Reciprocity and inequality in social dilemmas

by

Valentin Hübner

June, 2025

A thesis submitted to the
Graduate School
of the
Institute of Science and Technology Austria
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Committee in charge:

Sandra Siegert, Chair
Krishnendu Chatterjee
Krzysztof Pietrzak
Alexander Stewart

ISTA Institute of Science and Technology Austria

The thesis of Valentin Hübner, titled *Reciprocity and inequality in social dilemmas*, is approved by:

Supervisor: Krishnendu Chatterjee, ISTA, Klosterneuburg, Austria

Signature: ______________________________

Committee Member: Krzysztof Pietrzak, ISTA, Klosterneuburg, Austria

Signature: ______________________________

Committee Member: Alexander Stewart, Indiana University Bloomington, USA

Signature: ______________________________

Defense Chair: Sandra Siegert, ISTA, Klosterneuburg, Austria

Signature: ______________________________

Signed page is on file

ISTA Thesis, ISSN: 2663-337X

I hereby declare that this thesis is my own work and that it does not contain other people's work without this being so stated; this thesis does not contain my previous work without this being stated, and the bibliography contains all the literature that I used in writing the dissertation.

I accept full responsibility for the content and factual accuracy of this work, including the data and their analysis and presentation, and the text and citation of other work.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee, and that this thesis has not been submitted for a higher degree to any other university or institution.

I certify that any republication of materials presented in this thesis has been approved by the relevant publishers and co-authors.

Signature: ____________________________

Valentin Hübner
June, 2025

Signed page is on file

Abstract

Cooperation, that is, one person paying a cost for another's benefit, is a fundamental principle without which no form of society could exist. The extent to which humans cooperate with each other is also an essential feature that differentiates them from other animals. Cooperation occurs even in the absence of altruistic motivations, when it is selfishly incentivised by the expectation of a future reward. For example, many economic interactions are well described that way. This kind of cooperation requires that people exhibit reciprocal behaviour that acts as a mechanism that rewards cooperation.

With game-theoretic models, it is possible to formally study potential such mechanisms and under what conditions they can exist. This thesis contributes to this effort by analysing recently introduced models of cooperation that advance on previous work by taking into account the potential for pre-existing inequality among cooperating individuals as well as the different forms that reciprocity can take.

Individuals may differ both intrinsically, in their abilities, as well as extrinsically, in the amount of resources they have available. Allowing for such differences in a model of cooperation helps to understand how inequality affects the potential for, and outcomes of, cooperation among unequals. In this thesis, it is shown that in the presence of intrinsic inequality, a similar unequal distribution of resources can increase the potential for cooperation. This effect is stronger the smaller the group is in which cooperation takes place. It is also shown that under particular assumptions, if the unequal members of a group vary the size of their contributions to a cooperative effort over time, they can thereby increase their efficiency and improve the collective outcome.

Cooperative behaviour in a two-person interaction can be rewarded either by direct reciprocation whenever the same two people interact again, or indirectly by a third party who observed the interaction. In the latter case of indirect reciprocity, individuals are proximally rewarded by a good reputation, which ultimately translates to being rewarded with cooperative behaviour by others. This mechanism can enable selfishly motivated cooperation even in circumstances where individuals are unlikely to meet again, akin to how money facilitates trade. While these two forms of reciprocity have mostly been studied in isolation, this thesis analyses both direct and indirect reciprocity in a general model in order to compare their relative effectiveness under different circumstances. The contribution of this thesis is an extension of previous work regarding a specific kind of interaction, whose parameters allow for convenient mathematical analysis, to the most general set of possible interactions.

Acknowledgements

Zuallererst möchte ich meinem Betreuer, Krishnendu Chatterjee, danken. Für ihn stehen die Menschen in seiner Forschungsgruppe immer an erster Stelle. Er gibt uns die Möglichkeit, nicht nur seinem Forschungsprogramm, sondern vor allem unserer eigenen Neugierde zu folgen, und unterstützt uns auf diesem Weg.

Weiters bin ich Christian Hilbe für seine dauerhafte fachliche Unterstützung, aber auch die bisweilen nötige menschliche Ermutigung, außerordentlich dankbar. Er hat damit für mich in den letzten vier Jahren selbst aus weiter Ferne einen großen Unterschied gemacht. Seine ansteckende Begeisterung für unser Forschungsgebiet hat mir geholfen, meine Arbeit auch aus einem positiven Blickwinkel zu sehen.

Darüber hinaus danke ich Esra Ceylan, Gavin T. Rees, Jiří Srba, Laura Schmid, Manuel Staab, Marija Kleschnina, Matthias Bentert und Stefan Schmid. Es war eine große Freude mit euch zusammenzuarbeiten und von euch zu lernen!

Ein besonderer Dank geht an meine Familie, an Nyasha und an meine Freunde für ihre Unterstützung und Teilnahme.

Finanzierung. Die Forschungsarbeit für diese Dissertation wurde durch den Europäischen Forschungsrat (Finanzhilfevereinbarungen Nr. 863818 und Nr. 850529), das Forschungs- und Innovationsprogramm Horizont 2020 der Europäischen Union (Marie-Skłodowska-Curie-Finanzhilfevereinbarung Nr. 754411), den Österreichischen Wissenschaftsfonds (Finanzhilfevereinbarung DOI 10.55776/COE12), die französische Agence Nationale de la Recherche im Rahmen des Programme d'investissements d'avenir (Projektreferenz 17-EURE-0010) und die australische Regierung durch das Australian Research Council (Finanzhilfevereinbarung Nr. SR200100005, "Securing Antarctica's Environmental Future") unterstützt.

First and foremost, I want to thank my supervisor, Krishnendu Chatterjee. He is an exemplary research manager in the way that he always puts people first. Because he believes in curiosity-driven research, he gives us all the freedom we desire and all the guidance we need.

I am extremely grateful to Christian Hilbe, whose advice, mentorship and, where needed, moral support made a big difference for me throughout the last four years. If anything, it was his own enthusiasm for our field of research that made me appreciate it too.

Furthermore, I thank Esra Ceylan, Gavin T. Rees, Jiří Srba, Laura Schmid, Manuel Staab, Maria Kleshnina, Matthias Bentert, and Stefan Schmid. It was a great pleasure to work with you and learn from you!

Finally, I am immensely grateful to my family, to Nyasha, and to my friends for their presence and support.

Funding. The research for this thesis was supported by the European Research Council (grant agreements No. 863818 and No. 850529), the European Union's Horizon 2020 research and innovation programme (Marie Skłodowska-Curie grant agreement No. 754411), the Austrian Science Fund (grant DOI 10.55776/COE12), the French Agence Nationale de la Recherche under the Programme d'investissements d'avenir (project reference 17-EURE-0010) and the Australian Government through the Australian Research Council (grant No. SR200100005, "Securing Antarctica's Environmental Future").

About the Author

Valentin Hübner completed a combined MMath+BA degree at the University of Cambridge (Pembroke College). In September 2020, he joined ISTA's interdisciplinary graduate school, initially with an unspecific plan of finding a research topic in theoretical biology that would connect his mathematical training with his interest in cell biology. Ultimately however, he settled on joining the Chatterjee group to study cooperation in social dilemmas through game theory.

His contact details are on his website: `https://vmh.name/kontakt/`

List of Collaborators and Publications

This thesis is based on the following publications:

- Valentin Hübner, Manuel Staab, Christian Hilbe, Krishnendu Chatterjee, and Maria Kleshnina. Efficiency and resilience of cooperation in asymmetric social dilemmas. *Proceedings of the National Academy of Sciences*, 121(10), February 2024
- Valentin Hübner, Christian Hilbe, Manuel Staab, Maria Kleshnina, and Krishnendu Chatterjee. Time-dependent strategies in repeated asymmetric public goods games. *Dynamic Games and Applications*, February 2025
- Valentin Hübner, Laura Schmid, Christian Hilbe, and Krishnendu Chatterjee. Stable strategies of direct and indirect reciprocity across all social dilemmas. *PNAS Nexus*, May 2025

I am grateful to the co-authors for their essential contributions to our collaborative work.

Table of Contents

List of Figures

CHAPTER 1

Introduction

A few weeks after a dramatic heist took place at the art museum, a shady business tycoon receives a call. The gang of thieves are proposing to sell him a rare masterpiece—the kind that money seldom buys—, to be exchanged for a large sum in cash at a quiet location outside of the city. It is a deal too good to turn down. But can it be done? What if the thieves were to set up an ambush, hoping to take home a suitcase full of cash without even bringing the painting to the scene? And wouldn't he be better off too, no matter what the thieves are planning, to send his own men with guns instead of money? If both parties act strategically in this way, they are headed for a stalemate at best, if not a violent clash. Even though the proposed deal offers mutual benefit, the temptation to double-cross each other may yet render it impossible to fulfill.

This is a classic case of a social dilemma, where one possible outcome is preferred by all, but paradoxically, by following their selfish interests, they will nonetheless act so as to collectively produce another, inferior outcome. In other words, there exists a conflict between the collective and the individual interest. We refer to the action that produces the desirable outcome as cooperation, and to the action that produces the undesirable outcome as defection. Sometimes, people choose to cooperate in a social dilemma, while at other times, they choose to defect.

Under the assumption that individuals act selfishly as well as rationally, these situations lend themselves to formal analysis. The tool for this analysis is game theory, which is the formal science of strategic decision making. In game-theoretic analysis, one creates a formal model of a strategic interaction, which specifies the possible actions available to each individual, the outcome that each combination of actions produces, and the individual preferences regarding the possible outcomes. Chapter 2 presents a discussion of what it means to assume that individuals act selfishly and why this assumption cannot be omitted.

The first result of game theory is that when selfish, rational individuals engage in an isolated interaction that presents a social dilemma, they will always defect, to their collective disadvantage. To take the example presented initially, the thieves always profit from not bringing the painting, independently of whether or not the buyer intends to double-cross them. So, if they are rational and selfish, that is what they will do. The buyer's calculus is similar. However, another fundamental result from game theory, which is called the folk theorem of repeated games, also shows that if individuals get to interact repeatedly, then sometimes there is potential for them to cooperate, even just based on selfish motives. When and in what form such a potential exists in a given situation is the subject of active research.

Within that broad field, this thesis focusses on two main aims: The first aim is to study how inequality affects cooperation in collective-action problems. For this, we use an asymmetric version of a classical model known as the public goods game (see Section 1.1 below). The second aim relates to a theoretical framework that unifies two common models of reciprocal behaviour, called direct and indirect reciprocity. Our aim is to extend the analysis of this framework to all possible forms of social dilemmas (see Section 1.2).

1.1 The asymmetric public goods game

In economics, a non-excludable good is any good whose use cannot be restricted. Two particular forms of social dilemmas, or collective-action problems, arise in relation to non-excludable goods: The "tragedy of the commons" and the "free-rider problem". Both can be formally analysed with a game-theoretic model called the public goods game.

Firstly, when a non-excludable good is rivalrous, which means that its use by one person diminishes its value to another person, it is liable to be overused to the detriment of all. Such a non-excludable, rivalrous good is called a common good, and the problem of overuse is known as the tragedy of the commons [Har68]. For example, fish stock in international waters are a common good, and their existence is threatened by overfishing. In some cases, such problems can be avoided by privatisation or nationalisation (that is, by assigning property rights), or by state regulation.

Secondly, even when a non-excludable good is non-rivalrous, in which case it is called a public good, it may need to be provided or sustained. Examples of such public goods are clean air, public safety, or basic research. Since the provider of a public good cannot prevent anyone from using it, they cannot demand remuneration for the provision, which makes the goods liable to be underprovisioned. This is called the "free-rider problem" [Ols71]. A common solution is tax-funded provision of public goods.

In summary, both the conservation of a common good as well as the provision of a public good require cooperative collective behaviour, which is susceptible to exploitation by uncooperative individuals. While the classical solutions to these collective-action problems all rely on state enforcement, whether of regulation, property rights, or taxation, analysis of the public goods game shows that they can also be solved non-coercively through self-enforcing agreements. This is highly relevant not only but also in the context of the community of sovereign states, which is anarchically organised: It has a common interest in cooperation in areas such as climate change mitigation and other forms of environmental protection, or corporate taxation. Just like social dilemmas among humans, these social dilemmas among states can be formally analysed with game-theoretic models on the basis of the assumption that the states act selfishly and rationally.

1.1.1 State of the field

While it is given that selfish, rational agents will defect in any interaction that forms a social dilemma, a fundamental result from classical game theory states that when they face such an interaction repeatedly, there can be the potential for cooperation. Put simply, they can be in a state where they all expect each other to cooperate repeatedly, yet are prepared to switch to defection as soon as someone else deviates from the agreement. Depending on the specific parameters, such a mechanism may align the individual interest with the collective interest and thus make cooperation self-enforcing. The so-called folk theorem of repeated games states that when an interaction is repeated indefinitely, any kind of behaviour can be made self-enforcing as long as it is to everyone's benefit [Fri71, FM86a]. It is worth noting that this theory makes no prediction for what kind of self-enforcing behaviour could be expected in practice.

Upon introducing greater realism into the model by considering an interaction that is repeated for a finite amount of time, rather than indefinitely, some forms of cooperation may still be possible, while others may not be. Mathematical analysis can be used to examine the space of all patterns of behaviour that are still consistent with the assumptions of selfishness and rationality.

1.1.2 Our contributions

Normally, in the obvious interest of simplicity, individuals are given identical properties. This is also natural given that the form of cooperation that is being studied is sustained by a mechanism of reciprocal behaviour: the agents cooperate because they expect that others will treat them accordingly in the subsequent rounds. Reciprocating one another's actions is conceptually most simple when there are no intrinsic differences among individuals.

Yet there is a lot of interesting potential in exploring model variants where the players of the game are asymmetric. For example, this is again very relevant to international cooperation among sovereign states, who differ by orders of magnitude in the amount of resources at their disposal and in the degree to which they are able to influence each other's outcomes. It is such an asymmetric public goods game model that is analysed in this thesis. Specifically, we study a model first introduced by Hauser et al. [HHCN19], in which agents interact who either have unequal amounts of resources that they can contribute to a public good, or differ in how productively they can contribute their resources. The inequality in resources, which we call endowment inequality, affects how much each player can either contribute towards the public good or alternatively consume privately. The productivity inequality affects how effective their contributions to the public good are, that is, how much of a benefit they provide to the group as a whole. We do not study public goods games in which these relationships are nonlinear (i.e. with synergistic or diminishing effects), or in which the public good is divided unequally.

Given the individuals' productivities, we ask which allocation of endowments is optimal for cooperation. To this end, in Chapter 3, we consider two notions of optimality. The first notion focusses on the resilience of cooperation. The respective endowment distribution ensures that full cooperation is feasible even under the most adverse conditions. The second notion focusses on efficiency. The corresponding endowment distribution maximises group welfare. Using analytical methods, we fully characterise these two endowment distributions. This analysis reveals that both optimality notions favour some endowment inequality: more productive players ought to get higher endowments. Yet the two notions disagree on how unequal endowments are supposed to be. A focus on resilience results in less inequality. With additional computational simulations, we show that the optimal endowment allocation needs to account for both the resilience and the efficiency of cooperation.

Still, while that work explores to which extent asymmetry allows for full cooperation, such that players contribute their full endowment each round, by design it is therefore limited to equilibria where individuals make the same contribution in each round. In Chapter 4 therefore, we also consider players whose contributions along the equilibrium path can change from one round to the next. We do so for three different models – one without any budget constraints, one with endowment constraints, and one in which individuals can save their current endowment to be used in subsequent rounds. In each case, we explore two key quantities: the welfare and the resource efficiency that can be achieved in equilibrium. Welfare corresponds to the sum of all players' payoffs. Resource efficiency relates this welfare to the total contributions made by the players. Compared to constant contribution sequences, we find that time-dependent contributions can improve resource efficiency across all three models. Moreover, they can improve the players' welfare in the model with savings.

With that, we provide a comprehensive analysis of the impact of two of the main simplifying assumptions of classical public goods game models.

1.2 Direct and indirect reciprocity

Reciprocity, that is, responding in kind to how one is treated by others, is a commonly observed pattern of human behaviour in social dilemmas. The presence of reciprocity as a social norm is also a possible mechanism by which cooperation can be selfishly motivated. Furthermore, reciprocal behaviour requires comparatively little cognitive capacity. For all of these reasons, reciprocity is of great interest to game theory and the study of human behaviour in general.

In its simplest form, which is called direct reciprocity, it requires that individuals can expect to interact repeatedly, such as it would be the case among the members of any sufficiently small community. Unlike in the asymmetric public goods game, here, only dyadic interactions, that is, interactions between two players, are considered, but these may still occur among pairs of players within a larger group. When reciprocal behaviour is the norm in such a group, it can be in people's selfish interest to behave cooperatively, because they know that their behaviour will likely be reciprocated in the future. In a modern society however, it is common for strangers to interact who are unlikely to meet again. There, direct reciprocity fails as a mechanism that would selfishly incentivise cooperation. A similar mechanism can still work, however, if actions are publicly observed and confer a corresponding reputation on the agent, which then affects how they can expect to be treated by third parties in the future. This is called indirect reciprocity.

1.2.1 State of the field

Both the concepts of direct [Axe81, HCN18, GV18, GK21, RH23] and indirect [NS98b, NS05, Sig12, Oka20] reciprocity have received great attention in the literature. In order to jointly analyse both of them in one formal model, which had not previously been done, Schmid et al. [SCHN21] introduced their so-called unified framework of direct and indirect reciprocity. This unified framework is a generalisation of the usual models for both kinds of reciprocity and also allows for intermediate forms. The authors identify cooperative strategies of indirect reciprocity as well as of intermediate forms of reciprocity, and characterise the conditions under which they can maintain cooperative behaviour.

Schmid et al.'s model exclusively considers the donation game, which is the two-player version of the public goods game. In the donation game, the effects of individual behaviours are independent, which is not necessarily given in a social dilemma. By its particular properties, the donation game allows for convenient mathematical analysis. However, it is not representative of the full variety of social dilemmas that may occur between two persons. Not only is the donation game a special case of the more general class of prisoner's dilemmas, but there are also other classes of social dilemmas, which are qualitatively different from the prisoner's dilemma.

The seminal work of Press and Dyson [PD12], on the other hand, considers arbitrary kinds of social dilemmas but only in the context of direct reciprocity. They show that there exists a class of strategies, called the equalizer strategies, that completely decouple one's own welfare from the actions of others. This property makes any such strategy a stable social norm, which can potentially sustain cooperation. However, a generalisation of the work of Press and Dyson to indirect reciprocity and to the more general unified framework has remained has remained open so far.

1.2.2 Our contributions

In Chapter 5, this thesis presents a model that is a version of the unified framework of direct and indirect reciprocity for all kinds of social dilemmas. We use this model to generalise the theory of equalizer strategies from direct reciprocity to the unified framework, and thus as a special case also to indirect reciprocity. Thereby, we show how individuals can sustain socially optimal outcomes across all pairwise social dilemmas, using either direct or indirect reciprocity, and arbitrary mixtures thereof. We apply novel proof techniques to overcome the greater mathematical complexity of strategic interaction in other social dilemmas compared to in the donation game. The unified framework allows us to explore how individuals can integrate social information from different sources to solve collective-action problems, and to generally build further bridges between the hitherto separate literatures on direct and indirect reciprocity.

CHAPTER 2

Evolutionary game theory and the assumption of selfishness

Two basic assumptions in game theory are that individuals are rational and selfish. This preliminary chapter presents a critical discussion of the latter. Firstly, what is meant by rationality and selfishness? The understanding of these terms in game theory is based on a model of human agency most notably tracing back to David Hume's *Treatise of Human Nature* [Hum78], in which passions and reason together determine a person's actions. The passions, which are not subject to rational scrutiny, are whatever motivates a person to act. Reason is the cognitive process in which a person chooses their actions with the aim of satisfying those passions.

What Hume called "passions" is referred to as an individual's "preferences" in game theory, which are usually assumed to be consistent. A person is called rational if they act so as to optimally satisfy their preferences, whatever those may be. These assumptions have their limitations: Neither is it really the case that preferences cannot be rationally revised, nor are they always consistent [HHV02]. Finally, it is also clear that humans do not always act rationally. There is a large body of research on systematic departures from rationality in human behaviour, which are called "cognitive biases". Even the finiteness of human cognitive capacity is on its own enough to show that humans cannot be fully rational. But instrumental rationality, as it is called, is a useful model.

A person's preferences are called selfish if they are independent of how the person's actions affect others. To understand why one would make a general assumption that that were the case, we need to differentiate between classical game theory and evolutionary game theory.

In classical game theory, one aims to analyse one particular strategic interaction as it might occur among two or more individuals, such as in the examples mentioned initially. To that end, one creates a formal model of what choices are available to each individual, what outcomes they produce in combination, and what each individual's preferences over these outcomes are. With the preferences in particular, any modelling choice is bound to be to some degree arbitrary, since it represents people's inner motivations, which are difficult to observe and to quantify. It is thus convenient to assume that individuals are purely selfishly motivated. But this remains an assumption, which is accurate in some cases and inaccurate in others.

2.1 Evolutionary game theory

In contrast to classical game theory, evolutionary game theory is an application of game-theoretic analysis to the behaviour of all kinds of living organisms, from animals all the way to microbes and even single genes and viruses. Interestingly, some of game theory's greatest scientific achievements are from this realm of evolutionary biology, even though game theory originated with the study of human behaviour through classical game theory [MS82]. That is because as it turns out, the behaviour of comparatively simple biological organisms conforms far better to the assumptions of selfishness and rationality than that of humans.

But what does it mean for a bacterium to be rational, or for a gene to be selfish? These notions rely on an analogy to human agency by which the preferences of an individual from an evolving population, say, a bacterium, are defined as the equivalent of its Darwinian fitness. By definition, these preferences are selfish (other than towards close genetic relatives when we adopt a definition of inclusive fitness [Ham64]). Now, the interesting point is that insofar as the individual's behaviour is adaptive, it is also by definition rational. The simpler the behavior of a species is, the more powerful this model becomes. In particular, unlike in the study of human behaviour through classical game theory, there is no need to make any assumptions about the preferences of the agents.

All kinds of organisms encounter their own kinds of social dilemmas. In some of those cases, cooperative behaviour can be observed. For example, the bacterial pathogen *Pseudomonas aeruginosa* produces and secretes pyoverdine, which enables it to take up iron from its host environment [MIF+16]. This is a form of cooperation, because pyoverdine is diffusible and can therefore be taken up by any bacterial individual, not just the producer [RD00], who pays the high [GWB04] metabolic cost of producing it. One might expect that the only adaptive, or selfishly rational, behaviour is not to produce pyoverdine, that is, to free-ride on the contributions of the others. But this is not what is observed. Evolutionary game theory aims to explain observations such as this by modelling mechanisms through which selfish motives can give rise to cooperation. Examples of these are kin selection [Fra98, WPG02, FWR06], reciprocity [Tri71, AH81] (which requires a degree of cognitive ability not possessed by bacteria), and group selection [RR03, TN06].

Since humans are also an evolving species [Dar09], one may in principle also attempt to study human behaviour with the methods of evolutionary game theory. However, there are some important caveats. Firstly, human behaviour is uniquely complex [VGE08]. Secondly, while there is controversy about the degree to which some traits such as personality and intelligence are genetically heritable [Rid03], it is universally accepted that the majority of human behaviour is not [BR88a, Ric05, Pin10, BRH11, Ste12, CT13]. Besides a component of innate behaviour, which is called instinctive, human behaviour is largely determined by cognitive processes which form and change a person's behaviour throughout their lifetime, such as learning and reflection. It therefore cannot be the case that it is adaptive as a result of Darwinian evolution.

For some examples, consider dietary choices: (a) Spoilt food, which may carry pathogens, commonly elicits disgust, which is partly an instinctive reaction [CB01]. (b) An estimated 1.5 billion Muslims fast during Ramadan, motivated by a religious commandment [Pew12]. (c) Around 1% of the German population is estimated to follow a vegan diet [MBB16]. Of these, about 90% report to be motivated by animal welfare, and 47% by a desire to reduce their negative impact on the environment [JBRH16]. (d) The average consumption of sugar in Canada is about twice that of Japan, despite similar GDP (PPP) per capita and HDI [FAO24]. All of the above choices can impact an individual's fitness [MBB16, HMD+11]. Yet only the first example, disgust, is partly instinctive and thus a candidate for a rigorous explanation in terms of genetic evolution.

2.2 Learning

The fact that human behaviour is largely not innate seems to preclude the application of evolutionary game theory to most aspects of human behaviour. In order to circumvent this problem, mathematical models of learning, reflection, and other cognitive processes have been proposed in which humans are posited to continuously update their behaviour, for example by imitating the behaviour of others who are seen as successful, or by evaluating one's current behaviour against hypothetical alternatives through counterfactual reasoning [SBM+09]. According to this theory, given any new situation, humans learn and use reason to behave in the interest of their own reproductive fitness [Ale17].

This theory has multiple problems. Firstly, whatever processes are assumed to operate at the cognitive level (whether conscious or unconscious) to optimise human behaviour for reproductive fitness as an extension of genetic evolution cannot do so directly, but must instead optimise for a proxy measure that is cognitively represented while still being genetically determined. That is, they must optimise for alignment with instinctive desires. One should therefore expect that altruistic behaviour evolves only at the cognitive level, while the instinctive desires are completely selfish. Otherwise, if a general concern for the welfare of others were simply part of human nature, what would stop the processes of learning and reflection from producing behaviour by which one would indeed sometimes compromise one's own fitness for nothing but the benefit of others?

However, altruism is in fact partly innate. Neurological research shows that empathy is a natural human function [Ado99, PdW02, DVS06]. In other words, humans can feel joy as a result of perceiving others to be joyful, or pain as a result of perceiving others to be in pain, which motivates altruistic behaviour, and this phenomenon is (at least partly) innate [Slo17]. This aligns with the presence of altruism in animals whose behaviour is essentially determined by instinct, and may be seen as a parallel to altruistic behaviour in species without cognitive ability. While there is debate over the mechanism by which such an altruistic instinct may have evolved, it is acknowledged that this aspect of human nature is a necessary component in explaining human behaviour on the basis of its motivations [RN13].

As a second problem, empirical evidence also suggests that the full scope of human cooperation is inconsistent with the assumption of fitness-adaptivity. This is not a self-evident fact, because the majority of human cooperation is indeed based on selfish, not altruistic motivation. Take the example of an artisanal baker. She bakes bread not just for herself, but for hundreds of people every day. She does this not primarily out of compassion for those people's desire to eat good bread, but rather because she receives a monetary reward in return. This is so obvious that we do not consider it to be altruistic behaviour. Starting from observations like this, it is tempting to conjecture that in fact all forms of human cooperation can be explained in a similar fashion: we may aim to explain helpful acts by a desire for them to be reciprocated, and even generosity towards strangers by a desire for the eventual benefits of a good reputation.

Some of these explanations will be accurate, and others will not be. Given the complexity of human social life, it is hard to test them. However, behavioural experiments have shown that the general hypothesis, that all of human cooperation is motivated in this way, is false. A well-known example is the so-called dictator game, a very simple artificially created situation where one of two participants is given some amount of money and may choose freely how to divide it between themselves and another participant [KKT86]. Hoffman et al. [HMS96] conducted a double-blind dictator experiment in which the participants could make their decisions in complete and lasting anonymity from each other and the experimenters. They found that while anonymity results in more selfish behaviour, still 36% of participants gave away some share of the money they received. This behaviour can only be assumed to be to their own reproductive disadvantage. The most parsimonious explanation here is that participants were motivated by a real desire for fairness, which was present also in the absence of an expected reward for cooperative behaviour. Other experiments paint a similar picture [Cam03].

The third problem, finally, is that even if the theory of fitness-adaptive human behaviour were true, it would not be practically useful for the purpose of formally (or otherwise) studying social interactions. For example, consider a person's decision in a social dilemma, such as the decision whether or not to reduce their air travel in order to help mitigate climate change. Even if we could soundly assume that that decision will be whatever is ultimately most beneficial for the person's reproductive fitness, we still don't know what that choice should therefore be. To find out, we would have to balance the financial cost and the forgone relaxation that a trip by plane could have brought with the gain in social reputation that might be obtained by abstaining from it, all of which depend entirely on the specific circumstances of that individual person and are impossible to quantify even approximately. Trying to give a scientific answer to this question is futile.

2.3 Interpretation

So we must discard the hypothesis that human behaviour is essentially fitness-adaptive. This may be surprising, given that human behaviour did in some sense arise through genetic evolution, a process that continuously increases fitness [Fis30]. But this is the nature of emergence, of which not just the human mind, but also human society are paradigmatic examples. Just like bacterial evolution cannot be understood through the laws of chemical reactions, much of human behaviour cannot be understood through the evolutionary process, even though it emerged from it. To illustrate this fact: It is entirely possible that future human generations will start to form the genomes of their successors by synthetic design rather than through recombination of their own [Fuk03], which would set their own Darwinian fitness to zero. There is no reason to believe that human behaviour in areas other than reproduction should generally be any more restricted by evolutionary forces.

So what role should evolutionary game theory play? It is certainly an excellent tool in behavioural ecology, for the study of the simpler behaviour of non-human species, where it is very successful. In a more speculative way, it can also be used for hypothesising about the evolutionary origins of human instincts, such as, most notably, to address the question of how human altruistic instincts could have evolved.

However, when we study any specific human interaction, it is my view as argued above that we may not simply assume that behaviour is consistent with the reproductive interest.

Such an assumption is often implicitly made when real-world examples of cooperation are described a priori as surprising or paradoxical, or as being necessarily in need of explanation by a game-theoretic mechanism. In reality, any such instance of cooperation may also just be the result of existing unspecific altruistic preferences.

So, if we wish to study specific strategic human interactions (and this may include interactions among organisations as well as among individuals) we must make hypotheses about the agents' preferences in the given context and create classical game-theoretic models based on such assumptions. Wherever their predictions fail the empirical test, the hypotheses must be revised. The work in this thesis should be seen in that light as being predicated on the assumption that agents are to some degree of approximation selfishly motivated in whatever social interaction is being considered. Where that is not given, the models and results are not applicable.

CHAPTER 3

Efficiency and resilience of cooperation in asymmetric social dilemmas

This chapter was originally published as: Valentin Hübner, Manuel Staab, Christian Hilbe, Krishnendu Chatterjee, and Maria Kleshnina. Efficiency and resilience of cooperation in asymmetric social dilemmas. *Proceedings of the National Academy of Sciences*, 121(10), February 2024

Everyday life is rich in situations where individuals have to decide whether to act for the benefit of the group or their individual gain. These situations are commonly referred to as social dilemmas [KGSF04, N^{+}12]. For pro-social behaviour to be maintained in these settings, it takes some mechanism that enables individuals to overcome selfish interests [Now06b]. Direct reciprocity is one such mechanism [Tri71, Axe81, Sig10]. It requires individuals to interact repeatedly, so that previous actions may shape future decisions. As a result, even in the absence of explicit punishments, cooperation can evolve and be stable.

Most previous studies on the evolution of reciprocity focus on fully symmetric interactions [NS92, Fre94, KDK99, IN10, KWI10, GT12, PD12, VSPLS12, SP14c, TRS14, SP14a, SP15, DLTZ15, PHRZ15, BJHN16, MH16a, RHR^{+}18, IM18, HCN18, SHCN22]. In these studies, individuals are perfectly interchangeable. This assumption plays a critical role in the context of direct reciprocity, because it implies that the ability to increase or reduce an opponent's payoff is identical across individuals. If, however, individuals differ in their costs and benefits of cooperation, some individuals might be harder to discipline than others, making cooperation more difficult to sustain. This has, for instance, been shown in the context of endowment inequality [HHCN19, CKS05, MSMS21, ACBS^{+}18, BBV86, VSPL14, ACT14, Aki15, Zel03, DBF18, BC06, HRS16, GMTV17, NSRC15, Gäc15, CMMM99, KMS17]. Endowment inequality is often identified with real-world inequalities in income or wealth, which have a negative impact on social outcomes more generally [GVDDVD20, HG23, SN23, MSMP22, SMB20]. Such observations suggest that if individuals are otherwise symmetric, endowment inequality ought to be as small as possible to promote cooperation. However, in addition to endowment inequality, individuals often differ along multiple other dimensions, such as their level of skill. In such a context, recent studies suggest that a perfectly equal endowment distribution may not be optimal for cooperation either [MW22, YWH22, NAKP20]. This raises the question what the optimal level of endowment inequality is. This paper aims to provide an answer to that question.

Taking the framework of Hauser *et al.* [HHCN19] as a starting point, we study repeated linear public good games among asymmetric players. Our baseline model contains two sources of asymmetry. First, players may differ in their endowments, which influences how much they can contribute to the public good. Second, players may differ in their productivities, which influences how effective contributions are. Given the players' productivities, we ask how endowments should be optimally allocated. To tackle that question, we introduce two notions of optimality. First, we characterise the endowment distribution that results in the highest *resilience of cooperation*. Here, individuals are able to enforce cooperation even as the game's continuation probability approaches the theoretical minimum. Second, we characterise the endowment distribution that exhibits the highest *efficiency of cooperation*. This distribution maximises social welfare in the best possible equilibrium. We identify those two optimal distributions for any form of heterogeneity in individual productivities and any group size.

We find that according to both notions, more productive players ought to get higher endowments. The exact magnitude of this optimal endowment inequality, however, depends on which notion is used, and on the parameters of the game. In particular, we identify scenarios where resilience of cooperation requires endowments to be almost equal even though productivities are not. Conversely, we also describe scenarios in which minor differences in productivities result in major differences in endowments. As a general rule, we find that the efficiency-maximising endowment distribution is always more unequal than the resilience-maximising one. This suggests that there is a non-trivial trade-off between the resilience of cooperation and efficiency. To further study this trade-off, we simulate learning dynamics among interacting individuals. These simulations suggest that the endowment distribution that performs best lies on a Pareto frontier between efficiency and resilience. Where exactly that point is located depends on the chosen parameter values. When parameters are generally favourable to cooperation, payoffs are highest when endowments are close to the efficiency-maximising distribution. In contrast, in noisy environments in which cooperation is generally difficult to sustain, payoffs are higher when the endowment distribution prioritises resilience.

This work highlights how different objectives, such as efficiency or resilience, have different implications for the optimal allocation of endowments within groups. While both objectives tolerate some endowment inequality, this inequality needs to be smaller when cooperation is to be resilient.

