Modelling Enduring Institutions: The Complementarity of Evolutionary and Agent-based Approaches

Simon T. Powers[a], Anikó Ekárt[b], Peter R. Lewis[b]

[a]*Edinburgh Napier University, UK*
[b]*Aston University, UK*

Abstract

Empirical work has shown that societies can sometimes avoid antisocial outcomes, such as the Tragedy of the Commons, by establishing institutional rules that govern their interactions. Moreover, groups are more likely to avoid antisocial outcomes when they design and enforce their own rules. But this raises the question: when will group members put effort into maintaining their institution so that it continues to provide socially beneficial outcomes? Ostrom derived a set of empirical principles that predict when institutions will endure, which have subsequently been formalised in agent-based models that are based on an executable description of the content of an individual's behaviour. Here we show how these models can be complemented by evolutionary game theory, which focuses on the value or payoff of different behaviours, rather than on the mechanistic content of the behaviour. Using such a value-based model, we determine exactly when individuals will be incentivised to maintain their institution and enforce its rules, including the critical amount that a group must invest into incentivising agents to monitor rule compliance. We highlight the complementarity of content-based and value-based modelling approaches, and therefore provide a step towards unifying theoretical and empirical approaches to understanding enduring institutions and other social phenomena.

Keywords: institutions, evolutionary game theory, agent-based modelling

Email addresses: `s.powers@napier.ac.uk` (Simon T. Powers), `a.ekart@aston.ac.uk` (Anikó Ekárt), `p.lewis@aston.ac.uk` (Peter R. Lewis)

1. Introduction

Cooperation can be defined as a behaviour that provides a benefit to other individuals, i.e. increases the social welfare of the group. Under the assumption of self-interested behaviour, micro-economic theory demonstrates that if agents are to cooperate, then there needs to be the provision of individual incentives for them to do so (Oliver, 1980; Olson, 1965). Increasing social welfare in and of itself is not sufficient; individuals must gain more from cooperating than from defecting. Left unchecked, this leads to the phenomenon known as the Tragedy of the Commons (Hardin, 1968), in which antisocial outcomes pervade, such as the depletion of common-pool resources. This result has been the prevailing starting point for many socio-economic policy decisions, as well as many distributed computing design decisions, for several decades.

However, the conclusions of the Tragedy of the Commons rest on the assumption that individuals are playing a particular game form, corresponding to an n-player version of the single-shot Prisoner's Dilemma (Ostrom, 1990). In reality, individuals typically have the potential to change the rules of their social interactions (North, 1990; Reiter, 1996), by reasoning through the situation in which they find themselves. In economics, an institution is defined as a family of game forms (strategies and the mappings between strategies and material outcomes) that individuals can choose between, given the state of the physical environment (e.g. their resource endowments) and their current technology (Hurwicz, 1996). More informally, we can think of a game form as the "rules of the game", and hence of individuals as being able to choose the rules of their game by creating an institution.

There are many empirical examples of societies being able to avoid anti-social outcomes by devising institutional rules that govern their interactions in the use of common-pool resources such as grazing lands, fisheries, and irrigation systems (e.g. Ostrom, 1990). Example rules include how much water an individual may take from a shared irrigation system, when they may take it, how often they must perform maintenance, etc. Furthermore, the empirical work suggests that these rules are self-enforcing (Greif, 2006), in the sense that it pays both for individuals to follow them, and to take actions that encourage others to follow them.

From a theoretical viewpoint, the creation of these rules changes the game form into one where self-interested individuals do best by cooperating (Greif, 2006; Hurwicz, 1996; North, 1990). The Folk Theorem of game theory explains why this can work (Binmore, 2014): when interactions are repeated, cooperation can be sustained as an equilibrium by conditional strategies that respond to the past behaviour of other agents. One example of such a strategy is Tit-for-Tat (Axelrod, 1984): cooperate on the first round, and thereafter mirror what the other agent did on the previous round. But this is just one example. In general, the Folk Theorem shows that any strategy that gives an agent more than the minimax payoff can be sustained as an equilibrium amongst self-interested agents. The minimax payoff is the largest payoff that an agent can receive if its opponent tries to minimise the agent's payoff, which in the Prisoner's Dilemma corresponds to the payoff received when the opponent defects. Therefore, any strategy that gives the agent a higher payoff than always being defected against will be an equilibrium when adopted by all of the agents, since if the agent deviated from this strategy then it could have its payoff reduced to the minimax payoff by its co-players. Importantly, this result also holds where N agents interact simultaneously (Fudenberg & Maskin, 1986), e.g. in the management of common-pool resources.

However, cooperation between self-interested agents under the Folk Theorem requires that the agents value future payoffs, do not know when their interactions will end, and have sufficient information about how other agents have behaved in the past. By creating institutional rules, individuals can create a social environment that satisfies these conditions (Guala, 2012), e.g. by setting up systems of monitoring (Ostrom, 1990), facilitating the spread of reputation (Hardy & Norgaard, 2015; Milgrom, North & Weingast, 1990), and decreasing the outside options of the agents so that they do indeed value future payoffs and do not know when their interactions will end (Casari, 2007). Furthermore, the creation of institutional rules helps agents to coordinate their behaviour onto one of the many possible equilibria, by creating shared expectations about how other agents will behave (Greif, 2006).

Creating, updating and implementing these institutional rules requires time and effort. Without this they are likely to collapse and individuals will revert back to the default game form where cooperation is not favoured. Ostrom's field studies suggest that institutions are more likely to endure and maintain socially beneficial outcomes in the long term when the institutional rules are both created and implemented by the same agents whose economic interactions are affected by those rules. This then raises the question: under what conditions will self-interested agents be willing to put the effort into doing this, by taking on various *institutional roles*? Examples of institutional roles include acting as a monitor to check for rule compliance, or organising votes on rule changes. We cannot predict whether institutions will endure in the long term without examining the incentives for agents to take on institutional roles.

In order to examine the conditions under which institutions can endure and maintain cooperation, researchers have recently formalised Ostrom's principles of enduring institutions using agent-based models (e.g. Pitt, Schaumeier & Artikis, 2012; Smajgl, Izquierdo & Huigen, 2008, 2010). Agent-based modelling provides a highly effective method with which to conduct experimental studies on the consequences of different assumptions about behaviour (Di Paolo, Noble & Bullock, 2000); in the humanities and social sciences, they have been referred to as *digital Petri dishes* (Gavin, 2014). Agent-based modelling is a highly attractive approach, primarily due to its ability to capture complex behaviours and interactions in executable form, and to explore emergent phenomena simply by "running" variants of the model (Bonabeau, 2002; Epstein & Axtell, 1996). This is particularly helpful when building intuition or illustrating counter-examples. However, due to the complexity of the formal description required, there is also a limit to its explanatory power. This is particularly true when answering questions related to incentivisation and critical values of parameters in a rigorous way.

As an alternative, evolutionary game theory (Maynard Smith, 1982) is a more descriptive modelling technique first established in theoretical biology to study the evolution of adaptive traits in populations of animals. It has since been applied in economics, sociology, anthropology, and elsewhere in biology, and is used to explore both genetic and cultural evolution.

In this paper, we explore how agent-based models, based on executing the *content of strategies*, can be complemented by evolutionary game theory, where a description of the *value of strategies* instead forms the basis. This allows us to draw on existing results and understanding from the evolutionary game theory literature, in order to provide additional insight. Specifically, we provide new analytical insight into the effects of different ways of incentivising agents to take on an institutional monitoring role, and on the optimal proportion of a group's resources that they should invest into monitoring.

The discussion and results in this paper therefore provide a step towards unifying theoretical and empirical approaches to understanding the formation of enduring institutions. Further, we anticipate that this will readily aid research into other questions of social and cultural nature.

2. The Complementarity of Content-based and Value-based Models

Both agent-based modelling (ABM) and evolutionary game theory (EGT) are well-established approaches to modelling social systems, especially for answering questions relating to population-level results arising from interactions between individuals with (potentially varying) behavioural strategies. We characterise these as instances of *content-based* and *value-based* modelling approaches, respectively, and in this section, explore their complementarity in general. Figure 1 illustrates this.

