

## Social and Asocial Learning in Collective Action Problems: The Rise and Fall of Socially-Beneficial Behaviour

Peter R. Lewis and Anikó Ekárt

*Aston Lab for Intelligent Collectives Engineering (ALICE)*

*Aston University, Birmingham, UK*

{p.lewis,a.ekart}@aston.ac.uk

**Abstract**—The allocation of common-pool resources is an important topic in technical and socio-technical systems, and when left unmanaged, such systems often collapse to highly unequal and unsustainable outcomes. Recent work has highlighted a role for electronic institutions in managing such resources, to ensure socially-beneficial outcomes in the long term. However, open self-organising multi-agent systems often involve agents that learn behaviours in order to meet their goals. In this paper we explore the interplay between institutional features and forms of social and asocial learning employed by self-interested agents. We show that, while recent results have associated social learning with sustainability, this is sensitive to the form of social learning used. We show that more realistic models that combine social and asocial learning are more likely to lead to unsustainable institutions and anti-social outcomes. However, a key role for pardons in the sanction mechanism of the institution is identified, which allows for tolerance of a range of behaviours associated with ongoing learning, including complacency and exploration.

### 1. Introduction

This paper studies one case of the application of social concepts in a technical system: the bottom-up formation and sustenance of institutions for the collective and socially-beneficial management of common-pool resources. Specifically, when an institution can be used to resolve a common-pool resource problem in some socially-desirable way (e.g., fairly), we ask: under what conditions do agents learn to comply with such an institution, and under what conditions does such learning itself cause the institution to fail? We consider this question in the case of different forms of purely social learning, as well as a combined social/asocial case.

Recent work by Pitt et al. [1], [2] has shown that, when faced with the problem of sustainable and fair common-pool resource management in (socio-)technical systems, the presence of an *electronic institution* with the requisite features for sustainability proposed by Ostrom [3] can avoid the free-riding problem. Pitt proposed a computational formalisation of Ostrom’s institutional requirements, which may be *designed in* to the system. As they put it, in this case there is “no need to ‘hope’ that an institution with the requisite features for sustainable management would evolve” [2]. In open and heterogeneous technical and socio-technical systems, however, can we do better than simply hoping a collective of learning agents evolve a sustainable institution

for socially-beneficial common-pool resource allocation? And if so, under what conditions does this happen?

This paper focusses on a common-pool resource appropriation problem, where agents have the ability to take a desired amount of resource from a common pool, but do so in a first-come-first-served manner. In this scenario, the first agents (in our case, the order is fixed) have a greater ability to take their desired amount, while later ones may be unsatisfied. We are interested in how such a system might self-organise to a more equal appropriation of the resource.

Therefore, how might this play out if instead agents’ compliance (or not) with the institution is learnt? How do different forms of learning affect the sustainability of an institution? In particular, we draw on understanding of social learning behaviour in animals, where the common assumption that claims learning by copying to be beneficial, is known to be flawed [4]. Theoretical analyses show that social learners have higher fitness than asocial learners only when copying is rare. As copying increases, copiers copy other copiers and the sampling of the environment by asocial learners is low. It is suggested that *individuals must be more selective in their copying and perform asocial learning some of the time*. Further, “individuals appear to switch between reliance on social and asocial sources of information in a flexible and facultative manner” [5, p369]. Laland [4] investigates strategies that animals adopt for copying through two aspects, the circumstances WHEN copying will happen and WHO, the choice of individual from whom to learn.

Our results highlight an important interplay between the institution’s rules and the agents’ learning behaviour. The way in which agents learn has an impact on the existence and sustainability of the institution. Conversely, the institution’s features can either tolerate or inhibit agent learning.

### 2. Problem Statement

In line with much work in multi-agent systems and the study of the evolution of cooperation, we assume that the system is defined by a population of agents, occupying some common environment, and self-interest on the part of agents, which may conflict with the group’s interests, such that this encourages free riding. We consider the case when resources are exogenous, i.e., they are provided by the environment over time. An example of this is agents appropriating water for their needs, from a reservoir [1]. The collective task facing the agents is to allocate this resource among them.

