# Max-Planck-Institut für Mathematik in den Naturwissenschaften Leipzig


## Group level effects of social versus individual learning

(revised version: January 2013)

by


*Jürgen Jost and Wei Li*



Preprint no.: 12 2013

# Group level effects of social versus individual learning

Jürgen Jost and Wei Li
Max Planck Institute for Mathematics in the Sciences,
Inselstr.22, 04103 Leipzig, Germany
(Dated: January 29, 2013)

We study the effects of learning by imitating others within the framework of an iterated game in which the members of two complementary populations interact via random pairing at each round. This allows us to compare both the fitness of different strategies within a population and the performance of populations in which members have access to different types of strategies. Previous studies reveal some emergent dynamics at the population level when players learn individually. We here investigate a different mechanism in which players can choose between two different learning strategies, individual or social. Imitating behavior can spread within a mixed population, with the frequency of imitators varying over generation time. When compared to a pure population with solely individual learners, a mixed population with both individual- and social learners can do better, independently of the precise learning scheme employed. We can then search for the best imitating strategy. Imitating the neighbor with the highest payoff turns out to be consistently superior. This is in agreement with findings in experimental and model studies that have been carried out in different settings.

## I. INTRODUCTION

Learning as the exploitation of past experiences to increase the reward in recurring situations is a basic adaptation in higher organisms, reaching its prime in humans. Likewise, in computer science, there is the important field of machine learning. Reinforcement learning, that is, to select actions on the basis of achieved or expected rewards, is a fundamental paradigm. Individual learning, however, may be costly, because it needs some trial and error to find good solutions. In groups of individuals facing similar situations, it might then become advantageous to exploit the learning results of others, that is, to simply copy the actions of successful fellows instead of going through the difficulties and risks of learning oneself [5]. This is called social learning [1]. This issue has been investigated, in particular, by P.Hammerstein and his collaborators [2].

Social learning may even be beneficial for the group, because in that way, successful strategies may spread quickly within the group. There could be a catch, however: When everybody tries to copy somebody else, then nobody will learn anything new [6]. Thus, ultimately, imitation might no longer constitute a group beneficial strategy. So, always some individuals seem to be needed that are willing to try and explore something new. There must be "information producers" (individual learners), not just "information scroungers" [7, 8]. These information producers, however, on average might perform worse than the imitators because the latter can reap the benefits without encountering the errors and risks. Therefore, with evolutionary selection based on accumulated rewards, imitation may spread in a population at the expense of exploration. Thus, the population may end up in the dilemma of generating an ever larger fraction of imitators.

However, in some game theoretical analyses, imitators perform better than individual learners only when fellow imitators are rare [8]. At the extreme, when there are no individual learners, no information will be sampled, which then makes copying a strategy that has lower fitness than individual learning [8, 9]. This might lead to the prediction that the relative fractions of individual and social learners may settle at some nontrivial equilibrium. Precisely the fraction of individual lerners at equilibrium equals the fraction of expected gain from individual learners. Whether this equilibrium is group beneficial, however, is another question.

In principle, one could also lift imitation to a higher level, that is, players need not simply imitate other players, but they might imitate strategies. A second order imitator would then use a learning or copying strategy depending on which looks more successful. Of course, this can then be iterated.

But let us return to the basic setting. How can we test and make quantitative such verbal reasoning as displayed above? Here, we employ the complementary population game scheme that we have developed in [3, 4] to analyze the interplay between individual optimization and emergent group level dynamics. In short, we have two populations whose members are repeatedly randomly paired in a symmetric game. This game possesses several Nash equilibria, some being better for the members of one, and the others for those of the other population. The members of each group compete with each other in terms of their fitness function that determines how many clones (with a suitable mutation rate) that can provide to the next generation. Individual fitness maximization, however, can lead to a global dynamics that leaves the own population at an inferior equilibrium.

Thus, we can test at the same time which learning strategy, exploration or imitation, might spread in a population, and how good or bad that will be for the population as a whole. In particular, we may break the symmetry between the two populations by allowing the respective members access to different types of strategy spaces, that is in the present setting, to different learning schemes.