2.1. What do ABM and EGT capture and what do they assume?

Game theory defines a *strategy* as a mapping from environmental context to actions (Fudenberg & Tirole, 1991). An agent's strategy therefore defines its

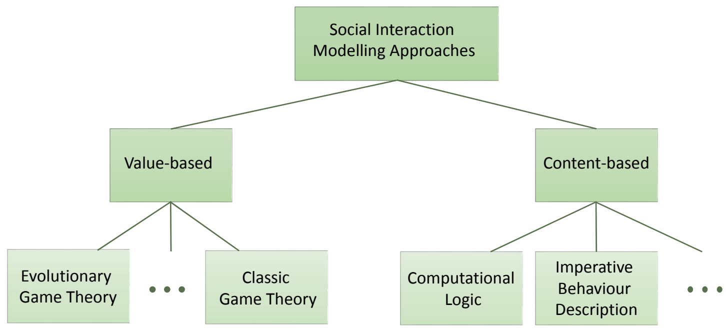

Figure 1: A sketch of a taxonomy of approaches for modelling systems of social interactions. The primary distinction made in this paper is between approaches that capture the value of different strategies, and those that capture the content of those strategies. A range of game theory variants, including EGT, can then be seen as value-based approaches. ABM, and other forms of executable simulation modelling, instead model the content of agent strategies. This can be implemented programatically in different ways, as shown in the right-hand side of the figure.

behaviour, as its environment (including the behaviour of other agents) changes. This is equivalent to the notion of Russell and Norvig's *agent function* (Russell & Norvig, 2010) in artificial intelligence: the mathematical object that maps a given percept sequence to an action, and is made explicit through an *agent program*. In ABM, many such agents are instantiated. Each agent's program is executed and is responsible for maintaining its own state over time. The environment is also modelled with a state that may be changed over time, both as a result of agent actions, or by a program that captures natural forces.

By contrast, in EGT, an explicit description of the content of a behaviour and its impact on the environment is not given. Instead, strategies are considered as *traits* that may be more or less prevalent in a population (which is typically assumed to be infinite and well-mixed), and compete with each other in an evolutionary sense. The task is then to write equations that describe the fitness (i.e., evolutionary value) of a strategy, given the current frequency of each of all possible strategies in the population. The dynamics of strategy frequency are then explored, under the assumption that the change in strategy frequency is correlated with the fitness of the strategy.

A key distinction, therefore, can be made in terms of what is captured and what is assumed in each case. In EGT the existence of a space of possible behaviours and their expected fitness is presented in a descriptive (equation-based) form. However, the content of the actions themselves that form part of the strategy, and lead to this fitness, are omitted. This omission includes any deliberative or developmental processes that are assumed to be included in the execution of the strategy; only the value of any such activity, in terms of its fitness, is given.

By contrast, in ABM a description of the content of the modelled actions is provided, typically in imperative or logical form, along with what effect they have on the world and other agents. Thus, it is possible to capture a deep and complex set of behaviours in an agent, based (for example) on learning, deliberative, and other cognitive processes. However, there is no explicit description of the value of carrying out the described activities, and furthermore, such a value is hard to arrive at, save by executing the agent programs and observing.

In summary, ABM and EGT each leave implicit what is made explicit in the other. ABM can capture rich behaviours, but struggles to support an analysis of their value. Conversely, EGT provides the necessary primitives to analyse the incentives and outcomes associated with different behaviours in a rigorous way, yet in doing so lacks the ability to capture what may be crucial details of the nature of the strategies themselves, and assumes that any value is accurately defined.

2.2. Producing a Justifiable Model

In ABM, model justification is usually done through calibration against observed phenomena (Janssen & Ostrom, 2006). First, one observes and captures micro-level behaviours (i.e., the behaviour of a single agent in a specific context), producing an agent program that replicates that behaviour. Second, one then observes macro-level behaviour for already well-observed phenomena, and the model is calibrated to ensure that observed emergent (global) outcomes are reproduced. Additional macro-level outcomes are then reported as predictions of the model.

As discussed in Section 2.1, to produce an EGT model, it is required that the modeller is able to arrive at the evolutionary 'value' of each possible strategy, in a way that justifiably drives strategy frequency. Such a justification is often plausible: in animal population studies, it is possible to identify which traits are correlated with greater numbers of successful offspring, and traits are assumed to be heritable; in economics, firms are more likely to copy the traits embodied by financially successful firms than those of bankrupt ones. However, in more complex social systems where growth in frequency of a strategy is likely to be strongly determined by human cognitive aspects, and not primarily driven by the copying of behaviour, we must be careful to ensure that the definition of evolutionary value in the model is justified.

2.3. Performing Analysis with ABM and EGT

The primary method of interrogation of ABMs is through experimentation on the effects of varying different parameters and behavioural rules. Thus, having expressed a set of executable behaviours, one needs to take an inductive scientific approach to arriving at claims. One varies parameters of the model (often both those within individual behaviours, as well as those concerning the world), and one can explore, in a black-box way, the outcomes the system produces. Typically, a full factorial or similar approach is taken, in order to build confidence in claims relating to the effect of varying each parameter.

In EGT, there is no requirement to execute the model, although the equations that form the model are often solved numerically through a computer program, in addition to being analysed in the classic sense. With EGT models, one is primarily looking for relationships and critical values that can be deduced by solving the equations algebraically. Numerical simulation is often used to validate or to solve these where tractability becomes an issue.

2.4. The Complementarity of ABM and EGT

As it is hopefully clear from the above discussion, neither ABM nor EGT is able to replace the other in terms of supporting the full breadth of analysis forms that the other provides; both bring something to the table for the modeller of social systems. Similarly to how, in software engineering, different language styles, e.g., imperative or functional, are used for different purposes, in the modelling of social systems, different modelling approaches are better suited to address different questions.

One significant benefit of content-based approaches like ABM is the ability to make the specification of the model its own execution. A further benefit is that it is often easier to discern and model the content of an agent's behaviour, rather than the value of that behaviour, and to capture this in a model. Deriving a payoff matrix from empirical observations can be difficult (although this has been done quantitatively in some studies, e.g. Gore, Youk & van Oudenaarden, 2009). Content-based approaches therefore lend themselves more to empirical study, *exploring* the outcome of observed behaviour. This can be achieved without the need to concern oneself with details of a method of analysis or solving the model, beyond running and interrogating a simulation. Content-based approaches vary in how they approach the description of agent behaviours. When using an imperative language (e.g., Lewis & Ekárt (2017) used Java), the solving is embedded in the description of the system itself. Alternatively, computational logic (e.g., Pitt et al. (2012) used Prolog) may be used to separate behaviour specification from behaviour execution, the latter being carried out through query resolution. Value-based approaches share this separation of concerns. However, the key benefit in the value-based case is that the description is already a statement that *quantifies* the outcome of carrying out a given behaviour. Therefore, critical parameters affecting the outcome are more readily accessible and they no longer need to be induced from the execution. Much of the rest of this paper, especially Section 5, is devoted to illustrating this benefit in the domain of institutional modelling.

2.5. Existing cross-fertilisation between Evolutionary Game Theory and Agent-Based Modelling

ABM and EGT have been successfully used to feed into each other. One of the first examples was Axelrod's tournament, where researchers were invited to submit different agent programs to play a repeated Prisoner's Dilemma game (Axelrod, 1984). This allowed Axelrod to empirically explore the space of different possible strategies, and their behavioural interactions with each other, rather than having to presuppose a fixed number in a model. However, analysis of the winning Tit-for-Tat strategy, in terms of the conditions under which it was stable and the conditions under which it could become established in a group, was eased by using a value-based evolutionary game theory approach (Axelrod & Hamilton, 1981). More recently, a similar tournament where researchers submitted agent programs containing different social learning strategies allowed the traditional assumptions of value-based models of social learning to be relaxed. This produced new insights that have in turn fed back into more descriptive value-based models of social learning (Rendell, Fogarty, Hoppitt, Morgan, Webster & Laland, 2011).

More generally, content-based ABMs have been used to expand results from EGT by relaxing assumptions such as only a small number of mutations being present at one time, and no communication between players (Adami, Schossau & Hintze, 2016). Going the other way, value-based models have provided insight into when individual strategies that punish non-cooperative behaviour can actually be stable (Lehmann, Rousset, Roze & Keller, 2007) that were difficult to achieve in simulation (Boyd, Gintis, Bowles & Richerson, 2003). In the remainder of this paper we examine how similar cross-fertilisation can benefit the study of institutions.

3. Institutions for Managing Common-pool Resources

Many individual behaviours are needed to sustain an institution. These include designing the rules, voting on them, monitoring the behaviour of group members, and sanctioning those found breaking the rules. It has been shown that if we abstract away from how these behaviours are carried out then institutions can both lead to stable cooperation (Pitt et al., 2012; Sasaki, Brännström, Dieckmann & Sigmund, 2012), and evolve *de novo* (Powers & Lehmann (2013)). In these models *institutional roles*, such as designing rules or monitoring rule compliance, are contracted out – it is assumed that some individuals will faithfully carry out these roles without shirking or free-riding. But to understand when institutions will be sustainable, we need to understand under what conditions it pays individuals to perform these roles. While many micro-level models of monitoring and sanctioning have been produced using classical and evolutionary game theory, these have not considered the context of institutional roles. How do evolving institutional rules affect individual incentives to monitor and sanction? In this study, we analyse different incentivisation mechanisms from both ABM and EGT perspectives.