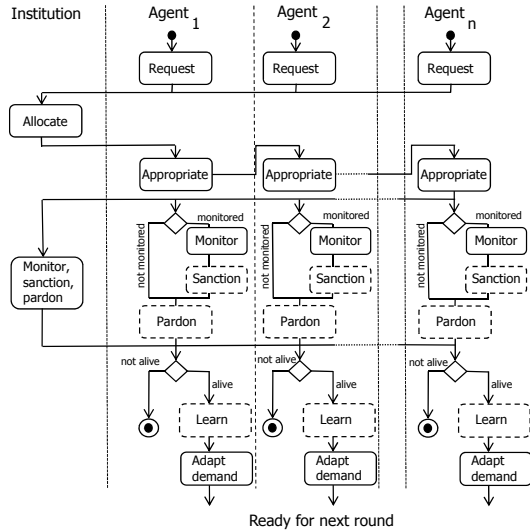


Figure 1. The operation of one round in our system

The included institutional principles [3] are:

- 1) Clearly defined boundaries: members can make demands on resources and then appropriate them; (living) non-member agents can appropriate resources.
- 2) Congruence between rules and the state of the prevailing environment: the *ration* rule is applied, where a member demanding more than its ration is allocated its ration, whereas a member demanding less than its ration will be allocated that.
- 3) Monitoring: the appropriation behaviour of a proportion (10%) of living agents is monitored.
- 4) Graduated sanctions: monitored agents who are caught not complying (i.e. appropriating more) are sanctioned according to their status: their appropriated amount is confiscated, and they are forbidden from demanding for increasing amounts of time, ultimately leading to a permanent ban / death. When institution also pardons agents, each sanction level expires after a period of time (10 rounds).

We consider a scenario based on [1], [3] as illustrated in Figure 1. A round starts by agents making their requests, followed by the institution making allocations using a ration rule, and the agents appropriating in a first-come-first-served manner, as determined by their individual compliance. Monitoring of compliance takes place next: when a monitored agent is caught not complying, graduated sanctions are applied. If the institution contains the pardon rule, expired sanctions are removed. In the learning scenarios, remaining live agents then learn according to the social/asocial strategy. Finally, live agents adapt their demand for the next round following a simple rule, whereby when the appropriated amount equals the demand, the demand is slightly increased for the next round with a small probability (0.1), otherwise the demand is similarly reduced with the same probability.

The problem is then to establish (1) the need for an institution under which there is an equitable appropriation in the long term, (2) the conditions under which agents learn to comply with such an institution, and (3) the conditions under which learning causes the institution to fail.

### 3. Approach

We investigate the WHEN and the HOW of learning, the WHO for social learning and the probability of compliance at the start of the system. The factors and levels of our full factorial experiment are summarised below (HOW and WHO are not applicable for WHEN is Never):

Factor	Level
Institution	None, No pardons, With pardons
WHEN	Never, All, If dissatisfied
HOW	Social, Social and Asocial
WHO	Majority, Satisfied
Starting compliance	0, 0.1, 0.9, 1

The types of studied learning strategies are motivated by ethology [4], [5]. We study conformity-biased social learning equivalent to always *copy-majority*, as well as *copy-satisfied-if-dissatisfied*, based on the combination of “copy successful individual”, “copy if better” and “copy if dissatisfied” strategies. Instead of considering the level of success of another agent we use its satisfaction. For asocial learning, we apply exploration via simple mutation. We apply a simple model to incorporate both social and asocial learning in the system: when an agent is dissatisfied with the appropriation, social or asocial learning are adopted with equal chance. Social learning involves modifying ones own strategy towards that of satisfied agents and asocial learning mutates ones strategy by a fixed amount in either direction.

The tracking of resource level over the long term allows us to observe the current amount of unappropriated resource. The number of members and living agents provide evidence for the sustainability or otherwise of the institution. The number of satisfied agents and mean compliance probability allow us to establish whether and what learning took place. For measuring how equitable the different scenarios are, we employ the Gini coefficient [6], which is commonly used in economics to express the inequalities in the wealth of a population (a value of zero indicates perfect equality and 1 corresponds to maximal inequality).

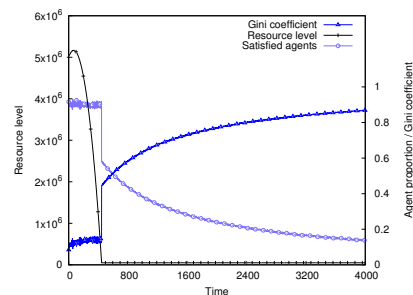


Figure 2. Effect of self-interested agent behaviour in the first-come-first-served exogenous resource appropriation scenario. Once the initial resource store is depleted, the appropriation per round becomes increasingly unequal.

