1 Reactive defecting Nash strategies in the donation game

In the previous sections, we characterized reactive partner strategies for the special case of the donation game and the general prisoner's dilemma. In the following section, we apply the same methods to characterize defecting Nash equilibria. For the case of reactive-1 strategies, we obtain the following characterization.

Theorem 1 (Reactive-1 defecting Nash strategies in the donation game)
A reactive-1 strategy $\mathbf{p}$ *is a defecting Nash strategy if and only if its entries satisfy the conditions*

$$p_C \leq \tfrac{c}{b} \qquad \textit{and} \qquad p_D = 0. \tag{1}$$

Theorem 2 (Reactive-2 defecting Nash strategies in the donation game)
A reactive-2 strategy $\mathbf{p}$ *is a defecting Nash strategy if and only if its entries satisfy the conditions*

$$p_{CC} \leq \frac{c}{b}, \qquad \frac{p_{CD}+p_{DC}}{2} \leq \frac{c}{2b}, \qquad p_{DD} = 0. \tag{2}$$

Theorem 3 (Reactive-3 defecting Nash strategies in the donation game)
A reactive-3 strategy $\mathbf{p}$ *is a defecting Nash strategy if and only if its entries satisfy the conditions*

$$\begin{aligned} p_{CCC} &\leq \frac{c}{b} \\ \frac{p_{CDC}+p_{DCD}}{2} &\leq \frac{1}{2}\cdot\frac{c}{b} \\ \frac{p_{CCD}+p_{CDC}+p_{DCC}}{3} &\leq \frac{2}{3}\cdot\frac{c}{b} \\ \frac{p_{CDD}+p_{DCD}+p_{DDC}}{3} &\leq \frac{1}{3}\cdot\frac{c}{b} \\ \frac{p_{CCD}+p_{CDD}+p_{DCC}+p_{DDC}}{4} &\leq \frac{1}{2}\cdot\frac{c}{b} \\ p_{DDD} &= 0. \end{aligned} \tag{3}$$

We repeat the same analysis for reactive counting strategies. We obtain the following results.

Theorem 4 (Reactive-2 defecting Nash counting strategies in the donation game)
A reactive-2 counting strategy $\mathbf{r}$ *is a defecting Nash strategy if and only if its entries satisfy the conditions*

$$r_2 \leq \frac{c}{b}, \qquad r_1 \leq \frac{1}{2}\cdot\frac{c}{b}, \qquad r_0 = 0. \tag{4}$$

Theorem 5 (Reactive-3 defecting Nash counting strategies in the donation game)
A reactive-3 counting strategy $\mathbf{r}$ *is a defecting Nash strategy if and only if its entries satisfy the conditions*

$$r_3 \leq \frac{c}{b}, \qquad r_2 \leq \frac{2}{3}\cdot\frac{c}{b}, \qquad r_1 \leq \frac{1}{3}\cdot\frac{c}{b}, \qquad r_0 = 0. \tag{5}$$

We can observe that for each value of n, the left-hand side of the conditions for cooperative and defective Nash are the same. Moreover, we see that, for a low cost-to-benefit ratio, the right-hand side of the defective Nash conditions is always strictly smaller than those of the cooperative Nash conditions. This means that within the space of feasible strategies, the volume of partner strategies is larger than the volume of defective Nash strategies. We also verify these analytical results numerically (Figure S1), and we also show that as memory increases, the number of cooperative and defective Nash strategies decreases, as the strategy spaces are larger. However, the decrease is more prominent in the case of defective Nash.

