

Agent-Based Modeling for Studying the Spontaneous Emergence of Money

Mattia Di Russo
Université Côte d’Azur, CNRS, I3S
Sophia Antipolis, France
dirusso.mattia@gmail.com

Zakaria Babutsidze
SKEMA Business School
Sophia Antipolis, France
zakaria.babutsidze@skema.edu

Célia da Costa Pereira
Université Côte d’Azur, CNRS, I3S
Sophia Antipolis, France
celia.da-costa-pereira@univ-cotedazur.fr

Maurizio Iacopetta
SKEMA Business School
Sophia Antipolis, France
maurizio.iacopetta@skema.edu

Andrea G. B. Tettamanzi
Université Côte d’Azur, Inria, CNRS, I3S
Sophia Antipolis, France
andrea.tettamanzi@univ-cotedazur.fr

Abstract—A central question in economics is how a society accepts money, defined as a commodity used as a medium of exchange, as an unplanned outcome of the individual interactions. This question has been approached theoretically in the literature and investigated by means of agent-based modeling. While an important aspect of the theory is the individual’s speculative behavior, that is, the acceptance of money despite a potential short-term loss, previous work has been unable to reproduce it with boundedly rational agents. We investigate the reasons for the failure of previous work to have boundedly rational agents learn speculative strategies. Starting with an agent-based model proposed in the literature, where the intelligence of the agents is guided by a learning classifier system that is shown to be capable of learning trade strategies (core strategies) that involve short sequences of trades, we test several modifications of the original model and we come up with a set of assumptions that enable the spontaneous emergence of speculative strategies, which explain the emergence of money even when the agents have bounded rationality.

Index Terms—Search and Money, Reinforcement Learning, Social Simulation

I. INTRODUCTION AND RELATED WORK

While there is disagreement among economists about the origin of money, some have conjectured that it emerged spontaneously from trading activities among selfish agents with bounded rationalities, without any particular supervision [9], [11], [13], [19], causing a certain number of goods to become acceptable to all actors in the trade, i.e., to become *money*.

Due to its complexity, this problem has been tackled using agent-based modeling, which is the computational study, carried out by means of simulation, of economies that are modeled as evolving systems of autonomous interacting agents [12], [15]–[17]. An agent-based model makes it possible to study the interaction among agents, even when these have considerable differences in preferences and skills.

In the seminal model of Kiyotaki and Wright [8] rational agents of different types trade their goods, with the ultimate goal of consuming their preferred good; stationary Nash equilibria are studied to check whether particular commodities can emerge as a medium of exchange. Duffey *et al.* [4] conducted

experiments aimed at providing as close an approximation of the Kiyotaki-Wright environment as possible in the laboratory. The results showed that “subjects tend to play ‘fundamental’ strategies rather than ‘speculative’ strategies, even in contexts where speculative strategies would yield better gains.” Marimon *et al.* [10] expand Kiyotaki and Wright’s study by assuming artificially intelligent agents instead of rational ones, and making them learn their strategies adaptively, by modeling them as classifier systems [6], [18]. They found that the system always converges to a stationary Nash equilibrium, but in economies with multiple equilibria, the only equilibrium that emerges is the one with low storage goods playing the role of money (Kiyotaki and Wright’s “fundamental equilibrium”).

Babutsidze and Iacopetta [1] study the conditions under which individuals can learn to play speculative strategies to accept a high-storage-cost good as money, whether the agents are fully or boundedly-rational. Their experimental study concluded that, contrary to what Kiyotaki and Wright suggested, when artificial agents or real people are asked to make decisions in such a trade environment, the low-return object is rarely accepted as money. They also pointed out some shortcomings of Marimon *et al.*’s model [10], which call for an analysis and motivate our proposal of an improvement of the agents’ learning capabilities so that they can learn complex strategies.

After accurately replicating the different models proposed in the literature, we thus extend them and make further experiments in order to improve the agents’ faculty of learning speculative strategies, an objective that is lacking in [1] and [10]. Our goal is to find out possible limitations of the agents’ artificial intelligence. The results we obtain are very promising, even if more can still be done on this subject.