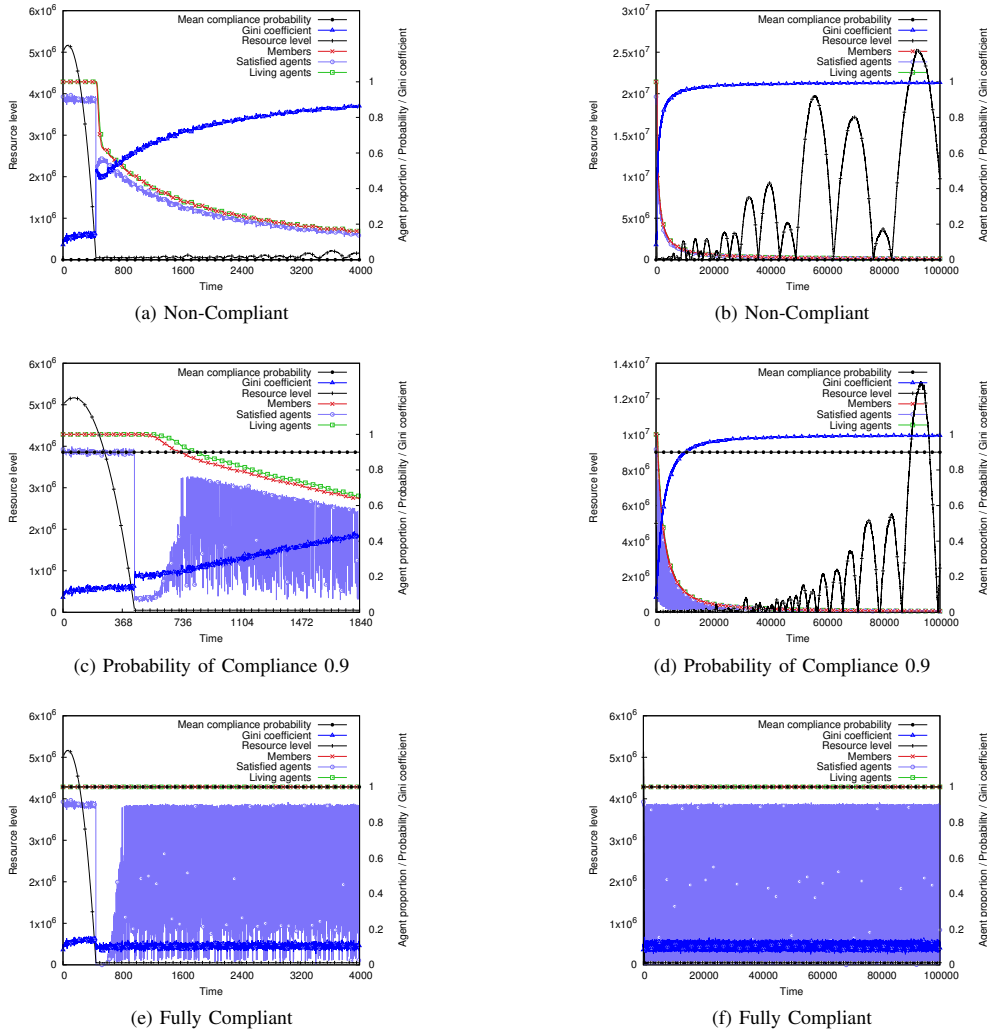


Figure 3. Fixed probability of compliance, no learning. The institution does not pardon; sanctions remain on the agents' permanent record. Unless there is full compliance by all agents, the institution removes and then eliminates its members, and thus collapses. Almost all cases lead to a highly unequal distribution of resources; only with full compliance is a more equal allocation achieved. Low probabilities were similar to non-compliance.

## 4. Results

Our first result, in Figure 2, shows the outcome with no institution. Hence there is no allocation made, no notion of compliance or strategy to learn. Agents take what they desire, and adapt their demand based on what they can get away with appropriating. Here, the early agents learn that they can take as much as they desire, and according to the demand adaptation mechanism (adopted from Pitt et al. [1]) soon realise they can appropriate the entire pool. Finally, the first agent adapts to take everything, and all others are left with zero; the Gini coefficient is approaching 1.