## II. THE POPULATION GAME AND THE STRATEGY SPACES

In the basic complementarity game of [3, 4], we have two populations, called buyers and sellers, whose members are randomly matched in each round. A buyer offers an amount $x$, and its partner at the present round, namely a random seller, asks for $y$, with both ranging between 0 and $K$ (an integer which can be taken as large as one wishes). If $x \geq y$ then a deal occurs, and the seller gains $y$ and the buyer $K - x$; otherwise no deal is concluded and both gain nothing. Thus, for instance a buyer should ideally offer as little as possible so that the deal is just concluded, that is, he should not offer more than the seller will ask or is expected to ask, and it would be good for the buyer to induce the seller to ask little. Analogously for the sellers. Thus, both players wish to drive as hard a bargain as possible, but if they push too hard the interaction could fail, which they certainly wish to avoid. If neither of them is allowed to learn or coordinate and all the conditions are equal, the game should be totally symmetric and might then stay at the symmetric Nash equilibrium $K/2$ forever. We should note, however, that any value between 0 and $K$ is a Nash equilibrium as it will pay off for neither player to deviate when the opponent sticks to that value. In that sense, the game is degenerate. Our key point is to break this degeneracy and the symmetry between the two populations by evolutionary elements. That is, the outcomes of some fixed rounds of interactions will be compared and accordingly the next generation will be constructed by some evolutionary scheme (operated by some evolutionary algorithm or a more elaborate genetic algorithm). There is plenty of room to break the intrinsic symmetry between the two populations. For instance, they can have different random mutation rates or they can be assigned different strategy options. By examining the resulting equilibrium value, one can check which side is doing better: if that value is larger than $K/2$ then the sellers are better off, otherwise the buyers. In general we find that simpler and more flexible strategies lead to superior results at the population level because they can process the information in a more efficient way, which will speed up the convergence rate. Also, a population with some bolder players may gain an advantage over one with timid players only.

For individual learning, we can also construct a number of strategies based on the types of information to be used. A rather naive one is that players select random offers which will be updated after some rounds of interactions. More elaborated strategies can use players' past experience, say the average value of what the opponents offer in the previous $m$ rounds. This is the direct use of information from another population. Or, they could also use friends' experience, say the average value of what their friends or friends' opponents offer in the previous rounds. (Here, "friends" are defined as neighbors in some network, and the population performance may depend on the topology of that network, see [3].) This is the indirect use of information from another population. We have observed that both direct and indirect use of information from another population can lead to an efficient population level performance . Also, each strategy can have two variants, either directly employing the value computed according to the chosen strategy as the next own offer, or using that value as the input in a look-up table whose output then is that next offer. The look-up table then is itself an object of evolution. In fact, since the look-up table has $K$ input entries and has to provide an output for each of them, evolution will take quite some time to test it out thoroughly. When $m$ is large, a simple averaging strategy without using look-up tables performs even better than the same type of strategy which uses look-up tables. When $m$ is 1, the strategy only takes into account what the opponents offer in the most recent round and such a strategy, named 1-round opponent for short, turns out to be a rather efficient strategy when employing look-up tables. In this strategy, only limited information is used. Without specific notation, in the present paper, the individual learners adopt the 1-round opponent strategy. To have a stable setting, in our simulations with evolving the look-up tables, the generation length, the selection percentage and the mutation rate are 100, 0.5 and 0.01, resp. The population size is chosen to be 1000 and the maximum offer available, i.e. $K$, is 50.

There are some issues for evaluating different learning strategies. First of all, it depends on the type of problem that has to be solved by learning. Here, this is given by the setting of the population game, with the twist, however, that the members of the opposite population may also learn, and so, learning has to chase a moving target. Secondly, it may depend on the individual composition of the populations. For simplicity, we initialize our populations to consist of equal numbers of individual learners and imitators. We keep the total population sizes constant, and so, the fractions of the two types will change by our fitness based selection scheme. Thirdly, it will depend on the individual learning strategies employed or in principle accessible within the strategy space assigned to each population. In particular, we can give the two populations different strategy spaces, that is, possible learning schemes, and then see which one is better for the population. Forthly, the copying scheme also makes a difference. The crucial question for a copyer is whom to copy. Again, different schemes are conceivable, and we can compare them within our setting.

More precisely, we shall evaluate the following copying strategies: (S1) copy a random fellow player; (S2) copy a random fellow player who was successful in the most recent round; (S3) copy from several random fellow players the one with the highest number of successful rounds; (S4) copy from several random fellow players the one with the highest payoff; (S5) copy the neighbor with the highest number of successful rounds; (S6) copy the neighbor with the highest payoff; (S7) copy from several random fellow players one whose number of successful rounds is above average; (S8) copy from several random fellow players one whose payoff is above average; (S9) copy from several random fellow players the one that has been imitated most; (S10) copy the neighbor who has been imitated most.