II. ECONOMIC MODEL

The Economic Model we use goes back to Kiyotaki and Wright [8]. For a more detailed discussion of the economic environment, we direct the reader to some relevant recent literature in Economics [3], [7].

Time is discrete and infinite and, at each time step, there are three indivisible commodities called *Good 1*, *Good 2* and *Good 3*. There is a continuum of infinitely lived agents of three types, with equal proportions of *Type 1*, *Type 2*, and *Type 3*, that specialize in both consumption and production: type i agents derive utility only from the consumption of Good i and are able to produce only Good $i^* = i + 1 \pmod{3} \neq i$.

All goods are storable at a cost, but agents can store only one unit at a time and, since goods are indivisible, only one good at a time. Storage costs are type- and good-specific. Let c_{ij} denote the cost (in terms of instantaneous disutility) to Type i of storing Good j . The assumption is made that $c_{i3} > c_{i2} > c_{i1} > 0$ for all i .

Type i 's expected discount lifetime utility is given by

$$E = \sum_{t=0}^{\infty} \beta^t \left[I_i^U(t)U_i - I_{i^*}^D(t)D_i + \sum_{j=1}^3 I_{ij}^c(t)c_{ij} \right], \quad (1)$$

where U_i denotes the instantaneous utility from consuming Good i , D_i the instantaneous disutility from producing good i^* , $\beta \in (0, 1)$ is the discount factor (common across types), $I_i^U(t)$ is an indicator function that equals one if the agent "eats" its consumption Good i , zero otherwise; $I_{i^*}^D(t)$ equals one if it produces its production Good i^* , zero otherwise; and $I_{ij}^c(t)$ equals one if it stores any good j , zero otherwise, at the end of period t .

We assume that the net utility of consuming plus producing, $u_i = U_i - D_i$, is large enough that agents will not want to drop out of the economy, which is always the case if

$$u_i > (c_{ii^*} - c_{ik})/(1 - \beta), \quad \text{for all } i, k. \quad (2)$$

At time t , if Type i is lucky enough to acquire its consumption Good i , it will consume it and produce a new unit of Good i^* . Thus each Type i always has an inventory of exactly one unit of one good other than Good i . At each period, agents are randomly matched in pairs and must decide whether or not to trade bilaterally. Trade always entails a one-for-one swap of inventories, given the physical environment, and occurs if and only if it is mutually agreeable.

A. Agent Behavior

Each individual chooses a trading strategy to maximize its expected discounted utility, taking as given the strategies of other agents and the inventory distribution $p(t)$. Notice that while this is true in the original model, our goal is to make the agents learn strategies when they are boundedly rational, that is, they don't have knowledge of the other agents' behaviors and of the distribution of goods among all the agents. A trading strategy is a rule determining the circumstances under which an agent is willing to trade.

B. Equilibrium

In this environment, the goal is to define the agent's behavior in order to reach a steady-state Nash equilibrium, i.e., a set of trading strategies, one for each type i , together with a steady-state distribution of inventories p , such that (i)

the strategy chosen by each individual maximizes its expected utility given the strategies of the others and the distribution p , and (ii) no agent has anything to gain by changing only their own strategy. The goal is to characterize equilibrium for different particular specifications of the production and consumption specialties.

Let $V_i(j)$ be the expected discounted utility for Type i when it exits a trading opportunity with Good j , given that it follows a maximizing strategy; that is, $V_i(j)$ is the indirect utility of leaving with good j . When i exits with its own consumption Good i , it consumes it and immediately produces a new unit of Good i^* , which yields the instantaneous utility $u_i = U_i - D_i$, plus the indirect utility of storing Good i^* , i.e., $V_i(i) = u_i + V_i(i^*)$. The indirect utility for Type i of storing Good $j \neq i$ is described by Bellman's equation of dynamic programming [2]:

$$V_i(j) = -c_{ij} + \max_j \beta E[V_i(j') | j], \quad (3)$$

where $E[V_i(j') | j]$ is the expectation of V_i at next period's random state j' , conditional on j , and the maximization is over strategies.

