Feature-Weighted Categorized Play across Symmetric Games

Appendix and Supplementary Material

A Appendix

A.1 Learning algorithms

We implemented the following parameter-free variants for the three major learning algorithms: 1) fictitious play (FP); 2) reinforcement learning (RL); and 3) experience weighted attraction (EWA).

FP: we set the probability that i plays H to $P_i(H) = 0.5$ until both actions H and L have been played at least once by the opponents; after that, we apply the standard algorithm. Let n_H and n_L be the number of times that $j = 3 - i$ has played H or L in the past for a game G in category K. We let $n_H = n_L = 0$ as initial condition. While $n_H + n_L = 0$, we set $P_i(H) = 0.5$. After the first round for a player i, we set i's current belief that j will play H to $\beta = \frac{n_H}{n_H+n_L}$ and compute i's expected utility U_i^G of playing H or L in game G. Maximizing expected utility, i plays H in game G with probability

$$P_i(H) = \begin{cases} 1 & \text{if } U_i^G(H) > U_i^G(L) \\ 0.5 & \text{if } U_i^G(H) = U_i^G(L) \\ 0 & \text{if } U_i^G(H) < U_i^G(L) \end{cases}$$

and L with the complementary probability.

RL: we set the probability that i plays H to $P_i(H) = 0.5$ until both actions H and L achieve a positive payoff at least once; after that, we apply the standard algorithm. Let α_H and α_L be i's attractions to H or L for a game G in category K. We let $\alpha_H = \alpha_L = 0$ as initial condition. While $\alpha_H \cdot \alpha_L = 0$, we set $P_i(H) = 0.5$. Once $\alpha_H \cdot \alpha_L > 0$, we apply reinforcement: if i plays an action s_i (H or L) and j plays s_j, we increase i's attraction for s_i to $\alpha'_{s_i} = \alpha_{s_i} + u_i(s_i, s_j)$. Given i's attractions, i plays H in game G with probability $P_i(H) = \frac{\alpha_H}{\alpha_H+\alpha_L}$ and L with the complementary probability.

EWA: sets $P_i(H) = 0.5$ until both H and L achieve a positive payoff at least once; after that, we apply the implementation in Chmura et al. (2012). Let α_H and α_L be i's attractions to H or L for a game G in category K. We let $\alpha_H = \alpha_L = 0$ as initial condition. Until $\alpha_H \cdot \alpha_L = 0$, we set $P_i(H) = 0.5$. Once $\alpha_H \cdot \alpha_L > 0$, we update attractions as follows: if i plays an action s_i (H or L) and j plays s_j, we change i's attraction for each choice s in $\{H, L\}$ to

$$A'_s = \frac{\phi N A_s + u_i(s_i, s_j)\,[\delta + (1-\delta)\mathbf{1}(s_i, s)]}{\phi N + 1},$$

where $\mathbf{1}(s, s')$ is the indicator function for $s = s'$ and the functions N, ϕ, δ are defined in (Chmura et al., 2012, p. 48) respectively as: N is the *experience weight* from Equation (8); ϕ is the *change-detector* from Equation (9); and δ is the *attention function* from Equation (13), with $W = 2$. Given i's attractions, i plays H in game G with probability $P_i(H) = \dfrac{A_H}{A_H + A_L}$ and L with the complementary probability.

A.2 Fitness measures

The tests for the five propositions A_1–A_5 are based on the following fitness measures.

A$_1$: we set $v = 1 - Q_2$, using available data for the 12 PD games in Table SM.2. The values for Q_2 appear in line 4 of Table 1). The (best) benchmark is NE. Then the (relative) rate of improvement for an algorithm α is

$$\frac{v_\alpha - v_0}{v_1 - v_0} = 1 - \frac{Q_2(\alpha)}{Q_2(\mathrm{NE})}$$

A$_2$: we set $v = 1 - Q_2$, using available data for the 12 SH games in Table SM.2. The values for Q_2 appear in line 5 of Table 1). The (best) benchmark is CF. Then the rate of improvement for an algorithm α is

$$\frac{v_\alpha - v_0}{v_1 - v_0} = 1 - \frac{Q_2(\alpha)}{Q_2(\mathrm{CF})}$$

A$_3$: we set $v = r$, where r is the Pearson correlation coefficient between the 23 round-by-round differences Δ_i in H-rates for the simulated data (averaged over all runs) generated by an algorithm α and the experimental data of sequence SQ_1 from Table SM.4. The (best) benchmark yields $r = 0$. Then the rate of improvement for an algorithm α is

$$\frac{v_\alpha - v_0}{v_1 - v_0} = r(\Delta_i(\alpha), \Delta_i(\mathrm{SQ}_1))$$

A$_4$: we set $v = Q_4$, where Q_4 is the mean square distance between simulated data and experimental data for the sequence SQ_2 in Table SM.4. Specifically,

$$Q_4 = \frac{1}{nm} \sum_{k=1}^{n} \sum_{j=1}^{m} \Big(\gamma(E_k) - \gamma(S_j)\Big)^2,$$

where n is the number of experiments, m is the number of simulation runs for the algorithm, E_k is the k-th experiment, S_j is the j-th simulation, and γ is the ratio $\frac{H_f(x<0.5) - H_i(x<0.5)}{H_f(x>0.5) - H_i(x>0.5)}$, where H_f and H_i are respectively the average H-rates over the final and initial five games with $x < 0.5$ or $x > 0.5$. The (best) benchmark is CF. Then the rate of improvement for an algorithm α is

$$\frac{v_\alpha - v_0}{v_1 - v_0} = 1 - \frac{Q_4(\alpha)}{Q_4(CF)}$$

$\mathbf{A}_5$: we set $v = 1 - Q_2$, using available data from Duffy and Fehr (2018) for the round-by-round H-rates in the four sequences SQ_3 to SQ_6 in Table SM.4. The (best) benchmark is CF. Then the rate of improvement for an algorithm α is

$$\frac{v_\alpha - v_0}{v_2 - v_0} = 1 - \frac{Q_2(\alpha)}{Q_2(\text{CF})}$$

A.3 Robustness checks

A.3.1 Learning rate

The update rule (2) modifies the current strengths (f, g, h) by a self-adjusting learning rate $\gamma = 1 - \max(P_i(H), P_i(L))$. This implies that the learning rate slows down as the propensity to play either H or L approaches 1. Our first robustness test checks the effect of using a fixed learning rate for FWC.

For each value of γ between 0.05 and 0.95 (with tick size 0.05), we compute the mean square deviation Q_2 for the round-by-round H-rates over the same dataset used to produce Figure 6. This smaller dataset assembles the 12 PD games and the 12 SH games for which round-by-round H-rates are available, and is used throughout this section.