3.1. Common-Pool Resource Allocation and the Tragedy of the Commons

A common-pool resource (CPR) is defined by Ostrom (1990, p.30) as "a natural or man-made resource system that is sufficiently large as to make it costly (but not impossible) to exclude potential beneficiaries from obtaining benefits from its use". Examples of such resource systems could be fisheries, various water resources ranging from groundwater basins to lakes and oceans, irrigation systems, bridges, and computer clusters. We study resource systems used by multiple individuals, who can appropriate or use resource units, such as tons of fish harvested from a fishery, cubic meters of water withdrawn from a water resource, number of crossings of a bridge, central processing units consumed on a cluster computer.

In a game-theoretic formulation of the common-pool resource allocation problem, at each time step, given the allocation of resource units to individuals, each individual can decide to comply and appropriate the allocated amount (cooperate) or not comply and appropriate the amount they wish (defect).

The Tragedy of the Commons (Hardin, 1968) is defined as the inevitable consequence of rational, self-interested individuals appropriating any number of resource units that they wish. Over time, as the individuals see the benefits of their own appropriations, they will increase their appropriations. The common-pool resource is expected to degrade and become depleted over time, due to the uncontrolled appropriations from the limited resource.

Historically, attempts to avoid the Tragedy of the Commons have involved centralisation or privatisation. With centralisation, an imposed institution would control the allocation of resource units to appropriators, monitor compliance and sanction non-compliance. In the case of privatisation, the resource is divided among individuals and they then become responsible for their share. Based on studies of small, closed CPR instances, such as fisheries, Ostrom pioneered new forms of institutions, where once the institution is in place, the individuals would self-organise and self-govern their resource in a way that prevents the Tragedy of the Commons from occurring.

3.2. Ostrom's principles for enduring institutions

Ostrom (1990) has extensively studied the governance of long-enduring, self-organised and self-governed CPRs, including fisheries, water irrigation systems and forests, some as old as 1000 years. The main studied aspects were the problems of commitment and mutual monitoring.

From this empirical study Ostrom derived eight principles for the design of long-enduring institutions:

1. *Clearly defined boundaries:* As a first step in organising for collective action, both the individuals who have the right to appropriate resource units from the CPR and the boundaries of the CPR must be clearly defined.
2. *Congruence between appropriation and provision rules and local conditions:* Having rules for appropriation and provision specific to the local conditions of the particular resource contributes to the endurance of CPRs. For example, in the Spanish huertas, substantially different rules must be applied in different regions for water irrigation, depending on local specificity, even though the water management problem is broadly similar.

3. *Collective-choice arrangements:* Appropriators can participate in the design of the institution by tailoring the rules over time. It must be noted that appropriators will not necessarily comply with good operational rules, when these exist, even if they took part in their design. Furthermore, even when reputation is important and individuals share the norm of honouring agreements, these are insufficient by themselves to ensure stable cooperation in the long term.
4. *Monitoring:* Monitors, who audit both state condition and appropriation behaviour, are part of or accountable to the appropriators. The cost of monitoring in long-enduring CPRs is often low. For example, in an irrigation system using a rotation appropriation rule, monitoring is a by-product: the individual nearing the end of their turn might wish to extend their turn, while the next individual ready to start their turn might wish to start earlier. They thus mutually monitor each other and ensure compliance with the rule by both.
5. *Graduated sanctions:* Appropriators, who do not respect community rules, are applied sanctions dependent on the seriousness of their offence, by appropriators or assigned officials accountable to appropriators, or both. The graduated sanctions will have to work hand-in-hand with monitoring to ensure sufficient level of rule-following and avoid increase in infractions.
6. *Conflict-resolution mechanisms:* There must exist cheap and easily accessible mechanisms to resolve conflicts between appropriators and officials or among appropriators. Although this by itself does not ensure enduring institutions, the maintenance of complex rule systems over time is helped by it.
7. *Minimal recognition of rights to organize:* External governmental officials do not challenge the right of appropriators to devise their own institutions. For example, in a fishery, local fishers can devise the rules determining who can use the fishing ground and with what equipment, without their authority being challenged by external governmental officials.
8. *Nested enterprises:* In the case of larger CPRs, organisation of all activities

is in the form of multiple layers of nested enterprises, with small, local CPRs at their bases.

Ostrom predicts that where these eight principles are satisfied institutions will be maintained and will continue to prevent over-exploitation of common-pool resources over a long time horizon.

3.3. Agent-based modelling of enduring institutions

In this section, we examine how Ostrom's empirical principles for enduring institutions have been formalised using ABM. Formalisation is necessary both to gain a deeper understanding of the conditions under which they are effective, and to allow their implementation in socio-technical systems that contain artificial as well as human agents. We highlight three studies contributing to the agent-based modelling of enduring institutions, starting from formal axiomatisation (Pitt et al., 2012) and continuing with the relationship between institutional features and forms of learning (Lewis & Ekárt, 2017) and relaxation of norms for sustainable institutions (Kurka & Pitt, 2017).

Pitt et al. (2012) develop a formal axiomatisation of Ostrom's first six principles for CPR in Event Calculus. They implement an executable test-bed and show that these principles support enduring institutions. They build gradually more complex and realistic tests for the principles. They find that when the agents comply with the rules for appropriation, the first three principles are sufficient for the institution to endure.[1] When the assumptions on compliance are relaxed, this is not the case any more and the next three principles become necessary. In their setting, these six principles ensure enduring institutions with high membership and resource sustainability. Thus, with this work, they establish the feasibility of an institution-based approach to dynamic resource allocation, specifically when long-term endurance is sought.

Lewis & Ekárt (2017) focus on the interplay between institutional features and forms of learning used by agents. They show that the way the agents learn

[1] Their experiments consider a lifespan of 500 time steps.

influences directly the existence and sustainability of the institution, and at the same time, the institution's features can either tolerate or inhibit learning. Institutional pardons in the sanctioning mechanism (Ostrom's principle 5) have a key role, as they allow for tolerance of behaviours associated with ongoing learning, such as complacency and exploration.

Kurka & Pitt (2017) study the relaxation of norms, in particular of sanctioning strategies for non-compliance in socio-technical systems, in a scenario where monitoring comes at a cost and also subjective and diverse behaviour of agents can be expected. They define principled violation of policy as "the active and intentional decision of an agent of not applying a policy to which it is entitled" (i.e. a sanction). They demonstrate via a series of experiments on CPR allocation that strategies of partially applying sanctions lead to more cost-effective solutions, that are flexible to different scenarios and behaviour.

So, ABM shows how both institutional pardons and partial sanction application are mechanisms that can lead to more sustainable institutions. But how can agents be incentivised to take on the roles that lead to sustainable institutions (such as monitoring behaviour or organising votes to change the rules)?

4. The Challenge of Predicting Conditions for Establishment and Sustenance of Cooperation-Promoting Institutions

Having established the complementarity of value-based and content-based models in general, and EGT and ABM in particular, in Section 2, our aim is to establish the value of each approach in understanding and controlling the behaviour of agents forming an institution to resolve common-pool resource allocation problems. The role of ABM has already been well demonstrated in prior work (as discussed already in this section), therefore, in the remainder of this paper we focus on illustrating additional insight that can be obtained by taking a value-based, EGT approach.

Using EGT, we focus on the challenge of predicting conditions for the formation and sustenance of cooperation-promoting institutions, when individual agents have to be incentivised to take on the institutional roles that are necessary for this. These predictions would be difficult to make from an agent-based model, other than by interrogating it rather laboriously in a black-box fashion. Here we aim to derive relations between parameters in order to answer the following questions:

1. How many agents need to take on a monitoring role in order to incentivise cooperation?
2. What level of investment into monitoring is necessary to incentivise this number of agents to become monitors?
3. What are the conditions for cooperation to become established given an initial state where no agent cooperates and no agent monitors?

5. Illustrating the Role and Benefits of Value-based Models

To illustrate the role and benefits of value-based models, we consider under what conditions agents can be incentivised to monitor each other's compliance with institutional rules. Previous work has recognised that monitoring rule compliance is necessarily costly. Monitoring can carry both physical costs, e.g., energy or CPU cycles, and opportunity costs where the time spent on monitoring is time lost carrying out other productive activities. This is true both in natural systems, such as irrigation systems (Weissing & Ostrom, 2000) and fisheries, and artificial systems such as community clouds (Khan, Freitag & Rodrigues, 2015) or community co-production energy systems (Torrent-Fontbona, López, Busquets & Pitt, 2016). Therefore, if self-interested agents are to be incentivised to monitor rule compliance then they need to be reimbursed for this cost somehow.