### 4.1. Static Probabilistic Compliance

Next, we consider stationary fixed and probabilistic compliance strategies. Figure 3 shows short and long term outcomes for varying compliance probabilities, without pardons. Figure 4 shows the same when pardons are present. In each case, every agent in the population was given the same probability (though of course their manifested behaviour varies). In all cases with low or zero compliance, the institution sanctioned and ultimately eliminated the agents, leading

to a highly unequal allocation, akin to that in Figure 2. Conversely, with full compliance, we unsurprisingly see no sanctions and an equal appropriation of resource maintained. Of particular interest is the 0.9 probability case, in comparing Figures 3c and 3d with Figures 4c and 4d, respectively. This shows the effect of the presence of pardons on a group, where each agent is likely to comply but not certain to. As discussed by Petruzzi et al. [7], forgiveness on the part of agents can enable a society to tolerate occasional infractions. Here, pardons can be seen as a mechanism for the codifying of forgiving behaviour within the institution. A high level of membership is therefore sustained, leading to the appropriation remaining highly equal.

### 4.2. Impact of Different Forms of Social Learning

Next we consider the effect of social learning of compliance probabilities. Figure 5 shows that sustainability is achieved with *copy-majority* learning, as all agents copy compliance behaviour until they have converged. This holds regardless of the initial probability, as agents learn by imitating even unintended apparent compliance when the resource