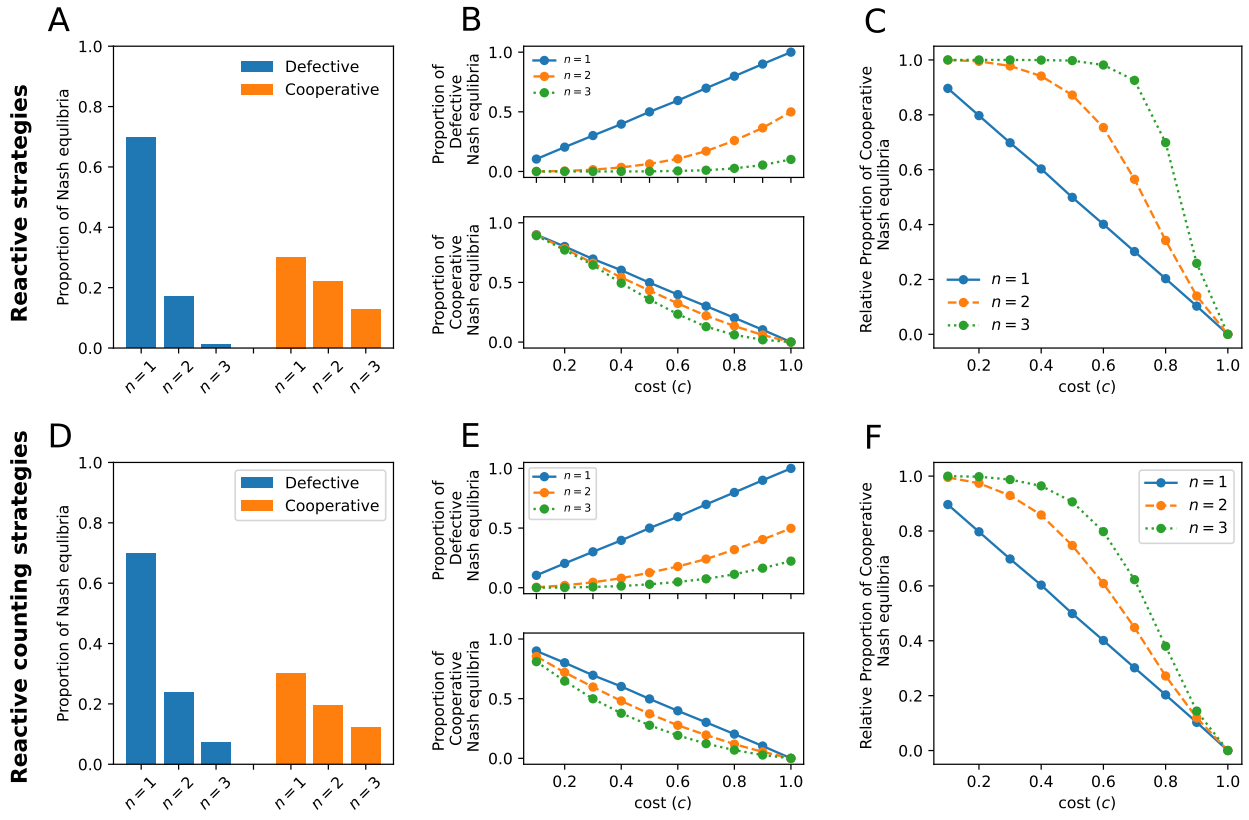

Figure S1: Proportions of cooperative and defective Nash. We draw 10^4 random strategies from the feasible space of strategies and create two copies of each strategy. For one copy, we set the probability of cooperating after full cooperation of the co-player to 1. For the second copy, we set the probability of cooperating after full defection of the co-player to 0. We then check if either copy is Nash: cooperative for the first copy and defective for the second. We set the benefit of cooperation to $b = 1$. **A, D** We plot the results for a given value of cost, $c = 0.5$. We do this for reactive strategies and reactive counting strategies. For $n = 1$, we can see the number of defective equilibria is higher than that of the cooperative. However, as we increase n, this is no longer true. The number of defective Nash decreases drastically. The number of partner strategies does as well, but to a lesser extent. **B, E** We plot the proportion of equilibria over different values of cost, for reactive and reactive counting strategies. As the cost increases, so does the proportion of defective equilibria, and the opposite is true for cooperative. As memory increases, we observe again a significant drop in the proportion of defective strategies, whereas there is a small decrease in the cooperative strategies. **C, D** We plot the relative proportion of cooperative Nash strategies. For this, we consider the sum of cooperative and defecting Nash strategies for each memory size and plot the number of cooperative Nash strategies over the total sum. This proportion increases as memory size increases.

2 Evolutionary Simulations

We repeated the evolutionary analysis of Figure 4 of the main text. This time, for panels **A** and **B**, we ran twenty independent simulations. We did this to check that the results, specifically the mean most abundant strategies, do not change.

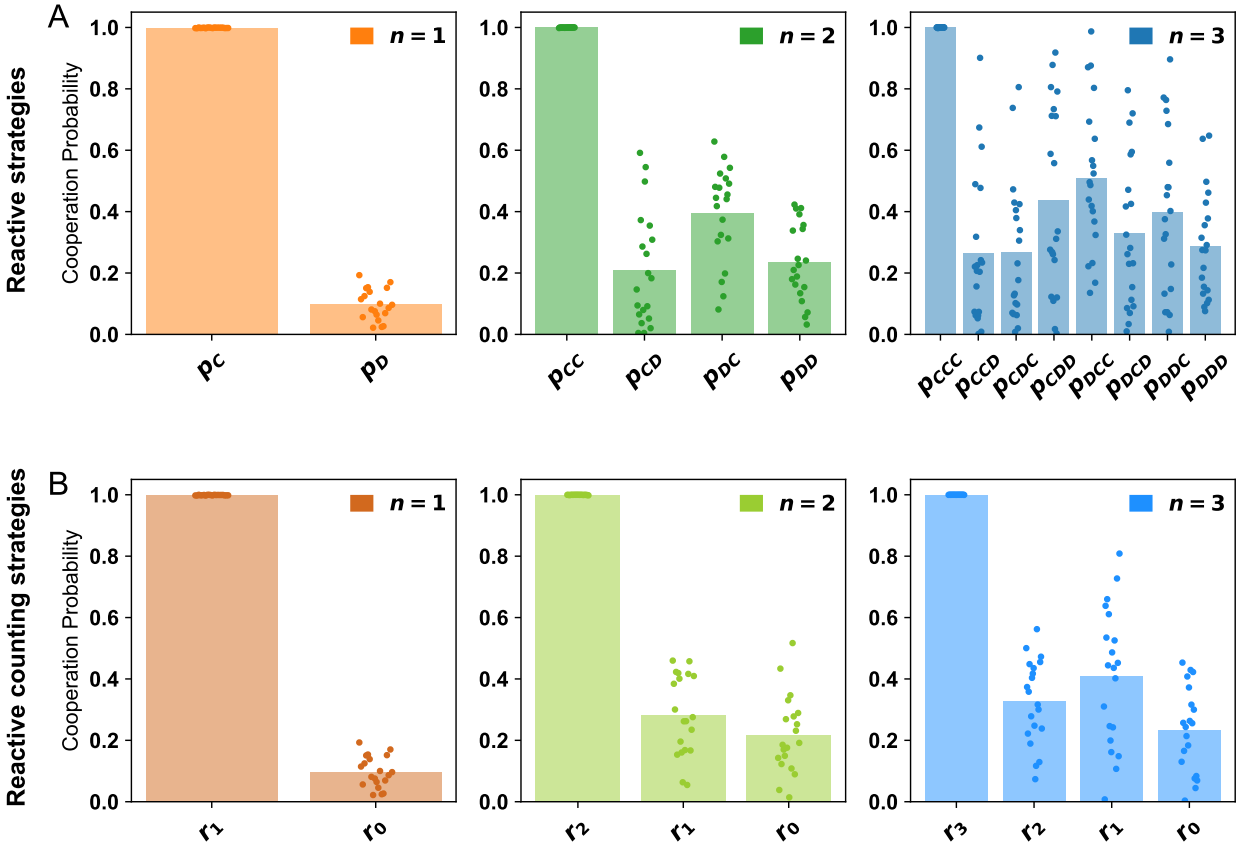

Figure S2: Evolutionary dynamics of reactive-n strategies: Most abundant strategies. The same process as that in the paper for Figure 4 is used, with the same parameters. Simulations are based on a donation game with $b = 1$, $c = 0.5$, a selection strength $\beta = 1$, and a population size $N = 100$. For n equal to 1 and 2, simulations are run for $T = 10^7$ time steps. For $n = 3$, we use $T = 2 \cdot 10^7$ time steps. This time, we run twenty independent simulations instead of ten.