We should point that in all these strategies an imitator can imitate a learner or another imitator. What counts is only the performance of the imitated player.

## III. SIMULATION RESULTS

Players who learn individually are abbreviated as learners and those who learn socially, imitators. In principle, we may allow the players to change their roles according to performance. They could be learners at one time and imitators at another time.

As stated in the previous sections, we have two major goals in studying the behavior of imitating in our game. First, will a mixed population with both imitators and learners do better than a pure population with only learners? How will the imitators spread inside the population? Second, which copying rule is the best? Can we rank their performance in a consistent way?

As stated, in the simulations of this work, learners will employ 1-round opponent strategy which can only do well when combined with look-up tables. Imitators, however, can decide whether to use look-up tables or not. A more straight consideration is that imitators just directly copy the values from those they are imitating, without wasting time in dealing with the values again by checking the look-up tables which also need time to evolve. For comparison we also check another version in which imitators use the copied values as inputs for outputs extracted from their respective look-up tables. These two different versions will of course lead to some difference but do not change radically our main conclusions.

We first let imitators directly copy the values of the players they are imitating. We compare a mixed population, with half learners and half imitators initially, to a pure population with only learners. For the mixed population, the relative frequencies of learners and imitators will change over generation time as players will only produce offsprings of their own types. That is, better performing imitators (learners) will generate more imitators (learners). But of course, the two populations depend on each other in the sense that learners will make use of the offers from their opponents and imitators will copy the offers from fellow learners. As seen in Fig. 1, when the buyers are a pure learning population and the imitators in the seller population employ imitating strategy S6, namely copying the neighbor with the highest fitness, the mixed population can achieve an advantage (the equilibrium of 27 is greater than 25, the symmetric, ideal equilibrium). We should also notice that the pure population is converging to the equilibrium in a more steady way. But for the mixed population the process is far more complicated as there are very steep ups and downs. These ups and downs may correspond to some good values and some bad values imitators happen to choose, as they do not copy fixed individuals. But the mixed population eventually achieves a better equilibrium for itself. Of course, we observe similar effects when roles of buyers and sellers are reversed, that is, when the buyers are a mixed population and the sellers are pure, see Fig. 2. The effect ia slso seen for other copying strategies.

As explained, by our selection scheme, the relative frequencies of imitators and learners will change according to their relative success. In Fig. 3, we show the change of frequency of imitators, who adopt imitating strategy S6 (copying the neighbor with the highest fitness), over the number of generations. Initially the frequency increases from 0.5 to around 0.88 (the first peak) after 60 generations, as learners still have to discover good solutions and imitators can copy the best available ones. Then the frequency decreases to around 0.43 (the first bottom) after 100 generations, as learners gradually improve. The frequency increases again to 0.65 (the second peak) after 125 generations, then decreases to 0.22 after 140 generations (the lowest ever) and then climbs up to 1.0 after nearly 170 generations. Eventually imitators take over the mixed population in which all the players are imitators. Does this contradict our previous prediction that a population full of imitators cannot preform well at the population level? In fact this is not the case. In Fig. 1 which has the same setting as Fig. 3, we see that the equilibrium has been reached after almost 80 generations (which corresponds to the first bottom in Fig. 3). Thus, when equilibrium is reached, the dominating players inside the mixed population are still learners. Since after this point there is no longer an average fitness difference between imitators and learners, the randomness of copying behavior may allow imitators to spread and take over the whole population. In Fig. 4, we see however that when copying strategy S5 (copying from several random individuals that with the highest number of successes) the learners will always stay in the majority.

In order to compare imitation strategies directly, we use pairwise competitions, so-called round-robin tournaments of which we need 45 for 10 strategies, or 90 if we want to make it symmetric between buyers and sellers. In any such competition, the winning strategy, i.e., that for which the equilibrium value is better, gets 2, the losing one 0

TABLE I: Win-Loss ratio for different pair of imitating strategies. Here the values in the brackets denote the numbers that the given two strategies tie among the 200 simulations. All the imitators do not employ look-up tables.

| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 |
|---|---|---|---|---|---|---|---|---|---|---|
| S1 | | 100:100 | 99:101 | 100:100 | 200:0 | 0:200 | 100:99(1) | 199:0(1) | 100:100 | 2:178(20) |
| S2 | 100:100 | | 146:54 | 0:100(100) | 102:97(1) | 99:101 | 199:1 | 100:100 | 100:100 | 1:199 |
| S3 | 101:99 | 54:146 | | 153:1(54) | 109:51(40) | 0:200 | 102:1(97) | 200:0 | 63:137 | 1:196(3) |
| S4 | 100:100 | 100:0(100) | 1:153(46) | | 200:0 | 100:100 | 100:0(100) | 200:0 | 100:100 | 102:33(65) |
| S5 | 0:200 | 97:102(1) | 51:109(40) | 0:200 | | 1:199 | 99:101 | 97:7(96) | 97:102(1) | 17:183 |
| S6 | 200:0 | 101:99 | 200:0 | 100:100 | 199:1 | | 200:0 | 200:0 | 100:100 | 101:0(99) |
| S7 | 99:100(1) | 1:199 | 1:102(97) | 0:100(100) | 101:99 | 0:200 | | 5:101(94) | 0:200 | 21:155(24) |
| S8 | 0:199(1) | 100:100 | 0:200 | 0:200 | 7:97(96) | 0:200 | 101:5(94) | | 0:200 | 101:96(3) |
| S9 | 100:100 | 100:100 | 137:63 | 100:100 | 102:97(1) | 100:100 | 200:0 | 200:0 | | 99:97(4) |
| S10 | 178:2(20) | 199:1 | 196:1(3) | 33:102(65) | 183:17 | 0:101(99) | 155:21(24) | 96:101(3) | 97:99(4) | |

TABLE II: Win-Loss ratio for different pair of imitating strategies. Here the values in the brackets denote the numbers that the given two strategies tie among the 200 simulations. All the imitators employ look-up tables.

| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 |
|---|---|---|---|---|---|---|---|---|---|---|
| S1 | | 83:75(42) | 106:48(46) | 110:43(47) | 48:100(52) | 41:100(49) | 67:74(59) | 73:72(55) | 112:40(48) | 44:108(48) |
| S2 | 75:83(42) | | 106:37(57) | 100:55(45) | 56:98(46) | 40:122(38) | 73:76(51) | 77:68(55) | 105:44(51) | 51:97(52) |
| S3 | 48:106(46) | 37:106(57) | | 83:74(43) | 62:72(56) | 62:82(56) | 58:96(46) | 52:110(38) | 97:75(28) | 64:81(55) |
| S4 | 43:110(47) | 55:100(45) | 72:62(56) | | 65:89(46) | 50:91(59) | 36:112(52) | 34:113(53) | 83:79(38) | 75:86(39) |
| S5 | 100:48(52) | 98:56(46) | 72:62(56) | 89:65(46) | | 66:89(45) | 105:48(47) | 128:45(37) | 98:53(49) | 82:67(51) |
| S6 | 100:41(49) | 122:40(38) | 82:62(56) | 91:50(59) | 89:66(45) | | 97:55(48) | 107:45(48) | 87:54(59) | 85:68(47) |
| S7 | 74:67(59) | 76:73(51) | 96:58(46) | 112:36(52) | 48:105(47) | 55:97(48) | | 92:62(46) | 113:43(44) | 42:115(43) |
| S8 | 72:73(55) | 68:77(55) | 110:52(38) | 113:34(53) | 45:128(37) | 45:107(48) | 62:92(46) | | 96:49(55) | 53:87(60) |
| S9 | 40:112(48) | 44:105(51) | 75:97(28) | 79:83(38) | 53:98(49) | 54:87(59) | 43:113(44) | 49:96(55) | | 60:96(44) |
| S10 | 108:44(48) | 97:51(52) | 81:64(55) | 86:75(39) | 67:82(51) | 68:85(47) | 115:42(43) | 87:53(60) | 96:60(44) | |

points, and in case of a tie, each gets 1. For each type of tournament we run 100 different simulations (with different realizations) so the total number of simulations is 9000. In the following table I we summarize the statistics of win-loss ratio for any pair of imitating strategies.

The statistics yields a ranking of performance of strategies according to our rule: the winner receives 2 points and the loser , 0; if they tie both receive 1. The resulting ranking (in decreasing order) is: S6, S4, S9, S10, S3, S2, S1, S5, S8 and S7. The best imitating strategy is copying the neighbor with the highest payoff and it almost defeats all other strategies (except that it ties S4). S4 is copying from several random fellow players the one with the highest payoff. The most successful imitating strategies are all payoff-biased, essentially because pay-off is our selection criterion. In fact, the payoff-biased imitating strategy has been confirmed in experiments [10] to be the second frequently used strategy (its percentage could reach more than 20 percent), only next to individual learning. S6 performs better better than S4 probably because it samples from a larger pool of other players, and perhaps also that it avoids that small groups of players learn from each other suboptimal strategies. S9 and S10 also do reasonably well, that is, copying a player that is copied by many others might make sense because such a player in the first place needs to survive manyselection rounds itself. Random copying strategies S3, S2 and S1 come next, whereas S8 and S7, that is, above-average-based copying, is worst.