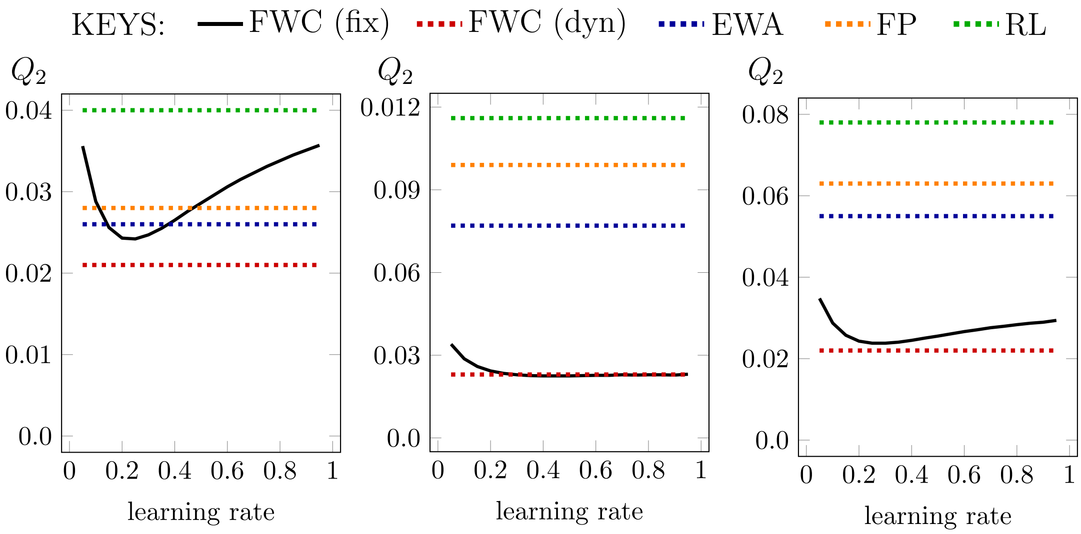
Figure 15 summarizes the results. The dotted lines depict the original Q_2 values from Figure 6 for the four algorithms: FWC (red), EWA (blue), FP (orange) and RL (green). The solid black line shows the Q_2 values obtained for FWC over a range of fixed learning rates in $(0.05, 0.95)$. The left and middle panel show the Q_2 values over 12 PD and 12 SH games, respectively. The right panel shows the Q_2 values over the 24 games combined. Note that the y-axes are scaled differently across panels, because all algorithms perform better (and thus have a lower Q_2) at describing round-by-round behavior in PD games than in SH games.

By visual inspection, it is clear that using a fixed learning rate entails no significant degradation in the performance of FWC over the harder task of matching behavior in the 12 SH games, as well as in the combined dataset. On the other hand, choosing an extreme fixed learning rate (very close to 0.05 or higher than 0.45) has a negative impact on FWC's ability to match round-by-round behavior.

A.3.2 Memory size

The update rule (2) for FWC acts only on the opponents' last action, so that agents (implicitly) use a memory of size 1. Our second robustness test checks the effect of using a longer memory.

Suppose that agents consider at most the last m rounds of play, and let k_t be the number of rounds already played before round t. Then $n_t = \min(m, k_t)$ is the actual



(a) 12 PD games (b) 12 SH games (c) 24 PD+SH games

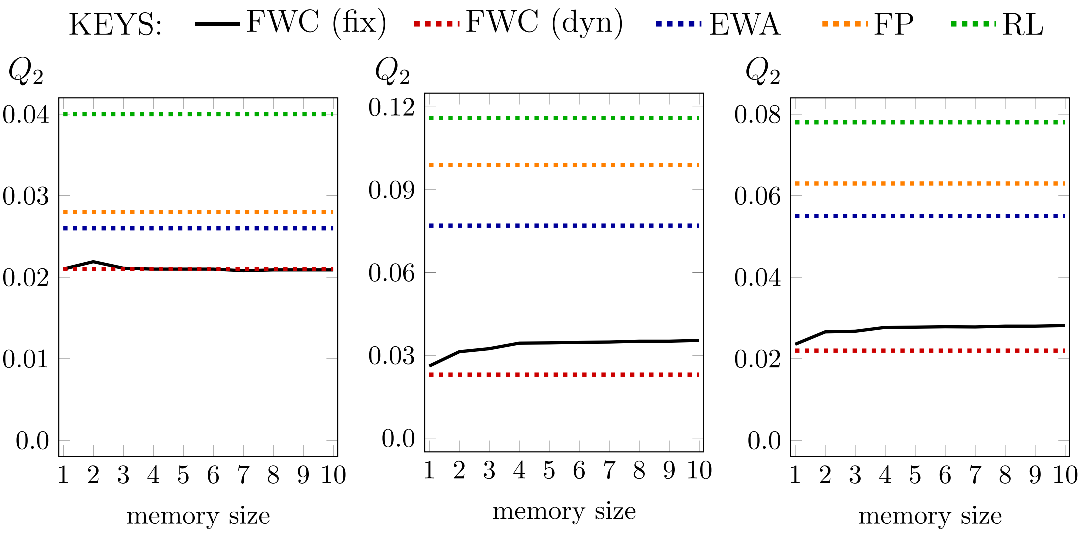
Figure 15: Robustness check for the learning rate. The dotted lines depict the Q_2 values from Figure 6; the solid black line shows the Q_2 values for FWC using different fixed learning rates.

number of rounds that can be used for the update rule in period t. We assume that the update rule is based on the most frequent opponents' action over the last n_t rounds. Let $H_i(n_t)$ be the number of times that the opponents of agent's i have played H in the last n_t rounds and define $\varphi_t = H_i(n_t)/n_t$ as the relative frequency of H-play over the last n_t rounds. If $\varphi_t < 1/2$, then L was played more frequently; if $\varphi_t > 1/2$, then H was played more frequently.

We generalize the original update rule (2) for FWC with $m = 1$ to an arbitrary memory of size m as follows:

$$(f', g', h')_i = \begin{cases} (f + \gamma_i,\ g,\ h)_i & \text{if } \varphi_t < \frac{1}{2} \\[2ex] (f,\ g + \gamma_i,\ h)_i & \text{if } \varphi_t > \frac{1}{2} \\[2ex] (f,\ g,\ h + \gamma_i)_i & \text{if } \varphi_t > \frac{1}{2} \text{ and } (H, H) \text{ gives} \\ & \quad \text{the highest possible payoff} \end{cases}$$

For each value of m between 1 and 10, we compute the mean square deviation Q_2 for the round-by-round H-rates in our dataset. Figure 16 summarizes the results, using the same conventions of Figure 15. It is apparent that FWC outperforms the other algorithms for any memory size. The only noticeable effect is that expanding the memory size slightly degrades FWC's performance in SH games.



(a) 12 PD games (b) 12 SH games (c) 24 PD+SH games

Figure 16: Robustness check for the memory size. The dotted lines depict the Q_2 values from Figure 6; the solid black line shows the Q_2 values for FWC using different memory sizes.

A.3.3 Softmax action selection

Our third and last robustness check is more laborious. The four plain-vanilla algorithms (FWC, EWA, FP, and RL) select actions probabilistically. We upgrade each of them to a one-parameter softmax version, where the parameter is used to tune up the probabilities with which an action is chosen; see Sutton and Barto (1998, Section 2.3). After separately calibrating the parameter for each algorithm over half of the dataset, we test its performance by computing Q_2 over the other half of the dataset.