2.1 Invasion Analysis

One question that arises is which of these top strategies are partner strategies, and why some partner strategies are selected more than others. For example, in the case of reactive-1 strategies, we consider a reactive strategy to be partner if $(p_C > 0.95)$ and $(p_D < \frac{c}{b})$. Here, we observe strategies with $(p_D \approx 0.1)$ and not those closer to the boundary. The reason for this is explained by Figure S3.

A similar question arises for reactive-2 strategies. More specifically, why is it that we observe partner strategies where p_{CD} is almost always strictly lower than p_{DC}? Why are these partners chosen more often by the

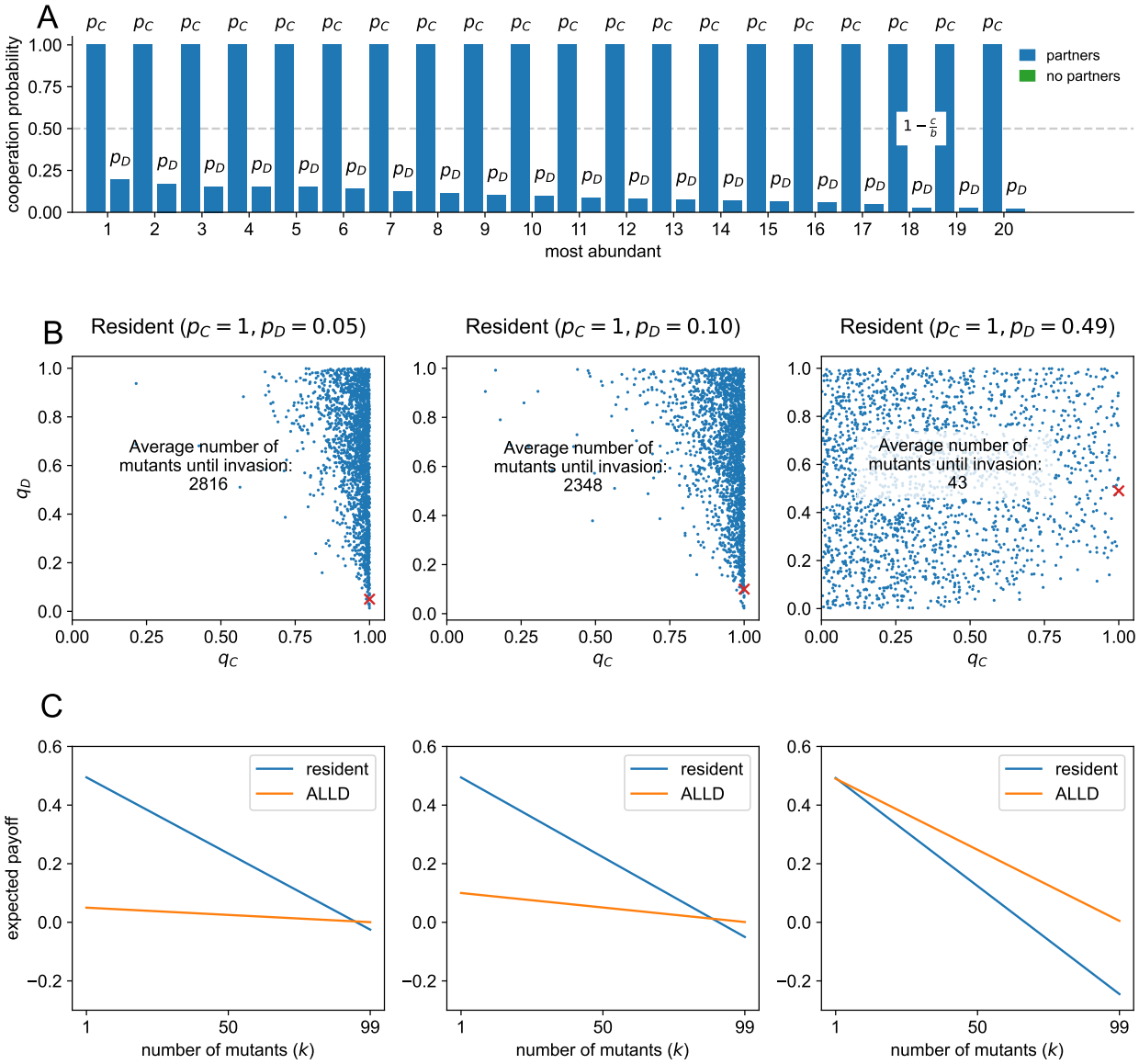

Figure S3: Invasion analysis for reactive-1. A The twenty most abundant reactive-1 strategies of Figure S2. We observe that all twenty of them are partner strategies, and that all have a very low p_D, typically less than 0.2. **B** We perform an invasion analysis. Namely, we select one resident and introduce mutants until a mutant becomes the new resident. We record how many mutants were introduced until this happens. We do this 10^3 times and report the average number of mutants over all the simulations. We consider three different residents. For each, we report the average number of mutants until invasion and the mutants that invaded as well (scatter points). We observed that a resident with a very low p_D is more resistant to invasion, which explains why we observe that the most abundant strategies have a low p_D. **C** To explain why a low p_D leads to more resistance to invasion, we look at the expected payoffs of each resident when when k ALLD mutants try to invade. We can see that for the far-left resident, the expected payoff is almost always strictly higher than that of the ALLD mutants; only when there are 99 mutants is the payoff the same.

evolutionary process?