When imitators also use look-up tables, the performance differences between strategies become smaller and the ranking changes somewhat, but S6 still remains best, see Table II.

## IV. CONCLUSION

In this paper we show the role of imitating behavior inside a complementarity game in which the members of two populations interact with each other via random pairing. Players can be learners or imitators, that is, either try to find a good solutions by themselves or to identify others that have found a good solution and simply copy them. We find that imitators can spread within a mixed population with both types of players, and consequently a mixed population outperforms a pure population with only learners. We also compare the efficiency of a list of 10 different imitating rules which represent different sources of information. Our round-robin simulations indicate that the best strategy is copying the neighbor with the highest payoff. Of course, we could also investigate other schemes, for instance, let different imitation schemes compete within a population. Probably, this will not affect the qualitative findings of this paper.


[1] L. Rendell, L. Forgarty and K.N. Laland, Rogers' paradox recast and resolved: population structure and the evolution of social learning strategies, Evolution **64-2**, 534 (2009).
[2] Hammerstein, P., ed. (2003). Genetic and Cultural Evolution of Cooperation. Cambridge, MA: MIT Press
[3] J. Jost and W. Li, Individual strategies in complementarity games and population dynamics, Physica A **345**, 245-266 (2005).
[4] J. Jost and W. Li, Learning, evolution and population dynamics, Advances in Complex Systems **11**, 901 (2008).
[5] R. Boyd and P.J. Richerson, *Culture and Evolutionary Process* (University of Chicago Press, Chicago, 1985).
[6] K.N. Laland, Social learning strategies, Learning and Behavior **32 (1)**, 4 (2004).
[7] L.-A. Giraldeau, T.J. Valone, and J.J. Templeton, Potential disadvantages of using socially acquired information, Philosophical Transactions of the Royal Society of London **357**, 1559 (2002).
[8] C.J. Barnard and R.M. Sibly, Producers and scroungers: a general model and its application to adptive flocks of house sparrows, Animal Behavior **29**, 543 (1981).
[9] A.R. Rogers, Does biology constrain culture? American Anthropologist **90**, 819 (1988).
[10] A. Mesoudi, An experimental comparison of human social learning strategies: payoff-biased social learning is adaptive but underused. Evolution and Human Behavior **32**, 334-342 (2011).