3.1 Results

3.1.1 Model

We consider a repeated linear public good game among n players. At every time step t, each player i receives an endowment e_i. Subsequently, each player decides how much of the endowment to contribute towards the public good, $c_i(t)$, and how much to consume individually, $e_i - c_i(t)$ (Fig. 3.1a). The contribution of a player i is multiplied by a productivity factor r_i, with $1 < r_i < n$, which may differ across players (Fig. 3.1b). The

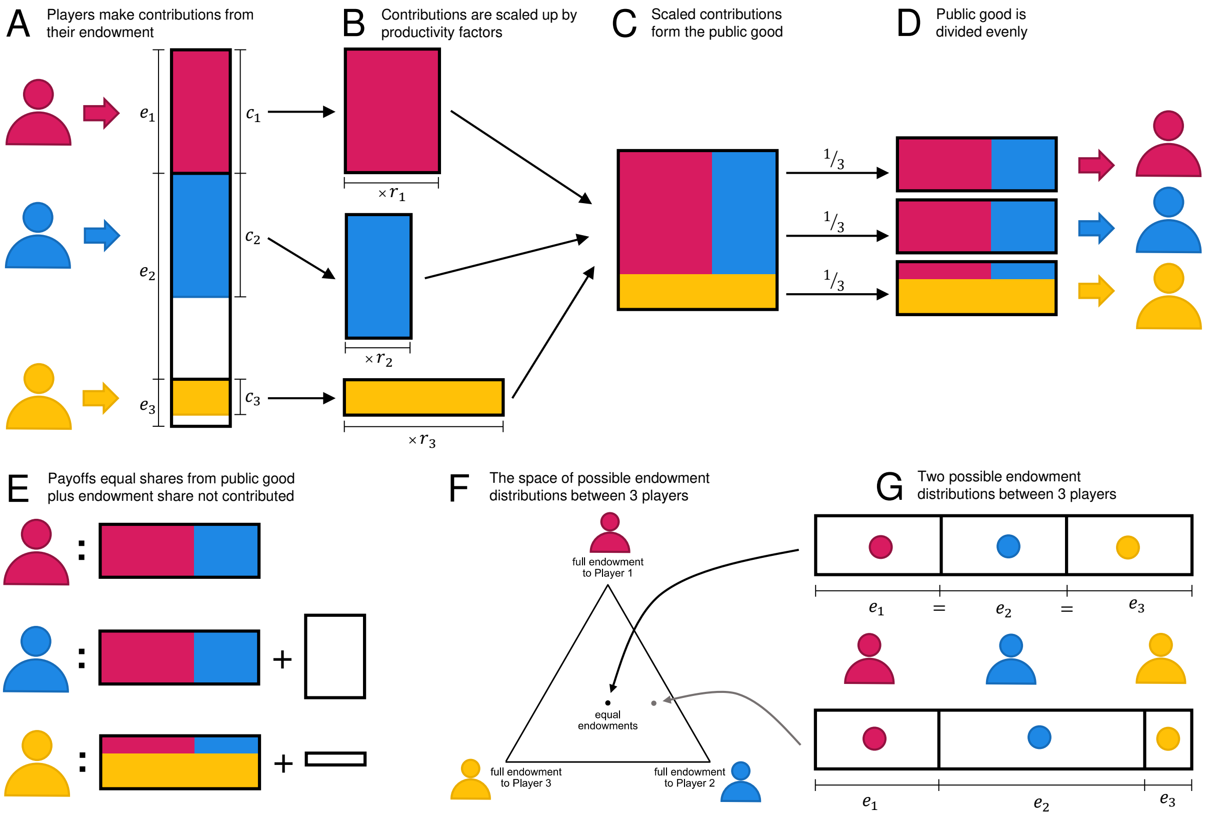

Figure 3.1: Schematic representation of the model. Players engage in a repeated linear asymmetric public good game. In every round, each player i receives an endowment e_i. A, Players choose how much to contribute towards the production of the public good, c_i, from their available endowment, e_i. B, All individual contributions are multiplied by individual productivity factors, r_i. C, The size of the public good is defined as the sum of all effective contributions. D, After its production, the public good is divided equally among all players. E, Individual payoffs are equal to the nth share of the public good plus the remaining share of the endowment that players did not contribute towards the public good. F, Without loss of generality, we assume $\sum_{i=1}^{n} e_i = 1$. In the case of a three-player game, we can represent the endowment distributions in a simplex, where each point corresponds to a vector $\mathbf{e} = (e_1, e_2, e_3)$. G, We aim to identify the optimal endowment distribution with respect to different objectives.

total amount of the public good produced equals the sum of all effective contributions $\sum_{i=1}^{n} r_i c_i(t)$. This amount is equally divided among all players, independently of their contributions (Fig. 3.1c–d). Individual payoffs $\pi_i(t)$ are determined by the quantity of the public good received, and the individually consumed shares of the endowments (Fig. 3.1e),

$$\pi_i(t) = e_i - c_i(t) + \frac{1}{n}\sum_{j=1}^{n} r_j c_j(t). \tag{3.1}$$

Without loss of generality, we assume that endowments are normalised such that $\sum_{i=1}^{n} e_i = 1$. Accordingly, we refer to the vector $\mathbf{e} = (e_1, \ldots, e_n)$ as an endowment distribution. This distribution summarises how endowments are allocated among the players (Fig. 3.1f–g). For example, the vector $\mathbf{e} = (1/n, \ldots, 1/n)$ describes an equal allocation.

If the above game is only played for a single round, full defection is the only equilibrium, for any endowment distribution. However, here we assume that after each round t, there

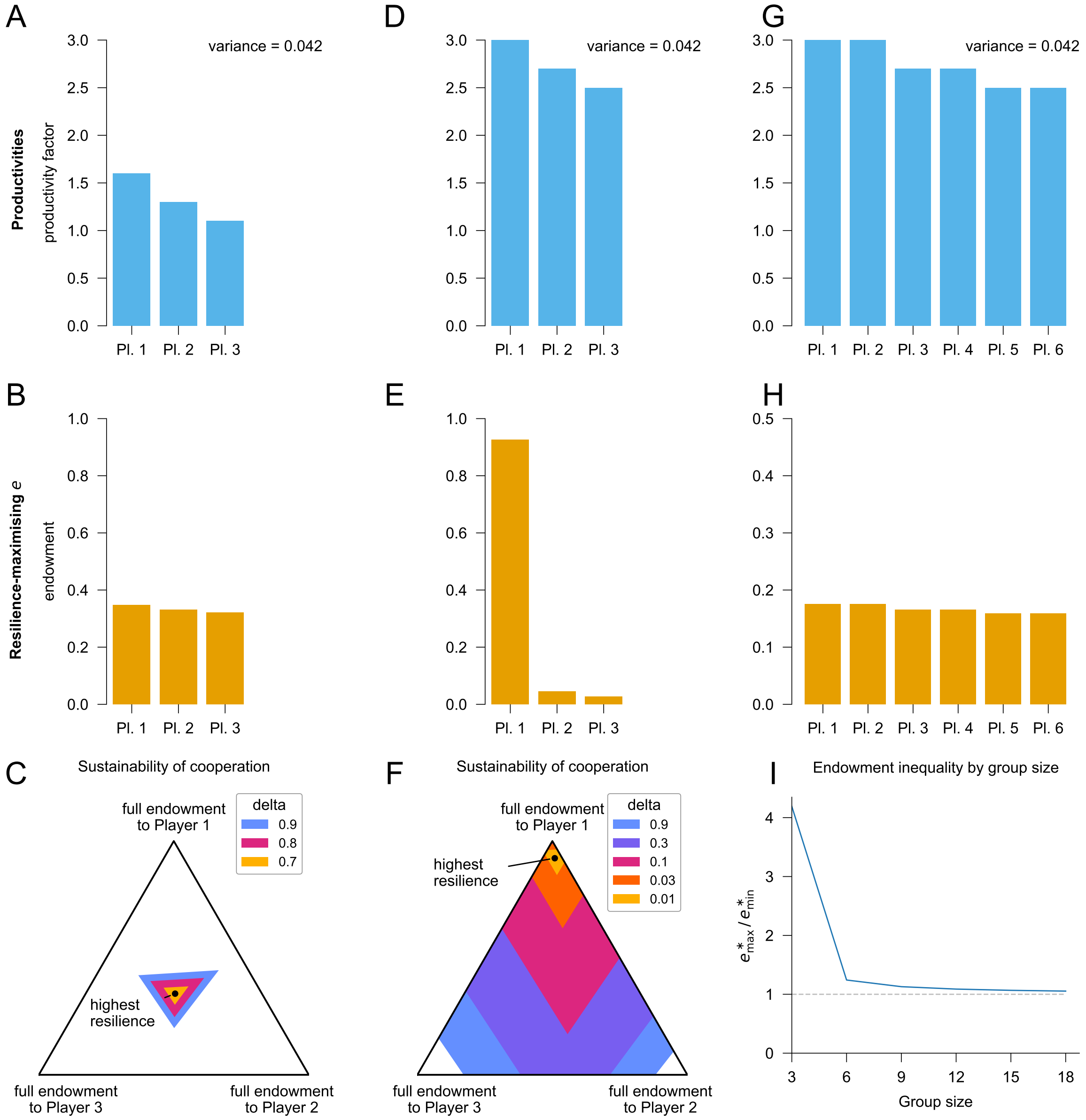

Figure 3.2: Resilience-maximising endowment distribution. A, To demonstrate the degree of inequality in the resilience-maximising endowment distribution, we construct three examples of games, all of which have the same level of heterogeneity in productivities. In this first example, three players with comparatively low productivities interact. B, When the ratio r_i/n is low, then the resilience-maximising endowment distribution is close to $(1/n, \ldots, 1/n)$. C, With productivities close to 1, cooperation is challenging: $\delta_{\min} = 0.620$. D–F, Cooperation becomes more attractive and much easier to sustain. $\delta_{\min} = 0.005$. The resulting degree of inequality of $\mathbf{e}^*$ increases. G–I, With high productivities, but more players, the endowment distribution is again close to $(1/n, \ldots, 1/n)$. The required continuation probability is now $\delta_{\min} = 0.238$, which is lower than in the previous example, but higher than in the example with low productivities. I, We demonstrate the general principle by systematically varying group size while keeping productivities fixed (up to multiplicity). We plot the degree of inequality of the resulting resilience-maximizing endowment distribution measured by the ratio between the highest and the lowest endowments in the allocation. As the ratio r_i/n decreases, the resilience-maximizing endowment allocation approaches $(1/n, \ldots, 1/n)$.

is another round with a fixed continuation probability $0 < \delta < 1$. Equivalently, one may also interpret our setup as that of a game with infinitely many rounds, and δ as the extent to which players care about their future payoffs [BDV+20]. In each case, Player i's expected payoff over all rounds, with a normalising factor of $1-\delta$, is given by

$$\hat{\pi}_i = (1-\delta)\sum_{t=0}^{\infty} \delta^t \pi_i(t). \tag{3.2}$$

For repeated games, the Folk theorem [Sel88, FT91] states that any individually rational outcome can be sustained in a Subgame Perfect Nash Equilibrium (SPNE), provided δ is sufficiently close to 1. In the following, we are particularly interested in equilibria with full cooperation, meaning that all players choose $c_i(t) = e_i$ in every single round t (By this definition, a player i who happens to get no endowment is considered fully cooperative, even though the player contributes $c_i(t)=e_i=0$ every round). In general, whether or not an equilibrium with full cooperation exists depends on how endowments are allocated. For large δ, there are generally many different endowment distributions for which a fully cooperative equilibrium exists. But as δ decreases, the respective set of endowment distributions shrinks. In that sense, endowment distributions differ in how resilient cooperation is to adverse circumstances. Moreover, due to the variation in individual productivities, different endowment distributions result in different levels of welfare even if everyone fully cooperates. We therefore investigate the resilience of cooperation and the resulting welfare implications.

More specifically, we distinguish two objectives: First, we aim to find the endowment distribution that sustains full cooperation at the lowest δ. We call this the *resilience-maximising endowment distribution*. Second, we are interested in the endowment distribution that allows for the highest group welfare in equilibrium. We refer to this as the *efficiency-maximising endowment distribution*.

3.1.2 The resilience-maximising endowment distribution

It turns out that there is a surprisingly elegant characterisation of the resilience-maximising endowment distribution. In the following, we provide a summary of our corresponding findings. All details and formal derivations are described in the Supporting Information.

First, we show that for an arbitrary (but fixed) set of productivities $\mathbf{r} = (r_1, \ldots, r_n)$, there is always an endowment distribution $\mathbf{e}$ and a continuation probability $\delta < 1$ such that full cooperation is possible. Next, if also the values of δ and $\mathbf{e}$ are fixed, we prove that full cooperation is sustainable in an SPNE exactly if

$$(\delta D - I_n)\mathbf{e} \geq \mathbf{0}. \tag{3.3}$$

Here, D is an $n \times n$ matrix with entries $D_{ij} = \frac{r_j}{n-r_i}$ for $i \neq j$ and $D_{ii} = 0$, and I_n is the $n \times n$ identity matrix. We observe that for the given endowment distribution $\mathbf{e}$, there exists a minimal continuation probability $\delta_{\min}(\mathbf{e})$ that satisfies (3.3). Hence full cooperation is sustainable in a SPNE if and only if $\delta \geq \delta_{\min}(\mathbf{e})$. Because δ can be interpreted as the patience of players, or how much they value their future payoffs, this lower bound on δ can be considered to be a measure of how hard it is to sustain cooperation with the given endowment distribution. The lower this minimum $\delta_{\min}(\mathbf{e})$, the easier it is to sustain cooperation.

Based on this observation, we define the resilience-maximising endowment distribution $\mathbf{e}^*$ to be the one with the smallest value of $\delta_{\min}(\mathbf{e})$, that is, $\mathbf{e}^* := \arg\min_{\mathbf{e}} \; \delta_{\min}(\mathbf{e})$. We use the notation $\delta^*_{\min} := \delta_{\min}(\mathbf{e}^*)$ for the corresponding minimal continuation probability. Using inequality (3.3), we can derive $\mathbf{e}^*$ and $\delta^*_{\min}$ for any number of players n and individual productivities $\mathbf{r}$. We show that $\mathbf{e}^*$ is exactly the Perron eigenvector of D, and the corresponding eigenvalue is equal to $(\delta^*_{\min})^{-1}$ (Section 3.4.3). This provides a simple method for calculating $\mathbf{e}^*$ for any set of parameters and any group size. For the special case of a two-player game, we recover Hauser *et al.*'s [HHCN19] result that the resilience-maximising endowment distribution is equal to

$$e_1^* = \frac{\sqrt{r_2(2-r_2)}}{\sqrt{r_1(2-r_1)} + \sqrt{r_2(2-r_2)}}$$

and

$$e_2^* = \frac{\sqrt{r_1(2-r_1)}}{\sqrt{r_1(2-r_1)} + \sqrt{r_2(2-r_2)}}.$$

Based on our characterisation for n players, we can derive general properties of the resilience-maximising endowment distribution $\mathbf{e}^*$. We find that the relationship between e_i^* and r_i is always order preserving. That is, more productive players always need to have a larger endowment than less productive players in order to guarantee the highest resilience of cooperation. Nonetheless, the degree of endowment inequality according to $\mathbf{e}^*$ may vary significantly. It depends on the ratio between players' individual productivities and the size of the group, r_i/n. The smaller the productivities are in relation to n, the more equal $\mathbf{e}^*$ is. In particular, fixing individual productivities at some level and increasing n results in $\mathbf{e}^*$ getting arbitrarily close to $(1/n, \ldots, 1/n)$. To see this, we note that for large n, the off-diagonal entries of the matrix D approach $D_{ij} = r_j/n$, which is independent of i. For the resulting matrix D, the uniform distribution is a right eigenvector.

We illustrate this effect in Fig. 3.2. When the players' productivities are comparably low (Fig. 3.2a), we observe a resilience-maximising endowment allocation that is approximately uniform (Fig. 3.2b). However, due to the low productivities, cooperation is not very resilient (Fig. 3.2c). Keeping the variance in productivities fixed while increasing their overall level (such that Player 1's productivity is almost equal to n, Fig. 3.2d) results in a very unequal endowment allocation. Now, player 1 receives almost all of the endowment (Fig. 3.2e). Yet, despite the high inequality in endowments, cooperation becomes more resilient (Fig. 3.2f). We further extend our argument by doubling the number of players, while keeping the productivities fixed (Fig. 3.2g). This reduces r_i/n by half. This new six-player game with identical variance in productivities again results in an almost uniform distribution $\mathbf{e}^*$ (Fig. 3.2h). The resilience of cooperation is intermediate in this case.

We can formalise this result by deriving an upper bound on the relative difference between players' endowments with respect to $\mathbf{e}^*$. This upper bound is given by

$$\frac{\max_i e_i^*}{\min_i e_i^*} \leq \frac{n-1}{n - \max_i r_i}. \tag{3.4}$$

In particular, if there is some number k such that productivities do not exceed n/k, then the absolute difference in endowments is bounded by $(\max_i e_i^* - \min_i e_i^*) \leq 1/(k-1)$.

3.1.3 Efficiency-maximising endowment distribution

The social-dilemma nature of the game implies that higher cooperation achieves greater group welfare. However, not all endowment allocations allow for full contributions in equilibrium and, due to the individual heterogeneity in productivity, full cooperation yields different levels of welfare with different endowment distributions.

We define welfare as the sum of the individual (expected) payoffs. If all players contribute fully, the group welfare can be expressed as a function of endowments:

$$\Phi(\mathbf{e}) := \sum_{i=1}^{n} \hat{\pi}_i = \sum_{i=1}^{n} r_i e_i, \tag{3.5}$$

where $\hat{\pi}_i$ is as defined in equation (3.2).

Maximisation of the group's welfare constitutes an optimisation problem of finding an endowment distribution $\mathbf{e}^\dagger$ under which full cooperation is sustainable and which maximises welfare $\Phi(\mathbf{e})$. We refer to $\mathbf{e}^\dagger$ as an *efficiency-maximising endowment distribution*. While finding an explicit expression for $\mathbf{e}^\dagger$ is not possible in general (Section 3.4.4), we can obtain numerically exact solutions for any group size n. Furthermore, we can fully characterise the general functional form of $\mathbf{e}^\dagger$ in the two-player case (considering without loss of generality $r_1 \geq r_2$) as

$$e_1^\dagger = \frac{\delta r_2}{2 - r_1 + \delta r_2} \quad \text{and} \quad e_2^\dagger = \frac{2 - r_1}{2 - r_1 + \delta r_2}. \tag{3.6}$$

As expected, the efficiency-maximising endowment distribution allocates larger shares of the endowment to more productive players. While this effect is similar to the effect of the resilience-maximising endowment distribution, the resulting degree of inequality differs (Fig. 3.3a–c). In fact, there exist parameters $\mathbf{r}$ and δ such that the efficiency-maximising distribution results in an exclusion of the least productive players by allocating them a zero-share of the total endowment. This is in stark contrast to the resilience-maximising endowment distribution; for $\mathbf{e}^*$, we prove that all players are always allocated a positive share (SI, Corollary 7).

3.1.4 Trade-off between efficiency and resilience

Since the endowment distributions $\mathbf{e}^*$ and $\mathbf{e}^\dagger$ are generally not the same, we further analyse the relation between them. There are two possible cases: First, the resilience-maximising endowment distribution $\mathbf{e}^*$ simultaneously also achieves maximal efficiency. This occurs exactly if $\delta = \delta^*_{\min}$ (in which case there is a unique endowment distribution that can sustain cooperation), or if all players have the same productivity (in which case every endowment distribution has the same efficiency). Second, in all other cases, we can prove that under any measure of inequality, the efficiency-maximising endowment distribution $\mathbf{e}^\dagger$ is always more unequal than the resilience-maximising distribution $\mathbf{e}^*$.

Given there is a trade-off between the two objectives in most settings, we combine them into a multi-objective optimisation setup. We visualise the resulting Pareto frontier between the resilience of cooperation and its efficiency in Fig. 3.3d–e. The pink line indicates the maximum welfare that can be sustained for any value of $\delta_{\min}(\mathbf{e})$ (Fig. 3.3d). Each point on this line corresponds to an endowment distribution $\mathbf{e}$ with the corresponding

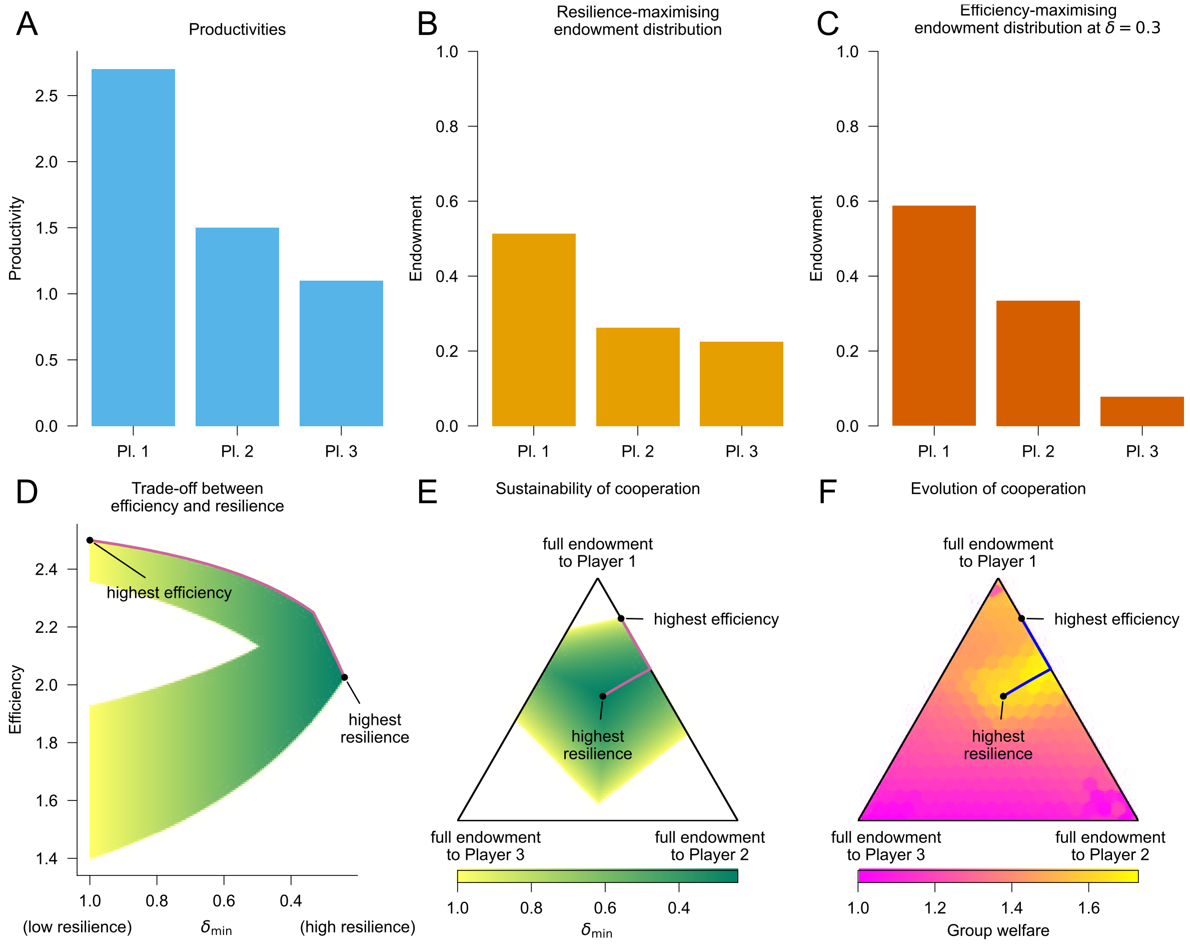

Figure 3.3: Trade-off between efficiency and resilience of cooperation. We demonstrate the difference between the resilience- and efficiency-maximising endowment distributions in a 3-player example. A, We chose the productivity vector to be $r = (2.7, 1.5, 1.1)$. B, The resulting resilience-maximising endowment distribution yields the total social welfare of $\Phi = 2.025$. C, We find that the efficiency-maximising endowment at $\delta = 0.3$ is more unequal and yields a total group payoff of $\Phi = 2.174$. D,E We formulate a multi-objective optimisation problem where both the resilience and efficiency are varied. The Pareto optimal values are shown by the pink line. F, We run simulations to test which of the endowment distributions performs best when players adopt strategies based on a stochastic learning process. We find that in general, the highest cooperation levels are achieved along the Pareto frontier. Indeed, the total maximum group payoff of 1.729 is achieved at $e_{\text{maxW}} = (0.65, 0.35, 0)$.

values of $\delta_{\min}(\mathbf{e})$ and $\Phi(\mathbf{e})$ (Fig. 3.3e). Generally, the most efficient endowment allocation $\mathbf{e}^{\dagger}$ is located at the boundary of the set of all endowment allocations that allow for full cooperation. Hence, while securing a maximal group payoff, it also poses the greatest strain on the resilience of cooperation. On the other hand, the resilience-maximising endowment allocation always requires higher equality while yielding lower welfare, signifying the trade-off between efficiency and resilience.

3.1.5 Dynamics of cooperation

To complement these static equilibrium results, in the following we explore when cooperation can emerge for a given set of endowments and productivities. To this end, we no longer assume that players act optimally from the outset. Rather, they adapt their strategies over time to optimise their payoffs. To model this, we use introspection dynamics [HHCN19, CGH22], a learning process where players are repeatedly selected at random to revise their strategies. When selected, players compare their current payoff with the payoff they could have obtained with a randomly generated alternative strategy. The higher the payoff of the alternative, the more likely players switch (as described in detail in the Methods section). In line with the literature on direct reciprocity [NS92, Fre94, KDK99, IN10, KWI10, GT12, PD12, VSPLS12, SP14a, SP15, SP14c, TRS14, DLTZ15, PHRZ15, BJHN16, MH16a, RHR+18, IM18, HCN18, SHCN22], we assume that individuals can only adopt pure memory-one strategies. That is, players condition their actions only on the outcome of the previous round [Sig10]. Moreover, in any given round, players either contribute their entire endowment or nothing at all. The resulting learning dynamics can be represented by a Markov chain. By computing its invariant distribution, we can infer the frequencies of each of the memory-one strategies in the long run.

To start with, we explore the simplest possible case of a game with two players with equal productivities. Here, the resilience-maximising and efficiency-maximising endowment distributions coincide at $e_1 = e_2 = 0.5$. In agreement with this prediction, we find that equal endowments are most favourable to the evolution of cooperation (Fig.3.4b). We also find, as expected, that higher values of the selection strength parameter allow for more cooperation (Fig. 3.4c). The effect of the error rate ϵ however is not monotonous (Fig. 3.4a). As has been documented in the past, a moderate amount of errors can be beneficial, because errors prevent the neutral invasion of conditionally cooperative strategies like ‘Win-Stay Lose-Shift’ by unconditional cooperators [Zha18]. Excessive errors, however, are always detrimental to cooperation, because they render conditionally cooperative strategies unstable.

Next, we look at a scenario with heterogeneity in individual productivities. We find that the endowment distribution that achieves the highest group welfare is located somewhere between the resilience-maximising and efficiency-maximising endowment distributions (Fig.3.4e). Its exact location depends on the the error rate and the selection strength. We observe that an increase in the selection strength results in a clear shift towards higher efficiency of the endowment distribution (Fig.3.4f). In contrast, the error rate can have varying effects (Fig.3.4e). As a rule of thumb, in more noisy settings (either because of a high error rate, or a low selection strength), allocations close to the resilience-maximising endowment distribution tend to result in a higher welfare.

In addition, Fig. 3.5 reports the resulting average cooperation rates. As can be seen, endowment allocations for which we observe the maximum group welfare are different from the endowment allocations that maximise cooperation (Fig.3.4e and Fig. 3.5e). To gain intuition for why this is the case, we look at the distribution of strategies for each endowment distribution of interest. Apart from the efficiency- and resilience-maximising endowment distributions, we also include the endowment distributions where we observe maximum group payoffs, $\mathbf{e}_{\max W}$, and maximum cooperation rates, $\mathbf{e}_{\max C}$ (Fig. 3.6), respectively. It appears that WSLS is the most abundant strategy for all of these endowment distributions apart from $\mathbf{e}^{\dagger}$, where it is not evolutionarily stable. WSLS

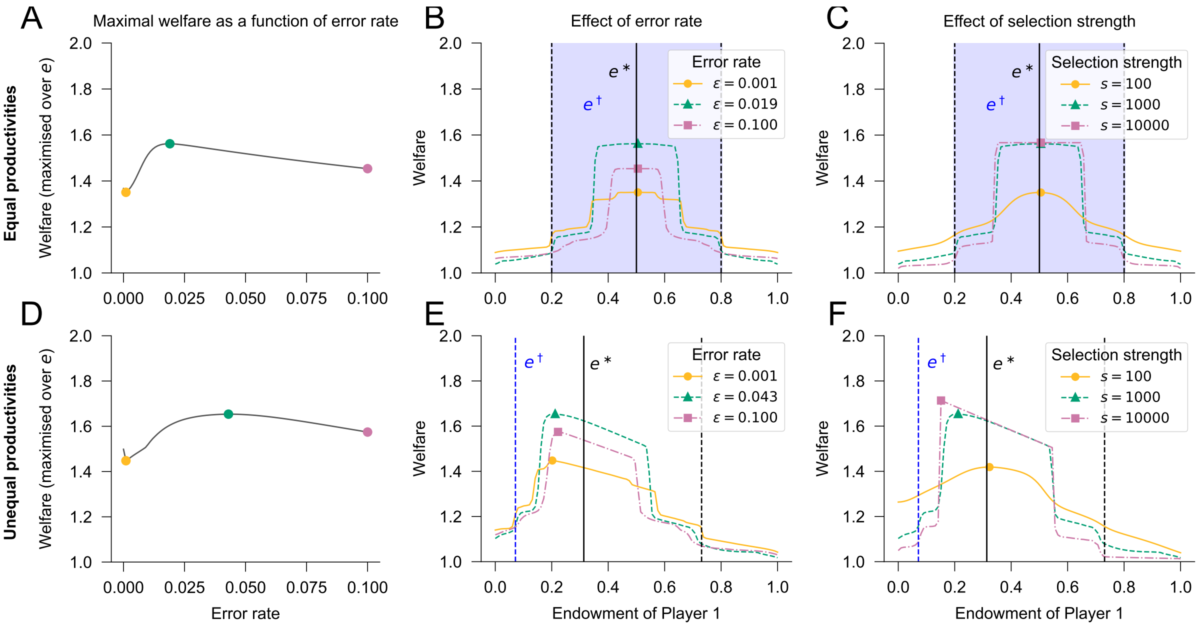

Figure 3.4: Evolutionary simulations of group welfare. In order to gain a deeper insight into the behaviour of the dynamics, we provide results of extensive simulations of a two-player game for a wide range of parameters of the dynamics, that is, the error rate and the intensity of selection. A,B,C, We first study the evolution of cooperation with equal productivities. Here, the dashed vertical lines bound the region where cooperation is sustainable at $\delta = 1$ according to the analytical model. We choose three error rate values for comparison: very rare errors, very frequent errors and the error rate $\varepsilon^* = 0.019$ that yields the highest welfare at $\mathbf{e} = \mathbf{e}^*$ (indicated by the solid vertical line). As can be seen, some (rare) errors can help the evolution of cooperation by ensuring stability of cooperative strategies such as WSLS [Zha18]. Here, there is no unique $\mathbf{e}^\dagger$, since all endowment distributions where full cooperation is sustainable (shaded in blue), yield identical welfare. Near the boundaries of this interval, we observe very low cooperation rates, while the highest group welfare is observed at the resilience-maximising endowment in the centre. D,E,F, Next, we consider a two-player game with unequal productivities given as $r_1 = 1.3$ and $r_2 = 1.9$. We employ the same logic for the choice of the parameters and obtain $\varepsilon^* = 0.043$. As can be seen in panel E, the highest group welfare is no longer attained at $\mathbf{e}^*$ but at a point in between $\mathbf{e}^\dagger$ and $\mathbf{e}^*$. For ε^*, we denote that point as $e_{\max\mathrm{W}}$. It is equal to $(0.21, 0.79)$. As with equal productivities, we find that higher selection strength increases welfare. Here, it also shifts the endowment $e_{\max\mathrm{W}}$ closer to $e^\dagger$.

is less prone to invasion at $\mathbf{e}^*$ and $\mathbf{e}_{\max C}$ for higher error rates (Fig. 3.7).

We also explore the dynamics of cooperation in a three-player game (Fig.3.3f). As predicted by the model, cooperation is higher when the most productive player obtains the largest endowment. Similarly, we observe that the endowment distributions with the highest group payoffs all lie close to the Pareto frontier between resilience and welfare, which connects the resilience-maximising and efficiency-maximising endowment distribution.

3.2 Discussion

Numerous studies have shown that wealth inequality can pose a challenge to cooperation [HHCN19, CKS05, MSMS21, ACBS+18, BBV86, VSPL14, ACT14, Aki15, Zel03, DBF18, BC06, HRS16, GMTV17, NSRC15, Gäc15, CMMM99, KMS17]. Since wealth inequality is abundant in many social settings, addressing it is often an important objective for policy makers. Perhaps one of the most straightforward ways of reducing inequality is through wealth redistribution [HL95, Lou81, BN93, Ben00, GZ93, Pik97, DD22]. As a naive generalisation of the theoretical and experimental findings, one could reach the conclusion that any wealth inequality is detrimental to cooperation and welfare, implying that redistribution should aim for equal allocations. However, our analysis shows that finding an allocation that most easily facilitates cooperation while maximising welfare is non-trivial.

We show that when productivity differs across people, an equal endowment allocation is neither optimal for the resilience of cooperation nor for the maximisation of welfare. Yet, consistent with previous findings [HHCN19, CKS05, MSMS21, ACBS+18, BBV86, VSPL14, ACT14, Aki15, Zel03, DBF18, BC06, HRS16, GMTV17, NSRC15, Gäc15, CMMM99, KMS17], we also find that excessive levels of endowment inequality cause cooperation to break down. We show how the optimal degree of endowment inequality varies with several parameters. In particular, it depends on the ratio between group size and individual productivities: if the players' productivities are fixed, larger groups require more equal endowments to maximise the resilience of cooperation, all the way to perfect equality in the limit of large group size (Fig. 3.2).