The softmax version of our algorithms computes the probability that an agent plays H as

$$P(H) = \frac{e^{\lambda \sigma_H}}{e^{\lambda \sigma_H} + e^{\lambda \sigma_L}}$$

This depends on a calibration (heat) parameter $\lambda \geq 0$ and on the strengths σ_H, σ_L. When the heat parameter is $\lambda = 0$, $P(H) = 1/2$; as λ increases, the probability to choose the action with the highest strength approaches 1. The respective strengths σ_H, σ_L are chosen separately for each algorithm in accordance with its characteristic features. Keeping the notation from Appendix A.1, they are given by: $\sigma_H = U_i^G(H)$ and $\sigma_L = U_i^G(L)$ for FP; $\sigma_H = \alpha_H$ and $\sigma_L = \alpha_L$ for RL; $\sigma_H = A_H$ and $\sigma_L = A_L$ for EWA; $\sigma_H = d_E + (d_T)^- + (d_R)^-$ and $\sigma_L = (d_T)^+ + (d_R)^+$ for FWC.

For each algorithm, we randomly split our dataset of 24 PD+SH games into two groups of equal size. First, we calibrate λ to the value λ^* that minimizes Q_2 over the first group;

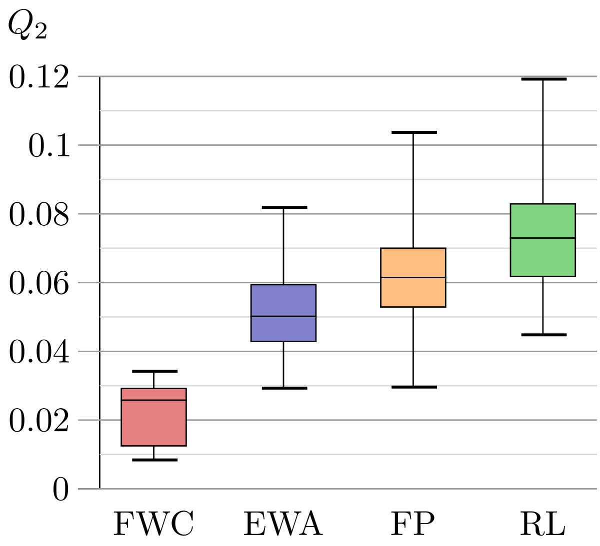


Figure 17: Robustness check for the calibrated softmax versions. Boxplots for the Q_2 values from 100 tests.

we search for the best value of λ over the interval $[0, 10]$ using a tick size of 0.01 in $[0, 0.2]$ and of 0.05 in $(0.2, 10]$. Second, we use λ^* to compute Q_2 over the second group. This approach simulates the calibration of the heat parameter over an existing sample, using the result to predict behavior out-of-sample.

We run the procedure 100 times for each algorithm, with independent draws for the group composition and all other variables. Figure 17 displays the results using a box plot for each algorithm. The calibrated softmax version of the FWC rule clearly outperforms the other algorithms, similarly to what we found for the plain-vanilla versions.

Supplementary Material (SM)

Contents:

Table SM.1: Database of 45 experiments over identical games (26 PD, 19 SH), published between 1992 and 2018 and ordered by year of publication. Each entry lists the game's payoffs (a, b, c, d), the game type (PD or SH), the average H-rate rate over all participants, the number of rounds, the number of trials, the number of participants per trial, and the source. We starred the 24 games (12 PD, 12 SH) for which the round-by-round average H-rates are available: these data are listed next as Table SM.2.

ID	Payoffs a, b, c, d	Type	Avg. H-rate	Rounds	Trials	Partic.	Source
01	100,0,80,80	SH	0.061	22	3	10	Cooper et al. (1992)
02*	7,0,12,4	PD	0.18	10	1	14	Andreoni and Miller (1993)
03*	100,0,80,80	SH	0.189	9	1	10	Straub (1995)
04*	55,25,35,35	SH	0.967	9	1	10	Straub (1995)
05*	100,20,60,60	SH	0.889	9	1	10	Straub (1995)
06*	100,20,80,80	SH	0.1	9	1	10	Straub (1995)
07*	80,10,70,30	SH	0.833	9	1	10	Straub (1995)
08*	800,0,1000,350	PD	0.224	10	1	40	Cooper et al. (1996)
09*	6,0,9,3	PD	0.192	15	8	6	Andreoni and Varian (1999)
10*	7,0,11,4	PD	0.325	15	8	6	Andreoni and Varian (1999)
11	45,0,35,40	SH	0.105	75	8	8	Battalio et al. (2001)
12	45,0,40,20	SH	0.392	75	8	8	Battalio et al. (2001)
13	45,0,42,12	SH	0.493	75	8	8	Battalio et al. (2001)
14*	1000,0,800,800	SH	0.17	10	2	10	Clark et al. (2001)
15*	1000,0,900,700	SH	0.19	10	2	10	Clark et al. (2001)
16*	1000,0,700,900	SH	0.165	10	1	10	Clark et al. (2001)
17	60,10,110,40	PD	0.11	4	2	8	Schmidt et al. (2001)
18	70,10,110,30	PD	0.28	4	2	8	Schmidt et al. (2001)
19	70,10,110,50	PD	0.08	4	2	8	Schmidt et al. (2001)
20	80,10,110,40	PD	0.24	4	2	8	Schmidt et al. (2001)
21	80,10,110,60	PD	0.11	4	2	8	Schmidt et al. (2001)
22	90,10,110,50	PD	0.24	4	2	8	Schmidt et al. (2001)
23*	100,20,60,60	SH	0.728	8	4	10	Schmidt et al. (2003)
24*	100,20,80,80	SH	0.681	8	4	10	Schmidt et al. (2003)
25*	100,60,80,80	SH	0.834	8	4	10	Schmidt et al. (2003)
26*	100,0,80,60	SH	0.438	8	4	10	Schmidt et al. (2003)
27	350,0,500,150	PD	0.21	5	3	8	Bohnet and Kuebler (2005)
28	350,100,500,150	PD	0.2	5	3	8	Bohnet and Kuebler (2005)
29	0.105,0.005,0.175,0.075	PD	0.61	10	3	22	Bereby-Meyer and Roth (2006)
30	800,100,850,150	PD	0.51	100	8	4	Grimm and Mengel (2009)
31	800,100,1100,400	PD	0.098	100	8	4	Grimm and Mengel (2009)
32	50,10,60,40	PD	0.18	10	18	16	Dal Bo et al. (2010)
33	45,0,42,12	SH	0.57	10	8	8	Dubois et al. (2012)
34	40,20,37,32	SH	0.34	10	8	8	Dubois et al. (2012)
35	44,4,38,28	SH	0.44	10	8	8	Dubois et al. (2012)
36	150,40,850,50	PD	0.46	10	1	24	Mengel (2018)
37	250,100,750,160	PD	0.26	10	1	24	Mengel (2018)
38*	150,5,850,95	PD	0.2	10	1	24	Mengel (2018)
39*	250,50,750,150	PD	0.129	10	1	24	Mengel (2018)
40*	10,1,90,5	PD	0.229	10	1	24	Mengel (2018)
41*	400,100,450,120	PD	0.429	10	1	24	Mengel (2018)
42*	400,100,450,200	PD	0.204	10	1	24	Mengel (2018)
43*	10,1,90,5	PD	0.085	10	1	24	Mengel (2018)
44*	10,1,110,9	PD	0.221	10	1	24	Mengel (2018)
45*	400,10,800,200	PD	0.108	10	1	24	Mengel (2018)

Table SM.2: Round-by-round average H-rates for the 24 identical-game experiments (12 PD, 12 SH) that are starred in Table SM.1.