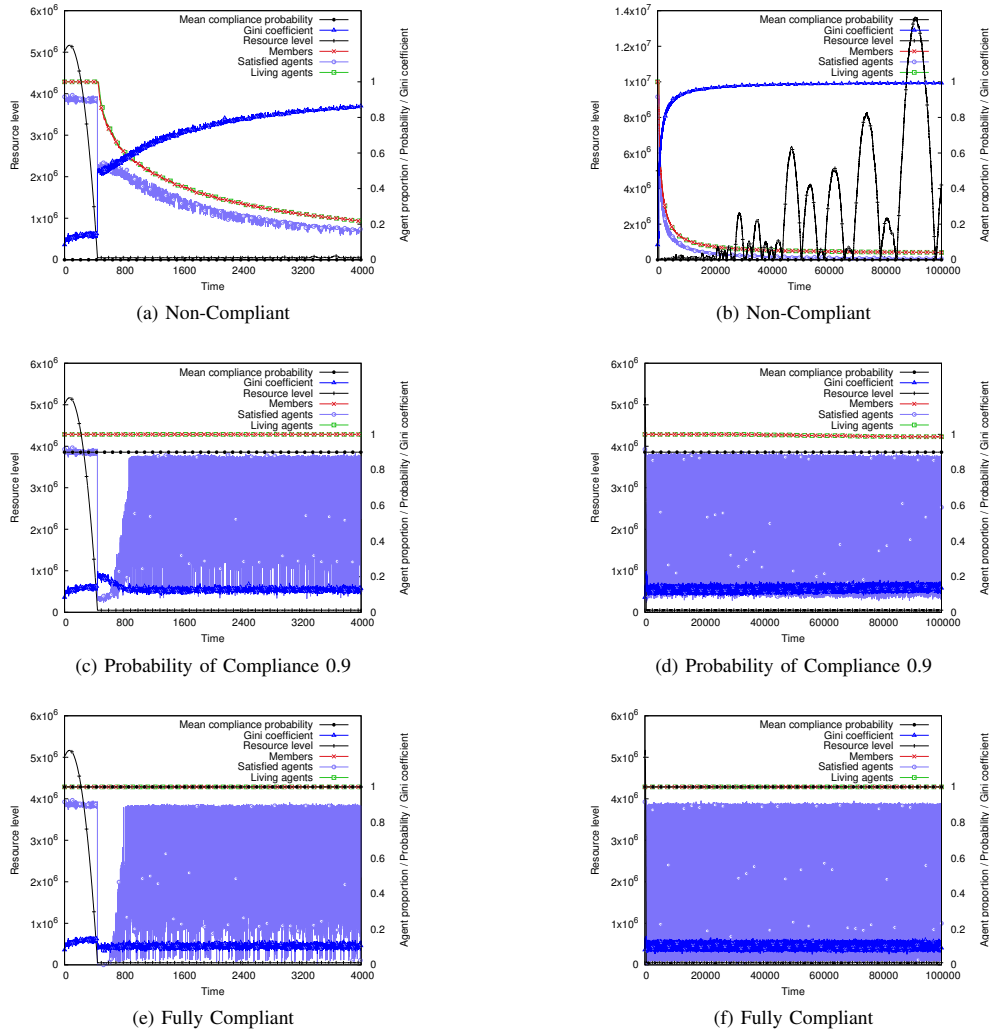


Figure 4. Fixed probability of compliance, no learning. The institution’s sanctions includes a pardon mechanism, whereby they expire after 10 rounds of not being observed to not comply. Following this, the agent return to the previous sanction level. Low probabilities were similar to non-compliance.

is plentiful. This aligns with similar results by Pitt et al. [8]. By contrast, the *copy-satisfied-if-dissatisfied* learning rule captures complacency of satisfied agents, who, due to their satisfaction, do not change their strategy. In this case, the institution is not sustained at the same level, particularly without pardons, as it is not tolerant of members who have not yet learned to behave in a socially beneficial way. The key messages from this result are that i) the form of social learning impacts on the sustainability of the institution, and ii) pardons can mitigate this, to an extent.

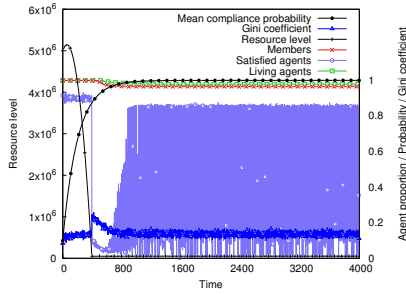
### 4.3. Mixed Social and Asocial Learning

Finally, we turn to the yet more realistic [5] model of learning, which combines both social and asocial aspects, and unlike the purely social variants, contains exploratory behaviour. Figure 6 shows the outcome of using this learning rule both with and without pardons. These results clearly show that without pardons in the institution’s sanctions, this exploratory behaviour is not tolerated, as it is manifested as non-compliance. Thus, the institution collapses relatively quickly, and the distribution of resources becomes

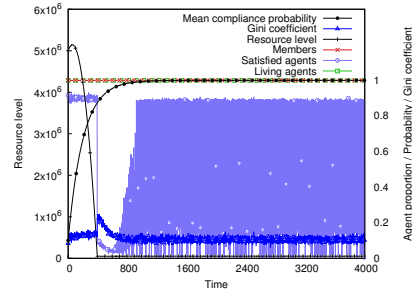
highly unequal. On the other hand, when the institution pardons sufficiently, ongoing exploratory asocial learning is tolerated. Hence, both the institution and the exploratory asocial learning remain sustainable in the long term and also the distribution of resources remains highly equitable. Note that even in Figure 6h, the institution is not sustained forever. Eventually agents will get unlucky enough cross the institution’s tolerance threshold, and will be eliminated. In this experiment, agents were initially eliminated at roughly a rate of one agent per five million rounds.

### 5. Conclusions and Future Work

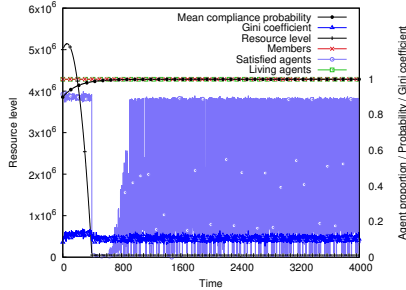
This paper studies a first-come-first-served common-pool resource appropriation problem. We exposed an interplay between pardons in the sanction mechanism, and forms of social and asocial learning used by agents. We showed that the way agents learn has a direct impact on the existence and sustainability of the institution, and at the same time, the institution’s features can either tolerate or inhibit agent learning. Exploration and complacency are shown to be key behaviours that institutions may need to explicitly tolerate,



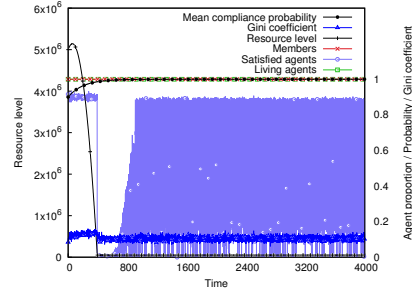
(a) Starting probability 0.1. No pardons. Copy-majority.