Our findings also point to a connection between the 'resilience of cooperation' and 'resilience of biological systems' [Hol73]. A higher resilience in our model means that full cooperation can be sustained for a wider range of continuation probabilities. If δ is seen as a parameter of the environment, then an endowment distribution with higher resilience allows for cooperation in environments that are less favourable to cooperation in the sense that they have a lower δ. Interpreted more loosely, when endowments are more resilient, cooperation can withstand greater perturbations of the environment.

To what degree resilience can be sacrificed for efficiency gains depends on the context. We explore this trade-off with introspection dynamics [CGH22]. We find that the endowment allocation that generates the highest welfare under these learning dynamics is located on the Pareto frontier between resilience and efficiency, balancing these two objectives. These observations are in line with a behavioural experiment conducted by Hauser *et al.* [HHCN19]. Interestingly, $\mathbf{e}_{\max W}$ and $\mathbf{e}_{\max C}$ are close to the endowments chosen for the treatments in the experiment. Similar to our numerical results, the authors find that cooperation rates for both of these allocations are roughly the same (approximately 73%) with a higher welfare achieved at the endowment allocation closer to $\mathbf{e}_{\max W}$.

Throughout this main text, we have focused on full cooperation in linear public good games with asymmetry in endowments and productivities. However, our theoretical results, as presented in the Supporting Information, are valid within a framework that is significantly more general in two aspects. First, we allow for the public good to be distributed unequally (Section 3.4.1). This weakens the public-good character of the game, but logically strengthens our results. Second, we consider arbitrary levels of contributions and derive all results with no restrictions on stochasticity, memory, or time-dependence of strategies (Section 3.4.3).

We believe our work makes at least two important contributions to the literature on the evolution of cooperation through direct reciprocity. First, we considerably extend earlier results by Hauser *et al.* [HHCN19]. Although they discuss which endowment distributions might maximise resilience (there: "endowment distribution most conducive to cooperation"), their analysis is restricted to groups of size two. Instead, here we provide an elegant formalism that allows us to compute $\mathbf{e}^*$ for any number of players. In this way, we can analyse the interaction between parameters of the game and the optimal degree of inequality. Our method can be used to further study the effects of inequality in more general settings, for example, in structured populations or when allowing for communication or signalling among players. Second, we study the effect of inequality on the interplay between the resilience and efficiency of cooperation. Our results for a general n-player linear public good game indicate that there exist non-trivial trade-offs, which need to be accounted for when deciding on the allocation of wealth. We explore these trade-offs using evolutionary simulations. Overall, our results suggest that a positive degree of inequality can be beneficial for cooperation, in particular in small-group interactions, while in other settings, almost perfect equality is optimal even in the face of intrinsic differences between individuals.

3.3 Methods

3.3.1 Model

Consider a game with n players who interact in an infinite sequence of rounds $t = 0, 1, 2, \dots$. Each player has a fixed positive endowment $e_i \geq 0$, where $\sum_i e_i = 1$ without loss of generality. In each round t, each player i chooses a contribution $c_i(t) \in [0, e_i]$ to make towards the public good. The productivity matrix R, a parameter of the game subject to below constraints, governs the relationship between contributions and payoffs as

$$\boldsymbol{\pi}(t) = \mathbf{e} - \mathbf{c}(t) + R\mathbf{c}(t), \text{ for all } t.$$

There are three constraints on R. First, $R_{ij} \geq 0$ for all i, j, ensuring that an increase in one player's contribution does not decrease any player's return from the public good. Second, $R_{ii} < 1$ for all i, so that a player's one-round payoff is higher the less they contribute. Third, $\sum_j R_{ji} > 1$ for all i, meaning that the group payoff of all players taken together is higher the more Player i contributes. The second and third condition create a tension between the individual and the collective interest, which makes this game a social dilemma. The restricted model discussed in the main text is the special case where R is of the form $R_{ij} = r_j/n$ for some $\mathbf{r} = (r_1, \dots, r_n)$.

3.3.2 Equilibrium analysis

We determine when it is possible for a given contribution sequence $(\mathbf{c}(t))_t$ to occur in a SPNE. We introduce a normal form of the productivity matrix R, which we call the zero-diagonal form and denote it by D. It is given by $D_{ij} = R_{ij}/(1 - R_{ii})$, for all $i \neq j$, and $D_{ii} = 0$, for all i. The game defined by D is equivalent to the game defined by R in the sense that the two games permit exactly the same equilibria, while the fact that the diagonal entries of D are zero simplifies the analysis. We show that a contribution sequence $(\mathbf{c}(t))_t$ is sustainable in a SPNE exactly if $\bar{\mathbf{c}}(t) \leq \delta D \bar{\mathbf{c}}(t+1)$ for all t (Theorem 1), where $\bar{\mathbf{c}}(t) = (1-\delta)\sum_{\tau=0}^{\infty} \delta^{\tau} \mathbf{c}(t+\tau)$ is the continuation contribution after round t. For a given contribution sequence $(\mathbf{c}(t))_t$, we define $\delta_{\min}((\mathbf{c}(t))_t)$ as the smallest continuation probability δ for which the sequence is sustainable.

3.3.3 Evolutionary analysis

We study introspection dynamics, a simple learning process [CGH22]. Players use pure memory-one strategies where in a given round they either contribute the entire endowment or nothing, that is, $c_i(t) \in \{0, e_i\}$. We represent strategies by a vector $\mathbf{p} = (p_{\mathbf{c}})_{\mathbf{c}}$, where $\mathbf{c}$ ranges over the 2^n possible outcomes of one round. Each component $p_{\mathbf{c}} \in \{0,1\}$ specifies whether after a round with outcome $\mathbf{c}$, a player with a strategy $\mathbf{p}$ contributes their entire endowment ($p_{\mathbf{c}} = 1$) or nothing ($p_{\mathbf{c}} = 0$). However, in our simulations players are also prone to making errors with a probability $\varepsilon > 0$, meaning that they sometimes play an action not prescribed by their strategy. Since that makes the process ergodic and we only focus on its asymptotic behaviour, our representation of strategies does not contain an initial move.

At every time step of the learning process, a player is chosen to consider switching their strategy. The player compares their current strategy to a randomly generated, alternative strategy in terms of average payoff and adopts the alternative with probability

$$\rho = \frac{1}{1 + \mathrm{e}^{s(\pi_{\mathrm{alt}} - \pi_{\mathrm{cur}})}},$$

where π_{cur} and π_{alt} are the payoffs of the current strategy and the alternative, respectively. The parameter $s \geq 0$ reflects the strength of selection. Higher values of s correspond to stronger selection.

We use different implementations of the learning process depending on n. For $n=2$, we calculate the asymptotic distribution by numerically representing the stochastic process as a transition matrix of a Markov chain. We then calculate the average payoffs as expected values under the invariant distribution as

$$\pi_i = \sum_{\mathbf{c}} v_{\mathbf{c}} \cdot \pi_i(\mathbf{c}),$$

where π_i is Player i's average payoff, $\mathbf{v}$ is the invariant distribution of the Markov chain, and $\pi_i(\mathbf{c})$ is Player i's payoff in a round with contribution vector $\mathbf{c}$. For $n = 3$, we run an agent-based simulation for $N = 10^6$ generations and report the average group payoffs.

3.3.4 Parameters used for figures

Figure 2a–c shows a three-player game with productivities $r_1 = 1.6$, $r_2 = 1.3$, $r_3 = 1.1$. In Figure 2d–f, the relative differences in productivities are identical, but all values are higher by 1.399: $r_1 = 2.999$, $r_2 = 2.699$, $r_3 = 2.499$. Figure 2g–h shows a game with six players. The productivity values are distributed as in Figure 2d–f, in the sense that for each player in the three-player game, there are two identical players in the six-player game: $r_1 = r_2 = 2.999$, $r_3 = r_4 = 2.699$, $r_5 = r_6 = 2.499$. In Figure 2i, we show the value of $e^*_{\max}/e^*_{\min}$ as a measure of inequality for groups of various sizes. In each case, one third each of players has the productivities 2.7, 1.5 and 1.1.

Figure 3 analyses a three-player game with productivities $r_1 = 2.7$, $r_2 = 1.5$, $r_3 = 1.1$. In Figure 3c, we report the efficiency-maximising endowment distribution for $\delta = 0.3$. In Figure 3f, we report evolutionary simulations for the same three-player game with the selection strength $s = 1000$.

Figure 4 presents data from evolutionary simulations of a two-player game. Here, players can either contribute their entire endowment e_i or defect by contributing 0 to the public good. In Figure 4a–c we report results for a symmetric two-player game with productivities $r_1 = r_2 = 1.6$. Figure 4a depicts the maximal welfare for each value of the error rate ε from 0 to 0.1 for varying endowments. Selection strength is set to $s = 1000$. Three points on the curve are highlighted: $\varepsilon = 0.001$, $\varepsilon = 0.019$, and $\varepsilon = 0.1$. The point $\varepsilon = 0.019$ is where the function attains its maximum, while the other two values are chosen arbitrarily for comparison. For these three values of ε, Figure 4b shows the welfare achieved for all possible endowment distributions, still with $s = 1000$. In this symmetric game, welfare is always maximised by $e = (0.5, 0.5)$. Finally, in Figure 4c, ε is held constant at 0.019 while three different values of the selection strength s are shown for comparison. Figure 4d–f follow the same pattern, except that productivities are unequal and set to $r_1 = 1.3$, $r_2 = 1.9$. In Figure 4d, the maximum is attained at $\varepsilon = 0.043$. The same parameter values are used in Figures 3.5–3.7.

3.4 Supplementary Information

3.4.1 Model

A group of $n \geq 2$ players interacts in a series $t = 0, 1, 2, \ldots$ of rounds. Each player has a fixed non-negative endowment e_i. Not all endowments are zero, and we let $\sum_i e_i = 1$ without loss of generality. In each round t, each player i chooses a contribution $c_i(t) \in [0, e_i]$ to make towards a jointly produced good. In the most general setting, players' payoffs $\pi_i(t)$ in round t are defined in terms of a productivity matrix R by

$$\boldsymbol{\pi}(t) = \mathbf{e} - \mathbf{c}(t) + R\mathbf{c}(t). \tag{3.7}$$

The matrix R must satisfy the following three criteria:

1. $R_{ij} \geq 0$ for all i, j. This means that no player gets less out of the public good when one player contributes more.

2. $R_{ii} < 1$ for all i. This means that in an individual round, a player's payoff is higher the less they contribute.

3. $\sum_j R_{ji} > 1$ for all i. This means that the group payoff of all players taken together is higher the more Player i contributes.

The antithesis between 2. and 3. constitutes the social dilemma that characterises this game.

If R is of the form $R = \mathbf{f}\,\mathbf{r}^\intercal$ for some $\mathbf{r}$ and $\mathbf{f}$, such that $\sum_i f_i = 1$ and $f_i \geq 0$ for all i, we can interpret $\mathbf{r}$ as the productivity vector and $\mathbf{f}$ as the sharing vector. That is, the contributions of each player i are enhanced by an individual productivity factor r_i. Then, all contributions are compounded to form the publicly produced good. Finally, the good is divided according to $\mathbf{f}$. The above conditions on R are then equivalent to $f_i r_i < 1$ and $1 < r_i$ for all i. We call this special case "origin-independent sharing" of the good.

We speak of "equal sharing" of the good, or of a "non-excludable public good" when additionally $\mathbf{f} = (1/n, \ldots, 1/n)^\intercal$. Our only requirement is then $1 < r_i < n$ for all i. Each player receives the same amount $n^{-1}\mathbf{r}^\intercal\mathbf{c}$ from the public good, so payoffs are

$$\pi_i(t) = e_i - c_i(t) + n^{-1}\mathbf{r}^\intercal\mathbf{c}(t).$$

After each round, it is decided independently at random whether or not the game is continued for another round, with a continuation probability of δ satisfying $0 < \delta < 1$. The probability that round t is played is therefore equal to δ^t.

We denote the game fully specified by R, $\mathbf{e}$, and δ by $\Gamma(R, \mathbf{e}, \delta)$. For the case of equal sharing, we simply write $\Gamma(\mathbf{r}, \mathbf{e}, \delta)$ to mean $\Gamma((1/n, \ldots, 1/n)^\intercal\,\mathbf{r}^\intercal, \mathbf{e}, \delta)$.

The continuation contributions $\bar{\mathbf{c}}(t)$ are defined as the expected sum of future contributions at time t multiplied by a normalising factor of $(1 - \delta)$:

$$\bar{\mathbf{c}}(t) = (1 - \delta) \sum_{\tau=0}^{\infty} \delta^\tau \mathbf{c}(t + \tau)$$

Similarly for continuation payoffs:

$$\bar{\boldsymbol{\pi}}(t) = (1 - \delta) \sum_{\tau=0}^{\infty} \delta^\tau \boldsymbol{\pi}(t + \tau)$$

We know that these sums converge, since $\mathbf{c}(t)$ is bounded above by $\mathbf{e}$. Since the definition of $\bar{\mathbf{c}}(t)$ depends on δ, we may also write $\bar{\mathbf{c}}_\delta(t)$, etc., whenever there is ambiguity.

We write $\hat{\boldsymbol{\pi}} = \bar{\boldsymbol{\pi}}(0)$ and $\hat{\mathbf{c}} = \bar{\mathbf{c}}(0)$ for the total payoffs and contributions, i.e., continuation payoffs and contributions at the beginning of the game. We have

$$\hat{\boldsymbol{\pi}} = \mathbf{e} - \hat{\mathbf{c}} + R\hat{\mathbf{c}} = \mathbf{e} + (R - I_n)(1 - \delta) \sum_{t=0}^{\infty} \delta^t \mathbf{c}(t). \tag{3.8}$$

Index of notation

r_i, $\mathbf{r}$	productivity
R	productivity matrix
e_i, $\mathbf{e}$	endowment
$\mathbf{e}^*$	resilience-maximising endowment distribution
$\mathbf{e}^\dagger$	efficiency-maximising endowment distribution
δ	continuation probability
$c_i(t)$, $\mathbf{c}(t)$	(absolute) contributions in round t
$(\mathbf{c}(t))_t$	the contribution sequence $(\mathbf{c}(0), \mathbf{c}(1), \mathbf{c}(2), \dots)$
$\bar{\mathbf{c}}(t)$ or $\bar{\mathbf{c}}_\delta(t)$	continuation contributions from time t, depends on δ
$\widehat{\mathbf{c}} = \bar{\mathbf{c}}(0)$	total contributions
$\boldsymbol{\pi}(t)$	payoffs in round t
D	the zero-diagonal form of the productivity matrix as defined on page 30
$\delta_{\min}$	resilience, the least δ for which a given $(\mathbf{c}(t))_t$ is sustainable with a given $\mathbf{r}$ or R. Note: A lower value of $\delta_{\min}$ means that $(\mathbf{c}(t))_t$ is more resilient.
$\delta^*_{\min}$	the least $\delta_{\min}$ of any $(\mathbf{c}(t))_t$ for a given $\mathbf{r}$ or R
$\Phi(\widehat{\mathbf{c}})$	efficiency (def. page 40) of a given total contribution vector $\widehat{\mathbf{c}}$

3.4.2 Sustainability of cooperation

In this section, we define the notion of sustainability for certain kinds of outcomes, i.e. functions of the game play. Simply put, an outcome is sustainable if it can occur in a subgame-perfect Nash equilibrium (SPNE). First, we define some standard notions.

Definition 1 (Strategy)**.** *A strategy for Player i is a function σ that assigns to every finite sequence of moves $(\mathbf{c}(t))_{t<T}$ some value $\sigma((\mathbf{c}(t))_{t<T}) \in \mathcal{P}([0, e_i])$, i.e. a probability distribution over the possible next moves by Player i.*

Player i is said to follow strategy σ if Player i's move $c(T)_i$ in round T is distributed as $\sigma((\mathbf{c}(t))_{t<T})$ under the condition that $(\mathbf{c}(t))_{t<T}$ was played in the first T rounds, for all T and $(\mathbf{c}(t))_{t<T}$.

A strategy profile $(\sigma_i)_i$ is an n-tuple of one strategy for each player.

The SPNE is a formal notion of rational behaviour in repeated interactions. In simple terms, a strategy profile $(\sigma_i)_i$ is in an SPNE if at no possible point in the game (i.e. after no finite sequence of moves) any player i has a positive incentive not to follow strategy σ_i.

Definition 2 (subgame-perfect Nash equilibrium)**.** *A strategy profile $(\sigma_i)_i$ is a subgame-perfect Nash equilibrium (SPNE, "equilibrium") when there are no t, $\mathbf{c}(0), \dots, \mathbf{c}(t-1)$, i, and σ_i^* such that the expected value of $\bar{\pi}_i(t)$ after the initial sequence of moves $\mathbf{c}(0), \dots, \mathbf{c}(t-1)$ is strictly greater when Player i follows σ_i^* compared to σ_i.*

Defining sustainability of cooperation

We are interested in the existence of equilibria that produce non-zero contributions, because they achieve higher payoffs than the always existing [ADS94] unconditionally defective (i.e. zero-contributing) SPNE. After establishing some preliminary results, we will look for equilibria that are optimal with respect to certain specific properties related to cooperation. Naturally, all of these properties are functions of the contribution sequence $(\mathbf{c}(t))_t$, while some even depend just on the total contribution vector $\widehat{\mathbf{c}}$.

We thus aim to characterise each contribution sequence regarding the existence of an SPNE strategy profile producing that sequence, so that we may henceforth directly analyse contribution sequences rather than the more complex strategy profiles.

Definition 3 (Contribution sequences and feasibility)**.** *A contribution sequence is any sequence $(\mathbf{c}(t))_{t\in\mathbb{N}}$ of elements in $\mathbb{R}^n$ that is non-negative and bounded. A total contribution vector is any non-negative $\widehat{\mathbf{c}} \in \mathbb{R}^n$.*

A contribution sequence is feasible in a given game $\Gamma(R, \mathbf{e}, \delta)$ if it satisfies $\mathbf{c}(t) \leq \mathbf{e}$ for all t. Similarly, a total contribution vector $\widehat{\mathbf{c}}$ is feasible if $\widehat{\mathbf{c}} \leq \mathbf{e}$.

Definition 4 (Sustainability)**.** *In a given game $\Gamma(R, \mathbf{e}, \delta)$, we say a contribution sequence $(\mathbf{c}(t))_t$ is sustainable if it is feasible and there is a SPNE strategy profile $(\sigma_i)_i$ whose sequence of expected contributions is exactly $(\mathbf{c}(t))_t$.*

A total contribution vector $\widehat{\mathbf{c}}$ is sustainable if it is the total contribution vector of a sustainable contribution sequence.

Grim trigger strategies. Proposition 1 below states that in order to determine whether a feasible contribution sequence is sustainable, it is enough to consider its associated Grim strategy profile. It is defined as follows.

Definition 5 (Grim strategies)**.** *The Grim strategy profile $G((\mathbf{c}(t))_t)$ for a feasible contribution sequence $(\mathbf{c}(t))_t$ is the pure strategy profile $G((\mathbf{c}(t))_t) = (\sigma_i)_i$ defined as follows: In each round t and for each i the strategy σ_i contributes $c_i(t)$ if all players have so far also played according to $(\mathbf{c}(t))_t$, but otherwise contributes* 0*.*

The Grim strategy profile $G((\mathbf{c}(t))_t)$ produces the contribution sequence $(\mathbf{c}(t))_t$, hence there is a bijective correspondence between feasible contribution sequences and Grim strategies.

Proposition 1. *In $\Gamma(R, \mathbf{e}, \delta)$, if a (mixed) strategy profile $(\sigma_i)_i$ is a SPNE, then there is a Grim strategy profile that is also a SPNE with the same expected contributions for each player in each round.*

Proof. We write $\pi_i(\mathbf{c})$ for Player i's payoff in a round with contributions $\mathbf{c}$. Below, $(\mathbf{c}(t))_t$ is the contribution sequence produced by $(\sigma_i)_i$, which is a random variable. Since $(\sigma_i)_i$ is a SPNE, we have for all i, t and any $\mathbf{x}(0), \ldots, \mathbf{x}(t-1)$ such that $\mathbb{P}(\forall s < t \quad \mathbf{c}(s) = \mathbf{x}(s)) > 0$, that

$$\begin{aligned} \mathbb{E}\Big(\pi_i(\mathbf{c}(t)_{-i}) \,\Big|\, \forall s < t \quad \mathbf{c}(s) = \mathbf{x}(s)\Big) + \sum_{k=t+1}^{\infty} \delta^{k-t}\pi_i(\mathbf{0}) \\ \leq \mathbb{E}\left(\sum_{k=t}^{\infty} \delta^{k-t}\pi_i(\mathbf{c}(k)) \,\middle|\, \forall s < t \quad \mathbf{c}(s) = \mathbf{x}(s)\right), \end{aligned} \tag{3.9}$$

where $\mathbf{c}(t)_{-i}$ is equal to $\mathbf{c}(t)$ with the ith component set to 0. This inequality represents the fact that Player i cannot profit from switching to zero cooperation starting in round t

after $\mathbf{x}(1), \ldots, \mathbf{x}(t-1)$ was played, and no more so if that would result in all the other players also switching to zero cooperation one round later.

By the law of total expectations,

$$\mathbb{E}\left(\pi_i(\mathbf{c}(t)_{-i})\right) + \sum_{k=t+1}^{\infty} \delta^{k-t}\pi_i(\mathbf{0}) \leq \mathbb{E}\left(\sum_{k=t}^{\infty} \delta^{k-t}\pi_i(\mathbf{c}(k))\right).$$

Since $\boldsymbol{\pi}(\mathbf{c})$ is affine linear, we may write this as

$$\pi_i\Big(\mathbb{E}\mathbf{c}(t)_{-i}\Big) + \sum_{k=t+1}^{\infty} \delta^{k-t}\pi_i(\mathbf{0}) \leq \sum_{k=t}^{\infty} \delta^{k-t}\pi_i\Big(\mathbb{E}\mathbf{c}(k)\Big).$$

This means that in $G((\mathbb{E}\mathbf{c}(t))_t)$, Player i cannot profit from deviating if so far everyone has played according to the sequence $(\mathbb{E}\mathbf{c}(t))_t$. But if someone has already deviated, then this is also the case, since in that case all players contribute 0 independently of each others' moves. Therefore, $G((\mathbb{E}\mathbf{c}(t))_t)$ is a SPNE. ☐

Sustainability of contribution sequences

Characterising sustainable contribution sequences with the zero-diagonal productivity matrix. Using Proposition 1, we will now find a useful characterisation of the sustainability of feasible contribution sequences.

Definition 6 (Productivity matrix in zero-diagonal form). *For a given productivity matrix R, define the zero-diagonal productivity matrix D by*

$$D_{ij} = \mathbb{1}_{i\neq j}\frac{R_{ij}}{1-R_{ii}},$$

where $\mathbb{1}_{i\neq j}$ is equal to 1 when $i \neq j$ and 0 otherwise.

It is easy to verify that D is indeed a valid productivity matrix, meaning it satisfies the three conditions on page 26. This alternative productivity matrix D has the property that for any game $\Gamma(D, \mathbf{e}, \delta)$ and any sequence of play $(\mathbf{c}(t))_t$, the resulting payoff vector $\boldsymbol{\pi}^D$ satisfies $\pi_i^D - e_i = (\pi_i^R - e_i)/(1 - R_{ii})$, where $\boldsymbol{\pi}^R$ is the payoff vector of the same contribution sequence in the game $\Gamma(R, \mathbf{e}, \delta)$. This fact can easily be verified using the payoff equation (3.7). Since this relationship between π^D and π^R is affine linear, $\Gamma(D, \mathbf{e}, \delta)$ admits exactly the same SPNEs as $\Gamma(R, \mathbf{e}, \delta)$. The fact that the diagonal entries of D are zero will make it a useful tool for equilibrium analysis. We can also check that the zero-diagonal form as defined above of a matrix whose diagonal is already zero is itself.

Under origin-independent sharing of the public good, $R = \mathbf{f}\mathbf{r}^\intercal$, the above definition takes the form

$$D_{ij} = \mathbb{1}_{i\neq j}\frac{r_j}{f_i^{-1} - r_i},$$

and under equal sharing of the public good, $R = n^{-1}\mathbf{1}\mathbf{r}^\intercal$, it becomes

$$D_{ij} = \mathbb{1}_{i\neq j}\frac{r_j}{n - r_i}.$$

Theorem 1. *In $\Gamma(R, \mathbf{e}, \delta)$, the feasible contribution sequence $(\mathbf{c}(t))_t$ is sustainable exactly if*

$$\forall t \quad \bar{\mathbf{c}}(t) \leq \delta D \bar{\mathbf{c}}(t+1). \tag{3.10}$$

Proof. By Proposition 1, $(\mathbf{c}(t))_t$ is sustainable if and only if $G((\mathbf{c}(t))_t)$ is a SPNE. This is the case if and only if

$$\pi_i\Big(\mathbf{c}(t)_{-i}\Big) + \sum_{k=t+1}^{\infty} \delta^{k-t}\pi_i(\mathbf{0}) \leq \sum_{k=t}^{\infty} \delta^{k-t}\pi_i\Big(\mathbf{c}(k)\Big).$$

for all i and t. We can rewrite this as

$$\begin{aligned}
\sum_{j \neq i} R_{ij} c_j(t) &\leq \sum_{k=0}^{\infty} \delta^k \left(\sum_j R_{ij} c_j(t+k) - c_i(t+k) \right) \\
\sum_{j \neq i} R_{ij} c_j(t) &\leq \sum_j R_{ij} \sum_{k=0}^{\infty} \delta^k c_j(t+k) - \sum_{k=0}^{\infty} \delta^k c_i(t+k)
\end{aligned}$$

for all i and t. By our definition of $\bar{\mathbf{c}}(t) = (1-\delta)\sum_{k=0}^{\infty} \delta^k \mathbf{c}(t+k)$, this can be written as

$$\begin{aligned}
\sum_{j \neq i} R_{ij}(1-\delta) c_j(t) &\leq \sum_j R_{ij}\bar{c}_j(t) - \bar{c}_i(t) \\
\bar{c}_i(t) + \sum_{j \neq i} R_{ij}\left((1-\delta)c_j(t) + \delta\bar{c}_j(t+1)\right) - \sum_j R_{ij}\bar{c}_j(t) &\leq \sum_{j \neq i} R_{ij}\delta\bar{c}_j(t+1) \\
\bar{c}_i(t) + \sum_{j \neq i} R_{ij}\bar{c}_j(t) - \sum_j R_{ij}\bar{c}_j(t) &\leq \sum_{j \neq i} R_{ij}\delta\bar{c}_j(t+1) \\
(1 - R_{ii})\bar{c}_i(t) &\leq \sum_{j \neq i} R_{ij}\delta\bar{c}_j(t+1) \\
\bar{c}_i(t) &\leq \sum_{j \neq i} \frac{R_{ij}}{1 - R_{ii}} \delta\bar{c}_j(t+1)
\end{aligned}$$

for all i and t. By our definition of D, this is equivalent to

$$\forall t \quad \bar{\mathbf{c}}(t) \leq \delta D \bar{\mathbf{c}}(t+1).$$

□

Corollary 1. *If for given (R, δ), a contribution sequence $(\mathbf{c}(t))_t$ is feasible with both endowment distributions $\mathbf{e}$ and $\mathbf{e}'$, then it is sustainable in $\Gamma(R, \mathbf{e}, \delta)$ exactly if it is sustainable in $\Gamma(R, \mathbf{e}', \delta)$.*

This follows immediately from Theorem 1, since inequality (3.10) does not depend on $\mathbf{e}$.

Corollary 2. *In any $\Gamma(R, \mathbf{e}, \delta)$, if a conical combination (a linear combination with non-negative coefficients) of either (a) sustainable contribution sequences $(\mathbf{c}(t))_t$, or (b) sustainable continuation contribution sequences $(\bar{\mathbf{c}}(t))_t$, or (c) sustainable total contribution vectors $\widehat{\mathbf{c}}$ is feasible, then it is also sustainable.*

Proof. For contribution sequences, this follows from homogeneity of (3.10). The statement directly translates to continuation contribution sequences and total contribution vectors by their definition. □

Corollary 3. *In any* $\Gamma(R, \mathbf{e}, \delta)$*, the set of sustainable contribution sequences* $(\mathbf{c}(t))_t$*, the set of sustainable continuation contribution sequences* $(\bar{\mathbf{c}}(t))_t$*, and the set of sustainable total contribution vectors* $\widehat{\mathbf{c}}$ *are convex. They contain* $(\mathbf{0})_t$ *and* $\mathbf{0}$*, respectively.*

Proof. Again, it is enough to show the statement for contribution sequences. Any convex combination of feasible contribution sequences is itself feasible. Both parts then follow from Corollary 2. □

Parameters allowing for cooperation. We say $(\sigma_i)_i$ is a non-defective strategy profile if it is not equal to $(ALLD)_i$. $ALLD$ is the strategy of always playing $c_i(t) = 0$. Note that $(\sigma_i)_i$ can be non-defective and still result in expected total contributions of $\mathbb{E}\widehat{\mathbf{c}} = \mathbf{0}$. We say a game $\Gamma(R, \mathbf{e}, \delta)$ allows for cooperation if it has a non-defective SPNE strategy profile. A game that allows for cooperation always has a strategy profile with $\mathbb{E}\widehat{\mathbf{c}} \neq \mathbf{0}$: Take any non-defective SPNE strategy profile and select an initial sequence of moves after which at least one player will not certainly play $c_i(t) = 0$. Then players' induced strategies in the corresponding subgame are a SPNE strategy profile with $\mathbb{E}\widehat{\mathbf{c}} \neq \mathbf{0}$. These strategies are also a SPNE when played in the root game, which is equal to the subgame. We can thus say that a game allows for cooperation exactly if $\mathbf{0}$ is not the only sustainable total contribution vector.

We say (R, δ) allows for cooperation if there is some $\mathbf{e}$ such that $\Gamma(R, \mathbf{e}, \delta)$ allows for cooperation.

Corollary 4. *If some* $\Gamma(R, \mathbf{e}, \delta)$ *allows for cooperation, then for all endowment distributions* $\mathbf{e}'$*,* $\Gamma(R, \mathbf{e}', \delta)$ *also allows for cooperation.*

This statement is a consequence of Corollaries 1 and 2.

We say a productivity matrix R allows for cooperation if there is a $\delta < 1$ such that (R, δ) allows for cooperation. As we will show in Theorem 2, every R allows for cooperation.

Sustainability of constant contribution sequences.

Corollary 5. *In a game* $\Gamma(R, \mathbf{e}, \delta)$*, feasible constant contributions of* $(\widehat{\mathbf{c}})_t$ *are sustainable in equilibrium exactly if*

$$\mathbf{0} \leq (\delta D - I_n)\widehat{\mathbf{c}}. \tag{3.11}$$

3.4.3 Resilience-maximising endowment distribution

With the results obtained in Section S2, we have characterised which patterns of cooperation (contribution sequences) we may observe in an interaction that is described by our model, and which patterns cannot occur. We should however not expect a qualitative shift between those strategies that are analytically stable, and those that are not, but rather for cooperation to become somehow less sustainable as we approach the boundary.

In order to also model this intuitive concept of a quantitative sustainability, we will use $\delta_{\min}$ as a measure of the resilience of cooperation, defined as the smallest δ that makes a given contribution sequence sustainable. This is supported by Lemma 1, which states that increasing δ can only expand the set of sustainable contribution sequences. We write $\delta^*_{\min}$ for the numerically lowest resilience attained by any strategy profile (thus the most resilient one) under any endowment distribution.

Using that definition, we show that there is exactly one endowment distribution $\mathbf{e}^*$ for which full cooperation is possible with resilience $\delta^*_{\min}$. Under this endowment distribution, players with higher productivity always have a higher endowment. In games where the number of players is not much larger than their productivities, this difference is significant, and the endowments e^*_i very roughly resemble $r_i/\sum_j r_j$. When on the other hand the number of players is large compared to their productivities, all players have almost the same endowment, i.e. $e^*_i \approx 1/n$.

Under every endowment distribution that has $e_i > 0$ for all i, forms of partial cooperation are possible with resilience $\delta^*_{\min}$. As $t \to \infty$, their contribution sequences all converge to multiples of $\mathbf{e}^*$.