One empirically grounded way in which the costs of monitoring can be reimbursed is by using a fraction of the group's common-pool resource to pay for monitoring. This fraction of the resource invested into monitoring is an *institutional fact*, i.e., it is determined by the current institutional rules. Several models have examined the effect of different levels of investment into monitoring at an abstract level (Balke, De Vos & Padget, 2013; Jaffe & Zaballa, 2010; Pitt & Schaumeier, 2012; Powers, 2018; Powers & Lehmann, 2013), by assuming that the probability that an agent is monitored for rule compliance is proportional to the amount of resource invested into monitoring. But these models did not examine what would happen if agents have to choose whether or not they will take on the monitoring role, and how the level of monitoring will consequently evolve over time. Here we take this theoretical work further by developing a micro-level model that considers agents explicitly choosing whether or not to take on a monitoring role when they must pay a cost for doing so.

In the following sections we develop a general descriptive model and then consider several variants in which monitoring is incentivised in different ways.

5.1. Base Model

We consider a model in which n agents take part in a linear public goods game to provision a common-pool resource. Each agent makes three decisions: i) whether or not to cooperate by provisioning the common pool at a cost to itself; ii) whether or not to pay a tax to support implementation of the institution; and iii) whether or not to monitor other agents to determine if they have contributed.

Agents that both did not provision to the common pool and were monitored (thus caught) are sanctioned, creating a cost to free-riding (C_{F}).

Provisioning the common- pool resource, as well as taking on the monitoring role, carries some cost to the agent (C_{C} and C_{M}, respectively). Monitors are reimbursed for their work according to two different schemes that we compare and contrast below. The process is then repeated for a number of rounds T.

Thus, the utility of an individual agent will be built up from a base utility, the individual's share of the common-pool resource, the individual's cost if they contribute to the common-pool, the individual's cost if instead they free-ride, the individual's net benefit if they take on a monitoring role, and the individual's cost of paying a tax to support the institution (C_{τ}).

More formally, the utility of agent i at round t is given explicitly by the following function:

$$u_i(t) = u_0 + B_{\mathrm{G}}(t) - \iota_{i\mathrm{C}} C_{\mathrm{C}} - (1 - \iota_{i\mathrm{C}}) C_{\mathrm{F}}(t) + \iota_{i\mathrm{M}} [B_{\mathrm{M}}(t) - C_{\mathrm{M}}] - \iota_{i\tau} C_\tau, \quad (1)$$

where $\iota_{i\mathrm{C}}$, $\iota_{i\mathrm{M}}$ and $\iota_{i\tau}$ are indicator variables that take the value 1 if the agent contributes to the common pool, monitors, and pays tax to support the institution, respectively, and 0 otherwise. In this utility function, u_0 is a baseline utility in the absence of social interactions. The term $B_{\mathrm{G}}(t)$ represents the individual's share of the common-pool resource, computed as:

$$B_{\mathrm{G}}(t) = \frac{1}{n} \times \alpha\, n_{\mathrm{C}}(t), \quad (2)$$

where $n_{\mathrm{C}}(t)$ is the number of agents that provisioned the common resource on round t (the number of agents with $\iota_{\mathrm{C}} = 1$) and α is a model parameter representing the amount of resource that each agent provides when they provision. The parameter C_{C} represents the cost to the agent of provisioning α units of common-pool resource. Following the definition of a linear public goods game, we assume that $C_{\mathrm{C}} < \alpha$, i.e. there is a benefit to agents of cooperating together to share their resources.

The term $C_{\mathrm{F}}(t)$ represents the cost of free-riding, i.e. of an agent not provisioning the common pool. This cost is paid by all agents with $\iota_{\mathrm{C}} = 0$. [2] The cost is calculated as the probability than an agent is monitored, multiplied by the sanction imposed if detected free-riding, s. This is computed as:

$$C_{\mathrm{F}}(t) = \frac{p\, n_{\mathrm{M}}(t)}{n} s, \quad (3)$$

where $n_{\mathrm{M}}(t)$ is the number of agents that take on the monitoring role at round t, i.e. the number of agents with $\iota_{\mathrm{M}} = 1$, and p is the number of agents monitored by each monitor. We assume that each monitor monitors a different, non-overlapping, set of agents, and that an agent is not monitored more than once. This corresponds to an assumption that agents have the technology to perfectly coordinate their monitoring.

[2] In the remainder of the paper we do not use the i index, as it would not affect the analysis and makes the formulas easier to read.

The term $B_{\mathrm{M}}(t)$ represents the amount that monitors are reimbursed for their monitoring work. We examine different ways in which monitoring can be paid for, and hence different expressions for $B_{\mathrm{M}}(t)$, in Sections 5.2 and 5.3.

The term C_{M} represents the cost to agent i of monitoring other agents. The cost of monitoring a single agent is δ, so the total cost to an agent of monitoring on one round is

$$C_{\mathrm{M}} = p\delta. \tag{4}$$

Finally, the parameter C_τ represents the tax paid each round to support implementation of the institutional arrangements, which is paid by all agents with $\iota_\tau = 1$.

The costs of monitoring, contributing to the common pool, and paying tax to support implementation of the institutional arrangements are constant every round, depending only on model parameters. By contrast, the individual's share of the common-pool resource, the benefit of monitoring, and the cost of free-riding are dynamic variables that depend on the values of the model state variables $n_{\mathrm{C}}(t)$, $n_{\mathrm{M}}(t)$ and $n_\tau(t)$ during that round.

We are interested in the conditions under which agents will create a system of monitoring that incentivises cooperation, i.e. that makes the cost of provisioning less than the cost of freeriding ($C_{\mathrm{C}} < C_{\mathrm{F}}(t)$). To determine this, we consider the evolution of the three agent behavioural traits ι_{C}, ι_{M}, and ι_τ when agents with those traits are in competition with each other (Maynard Smith, 1982).

An EGT analysis considers that there are eight possible types of agents depending on the values of their ι traits, and tracks the frequency of each type in the population. All agents with the same type are assumed to have the same utility. Specifically, we take an agent type's utility in round t from Equation 1 as the fitness of that type of agent in generation t, i.e. one round corresponds to one generation. The frequency of an agent type in the next generation is then proportional to its fitness (i.e. fitness proportionate selection), as described by the standard replicator equation (see, e.g. Maynard Smith, 1982).

However, direct analysis by means of the replicator equation is complicated because of the large number of types, and the possible effects of covariance between the different traits. To ease analysis, we therefore consider each trait independently, asking when an agent will gain fitness by switching the corresponding ι value from 0 to 1 (or vice versa).

Importantly, an EGT analysis does not assume genetic transmission of traits. Rather, it can be used to capture social learning where agents imitate the traits of other agents, and are more likely to imitate traits that they observe to bring a higher payoff – so-called payoff-biased social learning (Boyd & Richerson, 1985; Cavalli-Sforza & Feldman, 1981). We proceed by analysing the above equations to determine inequalities capturing the conditions under which provisioning is favoured (i.e. individuals evolve an ι_{C} value of 1).

5.2. Variant 1: Individuals make a unilateral decision about whether to contribute to a separate pool of monitoring fees

In the first variant of the model, monitors take their payment from the separate pool of institutional taxes paid by agents with $\iota_\tau = 1$. Specifically, B_{M} is computed as:[3]

$$B_{\mathrm{M}} = \frac{\beta n_\tau C_\tau}{n_{\mathrm{M}}}, \tag{5}$$

where β is the proportion of institutional taxes that are invested into monitoring and n_τ is the number of agents that pay the institutional taxes (i.e that have $\iota_\tau = 1$).

This model represents each agent making a unilateral decision about whether to make a separate contribution to sustain implementation of the institutional rules or not, in a manner similar to pool punishment models studied in evolutionary biology (Sigmund, De Silva, Traulsen & Hauert, 2010; Sigmund, Hauert, Traulsen & Silva, 2011; Traulsen, Röhl & Milinski, 2012).

[3]For the purpose of analysis, we do not use the time step in the remainder of the paper. As we are not interested in the evolution over time, but the analysis at a given moment in time, this makes the expressions easier to read.

The first question that we can ask from our value-based EGT model is: when does it pay an agent to cooperate, i.e. when will the fitness (utility) of an agent be greater if they cooperate than if they do not? In other words, when is cooperation incentivised, such that agents with $\iota_{\mathrm{C}} = 1$ outcompete agents with $\iota_{\mathrm{C}} = 0$? Cooperation will be incentivised when the cost of cooperating is less than the cost of free-riding, i.e. $C_{\mathrm{C}} < C_{\mathrm{F}}$. This occurs when $C_{\mathrm{C}} < \frac{psn_{\mathrm{M}}}{n}$, which entails that the proportion of monitors must satisfy the inequality:

$$\frac{n_{\mathrm{M}}}{n} > \frac{C_{\mathrm{c}}}{ps}. \tag{6}$$

We can see from this that increasing C_{C} will increase the number of monitors that are necessary to incentivise cooperation, while increasing either the number of agents that each monitor monitors for rule compliance (p) or the sanction imposed on a free-riding agent when they are monitored (s) will decrease the number of monitors that are necessary. As such, the value-based model makes clear and precise predictions about the amount of monitoring that is necessary. This is in contrast to executable content-based models of institutions (e.g. Pitt & Schaumeier, 2012), where large numbers of experiments have to be run to attempt to derive such inequalities by brute force numeric searching of the effects of model parameter values.