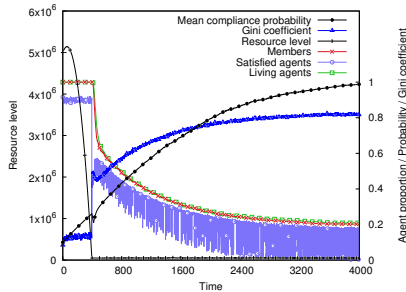
(b) Starting probability 0.1. With pardons. Copy-majority.



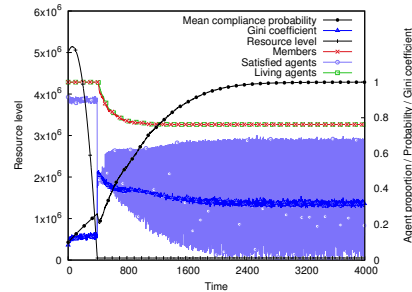
(c) Starting probability 0.9. No pardons. Copy-majority.



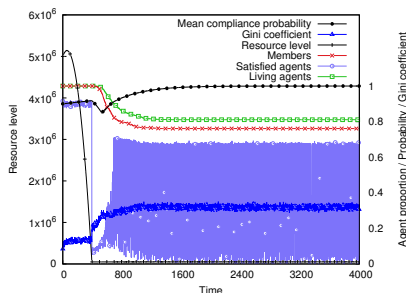
(d) Starting probability 0.9. With pardons. Copy-majority.



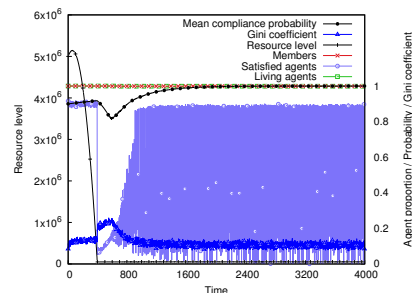
(e) Starting probability 0.1. No pardons. Copy-satisfied-if-dissatisfied.



(f) Starting probability 0.1. With pardons. Copy-satisfied-if-dissatisfied.



(g) Starting probability 0.9. No pardons. Copy-satisfied-if-dissatisfied.



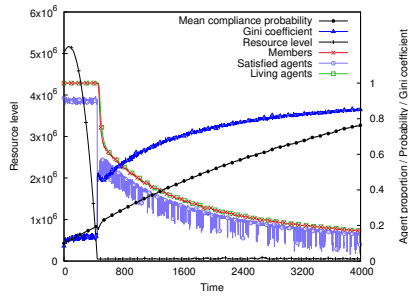
(h) Starting probability 0.9. With pardons. Copy-satisfied-if-dissatisfied.

Figure 5. 5a–5d: All agents modify their strategy each round to behave more like the observed majority behaviour. 5e–5h: The *copy-satisfied-if-dissatisfied* social learning rule means agents who appropriate their desired amount were “complacent”, and did not learn in that round. This can lead to institutional collapse to varying degrees, especially when the initial probability of compliance is low, and/or without pardons.

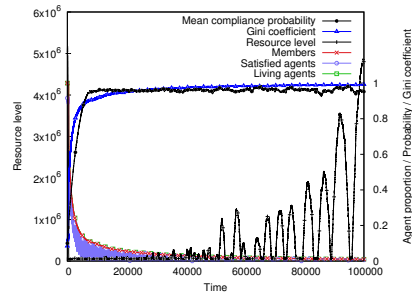
and we identify a key role for pardons in the institution’s sanction mechanism in doing so. Pardons allow the institution to tolerate complacency and exploration, which are behaviours associated with ongoing learning.

## References

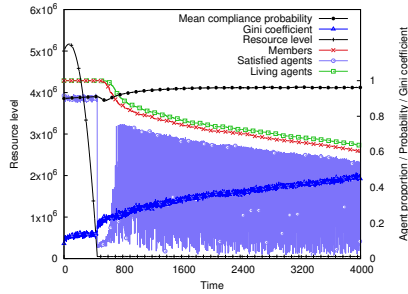
- [1] J. V. Pitt, J. Schaumeier, and A. Artikis, “Axiomatization of socio-economic principles for self-organizing institutions: Concepts, experiments and challenges,” *ACM Trans. Auton. Adapt. Syst.*, vol. 7, no. 4, pp. 39:1–39:39, 2012.