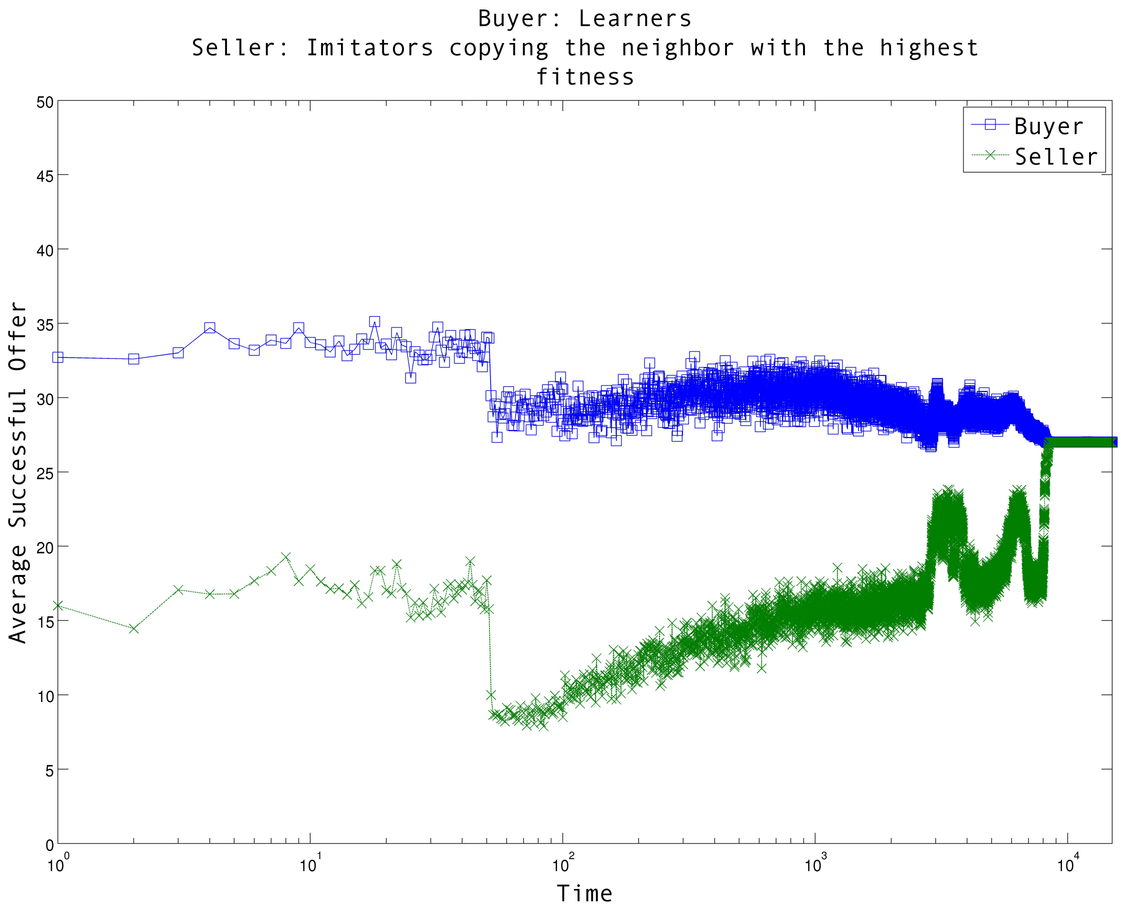


FIG. 1: Average successful offer versus time. The square dots represent buyers and the stars represent sellers. Buyers are pure learners and sellers are mixed with learners and imitators. Here imitators directly copy the neighbor with the highest payoff. The eventuall equilibrium is 27, which is more advantageous to sellers, the mixed population. The pure population learns more quickly than the mixed one does as there are ups and downs due to imitating behaviors. But the mixed population still manages to defeat the pure one.