The next question that is important to ask is: when will this level of monitoring be sufficiently incentivised, such that it individually pays all of these agents (n_{M}) to take on the monitoring role? Performing monitoring will be advantageous for an agent when $B_{\mathrm{M}} > C_{\mathrm{M}}$, that is when $\frac{\beta n_\tau C_\tau}{n_{\mathrm{M}}} > p\delta$. We can rearrange this to highlight the relationship between the frequency of tax payers and the frequency of monitors:

$$\frac{n_\tau}{n}\frac{\beta C_\tau}{p\delta} > \frac{n_{\mathrm{M}}}{n}. \tag{7}$$

This means that, to incentivise monitoring, the frequency of tax payers multiplied by the amount $\frac{\beta C_\tau}{p\delta}$ needs to be greater than the frequency of monitors. If this amount is less than 1 – assuming that all of the agents are self-interested – then not every tax payer can be a monitor. Although there are possibilities to make this amount larger than 1 (i.e. by setting C_τ to a large value or having a low cost of monitoring δ), we are most interested in the case when this is less than 1, because then there is a decision to be made, whether to monitor or not.

As inequalities 6 and 7 are both expressed in terms of the proportion of agents in the population performing monitoring, we can combine them to obtain the inequality $\frac{n_\tau}{n}\frac{\beta C_\tau}{p\delta} > \frac{C_\mathrm{C}}{ps}$ that must hold irrespective of the value of $\frac{n_\mathrm{M}}{n}$. By rearranging, we obtain

$$\frac{\delta}{s} < \frac{1}{n}\frac{\beta n_\tau C_\tau}{C_\mathrm{C}}. \tag{8}$$

So, the ratio of the monitoring cost (δ) to the sanction for free-riding (s) needs to be less than the ratio of one agent's share of the monitoring pool tax ($\frac{1}{n}\beta n_\tau C_\tau$) to the cost of cooperation (C_C). Of these, s, β and C_τ are likely to be at least partly under the control of the agents themselves, i.e. they are institutional facts. Choosing values for these accordingly ensures that it pays for n_M agents to do monitoring.

Finally, we need to examine the incentives to pay the institutional taxes, which in turn pay for some agents to monitor by providing βC_τ units of resource for monitoring. Exactly as for our analysis for the traits ι_C and ι_M, for tax paying to be incentivised, the cost of the tax needs to be less than the benefit to the individual agent of paying the tax. But we can see from Equation 1 that there is no individual benefit to paying the tax, i.e. there is no B_τ term. The benefits of paying tax are manifest through their use in incentivising monitoring and hence cooperation. But these benefits are shared equally with all of the agents, since the common-pool resource that is provisioned through cooperation is shared equally by all agents (Equation 2). Therefore, in this model self-interested agents will not pay institutional taxes, which means that there will be no resources invested into monitoring, and hence self-interested agents will not monitor. Then, in the absence of monitoring self-interested agents will not cooperate. In other words, monitoring itself becomes subject to a second-order tragedy of the commons (Axelrod, 1986; Boyd & Richerson, 1992; Fowler, 2005; Perc, 2012).

This problem is clearly highlighted by the equations of this model, since it is specified in terms of the value of each strategy. This shows that monitoring cannot be favoured *for any set of parameters*. Relying solely instead on a model that captured the content of behaviours, and not their value, would mean that an exhaustive search of parameter settings would need to be carried out in order to be sure that the lack of monitoring and cooperation was not an artefact of the particular parameter values chosen.

We now turn to investigate other ways in which monitoring can be incentivised.

5.3. Variant 2: Monitoring is paid for from the common-pool resource

In this variant, monitors take their payment directly from the common-pool resource according to a parameter β, which represents the proportion of the group's common-pool resource that is invested into monitoring. This is an institutional fact, i.e. part of the institutional rules. It corresponds more closely to several of the empirical examples given by Ostrom (1990), where agents use their common resources to either hire monitors that are accountable to themselves, or to reward certain group members for taking on the monitoring role. This involves agents making a collective decision about how much their group invests into monitoring (Conradt & List, 2009; Conradt & Roper, 2003), in contrast to the unilateral decision in Variant 1 of the model. In other words, agents play a political game in which they bargain and negotiate over the institutional rules and how to enforce them (Hurwicz, 1996; Reiter, 1996). This political game would result in setting the value of β in our model. The evolutionary dynamics of individual agent preferences for the value of β have been studied elsewhere (Powers, 2018; Powers & Lehmann, 2013). Here we do not consider the dynamics of exactly how β is set by a political game, but we instead focus on the effects of β on the level of monitoring that is incentivised.

In Variant 2 of the model, C_τ is set to 0, since monitoring is now paid for from the common-pool resource. This means that we no longer have to consider the evolution of ι_τ (it is a neutral trait). Since a fraction β of the common-pool resource is now used to pay for monitoring, the remaining fraction $1-\beta$ is distributed amongst all of the agents. Thus Equation 2 becomes:

$$B_{\mathrm{G}}(t) = (1-\beta)\frac{1}{n} \times \alpha\, n_{\mathrm{C}}(t) \tag{9}$$

The individual benefit of monitoring, B_{M}, is then computed as:

$$B_{\mathrm{M}} = \frac{\alpha\beta n_{\mathrm{C}}}{n_{\mathrm{M}}}. \tag{10}$$

The inequality for cooperation to be favoured remains the same as in Variant 1, i.e. $\frac{n_{\mathrm{M}}}{n} > \frac{C_{\mathrm{c}}}{ps}$. Monitoring, however, will now be incentivised when $\frac{\alpha\beta n_{\mathrm{C}}}{n_{\mathrm{M}}} > p\delta$. Rearrangement of this highlights the relationship between the frequency of monitors and the frequency of cooperation that is necessary to provide a sufficient amount of common-pool resource to pay for these monitors:

$$\frac{n_{\mathrm{M}}}{n} < \frac{n_{\mathrm{C}}}{n}\frac{\alpha\beta}{p\delta}. \tag{11}$$

From this we can draw out the roles of the parameters α, β, p and δ.

For full cooperation (i.e. every agent cooperates) to be an equilibrium, the largest frequency of monitors for which monitoring is individually incentivised in (11) needs to be greater than the frequency of monitors that is necessary to sustain full cooperation. This means that the following condition must hold, based on (11) and (6):

$$\frac{n_{\mathrm{C}}}{n} > \frac{C_{\mathrm{C}}}{s}\frac{\delta}{\alpha\beta} \tag{12}$$

where by setting $\frac{n_{\mathrm{C}}}{n} = 1$ we obtain:

$$\frac{\alpha\beta}{\delta} > \frac{C_{\mathrm{c}}}{s}. \tag{13}$$

The parameter p, the number of agents monitored by each monitor, appears on the denominator of both sides and so cancels out. This is a result of the assumption that monitors sample agents to monitor without replacement (Equation 3), and so doubling p means that half as many monitors are needed to sample the same number of agents.

When the relationship among the parameters in (13) holds then full cooperation will be an equilibrium. At this equilibrium, monitoring will go to the maximum frequency at which it is incentivised, which is when $\frac{n_{\mathrm{M}}}{n} = \frac{\alpha\beta}{p\delta}$ (subject to the constraint that $\frac{n_{\mathrm{M}}}{n}$ cannot exceed 1). When $\alpha\beta < p\delta$ then this will be less than 1, and so selection on ι_{M} will depend on the frequency of monitoring already in the population, leading to an interior equilibrium for the frequency of monitoring. Conversely, when $\alpha\beta \geq p\delta$ and (13) hold then there is an equilibrium in which every agent cooperates and every agent monitors. From (11) it follows that if $\alpha\beta < p\delta$ then monitoring and cooperation cannot be linked (or the same) traits, since the number of incentivised monitors is less than the number of incentivised cooperators by a fraction $\frac{\alpha\beta}{p\delta}$. Thus if we force every cooperator to monitor then self-interested agents will neither cooperate or monitor if $\alpha\beta < p\delta$. To promote cooperation in this kind of environment we should not, therefore, promote a policy in which every agent should both cooperate and monitor. This is in contrast to the findings of models of "peer punishment", where each agent makes a unilateral decision about whether or not to monitor and punish other agents and pays a unilateral cost for doing so. In these models, monitoring and punishment are promoted if cooperation and monitoring are linked traits, such that agents copy them as a pair (Boyd & Richerson, 1992; Lehmann et al., 2007). Thus, changing from unilateral to collective decision-making about how much to invest into monitoring changes whether or not we should try to force all agents to monitor, or only a subset.