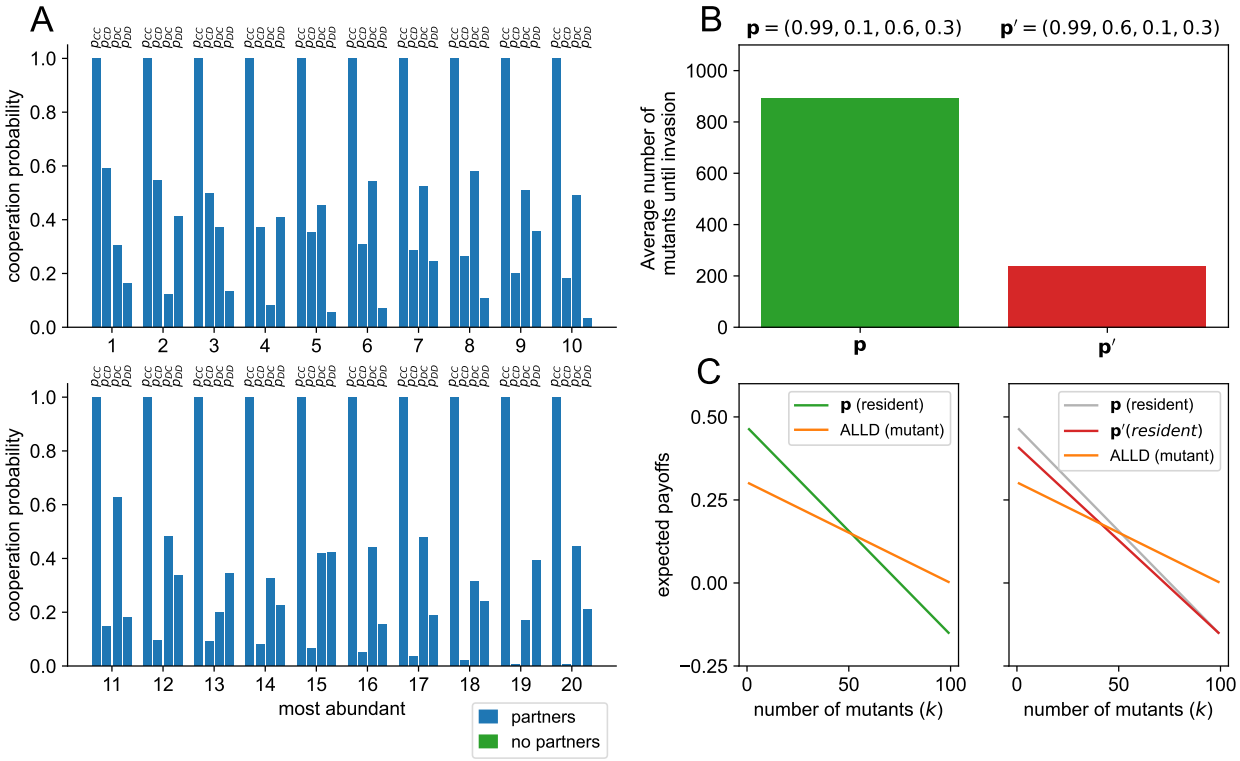

Figure S4: Invasion analysis for reactive-2. A, We repeated the same analysis for the case of reactive-2 strategies. The twenty top-performing strategies are partner strategies. We observe that the partner strategies selected by the evolutionary process almost always have a smaller p_{CD} than p_{DC} and a low p_{DD}. **B,** We already know why residents with low p_{DD} are more abundant. To understand the entries p_{CD} and p_{DC}, we will consider two example strategies: $\mathbf{p} = (0.99, 0.10, 0.60, 0.30)$ and $\mathbf{p}' = (0.99, 0.60, 0.10, 0.30)$. We perform an invasion analysis when either $\mathbf{p}$ or $\mathbf{p}'$ are the resident, and we introduce mutants until a mutant becomes the new resident. We repeat this process 10^4 times and report the average time until invasion over all the repetitions. We observe that $\mathbf{p}'$ is invaded much faster. **C,** We calculate the expected payoffs of $\mathbf{p}$ and $\mathbf{p}'$ when k `ALLD` mutants are present. We observe that the self-payoffs of the strategies are different, which in turn means that fewer `ALLD` mutants are required for `ALLD` to achieve a higher payoff compared to $\mathbf{p}'$.