In equilibrium, agents of the same type never trade, since neither can prefer what the other has. Under Assumption 2, each Type i will accept Good i , consume it, and produce a new unit of Good i^* whenever it has the opportunity. That is, for all i , $\max_j V_i(j) = V_i(i) = u_i - V_i(i)$. Since Type i always wants to consume Good i and produce Good i^* , a trade always occurs whenever Type i with Good j meets Type j with Good i .

An algorithm for finding equilibria is described in [8], which essentially performs an exhaustive search over the strategy profiles to completely characterize the set of equilibria. For certain parameter values, there will exist one equilibrium, referred to as a *fundamental* equilibrium since agents always prefer a lower-storage-cost commodity to a higher-storage-cost commodity, unless the latter is their own consumption good. For other parameter values there exists another equilibrium, referred to as a *speculative* equilibrium, since sometimes agents trade a lower- for a higher-storage-cost commodity not because they wish to consume it, but because they rationally expect that this is the best way to ultimately trade for another good that they do want to consume, that is because it is more marketable.

C. Fundamental Strategy

The fundamental strategies are described by $V_i(i) = \max_j V_i(j)$ for all i (agents always prefer their consumption good) and the inequalities $V_1(2) > V_1(3)$, $V_2(1) > V_2(3)$, and $V_3(1) > V_3(2)$ (otherwise they prefer higher-storage-cost goods).

Fundamental strategies are always the best response for Types 2 and 3 agents that only look at fundamentals for and best for Type 1 iff $c_{13} - c_{12} > (p_{31} - p_{21})bu_1$, with $b = \beta/3$ and p_{ij} equal to the probability for an agent of Type i to own a good of type j .

The steady-state inventory distribution p may be summarized here by three numbers (since $p_{ii} = 0$ and $\sum_j p_{ij} = 1$

for all i). For these fundamental strategies, this is given by $(p_{12}, p_{23}, p_{31}) = (1, \frac{1}{2}, 1)$, and therefore these strategies constitute an equilibrium iff $c_{13} - c_{12} > \frac{1}{2}bu_1$.

Types 1 and 3 always keep their production goods until they can trade directly for their consumption goods, never using indirect trade. Type 2 agents trade their production Good 3 for Good 1 whenever possible, and end up holding each exactly half of the time. They thereby act as middlemen, transferring Good 1 from Type 3 to Type 1 agents.

Good 1 is the unique medium of exchange, or commodity money, in this equilibrium.

D. Speculative Strategy

On the other hand, if $c_{13} - c_{12} < (p_{31} - p_{21})bu_1$, fundamental play by all agents does not constitute an equilibrium. The best response by Type 1 to fundamental play, in this case, is to speculate by attempting to trade Good 2 for Good 3, which has a higher storage cost but is also more marketable. Fundamental play is still the best response by Types 2 and 3 and, therefore, the strategies corresponding to $V_i(i) = \max_j V_i(j)$ and the inequalities $V_1(2) < V_1(3)$, $V_2(1) > V_2(3)$, and $V_3(1) > V_3(2)$ also constitute an equilibrium in some other region of the parameter space.

To sum up, Types 2 and 3 should indeed use fundamental strategies when Type 1 speculates, and Type 1 should speculate iff $c_{13} - c_{12} < (p_{31} - p_{21})bu_1$. The inventory distribution implied by these strategies is $(p_{12}, p_{23}, p_{31}) = (\frac{\sqrt{2}}{2}, \sqrt{2}-1, 1)$, and so speculative equilibrium obtains iff $c_{13} - c_{12} < (\sqrt{2} - 1)bu_1$.

By engaging in speculation, Type 1 agents now also play the role of middlemen in some trades, transferring Good 3 from Type 2 to Type 3. Type 2 agents are still middlemen in other trades, and they continue to use Good 1 as a medium of exchange, while Type 1 uses Good 3 as a medium of exchange. In this equilibrium, we therefore have dual commodity monies, with both the most storable and the least storable goods (i.e., Goods 1 and 3) used to achieve indirect trade in different instances by different individuals.