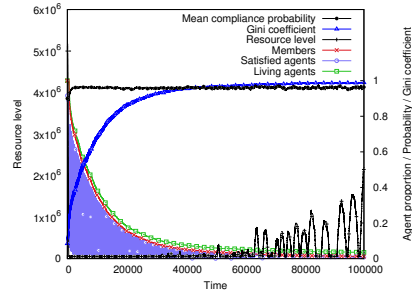
(a) Starting probability 0.1.



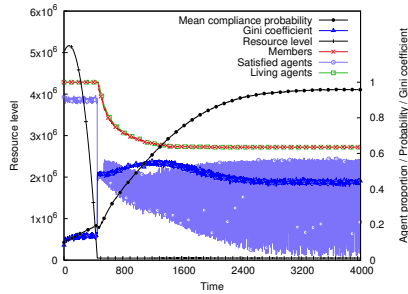
(b) Starting probability 0.1.



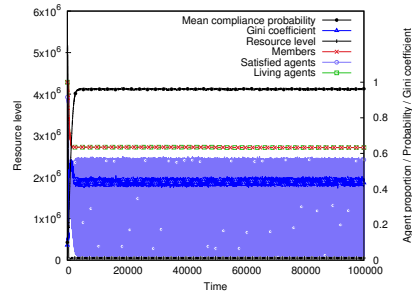
(c) Starting probability 0.9.



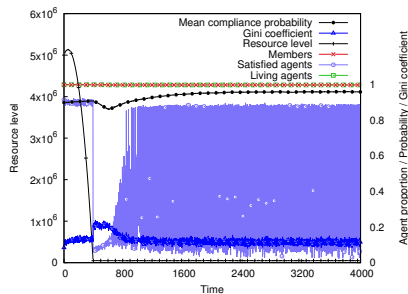
(d) Starting probability 0.9.



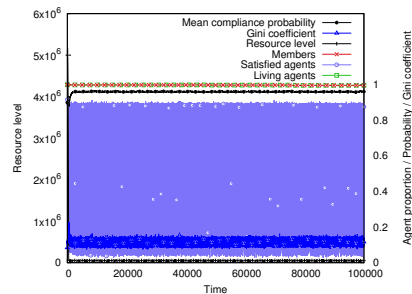
(e) Starting probability 0.1. With pardons.



(f) Starting probability 0.1. With pardons.



(g) Starting probability 0.9. With pardons.



(h) Starting probability 0.9. With pardons.

Figure 6. Agents use both social and asocial learning when dissatisfied. Sanctions are graduated, and in 6e to 6h they additionally contain pardons.

- [2] J. V. Pitt, D. Busquets, A. Diaconescu, A. Nowak, A. Rychwalska, and M. Roszczyńska-Kurasińska, "Algorithmic self-governance and the design of socio-technical systems." CEUR-WS, 2014, pp. 262–273.
- [3] E. Ostrom, *Governing the Commons*. Cambridge, UK: Cambridge University Press, 1990.
- [4] K. N. Laland, "Social learning strategies," *Learning & Behavior*, vol. 32, pp. 4–14, 2004.
- [5] I. Coolen and K. N. Laland, "Trade-offs in the adaptive use of social and asocial learning," *Advances in the Study of Behavior*, vol. 35, pp. 333–379, 2005.
- [6] L. Ceriani and P. Verme, "The origins of the Gini index: extracts from *Variabilità e Mutabilità* (1912) by Corrado Gini," *J Econ Inequal*, vol. 10, pp. 421–443, 2012.
- [7] P. E. Petrucci, D. Busquets, and J. V. Pitt, "A generic social capital framework for optimising self-organised collective action," in *IEEE 9th International Conference on Self-Adaptive and Self-Organizing Systems*, 2015, pp. 21–30.
- [8] J. V. Pitt, J. Schaumeier, and A. Artakis, "Coordination, conventions and the self-organisation of sustainable institutions," in *PRIMA*, D. K. et al., Ed., 2011, pp. 202–217.