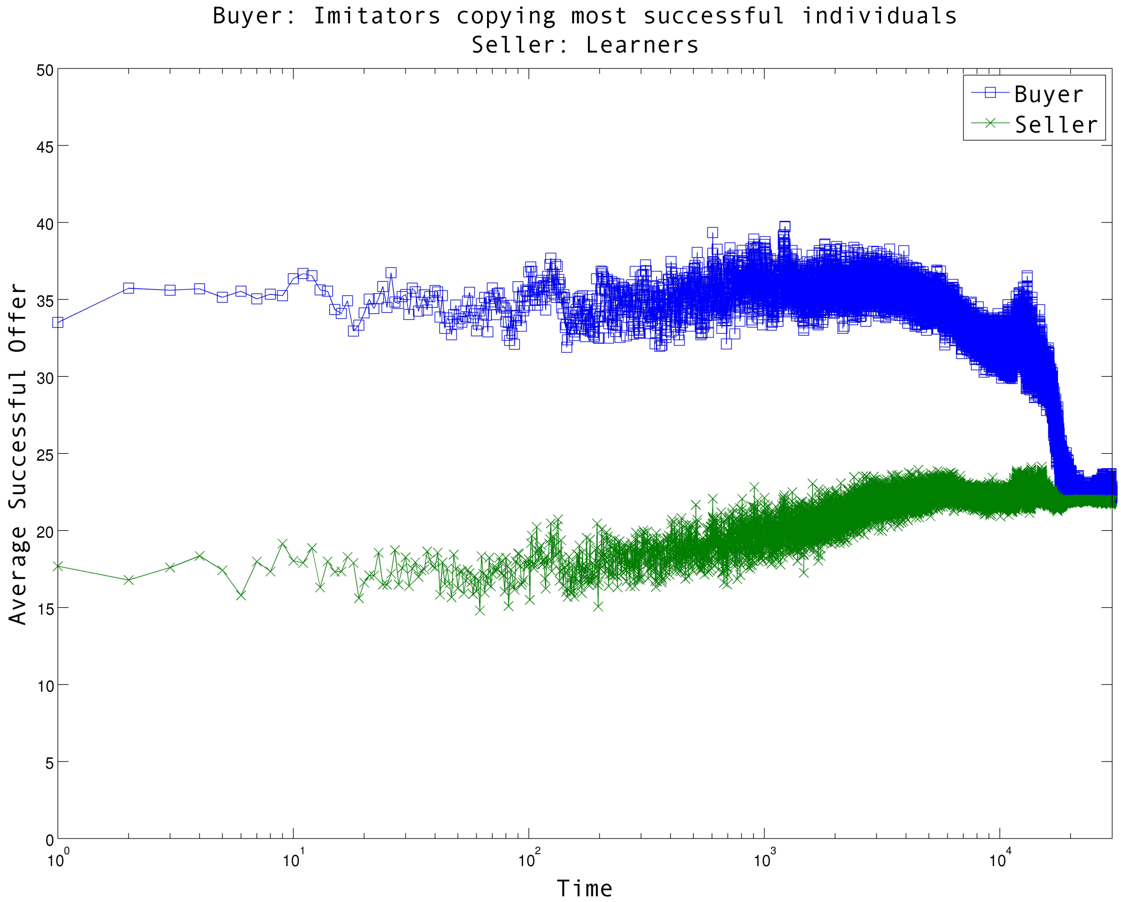


FIG. 2: Average successful offer versus time. The square dots represent buyers and the stars represent sellers. Sellers are pure learners and buyers are mixed with learners and imitators. Here imitators directly copy from random fellow players the one with the highest number of being successful. The eventual equilibrium is 23, which is more advantageous to buyers, the mixed population. Similar to Fig. 1, the mixed population defeats the pure one.