ID	Type	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15
02	PD	0.36	0.28	0.21	0.2	0.16	0.15	0.14	0.14	0.14	0.08	-	-	-	-	-
03	SH	0.4	0.2	0.4	0.1	0.3	0.1	0.1	0.1	0	-	-	-	-	-	-
04	SH	0.9	0.8	1	1	1	1	1	1	1	-	-	-	-	-	-
05	SH	0.6	0.7	0.9	0.8	1	1	1	1	1	-	-	-	-	-	-
06	SH	0.2	0.3	0.2	0.1	0	0.1	0	0	0	-	-	-	-	-	-
07	SH	0.9	0.8	0.8	0.7	0.9	0.8	0.8	0.9	0.9	-	-	-	-	-	-
08	PD	0.3	0.35	0.3	0.1	0.2	0.25	0.23	0.17	0.2	0.14	-	-	-	-	-
09	PD	0.39	0.25	0.33	0.12	0.17	0.21	0.12	0.29	0.05	0.12	0.12	0.25	0.17	0.12	0.17
10	PD	0.26	0.42	0.33	0.37	0.33	0.25	0.37	0.33	0.42	0.33	0.21	0.21	0.33	0.29	0.42
14	SH	0.3	0.35	0.175	0.225	0.2	0.075	0.075	0.05	0.1	0.15	-	-	-	-	-
15	SH	0.525	0.25	0.275	0.2	0.225	0.15	0.075	0	0.1	0.1	-	-	-	-	-
16	SH	0.5	0.4	0.3	0.3	0.1	0.05	0	0	0	0	-	-	-	-	-
23	SH	0.65	0.7	0.675	0.65	0.7	0.775	0.825	0.85	-	-	-	-	-	-	-
24	SH	0.575	0.775	0.675	0.675	0.725	0.725	0.65	0.65	-	-	-	-	-	-	-
25	SH	0.75	0.775	0.8	0.775	0.825	0.9	0.925	0.925	-	-	-	-	-	-	-
26	SH	0.475	0.475	0.575	0.55	0.475	0.375	0.35	0.225	-	-	-	-	-	-	-
38	PD	0.458	0.25	0.208	0.333	0.208	0.125	0.125	0.083	0.167	0.042	-	-	-	-	-
39	PD	0.25	0.208	0	0.25	0.125	0.125	0	0.167	0.083	0.083	-	-	-	-	-
40	PD	0.208	0.292	0.417	0.208	0.333	0.125	0.167	0.167	0.167	0.208	-	-	-	-	-
41	PD	0.708	0.458	0.333	0.542	0.625	0.417	0.458	0.292	0.25	0.208	-	-	-	-	-
42	PD	0.583	0.375	0.208	0.083	0.208	0.167	0.25	0.042	0.083	0.042	-	-	-	-	-
43	PD	0.2	0.15	0.1	0.1	0.05	0.15	0	0.05	0	0.05	-	-	-	-	-
44	PD	0.375	0.25	0.333	0.333	0.083	0.25	0.333	0.042	0.125	0.083	-	-	-	-	-
45	PD	0.333	0.083	0.208	0.125	0.042	0.125	0.042	0.042	0.083	0	-	-	-	-	-

Table SM.3: Sequences of similar games in multi-game environments. The payoffs for the games in sequence SQ_2 from Rankin et al. (2000) have been obtained from the appendix of the working paper version for Van Huyck and Stahl (2018). The length of the subsequences for SQ_3 to SQ_6 from Duffy and Fehr (2018) may change due to the specifics of the experimental setup; see original paper for information.

Space of games	Sequences	Source
$PD_1 = (80, 10, 110, 40)$, $PD_2 = (70, 10, 110, 50)$, $PD_3 = (90, 10, 110, 50)$, $PD_4 = (80, 10, 110, 60)$, $PD_5 = (70, 10, 110, 30)$, $PD_6 = (60, 10, 110, 40)$	$SQ_1 = [PD_1, PD_2, PD_3, PD_4, PD_5, PD_6, PD_1, PD_2, PD_3, PD_4, PD_5, PD_6, PD_1, PD_2, PD_3, PD_4, PD_5, PD_6, PD_1, PD_2, PD_3, PD_4, PD_5, PD_6]$	Schmidt et al. (2001)
$\{(370+\varepsilon, \varepsilon, n+\varepsilon, n+\varepsilon) \mid 1 \leq n \leq 369, 0 \leq \varepsilon \leq 50\}$	$SQ_2 = [(376, 6, 299, 299), (415, 45, 57, 57), (376, 6, 313, 313), (407, 37, 57, 57), (417, 47, 415, 415), (392, 22, 149, 149), (411, 41, 256, 256), (372, 2, 356, 356), (418, 48, 242, 242), (389, 19, 175, 175), (412, 42, 108, 108), (371, 1, 138, 138), (388, 18, 229, 229), (404, 34, 84, 84), (397, 27, 225, 225), (400, 30, 173, 173), (419, 49, 62, 62), (408, 38, 315, 315), (370, 0, 112, 112), (379, 9, 330, 330), (407, 37, 377, 377), (405, 35, 388, 388), (402, 32, 320, 320), (419, 49, 309, 309), (399, 29, 77, 77), (390, 20, 247, 247), (401, 31, 234, 234), (373, 3, 114, 114), (398, 28, 117, 117), (384, 14, 235, 235), (376, 6, 322, 322), (401, 31, 85, 85), (419, 49, 237, 237), (400, 30, 309, 309), (419, 49, 134, 134), (388, 18, 269, 269), (416, 46, 300, 300), (377, 7, 186, 186), (402, 32, 96, 96), (401, 31, 120, 120), (370, 0, 305, 305), (399, 29, 38, 38), (416, 46, 190, 190), (390, 20, 167, 167), (377, 7, 287, 287), (386, 16, 348, 348), (384, 14, 165, 165), (396, 26, 195, 195), (394, 24, 156, 156), (379, 9, 252, 252), (373, 3, 17, 17), (379, 9, 150, 150), (414, 44, 145, 145), (393, 23, 145, 145), (370, 0, 199, 199), (392, 22, 372, 372), (419, 49, 157, 157), (417, 47, 307, 307), (391, 21, 374, 374), (391, 21, 175, 175), (412, 42, 400, 400), (389, 19, 366, 366), (404, 34, 207, 207), (378, 8, 171, 171), (402, 32, 283, 283), (372, 2, 201, 201), (412, 42, 109, 109), (409, 39, 259, 259), (382, 12, 297, 297), (404, 34, 103, 103), (407, 37, 150, 150), (395, 25, 106, 106), (381, 11, 367, 367), (400, 30, 231, 231), (415, 45, 67, 67)]$	Rankin et al. (2000)
$PD_1 = (20, 0, 30, 10)$, $SH_1 = (20, 0, 10, 10)$, $PD_2 = (20, 0, 25, 10)$, $SH_2 = (20, 0, 15, 10)$	$SQ_3 = [37\times\ PD_1, 31\times\ SH_1, 33\times\ PD_1]$, $SQ_4 = [33\times\ SH_1, 46\times\ PD_1, 28\times\ SH_1]$, $SQ_5 = [47\times\ PD_2, 42\times\ SH_2, 31\times\ PD_2]$, $SQ_6 = [28\times\ SH_2, 29\times\ PD_2, 33\times\ SH_2]$	Duffy and Fehr (2018)
Data source for sequence SQ_2: `https://www.semanticscholar.org/paper/Learning-Conditional-Behavior-in-Similar-Stag-Hunt-Stahl-Huyck/7a30888b650ea8696b24ff50b689c60657a9f1e6`		
Data source for sequences SQ_3 to SQ_6:	`https://sites.google.com/site/dietmarfehr`	