No other set of strategies is consistent with equilibrium and, in the intermediate region, where $(\sqrt{2} - 1)bu_1 < c_{13} - c_{12} < \frac{1}{2}bu_1$, no pure strategy, steady-state equilibrium exists in which all agents of the same type play the same strategy.

E. Boundedly Rational Agents

Until now we have talked about an environment where individuals are fully informed about what other agents do, and on the current state of the economy. Now, following [1], we will analyze the behavior of agents that do not have perfect foresight. Hence, agents are thought of as having a limited capacity to take actions on the basis of forecasted future states of the economy.

One way to implement agents with bounded rationality is to have them make decisions through a classifier system, as in Marimon et al.'s [10] implementation of Kiyotaki and Wright's model.

In the simplest version of Marimon et al.'s [10] model, agents possess two interdependent classifier systems, one to guide their trading decisions, and the other to guide their consumption decisions. The exchange classifier maps the pre-trade system into a trading decision, while the consumption classifier maps the post-trade state into a consumption decision.

Conditions on the pre-trade asset holdings of the two agents can be described by a ternary string of 6 digits. Each digit expresses if the given agent holds (1) or does not hold (0) the given product; the "don't care" symbol # matches either state. The first three digits describe the inventory of one trader, the following three describe the inventory of its counterpart. For example, string 100 0## means the first trader has Good 1 in stock, while its counterpart does not have Good 1. After eliminating the redundant alternatives, all possibilities are captured by 72 trade classifiers.

Consumption classifiers are modeled similarly to the trade classifiers. They represent a string of ternary digits describing the post-trade holdings of the consumer, plus a binary value for the "consume / do not consume" decision. Elimination of the redundant alternatives yields 12 distinct consumption classifiers.

The time-line of one period in the computational model is as follows. An agent is matched with a randomly selected counterpart. It observes its own asset holdings as well as those of the counterpart. All matching trade classifiers are short-listed and enter the "auction" system, which ranks the classifiers according to their strength and generality (i.e., how many #s they contain). The agent takes the action indicated by the top-ranking exchange classifier. It either proposes to exchange the asset holdings, or refuses to do so. In case both agents agree to exchange their assets, a trade takes place. Following this, post-trade holding is realized and consumption classifiers enter a similar "auction". The top-ranking classifier gives directions on whether to consume or not. In the case of consumption, the agent costlessly produces its production good.

Storage costs (c_i) for the good the agent holds at the end of the period are subtracted from the utility (u_i) the agent has received during this time period. This net utility affects the strengths of the classifiers that brought the system to its current condition. Most notably, these are the current consumption classifier and the current trade classifier. However, this might also involve other trade classifiers, in particular, those that contributed toward the trade sequence that culminated in the current consumption decision. At this point the period is over and agents enter a new time period with updated asset holdings and classifier strength vectors ($S_{a_i}(t+1)$).

A genetic algorithm [5] is used to optimize a non-exhaustive set of classifiers, which makes it possible to maintain only a subset of all the possible rules. The algorithm adds to the classifier system a random initialization of the population of rules and a mechanism by which such population evolves, which results from the application of four operators: *creation*, which creates a new classifier from scratch when no classifier matches the current state; *mutation*, designed

to inject sufficient diversity into the range of actions called for by different classifiers in a given situation; *specialization* and *generalization*, called randomly with a probability that decreases over time, just after the winning bid has been determined.

This completes the description of the economic model studied in [8], [1] and [10].

III. LEARNING SPECULATIVE STRATEGIES

Experiments conducted to verify that boundedly-rational agents equipped with a Learning Classifier System could learn fundamental and speculative strategies under the right circumstances, as the theoretical result suggest, gave mixed results [1], [10]: in an economy with low utility reward from consumption, where their behavior should converge to a fundamental equilibrium, the agents are able to learn how to exchange correctly, while in a high-utility scenario, where some agents should learn speculative strategies, they failed at this task.