We can now ask, what is the minimum value of β necessary to make full cooperation an equilibrium? This can be derived from rearranging (13):

$$\beta > \frac{\delta C_{\mathrm{C}}}{\alpha s}. \tag{14}$$

When this inequality holds, and the agents are all cooperating, then a sufficient level of monitoring is incentivised to maintain full cooperation. This allows us to answer the important practical question: how much of their resources should a group invest into monitoring? The proportion of their common-pool resources, β, that they should invest in order to maintain cooperation is the smallest value that satisfies (14). Investing any more than this is wasteful. This highlights how value-based models can produce precise predictions about how to control a system.

So far our analysis has focussed on the conditions under which full cooperation will be an equilibrium. However, a separate question is under what conditions a group of agents will reach this equilibrium if they start out with no cooperation and no monitoring. We first ask what frequency of monitoring is necessary to incentivise cooperation when there are no cooperators in the group? From the previous results this is given by (6), which is independent of the frequency of cooperators. We then need to ask if this level of monitoring is incentivised when there are no cooperators in the population. Monitoring is incentivised when condition (11) is met. We can immediately see that this cannot be satisfied when $\frac{n_C}{n} = 0$, i.e. when no agents are currently cooperating. Consequently, there is also an equilibrium in which no agent monitors and no agent cooperates (6 and 11).

This equilibrium in which no agent cooperates or monitors represents a natural starting point when considering the origin of institutions. How, then, might a group break free from this equilibrium and move to the cooperate and monitor equilibrium that increases social welfare? Moving away from this equilibrium will initially require some agents to monitor for free, i.e. to discount the cost of monitoring in their utility functions. The critical fraction of monitors to select for an increase in cooperation is $\frac{C_C}{ps}$ (as per (6)). Therefore, initially at least this proportion of agents needs to start monitoring while ignoring the costs. Then, as some agents start to cooperate, the cost of monitoring will start to be repaid. For a given non-zero frequency of cooperators, a greater frequency of monitoring costs will be repaid when the proportion of CPR used for monitoring (β) is greater. This suggests that in order to reduce the amount of "charity" that monitors must initially perform, a group should initially set its β to a large value. This can then be reduced down to that given by (14) once full cooperation is reached (see also Chen, Sasaki, Brännström & Dieckmann, 2015 for a similar argument concerning switching from rewards to punishments in order to incentive cooperation more efficiently once cooperation becomes common).

The frequency of cooperation required to fully pay for monitoring is given by the limit in (12), $\frac{n_C}{n} = \frac{\delta}{\alpha\beta}\frac{C_C}{s}$. When this is less than 1 then there will be excess funds available at the full cooperation equilibrium that can be used to reimburse agents that initially suffered a cost for their monitoring, such that they do not pay a net lifetime cost even if they initially perform monitoring for free. This effect could likely be captured to some extent in a value-based model using strategies that make a commitment (Han, Pereira & Lenaerts, 2017), or that incorporate reinforcement learning of payoffs from imagined actions (Dridi & Lehmann, 2014). However, because we are describing a behaviour that requires agents to be forward-looking to some degree, it cannot be fully captured in an evolutionary game theory model where individuals' cognition is completely myopic. It could, however, be explored easily in an executable content-based model that implements cognitive theories of agent behaviour.

6. Recommendations

In this section, we offer some recommendations arising from this study, aimed at those using modelling approaches to understand, control and design social and socio-technical systems.

6.1. Recommendation 1: Use both content-based and value-based approaches

In this study, we have shown that existing results, found in the literature, and obtained from ABM techniques, can be complemented with those obtained by taking an EGT approach. The results in Section 5 would have been difficult to obtain empirically. However, the EGT approach would also struggle to obtain results concerning the interactions of more complex cognitive agent behaviours, such as those associated with richer human social interactions. Therefore, in order to benefit from the complementarity that each provides, our first recommendation is to use both content-based and value-based modelling approaches to build understanding of a social or socio-technical system.

The risk associated with not doing this is that it is easy otherwise for the resulting understanding to be limited by the assumptions present in one modelling form. By taking only a content-based approach, it is unlikely that the modeller will arrive at statements concerning the utility of a particular behaviour in a particular context, even though these might in some cases be quite obvious, once considered. Conversely, taking a purely value-based approach may discourage consideration of the effect of interacting cognitive agents. As an example of the latter, Ostrom's work highlights that, while the Tragedy of the Commons is a perfectly valid result given the assumed game rules and behaviour model, humans in practice are able to reflect on this situation, and put measures in place to change the rules of the game. Such a solution does not naturally arise within, say, a purely game theoretic framing of the problem. However, by viewing the problem from multiple theoretical standpoints, assumptions become more apparent and therefore open to challenge.

6.2. Recommendation 2: Don't worry if the value-based model is not complete

It is tempting to think that, unless one has a complete value-based model of the system, any model that has been produced would have limited value. While it is true that we can obtain more complete results from more complete models, even partial value-based models can expose inequalities that provide valuable insight.

For example, a more complete analysis of the common-pool resource allocation problem studied in this article would consider selection on both cooperation and monitoring at the same time, and arrive at statements accounting for the co-variance between them. However, even without doing this, we have been able to arrive at useful analytical results that provide insight beyond what was readily obtainable using agent-based methods.

6.3. Recommendation 3: Go for the qualitatively equivalent, but more tractable alternative

Often seemingly innocuous changes to a model can drastically change the tractability of value-based models. An example is provided by the assumption here that each monitor perfectly coordinates to monitor a non-overlapping set of agents, i.e. that sampling from the pool of agents to be monitored is without replacement. An alternative would be to assume that this sampling is with replacement, so that different monitors may end up monitoring the same agent in the same round, because their monitoring actions are uncoordinated. This then means that the proportion of agents monitored does not increase linearly with the number of monitors, but instead increasing the number of monitors produces diminishing marginal returns in terms of the proportion of agents covered.

This assumption would be operationalised in the model by changing Equation 3 to $C_{\mathrm{F}}(t) = [1-(1-\frac{p}{n})^{n_{\mathrm{M}}(t)}]s$. This would leave our results concerning the number of agents that are incentivised to monitor (inequality 11) unchanged. However, it would change the level of monitoring that is necessary to incentivise cooperation (inequality 6). But presenting this revised inequality in an intuitive form in terms of $\frac{n_{\mathrm{M}}}{n}$ is now much more difficult. Consequently, it is much harder to gain insight into how cooperation is likely to change with investment into monitoring, and harder to gain insight into the conditions under which cooperation and monitoring can become established in a group.

In reality, groups are likely to lie somewhere on a continuum between perfectly coordinated monitoring and completely uncoordinated monitoring, with their position depending on the monitoring technology available to them. This suggests that assuming perfectly coordinated monitoring, as in Equation 3, is as reasonable as assuming completely uncoordinated monitoring, but has the crucial advantage of providing intuitive insight. More generally, it is often possible to tweak model assumptions such that the qualitative insight of the model is still valid, but the analysis is both more tractable and more intuitive.

7. Discussion

In this article we have demonstrated and explored the complementarity of ABM and EGT modelling approaches for social and socio-technical systems, which we characterised as instances of *content-based* and *value-based* approaches, respectively. We have shown that each approach brings with it different assumptions, and also offers the potential for different insights, and hence both provide value.

7.1. Implications for enduring institutions

Our results suggest that how agents decide on the amount that their group should invest into monitoring is critical to whether or not a sufficient investment to promote cooperation will be achieved. If each agent makes a completely unilateral decision about the amount of its resources to invest, then the model predicts that agents are unlikely to produce a sufficient investment. This accords with the findings of peer-punishment (Boyd & Richerson, 1992; Lehmann et al., 2007) and pool-punishment (Perc, 2012; Sigmund et al., 2010, 2011) models from evolutionary biology. Various suggestions have been made to overcome this problem, including punishment of individuals that do not invest into monitoring (i.e. second-order punishment Axelrod, 1986; Boyd & Richerson, 1992; Perc, 2012), signalling an intention to punish beforehand (Boyd, Gintis & Bowles, 2010), and the proposition that agents do conformity-biased social learning and so will tend to imitate behaviour to invest into monitoring when the majority of other agents are already investing (Boyd et al., 2003).

There is likely to be some element of conformity bias in human groups (but see also Binmore, 2005; Burton-Chellew, Mouden & West, 2017; Burton-Chellew, Nax & West, 2015; Lamba, 2014; Lamba & Mace, 2011 for critiques of experiments that argue for conformity in collective action situations). However, field studies suggest that real collective-action problems tend to be solved by the creation of institutional rules that promote cooperation and monitoring (Ostrom, 1990). These often involve groups making a collective decision to invest a share of their common-pool resources to either hire monitors, or to incentivise group members themselves to act as monitors. Where this occurs, then our micro-level model demonstrates that sufficient monitoring can be incentivised (Variant 2), in contrast to the case where the decision is unilateral (Variant 1).