Table SM.4: Experimental results from multi-game environments: the sequence IDs match those in Table SM.3. All data are accessible from their respective source.

Seq.	Partic.	Type of data	Data	Source
SQ_1	8	Round-by-round H-rates, all 24 rounds	[0.396, 0.208, 0.229, 0.125, 0.271, 0.104, 0.125, 0, 0.208, 0.063, 0.25, 0.063, 0.229, 0.125, 0.396, 0.062, 0.271, 0.042, 0.083, 0.042, 0.084, 0.042, 0.146, 0.021]	Schmidt et al. (2001)
SQ_2	8	H-rates averaged over first 10 rounds	0.84 for low-x games, 0.63 for high-x games	Rankin et al. (2000)
		H-rates averaged over last 10 rounds	1.0 for low-x games, 0.91 for high-x games	
SQ_3	10 or 20	round-by-round H-rates, max 101 rounds	Fig. SM.3, data source: see Tab. SM.3, last row	Duffy and Fehr (2018)
SQ_4	10 or 20	round-by-round H-rates, max 107 rounds	Fig. SM.3, data source: see Tab. SM.3, last row	
SQ_5	10 or 20	round-by-round H-rates, max 120 rounds	Fig. SM.3, data source: see Tab. SM.3, last row	
SQ_6	20	round-by-round H-rates, max 90 rounds	Fig. SM.3, data source: see Tab. SM.3, last row	

Table SM.5: Database of 6 experiments over identical CG games. Round-by-round average H-rates are available for the 3 starred games. For the other 3 games there are only average H-values cumulated over a number of rounds.

ID	Payoffs a, b, c, d	Type	Avg. H-rate	Rounds	Trials	Partic.	Source
46	2,2,5,0	CG	0.431	40	1	10	Bornstein et al. (1997)
47	39, 9, 48, 3	CG	0.455	60	4	ca. 13**	Cason and Sharma (2007)
48*	200,100,300,0	CG	0.477	12	8	6	Kümmerli et al. (2007)
49	7,3,9,0	CG	0.554	20	16	12	Duffy and Feltovich (2010)
50*	160,80,200,20	CG	0.562	20	6	ca. 15***	Feltovich (2011)
51*	120,40,160,-20	CG	0.650	20	6	ca. 15***	Feltovich (2011)

** Authors indicate 4 sessions with 52 participants in total

*** Author indicates 6 sessions with 90 participants in total

Table SM.6: Round-by-round average H-rates for games 48, 50 and 51; cumulated H-rates at every 10 rounds for games 46 and 47, and at every 5 rounds for game 49.

ID	Type	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20
48	CG	0.75	0.646	0.458	0.396	0.396	0.333	0.563	0.542	0.417	0.396	0.396	0.438	-	-	-	-	-	-	-	-
50	CG	0.675	0.6	0.675	0.53	0.6	0.52	0.635	0.56	0.54	0.57	0.5	0.49	0.56	0.575	0.52	0.53	0.51	0.52	0.56	0.53
51	CG	0.82	0.75	0.67	0.725	0.68	0.625	0.625	0.6	0.65	0.64	0.66	0.59	0.64	0.63	0.63	0.61	0.62	0.61	0.61	0.57

ID	Type	R1-10	R11-20	R21-30	R31-40	R41-50	R51-60
46	CG	0.34	0.395	0.47	0.52	-	-
47	CG	0.57	0.47	0.43	0.47	0.43	0.39

ID	Type	R1-5	R6-10	R11-15	R16-20
49	CG	0.653	0.587	0.602	0.543

Table SM.7: Database of 4 experiments over identical 3×3 games from Cooper et al. (1990), where they are labeled as games 3, 4, 5, 6 and shown in Figure 4 on page 222.

ID	Payoffs from top left to bottom right, by row									Avg. H, M, L rates			Rounds	Trials	Partic.	Source
52	600	0	0	0	550	250	1000	350	350	0.086,	0.2,	0.714	22	1	11	Cooper et al. (1990)
53	600	0	0	0	550	250	700	350	350	0.15,	0.141,	0.709	22	1	11	Cooper et al. (1990)
54	600	0	0	1000	550	250	700	350	350	0.059	0.918	0.023	22	1	11	Cooper et al. (1990)
55	600	0	0	650	550	250	700	350	350	0.145,	0.786,	0.068	22	1	11	Cooper et al. (1990)

Table SM.8: Round-by-round average H-, M- and L-rates for the four identical 3x3 games listed in Table SM.7.