Given those premises, our objective was to investigate the assumptions that enable the emergence of speculative strategies in all scenarios.

A. Validation

As a preliminary step, we tested the framework we implemented based on the model described above to check if the agents have the correct behavior when they have to learn the fundamental strategy. Those experiments were performed by assuming that the agents are equally distributed, i.e., the population consists of 1/3 of agents of Type 1, 1/3 of agents of Type 2 and 1/3 of agents of Type 3.

Figure 1 shows the distribution of goods owned by fully-rational agents of Type 1, 2 and 3, respectively, showing convergence to $(p_{12}, p_{23}, p_{31}) = (1, \frac{1}{2}, 1)$, as expected per Section II-C.

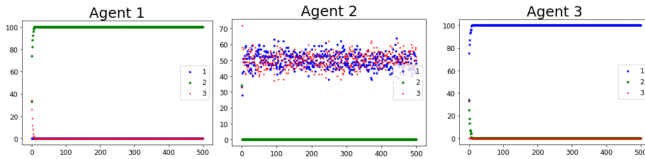


Fig. 1. Distribution of goods for the three agents using fundamental strategies.

On the x -axis we have the number of agents holding a certain good, while on the y -axis we have the time period. For Type 2 we do not have a flat line corresponding to exactly 50% of the agents holding Good 1 and 50% holding Good 3, as the theoretical study predicts, because, contrary to the theoretical model, here we have a finite number of agents (100 agents of each type).

Figure 2 shows the distributions when agents of Type 1 play speculative while Type 2 and 3 play fundamental (which is the result we aim at).

As already mentioned in Section II-D, when agents play those kinds of strategies, the system has reached an equilibrium on the state $(p_{12}, p_{23}, p_{31}) = (\frac{\sqrt{2}}{2}, \sqrt{2}-1, 1)$. This means

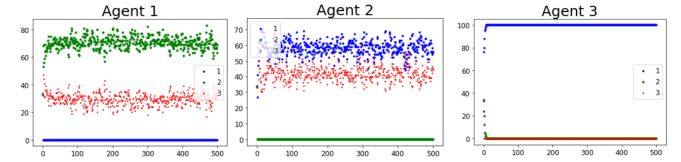


Fig. 2. Distribution of goods for the three types of agents when agents of Type 1 play speculative while Type 2 and 3 play fundamental.

that after reaching the Nash equilibrium, agents of Type 1 should hold Good 2 with 70% probability (consequently, they should hold Good 3 with 30% probability, since agents cannot hold their consumption good), agents of Type 2 should hold Good 3 with 41% probability, and agents of Type 3 should hold their production Good 1 100% of the time.

Let's now move on to the experiments with boundedly-rational agents. As anticipated, the classifier system itself works well when it has to make agents learn fundamental strategies only. This is shown in Figures 3 and 4. After 500 time periods and 30 model runs, we obtain an overall decision accuracy of 71% without the genetic algorithm and 78% with the genetic algorithm. This outcome is explained by the fact that there is some over-consumption in agents of Type 1 and 3. By accuracy, we mean the amount of right decisions over the total decisions made.

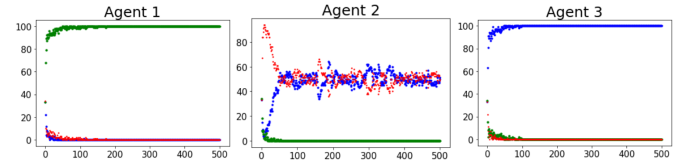


Fig. 3. Distribution after a model run using only classifier system

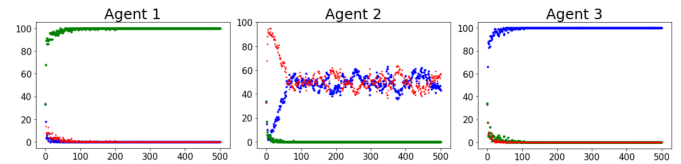


Fig. 4. Distribution after a model run using only classifier system and genetic algorithm

B. Simplifications

The shortcoming of this classifier system (including when using the genetic algorithm under incomplete enumeration of rules) is that it gives the results shown above, even when we know that agents of Type 1 should learn to play speculative strategies. We thus began by changing some assumptions as proposed by [1], aiming at simplifying the implementation described in Section II-E.