By explicitly modelling incentivisation using EGT, we can make a precise prediction about the proportion of its resources that a group should invest into monitoring (relation 14). Furthermore, the model suggests that a group should invest more into monitoring when an institution is trying to become established from an initial state with little cooperation. Finally, the model predicts that cooperation will not become established unless some agents initially monitor "for free", i.e. discount the cost of monitoring in their utility function. Then, as cooperation starts to become established more and more of this monitoring will become incentivised. Moreover, we showed conditions under which when agents are at the full cooperation equilibrium then there is sufficient investment into monitoring not only to pay for monitoring at that time, but to reimburse agents that initially monitored for free, such that they do not pay a lifetime cost for this.

7.2. Implications for modelling social and socio-technical systems

Figure 1 illustrated that, even within the family of value-based approaches, there are a variety of alternatives available, and the potential for others to be developed. Exploring these, and characterising the assumptions between classic game theory and evolutionary game theory, we can see two ends of a possible spectrum where different levels of (bounded) rationality are captured. Game theory, in its various forms, provides a natural way to examine issues related to the incentivisation of behaviour. It has proven to be useful across both the natural and social sciences, from biology through to anthropology, sociology, economics and computer science. A key point is what the various types of game theory assume about the cognition of agents.

On the one hand, classic game theory assumes that agents are both rational and fully forward-looking, being able to work out the consequences of their actions for an infinite number of rounds in the future. It is recognised that neither human nor artificial agents have the computational power or sufficient information about the consequences of their actions to do this. Consequently, this assumption has been relaxed to some extent with models of bounded rationality (Gigerenzer & Selten, 2001). On the other hand, EGT assumes that agents are completely myopic, only caring about their payoff in that "generation" (Maynard Smith, 1982). For this reason, EGT is often seen as a safe minimal assumption to make about the cognition of agents. Most formal models of cultural evolution theory also rest on this assumption of myopia and extremely limited cognition, which they operationalise by using equations from population genetics to model the spread of cultural traits by imitation (Boyd & Richerson, 1985; Cavalli-Sforza & Feldman, 1981). Put bluntly, humans are assumed to copy others because they are unable to attempt to calculate what they should do, or it is too costly for them to do so (Richerson & Boyd, 2005).

Our model suggests that an assumption of complete myopia is problematic for explaining the origin of cooperation-promoting institutions. Our results imply that some agents initially need to take on a monitoring role while ignoring the immediate costs, since this will lead to an equilibrium where these costs can be more than repaid. But if individuals are completely myopic, monitoring will not get off the ground unless we assume forces exogenous to the model such as "stochastic shocks" that induce a proportion of agents to simultaneously start cooperating (Foster & Young, 1990), or large numbers of cooperating agents arriving from other groups (Boyd et al., 2003). While both of these forces can theoretically produce the result where agents reach the full cooperation equilibrium, they do not correspond particularly well to human behaviour in many common-pool resource situations in the field (Ostrom, 1990). Rather, they are a way of forcing equilibrium shifts into a myopic model.

This suggests that a more natural way to model the origin of cooperation-promoting institutions is needed, for example by using content-based models as a complement. Content-based models allow us to capture different theories of cognition in the agent's architecture and examine the result of interactions between agents based on those theories. For example, BDI (Rao & Georgeff, 1995), HCogAff (Sloman, 2001), ACT-R (Anderson, Bothell, Byrne, Douglass, Lebiere & Qin, 2004), SOAR (Laird, 2012), or the range of agent architectures discussed by Russell & Norvig (2010), are all viable approaches to capture bounded reasoning processes as well as, in some cases, human emotions and other qualitative states and values such as trust, fairness and justice (Pitt, 2016).

In the short term, however, we believe that it is important for modellers to provide clarity concerning whether their models either assume or explore the extent to which agents engage in cognition, or if they assume that agents simply 'behave'. This is important, because model predictions may vary drastically as a result, and thus it provides the context for any resulting insight.

Finally, another line of research would be to consider whether EGT, or other value-based approaches, can be extended to capture more complex agent behaviour, where the value of a behaviour is not readily obtainable in general. One idea could be to induce the value of behaviours empirically, perhaps as a second layer in a content-based model. How, for example, might the assumption of bounded rationality be parametrised in order to capture varying levels of agents' capacities for knowledge gathering and reasoning, with this being linked to the *value* (e.g., fitness) of carrying out such cognitive behaviour? An architectural schema or styles perspective (Lewis, Platzner, Rinner, Tørresen & Yao, 2016; Russell & Norvig, 2010; Sloman, 2001) provides one way to explore this space, and combining this with the evolution of traits may provide a way of exploring the extent to which agents faced with a social dilemma can be expected to engage in cognitive behaviour to reason through their situation.

Acknowledgements

This work was partially supported by a research grant from Edinburgh Napier University to S. T. Powers. We thank two anonymous reviewers for their feedback and insight.

References

Adami, C., Schossau, J., & Hintze, A. (2016). Evolutionary game theory using agent-based methods. *Physics of Life Reviews*, *19*, 1–26. doi:10.1016/j.plrev.2016.08.015.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*, 1036–1060.

Axelrod, R. (1984). *The Evolution of Cooperation*. NY: Basic Books, Inc.

Axelrod, R. (1986). An evolutionary approach to norms. *American Political Science Review*, *80*, 1095–1111.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*, 1390–1396.

Balke, T., De Vos, M., & Padget, J. (2013). I-ABM: combining institutional frameworks and agent-based modelling for the design of enforcement policies. *Artificial Intelligence and Law*, *21*, 371–398. doi:10.1007/s10506-013-9143-1.

Binmore, K. (2005). Economic man ? Or straw man? *Behavioral and Brain Sciences*, *28*. doi:10.1017/s0140525x05230145.

Binmore, K. (2014). Bargaining and fairness. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 10785–10788. doi:10.1073/pnas.1400819111.

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 7280–7287. doi:10.1073/pnas.082080899.

Boyd, R., Gintis, H., & Bowles, S. (2010). Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science*, *328*, 617–620. doi:10.1126/science.1183665.

Boyd, R., Gintis, H., Bowles, S., & Richerson, P. J. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 3531–3535.

Boyd, R., & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, *13*, 171–195. doi:10.1016/0162-3095(92)90032-y.

Burton-Chellew, M. N., Mouden, C. E., & West, S. A. (2017). Social learning and the demise of costly cooperation in humans. *Proceedings of the Royal Society of London B: Biological Sciences*, *284*, 20170067+. doi:10.1098/rspb.2017.0067.

Burton-Chellew, M. N., Nax, H. H., & West, S. A. (2015). Payoff-based learning explains the decline in cooperation in public goods games. *Proceedings of the Royal Society of London B: Biological Sciences*, *282*, 20142678+. doi:10.1098/rspb.2014.2678.

Casari, M. (2007). Emergence of endogenous legal institutions: Property rights and community governance in the Italian Alps. *The Journal of Economic History*, *67*, 191–226. doi:10.1017/s0022050707000071.

Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, New Jersey: Princeton University Press.

Chen, X., Sasaki, T., Brännström, A., & Dieckmann, U. (2015). First carrot, then stick: how the adaptive hybridization of incentives promotes cooperation. *Journal of The Royal Society Interface*, *12*, 20140935+. doi:10.1098/rsif.2014.0935.

Conradt, L., & List, C. (2009). Group decisions in humans and animals: A survey. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 719–742. doi:10.1098/rstb.2008.0276.

Conradt, L., & Roper, T. J. (2003). Group decision-making in animals. *Nature*, *421*, 155–158. doi:10.1038/nature01294.

Di Paolo, E. A., Noble, J., & Bullock, S. (2000). Simulation models as opaque thought experiments. In M. Bedau, J. S. McCaskill, N. H. Packard, & S. Rasmussen (Eds.), *Artificial Life VII: Proceedings of the Seventh International Conference on Artificial Life* (pp. 497–506). Cambridge, MA: MIT Press.

Dridi, S., & Lehmann, L. (2014). On learning dynamics underlying the evolution of learning rules. *Theoretical Population Biology*, *91*, 20–36. doi:10.1016/j.tpb.2013.09.003.

Epstein, J. M., & Axtell, R. (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. Cambridge, MA: MIT Press.

Foster, D., & Young, P. (1990). Stochastic evolutionary game dynamics. *Theoretical population biology*, *38*, 219–232.

Fowler, J. H. (2005). Human cooperation: Second-order free-riding problem solved? *Nature*, *437*, E8. doi:10.1038/nature04201.

Fudenberg, D., & Maskin, E. (1986). The folk theorem in repeated games with discounting or with incompete information. *Econometrica*, *54*, 533–554.

Fudenberg, D., & Tirole, J. (1991). *Game Theory*. Cambridge, MA: MIT Press.

Gavin, M. (2014). Agent-Based modeling and historical simulation. *Digital Humanities Quarterly*, *8*.