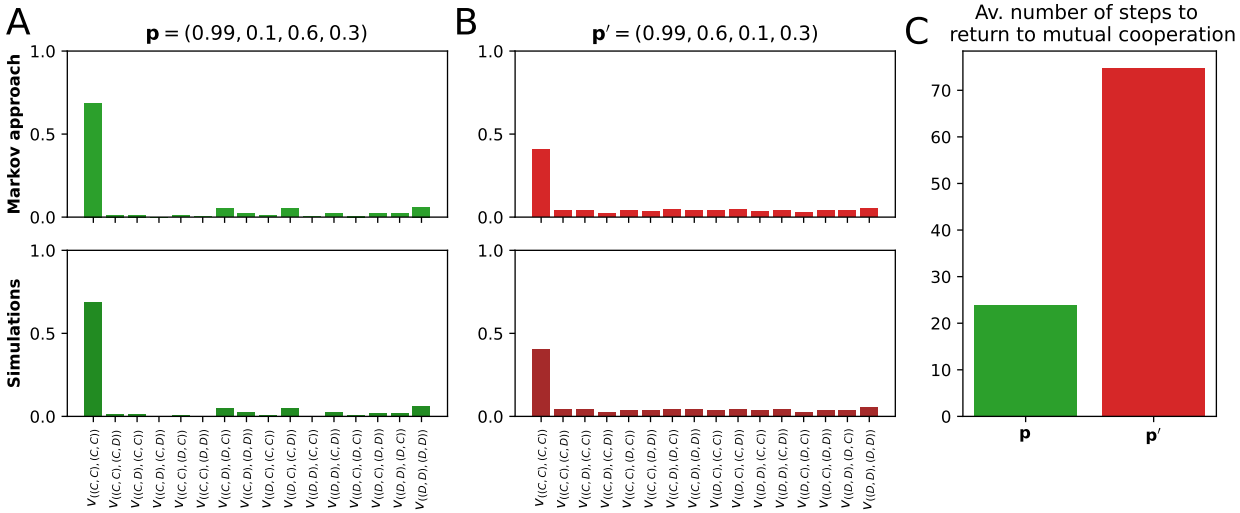

Figure S5: The difference between $\mathbf{p}$ and $\mathbf{p}'$. A-B, The self-payoffs $\pi(\mathbf{p}, \mathbf{p})$ and $\pi(\mathbf{p}', \mathbf{p}')$ would be equal if the strategies were nice. However, such strategies are never sampled by the evolutionary process. To understand the effect of stochastic p_{CC} and p'_{CC}, we calculate the stationary distribution of the self-play of p_{CC} and p'_{CC}. We do this in two ways: numerically and with simulations. For simulating the interactions, we use the open-source package Axelrod. Each interaction is run for 10^6 steps and repeated 20 times. The numerical results and the results from the simulations are in agreement. The stationary distribution of the self-plays shows that p_{CC} spends significantly more time in the $((C, C), (C, C))$ state than p'_{CC}, which explains the difference in the payoffs. **C,** In a self-interaction, the strategies $\mathbf{p}$ and $\mathbf{p}'$ spend most of their interactions in the $((C, C), (C, C))$ state. Once in this state, each player cooperates with a probability of 0.99. However, with a small probability of 0.01, they defect. We can calculate the expected time until the strategies return to the $((C, C), (C, C))$ state once such a "mistake" has occurred. We do this by simulating the interactions and recording the number of steps it takes to return to the $((C, C), (C, C))$ state. We observe that $\mathbf{p}$ needs fewer steps to return to the $((C, C), (C, C))$ state than $\mathbf{p}'$.

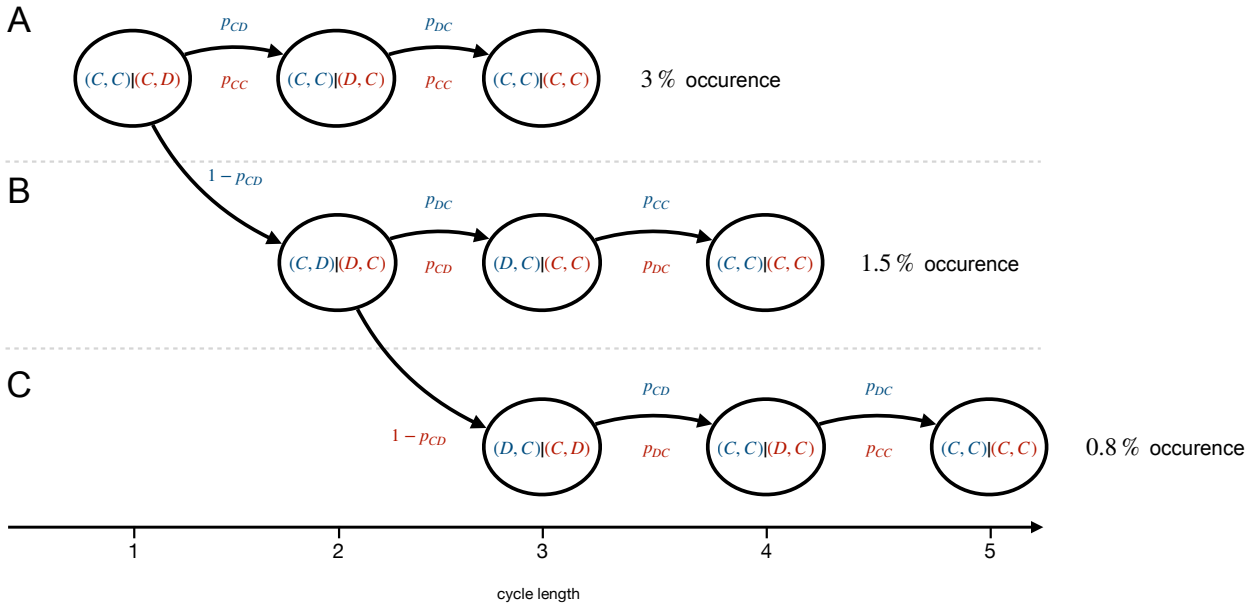

Figure S6: Cycle examples for $\mathbf{p} = (0.99, 0.10, 0.60, 0.30)$. We consider a cycle to be a sequence of states that the strategies go through until they return to the state $((C, C), (C, C))$. We consider two players (blue and red), who are using strategy $\mathbf{p}$. While in the mutual cooperation state, the red player defects. **A,** Cycle A is the most common and shortest possible cycle for both strategies $\mathbf{p}$ and $\mathbf{p}'$. The same is true for the cycle where the blue player defects; both cycles occur with the same frequency. Following a defection by the red player, the fastest way to mutual cooperation is for both players to cooperate twice in a row. Thus, two instances of forgiveness must occur from both players. **B,** In Cycle A, the less likely forgiveness is that of the blue player following the initial defection because p_{CD} is the lowest of all entries with a cooperation probability of 0.1. In Cycle B, we assume that the blue player indeed defected. As a result, this means that the counting towards two forgiveness rounds, for blue, is still at zero. From here, both players could cooperate twice, but an extra step has been added to the cycle, meaning that Cycle B has one step more than Cycle A. **C,** Similarly, in Cycle B, the less likely forgiveness is that of the red player when they cooperate with a probability of p_{CD}. If the red player defects, then their forgiveness counter resets to zero, even if the blue player already forgave the red player once. The fastest thing that can happen is regaining mutual cooperation until $(C, C), (C, C)$, but Cycle C has one extra step compared to Cycle B.