The first simplification was to eliminate the agent's learning about consumption preferences, as it is more interesting to analyze the agents' trading behavior, while it is realistic to assume that an agent knows what good it should consume.

After removing the learning of consumption preferences, an interesting fact emerged: it was over-consumption that caused

agent ID	agent type	own	counter-part	decision	strength	firing count
100	2	#0#	0##	0	-0.05	1
100	2	#0#	0##	1	9.38	207

TABLE I

STRENGTH AND FIRING COUNT OF TWO SAMPLE RULES OF A TYPE 2 AGENT, WITH ANTECEDENT #0# 0## AND TWO OPPOSITE CONSEQUENTS (DECISIONS).

the distributions seen in Figures 3 and 4. It turns out that the reason lies behind the nature of the learning classifier system itself. Rules like #0# 0## 1, i.e., “if I don’t hold Good 2 and my counterpart does not hold Good 1, then I’m willing to trade”, tend to be misleading for an agent, because there are situations in which this is a good rule and situations in which it is not. This situation is depicted in Table I, where we extracted information for a specific agent of Type 2 about the classifiers that reflect this situation. Here we expected that the two rules would have both been used a fair amount of times, while instead the one leading to the “no trade” decision has been used only once and then dropped. The rule leading to the “trade” decision has been used 207 times, even if we know that it is not always the best one. This limitation must be taken into account when interpreting the results. From now on, we will consider the results for the classifier system version that includes the genetic algorithm, since it always gives us the best results.

This issue notwithstanding, this simplification increases the overall accuracy from $\sim 75\%$ to an average of 91% , due the disappearance of incorrect consumption decisions. This accuracy is reached when the system is in a state in which every agent should play fundamental strategies. Figure 5 shows the distributions of goods after removing the consumption classifier system.

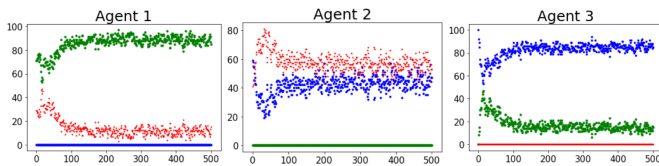


Fig. 5. Distribution after a model run after removing the consumption classifier system (fundamental equilibrium)

However, if we set the model with parameters and a distribution of goods that should lead to the use of speculative strategies, we do not get anything different from what we just saw in Figure 5, as everybody still plays fundamental. If we consider that they should play speculative strategies, the accuracy of Type 1 agent decisions is only 23% .

The second simplification is about the choices of Type 2 and 3 agents. As it is the behavior of Type 1 agents only that determines whether Good 3 emerges as money, we can impose that the trading choices of Type 2 and 3 agents always conform with the fundamental strategies. However, this simplification only yields a 3% gain in accuracy, as does making Type 1 agents always accept Good 1 (their consumption good) in order to reduce the amount of noise in the learning process:

the system still converges to a fundamental equilibrium like the one depicted in Figure 6.

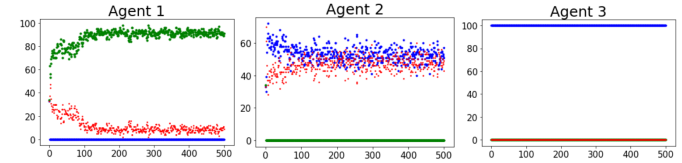


Fig. 6. Distribution after a model run after applying all the simplifications

C. Loop avoiding

These results suggest that a deeper cause for the failure of the model to learn speculative behaviors have to be sought for.

Since the learning process goes through a multitude of unrelated classifiers, agents can be “deceived”. For example, a Type 1 agent has to simultaneously learn not to trade Good 2 for 3, and to trade Good 3 for 2. Given the structure of the classifier system, these two learning processes go in parallel and there is no feedback across them. Specifically, learning a Nash strategy is decomposed into learning multiple (in this case two) actions simultaneously and can lead to conflicting actions. More generally, the design of the classifier prevents the agent from forming a global view of the relative trading opportunities and could induce it to miss profitable patterns.