ID	Choice	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20	R21	R22
	H	0.4	0.3	0.1	0.2	0.1	0.2	0.2	0	0	0.1	0.2	0	0	0	0	0	0	0	0	0	0.1	0
52	M	0.2	0.1	0.1	0.2	0.2	0.2	0	0.3	0.1	0.2	0.2	0.4	0.4	0.3	0.4	0.2	0.2	0.2	0.1	0.2	0.1	0.1
	L	0.4	0.6	0.8	0.6	0.7	0.6	0.8	0.7	0.9	0.7	0.6	0.6	0.6	0.7	0.6	0.8	0.8	0.8	0.9	0.8	0.8	0.9
	H	0.2	0.2	0.2	0.3	0.2	0.3	0.3	0.1	0.2	0.1	0	0.2	0	0.2	0.1	0.1	0	0	0.3	0.1	0.1	0.1
53	M	0.4	0.2	0.4	0.3	0.1	0.2	0.1	0	0	0.1	0.1	0.1	0.2	0.2	0.3	0.2	0	0	0.1	0.1	0	0
	L	0.4	0.6	0.4	0.4	0.7	0.5	0.6	0.9	0.8	0.8	0.9	0.7	0.8	0.6	0.6	0.7	1	1	0.6	0.8	0.9	0.9
	H	0	0.3	0.1	0.5	0	0	0	0.1	0	0.1	0	0	0	0.1	0	0	0	0	0	0	0	0.1
54	M	0.9	0.7	0.9	0.5	1	0.9	1	0.8	0.9	0.9	0.9	1	1	0.9	1	1	1	1	1	1	1	0.9
	L	0.1	0	0	0	0	0.1	0	0.1	0.1	0	0.1	0	0	0	0	0	0	0	0	0	0	0
	H	0.5	0.4	0.3	0.4	0.4	0.2	0.3	0.2	0.2	0	0	0.2	0.1	0	0	0	0	0	0	0	0	0
55	M	0.5	0.6	0.6	0.5	0.5	0.6	0.6	0.6	0.6	0.7	0.8	0.8	0.9	1	1	1	1	1	1	1	1	1
	L	0	0	0.1	0.1	0.1	0.2	0.1	0.2	0.2	0.3	0.2	0	0	0	0	0	0	0	0	0	0	0

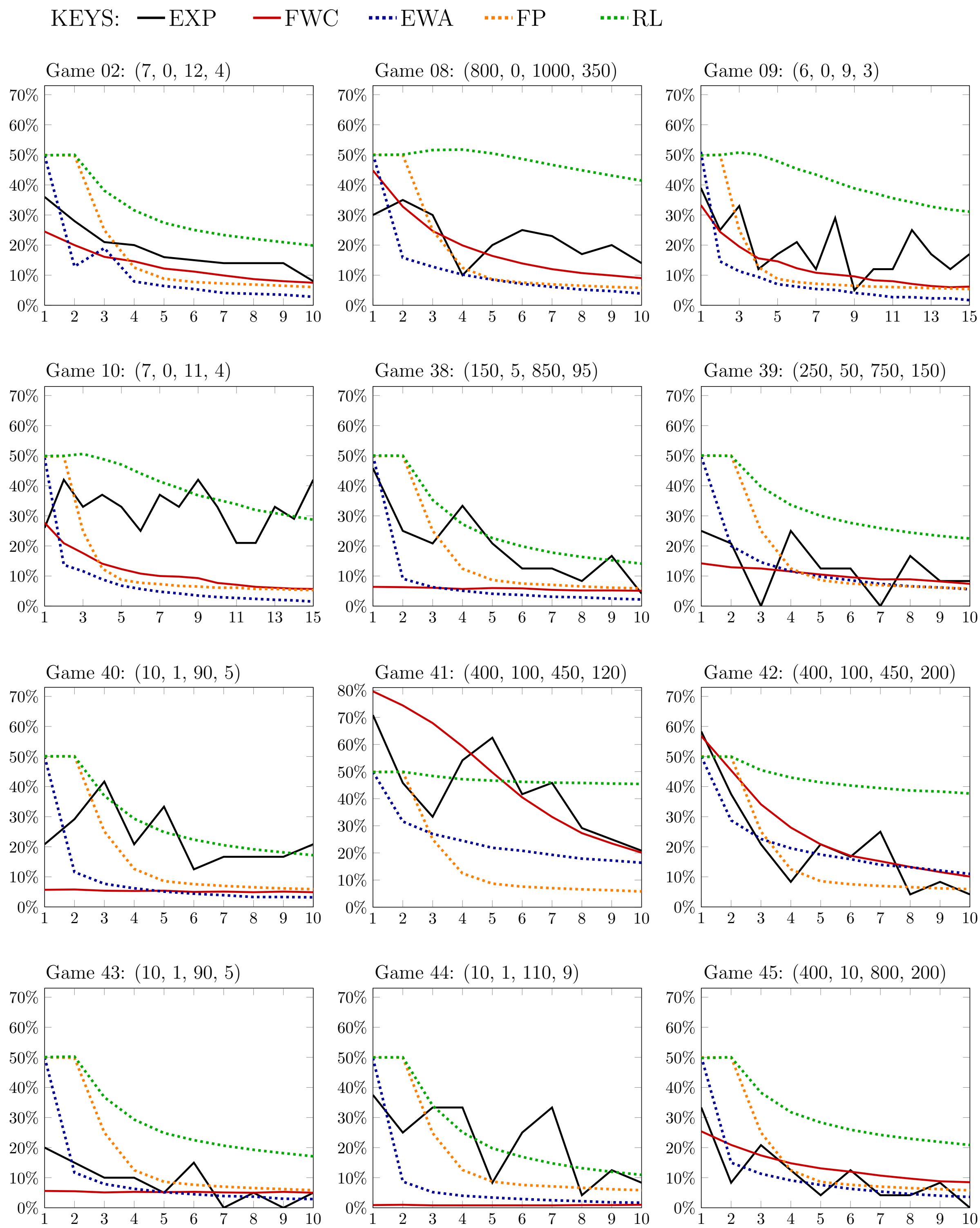


Figure SM.1: Round-by-round average H-rates for the experimental results over the 12 starred PD games from Table SM.1 (solid black line) and for the simulated H-rates averaged over 1000 simulation runs: AWC (solid red line), EWA (dotted blue line), FP (dotted orange line) and RL (dotted green line).