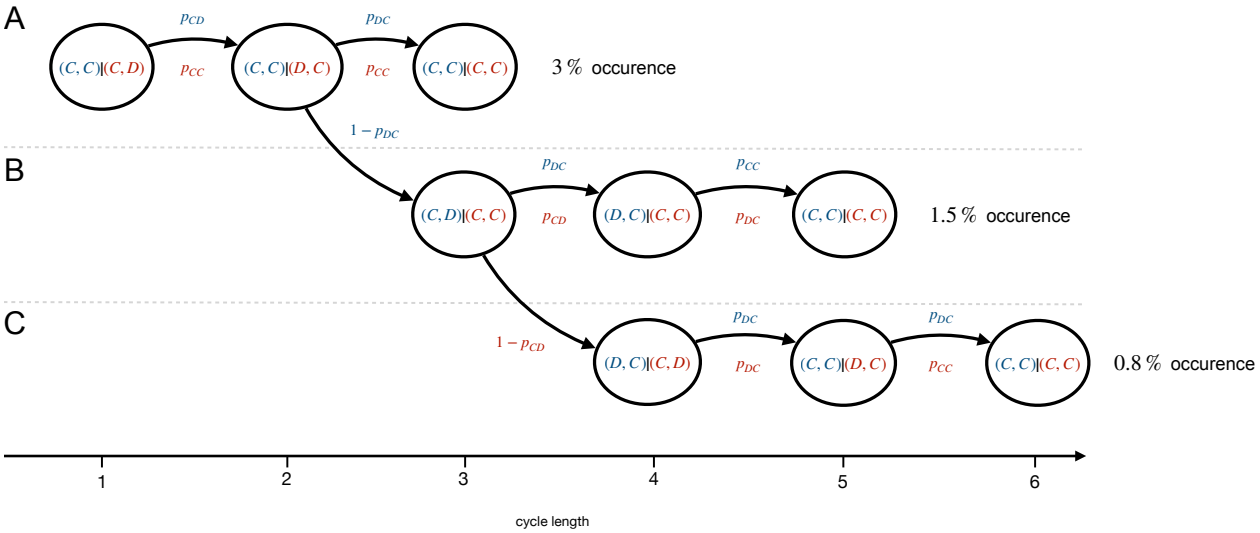

Figure S7: Cycle examples for $\mathbf{p}' = (0.99, 0.60, 0.10, 0.30)$. A similar but slightly different story is true for when players use $\mathbf{p}'$. In Cycle A, the less likely forgiveness to occur is when the blue player has already forgiven the red player once, because now p_{DC} is the lowest cooperation probability of the players. This leads to a new sequence, Cycle B, where the counter is reset, even though it was incremented by one, meaning that the new cycle has two extra steps. The same is true for Cycle C, which has two extra steps compared to Cycle B. In summary, the strategies need to cooperate twice to return to mutual cooperation. There are specific states they must navigate to reach $(C, C), (C, C)$ again. If they fail to cooperate in one of these states, they must try again. However, as these states form a path, the less likely cooperation is further down this path, the more time steps are required to return and attempt cooperation again. This is why $\mathbf{p}'$ takes longer to reach mutual cooperation once more.

3 Errors

So far, we have considered the case where there cannot be a mistake in the actions taken by a player; the actions of the players are realized without error. Here, we discuss what happens in the case where such an error is possible. More specifically, we consider that ϵ is the probability that a player makes a mistake in the action taken.

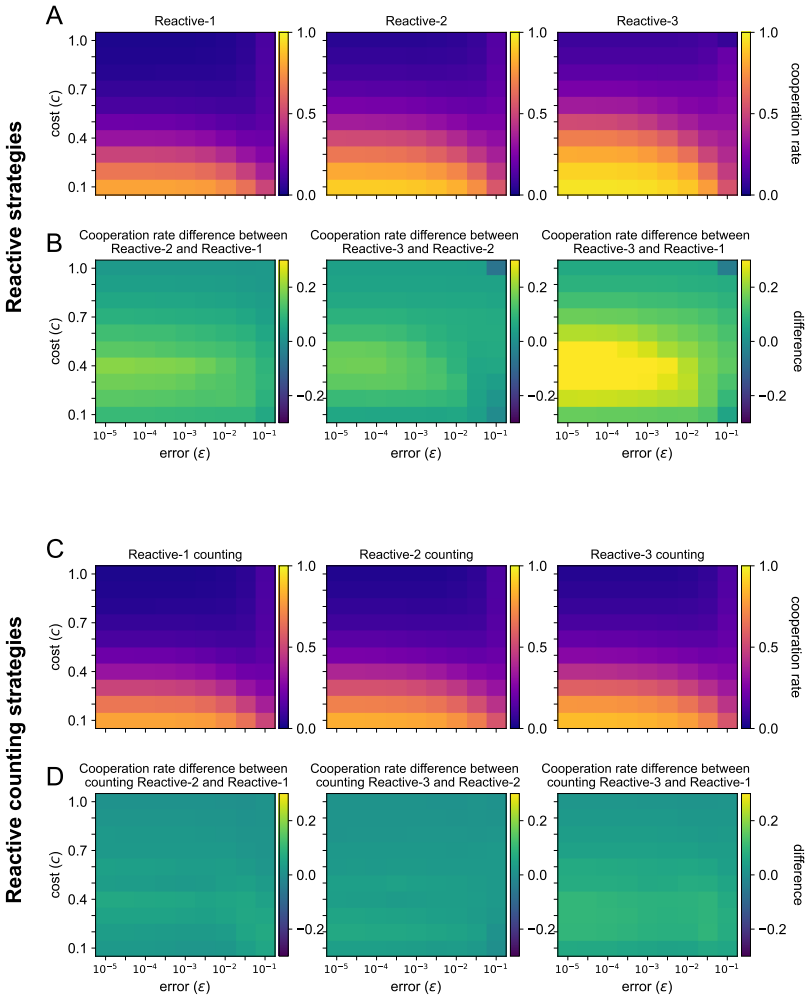

Figure S8: Cooperation rates with implementation errors. We simulate the evolutionary process, this time allowing for implementation errors. Specifically, we consider a probability ϵ that a player makes a mistake in the action taken. We calculate the average cooperation rate for different values of ϵ and c. We do this for reactive strategies **A** and reactive counting strategies **C**. **A, C** We plot the average cooperation rate for the different parameters when individuals use reactive-1, reactive-2, and reactive-3 strategies, respectively. **B, D** We plot the differences between the cooperation rates when individuals use different memory size strategies. From left to right, we show the differences between reactive-1 and reactive-2, reactive-2 and reactive-3, and reactive-1 and reactive-3 strategies.