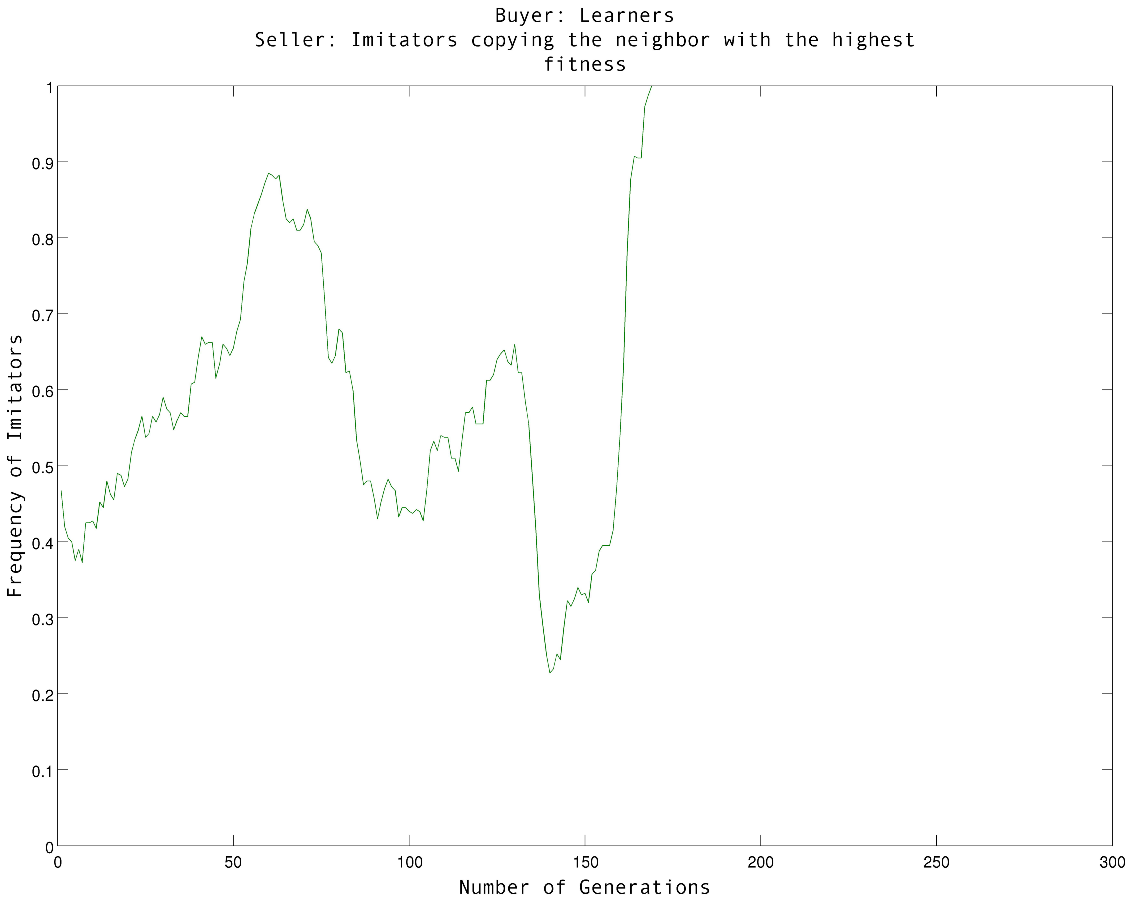


FIG. 3: Frequency change of imitators with respect to number of generations. Buyers are pure learners and sellers are mixed with learners and imitators. Here imitators directly copy the neighbor with the highest payoff. The frequency of imitators steadily increases from 0.5 to 0.88 after nearly 60 generations. The first bottom of frequency, 0.43, occurrs at the 100-th generation. The frequency decreases further down to 0.22 after 140 generations before climbing up to 1.0 after 170 generations. Eventually imitators take over the whole population.

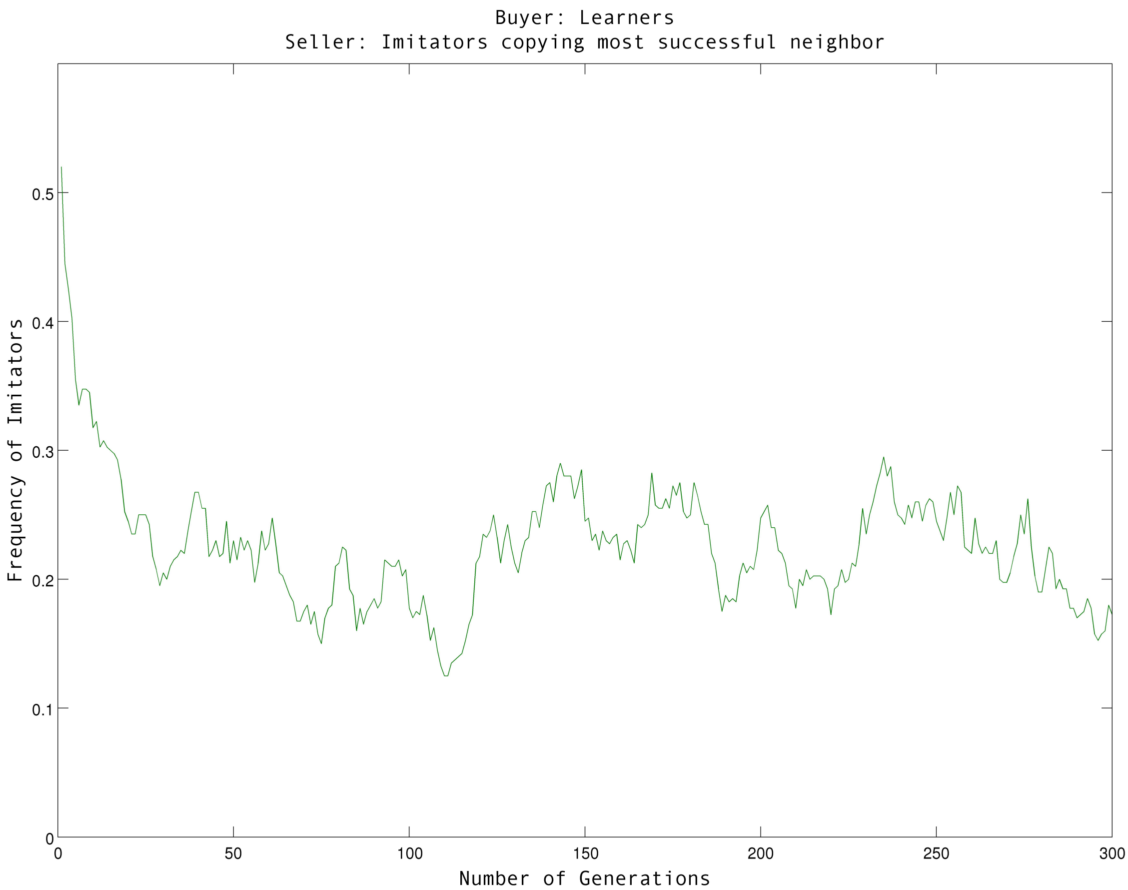


FIG. 4: Frequency change of imitators with respect to number of generations. Sellers are pure learners and buyers are mixed with learners and imitators. Here imitators directly copy from random fellow players the one with the highest number of being successful. Unlike in Fig. 3, the frequency of imitators decreases from 0.5 to 0.2 and then is never greater than 0.3. The reason is quite simple, that is, the strategy imitators employ is not efficient. Therefore individual learners are always the majority.