Clearly from Corollary 5, if some $(\widehat{\mathbf{c}})_t$, i.e. a constant contribution sequence, is sustainable for some δ, then it is also sustainable for all higher values of δ. Less obviously, the same is true for any sustainable contribution sequence:

Lemma 1. *Let $(\mathbf{c}(t))_t$ be sustainable in $\Gamma(R, \mathbf{e}, \delta)$. Then for all $\delta' \geq \delta$, $(\mathbf{c}(t))_t$ is also sustainable in $\Gamma(R, \mathbf{e}, \delta')$.*

Proof. Inequality 3.10, which holds for $(\mathbf{c}(t))_t$, δ, and all t, can be rearranged to

$$\sum_{k=1}^{\infty} \delta^k (D - I_n)\, \mathbf{c}(t+k) \geq \mathbf{c}(t). \tag{3.12}$$

We use this (twice) to derive

$$\begin{aligned}
&\sum_{k=1}^{\infty} (\delta')^k (D - I_n)\, \mathbf{c}(t+k) \\
&= \sum_{k=1}^{\infty} \left(\sum_{i=1}^{k} \left((\delta')^i \delta^{k-i} - (\delta')^{i-1} \delta^{k-i+1} \right) + \delta^k \right) (D - I_n)\, \mathbf{c}(t+k) \\
&= \sum_{k=1}^{\infty} \delta^k (D - I_n)\, \mathbf{c}(t+k) + \sum_{k=1}^{\infty} \sum_{i=1}^{k} \left((\delta')^i \delta^{k-i} - (\delta')^{i-1} \delta^{k-i+1} \right) (D - I_n)\, \mathbf{c}(t+k) \\
&\geq \mathbf{c}(t) + \sum_{k=1}^{\infty} \sum_{i=1}^{k} \left((\delta')^i \delta^{k-i} - (\delta')^{i-1} \delta^{k-i+1} \right) (D - I_n)\, \mathbf{c}(t+k) \\
&= \mathbf{c}(t) + (\delta' \delta^{-1} - 1) \sum_{i=1}^{\infty} (\delta')^{i-1} \sum_{k=i}^{\infty} \delta^{k-i+1} (D - I_n)\, \mathbf{c}(t+k) \\
&= \mathbf{c}(t) + (\delta' \delta^{-1} - 1) \sum_{i=0}^{\infty} (\delta')^{i} \sum_{k=1}^{\infty} \delta^{k} (D - I_n)\, \mathbf{c}(t+i+k) \\
&\geq \mathbf{c}(t) + (\delta' \delta^{-1} - 1) \sum_{i=0}^{\infty} (\delta')^{i}\, \mathbf{c}(t+i)
\end{aligned}$$

for all t. Writing $\bar{\mathbf{c}}_{\delta'}$ for the continuation contribution weighted by δ', we can state this as

$$\begin{aligned}\mathbf{c}(t) + (\delta'\delta^{-1} - 1)\sum_{k=0}^{\infty}(\delta')^k\,\mathbf{c}(t+k) &\le \sum_{k=1}^{\infty}(\delta')^k(D - I_n)\,\mathbf{c}(t+k)\\ \delta'\delta^{-1}\sum_{k=0}^{\infty}(\delta')^k\,\mathbf{c}(t+k) &\le \sum_{k=1}^{\infty}(\delta')^k D\,\mathbf{c}(t+k)\\ (\delta'\delta^{-1})\,\bar{\mathbf{c}}_{\delta'}(t) &\le \delta' D\bar{\mathbf{c}}_{\delta'}(t)\end{aligned}$$

or alternatively

$$\bar{\mathbf{c}}_{\delta'}(t) \le \delta D\bar{\mathbf{c}}_{\delta'}(t).$$

□

Based on this result, we can use the following

Definition 7 (Resilience of a strategy profile)**.** *Let $(R, \mathbf{e})$ and non-defective $(\sigma_i)_i$ be given, where $(\sigma_i)_i$ is a SPNE in $\Gamma(R, \mathbf{e}, \delta)$ for some $\delta < 1$. Then the resilience $\delta_{\min}$ of $(\sigma_i)_i$ is the minimal value such that $(\sigma_i)_i$ is a SPNE in $\Gamma(R, \mathbf{e}, \delta_{\min})$.*

We note that $\delta_{\min}$ is well defined, in the sense that a minimum as required in the definition is indeed attained. To see this, we resort to the definition of a SPNE: $(\sigma_i)_i$ is a SPNE if at no time t and after no previous play $(\mathbf{c}(0), \ldots, \mathbf{c}(t-1))$, any player i can gain a positive benefit from playing any alternative strategy σ'_i. Let $\mathbb{E}\widehat{\pi}_{i,\delta}(t)$ be the expected continuation payoff with strategy profile $(\sigma_i)_i$, and $\mathbb{E}\widehat{\pi}'_{i,\delta}(t)$ when Player i switches to strategy σ'_i. Both are continuous functions of δ, hence the range of δ where $\mathbb{E}\widehat{\pi}_{i,\delta}(t) \ge \mathbb{E}\widehat{\pi}'_{i,\delta}(t)$ is a closed subset of $[0, 1)$. Taking intersections over all i, σ'_i and t, we obtain the set of all δ such that $(\sigma_i)_i$ is a SPNE in $\Gamma(R, \mathbf{e}, \delta)$, which is therefore also a closed subset of $[0, 1)$. As required in the definition, it is non-empty, and, hence, it has a minimum.

Resilience of a contribution sequence.

Definition 8 (Resilience of a contribution sequence)**.** *For given $(R, \mathbf{e})$, the resilience of a non-zero contribution sequence that is sustainable in $\Gamma(R, \mathbf{e}, \delta)$ for some δ is the infimum of the resilience values of the strategy profiles producing that contribution sequence.*

By "non-zero contribution sequence", we mean any contribution sequence other than $(\mathbf{0})_t$.

The above definition is equivalent to saying that the resilience of $(\mathbf{c}(t))_t$ is equal to the resilience of its associated Grim strategy profile. This is because by Proposition 1, the Grim strategy profile is a SPNE at every δ where any other strategy profile producing $(\mathbf{c}(t))_t)$ is a SPNE. In particular, the infimum in the definition is attained.

From Corollary 1, we know that for given R, a non-zero contribution sequence that is sustainable for some δ with both endowment distributions $\mathbf{e}$ and $\mathbf{e}'$ has the same resilience with both $(R, \mathbf{e})$ and $(R, \mathbf{e}')$. So its resilience is independent of $\mathbf{e}$ (among those values of $\mathbf{e}$ where it is feasible).

As a shorthand, we speak of the resilience of an endowment distribution to mean the resilience of full cooperation under that endowment distribution, whenever it is sustainable:

Definition 9 (Resilience of an endowment distribution)**.** *For given $(R, \mathbf{e})$, if full cooperation, i.e. the constant contribution sequence $(\mathbf{e})_t$, is sustainable, we refer to the resilience of $(\mathbf{e})_t$ as the resilience of $\mathbf{e}$ itself under productivity matrix R.*

The endowment distribution that maximises resilience in this sense for a given R is, by definition, equal to the "endowment distribution most conducive to cooperation" as defined in [HHCN19]. We refer to it as the *resilience-maximising endowment distribution* and denote it by $\mathbf{e}^*$. In the following section, we show that it is unique, that full cooperation under $\mathbf{e}^*$ indeed has resilience $\delta^*_{\min}$, and derive the value of $\mathbf{e}^*$ for general n and R.

General case

Definition 10 (Minimal resilience)**.** *Let R be a productivity matrix that allows for cooperation. Then we define $\delta^*_{\min}$ as the infimum of all δ such that (R, δ) allows for cooperation.*

In other words, $\delta^*_{\min}$ is the infimum of the resilience values attained with productivity matrix R.

Theorem 2. *In a given game $\Gamma(R, \mathbf{e}, \delta)$, let $(\mathbf{c}(t))_t$ be a non-zero sustainable contribution sequence that has resilience $\delta^*_{\min}$. Let λ and $\mathbf{v}$ be the Perron eigenvalue and eigenvector of the zero-diagonal productivity matrix D associated with R. Then $\mathbf{c}(t)$ converges to a multiple of $\mathbf{v}$ as $t \to \infty$. Furthermore, $\delta^*_{\min} = \lambda^{-1}$, and all entries of $\mathbf{v}$ are strictly positive.*

Proof. Let $n \geq 3$. D has non-negative entries, and D^2 is positive. So D is a primitive matrix and we can apply the Perron-Frobenius Theorem [Mey00]: D has a unique, non-repeated eigenvalue λ on the spectral circle, which is real and positive (hence equal to the spectral radius). Furthermore, the associated eigenvector can be taken with positive entries, and

$$\lim_{m \to \infty} \frac{D^m}{\lambda^m} = \mathbf{v}\mathbf{w}^\intercal, \tag{3.13}$$

where $\mathbf{v} > \mathbf{0}$ and $\mathbf{w} > \mathbf{0}$ are the right and left eigenvectors, respectively, taken such that $\mathbf{w}^\intercal \mathbf{v} = ||\mathbf{v}|| = 1$.

Furthermore, we know that

$$\min_i \sum_j D_{ij} \leq \lambda \leq \max_i \sum_j D_{ij}. \tag{3.14}$$

A proof of (3.14) is found in [BP94] on page 37. We insert for D_{ij} and get

$$\begin{aligned} \min_i \sum_{j \neq i} \frac{R_{ij}}{1 - R_{ii}} &\leq \lambda \leq \max_i \sum_{j \neq i} \frac{R_{ij}}{1 - R_{ii}} \\ \min_i \frac{\sum_j R_{ij} - R_{ii}}{1 - R_{ii}} &\leq \lambda \leq \max_i \frac{\sum_j R_{ij} - R_{ii}}{1 - R_{ii}}. \end{aligned} \tag{3.15}$$

Since $\sum_j R_{ij} > 1$ for all i, we can conclude that $1 < \lambda$.

By assumption, the non-zero contribution sequence $(\mathbf{c}(t))_t$ is sustainable in $\Gamma(D, \mathbf{e}, \delta)$. From Theorem 1, we get that

$$\bar{\mathbf{c}}(0) \leq (\delta D)^t \bar{\mathbf{c}}(t)$$

for all t. But $\bar{\mathbf{c}}(t) \leq \mathbf{e}$ and D is non-negative, so

$$\bar{\mathbf{c}}(0) \leq (\delta D)^t \mathbf{e},$$

or alternatively

$$\bar{\mathbf{c}}(0) \leq (\delta\lambda)^t \frac{D^t}{\lambda^t} \mathbf{e},$$

for all t. From (3.13), we know that $\frac{D^t}{\lambda^t}$ converges as $t \to \infty$. So $\delta\lambda < 1$ implies $\bar{\mathbf{c}}(0) = 0$. We therefore know that (R, δ) does not allow for cooperation for any $\delta < \lambda^{-1}$.

Now take $\delta = \lambda^{-1} < 1$. We have $\mathbf{v} = \delta D \mathbf{v}$, so $(\mathbf{c}(t))_t = (\bar{\mathbf{c}}(t))_t = (\mathbf{v})_t$ is sustainable by Corollary 5. So R allows for cooperation and $\delta^*_{\min} = \lambda^{-1}$. We can now ask which non-zero contribution sequences other than $(\mathbf{v})_t$ are also maximally resilient, i.e. sustainable at $\delta = \lambda^{-1}$.

Let $(\mathbf{c}(t))_t$ be such a sequence. We have

$$\bar{\mathbf{c}}(t) \leq \frac{D^\tau}{\lambda^\tau} \bar{\mathbf{c}}(t + \tau)$$

for all t, τ, where $\bar{\mathbf{c}}(t)$ means $\bar{\mathbf{c}}_\delta(t)$ with $\delta = \lambda^{-1}$. So for all $\varepsilon > 0$, there is a T such that for all $\tau \geq T$,

$$\bar{\mathbf{c}}(t) \leq (1 + \varepsilon)\ \mathbf{v}\mathbf{w}^\intercal\ \bar{\mathbf{c}}(t + \tau).$$

This follows from (3.13). Let $L = \liminf_\tau\ \mathbf{w}^\intercal \bar{\mathbf{c}}(\tau)$. Then

$$\bar{\mathbf{c}}(t) \leq (1 + \varepsilon) L \mathbf{v}$$

for all $\varepsilon > 0$, hence $\bar{\mathbf{c}}(t) \leq L\mathbf{v}$.

Since $\mathbf{w} > \mathbf{0}$, we can multiply with $\mathbf{w}$ on both sides:

$$\mathbf{w}^\intercal \bar{\mathbf{c}}(t) \leq L \mathbf{w}^\intercal \mathbf{v} = L$$

for all t. Hence, $\mathbf{w}^\intercal \bar{\mathbf{c}}(t)$ converges to L, which we can write as $\mathbf{w}^\intercal (L\mathbf{v} - \bar{\mathbf{c}}(t)) \to 0$. Since $L\mathbf{v} - \bar{\mathbf{c}}(t) \geq \mathbf{0}$ for all t, we know that $L\mathbf{v} - \bar{\mathbf{c}}(t) \to \mathbf{0}$ as $t \to \infty$. Then, $\bar{\mathbf{c}}(t)$ converges to a multiple of $\mathbf{v}$. $(1 - \delta)\mathbf{c}(t) = \bar{\mathbf{c}}(t) - \delta \bar{\mathbf{c}}(t + 1)$, hence the same is true for $\mathbf{c}(t)$.

In the special case that $n = 2$, D is non-negative and irreducible, but not primitive. We can still find λ, $\mathbf{v}$ and $\mathbf{w}$. It is easy to ascertain manually that $\lambda > 1$. Instead of (3.13), we get

$$\lim_{m\to\infty} \frac{D^{2m}}{\lambda^{2m}} = \mathbf{v}\mathbf{w}^\intercal,$$

which is sufficient to prove $\delta^*_{\min} = \lambda^{-1}$ and that $\bar{\mathbf{c}}(2t)$ and $\bar{\mathbf{c}}(2t+1)$ converge to multiples of $\mathbf{v}$. Since

$$\bar{\mathbf{c}}(2t) \leq \lambda^{-1} D\bar{\mathbf{c}}(2t+1) \leq \lambda^{-2} D^2 \bar{\mathbf{c}}(2t+2),$$

they must be the same multiple. □

Resilience-maximising constant contribution sequence. In particular, we have

Corollary 6. *In a given game $\Gamma(R, \mathbf{e}, \delta)$, the constant sustainable non-zero contribution sequences which have resilience $\delta^*_{\min}$ are exactly those of the form $(\mu\mathbf{v})_t$, where $\mathbf{v}$ is the Perron eigenvector of D and $\mu > 0$ is such that $\mu v_i \leq e_i$ for all i.*

Proof. It is a direct consequence of Theorem 2 that any constant sustainable non-zero contribution sequence with $\delta^*_{\min}$ must be of the form $(\mu\mathbf{v})_t$ for some $\mu > 0$. It is feasible exactly if μ satisfies the specified condition.

To show the converse, we use the fact that $\delta^*_{\min} = \lambda^{-1}$, also from Theorem 2: Take some feasible $(\mu\mathbf{v})_t$. Then for $\delta = \delta^*_{\min}$, Inequality 3.10 of Theorem 1 holds, so $(\mu\mathbf{v})_t$ is sustainable. □

Resilience-maximising endowment distribution. Usually, in a public good games model, we are interested in finding conditions for full cooperation to be an equilibrium. Full cooperation means that all players contribute their whole endowment in every round, i.e. $(\mathbf{c}(t))_t = (\mathbf{e})_t$. Our Corollary 6 implies that the resilience-maximising endowment distribution is $\mathbf{e}^* = \mathbf{v}$:

Corollary 7. *For given (R, δ), the most resilient endowment distribution is $\mathbf{e}^* = \mathbf{v}$, where $\mathbf{v}$ is the Perron eigenvector of D (chosen such that $\sum_i v_i = 1$). In particular, all players are allocated a strictly positive endowment.*

Sustainability of positive contributions.

Corollary 8. *In any game $\Gamma(R, \mathbf{e}, \delta)$ with $(\forall i) e_i > 0$ that allows for cooperation, positive contributions (meaning $c_i(t) > 0$ for all i and all t) are sustainable.*

Proof. $\Gamma(R, \mathbf{e}, \delta)$ allows for cooperation exactly if $\delta \geq \delta^*_{\min}$. If this is the case, then $(\mathbf{c}(t))_t = (\mu\mathbf{v})_t$ is sustainable for some $\mu > 0$. The Perron eigenvector $\mathbf{v}$ is positive. □

Focusing on non-excludable public goods

So far, we have derived our results for general R. But we are mainly interested in the case of a non-excludable public good, or equal sharing, where $R_{ij} = n^{-1}r_j$. We will now restrict to equal sharing and make use of our main result, Theorem 2.

Proposition 2. *In a given game* $\Gamma(\mathbf{r}, \mathbf{e}, \delta)$*, the minimal resilience* $\delta^*_{\min}$ *is bounded by*

$$\frac{n - \min_i r_i}{\sum_i r_i - \min_i r_i} \leq \delta^*_{\min} \leq \frac{n - \max_i r_i}{\sum_i r_i - \max_i r_i} < \frac{n-1}{\sum_i r_i - 1}.$$

Proof. For $n > 2$, we get from (3.15) that

$$\begin{aligned}
\min_i \frac{\sum_j n^{-1}r_j - n^{-1}r_i}{1 - n^{-1}r_i} &\leq \lambda \leq \max_i \frac{\sum_j n^{-1}r_j - n^{-1}r_i}{1 - n^{-1}r_i} \\
\min_i \frac{n - r_i}{\sum_j r_j - r_i} &\leq \lambda^{-1} \leq \max_i \frac{n - r_i}{\sum_j r_j - r_i} \\
\frac{n - \min_i r_i}{\sum_i r_i - \min_i r_i} &\leq \delta^*_{\min} \leq \frac{n - \max_i r_i}{\sum_i r_i - \max_i r_i}.
\end{aligned}$$

since $\sum_i r_i > n$. The looser upper bound follows from $r_i > 1$ for all i.

It is easy to verify that the same inequalities also hold for $n = 2$. □

Characterisation of the relative inequality of $\mathbf{e}^*$. By Theorem 2, we have

$$\mathbf{e}^* = \delta^*_{\min} D \mathbf{e}^*.$$

This implies that

$$\forall i \quad e^*_i = \delta^*_{\min} \sum_{j \neq i} \frac{r_j}{n - r_i} e^*_j,$$

and hence

$$\forall i \quad e^*_i(n - (1 - \delta^*_{\min})r_i) = \delta^*_{\min} \mathbf{r}^\intercal \mathbf{e}^*.$$

We can conclude the following relationship between e^*_i and r_i.

$$e^*_i \propto \frac{1}{n - (1 - \delta^*_{\min})r_i}$$

Corollary 9. *The relationship between* e^*_i *and* r_i *is order preserving.*

We can further derive the maximum relative difference in endowments between players in the group to characterise the degree of inequality of $\mathbf{e}^*$ better.

Proposition 3. *The relative difference between two players' endowments under distribution* $\mathbf{e}^*$ *is at most*

$$\frac{\max_k r_k - 1}{n - \max_k r_k}.$$

By "relative difference" we mean that if $e_i^* \leq e_j^*$, then

$$\frac{e_j^*}{e_i^*} \leq 1 + \frac{\max_k r_k - 1}{n - \max_k r_k}.$$

Proof. Let i, j be such that $e_i^* \leq e_j^*$ and thus also $r_i \leq r_j$, as observed above. Then

$$\frac{e_j^*}{e_i^*} = \frac{n - (1 - \delta_{\min}^*) r_i}{n - (1 - \delta_{\min}^*) r_j} \leq \frac{n - r_i}{n - r_j} \leq \frac{n - 1}{n - \max_k r_k} = 1 + \frac{\max_k r_k - 1}{n - \max_k r_k}.$$

□

There are many scenarios where $\max_k r_k \ll n$. In particular, very large productivities are unrealistic when we interpret productivities as the ratio between the return of an investment in the public good vs. the return of investment in the best outside option. For such a case, Proposition 3 states that approximately equal contributions from all players are optimal. However, while approximately uniform, the optimal distribution is of course still order preserving, as stated by Corollary 9. This argument can also be extended to a more general case. It can be seen that for large n, the entries D_{ij} of the zero-diagonal productivity matrix become almost independent of i as

$$D_{ij} = \frac{1}{n} \frac{r_j}{1 - \frac{1}{n} r_i} = \frac{1}{n} r_j + \left(\frac{1}{n}\right)^2 r_i r_j + \mathcal{O}\left(\frac{1}{n}\right)^3,$$

implying more equal distributions for larger groups.

3.4.4 Efficiency-maximising endowment distribution

While resilience is certainly a desirable property, it cannot be the only objective of a policy maker deciding on wealth redistribution. For example, a non-defective SPNE might have very high resilience even though players only contribute small fractions of their endowments and thus that partial cooperation is of little value. To capture this, we consider the efficiency of cooperation, or total group welfare, defined as the sum of all players' payoffs.

We will ask both when it is maximised under a given endowment distribution, and when it is maximised over all endowment distributions. In what follows, we exclusively focus on equal sharing of the public good, i.e. $R_{ij} = n^{-1} r_j$.

We will show that in a given game $\Gamma(\mathbf{r}, \mathbf{e}, \delta)$, maximal efficiency can always be attained by a constant contribution sequence. The maximal welfare over all $\mathbf{e}$ can, and can only be attained by a constant contribution sequence.

Definition 11 (Welfare)**.** *The welfare $\Phi(\widehat{\mathbf{c}})$ of a contribution vector $\widehat{\mathbf{c}}$ under productivities $\mathbf{r}$ is defined as the sum of payoffs*

$$\Phi(\widehat{\mathbf{c}}) := \sum_i \widehat{\pi}_i = 1 + (\mathbf{r} - \mathbf{1})^\intercal \widehat{\mathbf{c}}.$$

Note that Φ is a function of $\widehat{\mathbf{c}} = \bar{\mathbf{c}}(0)$, but as a function of $(\mathbf{c}(t))_t$ it also depends on δ.

Definition 12 (Sustainability of welfare values)**.** *A welfare (a real number) Φ is sustainable in a given game $\Gamma(\mathbf{r}, \mathbf{e}, \delta)$ if there exists a sustainable contribution vector $\widehat{\mathbf{c}}$ such that $\Phi = \Phi(\widehat{\mathbf{c}})$.*

We write $\Phi_{\max}(\mathbf{e})$ for the maximal sustainable welfare in $\Gamma(\mathbf{r}, \mathbf{e}, \delta)$, and $\Phi^\dagger_{\max} = \max_{\mathbf{e}} \Phi_{\max}(\mathbf{e})$. Below, we show that these maxima are well defined.

Lemma 2. *In any given game $\Gamma(\mathbf{r}, \mathbf{e}, \delta)$, the sustainable welfare values attain a maximum $\Phi_{\max}(\mathbf{e})$, and this maximum is attained with a constant contribution sequence.*

Proof. We recursively define the sequence $(\mathbf{x}(s))_s$ by $\mathbf{x}(0) = \mathbf{e}$ and

$$x_i(s+1) = \min\{\ x_i(s)\ \ ,\ \ \delta(D\mathbf{x}(s))_i\ \}$$

for all i and all $s \geq 0$. (Note that $(\mathbf{x}(s))_s$ is not interpreted as a contribution sequence.) Take any sustainable contribution sequence $(\mathbf{c}(t))_t$. By induction on s, we have $\bar{\mathbf{c}}(t) \leq \mathbf{x}(s)$ for all t and s: It is clearly true for $s = 0$, since $\mathbf{c}(t) \leq \mathbf{e}$ for all t. Now assume it is true for all t and some fixed s. Then we can also write $\bar{\mathbf{c}}(t+1) \leq \mathbf{x}(s)$ for all t. Since D is a non-negative matrix, we have as well $\delta D\bar{\mathbf{c}}(t+1) \leq \delta D\mathbf{x}(s)$ for all t. So, by Theorem 1, $\bar{\mathbf{c}}(t) \leq \delta D\mathbf{x}(s)$ for all t. Together with $\bar{\mathbf{c}}(t) \leq \mathbf{x}(s)$, this means that $\bar{\mathbf{c}}(t) \leq \mathbf{x}(s+1)$. So the statement is true for all t and all s.

Since $\mathbf{x}(s)$ is a decreasing sequence and bounded below by $\mathbf{0}$, it converges to some limit $\mathbf{x}$. This limit $\mathbf{x}$ also has the property that $\bar{\mathbf{c}}(t) \leq \mathbf{x}$ for all t. For all s, we have $\mathbf{x} \leq \mathbf{x}(s+1) \leq \delta D\mathbf{x}(s)$. Since $\delta D\mathbf{x}(s) \to \delta D\mathbf{x}$ as $s \to \infty$, we can conclude $\mathbf{x} \leq \delta D\mathbf{x}$.

Because $\mathbf{x} \leq \mathbf{e}$, the constant contribution sequence $(\mathbf{x})_t$ is feasible under endowment distribution $\mathbf{e}$, and by Corollary 5 it is sustainable. Since $\bar{\mathbf{c}}(0) = \widehat{\mathbf{c}} \leq \mathbf{x}$, we have $\Phi(\widehat{\mathbf{c}}) \leq \Phi(\mathbf{x})$. So $(\mathbf{x})_t$ has maximal welfare under endowment distribution $\mathbf{e}$. □

Lemma 3. *For any pair $(\mathbf{r}, \delta)$ that allows for cooperation, there is an endowment distribution $\mathbf{e}^\dagger$ such that full cooperation is sustainable in $\Gamma(\mathbf{r}, \mathbf{e}^\dagger, \delta)$ and achieves a maximal sustainable welfare $\Phi^\dagger_{\max}$.*

This means that the choice of $\mathbf{e} = \mathbf{e}^\dagger$ and $(\mathbf{c}(t))_t = (\mathbf{e}^\dagger)_t$ maximises welfare over all possible $\mathbf{e}$ and $(\mathbf{c}(t))_t$ sustainable in $\Gamma(\mathbf{r}, \mathbf{e}, \delta)$. Note that in special cases, $\mathbf{e}^\dagger$ might not be unique.

Proof. Lemma 2 states that for all $\mathbf{e}$, the maximal welfare $\Phi_{\max}(\mathbf{e})$ in $\Gamma(\mathbf{r}, \mathbf{e}, \delta)$ is attained by a constant contribution sequence. Of course, this means that any sustainable welfare in $\Gamma(\mathbf{r}, \mathbf{e}, \delta)$ is attained by a sustainable constant contribution sequence, by scaling down the maximal-welfare sequence.

So the values of Φ that are sustainable in $\Gamma(\mathbf{r}, \mathbf{e}, \delta)$ are exactly the values $\Phi(\widehat{\mathbf{c}})$ where $(\widehat{\mathbf{c}})_t$ is a sustainable constant contribution sequence. The $\widehat{\mathbf{c}}$ which are sustainable for some $\mathbf{e}$ form the compact set $\{\widehat{\mathbf{c}} \in \mathbb{R}^n_{\geq 0} | \sum_i \widehat{c}_i \leq 1 \wedge \mathbf{0} \leq (\delta D - I_n)\widehat{\mathbf{c}}\}$. The welfare function $\Phi(\widehat{\mathbf{c}}) = 1 + (\mathbf{r} - \mathbf{1})^\intercal \widehat{\mathbf{c}}$ is continuous, so it attains a maximum on this set. So there is a constant sustainable contribution sequence $(\widehat{\mathbf{c}})_t$ that maximises Φ. Choose $\mathbf{e}$ such that $(\widehat{\mathbf{c}})_t$ is sustainable. We must have $\widehat{\mathbf{c}} = \mathbf{e}$ by maximality of Φ.

If a non-constant contribution sequence, sustainable in $\Gamma(\mathbf{r}, \mathbf{e}, \delta)$ for some $\mathbf{e}$, were to have the same maximal efficiency, then by Lemma 2 some constant $(\widehat{\mathbf{c}})_t$ with $\widehat{\mathbf{c}} \leq \mathbf{e}$ would also attain the maximum for $\mathbf{e}$. Again we must have $\widehat{\mathbf{c}} = \mathbf{e}$. The non-constant sequence however has strictly smaller contributions than $\mathbf{e}$ in at least one round, therefore its welfare is strictly smaller. That is a contradiction, hence no such contribution sequence exists. □

Linear program formulation

We refer to $\mathbf{e}^\dagger$ as *the efficiency-maximising endowment distribution*. To find $\mathbf{e}^\dagger$ and $\Phi^\dagger_{\max}$, we maximise $\Phi(\mathbf{e})$ over all endowment distributions $\mathbf{e}$ with which full cooperation is sustainable. We can express that optimisation as a linear program:

$$\begin{aligned} \Phi^\dagger_{\max} = \max_{\mathbf{e}} \quad & \mathbf{r}^\intercal \mathbf{e} \\ \text{s.t.} \quad & (\delta D - I_n)\mathbf{e} \geq \mathbf{0} \\ & \mathbf{1}^\intercal \mathbf{e} = 1 \\ & \mathbf{e} \geq \mathbf{0} \end{aligned}$$

The maximal welfare for a given endowment $\mathbf{e}$ is by Lemma 2 the solution to a similar linear program that has the constraint $\widehat{\mathbf{c}} \leq \mathbf{e}$ instead of $\mathbf{1}^\intercal \mathbf{e} = 1$:

$$\begin{aligned} \Phi = \max_{\widehat{\mathbf{c}}} \quad & 1 + (\mathbf{r} - \mathbf{1})^\intercal \widehat{\mathbf{c}} \\ \text{s.t.} \quad & (\delta D - I_n)\widehat{\mathbf{c}} \geq \mathbf{0} \\ & \widehat{\mathbf{c}} \leq \mathbf{e} \\ & \widehat{\mathbf{c}} \geq \mathbf{0} \end{aligned}$$

The efficiency-maximising endowment distribution has the property stated below. Informally summarised, the players with low productivity have an endowment of 0, while the players with high productivity are indifferent between cooperation and defection. In the general case, there is exactly one player whose productivity lies between those two groups. In edge cases, that player may not exist, or there may be multiple players who have an equal intermediate productivity.

Theorem 3. *Let w.l.o.g. $r_1 \leq \cdots \leq r_n$. Let $\mathbf{e}^\dagger$ be any endowment distribution such that $\Phi(\mathbf{e}^\dagger) = \Phi^\dagger_{\max}$. Then there exist integers $k_1, k_2 \in \{1, \ldots, n\}$ such that*

$$e^\dagger_i = 0 \quad \textit{for all} \quad i < k_1$$

$$r_i = r_j \quad \textit{for all} \quad i, j \quad \textit{such that} \quad k_1 \leq i \leq j \leq k_2,$$

and

$$((\delta D - I_n)\mathbf{e}^\dagger)_i = 0 \quad \textit{for all} \quad i > k_2.$$

In other words, out of the two inequality constraints on $\mathbf{e}^\dagger$, which are $\mathbf{e}^\dagger \geq \mathbf{0}$ and $(\delta D - I_n)\mathbf{e}^\dagger \geq \mathbf{0}$, the first one has equality in columns $1, \ldots, (k_1 - 1)$ and the second one has equality in columns $(k_2 + 1), \ldots, n$, while $r_{k_1} = \cdots = r_{k_2}$.

Proof. Given $\mathbf{e}^\dagger$ with $\Phi(\mathbf{e}^\dagger) = \Phi^\dagger_{\max}$, define k_1 as the maximal k_1 such that $e^\dagger_i = 0$ for all $i < k_1$ and k_2 as the minimal k_2 such that $((\delta D - I_n)\mathbf{e}^\dagger)_i = 0$ for all $i > k_2$. This is always possible, since these conditions are satisfied when $k_1 = 1$ and $k_2 = n$.

Now assume we have i, j such that $k_1 \leq i \leq j \leq k_2$. It is enough to show that $r_i = r_j$, so assume otherwise. We must have $r_i < r_j$. Since $i \geq k_1$, we have

$$e^\dagger_i > 0.$$

Since $j \leq k_2$, we have $((\delta D - I_n)\mathbf{e}^\dagger)_j > 0$, which we can also write as

$$(n - r_j)e^\dagger_j < \delta \sum_{l \neq j} r_l e^\dagger_l.$$

Choose some $\varepsilon > 0$ small enough such that when we define $\mathbf{e}$ by $e_i = e^\dagger_i - \varepsilon$ and $e_j = e^\dagger_j + \varepsilon$ and $e_l = e^\dagger_l$ in all other components, then $\mathbf{e}$ still satisfies the inequalities $e_i > 0$ and $(n - r_j)e_j < \delta \sum_{l \neq j} r_l e_l$.

Clearly, the constraint $\mathbf{e} \geq \mathbf{0}$ is satisfied in all components. Since $r_i < r_j$, we have $r_i e_i + r_j e_j > r_i e^\dagger_i + r_j e^\dagger_j$. For this reason, the constraint $(\delta D - I_n)\mathbf{e} \geq \mathbf{0}$, which can be written as $(n - r_l)e_l \leq \delta \sum_{l' \neq l} r_{l'} e_{l'}$ for all l, is satisfied not only in the jth component, which was by construction of ε, but also in all other components other than i. In the ith component, it is true for a simpler reason, which is that $(n - r_i)e_i < (n - r_i)e^\dagger_i$ while $\sum_{l \neq i} r_l e_l > \sum_{l \neq i} r_l e^\dagger_l$. Trivially, $\sum_l e_l = 1$. So $\mathbf{e}$ satisfies all the constraints of the linear program. However, since $r_i e_i + r_j e_j > r_i e^\dagger_i + r_j e^\dagger_j$, its welfare, which is the objective function of the linear program, is higher than that of $\mathbf{e}^\dagger$; a contradiction. Therefore, we must have $r_i = r_j$. □

If $n = 2$, the above linear program for $\Phi^\dagger_{\max}$ becomes:

$$\delta \frac{r_2}{2 - r_1} e_2 \geq e_1 \tag{3.16}$$

$$\delta \frac{r_1}{2 - r_2} e_1 \geq e_2 \tag{3.17}$$

$$e_1 + e_2 = 1 \tag{3.18}$$

$$e_1 \geq 0 \tag{3.19}$$

$$e_2 \geq 0 \tag{3.20}$$

and $r_1 e_1 + r_2 e_2$ is the objective function. The productivity vector $\mathbf{r}$ allows for cooperation exactly if $\delta \geq \delta^*_{\min}$, that is, if $\delta^2 \frac{r_1 r_2}{(2 - r_1)(2 - r_2)} \geq 1$, otherwise the optimisation problem is infeasible. Assuming that cooperation is indeed sustainable, and w.l.o.g. that $r_1 \geq r_2$, the maximum is attained at

$$e_1^\dagger = \frac{\delta r_2}{2 - r_1 + \delta r_2}, \qquad e_2^\dagger = \frac{2 - r_1}{2 - r_1 + \delta r_2}.$$

This is in the case that we are optimising over $\mathbf{e}$ as well as $\widehat{\mathbf{c}}$.

If instead $\mathbf{e}$ is given and we are finding $\Phi_{\max}(\mathbf{e})$ by varying $\widehat{\mathbf{c}}$, then (3.18) becomes

$$e_1 \geq \widehat{c}_1 \tag{3.21}$$

$$e_2 \geq \widehat{c}_2. \tag{3.22}$$

Again we assume $\delta \geq \delta^*_{\min}$. We define $\boldsymbol{\gamma}_i = (\delta D - I_2)_i$ for $i = 1, 2$, such that (3.16) and (3.17) are equivalent to $\boldsymbol{\gamma}_1 \widehat{\mathbf{c}} \geq 0$ and $\boldsymbol{\gamma}_2 \widehat{\mathbf{c}} \geq 0$, respectively. Clearly, one of (3.21) and (3.22) must be tight, since (3.16) and (3.17) are homogeneous. (3.21) can only be tight if $\boldsymbol{\gamma}_1 \widehat{\mathbf{c}} \geq 0$, and (3.22) only if $\boldsymbol{\gamma}_2 \widehat{\mathbf{c}} \geq 0$. Since we assumed that the problem is sustainable, at least one of those must be true. If exactly one is true, w.l.o.g. if $\boldsymbol{\gamma}_1 \widehat{\mathbf{c}} \geq 0$ and $\boldsymbol{\gamma}_2 \widehat{\mathbf{c}} < 0$, then the optimal solution is $\widehat{c}_1 = e_1$ and $\widehat{c}_2 = \delta \frac{r_1}{2 - r_2} e_1$. Finally, if both are true, then the optimal solution is $\widehat{\mathbf{c}} = \mathbf{e}$.