4 Memory-n

So far in the evolutionary simulations, we have considered reactive strategies and shown that strategies with larger memory allow for more cooperative populations. We repeat the evolutionary simulations, but this time using memory-n strategies. We get results for memory-1, memory-2, and memory counting strategies for n equal to $1, 2$, and 3. The main result still holds. For more memory, more cooperation evolves, however, not in the case of counting strategies.

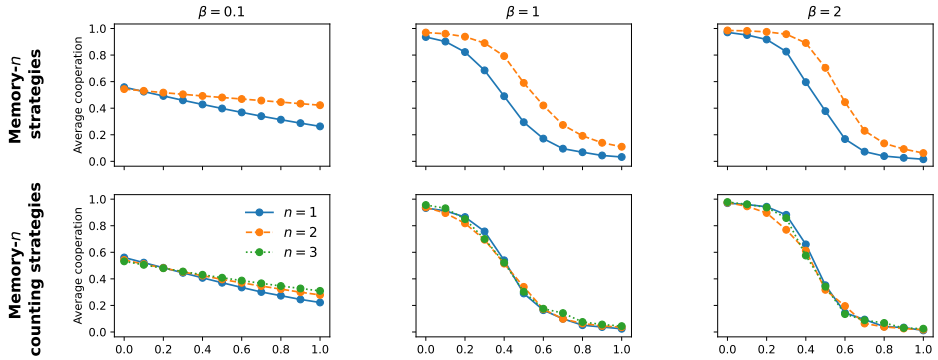

Figure S9: Memory-n simulations. We explore the cooperation rates of memory-n strategies and memory-n counting strategies over different values of cost and selection strength. Simulations are based on a donation game with $b = 1$, $c = 0.5$, a selection strength $\beta = 1$ and a population size $N = 100$, unless noted otherwise. For n equal to 1 and 2, simulations are run for $T = 10^7$ time steps.

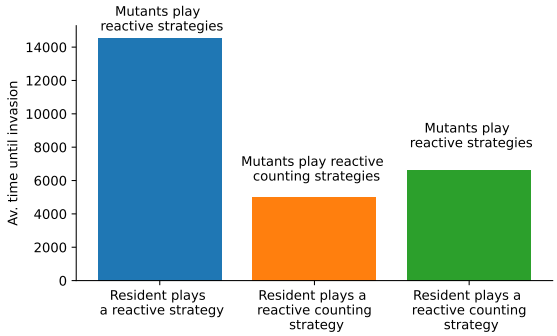

Figure S10: Invasion times for reactive and reactive counting. To understand why counting strategies do not allow for more cooperation, we focus on reactive-2 strategies and reactive-2 counting strategies. For $\beta = 1$ and $c = 0.3$, what we observe is that the top abundant strategies for both simulations are very cooperative strategies, achieving a cooperation rate of ≈ 1. From the top strategies, it is also clear that the average time until they were invaded is higher for reactive strategies. To test this, we perform the following exercise: we pick a top abundant reactive strategy and a reactive counting strategy. These are (0.999985, 0.160136, 0.553336, 0.035629) for the reactive strategy and (0.998204, 0.466715, 0.317605) for the reactive counting strategy. We run an invasion analysis when both strategies are residents. For the reactive counting strategy, we run two invasion analyses: one where only counting reactive mutants are introduced and one with reactive mutants. What we observe is that the reactive strategy can repel more mutants. However, we also see that for the counting strategy, when we consider reactive mutants, the average time until invasion increases. We run 10^3 simulations and take the average.

5 Proofs

5.1 Proof of Theorem 1: Reactive-1 defective Nash strategies in the donation game

Proof. The proof is similar to one for partner strategies. We again enumerate the four pure self-reactive-1 strategies $\tilde{\mathbf{p}}$ by interpreting the strategy as a binary number, we obtain the following payoffs.

$$\begin{aligned}
\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) &= 0 && \text{for } j \in \{0, 2\} \\
\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) &= \frac{b \cdot p_C - c}{p_C + 1} && \text{for } j \in \{1\} \\
\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) &= b \cdot p_c - c && \text{for } j \in \{3\}
\end{aligned}$$

Requiring the payoffs in this list to be at most the mutual defection payoff 0, we get the following unique conditions,

$$p_D \leq \frac{c}{b}.$$

□

5.2 Proof of Theorem 2: Reactive-2 defective Nash strategies in the donation game

Proof. The proof is similar to one for partner strategies. We again enumerate the sixteen pure self-reactive-2 strategies $\tilde{\mathbf{p}}$ by interpreting the strategy as a binary number, we obtain the following payoffs.

$$\begin{aligned}
\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) &= 0 && \text{for } j \in \{0, 2, 4, 6, 8, 10, 12, 14\} \\
\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) &= \frac{p_{CD} + p_{DC}}{3}\, b - \frac{c}{3} && \text{for } j \in \{1, 9\} \\
\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) &= \frac{p_{CC} + p_{CD} + p_{DC}}{4}\, b - \frac{c}{2} && \text{for } j \in \{3\} \\
\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) &= \frac{p_{CD} + p_{DC}}{2}\, b - \frac{c}{2} && \text{for } j \in \{4, 5, 12, 13\} \\
\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) &= \frac{p_{CC} + p_{CD} + p_{DC}}{3}\, b - \frac{2c}{3} && \text{for } j \in \{6, 7\} \\
\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) &= p_{CC}\, b - c && \text{for } j \in \{8, 9, 10, 11, 12, 13, 14, 15\}
\end{aligned}$$