For this reason, when we reward a rule, we want to remove a fraction of this reward from its opposite rule (which we will call OPP from now on). Our first try was then to subtract/add from OPP’s strength half the strength that is respectively added/subtracted from the considered rule. Adding a negative amount to the strength means reducing the possibility for the rule to be chosen. An example of a rule that gets its strength naturally reduced is when an exchange occurs, but there is no consumption: the agent incurs a cost but gets no utility.

Even though this OPP approach did not solve the problem, it increased the performance by 26% , yielding an accuracy of 59% . The results in Figure 7 may look similar to the ones in Figure 6, but it can be noticed that we are moving towards the expected distribution of Figure 2.

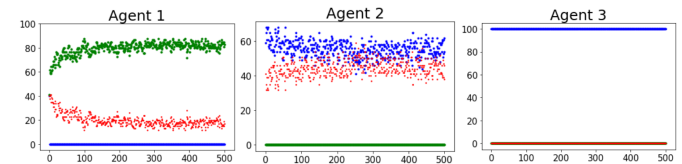


Fig. 7. Distribution after a model run after with OPP update

Since the problem appears to be that contradictory rules can be learned, we should look at its cause, which is payouts are rewarded only after a trading sequence eventually leads to consumption. Every rule that is used pays a bid that is summed to all the other bids from other rules. When consumption occurs, those bids are equally distributed across all classifiers that were involved in the trading process that led to consumption.

This leads to the insight that the reason why agents do not learn to play speculative strategies is that these involve longer trading chains, but rules that do not contribute or work against the final outcome are rewarded in the same way as the favorable rules are. This situation happens when there is a complete trade cycle. For instance, say Type 1 agent is holding Good 2 at some period. In the next period, it trades it for Good 3, then later it trades it back to Good 2 and finally it trades for Good 1 and consumes. Now, if the reward is equally distributed over all classifiers, the action to trade Good 2 for Good 3, which went clearly in the wrong direction, will be rewarded, whereas it should not. How to deal with such situations were never addressed in [10].

We propose to solve this issue by introducing an exponential smoothing of rewards, which assigns exponentially decreasing weights to older rule firings, according to the following formula:

$$y_{t+1|t} = \alpha y_t + \alpha(1 - \alpha)y_{t-1} + \alpha(1 - \alpha)^2 y_{t-2} \dots \quad (4)$$

The one-step-ahead forecast for time $t + 1$ is a weighted average of all of the past observations in series y_1, \dots, y_t . The rate at which the weights decrease is controlled by a smoothing parameter $0 \leq \alpha \leq 1$. If α is small (i.e., close to 0), more weight is given to observations from the more distant past. If α is large (i.e., close to 1), more weight is given to the more recent observations [14]. Using this approach, we get the best results with $\alpha = 0.8$, which yields an 87% of accuracy from Type 1 agents; the results are plotted in Figure 8.



Fig. 8. Distribution after a model run after applying exponential smoothing with $\alpha = 0.8$

The last step was to avoid the second and third simplifications (see Section III-B), to see if now the agents can learn autonomously when to use a fundamental or speculative strategy. The result is that the average accuracy of decisions when the parameters suggest that all the agents should play fundamental, is still near the one when we removed the consumption classifier (old: 91%; new: 90%). When Type 1 agents should play speculative while Type 2 and 3 agents should play fundamental, the accuracy is now 84%, with a slight reduction from 87%, because now Type 2 and 3 agents can make errors.

The introduction of the exponential smoothing of rewards seems to be the key assumption leading to the correct learning of speculative strategies, for it allows a majority of agents to learn the correct strategy also when a speculative equilibrium is optimal.

IV. CONCLUSIONS

We investigated why, in the models of [1] and [10], agents did not always learn speculative strategies. We identified an

issue with the distribution of rewards along the classifiers involved in a chain of trades. The exponential smoothing of rewards appears to be the key factor responsible for the model being able to reproduce speculative behavior as expected.