Gigerenzer, G., & Selten, R. (Eds.) (2001). *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press.

Gore, J., Youk, H., & van Oudenaarden, A. (2009). Snowdrift game dynamics and facultative cheating in yeast. *Nature*, *459*, 253–256.

Greif, A. (2006). *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*. Cambridge, UK: Cambridge University Press.

Guala, F. (2012). Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate. *Behavioral and Brain Sciences*, *35*, 1–15. doi:10.1017/s0140525x11000069.

Han, T., Pereira, L., & Lenaerts, T. (2017). Evolution of commitment and level of participation in public goods games. *Autonomous Agents and Multi-Agent Systems*, *31*, 561–583. doi:10.1007/s10458-016-9338-4.

Hardin, G. (1968). The tragedy of the commons. *Science*, *162*, 1243–1248.

Hardy, R. A., & Norgaard, J. R. (2015). Reputation in the internet black market: An empirical and theoretical analysis of the deep web. *Journal of Institutional Economics*, *12*, 515–539. doi:10.1017/s1744137415000454.

Hurwicz, L. (1996). Institutions as families of game forms. *The Japanese Economic Review*, *47*, 113–132. doi:10.1111/j.1468-5876.1996.tb00038.x.

Jaffe, K., & Zaballa, L. (2010). Co-Operative punishment cements social cohesion. *Journal of Artificial Societies and Social Simulation*, *13*. doi:10.18564/jasss.1568.

Janssen, M. A., & Ostrom, E. (2006). Empirically based, agent-based models. *Ecology and society*, *11*.

Khan, A. M., Freitag, F., & Rodrigues, L. (2015). Current trends and future directions in community edge clouds. In *2015 IEEE 4th International Conference on Cloud Networking (CloudNet)* (pp. 239–241). doi:10.1109/CloudNet.2015.7335315.

Kurka, D. B., & Pitt, J. (2017). The principled violation of policy: Norm flexibilization in open Self-Organising systems. In *2017 IEEE 2nd International Workshops on Foundations and Applications of Self* Systems (FAS*W)* (pp. 33–38). doi:10.1109/FAS-W.2017.117.

Laird, J. E. (2012). *The SOAR Cognitive Architecture*. MIT Press.

Lamba, S. (2014). Social learning in cooperative dilemmas. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20140417. doi:10.1098/rspb.2014.0417.

Lamba, S., & Mace, R. (2011). Demography and ecology drive variation in cooperation across human populations. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 14426–14430. doi:10.1073/pnas.1105186108.

Lehmann, L., Rousset, F., Roze, D., & Keller, L. (2007). Strong reciprocity or strong ferocity? A population genetic view of the evolution of altruistic punishment. *American Naturalist*, *170*, 21–36. doi:10.1086/518568.

Lewis, P. R., & Ekárt, A. (2017). Social and asocial learning in collective action problems: The rise and fall of Socially-Beneficial behaviour. In *2nd IEEE International Workshops on Foundations and Applications of Self* Systems, FAS*W 2017* (pp. 91–96).

Lewis, P. R., Platzner, M., Rinner, B., Tørresen, J., & Yao, X. (Eds.) (2016). *Self-aware Computing Systems*. Cham: Springer International Publishing. doi:10.1007/978-3-319-39675-0.

Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge University Press.

Milgrom, P. R., North, D. C., & Weingast, B. R. (1990). The role of institutions in the revival of trade: The Law Merchant, private judges, and the Champagne Fairs. *Economics & Politics*, *2*, 1–23. doi:10.1111/j.1468-0343.1990.tb00020.x.

North, D. C. (1990). *Institutions, Institutional Change and Economic Performance (Political Economy of Institutions and Decisions)*. Cambridge University Press.

Oliver, P. (1980). Rewards and punishments as selective incentives for collective action: Theoretical investigations. *American Journal of Sociology*, *85*, 1356–1375. doi:10.1086/227168.

Olson, M. (1965). *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard economic studies, v. 124 (revised ed.). Cambridge, MA: Harvard University Press.

Ostrom, E. (1990). *Governing the Commons : The Evolution of Institutions for Collective Action*. Cambridge, UK: Cambridge University Press.

Perc, M. (2012). Sustainable institutionalized punishment requires elimination of second-order free-riders. *Scientific Reports*, *2*.

Pitt, J. (2016). From trust and forgiveness to social capital and justice: Formal models of social processes in open distributed systems. In W. Reif, G. Anders, H. Seebach, J.-P. Steghöfer, E. André, J. Hähner, C. Müller-Schloer, & T. Ungerer (Eds.), *Trustworthy Open Self-Organising Systems* Autonomic Systems (pp. 185–208). Springer International Publishing. doi:10.1007/978-3-319-29201-4_7.

Pitt, J., & Schaumeier, J. (2012). Provision and appropriation of common-pool resources without full disclosure. In I. Rahwan, W. Wobcke, S. Sen, & T. Sugawara (Eds.), *PRIMA 2012: Principles and Practice of Multi-Agent Systems* (pp. 199–213). Springer Berlin Heidelberg volume 7455 of *Lecture Notes in Computer Science*. doi:10.1007/978-3-642-32729-2_14.

Pitt, J., Schaumeier, J., & Artikis, A. (2012). Axiomatization of socio-economic principles for self-organizing institutions: Concepts, experiments and challenges. *ACM Transactions On Autonomous And Adaptive Systems*, *7*. doi:10.1145/2382570.2382575.

Powers, S. T. (2018). The institutional approach for modeling the evolution of human societies. *Artificial Life*, *24*, 10–28. doi:10.1162/ARTL_a_00251. arXiv:https://doi.org/10.1162/ARTL_a_00251. PMID: 29369715.

Powers, S. T., & Lehmann, L. (2013). The co-evolution of social institutions, demography, and large-scale human cooperation. *Ecology Letters*, *16*, 1356–1364. doi:10.1111/ele.12178.

Rao, A. S., & Georgeff, M. P. (1995). BDI agents: From theory to practice. In *Proceedings of the First International Conference On Multi-agent Systems (ICMAS-95)* (pp. 312–319). AAAI.

Reiter, S. (1996). On endogenous economic regulation. *Economic Design*, *2*, 211–243. doi:10.1007/bf02499134.

Rendell, L., Fogarty, L., Hoppitt, W. J. E., Morgan, T. J. H., Webster, M. M., & Laland, K. N. (2011). Cognitive culture: Theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences*, *15*, 68–76.

Richerson, P. J., & Boyd, R. (2005). *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago: University of Chicago Press.

Russell, S. J., & Norvig, P. (2010). *Artificial Intelligence - A Modern Approach*. (3rd ed.). Pearson Education.

Sasaki, T., Brännström, A., Dieckmann, U., & Sigmund, K. (2012). The take-it-or-leave-it option allows small penalties to overcome social dilemmas. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 1165–1169. doi:10.1073/pnas.1115219109.

Sigmund, K., De Silva, H., Traulsen, A., & Hauert, C. (2010). Social learning promotes institutions for governing the commons. *Nature*, *466*, 861–863. doi:10.1038/nature09203.

Sigmund, K., Hauert, C., Traulsen, A., & Silva, H. (2011). Social control and the social contract: The emergence of sanctioning systems for collective action. *Dynamic Games and Applications*, *1*, 149–171. doi:10.1007/s13235-010-0001-4.

Sloman, A. (2001). Varieties of affect and the CogAff architecture schema. In C. Johnson (Ed.), *Proceedings of the Symposium on Emotion, Cognition, and Affective Computing AISB'01 Convention* (pp. 39–48).

Smajgl, A., Izquierdo, L. R., & Huigen, M. (2008). Modeling endogenous rule changes in an institutional context: The ADICO sequence. *Advances in Complex Systems*, *11*, 199–215. doi:10.1142/S021952590800157X.

Smajgl, A., Izquierdo, L. R., & Huigen, M. (2010). Rules, knowledge and complexity: How agents shape their institutional environment. *Journal of Modelling & Simulation of Systems*, *1*.

Torrent-Fontbona, F., López, B., Busquets, D., & Pitt, J. (2016). Self-organising energy demand allocation through canons of distributive justice in a microgrid. *Engineering Applications of Artificial Intelligence*, *52*, 108–118. doi:https://doi.org/10.1016/j.engappai.2016.02.010.

Traulsen, A., Röhl, T., & Milinski, M. (2012). An economic experiment reveals that humans prefer pool punishment to maintain the commons. *Proceedings of the Royal Society B: Biological Sciences*, *279*, rspb20120937–3721. doi:10.1098/rspb.2012.0937.

Weissing, F., & Ostrom, E. (2000). Irrigation institutions and the games irrigators play: Rules enforcement in government- and Farmer-Managed systems. In *Polycentric Games and Institutions: Readings from the Workshop in Political Theory and Policy Analysis* (pp. 366–398). Ann Arbor: University of Michigan Press.