3.4.5 Trade-off between efficiency and resilience

The efficiency-maximising endowment distribution $\mathbf{e}^\dagger$ is always weakly more unequal than the resilience-maximising endowment distribution $\mathbf{e}^*$. Formally, we state that using the concept of Lorenz dominance.

Lorenz dominance of $\mathbf{e}^*$ over $\mathbf{e}^\dagger$ means that if we define permutations σ_1 and σ_2 such that $(e^*_{\sigma_1(i)})_i$ and $(e^\dagger_{\sigma_2(i)})_i$ are increasing, then

$$\sum_{i=1}^{k} e^*_{\sigma_1(i)} \geq \sum_{i=1}^{k} e^\dagger_{\sigma_2(i)} \tag{3.23}$$

for all $k \leq n$, with strict inequality for some k. It implies that $\mathbf{e}^\dagger$ is more unequal than $\mathbf{e}^*$ under any inequality measure [DSS73]. For example, the Gini coefficient of $\mathbf{e}^\dagger$ is necessarily higher than or equal to the Gini coefficient of $\mathbf{e}^*$.

Theorem 4. *For any given $(\mathbf{r}, \delta)$ allowing for cooperation, either $\mathbf{e}^*$ Lorenz-dominates $\mathbf{e}^\dagger$ or $\mathbf{e}^* = \mathbf{e}^\dagger$. The latter is the case for some choice of $\mathbf{e}^\dagger$ exactly if $\delta = \delta^*_{\min}$ or all productivities are equal.*

Note that while $\mathbf{e}^*$ is always unique, there are edge cases where $\mathbf{e}^\dagger$ is not unique. Of course, "$\mathbf{e}^* = \mathbf{e}^\dagger$ is the case for some choice of $\mathbf{e}^\dagger$" simply means that $\Phi(\mathbf{e}^*) = \Phi^\dagger_{\max}$.

Proof. First, we will show that $\mathbf{e}^*$ weakly Lorenz-dominates $\mathbf{e}^\dagger$, and that in fact we have either strict dominance or $\mathbf{e}^* = \mathbf{e}^\dagger$. Then, we will show the conditions for $\mathbf{e}^* = \mathbf{e}^\dagger$.

Assume w.l.o.g. that $(r_i)_i$ is non-decreasing. Then, by Corollary 9, so is $(e^*_i)_i$. Choose k_1 and k_2 as in Theorem 3.

First, we want to show that $(e_i^\dagger)_i$, too, is non-decreasing. Trivially, if $i < k_1$, then $e_i^\dagger \leq e_j^\dagger$ for all j. The same is also true for all $i \leq j$ when $k_2 < j$: We have $r_i < r_j$ and

$$(n - (1-\delta)r_i)e_i^\dagger \leq \delta \sum_l r_l e_l^\dagger$$

and

$$(n - (1-\delta)r_j)e_j^\dagger = \delta \sum_l r_l e_l^\dagger,$$

while $r_i \leq r_j$, so indeed $e_i^\dagger \leq e_j^\dagger$. To exhaust all (i, j) with $i \leq j$, what remains is the case $k_1 \leq i \leq j \leq k_2$. Here, we may assume it w.l.o.g. to be true. So $e_i^\dagger \leq e_j^\dagger$ for all $i \leq j$. We may thus choose $\sigma_1 = \sigma_2 = \mathrm{id}_{[n]}$ in (3.23).

We can now show Lorenz dominance of $\mathbf{e}^*$ over $\mathbf{e}^\dagger$ by verifying that (3.23) holds for all k. For $1 \leq k < k_1$, we have

$$\sum_{i=1}^{k} e_{\sigma_2(i)}^\dagger = \sum_{i=1}^{k} e_i^\dagger = 0,$$

so (3.23) holds.

For $k_2 \leq k \leq n$, we have

$$\sum_{i=1}^{k} e_{\sigma_2(i)}^\dagger = 1 - \sum_{i=k+1}^{n} e_i^\dagger \leq 1 - \sum_{i=k+1}^{n} e_i^* = \sum_{i=1}^{k} e_{\sigma_1(i)}^* \tag{3.24}$$

For all $i > k_2$, we have that

$$(n - r_i)e_i^* = \delta_{\min}^* \sum_{j \neq i} r_j e_j^*$$

and

$$(n - r_i)e_i^\dagger = \delta \sum_{j \neq i} r_j e_j^\dagger,$$

which we can also write as

$$({\delta_{\min}^*}^{-1}(n - r_i) + r_i)e_i^* = \Phi(\mathbf{e}^*)$$

and

$$(\delta^{-1}(n - r_i) + r_i)e_i^\dagger = \Phi(\mathbf{e}^\dagger).$$

By definition, $\delta^{-1} \leq {\delta_{\min}^*}^{-1}$ and $\Phi(\mathbf{e}^*) \leq \Phi(\mathbf{e}^\dagger)$. So $e_i^* \leq e_i^\dagger$, with equality exactly if $\delta = \delta_{\min}^*$ and $\Phi(\mathbf{e}^*) = \Phi_{\max}^\dagger$.

Now assume there is some k with $k_1 \leq k \leq k_2$ such that (3.23) does not hold. Take the smallest such k. By minimality, we must have $e^*_k < e^\dagger_k$. For every l such that $k \leq l \leq k_2$, we have $r_k = r_l$, so also $e^*_k = e^*_l$ by uniqueness of $\mathbf{e}^*$. Since $e^\dagger_k \leq e^\dagger_l$, it follows that $e^*_l < e^\dagger_l$. But then

$$\sum_{i=1}^{k_2} e^*_{\sigma_1(i)} = \sum_{i=1}^{k} e^*_i + \sum_{i=k+1}^{k_2} e^*_i < \sum_{i=1}^{k} e^\dagger_i + \sum_{i=k+1}^{k_2} e^\dagger_i = \sum_{i=1}^{k_2} e^\dagger_{\sigma_2(i)},$$

which is a contradiction to (3.24). So there can be no such k; (3.23) holds for all k.

We have shown weak Lorenz dominance. The distribution $\mathbf{e}^*$ weakly but not strictly dominates $\mathbf{e}^\dagger$ exactly if it is a permutation of $\mathbf{e}^\dagger$ [DSS73]. But since we have been able to assume w.l.o.g. that both vectors are non-decreasing with the order of their indices, that is equivalent to $\mathbf{e}^* = \mathbf{e}^\dagger$.

It remains to determine the conditions for $\mathbf{e}^* = \mathbf{e}^\dagger$. Assume equality holds. Then $\Phi(\mathbf{e}^*) = \Phi^\dagger_{\text{max}}$, so by Theorem 3, we can select k_1 and k_2 with the properties stated in the Proposition. By Corollary 7, $\mathbf{e}^*$ is positive, so $k_1 = 1$. We distinguish two cases. If $k_2 < n$, then we have equality in the nth component of the constraint $(\delta D - I_n)\mathbf{e}^\dagger \geq \mathbf{0}$. But also $(\delta^*_{\text{min}} D - I_n)\mathbf{e}^* = \mathbf{0}$. So we must have $\delta = \delta^*_{\text{min}}$. If on the other hand $k_2 = n$, then by the properties of k_1 and k_2, all productivities are equal.

To show the converse, consider first the case that $\delta = \delta^*_{\text{min}}$. We know that when $\delta = \delta^*_{\text{min}}$, $\mathbf{e}^*$ is the only endowment distribution that allows for full cooperation, so it has maximal welfare. Consider instead the case that all productivities are equal. Now, $\Phi(\mathbf{e})$ is independent of $\mathbf{e}$, so again, $\mathbf{e}^*$ has maximal welfare. So the conditions are exact. □

Data and software availability

Figures in this chapter are based on simulation averages over many independent runs of the respective simulation. Results were analysed and visualised with Python and Matlab R2023a. The computer code is published in [HK24]. The parameters used are described in Section 3.3.

Author contribution statement

All authors designed research; V.H., M.S., and M.K. performed research; V.H. and M.K. analysed data; and all authors wrote the manuscript. K.C. and M.K. contributed equally to this work.

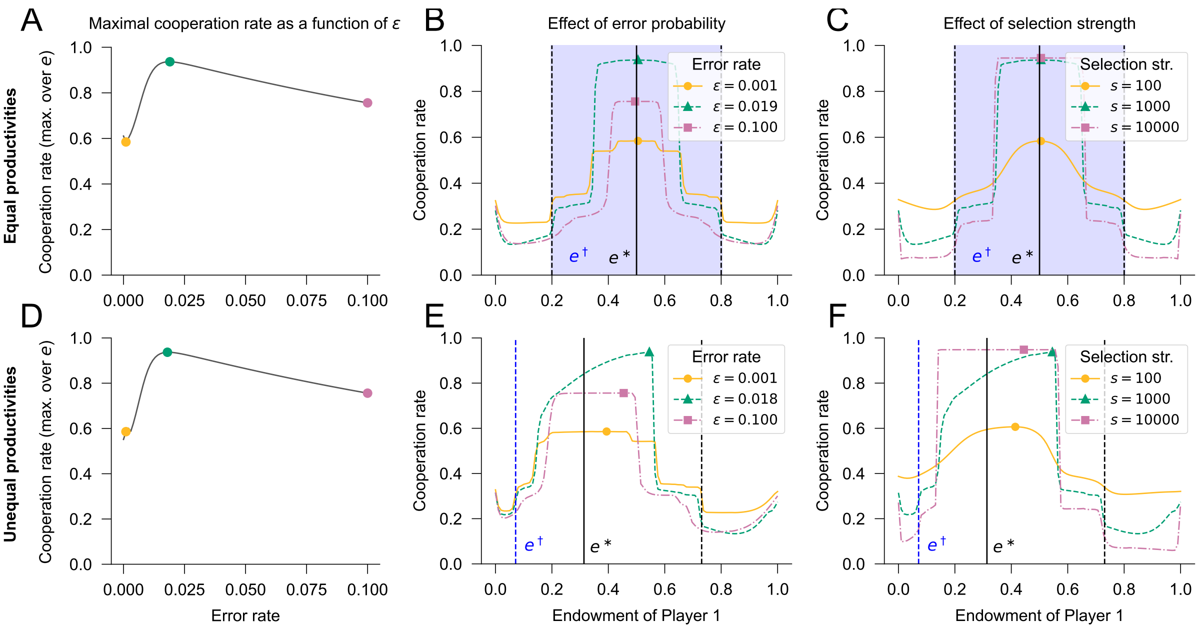

Figure 3.5: Evolutionary simulations of the group cooperation rates. We report similar data from the evolutionary simulations as in Fig. 4, but for cooperation rates rather than welfare. A,B,C, The case with equal productivities shows qualitatively similar results as for the group welfare. However, the case of unequal productivities differs with respect to where the maxima are observed. For example, as a function of error rates, the cooperation is highest at $e_{\text{maxC}} = (0.55, 0.45)$. D,E,F, We find that cooperation is most likely to evolve among two sufficiently equal players even if their productivities differ.

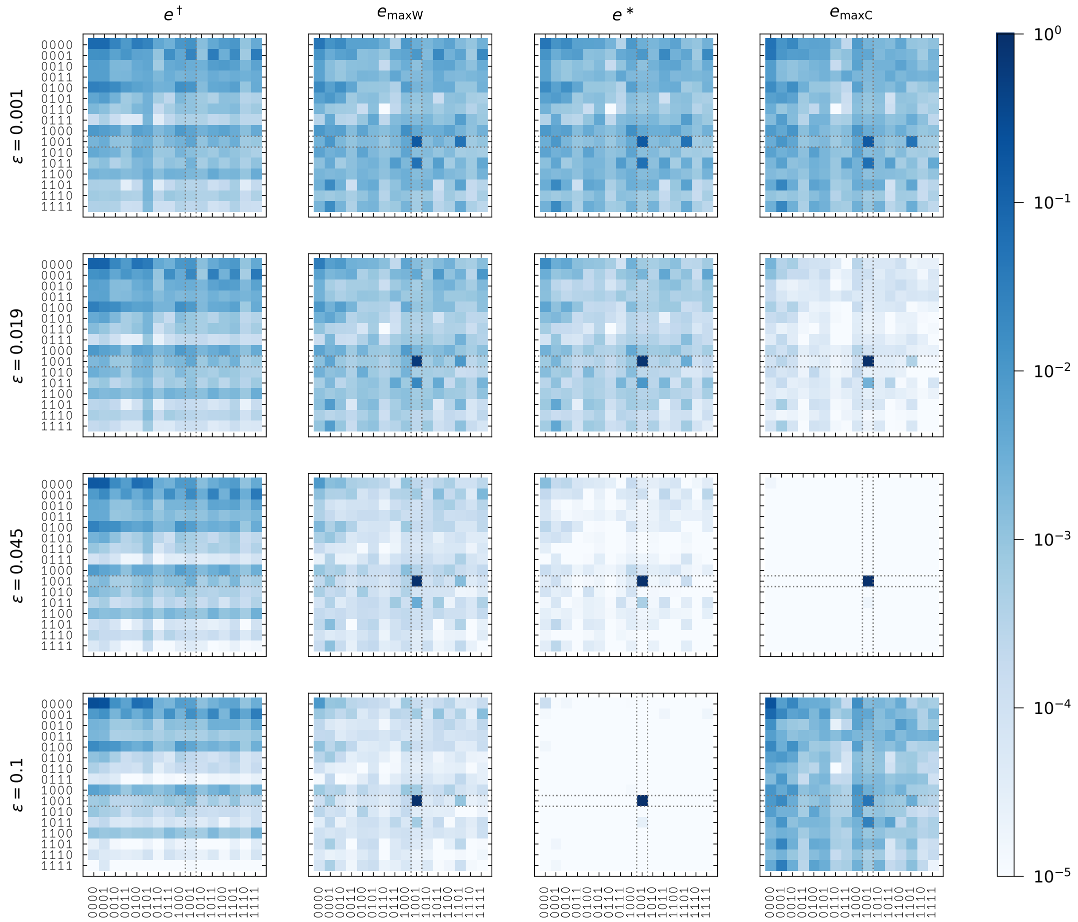

Figure 3.6: Evolving strategies. To gain better understanding of the endowment distributions that maximise cooperation, we analyse evolving strategies for the parameters used in Fig. 4 and 3.5. We first report evolving strategies for the efficiency-maximising endowment distribution at $\delta = 1$ for three different error rates. We find that non-cooperating strategies are very likely to evolve with a very low frequency of WSLS players. This explains low cooperation. We compare this to the endowment where actual maximal social welfare was observed. While we do observe higher frequency of WSLS players, there is also a lot of noise from other strategies. For the resilience-maximising endowment distribution the noise from other strategies is decreasing and the frequency of WSLS players is increasing. Yet, the highest frequency of WSLS players is observed at the endowment distribution with the highest cooperation rates.

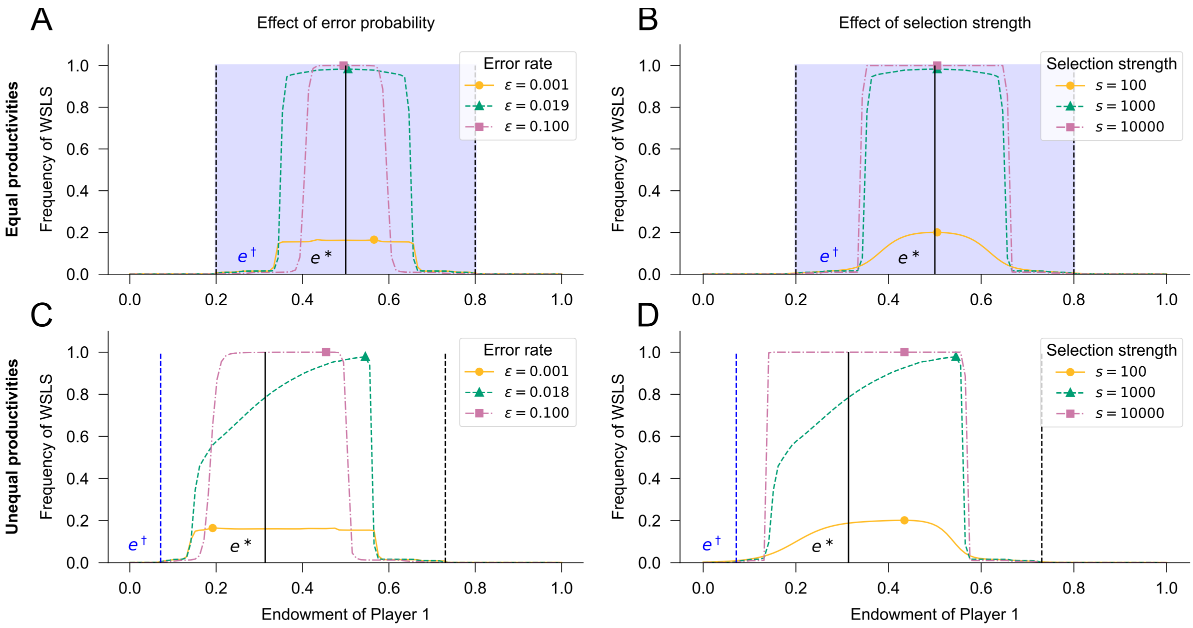

Figure 3.7: The frequency of WSLS. We report the frequency of the strategy profile (WSLS, WSLS), that is where both players play the strategy Win-Stay-Lose-Shift. The parameters are the same as in Figure S2, with intermediate error rates chosen to maximise cooperation rate.

CHAPTER 4

Time-dependent strategies in repeated asymmetric public goods games

This chapter was originally published as: Valentin Hübner, Christian Hilbe, Manuel Staab, Maria Kleshnina, and Krishnendu Chatterjee. Time-dependent strategies in repeated asymmetric public goods games. *Dynamic Games and Applications*, February 2025

Cooperation is typically conceptualised as a behaviour that is costly for the individual but beneficial to the group [RN13]. Examples of cooperation abound, ranging from small favours among friends to collective efforts to mitigate climate change. These cooperative interactions can, and have been, described with game theory [BR13, LM23]. This literature has produced rich predictions about potential mechanisms that can sustain cooperation [Now06a, KI09]. One such mechanism is direct reciprocity [Tri71, Axe81, BR88b, Sig10]. Here, individuals are assumed to engage in the same interaction repeatedly, over many rounds. Repeated interactions allow players to condition their behaviour on the previous history of play. In this way, they can enforce mutual cooperation despite any short-run temptations to free-ride [Fri71, FM86b].

Traditionally, many models of direct reciprocity, especially in the evolutionary game theory literature, assume that interactions are symmetric [NS92, HS97, KDK99, DH05, IN10, KWI10, AS11, VSPLS12, PGGS$^+$13, SP14a, PHRZ15, SP16, HCN18, GFGGvV24]. This means that players are completely interchangeable with respect to their actions and feasible payoffs. More recently, however, the evolution of cooperation among asymmetric players has received more attention [HTN14, VSPL14, ACT14, Aki15, MH15, Kur16, ACBS$^+$18, MSMS21, CGH22, CP23]. This interest has also been spurred by empirical studies that explore the role of inequality in controlled experiments [CMMM99, BC06, NSRC15, HRS16, GMTV17, KMS17, SSG$^+$23, WCW$^+$23].

Oftentimes, these studies are based on some variation of the linear public goods game. In this game, players obtain their fixed endowments in the beginning of each round. Then they independently decide how much of their endowment they wish to contribute to the public good. Contributions are multiplied by some productivity factor, and the resulting amount is evenly split among all group members. There are various ways to allow for asymmetry in this game. For example, players may have unequal endowments, unequal productivities, or both. The main takeaway from the above-mentioned studies is that endowment inequality tends to be detrimental to cooperation. However, as shown by [HHCN19] and [HSH+24], there can be exceptions. If individuals already differ in their productivity, it can become easier to sustain full cooperation if they also differ in their endowments. As a rule of thumb, a player's endowment ought to be larger the more productive that player is.

However, the studies of [HHCN19] and [HSH+24] consider a rather restricted question. They ask: Under which conditions are there subgame perfect equilibria in which all players contribute their full endowment in every round? In particular, they thereby only consider equilibria whose resulting contribution sequence along the equilibrium path is constant. Instead, in the following we are interested in contribution sequences that can vary in time. We ask: Once individuals have the ability to make time-dependent contributions along the equilibrium path, to which extent can they achieve outcomes that are infeasible with constant contribution sequences?

To this end, we study three different but related models of public good provision. The first is most convenient from a mathematical perspective. Here, players can make arbitrary (non-negative) contributions each round. In particular, contributions are not constrained by any endowments that individuals might have in that round. Instead, we only require that the players' overall discounted contributions over the entire game are bounded (with the upper bound being arbitrary). We refer to this model as the 'base model'. The second model reproduces typical public goods game models, such as the ones considered in [HHCN19] and [HSH+24]. Here, a player's contribution each round is bounded by the player's assigned endowment. Accordingly, we speak of the 'endowment model'. Finally, the third model is a hybrid of the first two. Here, players obtain a fixed and constant endowment each round. But now they can decide to deposit some of this endowment into a savings account, which is then available in the next round. Under this assumption, contributions each round are not bounded by the players' endowments anymore. Instead, they are bounded by the players' accumulated endowments up to that point. We call this the 'saving model'.

For all three models, we consider the equilibrium outcomes that can be achieved with time-dependent contributions. We compare them to the possible equilibrium outcomes when players are required to make a fixed and constant contribution along the equilibrium path. We make this comparison based on two key quantities. One quantity is the group's welfare in equilibrium (the total sum of the players' payoffs). As we discuss in more detail below, this quantity is particularly relevant in the endowment model and in the saving model. The other quantity is an equilibrium's resource efficiency (the ratio of the group's welfare relative to the players' total contributions). This quantity is relevant for all three models.

We characterise under which condition time-dependent contributions allow for equilibria with larger resource efficiency (compared to equilibria based on constant contributions). We do so for arbitrary discount factors. In particular, we do not require that players are sufficiently patient, as often done in the classical folk theorem literature [Fri71, FM86b]. Our results depend on the group size and on the number of players with the highest productivity. In particular, we find that when there is a unique player with maximum productivity, time-dependent contributions provide an advantage. With respect to welfare maximisation, we find a similar result – but only for the saving model.

4.1 The base model

4.1.1 Model setup

We start with the base model (as illustrated in Fig. 4.1), which we will use to derive our first results. These results will also have important implications for the other two models studied subsequently.

In the base model, a group of $n \geq 2$ players interacts for an indefinite sequence of rounds. In every round t, each player i decides which non-negative amount $c_i(t)$ to contribute towards the public good. There is no limit on how much players may contribute. Hence, $c_i(t) \in \mathbb{R}_{\geq 0}$. These individual contributions can be collected in a vector, $\mathbf{c}(t) = (c_1(t), \ldots, c_n(t))^\intercal$. We refer to a sequence $\big(\mathbf{c}(t)\big)_t$ of contribution vectors as a contribution sequence, or as a 'play' of the game. If $\mathbf{c}(t) = \mathbf{c}(0)$ for all times t, the contribution sequence is called constant. Otherwise, it is time-dependent.

Contributions of each player i are multiplied by their productivity factor r_i and added to the public good. The total public good is then evenly shared among all players. This is a slight generalisation of the standard formulation of the public goods game, according to which every player has the same productivity. In the following, we use $\mathbf{r} = (r_1, \ldots, r_n)^\intercal$ to denote the vector of all productivities. Based on this notation, we can write payer i's payoff $\pi_i(t)$ in round t as

$$\pi_i(t) = \frac{1}{n}\mathbf{r}^\intercal \mathbf{c}(t) - c_i(t). \tag{4.1}$$

We assume productivities satisfy $1 < r_i < n$ for all players i. The first inequality $r_i > 1$ ensures that the group's total payoff (across all members) is increasing in each player's contributions. The second inequality, $r_i < n$, on the other hand, ensures that each individual is tempted to give as little as possible. Together, these two inequalities render the game a social dilemma. For each player, there is a conflict between their private interest and the collective interest of the group.

To define the players' payoffs over the entire repeated game, we assume players value each subsequent round at a discount of δ, with $0 < \delta < 1$. Accordingly, when the contribution sequence $\big(\mathbf{c}(t)\big)_t$ is bounded, we define the total payoff of each player as the weighted sum of their payoffs each round,

$$\hat{\pi}_i = (1-\delta)\sum_{t=0}^{\infty} \delta^t \pi_i(t).$$

Here, the term $1 - \delta$ serves as a normalising factor. It ensures that repeated-game payoffs are comparable to the game's one-shot payoffs. Because we assumed the contribution sequence to be bounded, the above sum is guaranteed to converge. We do not define a total payoff for unbounded contribution sequences.

In a common alternative interpretation of repeated games with discounting, which is applicable to our base model and model variation I, but not model variation II, the number of rounds is finite and random, with δ being not a discount factor but the round-wise continuation probability. This means that after each round, with probability δ the game continues for at least one more round and with probability $1 - \delta$ it ends. In that interpretation, all rounds have equal value given that they are played, and the expected number of rounds is $1/(1 - \delta)$. Intuitively, this suggests that higher values of δ, which mean longer games, are conducive to cooperation.

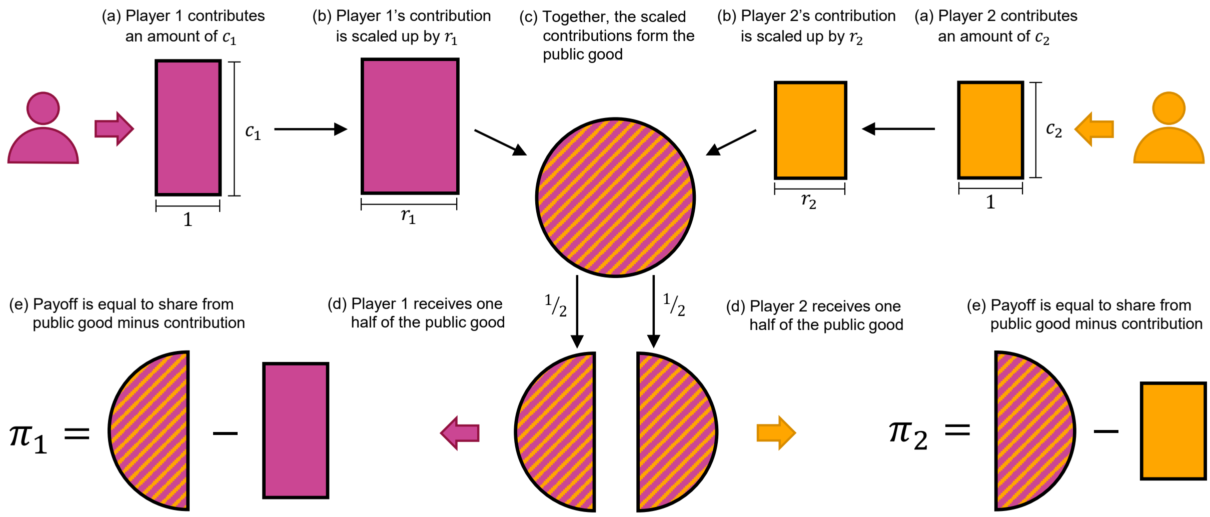

Figure 4.1: A one-round asymmetric public goods game. To illustrate the base model, we consider $n=2$ players. They freely choose the size of their contributions, c_1 and c_2. Contributions are enhanced by the individual productivity factors, which are $r_1=1.5$ and $r_2=1.1$ in this example. The size of the public good is the sum of all enhanced contributions, $\mathbf{r}^\intercal\mathbf{c}=r_1c_1+r_2c_2$. Each player receives an equal share of this sum. The payoff of each player then equals their share of the public good minus their contribution.

Notation 1. *In this base model, a game is fully specified by the players' productivities* $\mathbf{r}$ *and by the discount factor* δ. *We denote the corresponding game as* $\Gamma_{\mathrm{B}}(\mathbf{r},\delta)$.

For our subsequent analysis, it will be useful to consider the weighted sum of a player's contributions after a given time t. Formally, these continuation contributions of player i are defined as

$$\bar{c}_i(t) = (1-\delta)\sum_{\tau=0}^{\infty}\delta^\tau c_i(t+\tau). \tag{4.2}$$

One can also define a sequence that collects the respective continuation contributions for each round, $\big(\bar{\mathbf{c}}(t)\big)_t = \big(\bar{\mathbf{c}}(0), \bar{\mathbf{c}}(1), \bar{\mathbf{c}}(2), \dots\big)$. We call $\big(\bar{\mathbf{c}}(t)\big)_t$ the continuation contribution sequence associated with contribution sequence $(\mathbf{c}(t))_t$. Every contribution sequence uniquely specifies a continuation contribution sequence and vice versa. Analogously, we can also define continuation payoffs at time t,

$$\bar{\pi}_i(t) = (1-\delta)\sum_{\tau=0}^{\infty}\delta^\tau \pi_i(t+\tau).$$

We write $\bar{\boldsymbol{\pi}}(t) = \big(\bar{\pi}_1(t), \dots, \bar{\pi}_n(t)\big)$ for the respective vector, and note that $\hat{\boldsymbol{\pi}} = \bar{\boldsymbol{\pi}}(0)$. Similarly, we write $\hat{\mathbf{c}} = \bar{\mathbf{c}}(0)$ for the continuation contributions at time zero. We call $\hat{\mathbf{c}}$ the total contribution vector. By linearity of the one-round payoffs (4.1), we have

$$\hat{\pi}_i = \frac{1}{n}\mathbf{r}^\intercal\hat{\mathbf{c}} - \hat{c}_i. \tag{4.3}$$

That is, each player's total payoff is uniquely determined by the total contribution vector. By definition, this payoff $\hat{\pi}_i$ is the quantity that player i aims to maximise.

Players make their decisions based on their strategies. A strategy σ_i for player i is a function that assigns to each initial contribution sequence $\big(\mathbf{c}(0), \mathbf{c}(1), \dots, \mathbf{c}(t-1)\big)$ a next contribution value $c_i(t)=\sigma_i\Big((\mathbf{c}(\tau))_{\tau<t}\Big)\in\mathbb{R}_{\geq 0}$. A strategy is called bounded if it always produces a bounded contribution sequence (irrespective of the co-players' strategies). An assignment of one strategy to each player $(\sigma_i)_i$ is called a strategy profile. A strategy profile is bounded if all its strategies are bounded.

4.1.2 Sustainable contribution sequences

In the following, we are particularly interested in those strategy profiles that form a subgame perfect equilibrium [SPE, or 'equilibrium', see FT91]. We say a bounded strategy profile is in equilibrium when no player has an incentive to deviate, after no finite sequence of moves. Formally, $(\sigma_i)_i$ is in equilibrium if there is no initial contribution sequence $\mathbf{c}(0),\ldots,\mathbf{c}(t-1)$ such that some player j could get a larger payoff by deviating towards another bounded strategy σ_j^* after that time t. For a given game $\Gamma_{\mathrm{B}}(\mathbf{r},\delta)$, we call a contribution sequence sustainable if it is the contribution sequence of some equilibrium strategy profile. A total contribution vector $\hat{\mathbf{c}}$ is sustainable if it is the total contribution vector of a sustainable contribution sequence.

To derive our main results, we make extensive use of the previously published Theorem 5 below. This theorem gives us a comfortable characterisation of sustainable contribution sequences.

Theorem 5 ([HSH+24])**.** *For a given game $\Gamma_{\mathrm{B}}(\mathbf{r},\delta)$, define an associated $n\times n$ matrix $D=(D_{ij})$, called the productivity matrix in zero-diagonal form, by*

$$D_{ij}=\begin{cases} r_j/(n-r_i) & \text{if } i\neq j\\ 0 & \text{if } i=j.\end{cases} \tag{4.4}$$

Then a contribution sequence $(\mathbf{c}(t))_t$ is sustainable if and only if the associated continuation contributions satisfy

$$\bar{\mathbf{c}}(t)\leq\delta D\bar{\mathbf{c}}(t+1)\quad\text{for all } t. \tag{4.5}$$

With Theorem 1, we can determine whether a given contribution sequence is sustainable by checking if the associated continuation contribution sequence $(\bar{\mathbf{c}}(t))_t$, as defined by (4.2), satisfies (4.5) for all t. The two following corollaries are immediate consequences of this theorem.

Corollary 10. *For any game $\Gamma_{\mathrm{B}}(\mathbf{r},\delta)$, the set of sustainable contribution sequences $(\mathbf{c}(t))_t$, the set of sustainable continuation contribution sequences $(\bar{\mathbf{c}}(t))_t$, and the set of sustainable total contribution vectors $\hat{\mathbf{c}}$ are closed under addition and multiplication by a non-negative scalar (that is, they are convex cones).*

This convexity result implies that whether or not a contribution sequence is sustainable only depends on the relative magnitude of the players' contributions. This result holds because payoffs depend linearly on contributions. Therefore, scaling all contributions up or down by the same positive factor does not affect whether or not the equilibrium conditions are satisfied.

Corollary 11. *In a given game $\Gamma_{\mathrm{B}}(\mathbf{r},\delta)$, a constant contribution sequence of $(\hat{\mathbf{c}})_t$ is sustainable if and only if*

$$\hat{\mathbf{c}}\leq\delta D\hat{\mathbf{c}}. \tag{4.6}$$

So we have a set of n linear constraints that defines the set of feasible constant contribution sequences. We can use Corollary 11 to derive a version of the folk theorem, applied to our setup. The folk theorem famously relates the possible equilibrium payoffs in the repeated game to the properties of the one-shot payoffs [Fri71, FM86b]. To state our version, we note that in our public goods game, players can always guarantee a non-negative payoff (by not contributing anything). Hence, we say a contribution vector $\hat{\mathbf{c}}$ for the one-shot game is individually rational if it yields a non-negative payoff to each player.