Requiring the payoffs in this list to be at most the mutual defection payoff 0, we get the following unique conditions,

$$p_{CC} \leq \frac{c}{b}, \qquad \frac{p_{CD} + p_{DC}}{2} \leq \frac{1}{2} \cdot \frac{c}{b}, \qquad \frac{p_{CD} + p_{DC} + p_{CC}}{3} \leq \frac{2}{3} \cdot \frac{c}{b}.$$

Because the last condition is implied by the first two, we end up with the conditions in (2). □

5.3 Proof of Theorem 3: Reactive-3 defective Nash strategies in the donation game

Proof. The proof is similar to the previous one. Again, enumerating the 256 pure self-reactive 3 strategies $\tilde{\mathbf{p}}$ by interpreting the strategy as a binary number, we obtain the following payoffs.

$$\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) = 0 \qquad \text{for } j \in \{0, 2, 4, 6, \ldots, 250, 252, 254\}$$

$$\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) = \frac{p_{CDD} + p_{DCD} + p_{DDC}}{4}\, b - \frac{1}{4}\, c \qquad \text{for } j \in \{1, 9, 33, 41, 65, 73, 97, 105, 129, 137, 161, 169, 193, 201, 225, 233\}$$

$$\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) = \frac{p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC}}{5}\, b - \frac{2}{5}\, c \qquad \text{for } j \in \{3, 7, 35, 39, 131, 135, 163, 167\}$$

$$\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) = \frac{p_{CDC} + p_{DCD}}{2}\, b - \frac{1}{2}\, c \qquad \text{for } j \in \{4-7, 12-15, 20-23, 28-31, 68-71, 76-79, 84-87, 92-95, 132-135, 140-143, 148-151, 156-159, 196-199, 204-207, 212-215, 220-223\}$$

$$\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) = \frac{p_{CCC} + p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC}}{6}\, b - \frac{1}{2}\, c \qquad \text{for } j \in \{11, 15, 43, 47\}$$

$$\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) = \frac{p_{CDD} + p_{DCD} + p_{DDC}}{3}\, b - \frac{1}{3}\, c \qquad \text{for } j \in \{16, 17, 24, 25, 48, 49, 56, 57, 80, 81, 88, 89, 112, 113, 120, 121, 144, 145, 152, 153, 176, 177, 184, 185, 208, 209, 216, 217, 240, 241, 248, 249\}$$

$$\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) = \frac{p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC}}{4}\, b - \frac{1}{2}\, c \qquad \text{for } j \in \{18, 19, 22, 23, 50, 51, 54, 55, 146, 147, 150, 151, 178, 179, 182, 183\}$$

$$\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) = \frac{p_{CCC} + p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC}}{5}\, b - \frac{3}{5}\, c \qquad \text{for } j \in \{26, 27, 30, 31, 58, 59, 62, 63\}$$

$$\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) = \frac{p_{CCD} + p_{CDC} + p_{CDD} + p_{DCC} + p_{DCD} + p_{DDC}}{7}\, b - \frac{3}{7}\, c \qquad \text{for } j \in \{37, 67, 165, 195\}$$

$$\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) = \frac{p_{CCC} + p_{CCD} + p_{CDC} + p_{CDD} + p_{DCC} + p_{DCD} + p_{DDC}}{8}\, b - \frac{1}{2}\, c \qquad \text{for } j \in \{45, 75\}$$

$$\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) = \frac{p_{CCD} + p_{CDC} + p_{CDD} + p_{DCC} + p_{DCD} + p_{DDC}}{6}\, b - \frac{1}{2}\, c \qquad \text{for } j \in \{52, 53, 82, 83, 180, 181, 210, 211\}$$

$$\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) = \frac{p_{CCC} + p_{CCD} + p_{CDC} + p_{CDD} + p_{DCC} + p_{DCD} + p_{DDC}}{7}\, b - \frac{4}{7}\, c \qquad \text{for } j \in \{60, 61, 90, 91\}$$

$$\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) = \frac{p_{CCD} + p_{CDC} + p_{DCC}}{3}\, b - \frac{2}{3}\, c \qquad \text{for } j \in \{96-103, 112-119, 224-231, 240-247\}$$

$$\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) = \frac{p_{CCC} + p_{CCD} + p_{CDC} + p_{DCC}}{4}\, b - \frac{3}{4}\, c \qquad \text{for } j \in \{104-111, 120-127\}$$

$$\pi^1(\tilde{\mathbf{p}}_\mathbf{j}, \mathbf{p}) = p_{CCC}\, b - c \qquad \text{for } j \in \{128, 129, 130, \ldots, 255\}$$

Requiring these payoffs to be at most equal to the mutual defection payoff 0 gives several conditions. The necessary conditions for a reactive-3 strategy to be a defective Nash strategy are the following:

$$p_{CCC} \leq \frac{c}{b}, \qquad \frac{p_{CDC} + p_{DCD}}{2} \leq \frac{1}{2} \cdot \frac{c}{b}, \qquad \frac{p_{CDD} + p_{DCD} + p_{DDC}}{3} \leq \frac{1}{3} \cdot \frac{c}{b},$$

$$\frac{p_{CCD} + p_{CDC} + p_{DCC}}{3} \leq \frac{2}{3} \cdot \frac{c}{b}, \qquad \frac{p_{CCD} + p_{CDD} + p_{DCC} + p_{DDC}}{4} \leq \frac{1}{2} \cdot \frac{c}{b}.$$

□