While the modified model reaches a satisfactory level of accuracy (87%), it is still not perfect. Other adjustments might enable it model to get even closer to 100% accuracy. One promising approach would be to increase the expressivity of the classifiers, for instance to enable them to recognize which type of agent the counterpart is. Indeed, rules that prove to be beneficial when trading with one type of agent may turn out to be deleterious when trading with another type of agent.

Investigating the influence of the size of the economy (i.e., the number of agents) and different mixes of the three types on the convergence to the equilibria is another interesting direction, as are various generalizations of the model.

REFERENCES

- [1] Zakaria Babutsidze and Maurizio Iacopetta. The Emergence of Money: Computational Approaches with Fully and Boundedly Rational Agents. *Computational Economics*, March 2019.
- [2] Dimitri P. Bertsekas. *Dynamic Programming and Stochastic Control*. Academic Press, Inc., USA, 1976.
- [3] F. Bonetto and M. Iacopetta. A dynamic analysis of Nash equilibria in search models with fiat money. *Journal of Mathematical Economics*, 84(C):207–224, 2019.
- [4] John Duffy and Jack Ochs. Emergence of money as a medium of exchange: An experimental study. *American Economic Review*, 89(4):847–877, 1999.
- [5] David E. Goldberg. *Genetic Algorithms in Search, Optimization & Machine Learning*. Addison-Wesley, Reading, MA, 1989.
- [6] John H. Holland. *Adaptation in Natural and Artificial Systems*. The University of Michigan Press, Ann Arbor, 1975.
- [7] Maurizio Iacopetta. The emergence of money: A dynamic analysis. *Macroeconomic Dynamics*, 23(7):2573–2596, 2019.
- [8] Nobuhiro Kiyotaki and Randall Wright. On money as a medium of exchange. *Journal of Political Economy*, 97(4):927–954, 1989.
- [9] Masaaki Kunigami, Masato Kobayashi, Satoru Yamadera, Takashi Yamada, and Takao Terano. A doubly structural network model and analysis on the emergence of money. In *Simulating Interacting Agents and Social Phenomena*, pages 137–149. Springer Japan, 2010.
- [10] Ramon Marimon, Ellen McGrattan, and Thomas Sargent. Money as a medium of exchange in an economy with artificially intelligent agents. *Journal of Economic Dynamics and Control*, 14(2):329–373, 1990.
- [11] Carl Menger. On the origins of money. *History of Economic Thought Articles*, 2:239–255, 1892.
- [12] David Minarsch, Marco Favorito, Ali Hosseini, and Jonathan Ward. Trading agent competition with autonomous economic agents. In *AAMAS*, pages 2107–2110. International Foundation for Autonomous Agents and Multiagent Systems, 2020.
- [13] André Orléan. *The Origin of Money*, pages 113–143. Springer Netherlands, Dordrecht, 1992.
- [14] Simple exponential smoothing.
- [15] Paolo Serafino, Carmine Ventre, and Angelina Vidali. Truthfulness on a budget: trading money for approximation through monitoring. *Auton. Agents Multi Agent Syst.*, 34(1):5, 2020.
- [16] Mike Shann, Alper T. Alan, Sven Seuken, Enrico Costanza, and Sarvapali D. Ramchurn. Save money or feel cozy?: A field experiment evaluation of a smart thermostat that learns heating preferences. In *AAMAS*, pages 1008–1016. ACM, 2017.
- [17] Leigh Tesfatsion. Agent-based computational economics: modeling economies as complex adaptive systems. *Information Sciences*, 149(4):263–269, 2003.
- [18] Ryan Urbanowicz and Jason Moore. Learning classifier systems: A complete introduction, review, and roadmap. *Journal of Artificial Evolution and Applications*, 2009, 09 2009.
- [19] Ayumu Yasutomi. The emergence and collapse of money. *Physica D: Nonlinear Phenomena*, 82(1):180–194, 1995.