Theorem 6 (Folk theorem of repeated games)**.** *Constant contributions* $(\hat{\mathbf{c}})_t \in \mathbb{R}_{\geq 0}$ *are sustainable in the game* $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$ *for sufficiently large* δ *if and only if* $\hat{\mathbf{c}}$ *is individually rational.*

The condition for a contribution vector $\hat{\mathbf{c}}$ to be individually rational can be written as

$$\max_i \hat{c}_i \leq \frac{1}{n}\mathbf{r}^\intercal \hat{\mathbf{c}}. \tag{4.7}$$

An equivalent formulation of the folk theorem is therefore: the constant contribution sequence $(\hat{\mathbf{c}})_t$ is sustainable for sufficiently large δ if and only if Eq. (4.7) holds.

The above results allow us to characterise the properties of sustainable contribution sequences. Perhaps one of the most important properties is whether or not the contribution sequence entails at least some cooperation. More specifically, we define a play to be non-defective if at least one player makes a positive contribution in at least one round (i.e., $\hat{\mathbf{c}} \neq \mathbf{0}$). Otherwise we call the play defective. It is easy to see that for non-defection to be sustainable, not just one, but at least two players have to make positive contributions. This is because a hypothetical lone non-defector would benefit from deviating towards full defection. If non-defection is sustainable in a given game $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$, then we say that $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$ allows for non-defection. Any productivity vector $\mathbf{r}$ (satisfying the general requirement $r_i > 1$ for all i) allows for non-defection when δ is sufficiently large [HHCN19, Supplementary Information, Proposition 2].

4.1.3 Welfare and resource efficiency

While the binary distinction between defection and non-defection is useful, not all forms of non-defection are equally desirable. After all, even non-defective contribution sequences might result in payoffs arbitrarily close to the full defection payoff of zero. Therefore, in the following we introduce two other key metrics of interest. The first metric is the (overall) welfare W of a given play, which equals the sum of all payoffs,

$$W = \sum_{i=1}^{n} \hat{\pi}_i. \tag{4.8}$$

This welfare can be expressed as a function of the total contribution vector $\hat{\mathbf{c}}$ as

$$W(\hat{\mathbf{c}}) = (\mathbf{r} - \mathbf{1})^\intercal \hat{\mathbf{c}}. \tag{4.9}$$

By this equation, non-defective plays have $W > 0$, whereas defective plays have $W = 0$.

By Corollary 10, any non-defective contribution sequence can be scaled arbitrarily without affecting their sustainability. It follows that our base model allows for arbitrary welfares.

As an alternative metric that is still relevant with unlimited resources, we measure how efficiently the players are able to use them. The resource efficiency E of a non-defective play is defined as the sum of all payoffs divided by the sum of all contributions,

$$E = \frac{\sum_{i=1}^{n} \hat{\pi}_i}{\sum_{i=1}^{n} \hat{c}_i}. \tag{4.10}$$

Resource efficiency, too, is a function of the total contribution vector:

$$E(\hat{\mathbf{c}}) = \frac{\mathbf{r}^\intercal \hat{\mathbf{c}}}{\mathbf{1}^\intercal \hat{\mathbf{c}}} - 1. \tag{4.11}$$

The first term on the right hand side of Eq. (4.11) can be slightly rewritten, as

$$\frac{\mathbf{r}^\intercal \hat{\mathbf{c}}}{\mathbf{1}^\intercal \hat{\mathbf{c}}} = \sum_{i=1}^{n} \frac{c_i}{c_1 + \ldots + c_n} \cdot r_i.$$

This representation allows us to interpret this term as a weighted mean of the players' productivities; the weights correspond to the players' contributions. In particular, this observation implies that $E(\hat{\mathbf{c}})$ is always in between $\min_i r_i - 1$ and $\max_i r_i - 1$. Whether or not the upper bound (or equivalently, the lower bound) can be realised depends on which players contribute in equilibrium. For example, the upper bound can be realised if and only if there is an equilibrium in which only those players i with $r_i = \max_j r_j$ make a contribution. That is only possible if there are multiple players with maximum productivity.

Example. It is instructive to illustrate these concepts with a two-player game (which we continue to use throughout this article). Consider the game $\Gamma_{\mathrm{B}}\big((1.5, 1.1)^\intercal, 0.9\big)$. That is, there are $n=2$ players with productivities $r_1=1.5$ and $r_2=1.1$ (as in Fig. 4.1), and the discount factor is $\delta=0.9$. In this example, the value of the matrix D is

$$D = \begin{pmatrix} 0 & r_2/(2-r_1) \\ r_1/(2-r_2) & 0 \end{pmatrix} = \begin{pmatrix} 0 & 11/5 \\ 5/3 & 0 \end{pmatrix} \approx \begin{pmatrix} 0 & 2.2 \\ 1.666 & 0 \end{pmatrix}.$$

Suppose player 1 makes a constant contribution $\hat{c}_1=7$ in every round, whereas player 2 makes the constant contribution $\hat{c}_2=5$. It follows that the total size of the public good is $\mathbf{r}^\intercal\hat{\mathbf{c}}=1.5\cdot 7 + 1.1\cdot 5=16$. Thus, each player's share of the public good is 8, and their payoffs according to Eq. (4.3) are $\hat{\pi}_1=8-7=1$ and $\hat{\pi}_2=8-5=3$. Because both payoffs are non-negative, the respective constant contribution sequences are individually rational. Hence, by the folk theorem, they are sustainable in the repeated game for sufficiently large δ. For this case of $n=2$, Eq. (4.6) in Corollary 11 takes the form of the following system of inequalities:

$$\hat{c}_1 \le \delta \frac{r_2}{2-r_1} \hat{c}_2 \tag{4.12}$$

$$\hat{c}_2 \le \delta \frac{r_1}{2-r_2} \hat{c}_1 \tag{4.13}$$

With these, we can verify that the given discount factor $\delta=0.9$ is indeed sufficiently large. According to Eq. (4.8), the resulting welfare is $W=3+1=4$, and according to Eq. (4.10), resource efficiency is $E=4/12\approx 0.333$. If contributions were ten times larger, welfare would increase to 40 but the resource efficiency would remain the same.

Instead, suppose now that the two players contribute equal constant amounts, say $\hat{c}_1 = \hat{c}_2 = 6$. By the inequalities (4.12–4.13), this contribution vector is also sustainable. It yields a payoff of 1.8 for each player. Thus, the welfare is $W = 3.6$, whereas resource efficiency is $E = 0.3$. We conclude that equal contributions are less resource efficient, compared to the previous example with unequal contributions. This is intuitive because in the previous example, the more productive player 1 contributed a larger share.

In light of these observations, it is natural to ask what the optimal ratio of the two player's contributions is, if we aim to maximise resource efficiency in equilibrium. Again from the inequalities (4.12–4.13), we see that when $r_1 > r_2$ (as in our example), this ratio is given by

$$\frac{\hat{c}_1}{\hat{c}_2} = \delta \frac{r_2}{2 - r_1}.$$

If for example $\hat{c}_1 = 10$, then $\hat{c}_2 = 500/99 \approx 5.051$, which yields payoffs of $\hat{\pi}_1 \approx 0.278$ and $\hat{\pi}_2 \approx 5.227$. The resulting welfare is $W \approx 5.505$ and resource efficiency is $E^{\mathrm{c}}_{\mathrm{sup}} \approx 0.366$. Note that this maximum resource efficiency depends on the discount factor δ. If δ were larger, even higher efficiencies would be possible. In contrast, if δ were lower, either $E^{\mathrm{c}}_{\mathrm{sup}}$ would be lower, or the game might not allow for non-defection at all.

In the above example, we only considered the simple case of constant contributions. We now address the question of whether we can achieve higher resource efficiency when contribution sequences are allowed to be time-dependent.

4.1.4 Efficiency with time-dependent contributions

If we only consider the binary distinction between whether or not a game allows for non-defection, then constant and time-dependent contributions are equally effective. Specifically, if a game has any non-defective equilibrium at all, then it also has a non-defective equilibrium with a constant contribution sequence [Corollary 6 of HSH+24]. However, below we show that within the space of non-defective equilibria, time-dependent contributions can indeed enable outcomes that are not sustainable otherwise. To this end, we first introduce some notation.

Notation 2. *For two vectors* $\mathbf{v}$ *and* $\mathbf{w}$*, we write* $\mathbf{v} \leq_1 \mathbf{w}$ *if* $v_i \leq w_i$ *for all* i *and* $v_i = w_i$ *for at most one* i*. In other words,* $\mathbf{v} \leq_1 \mathbf{w}$ *corresponds to* $\mathbf{v} \leq \mathbf{w}$ *with equality in at most one component.*

Using this notation, we can characterise which total contribution vectors are sustainable with time-dependent contribution sequences.

Theorem 7. *Let* $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$ *allow for non-defection. Then the total contribution vector* $\hat{\mathbf{c}}$ *is sustainable in game* $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$ *if and only if either* $\hat{\mathbf{c}} = \mathbf{0}$ *or*

$$\hat{\mathbf{c}} \leq_1 D\hat{\mathbf{c}}. \tag{4.14}$$

Furthermore, all sustainable total contribution vectors $\hat{\mathbf{c}}$ *are sustainable with a continuation contribution sequence* $(\bar{\mathbf{c}}(t))_t$ *that satisfies* $\hat{\mathbf{c}} \leq \bar{\mathbf{c}}(t)$ *for all* t*.*

The proof of Theorem 7, and of all subsequent results, is provided in the Appendix.

Theorem 7 characterises the set of total contribution vectors that are sustainable in a given game $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$. The theorem is analogous to Corollary 11, which allowed for constant contribution sequences only. Interestingly, however, condition (4.6) in Corollary 11 depends on the discount rate δ. In contrast, condition (4.14) is independent of δ; the only requirement of the theorem is that δ be sufficiently large to allow for non-defection in the first place. That is, suppose the discount factor δ is large enough for the game to allow for *some* non-defective total contribution vector $\hat{\mathbf{c}}$ to be sustainable. Then it automatically allows for *all* vectors that satisfy $\hat{\mathbf{c}} \leq_1 D\hat{\mathbf{c}}$.

The last part of the theorem states that the relevant contribution sequences can be chosen such that future contributions are always at least as large as the past contributions. This statement is included as a technical result that will be useful later on (in the proof of Proposition 6).

As a special case we obtain the following result on the sustainability of equal contributions across all players.

Corollary 12. *Let $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$ allow for non-defection. Then for any $\lambda \geq 0$, the total contribution vector $\hat{\mathbf{c}} = \lambda\mathbf{1}$ is sustainable.*

Importantly, for this statement to be true, time-dependent contributions are essential. If instead players are restricted to make constant contributions along the equilibrium path, there exist games $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$ where no equal contribution vector $\lambda\mathbf{1}$ is sustainable [HHCN19]. Fig. 4.2 illustrates these results. For our 2-player game with $r_1 = 1.5$ and $r_2 = 1.1$, it shows the region of total contribution vectors that are sustainable with time-dependent contribution sequences (Fig. 4.2a), which is independent of δ, and the region of total contribution vectors that are sustainable with constant contribution sequences for a fixed value of δ (for $\delta = 0.55$ in Fig. 4.2b and $\delta = 0.9$ in Fig. 4.2c). Fig. 4.2a was obtained from Eq. (4.14), which is independent of δ. Fig. 4.2b–c was obtained from the analogous condition for constant contribution sequences, Eq. (4.6), which depends on δ. The figure shows that for a low discount factor like $\delta = 0.55$, equal total contributions are feasible with a time-dependent contribution sequence, as predicted by Corollary 12, but not with a constant contribution sequence.

The contribution vectors $\hat{\mathbf{c}}$ that are individually rational in the one-round game (in the sense of the folk theorem) are those where $\hat{\mathbf{c}} \leq D\hat{\mathbf{c}}$. Theorem 7 shows that almost all of those (all except for a set with measure zero) are sustainable once the discount factor δ is sufficiently large to allow for *any* non-defective equilibrium. This requirement is considerably weaker than the one typically used in the folk theorem literature, where δ is thought to approach 1. In the case of two players, the above result holds for strictly all individually rational contribution vectors, as stated by Corollary 13 below.

Corollary 13. *Let $n = 2$ and let $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$ allow for non-defection. Then the total contribution vector $\hat{\mathbf{c}}$ is sustainable exactly if $\hat{\mathbf{c}} \leq D\hat{\mathbf{c}}$.*

Corollary 13 follows in two steps. First, we note that equality, $\hat{\mathbf{c}} = D\hat{\mathbf{c}}$, is impossible unless $\hat{\mathbf{c}} = \mathbf{0}$. To see why, first observe that $\hat{\mathbf{c}} = D\hat{\mathbf{c}}$ is equivalent to $n\hat{c}_i = \sum_j r_j\hat{c}_j$ for all players i. Summing up over all i and dividing by the group size n gives $\sum_j \hat{c}_j = \sum_j r_j\hat{c}_j$. Unless $\hat{\mathbf{c}} = \mathbf{0}$, this is a contradiction, because all r_j are larger than one. The second step is straightforward: given that $\hat{\mathbf{c}} \neq D\hat{\mathbf{c}}$, the statements $\hat{\mathbf{c}} \leq D\hat{\mathbf{c}}$ and $\hat{\mathbf{c}} \leq_1 D\hat{\mathbf{c}}$ are equivalent for $n = 2$.

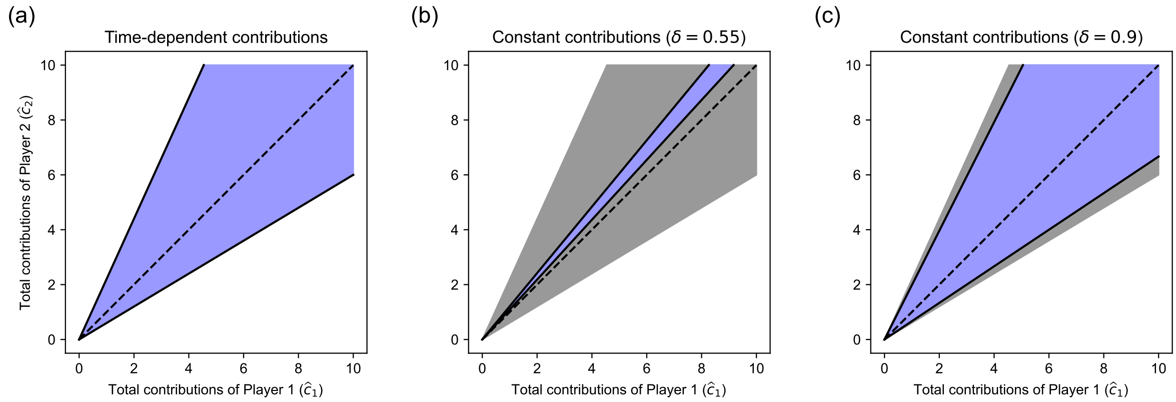

Figure 4.2: The region of sustainable total contribution vectors in a two-player game with $r_1 = 1.5$ and $r_2 = 1.1$. (a) The blue area shows the total contribution vectors that are sustainable with time-dependent contributions as long as δ is sufficiently large to allow any kind of non-defection, which in this case means $\delta \geq 0.522$. The dashed line represents equal contributions. (b) The blue area shows the total contribution vectors that are sustainable with constant contributions in the game $\Gamma_\mathrm{B}(\mathbf{r}, 0.55)$. The grey area shows the total contribution vectors sustainable with time-dependent contributions, for comparison. (c) Like (b), but for the game $\Gamma_\mathrm{B}(\mathbf{r}, 0.9)$. The higher value of δ allows for a larger set of total contribution vectors to be sustainable with constant contributions.

With Theorem 7, we have established that time-dependent contribution sequences can achieve total contribution vectors that are not sustainable with constant contributions. Because payoffs, resource efficiency, and welfare are all functions of the total contribution vector, we can use Theorem 7 to analyse the possible outcomes in terms of those quantities. In particular, the following result suggests that time-dependent contributions can indeed allow for a higher resource efficiency.

Theorem 8. *Let $\Gamma_\mathrm{B}(\mathbf{r}, \delta)$ allow for non-defection, let $r_\mathrm{max} = \max_i r_i$, and let m be the number of players with maximum productivity r_max. There exists a sustainable time-dependent contribution sequence that is more resource-efficient than all sustainable constant contribution sequences if and only if*

$$\Big(1 + \delta(m-1)\Big) \cdot r_\mathrm{max} < n. \tag{4.15}$$

In particular, such a sequence exists if there is only a single player with productivity r_max.

To understand the statement of Theorem 8, consider the case that there is a single player with maximum productivity (i.e., $m = 1$). In that case, (4.15) certainly holds, so the statement of the theorem is simply that resource efficiency cannot be optimised with constant contribution sequences.

Let us also provide some intuition for why the theorem holds, again by considering the case $m = 1$. Without loss of generality, let the unique most productive player be player 1. Let the group play the most resource-efficient sustainable constant contribution sequence (the maximum can indeed be attained). As shown in [HSH+24], this means that a player's contribution is larger the more productive that player is; in particular, player 1's contribution is the largest and hence positive. But in principle, player 1 could deviate and contribute nothing instead. That way, player 1 would get a positive payoff from the other players' contributions in the first round, and a non-negative payoff thereafter. That

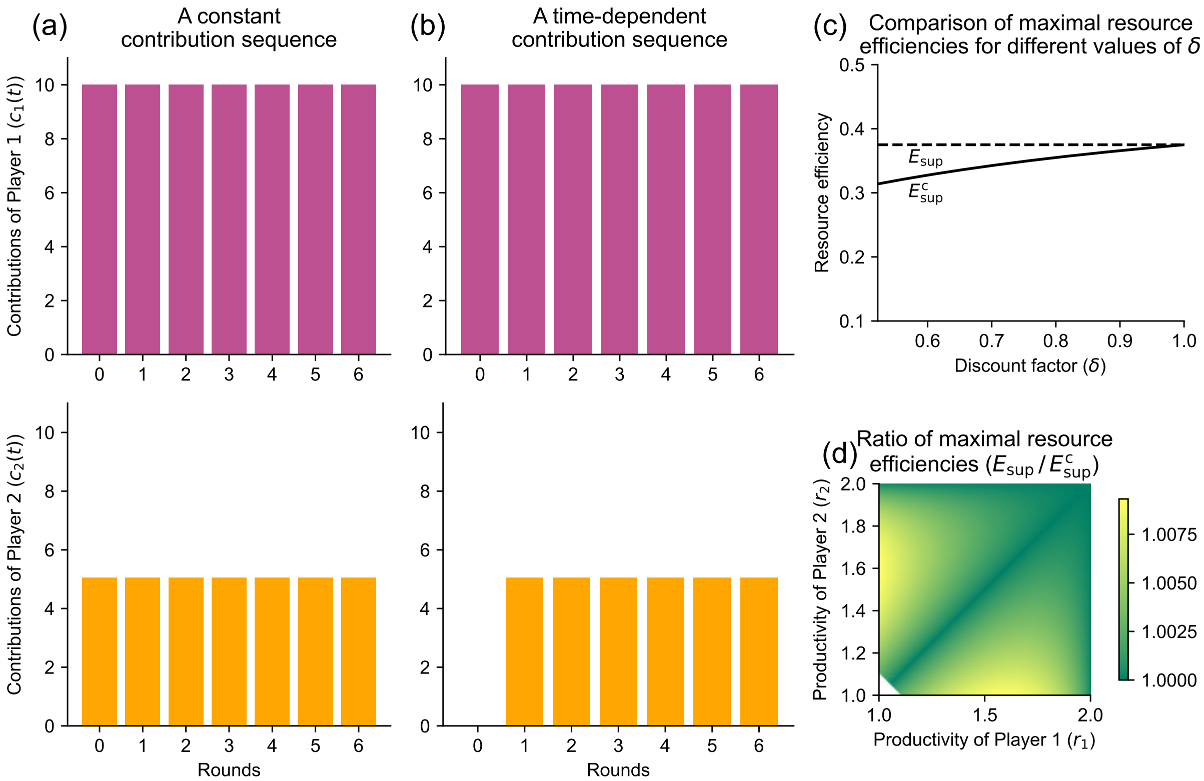

Figure 4.3: Time-dependent contributions allow for greater resource efficiency than constant contributions. Here, we illustrate this result with the game $\Gamma_B\big((1.5, 1.1)^\intercal, 0.9\big)$. (a) When players are required to make constant contributions each round, maximum resource efficiency is attained for $\hat{c}_1 = 10$ and $\hat{c}_2 = 500/99 \approx 5.051$ (or any positive multiple thereof). The attained efficiency is $E^c_{\text{sup}} \approx 0.366$. (b) Players can achieve a better resource efficiency of $E_{\text{sup}} = 0.375$ by switching to a time-dependent contribution sequence. In this example, player 1's contributions are still constant, while player 2's contributions are constant after the first round. (c) For the same productivities as in (a) and (b), we show the optimal resource efficiency with constant contributions (E^c_{sup}, solid line) and with any contributions (E_{sup}, dashed line). (d) Here we depict the ratio $E_{\text{sup}}/E^c_{\text{sup}}$ for different productivities r_1 and r_2, for a discount factor of $\delta = 0.9$. The white area in the bottom left corner is where cooperation is not possible for this value of δ.

is, player 1 would obtain a positive payoff overall. However because of our assumption that the contribution sequence is sustainable, player 1 does not benefit from such a deviation. It follows that player 1's payoff from the constant contribution sequence is as least as good, hence positive. This means that the other players could decide to extract parts of this positive payoff from player 1. For example, they might require that in an initial extra round, player 1 make a solitary contribution. As long as the demanded contribution does not exceed player 1's positive benefit from all subsequent contributions, player 1 would be compelled to go along. Since this player is most productive, and now accounts for a greater share of the contributions, this new scheme increases resource efficiency.

Example. Again, it is instructive to illustrate these considerations with our two-player example, game $\Gamma_B\big((1.5, 1.1)^\intercal, 0.9\big)$. For constant contribution sequences, we have shown that the maximum resource efficiency is $E^c_{\text{sup}} \approx 0.366$. This maximum is attained, for example, when the two players make contributions of $\hat{c}_1 = 10$ and $\hat{c}_2 = 500/99 \approx 5.051$ each round (or any positive multiple thereof), see Fig. 4.3a. By Theorem 8, time-dependent contribution sequences can achieve a superior resource efficiency, because there is a single player with maximum productivity. Fig. 4.3b shows an example of such a time-dependent contribution sequence. Here, in round zero, player 1 contributes $c_1(0) = 10$ whereas player 2 contributes nothing. In all subsequent rounds, players use the numbers of the earlier constant contribution sequence, $c_1(t) = 10$ and $c_2(t) = 500/99 \approx 5.051$ for $t \geq 1$. By Theorem 5, this contribution sequence is sustainable. Moreover, it achieves the best possible resource efficiency $E_{\text{sup}} = 0.375$. This value of E_{sup} does not depend on the discount factor δ, provided δ is sufficiently large to allow for non-defection in the first place (Fig. 4.3c). For the given $\mathbf{r}$, one can compute this minimum discount factor to be $\delta_{\min} \approx 0.522$ [In HSH$^+$24, we show that $\delta_{\min}$ can be computed as the inverse of the largest eigenvalue of matrix D, as defined by Eq. (4.4)].

Interestingly, Theorem 8 also shows that in case of equal productivities for all players, there is no advantage from time-dependent contributions. To see why, observe that for $m = n$ and $\mathbf{r} = (r, r, \ldots, r)^\intercal$, condition (4.15) becomes

$$\Big(1 + \delta(n-1)\Big) \cdot r < n. \tag{4.16}$$

However, by Theorem 2 in the SI Appendix of [HSH$^+$24], $\Gamma_{\text{B}}(r, \delta)$ only allows for non-defection if

$$\delta \geq \frac{n - r}{r(n-1)}, \tag{4.17}$$

Because condition (4.17) is the negation of (4.16), these two conditions are incompatible.

4.2 Model variation I: Endowment constraints

So far, we considered the base model in which players were free to make arbitrary contributions each round. While this model has been convenient to work with, it has at least two disadvantages. First, it requires players to have access to arbitrary amounts of resources, which seems unrealistic. Second, it renders any attempt to optimise the players' welfare meaningless, since players can arbitrarily scale up their contributions (and hence their welfare). In the following, we aim to show how the base model's results extend to two more realistic model variants.

First, we consider a model variation called the endowment model. Here, there is a constant upper limit on player i's contribution $c_i(t)$ in any given round. We refer to this upper limit as the player's endowment, e_i, and we use the notation $\mathbf{e} = (e_1, \ldots, e_n)$ to refer to the vector of all endowments. Endowments are positive and, without loss of generality, normalised such that $\sum_i e_i = 1$. In each round t, each player i chooses which part of their endowment to contribute. That is, they choose some $c_i(t) \in [0, e_i]$. Contributions have the same effect as in the base model. Overall, the players' payoffs consist of their share of the public good, and of their remaining endowment (see Fig. 4.4),

$$\pi_i(t) = \frac{1}{n}\mathbf{r}^\intercal \mathbf{c}(t) + \Big(e_i - c_i(t)\Big).$$

Note that in the case of zero contributions, we now have payoffs of $\hat{\boldsymbol{\pi}} = \mathbf{e}$, as opposed to the base model, where zero contributions give $\hat{\boldsymbol{\pi}} = \mathbf{0}$. However, the definition of welfare

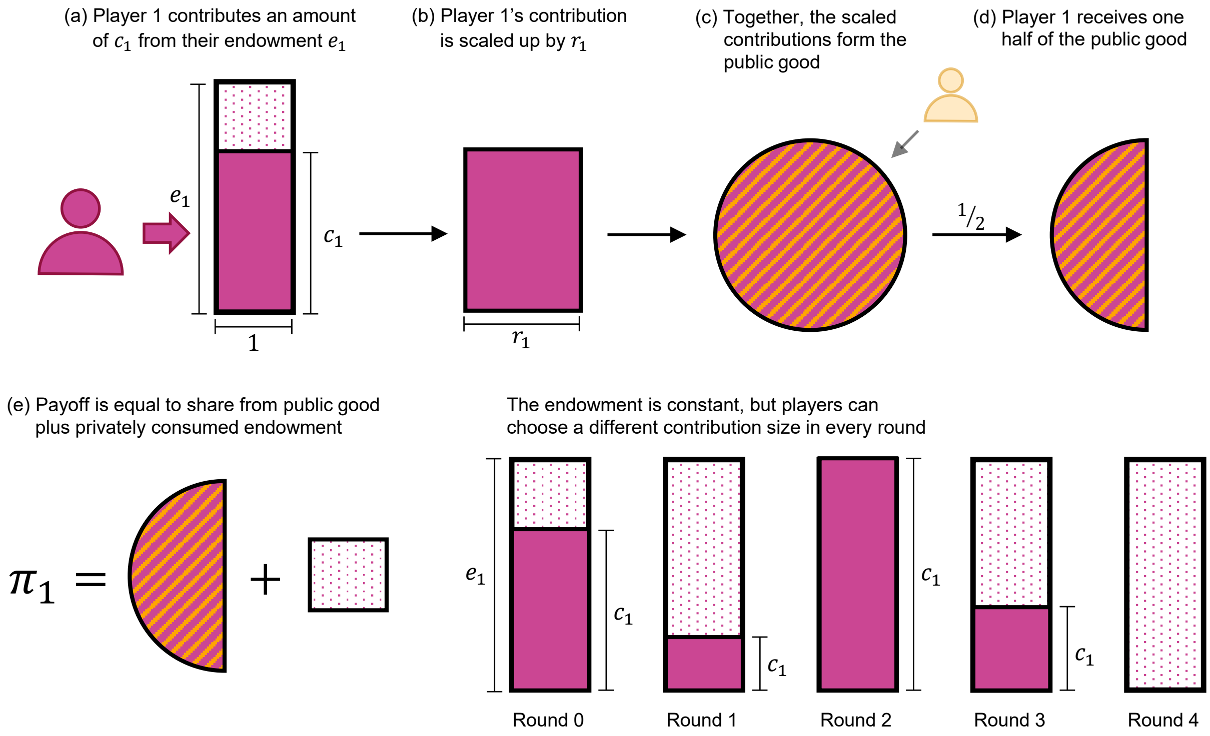

Figure 4.4: The endowment model. Here, the players' contributions each round are constrained by their fixed endowment e_i. (a) In each round, players decide how much of their endowment they wish to contribute to the public good. The remaining amount $e_i - c_i$ is consumed privately. (b–d) The following steps are identical to the base model. In general, player 1 receives the nth share of the public good; in this example, $n = 2$. (e) Players derive a payoff from their share of the public good, and from their private consumption of the remaining endowment.

remains the same as before, $W = \sum_{i=1}^{n} \hat{\pi}_i$. We define resource efficiency for this model as

$$E = \frac{\sum_{i=1}^{n} (\hat{\pi}_i - e_i)}{\sum_{i=1}^{n} c_i},$$

that is, sum of obtained payoffs minus full defection payoffs, divided by the sum of all contributions. Eq. (4.11) remains valid.

Notation 3. *In the endowment model, a game is specified by the players' productivities* $\mathbf{r}$*, their endowments* $\mathbf{e}$*, and the discount factor* δ*. We denote the respective game as* $\Gamma_{\mathrm{E}}(\mathbf{r}, \mathbf{e}, \delta)$*.*

We can now translate the results of the base model to the endowment model. Of Corollary 12, we obtain a weaker version:

Corollary 14. *Let* $\Gamma_{\mathrm{E}}(\mathbf{r}, \mathbf{e}, \delta)$ *allow for non-defection. Then there exists a* $\lambda > 0$ *such that equal contributions of* $\hat{\mathbf{c}} = \lambda \mathbf{1}$ *are feasible and sustainable.*

In other words, if any kind of non-defection is possible, then equal total contributions are also possible. However, now these contributions need to be sufficiently small so that in each round, the players do not exceed their endowment limits. Similar to the base model, the statement of Corollary 9 does no longer hold when we require contributions to be constant [HHCN19]. That is, there are cases when equal total contributions are possible, but only with time-dependent contribution sequences.

Trivially from Corollary 10, whether or not non-defection is possible in $\Gamma_{\mathrm{E}}(\mathbf{r}, \mathbf{e}, \delta)$ does not depend on $\mathbf{e}$; rather, it is possible if and only if it is possible in $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$. If that is the case, we may simply say that the pair $(\mathbf{r}, \delta)$ allows for non-defection, without referring to a particular one of the models.

Since resource efficiency is invariant under linear scalings of the contributions, all sustainable resource efficiencies can be achieved with arbitrarily small contributions. The set of sustainable resource efficiencies is thus unaffected by endowment constraints. This insight trivially allows us to extend the results about resource efficiency from the base model to the endowment model, as formally stated in the following proposition.

Proposition 4. *Take any game* $\mathbf{r}$*,* $\mathbf{e}$*,* δ*.*

1. *The set of sustainable resource efficiencies in* $\Gamma_{\mathrm{E}}(\mathbf{r}, \mathbf{e}, \delta)$ *is the same as in* $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$*.*
2. *The set of resource efficiencies sustainable with constant contributions in* $\Gamma_{\mathrm{E}}(\mathbf{r}, \mathbf{e}, \delta)$ *is the same as in* $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$*.*

It follows that Theorem 8 equally applies to the endowment model. In particular, for any game $\Gamma_{\mathrm{E}}(\mathbf{r}, \mathbf{e}, \delta)$, time-dependent contributions can enable higher resource efficiencies than constant contributions.

In addition to resource efficiency, in the endowment model it is also meaningful to study the player's welfare. However, with respect to welfare, the following (existing) result shows that time-dependent contributions provide no advantage.

Proposition 5 ([HSH+24], Supplementary Information Lemma 2)**.** *In any game* $\Gamma_{\mathrm{E}}(\mathbf{r}, \mathbf{e}, \delta)$*, sustainable welfare attains a maximum, and that maximum is attained with constant contributions.*

Example. To illustrate how the endowment model relates to the base model, we consider a game with the same parameters as in the previous example, that is, with $\mathbf{r} = (1.5, 1.1)^{\intercal}$ and $\delta = 0.9$. As the endowment distribution, we choose $\mathbf{e} = (0.2, 0.8)^{\intercal}$. So we have the game $\Gamma_{\mathrm{E}}((1.5, 1.1)^{\intercal}, (0.2, 0.8)^{\intercal}, 0.9)$.

In the previous example, we saw that the highest attainable resource efficiency in $\Gamma_{\mathrm{B}}((1.5, 1.1)^{\intercal}, 0.9)$ is $E_{\mathrm{sup}} = 0.375$ and that it is only attainable with a time-dependent contribution sequence. By Proposition 4, these statements carry over identically to the endowment model independently of our choice of $\mathbf{e}$. The sequence we constructed earlier to obtain resource efficiency 0.375 was given by $c_1(0) = c_1(1) = c_1(2) = \ldots = 10$ and $c_2(0) = 0$ and $c_2(1) = c_2(1) = \ldots = 500/99$. Of course, contributions of that size would by far exceed both players' endowments. But we can scale the sequence down by a factor of 50 to get $c_1(0) = c_1(1) = c_1(2) = \ldots = 0.2$ and $c_2(0) = 0$ and $c_2(1) = c_2(1) = \ldots = 10/99 \approx 0.101$, which is feasible with respect to the endowments, and sustainable in equilibrium. This sequence has the same resource efficiency of 0.375.

However, while player 1 contributes their entire endowment of 0.2 in every round, the largest part of player 2's endowment is unproductive in this contribution sequence. This is reflected in the comparatively low welfare. In round 0, payoffs are $\pi_1(0) = 0.15 + (0.2 - 0.2) = 0.15$ and $\pi_2(0) = 0.15 + (0.8 - 0) = 0.95$. In the subsequent rounds, payoffs are $\pi_1(t) \approx 0.205 + (0.2 - 0.2) \approx 0.205$ and $\pi_2(t) \approx 0.205 + (0.8 - 0.101) \approx 0.904$. This gives an overall welfare of $W \approx 1.108$ (remember that here, the welfare of full defection is $W = 1$). Compare that to constant contributions of $\hat{c}_1 = 0.2$ and $\hat{c}_2 = 0.3$, which are sustainable and achieve an optimal welfare of $W = 1.13$. Proposition 5 states that the welfare optimum can always be attained with constant contributions.

There is an intuitive general reason why time-dependent contributions have no positive welfare effects: In the endowment model, the more a contribution sequence varies, the more resources are non-productively withheld in some of the rounds. In particular, only constant contributions can achieve full cooperation, which by definition means $c_i(t) = e_i$ for all t and i. (With the right endowment distribution, a maximal welfare of $W = E^{\mathrm{c}}_{\mathrm{sup}} + 1$ can be realised that way; in the above example that is $W \approx 1.366$.) This observation suggests that the endowment model is inherently geared towards constant contributions. Instead, we would like to consider a setup in which overall contributions are constrained, yet players may freely choose how to allocate their contributions over time. To do this, we introduce another model variant.