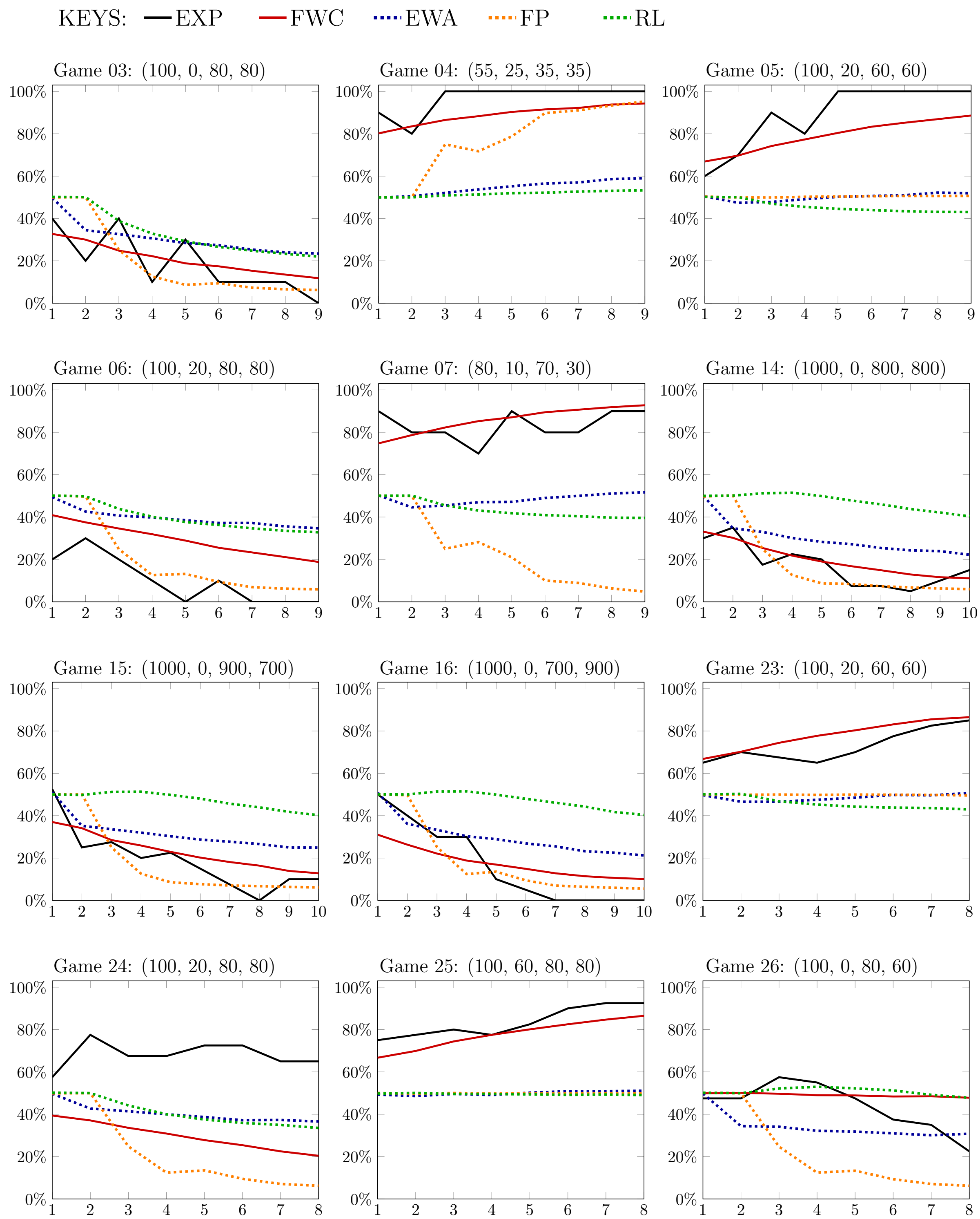


Figure SM.2: Round-by-round average H-rates for the experimental results over the 12 starred SH games from Table SM.1 (solid black line) and for the simulated H-rates averaged over 1000 simulation runs: AWC (solid red line), EWA (dotted blue line), FP (dotted orange line) and RL (dotted green line).

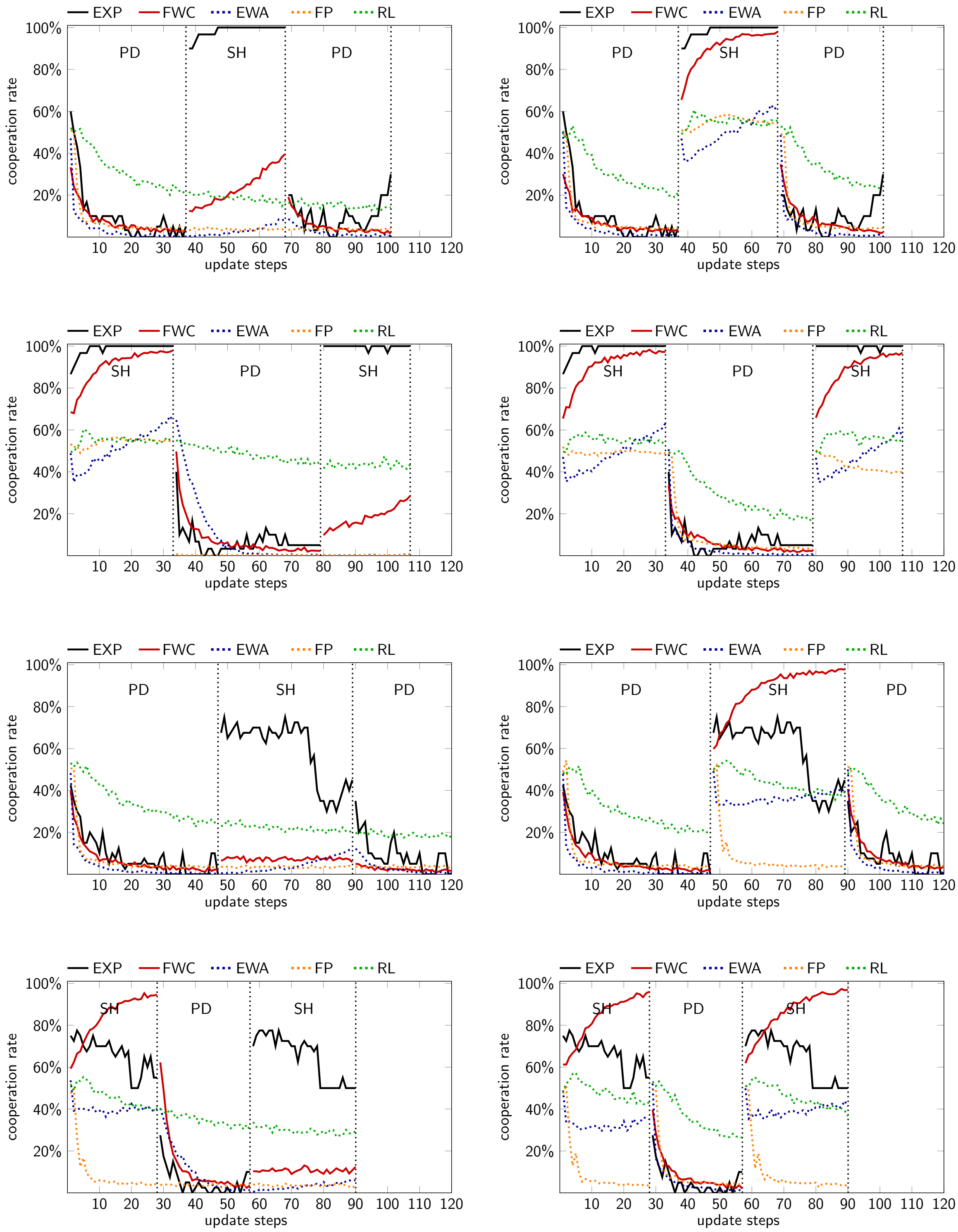


Figure SM.3: Round-by-round average H-rates for the experimental results over the four sequences SQ_3 (first row), SQ_4 (second row), SQ_5 (third row) and SQ_6 (fourth row), as given in Table SM.3. Left (right) panels show results without (with) reset of the learning rules after a switch in the game class.

