajit, SenGupta, Sonali, 2020. Nash versus coarse correlation. Exp. Econ., 1–27.

Hart, Sergiu, Mas-Colell, Andreu, 2000. A simple adaptive procedure leading to correlated equilibrium. Econometrica 68 (5), 1127–1150.

Hart, Sergiu, Mas-Colell, Andreu, 2001. A reinforcement procedure leading to correlated equilibrium. In: Economics Essays. Springer, pp. 181–200.

Metzger, Lars P., 2018. Evolution and correlated equilibrium. J. Evol. Econ. 28 (2), 333–346.

Moreno, Diego, Wooders, John, 1998. An experimental study of communication and coordination in noncooperative games. Games Econ. Behav. 24 (1–2), 47–76.

Moulin, Hervé, Vial, J-P., 1978. Strategically zero-sum games: the class of games whose completely mixed equilibria cannot be improved upon. Int. J. Game Theory 7 (3–4), 201–221.

Nax, Heinrich H., Burton-Chellew, Maxwell N., West, Stuart A., Young, H. Peyton, 2016. Learning in a black box. J. Econ. Behav. Organ. 127, 1–15.

Palfrey, Thomas R., Pogorelskiy, Kirill, 2019. Communication among voters benefits the majority party. Econ. J. 129 (618), 961–990.

Pradelski, Bary S.R., Young, H. Peyton, 2012. Learning efficient Nash equilibria in distributed systems. Games Econ. Behav. 75 (2), 882–897.

Young, H. Peyton, 2009. Learning by trial and error. Games Econ. Behav. 65 (2), 626–643.

Zhao, Shuchen, 2021. Taking turns in continuous time. J. Econ. Behav. Organ. 191, 257–279.