4.3 Model variation II: A model with savings

The saving model builds on the earlier endowment model. Again, each player i obtains a fixed endowment e_i every round. However, now players have three options for how to spend their endowment, rather than two. They can contribute to the public good, consume parts or all of the endowment privately, or they can make a deposit into a savings account. Savings pay interest at the rate of $(\delta^{-1} - 1)$ per round – which exactly corresponds to the time value of money at the discount factor δ. These savings can then be spent in future rounds, either to contribute to the public good or for private consumption.

More specifically, each round proceeds as follows. In the beginning of each round t, players receive an endowment e_i. In addition, they have access to an amount of $s_i(t)$ on their savings account (in the very first round, savings are set to zero). Players then decide which amount $p_i(t)$ to consume privately, which amount $c_i(t)$ to contribute to the public good, and which amount $d_i(t)$ to deposit into the savings account. These variables need to satisfy the budget constraint $e_i + s_i(t) = p_i(t) + c_i(t) + d_i(t)$. Contributions to the public good and private consumption directly enter the player's payoff function,

$$\pi_i(t) = \frac{1}{n}\mathbf{r}^{\intercal}\mathbf{c}(t) + p_i(t).$$

Deposits, on the other hand, determine a player's savings in the beginning of the next round, $s_i(t+1) = \delta^{-1} d_i(t)$. At time $t+1$, the process is repeated; individuals again have to decide how much to consume, to contribute, and to save, see Fig. 4.5. Total payoffs (across all rounds), welfare, and resource efficiency are then defined as in the previous two models.

Notation 4. *The saving model uses the same parameters as the endowment model: productivities* $\mathbf{r}$*, endowments* $\mathbf{e}$*, and the discount factor* δ*. We denote the game as* $\Gamma_{\mathrm{S}}(\mathbf{r}, \mathbf{e}, \delta)$*.*

We make the following observations about the saving model: First, in our implementation of this model, savings are payoff-neutral: The interest earned over one round is exactly offset by a player's discounting of future rewards. Second, without the opportunity for

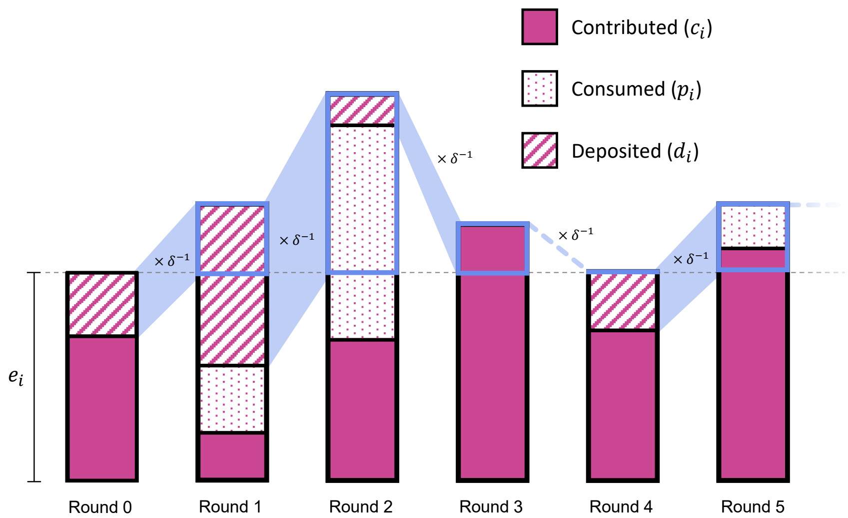

Figure 4.5: The saving model. Here, we depict the first six rounds of a player's gameplay. In each round, the player not only has their endowment e_i available, but also whatever resources they deposited in the previous round, plus interest at the rate $\delta^{-1}-1$. They decide which part of that they want to contribute, which part to consume, and which part to deposit to the savings account.

saving, this model recovers the endowment model as discussed in the previous section and studied in [HHCN19]. Third, the saving model is equivalent to saying that endowments only apply as a constraint to the cumulative contributions. That is, each player i is required to play such that $\sum_{\tau=0}^{t} \delta^\tau c_i(t) \leq \sum_{\tau=0}^{t} \delta^\tau e_i$ for all t, but without the stronger requirement that $c_i(t) \leq e_i$ for all t. Finally, since resource efficiency is equal to surplus welfare divided by total contributions, maximising resource efficiency in the base model is equivalent to maximising welfare in the saving model when the endowments are also an optimisation variable.

To state our main results for this section, we first define the notion of a welfare supremum with and without savings. For given parameters $\mathbf{r}, \mathbf{e}, \delta$, the welfare supremum with savings $W^{\mathrm{s}}_{\sup}(\mathbf{e})$ is the supremum of welfare over all equilibria of the game $\Gamma_{\mathrm{S}}(\mathbf{r}, \mathbf{e}, \delta)$. It quantifies the maximum value that the group can derive from cooperation. Similarly, we define the welfare supremum without savings, $W_{\sup}(\mathbf{e})$, as the supremum of welfare over all equilibria that satisfy $d_i(t) = 0$ for all i and t. It quantifies the maximum value that the group can derive without ever saving any amount. Equivalently, it corresponds to the welfare supremum of the endowment model, $\Gamma_{\mathrm{E}}(\mathbf{r}, \mathbf{e}, \delta)$. We interpret the difference $W^{\mathrm{s}}_{\sup}(\mathbf{e}) - W_{\sup}(\mathbf{e})$ as the (positive or zero) advantage that savings can provide.

Theorem 9. *Let $(\mathbf{r}, \delta)$ allow for non-defection. Take any endowment distribution $\mathbf{e}$. Then savings provide no advantage (i.e. $W^{\mathrm{s}}_{\sup}(\mathbf{e}) = W_{\sup}(\mathbf{e})$), if and only if*

$$\mathbf{e} \leq \delta D \mathbf{e}. \tag{4.18}$$

Without savings, $\hat{\mathbf{c}} = \mathbf{e}$ requires constant contributions. So by Corollary 11, Eq. (4.18) is equivalent to $\hat{\mathbf{c}} = \mathbf{e}$ being sustainable without savings. Therefore, Theorem 9 states that exactly one of the following is the case: Either full contributions are sustainable without savings, or savings provide an advantage for welfare. In particular, for any $\mathbf{r}$ and $\mathbf{e}$, when δ is sufficiently low, savings provide an advantage. Alternatively, for any $\mathbf{r}$ and δ, when $\mathbf{e}$ is sufficiently unequal, savings provide an advantage ([HHCN19], Supplementary Information Proposition 3).

Savings also provide an advantage from the perspective of a social planner who chooses an endowment distribution with the aim of maximising welfare. To assess this, we consider the suprema of $W^{\mathrm{s}}_{\sup}(\mathbf{e})$ and $W_{\sup}(\mathbf{e})$ over all possible endowment distributions $\mathbf{e}$. From Theorem 8, we can derive the following result:

Theorem 10. *Let $(\mathbf{r}, \delta)$ allow for non-defection. Let $r_{\max} = \max_i r_i$, and let m be the number of players with productivity $r_{\max}$. Then $\sup_{\mathbf{e}} W^{\mathrm{s}}_{\sup}(\mathbf{e}) > \sup_{\mathbf{e}} W_{\sup}(\mathbf{e})$ if and only if*

$$\Big(1 + \delta(m-1)\Big) \cdot r_{\max} < n. \tag{4.19}$$

To understand the theorem intuitively, consider again the case that there is only a single player with maximum productivity $r_{\max}$, i.e. that $m = 1$. Then Theorem 10 states that strictly better welfare is possible when players are permitted to save part or all of their endowment for later rounds, compared to when they are not. This is because with saving, they can play a more resource-efficient time-dependent contribution sequence and still productively use all of their endowment, whereas without, they are restricted to make constant contributions in order to maximise welfare. Conversely, as another special case of Theorem 10, we conclude that savings never provide a welfare advantage if all players have the same productivity.

Example. We revisit the same example as before, now in the saving model: $\Gamma_{\mathrm{S}}((1.5, 1.1)^\intercal, (0.2, 0.8)^\intercal, 0.9)$. In the endowment model, we had to scale the optimally resource-efficient contribution sequence to $c_1(0) = 0.2$ so that players do not exceed their endowment limits. With savings, it is enough that at no point in time their cumulative contributions exceed the endowment limit. A simple example of a superior contribution sequence is as follows. In round 0, player 1 contributes and consumes nothing ($c_1(0) = p_1(0) = 0$) and deposits everything ($d_1(0) = e_1 = 0.2$). In round 1, with the interest received, player 1 has savings of approximately 0.222 and again receives an endowment of 0.2, which makes for a total available amount of 0.422. Of this, player 1 contributes $c_1(1) \approx 0.222$ and again deposits $d_1(1) = 0.2$. In round 2, savings with interest again make up 0.222, and player 1 continues with contributing 0.222 and depositing 0.2 in every subsequent round. Player 2, on the other hand, from the beginning simply contributes $c_2(t) \approx 0.333$ in every round and privately consumes the rest, $p_2(t) \approx 0.467$, without depositing anything. The welfare of this contribution sequence is $W \approx 1.164$, which is more than the optimal value without saving, $W = 1.13$. Theorem 9 predicts that saving provides an advantage like this as long as full contributions are not sustainable, which is the case here.

The optimal welfare over all endowment distributions, $\sup_{\mathbf{e}} W^{\mathrm{s}}_{\sup}(\mathbf{e})$, requires the endowment distribution $\mathbf{e} = ({}^{11}\!/_{16}, {}^{5}\!/_{16})^\intercal$. This is exactly the endowment distribution at which the maximally resource efficient contribution sequence can be played in such a way that all endowments are eventually contributed: Player 1 contributes $c_1(t) = 11/16$ in every round. Player 2 deposits everything in round 0 ($d_2(0) = e_2 = 5/16$). Thereafter, player 2 contributes $50/(16 \cdot 9)$ and deposits $5/16$ in every round. (The sequence of contributions is identical to that of the earlier example in Section 4.1.4, up to rescaling by a factor of $160/11$, and thus also maximally resource efficient.) In this sequence, all resources are used productively and none are consumed privately, which means that the optimal resource efficiency also translates to optimal welfare. Indeed, the welfare is $W = \sup_{\mathbf{e}} W^{\mathrm{s}}_{\sup}(\mathbf{e}) = 1 + E_{\sup} = 1.375$.

The fact that this is greater than the optimum without saving over all endowment distributions, which by Proposition 5 is $\sup_{\mathbf{e}} W_{\text{sup}}(\mathbf{e}) = 1 + E^{\text{c}}_{\text{sup}} \approx 1.366$, is predicted by Theorem 10.

4.4 Discussion

The repeated public good game is one of the major models in (evolutionary) game theory to understand cooperation in groups. This literature describes how individuals can use conditionally cooperative strategies to sustain outcomes that are infeasible in one-shot encounters. Yet when describing the possible equilibrium outcomes, many previous studies implicitly restrict their analysis to the case that players make the same constant contribution each round [e.g. HHCN19, HSH+24]. Instead, here we study the effect of time-dependent contributions. Contrary to many other models of reciprocity, we allow players to select their actions from a continuum between full defection and full cooperation. We explore to which extent individuals can obtain better outcomes (e.g., a better resource efficiency or welfare) when they are able to vary their contributions along the equilibrium path.

From the outset, it is not clear whether time-dependent contributions provide any substantial advantage at all. After all, suppose players could achieve a superior outcome with contributions $\big(\mathbf{c}(t)\big)_t$ that vary in time. Then players might achieve just the same outcome by instead making a constant contribution $\hat{\mathbf{c}}$ each round, where $\hat{\mathbf{c}}$ is the appropriate (time-discounted) average contribution per round, $\hat{\mathbf{c}} = (1-\delta)\sum_t \delta^t c(t)$. With respect to their payoff implications, the two sequences $\big(\mathbf{c}(t)\big)_t$ and $(\hat{\mathbf{c}})_t$ are identical. After all, by Eq. (4.3), payoffs only depend on the players' total contributions across all rounds. As a result, the two sequences generate the same resource efficiency and welfare. However, as we show in this article, the two sequences may differ in their sustainability. There are instances in which the time-dependent sequence $\big(\mathbf{c}(t)\big)_t$ can be realised by a subgame perfect equilibrium, whereas the constant sequence $(\hat{\mathbf{c}})_t$ cannot.

To make this point, we study three different models: a base model, a model with endowment constraints, and a model with savings. In the base model, players are allowed to make arbitrary contributions each round (the only requirement is that the sequence of contributions does not diverge). This setup imposes minimal constraints on the players' behaviour, and it is convenient to work with mathematically. In contrast, the other two models are perhaps more realistic (and hence they have been studied more frequently). For example, the endowment model corresponds to the classical setup that is also frequently used in experiments [e.g. CMMM99, BC06, HRS16, KMS17]. Here, contributions are constrained by the endowments that the players receive each round. The saving model is similar, but in addition it allows players to (payoff-neutrally) transfer some of their endowments to future rounds. Interestingly, many of our results for these last two model variants are directly related to our findings in the base model. As an example, with Theorem 10, we characterise under which circumstances savings provide a welfare advantage in the saving model. The respective result is directly related to whether or not time-dependent contributions provide an advantage in the base model, Theorem 8. These similarities between those theorems highlight how several findings in the more abstract base model carry over to more applied settings.

Interestingly, the respective theorems also suggest that for our results, some asymmetry among players is crucial. Specifically, when players are identical with respect to their productivities, Theorem 8 shows that time-dependent contributions do not grant any advantage. Any resource efficiency that can be sustained with time-dependent contributions can already be sustained with constant contributions. But once players differ in their productivities, it becomes fairly easy for time-dependent contributions to be superior. In fact, such an advantage is guaranteed when the group contains a single player whose productivity exceeds everyone else's.

Overall, our findings highlight the impact of variable contributions on resource efficiency and, more generally, on the sustainability of cooperation. They suggest that by focussing solely on constant contributions, we may overlook important equilibria that can arise in dynamic settings.

4.5 Proofs

4.5.1 Proof of Theorem 7

Proof of Theorem 7. Clearly $\hat{\mathbf{c}} = \mathbf{0}$ is sustainable, so it is sufficient to show the statement for $\hat{\mathbf{c}} \neq \mathbf{0}$.

First we will show that any sustainable $\hat{\mathbf{c}} \neq \mathbf{0}$ satisfies $\hat{\mathbf{c}} \leq_1 D\hat{\mathbf{c}}$. By Theorem 5, we have for any sustainable $\bar{\mathbf{c}}(t)$ that

$$\delta\bar{\mathbf{c}}(1) \leq \bar{\mathbf{c}}(0) \leq \delta D\bar{\mathbf{c}}(1). \tag{4.20}$$

Left-multiplying with the non-negative matrix D on both sides of the first inequality, we obtain

$$\delta D\bar{\mathbf{c}}(1) \leq D\bar{\mathbf{c}}(0)$$

with equality in the ith component exactly if $\delta c_j(1) = c_j(0)$ for all $j \neq i$.

Together with the second inequality of (4.20), this gives

$$\bar{\mathbf{c}}(0) \leq D\bar{\mathbf{c}}(0),$$

or equivalently

$$\hat{\mathbf{c}} \leq D\hat{\mathbf{c}}, \tag{4.21}$$

where equality in the ith component requires $\delta c_j(1) = c_j(0)$ for all $j \neq i$.

If (4.21) has equality in at least two components, then $\delta c_j(1) = c_j(0)$ for all j, so $\delta\mathbf{c}(1) = \mathbf{c}(0)$. Analogously, equality in two components requires $\delta\mathbf{c}(2) = \mathbf{c}(1)$, etc., so the sequence $(\mathbf{c}(t))_t$ diverges and is not a valid contribution sequence. That is a contradiction, so we can have $\hat{c}_i = (D\hat{\mathbf{c}})_i$ for at most one i.

Now we will show that if some $\hat{\mathbf{c}} \neq \mathbf{0}$ satisfies $\hat{\mathbf{c}} \leq_1 D\hat{\mathbf{c}}$, then $\hat{\mathbf{c}}$ is sustainable with a continuation contribution sequence $(\bar{\mathbf{c}}(t))_t$ that satisfies $\hat{\mathbf{c}} \leq \bar{\mathbf{c}}(t)$ for all t. Assume first that the stronger condition $\mathbf{0} < \hat{\mathbf{c}} < D\hat{\mathbf{c}}$ holds. Take $\varepsilon > 0$ such that $1 + \varepsilon < \delta^{-1}$ and $(1+\varepsilon)\mathbf{x} \leq D\mathbf{x}$. Let $\mathbf{v}$ be the Perron eigenvector of D, scaled so that $\mathbf{v} \leq \mathbf{x}$. Let finally

$$T = \left\lceil \frac{\delta_{\min} \max_i \frac{x_i}{v_i} - \min_i \frac{x_i}{v_i}}{\varepsilon(1-\delta_{\min})} \right\rceil$$

or $T = 0$, whichever is larger.

Let $\mathbf{x} = \hat{\mathbf{c}}$ for a given $\hat{\mathbf{c}}$ with $\mathbf{0} < \hat{\mathbf{c}} < D\hat{\mathbf{c}}$. We define a continuation contribution sequence $(\bar{\mathbf{c}}(t))_t$ by

$$\bar{\mathbf{c}}(t) = ((1+\varepsilon)\delta)^{-t}(\mathbf{x} + \varepsilon t\mathbf{v})$$

for all $0 \leq t \leq T$ and

$$\bar{\mathbf{c}}(t) = (1+\varepsilon)^{-T}\delta^{-(T+1)}\left(\min_i \frac{x_i}{v_i} + \varepsilon T\right)\mathbf{v}$$

for all $t > T$. If we show that it obeys (4.5) for all t, we know it is sustainable. Since $\bar{\mathbf{c}}(0) = \mathbf{x}$, that is enough to prove the first statement of the present theorem. We also see that $(\bar{\mathbf{c}}(t))_t$ is non-decreasing. Therefore, $\hat{\mathbf{c}} = \bar{\mathbf{c}}(0) \leq \bar{\mathbf{c}}(t)$ for all t, which is the second statement of the theorem.

For $0 \leq t < T$, the first inequality, $\delta\bar{\mathbf{c}}(t+1) \leq \bar{\mathbf{c}}(t)$, follows from

$$\mathbf{v} \leq \mathbf{x}$$

as follows. First, we multiply with ε and add $\varepsilon^2 t\mathbf{v}$, which is non-negative, on the right-hand side:

$$\varepsilon\mathbf{v} \leq \varepsilon\mathbf{x} + \varepsilon^2 t\mathbf{v}$$

We add $\mathbf{x} + \varepsilon t\mathbf{v}$ on both sides and factor out:

$$\mathbf{x} + \varepsilon(t+1)\mathbf{v} \leq (1+\varepsilon)(\mathbf{x} + \varepsilon t\mathbf{v})$$

We multiply with $\delta((1+\varepsilon)\delta)^{-(t+1)}$ on both sides:

$$\delta((1+\varepsilon)\delta)^{-(t+1)}(\mathbf{x} + \varepsilon(t+1)\mathbf{v}) \leq ((1+\varepsilon)\delta)^{-t}(\mathbf{x} + \varepsilon t\mathbf{v})$$

This is equivalent to

$$\delta\bar{\mathbf{c}}(t+1) \leq \bar{\mathbf{c}}(t)$$

by the definition of $\bar{\mathbf{c}}(t)$.

The second inequality, $\bar{\mathbf{c}}(t) \leq \delta D\bar{\mathbf{c}}(t+1)$, follows from

$$\mathbf{v} = \delta_{\min} D\mathbf{v}.$$

We replace $\delta_{\min}$ with δ, which is larger or equal, on the right-hand side:

$$\mathbf{v} \leq \delta D\mathbf{v}$$

By construction, $1 - \varepsilon < \delta^{-1}$, so we can write:

$$(1+\varepsilon)\mathbf{v} < D\mathbf{v}$$

Multiply by εt on the left-hand side and by $\varepsilon(t+1)$ on the right-hand side:

$$(1+\varepsilon)\varepsilon t\mathbf{v} < \varepsilon(t+1)D\mathbf{v}$$

Sum with the inequality $(1+\varepsilon)\mathbf{x} \leq D\mathbf{x}$, which also holds by construction of ε:

$$(1+\varepsilon)(\mathbf{x}+\varepsilon t\mathbf{v}) < D(\mathbf{x}+\varepsilon(t+1)\mathbf{v})$$

Multiply by $\delta((1+\varepsilon)\delta)^{-(t+1)}$ on both sides:

$$((1+\varepsilon)\delta)^{-t}(\mathbf{x}+\varepsilon t\mathbf{v}) < \delta((1+\varepsilon)\delta)^{-(t+1)}D(\mathbf{x}+\varepsilon(t+1)\mathbf{v})$$

This is equivalent to

$$\bar{\mathbf{c}}(t) \leq \delta D\bar{\mathbf{c}}(t+1).$$

Next, we consider $t = T$: The first inequality follows from

$$\mathbf{x}+\varepsilon T\mathbf{v} \leq (1+\varepsilon)(\mathbf{x}+\varepsilon T\mathbf{v}).$$

First, we replace $\mathbf{x}$ by $\min_i \frac{x_i}{v_i}\mathbf{v}$, which is at most as large in each component:

$$\min_i \frac{x_i}{v_i}\mathbf{v}+\varepsilon T\mathbf{v} \leq (1+\varepsilon)(\mathbf{x}+\varepsilon T\mathbf{v})$$

Then, we multiply by $\delta((1+\varepsilon)\delta)^{-(T+1)}$ on both sides:

$$\delta((1+\varepsilon)\delta)^{-(T+1)}\left(\min_i \frac{x_i}{v_i}+\varepsilon T\right)\mathbf{v} \leq ((1+\varepsilon)\delta)^{-T}(\mathbf{x}+\varepsilon T\mathbf{v})$$

This is by definition equivalent to

$$\delta\bar{\mathbf{c}}(T+1) \leq \bar{\mathbf{c}}(T).$$

The second inequality follows from

$$\left\lceil \frac{\delta_{\min}\max_i \frac{x_i}{v_i} - \min_i \frac{x_i}{v_i}}{\varepsilon(1-\delta_{\min})} \right\rceil \leq T,$$

which is by definition of T. We remove the ceiling function and multiply by $\varepsilon(1-\delta_{\min})$ on both sides:

$$\delta_{\min}\max_i \frac{x_i}{v_i} - \min_i \frac{x_i}{v_i} \leq (1-\delta_{\min})\varepsilon T$$

We move $\min_i \frac{x_i}{v_i}$ to the right-hand side, $\delta_{\min}\varepsilon T$ to the left, and then multiply by $\delta_{\min}^{-1}\mathbf{v}$:

$$\max_i \frac{x_i}{v_i}\mathbf{v}+\varepsilon T\mathbf{v} \leq \left(\min_i \frac{x_i}{v_i}+\varepsilon T\right)\delta_{\min}^{-1}\mathbf{v}$$

Now, we can replace $\max_i \frac{x_i}{v_i}\mathbf{v}$ by $\mathbf{x}$, which is at most as large in each component, and $\delta_{\min}^{-1}\mathbf{v}$ by $D\mathbf{v}$, which is equal:

$$\mathbf{x}+\varepsilon T\mathbf{v} \leq \left(\min_i \frac{x_i}{v_i}+\varepsilon T\right)D\mathbf{v}$$

We multiply by $((1+\varepsilon)\delta)^{-T}$ on both sides:

$$((1+\varepsilon)\delta)^{-T}(\mathbf{x}+\varepsilon T\mathbf{v}) \leq \delta D\left((1+\varepsilon)^{-T}\delta^{-(T+1)}\left(\min_i \frac{x_i}{v_i}+\varepsilon T\right)\mathbf{v}\right)$$

This is by definition equivalent to

$$\bar{\mathbf{c}}(t) \leq \delta D\bar{\mathbf{c}}(t+1).$$

Finally, for $t > T$, we have $\bar{\mathbf{c}}(t) = \bar{\mathbf{c}}(t+1)$. The first inequality is trivially true, and the second inequality is true since $\bar{\mathbf{c}}(t)$ is a multiple of $\mathbf{v}$. So we have shown that $\hat{\mathbf{c}}$ is sustainable if $\mathbf{0} < \hat{\mathbf{c}} < D\hat{\mathbf{c}}$.

Now we will reduce the more general case of $\hat{\mathbf{c}} \neq \mathbf{0}$ and $\hat{\mathbf{c}} \leq_1 D\hat{\mathbf{c}}$ to this in two sequential steps.

First, let some $\hat{\mathbf{c}} \neq \mathbf{0}$ have $\mathbf{0} < \hat{\mathbf{c}} \leq_1 D\hat{\mathbf{c}}$ with equality of the weak inequality in exactly one component i. Let $\mathbf{x}(\varepsilon) = \hat{\mathbf{c}} - \varepsilon\mathbf{u}_i$, where $\mathbf{u}_i$ is the ith standard unit vector. Since $\mathbf{x}(\varepsilon) \to \hat{\mathbf{c}}$ as $\varepsilon \to \infty$ and $\hat{c}_j < (D\hat{\mathbf{c}})_j$ for all $j \neq i$, we can choose $\varepsilon > 0$ sufficiently small such that $\hat{c}_j = x_j(\varepsilon) < (D\mathbf{x}(\varepsilon))_j$ for all $j \neq i$ as well, and $0 < x_i(\varepsilon)$. We then have $x_i(\varepsilon) < \hat{c}_i$, but $(D\mathbf{x}(\varepsilon))_i = (D\hat{\mathbf{c}})_i$, so also $x_i(\varepsilon) < (D\mathbf{x}(\varepsilon))_i$. Hence $\mathbf{x}(\varepsilon)$ satisfies $\mathbf{0} < \mathbf{x}(\varepsilon) < D\mathbf{x}(\varepsilon)$ and $\hat{\mathbf{c}} \leq D\mathbf{x}(\varepsilon)$. By the above result, we can choose a continuation contribution sequence $(\mathbf{c}^*(t))_t$ such that $\hat{\mathbf{c}}^* = \mathbf{x}(\varepsilon)$. Now consider the sequence $(\hat{\mathbf{c}}, \delta^{-1}\bar{\mathbf{c}}^*(0), \delta^{-1}\bar{\mathbf{c}}^*(1), \delta^{-1}\bar{\mathbf{c}}^*(2), \dots)$. By definition, $\bar{\mathbf{c}}^*(0) = \mathbf{x}(\varepsilon)$, and we have

$$\delta(\delta^{-1}\mathbf{x}(\varepsilon)) \leq \hat{\mathbf{c}} \leq \delta D(\delta^{-1}\mathbf{x}(\varepsilon)),$$

so by Theorem 5, the sequence we constructed is a sustainable continuation contribution sequence. It begins with $\hat{\mathbf{c}}$, so the statement of the Lemma holds for $\hat{\mathbf{c}}$.

Now, the only case that remains is $\hat{\mathbf{c}} \leq_1 D\hat{\mathbf{c}}$ and $\mathbf{0} \not< \hat{\mathbf{c}} \neq \mathbf{0}$. In the definition of the public goods game, we imposed the conditions $1 < r_i$ and $r_i < n$ for all i. We used $r_i < n$ in the proofs about sustainability (it is necessary for D being well defined and positive), but $1 < r_i$ was only used to prove statements about maximal welfare. So we can consider games that only satisfy $\mathbf{r} > \mathbf{0}$ instead of $\mathbf{r} > \mathbf{1}$, and the same results about sustainabililty will apply, including the ones from this proof.

Let $\hat{\mathbf{c}} \neq \mathbf{0}$ be any total contribution vector satisfying $\hat{\mathbf{c}} \leq_1 D\hat{\mathbf{c}}$. Let n' be the number of non-zero components of $\hat{\mathbf{c}}$. W.l.o.g. let these be the first n' components of $\hat{\mathbf{c}}$. Necessarily $n' > 1$. Let $\hat{\mathbf{c}}'$ be the first n' components of $\hat{\mathbf{c}}$, and let $\mathbf{r}'$ be the n'-vector defined by $r'_i = n'n^{-1}r_i$ for all $1 \leq i \leq n'$. Consider the game $\Gamma_{\mathrm{B}}(\mathbf{r}', \delta)$ and its zero-diagonal productivity matrix D'. We have $D'_{ij} = D_{ij}$ for all $1 \leq i, j \leq n'$. Since $\hat{\mathbf{c}}$ satisfies $\hat{\mathbf{c}} \leq_1 D\hat{\mathbf{c}}$, consequently $\hat{\mathbf{c}}'$ also satisfies $\hat{\mathbf{c}}' \leq_1 D'\hat{\mathbf{c}}'$. But $\hat{\mathbf{c}}'$ additionally satisfies $\mathbf{0} < \hat{\mathbf{c}}'$. So by the case handled above, there is a sustainable continuation contribution sequence $(\mathbf{c}'(t))_t$ starting with $\hat{\mathbf{c}}'$. Let $(\mathbf{c}(t))_t$ be the n-player sequence with $c'_i(t) = c_i(t)$ for all $i \leq n'$ and all t, and $c'_i(t) = 0$ for all $i > n'$ and all t. Using Equation 4.5, we can see that sustainability of $(\mathbf{c}(t))_t$ follows trivially from sustainability of $(\mathbf{c}'(t))_t$. So we have found a continuation contribution sequence starting with $\hat{\mathbf{c}}$.

We have thus shown in the most general case that if some $\hat{\mathbf{c}} \neq \mathbf{0}$ satisfies $\hat{\mathbf{c}} \leq_1 D\hat{\mathbf{c}}$, then $\hat{\mathbf{c}}$ is sustainable. Together with the converse, which we already showed, this completes the proof.

□

4.5.2 Proof of Theorem 8

Proof of Theorem 8. Theorem 7 says that a total contribution vector of $\hat{\mathbf{c}}$ is sustainable exactly if

$$\hat{\mathbf{c}} \leq_1 D\hat{\mathbf{c}}.$$

Corollary 11 says that constant contributions of $\hat{\mathbf{c}}$ are sustainable exactly if

$$\hat{\mathbf{c}} \leq \delta D\hat{\mathbf{c}}.$$

Let $\mathcal{F}$ be the region defined by $\hat{\mathbf{c}} \leq_1 D\hat{\mathbf{c}}$. Let

$$E_{\text{sup}} := \sup_{\hat{\mathbf{c}} \in \mathcal{F} \setminus \{\mathbf{0}\}} E(\hat{\mathbf{c}}) = \sup_{\hat{\mathbf{c}} \in \overline{\mathcal{F}} \setminus \{\mathbf{0}\}} E(\hat{\mathbf{c}}),$$

where $\overline{\mathcal{F}}$ is the closure of $\mathcal{F}$, which is given by $\hat{\mathbf{c}} \leq D\hat{\mathbf{c}}$. Firstly, note that E_{sup} is well defined, since $E(\hat{\mathbf{c}})$ is bounded above by $\max_i r_i$. We observe that $E(\hat{\mathbf{c}})$ attains a maximum on $\overline{\mathcal{F}} \setminus \{\mathbf{0}\}$: We have

$$\{E(\hat{\mathbf{c}}) \mid \hat{\mathbf{c}} \in \overline{\mathcal{F}} \setminus \{\mathbf{0}\}\} = \{E(\hat{\mathbf{c}}/\|\hat{\mathbf{c}}\|) \mid \hat{\mathbf{c}} \in \overline{\mathcal{F}} \setminus \{\mathbf{0}\}\} = \{E(\hat{\mathbf{c}}) \mid \hat{\mathbf{c}} \in \overline{\mathcal{F}} \cap S^{n-1}\},$$

where $S^{n-1} = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = 1\}$. Since $\overline{\mathcal{F}}$ is closed, $\overline{\mathcal{F}} \cap S^{n-1}$ is a compact set. So we can write

$$E_{\text{sup}} = \max_{\hat{\mathbf{c}} \in \overline{\mathcal{F}} \setminus \{\mathbf{0}\}} E(\hat{\mathbf{c}}).$$

Now we are ready to prove the statement of the theorem. But instead of showing that a sustainable time-dependent contribution sequence that is more resource-efficient than all constant contribution sequences exists exactly if

$$(\delta(m-1)+1)r_{\max} < n,$$

as stated in the theorem, we will show the following equivalent statement: A constant contribution sequence with resource efficiency E_{sup} exists exactly if

$$n \leq (\delta(m-1)+1)r_{\max}. \tag{4.22}$$

The following statements are equivalent for $\hat{\mathbf{c}} \in \mathbb{R}^n$:

$$\hat{\mathbf{c}} \in \overline{\mathcal{F}}$$
$$\hat{\mathbf{c}} \leq D\,\hat{\mathbf{c}}$$
$$\forall i \quad \hat{c}_i \leq \sum_{j \neq i} \frac{r_j}{n - r_i} \hat{c}_j \tag{4.23}$$
$$\forall i \quad n\hat{c}_i \leq \sum_j r_j \hat{c}_j. \tag{4.24}$$

Take any $\hat{\mathbf{c}} \in \overline{\mathcal{F}} \setminus \{\mathbf{0}\}$ such that $E(\hat{\mathbf{c}}) = E_{\text{sup}}$ that is sustainable with a constant contribution sequence, which we assume exists. This $\hat{\mathbf{c}}$ maximises $E(\hat{\mathbf{c}})$ over all $\hat{\mathbf{c}} \neq \mathbf{0}$ satisfying the above inequality (4.24). This implies the following statement ($*$): There are no i_1, i_2 such that $r_{i_1} > r_{i_2}$ and $n\hat{c}_{i_1} < \sum_j r_j \hat{c}_j$ and $\hat{c}_{i_2} > 0$. Otherwise, we could increase $\hat{c}_{i_1}$ and decrease $\hat{c}_{i_2}$ by equal amounts to increase $E(\hat{\mathbf{c}})$ while also staying within $\overline{\mathcal{F}} \setminus \{\mathbf{0}\}$.