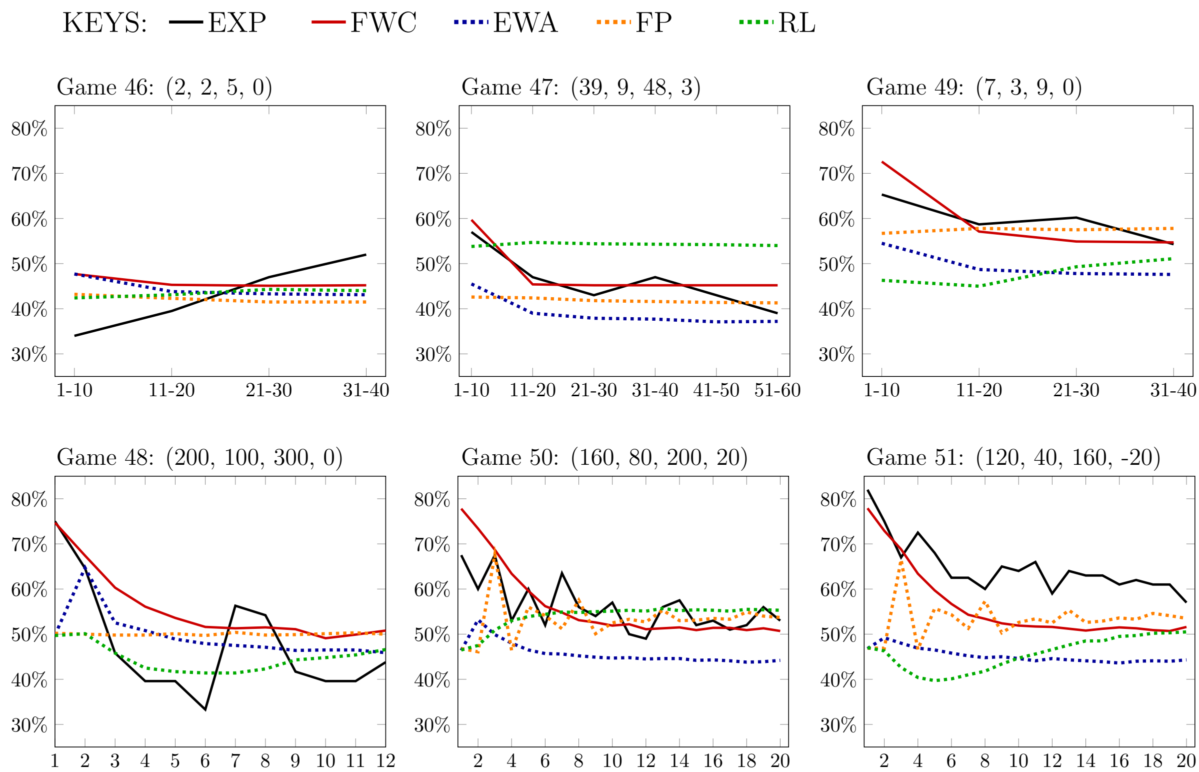


Figure SM.4: Block-by-block (top panels) and round-by-round (bottom panels) average H-rates for the experimental results over the 6 chicken games from Table SM.5 (solid black line) and for the simulated H-rates averaged over 1000 simulation runs: FWC (solid red line), EWA (dotted blue line), FP (dotted orange line) and RL (dotted green line).

References

J. Andreoni and J. H. Miller. Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. *Economic Journal*, 103(418):570–585, 1993.

J. Andreoni and H. Varian. Preplay contracting in the prisoner's dilemma. *Proceedings of the National Academy of Sciences*, 96:10933–10938, 1999.

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

Y. Bereby-Meyer and A. Roth. The speed of learning in noisy games: Partial reinforcement and the sustainability of cooperation. *American Economic Review*, 96(4):1029–1042, 2006.

I. Bohnet and D. Kuebler. Compensating the cooperators: Is sorting in the prisoner's dilemma possible? *Journal of Economic Behavior and Organization*, 56:61–76, 2005.

G. Bornstein, D. Budescu, and S. Zamir. Cooperation in intergroup, n-person, and two-person games of chicken. *Journal of Conflict Resolution*, 41(3):384–406, 1997.

T. N. Cason and T. Sharma. Recommended play and correlated equilibria: an experimental study. *Economic Theory*, 33:11–27, 2007.

T. Chmura, S. J. Goerg, and R. Selten. Learning in experimental 2×2 games. *Games and Economic Behavior*, 76:44–73, 2012.

K. Clark, S. Kay, and M. Sefton. When are Nash equilibria self-enforcing? An experimental analysis. *International Journal of Game Theory*, 29:495–515, 2001.

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

R. Cooper, D. V. DeJong, R. Forsythe, and T. W. Ross. Communication in coordination games. *Quarterly Journal of Economics*, 107(2):739–771, 1992.

R. Cooper, D. V. DeJong, R. Forsythe, and T. W. Ross. Cooperation without reputation: Experimental evidence from prisoner's dilemma games. *Games and Economic Behavior*, 12 (2):187–218, 1996.

P. Dal Bo, A. Foster, and L. Putterman. Institutions and behavior: Experimental evidence on the effects of democracy. *American Economic Review*, 100(5):2205–2229, 2010.

D. Dubois, M. Willinger, and P. Van Nguyen. Optimization incentive and relative riskiness in experimental stag-hunt games. *International Journal of Game Theory*, 41:369–380, 2012.

J. Duffy and D. Fehr. Equilibrium selection in similar repeated games: Experimental evidence on the role of precedents. *Experimental Economics*, 21:573–600, 2018.

J. Duffy and N. Feltovich. Correlated equilibria, good and bad: An experimental study. *International Economic Review*, 51(3):573–600, 2010.

N. Feltovich. The effect of subtracting a constant from all payoffs in a hawk-dove game: Experimental evidence of loss aversion in strategic behavior. *Southern Economic Journal*, 77(4): 814–826, 2011.

V. Grimm and F. Mengel. Cooperation in viscous populations–experimental evidence. *Games and Economic Behavior*, 66:202–220, 2009.

R. Kümmerli, C. Colliard, N. Fiechter, B. Petitpierre, F. Russier, and L. Keller. Human cooperation in social dilemmas: comparing the snowdrift game with the prisoner's dilemma. *Proceedings of the Royal Society B*, 274:2965–2970, 2007.

F. Mengel. Risk and temptation: A meta-study on prisoner's dilemma games. *Economic Journal*, 128(616):3182–3209, 2018.

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

D. Schmidt, R. Shupp, J. Walker, T. Ahn, and E. Ostrom. Dilemma games: Game parameters and matching protocols. *Journal of Economic Behavior and Organization*, 46:357–377, 2001.

D. Schmidt, R. Shupp, J. M. Walker, and E. Ostrom. Playing safe in coordination games: The roles of risk dominance, payoff dominance, and history of play. *Games and Economic Behavior*, 42:281–299, 2003.

P. G. Straub. Risk dominance and coordination failures in static games. *Quarterly Review of Economics and Finance*, 35(4):339–363, 1995.

R. Sutton and A. Barto. *Reinforcement Learning: An Introduction.* The MIT Press, Cambridge, MA, 1998.

J. B. Van Huyck and D. O. Stahl. Conditional behavior and learning in similar stag hunt games. *Experimental Economics*, 21:513–526, 2018.