Since $\hat{\mathbf{c}}$ is sustainable with a constant contribution sequence, we have $\hat{\mathbf{c}} \leq \delta D\hat{\mathbf{c}}$, which is equivalent to

$$\forall i \quad \hat{c}_i \leq \delta \sum_{j \neq i} \frac{r_j}{n - r_i} \hat{c}_j. \tag{4.25}$$

At least two $\hat{c}_i$ must be strictly positive, so $\sum_{j \neq i} \frac{r_j}{n-r_i} \hat{c}_j > 0$ for every i. Since $\delta < 1$, (4.25) implies

$$\forall i \quad \hat{c}_i < \sum_{j \neq i} \frac{r_j}{n - r_i} \hat{c}_j.$$

So the statement (∗) simplifies to: There are no i_1, i_2 such that $r_{i_1} > r_{i_2}$ and $\hat{c}_{i_2} > 0$. This means that for any i, if $\hat{c}_i > 0$, then $r_i = r_{\max}$.

From (4.25), we get

$$\forall i \quad (n - (1 - \delta) r_i)\hat{c}_i \leq \delta \sum_j r_j \hat{c}_j, \tag{4.26}$$

where

$$\sum_j r_j \hat{c}_j = r_{\max} \sum_j \hat{c}_j.$$

Exactly m components of $\hat{\mathbf{c}}$ are non-zero. So choose some i such that $\hat{c}_i > 0$ and

$$\hat{c}_i \geq \frac{1}{m} \sum_j \hat{c}_j.$$

Inserting into (4.26), we get

$$(n - (1 - \delta) r_{\max}) \frac{1}{m} \sum_j \hat{c}_j \leq \delta r_{\max} \sum_j \hat{c}_j.$$

The $\hat{c}_j$ sum to 1. We multiply with m on both sides and get

$$n - (1 - \delta) r_{\max} \leq \delta r_{\max} m.$$

By simple rearrangement, this is equivalent to (4.22). So (4.22) being false is a necessary condition for the existence of a constant contribution sequence with resource efficiency E_{sup}, which was our only assumption.

Now instead assume conversely that (4.22) holds. Let $\hat{c}_i = 1$ for all i such that $r_i = r_{\max}$, and let $\hat{c}_i = 0$ for all other i. We can check easily from the definitions that $\hat{\mathbf{c}}$ satisfies $\hat{\mathbf{c}} \leq \delta D\hat{\mathbf{c}}$, so it is sustainable with a constant contribution sequence. The resource efficiency of $\hat{\mathbf{c}}$ is $E(\hat{\mathbf{c}}) = r_{\max}$, so it is maximal. Therefore, (4.22) is also a sufficient condition, and hence an exact condition, for the existence of a constant contribution sequence with resource efficiency E_{sup}. □

4.5.3 Proof of Theorem 9

Proof of Theorem 9. We will show that if $\mathbf{e} \leq \delta D\mathbf{e}$, then $W^{\text{s}}_{\text{sup}}(\mathbf{e}) = W_{\text{sup}}(\mathbf{e})$, and if $\mathbf{e} \not\leq \delta D\mathbf{e}$, then $W^{\text{s}}_{\text{sup}}(\mathbf{e}) > W_{\text{sup}}(\mathbf{e})$.

Firstly, if $\mathbf{e} \leq \delta D\mathbf{e}$, then by Corollary 11, the constant contribution sequence $(\mathbf{c}(t))_t = (\mathbf{e})_t$ is sustainable. So the total contribution vector $\hat{\mathbf{c}} = \mathbf{e}$ is sustainable without saving. Since

$\mathbf{e}$ is an upper bound on the total contribution vector, $\mathbf{r}^\intercal\mathbf{e}$ is also an upper bound on welfare in general, and we have $W^{\mathrm{s}}_{\sup}(\mathbf{e}) = W_{\sup}(\mathbf{e}) = \mathbf{r}^\intercal\mathbf{e}$.

Now, assume that $\mathbf{e} \not\leq \delta D\mathbf{e}$. By Proposition 5, there is always a total contribution vector $\hat{\mathbf{c}}$ that satisfies $W(\hat{\mathbf{c}}) = W_{\sup}(\mathbf{e})$ and is sustainable with a constant contribution sequence in $\Gamma_{\mathrm{E}}(\mathbf{r}, \mathbf{e}, \delta)$, i.e. without saving. Take such a $\hat{\mathbf{c}}$. By Corollary 11, we have $\hat{\mathbf{c}} \leq \delta D\hat{\mathbf{c}}$. So $\hat{\mathbf{c}} \neq \mathbf{e}$, meaning there is some i such that $\hat{c}_i < e_i$. Fix such an i.

Define $\hat{\mathbf{c}}(\varepsilon) = \hat{\mathbf{c}} + \varepsilon(\mathbf{e} - \hat{\mathbf{c}})$ for $\varepsilon \geq 0$. Clearly $\hat{\mathbf{c}}(\varepsilon) \to \hat{\mathbf{c}}$ as $\varepsilon \to 0$. Since $\hat{\mathbf{c}} \leq \delta D\hat{\mathbf{c}}$, since $D\hat{\mathbf{c}}$ is a positive vector, and since $\delta < 1$, we have $\hat{\mathbf{c}} < D\hat{\mathbf{c}}$. So we can choose $\varepsilon > 0$ sufficiently small such that $\hat{\mathbf{c}}(\varepsilon) < D\hat{\mathbf{c}}(\varepsilon)$ as well. By Theorem 7, $\hat{\mathbf{c}}(\varepsilon)$ is thus a sustainable total contribution vector in $\Gamma_{\mathrm{S}}(\mathbf{r}, \mathbf{e}, \delta)$. But $\hat{\mathbf{c}}(\varepsilon) \geq \hat{\mathbf{c}}$ and $\hat{c}(\varepsilon)_i > \hat{c}_i$. So $W(\hat{\mathbf{c}}(\varepsilon)) > W(\hat{\mathbf{c}})$. Consequently, $W^{\mathrm{s}}_{\sup}(\mathbf{e}) > W_{\sup}(\mathbf{e})$. □

4.5.4 Proof of Theorem 10

Proposition 6. *Take any game $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$ that allows for non-defection. Then for any $\hat{\mathbf{c}} \in \mathbb{R}^n_{\geq 0}$ satisfying $\sum_{i=1}^n \hat{c}_i \leq 1$, the following are equivalent:*

1. *A total contribution vector of $\hat{\mathbf{c}}$ is sustainable in the game $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$.*
2. *There exists an endowment distribution $\mathbf{e}$ such that a total contribution vector of $\hat{\mathbf{c}}$ is sustainable in the game $\Gamma_{\mathrm{S}}(\mathbf{r}, \mathbf{e}, \delta)$.*

In a game $\Gamma_{\mathrm{S}}(\mathbf{r}, \mathbf{e}, \delta)$, a contribution sequence $(\mathbf{c}(t))_t$ is called sustainable if there is a play that results in contribution sequence $(\mathbf{c}(t))_t$.

It is easy to check that a contribution sequence $(\mathbf{c}(t))_t$ is sustainable in $\Gamma_{\mathrm{S}}(\mathbf{r}, \mathbf{e}, \delta)$ if and only if

$$\hat{\mathbf{c}} \leq \delta^t\mathbf{e} + \delta^{t+1}\bar{\mathbf{c}}(t+1) \tag{4.27}$$

for all $t \geq 0$. Inequality 4.27 simply states that every player i has enough resources in every round t in order to make a contribution of $c_i(t)$ as long as they never consume any of their available resources privately.

In the game $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$, the Grim strategy profile $G((\mathbf{c}(t))_t)$ for a contribution sequence $(\mathbf{c}(t))_t$ is the pure strategy profile $G((\mathbf{c}(t))_t) = (\sigma_i)_i$ defined as follows: In each round t and for each i the strategy σ_i contributes $c_i(t)$ if all players have so far also played according to $(\mathbf{c}(t))_t$, but otherwise contributes 0.

In the base model, a contribution sequence $(\mathbf{c}(t))_t$ is sustainable in a given game if and only if its associated Grim strategy profile $G((\mathbf{c}(t))_t)$ is a SPE of the game [HSH+24].

In the game $\Gamma_{\mathrm{B}}(\mathbf{r}, \mathbf{e}, \delta)$, the Grim strategy profile $G^{\mathrm{s}}((\mathbf{c}(t))_t)$ for a contribution sequence $(\mathbf{c}(t))_t$ is defined as follows.

First, we recursively construct a deposit sequence $(\mathbf{d}(t))_t$. For each t and each i, let $d_i(t)$ be the minimal d such that there is a play that results in contribution sequence $(\mathbf{c}(t))_t$ and an initial deposit sequence for player i of $d_i(0)$, …, $d_i(t-1)$, d. We can check that a minimum is indeed attained. Then there also exists a play that results in contribution sequence $(\mathbf{c}(t))_t$ and deposit sequence $(\mathbf{d}(t))_t$. This also uniquely determines the private consumption sequence $(\mathbf{p}(t))_t$.

Now the Grim strategy $G^{\mathrm{s}}((\mathbf{c}(t))_t)_i$ for player i plays as follows. In each round t, if all players have so far played according to $G^{\mathrm{s}}((\mathbf{c}(t))_t)$, then player i contributes $c_i(t)$, privately consumes $p_i(t)$, and deposits $d_i(t)$. Otherwise, player i contributes 0, deposits 0, and privately consumes the entire available amount.

We see that in both models, the Grim strategy profile $G((\mathbf{c}(t))_t)$ or $G^{\mathrm{s}}((\mathbf{c}(t))_t)$, respectively, produces the contribution sequence $(\mathbf{c}(t))_t$. Hence, in both models, there is a bijective correspondence between sustainable contribution sequences and Grim strategies.

Proof of Proposition 6. Take any $\hat{\mathbf{c}} \in \mathbb{R}^n_{\geq 0}$ satisfying $\sum_{i=1}^n \hat{c}_i \leq 1$.

Assume that statement 1 of Proposition 6 holds. By Theorem 7, choose a sustainable contribution sequence $(\mathbf{c}(t))_t$ such that $\hat{\mathbf{c}} = \bar{\mathbf{c}}(0)$ and $\hat{\mathbf{c}} \leq \bar{\mathbf{c}}(t)$ for all t. We know that $G((\mathbf{c}(t))_t)$ is a SPE of the game $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$.

Set $\mathbf{e} = \hat{\mathbf{c}}$. We want to show that in the game $\Gamma_{\mathrm{S}}(\mathbf{r}, \mathbf{e}, \delta)$, the Grim strategy profile $G^{\mathrm{s}}((\mathbf{c}(t))_t)$ is a SPE, which is sufficient to show that statement 2 holds. Let $(\mathbf{d}(t))_t$ be the deposit sequence produced by $G^{\mathrm{s}}((\mathbf{c}(t))_t)$. So consider a strategy profile where all players play according to $G^{\mathrm{s}}((\mathbf{c}(t))_t)$ with the execption of one mutant, i, who plays according to a general strategy σ. Let $(\mathbf{c}'(t))_t$ be the contribution sequence, $(\mathbf{p}'(t))_t$ the private consumption sequence, and $(\mathbf{d}'(t))_t$ the deposit sequence produced by this strategy profile. Let ${\pi^{\mathrm{s}}}'_i$ be the payoff of player i, and let π^{s}_i be the payoff of player i if player i were also playing according to $G^{\mathrm{s}}((\mathbf{c}(t))_t)$. We need to show that ${\pi^{\mathrm{s}}}'_i \leq \pi^{\mathrm{s}}_i$.

Now, in the game $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$, consider a strategy profile where all players play according to $G((\mathbf{c}(t))_t)$ with the execption of one mutant, i, who unconditionally plays the contribution sequence $(c'_i(t))_t$. Clearly, this will also result in the same contribution sequence $(\mathbf{c}'(t))_t$. Let π'_i be the payoff of player i in this strategy profile, and π_i the payoff of player i if player i were also playing according to $G((\mathbf{c}(t))_t)$.

We have $\sum_{t=0}^{\infty} \delta^t p'_i(t) \leq \sum_{t=0}^{\infty} \delta^t d'_i(t)$, so ${\pi^{\mathrm{s}}}'_i \leq \pi'_i$. By the minimality of the deposit sequence $(\mathbf{d}(t))_t$ according to the definition of Grim strategies in the saving model, we have $\sum_{t=0}^{\infty} \delta^t p_i(t) = \sum_{t=0}^{\infty} \delta^t d_i(t)$. So $\pi^{\mathrm{s}}_i = \pi_i$. Since $G((\mathbf{c}(t))_t)$ is a SPE of the game $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$, we have $\pi'_i \leq \pi_i$. We therefore have that ${\pi^{\mathrm{s}}}'_i \leq \pi^{\mathrm{s}}_i$.

Now instead assume that statement 2 holds. Fix any $\mathbf{e}$ and any SPE of the game $\Gamma_{\mathrm{S}}(\mathbf{r}, \mathbf{e}, \delta)$ which produces some contribution sequence $(\mathbf{c}(t))_t$, sequence of deposits $(\mathbf{d}(t))_t$ and the total contribution vector $\hat{\mathbf{c}}$. By the definition of an SPE, no player can at any time profit from switching to privately consuming the entire amount that they have at their disposal in each round for the rest of the game.

Take any player i and any time t. The continuation payoff of player i at time t is ordinarily

$$\begin{aligned}
\bar{\pi}_i(t) &= (1-\delta) \sum_{\tau=t}^{\infty} \delta^{\tau-t} \pi_i(\tau) \\
&= (1-\delta) \sum_{\tau=t}^{\infty} \delta^{\tau-t} (p_i(\tau) + n^{-1} \mathbf{r}^{\intercal} \mathbf{c}(\tau)) \\
&= (1-\delta) \sum_{\tau=t}^{\infty} \delta^{\tau-t} (e_i + s_i(\tau) - c_i(\tau) - d_i(\tau) + n^{-1} \mathbf{r}^{\intercal} \mathbf{c}(\tau)) \\
&= e_i + (1-\delta) s_i(t) - \bar{c}_i(t) + n^{-1} \mathbf{r}^{\intercal} \bar{\mathbf{c}}(t),
\end{aligned}$$

where the last equality holds because $\delta s_i(\tau+1) = d_i(\tau)$.

If the player switched at time t to $c_i'(\tau) = 0$ and $p_i'(\tau) = e_i + s_i'(\tau)$ for all $\tau \geq t$, then the continuation payoff would be

$$\begin{aligned}
(1-\delta)^{-1}\bar{\pi}_i'(t) &= \pi_i'(t) + \sum_{\tau=t+1}^{\infty} \delta^{\tau-t}\pi_i(\tau) \\
&= p_i'(t) + n^{-1}\mathbf{r}^\intercal \mathbf{c}'(t) + \sum_{\tau=t+1}^{\infty} \delta^{\tau-t}(p_i'(\tau) + n^{-1}\mathbf{r}^\intercal \mathbf{c}'(\tau)) \\
&= e_i + s_i(t) + n^{-1}(\mathbf{r}^\intercal \mathbf{c}(t) - r_i c_i(t)) + \sum_{\tau=t+1}^{\infty} \delta^{\tau-t}(e_i + n^{-1}\mathbf{r}^\intercal \mathbf{c}'(\tau)) \\
&= (1-\delta)^{-1}e_i + s_i(t) + n^{-1}(\mathbf{r}^\intercal \mathbf{c}(t) - r_i c_i(t)) + \sum_{\tau=t+1}^{\infty} \delta^{\tau-t} n^{-1}\mathbf{r}^\intercal \mathbf{c}'(\tau) \\
&\geq (1-\delta)^{-1}e_i + s_i(t) + n^{-1}(\mathbf{r}^\intercal \mathbf{c}(t) - r_i c_i(t)).
\end{aligned}$$

So $\bar{\pi}_i'(t) \leq \bar{\pi}_i(t)$ implies the following equivalent statements:

$$\begin{aligned}
e_i + (1-\delta)s_i(t) + (1-\delta)n^{-1}(\mathbf{r}^\intercal \mathbf{c}(t) - r_i c_i(t)) &\leq e_i + (1-\delta)s_i(t) - \bar{c}_i(t) + n^{-1}\mathbf{r}^\intercal \bar{\mathbf{c}}(t) \\
n\bar{c}_i(t) + (1-\delta)\sum_{j\neq i} r_j c_j(t) &\leq \sum_j r_j \bar{c}_j(t) \\
n\bar{c}_i(t) + (1-\delta)\sum_{j\neq i} r_j c_j(t) &\leq r_i\bar{c}_i(t) + (1-\delta)\sum_{j\neq i} r_j c_j(t) \\
&\qquad + \delta \sum_{j\neq i} r_j \bar{c}_j(t+1) \\
(n - r_i)\bar{c}_i(t) &\leq \delta \sum_{j\neq i} r_j \bar{c}_j(t+1) \\
\bar{c}_i(t) &\leq (\delta D \bar{\mathbf{c}}(t+1))_i.
\end{aligned}$$

So the continuation contribution sequence $(\bar{\mathbf{c}}(t))_t$ satisfies $\bar{\mathbf{c}}(t) \leq \delta D \bar{\mathbf{c}}(t+1)$ for all t. By Theorem 5 therefore, it is sustainable in the game $\Gamma_{\mathrm{B}}(\mathbf{r}, \delta)$. So the total contribution vector $\hat{\mathbf{c}}$ is also sustainable in that setting.

So the two statements are equivalent. □

Proof of Theorem 10. Let (r, δ) allow for non-defection.

Write $\mathcal{W}^{\mathrm{s}}$ for the set of welfare values attainable with sustainable contribution sequences in the saving model. Write $\mathcal{E}$ for the set of resource-efficiencies of non-defective SPEs in the base model. We can show that $\sup \mathcal{W}^{\mathrm{s}} = \sup \mathcal{E} + 1$. First, assume that $W \in \mathcal{W}^{\mathrm{s}}$. Then, in the saving model, there is $\mathbf{e}$ and a sustainable $\hat{\mathbf{c}}$ such that $W = W_{\mathbf{e}}(\hat{\mathbf{c}})$. We have

$$W = W_{\mathbf{e}}(\hat{\mathbf{c}}) = \sum_i \hat{\pi}_i = \sum_i (\hat{\pi}_i - e_i) + 1 \leq \frac{\sum_i (\hat{\pi}_i - e_i)}{\sum_i \hat{\mathbf{c}}_i} + 1 = E(\hat{\mathbf{c}}) + 1.$$

The value of $E(\hat{\mathbf{c}})$ is the same in both models. So $W \leq \sup \mathcal{E} + 1$.

Now take some $E \in \mathcal{E}$. In the base model, take a sustainable, non-defective $\hat{\mathbf{c}}$ such that $E = E(\hat{\mathbf{c}})$. By Corollary 10 and linearity of $E(\hat{\mathbf{c}})$, we can assume $\sum_i \hat{c}_i = 1$. By Proposition 6, take $\mathbf{e}$ such that $\hat{\mathbf{c}}$ is sustainable in the saving model. Now,

$$E + 1 = E(\hat{\mathbf{c}}) + 1 = \frac{\sum_i (\hat{\pi}_i - e_i)}{\sum_i \hat{\mathbf{c}}_i} + 1 = \sum_i \hat{\pi}_i = W_{\mathbf{e}}(\hat{\mathbf{c}}).$$

So $E + 1 \leq \sup \mathcal{W}^{\mathrm{s}}$. We therefore have $\sup \mathcal{W}^{\mathrm{s}} = \sup \mathcal{E} + 1$. We write this as $W^{\mathrm{s}}_{\mathrm{sup}} = E_{\mathrm{sup}} + 1$.

Write $\mathcal{W}$ for the set of welfare values that can be achieved with any $\mathbf{e}$ in the endowment model. We already defined $W_{\mathrm{sup}} = \sup \mathcal{W}$. Write $\mathcal{E}^{\mathrm{c}}$ for the set of resource-efficiencies of non-defective SPEs in the base model that are achieved with constant contribution sequences, and $E^{\mathrm{c}}_{\mathrm{sup}}$ for its supremum. We now show that $W_{\mathrm{sup}} = E^{\mathrm{c}}_{\mathrm{sup}} + 1$.

[HSH+24] show (their Lemma 3) that W_{sup} is attained, and that it is attained with a sustainable constant contribution sequence. Naturally, this requires that $\hat{\mathbf{c}} = \mathbf{e}$, that is, full contribution. But at full contribution, $W_{\mathbf{e}}(\hat{\mathbf{c}}) = E(\hat{\mathbf{c}}) + 1$. So $W_{\mathrm{sup}} \leq E^{\mathrm{c}}_{\mathrm{sup}} + 1$.

Now take some $E \in \mathcal{E}^{\mathrm{c}}$. In the base model, take a non-defective $\hat{\mathbf{c}}$ such that $E = E(\hat{\mathbf{c}})$, which is sustainable with a constant contribution sequence. Set $\mathbf{e} = \hat{\mathbf{c}}$. Since the constant sequence $(\mathbf{e})_t$ is sustainable in the base model, $\hat{\mathbf{c}}$ is also sustainable with the endowment constraint $\mathbf{e}$. Again, we have $E + 1 = W_{\mathbf{e}}(\hat{\mathbf{c}})$. So $E + 1 \leq W_{\mathrm{sup}}$. We therefore have $W_{\mathrm{sup}} = E^{\mathrm{c}}_{\mathrm{sup}} + 1$.

Making use of both of the identities that we derived, we have that $W^{\mathrm{s}}_{\mathrm{sup}} > W_{\mathrm{sup}}$ is equivalent to $E_{\mathrm{sup}} > E^{\mathrm{c}}_{\mathrm{sup}}$. The statement now follows from Theorem 8. □

Data and software availability

The example data shown in Figure 4.3 were generated with GNU Octave 8.4 and Python 3.12. The computer code used is published in in [H24̈b].

Author contribution statement

All authors conceived and discussed the study; V.H. analysed the model; V.H., C.H. and M.K. wrote the manuscript; all authors discussed the results and edited the manuscript. M.K. and K.C. contributed equally to this work.

CHAPTER 5

Strategies of direct and indirect reciprocity in general social dilemmas

This chapter was originally published as: Valentin Hübner, Laura Schmid, Christian Hilbe, and Krishnendu Chatterjee. Stable strategies of direct and indirect reciprocity across all social dilemmas. *PNAS Nexus*, May 2025

People frequently encounter situations in which individually optimal behaviours diminish the welfare of others. Such social dilemmas may, for example, lead individuals to put too little effort into group projects, or to overuse public resources [Ols71, Har68, BDV$^+$20, KHv$^+$23]. These types of conflict can be analysed using the mathematical framework of (evolutionary) game theory [FT91, Sig10, BR13]. This framework provides tools to describe individuals who, consciously or subconsciously, make decisions that affect others' well-being. In particular, this literature describes several mechanisms that help individuals to cope with their social dilemmas [Now06b]. One prominent mechanism, especially in the context of pairwise interactions, is reciprocity. According to this mechanism, individuals have more of an incentive to act in the interests of others if their actions today may be reciprocated in the future.

The literature on evolutionary game theory distinguishes several types of reciprocity. One type is direct reciprocity [Axe81, HCN18, GK21, RH23]. Here, individuals decide how to act based on their previous experience with the respective interaction partner. That is, when Alice decides how to treat Bob, she considers how Bob treated her in the past. Such conditional behaviours have primarily been explored in the context of the prisoner's dilemma [HS97, SP13, SP14b, Aki16] (Fig. 5.1A). In this game, the socially optimal choice of cooperation is dominated by defection. However, once individuals interact repeatedly, reciprocal strategies such as Tit-for-Tat [Axe81] can help sustain cooperation. Even though the repeated prisoner's dilemma has been the main model to study direct reciprocity, the very same mechanism can also be effective in weaker forms of social conflict [MVCS12, SP14a].

Another type of reciprocity is indirect reciprocity [NS98b, NS05, Sig12, Oka20]. Here, when Alice decides how to treat Bob, she takes into account Bob's overall behaviour, including how he acted towards Charlie or Dave. That is, she takes into account Bob's general reputation. Unlike direct reciprocity, this mechanism does not require repeated interactions among the same two individuals. It merely requires that individuals repeatedly interact within a larger community. With a few exceptions [NO14, CFW20], researchers study this type of reciprocity with an even more restricted type of social dilemma, the donation game (a special case of the prisoner's dilemma, see Fig. 5.1B). This literature suggests that cooperation can be sustained with a variety of strategies, most notably the 'leading-eight' norms [OI04, MH23].

Although direct and indirect reciprocity are based on a similar premise, most theoretical and experimental studies [FGF01, MPGFDL22, WM00, OYS+18] either consider one or the other. Such a reductionist approach is useful to clarify whether either mechanism can be effective on its own. At the same time, however, it renders many interesting research questions infeasible. For example, such models cannot explain how individuals would cope with conflicting pieces of evidence [e.g., when Alice's personal impression of Bob runs counter to his public reputation MBE13, MSA20]. Similarly, such models cannot explain why in direct reciprocity, cooperation can be maintained with comparably simple strategies, whereas indirect reciprocity seems to require strategies of greater complexity [LH01, BS04, BG16, SSP18, RSP19, MMKS+24]. Only more recently, researchers have begun to describe different types of reciprocity within a single framework [Rob07, SCHN21, SUOY24, PHG24]. The corresponding studies explore when people would rather adopt one type of reciprocity instead of the other. Unfortunately, however, these studies are restricted to the analysis of donation games only. As a result, they cannot capture synergistic interactions, as in the stag hunt game [Sky03, PSSS09] (Fig. 5.1C). Similarly, they cannot capture cases in which one individual's cooperation crowds out the need for others to cooperate, as in the volunteer's dilemma [Die85], the snowdrift game [DH05] or other classes of hawk-dove games (Fig. 5.1D). To describe the effects of reciprocity in full generality, it takes models that allow for all kinds of social dilemmas. We present such a model herein.

Such a generalisation is not straightforward. In donation games, the payoff consequences of one individual's cooperation are independent of whether or not their interaction partner cooperates too. This independence allows researchers to compute the players' payoffs explicitly, by solving a low-dimensional system of linear recursions [SCHN21]. Beyond donation games, this simple recursion no longer applies. Hence, an analysis of direct and indirect reciprocity across all social dilemmas requires a new set of proof techniques, which we summarise below (and which we discuss in full detail in the **Supplementary Information**).

To characterise whether socially optimal outcomes can be sustained with either direct or indirect reciprocity (or both), we extend the notion of so-called *equalizer* strategies. These strategies have been first introduced in the context of direct reciprocity [BNS97, PD12]. By implementing an equalizer strategy, individuals can unilaterally set their opponent's payoff to a fixed value. That is, opponents always get the same payoff, irrespective of their own behavior. Once all other players adopt an equalizer strategy, the remaining player thus neither has an advantage, nor a disadvantage, from deviating. This property makes equalizers a useful tool to prove the abstract existence of Nash equilibria. The resulting set of equalizers includes several well-known strategies, such as Generous Tit-for-Tat [Mol85, NS92] and generalisations thereof [GANH24]. By building on these ideas, Schmid *et al* [SCHN21] have shown that equalizers can also be used to sustain full cooperation in models of indirect reciprocity – provided the game at hand is a donation game. Herein, we characterise when such equalizers exist in arbitrary pairwise social dilemmas, for both direct and indirect reciprocity (and arbitrary mixtures). Along the

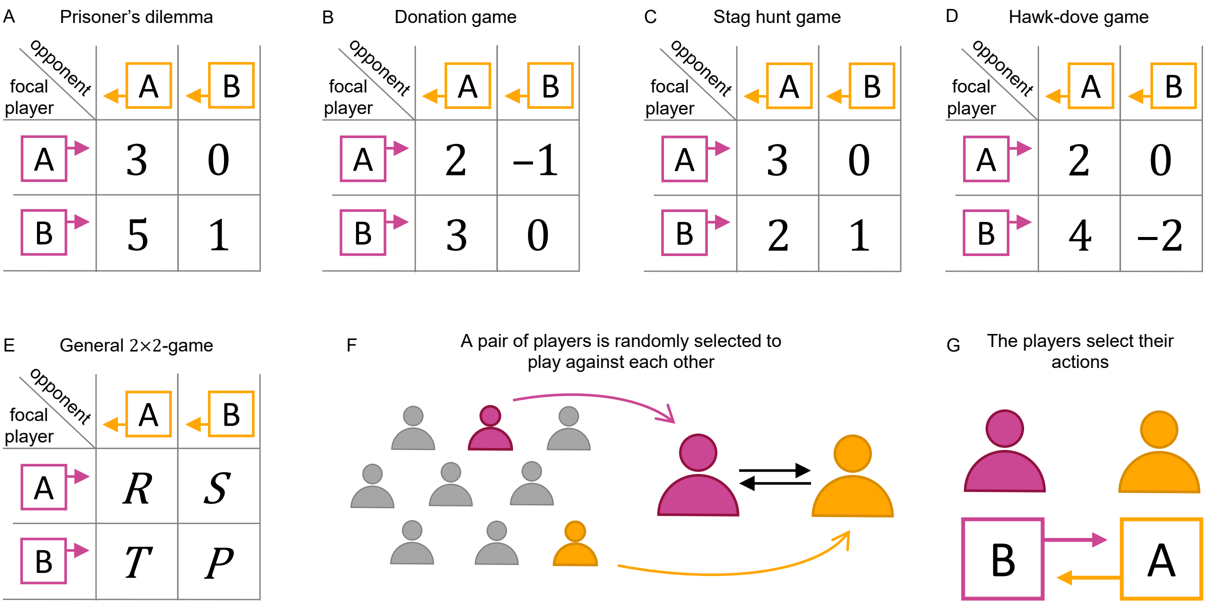

Figure 5.1: Illustration of the game dynamics. We consider populations of players who engage in pairwise social dilemmas with two actions A and B. A, In the prisoner's dilemma, the best outcome for the group is for both players to choose A. Yet individually, B is a dominant action. B, The donation game is a special case of a prisoner's dilemma. Here, action A can be interpreted as paying a cost $c>0$ for the co-player to receive a benefit $b>c$. Action B means to do nothing. In the depicted example, $b=3$ and $c=1$. C, The stag hunt game captures a dilemma in which players may fail to coordinate on the most profitable equilibrium. D, In the hawk-dove game, there are two (pure) equilibria. In each equilibrium one player chooses A and the other one B. Each player prefers to be the one who chooses B. E, In general, payoffs are often denoted by the letters R, S, T, P, respectively. F, The game unfolds over many rounds. Each round, two players are randomly drawn to interact with each other in the given social dilemma. G, Each of the players chooses action A or action B. Their choice might depend on their co-player's previous interactions.

way, we also prove that in general, equalizers need to be more complex than previously appreciated (in technical terms: in donation games, equalizers can be implemented with simple *reactive* strategies [SCHN21]. For more general social dilemmas, it takes the richer set of *memory-1* strategies instead).

Our model combines direct and indirect reciprocity within a single framework, irrespective of the considered social dilemma. This framework can serve as an important bridge to transfer insights from one field to another. Herein, we use this bridge, for example, to incorporate the well-known memory-one strategies from direct reciprocity into models of indirect reciprocity. In this way, we can prove that even in indirect reciprocity, fully cooperative outcomes can be sustained as a Nash equilibrium. In the past, such rigorous results for indirect reciprocity have been difficult to establish; these difficulties have been especially pronounced in the case of 'private information', when individuals are allowed to disagree on each others' reputations [HST$^+$18]. Instead, here we prove the existence of such Nash equilibria for arbitrary social dilemmas, and for players who are allowed to discount the future – even when information is private.

5.1 Results

Game setup. We consider a population of n individuals, referred to as players. These players repeatedly interact in a pairwise interaction. More specifically, in each round, two players are selected at random (Fig. 5.1F). Each player then chooses one of two actions, A or B (Fig. 5.1G). Their choices determine the payoffs they get, according to the given payoff matrix (Fig. 5.1A-E). All other population members observe the interaction, but they may independently misperceive each player's action with some probability $\varepsilon < 1/2$. Instead of correctly identifying a player's action as, say, action A, they perceive it to be B, or vice versa. After the interaction has taken place, there is another round with continuation probability d. In that case, a new pair of players is randomly drawn to interact with one another. With the converse probability $1-d$, the game ends. The total payoff of each player is defined as the sum of the payoffs they obtained in each round, times a normalisation constant. Equivalently, one may also interpret this setup as an infinitely repeated interaction in which players discount future rounds by a constant factor d. In the limit $d \to 1$, we recover the classical case of an infinitely repeated game without discounting, as for example in Press & Dyson [PD12]; see Methods for details.

The exact nature of the game played each round depends on the four entries R, S, T, P of the payoff matrix (Fig. 5.1E). In the following, we are particularly interested in games that can be interpreted as social dilemmas. Based on the 'individual-based' interpretation in Kerr *et al* [KGSF04], this means payoffs satisfy the following constraints. First, players prefer mutually choosing A to mutually choosing B, such that $R > P$ (except for the degenerate case of $R = P$, this assumption is without loss of generality; otherwise we just need to relabel the two actions). Second, players always prefer their co-player to choose A, implying $R > S$ and $T > P$. Third, in mixed pairs, the player who chooses B gets the higher payoff, such that $T > S$. Together these assumptions ensure that on a collective level, individuals have some incentive to choose A; yet on an individual level, they may want to choose B. Accordingly, we interpret action A as cooperation, and we associate B with defection (however, we use the more neutral letters A and B, rather than the usual letters C and D, to highlight that our framework is not restricted to the prisoner's dilemma).

The notion of a social dilemma captures several classical games. (i) In the prisoner's dilemma, payoffs satisfy the inequalities $T > R > P > S$ and $2R > T + S$, as in Fig. 5.1A. (ii) The donation game additionally requires $R + P = S + T$, see Fig. 5.1B. Such games are sometimes called 'additive' [MRH21, CP23]. (iii) The stag hunt game satisfies $R > T > P > S$, as depicted in Fig. 5.1C. (iv) Finally, the hawk-dove game satisfies $T > R > S > P$, as in Fig. 5.1D. The exact payoff ranking determines the severity of the dilemma. Among the above examples, players arguably face the strongest conflict between cooperation and defection in the prisoner's dilemma and the donation game. However, also the other two games entail some conflict. Players may either have difficulties to coordinate on the equilibrium that is better for both (as in the stag hunt game), or they may prefer different equilibria altogether (as in the hawk-dove game).

Reactive and memory-1 strategies. When playing the above games, players make their decisions based on their strategies. Strategies are recipes that tell the player what to do, depending on the outcome of previous interactions.

In order to allow for an explicit analysis, researchers often consider a restricted space

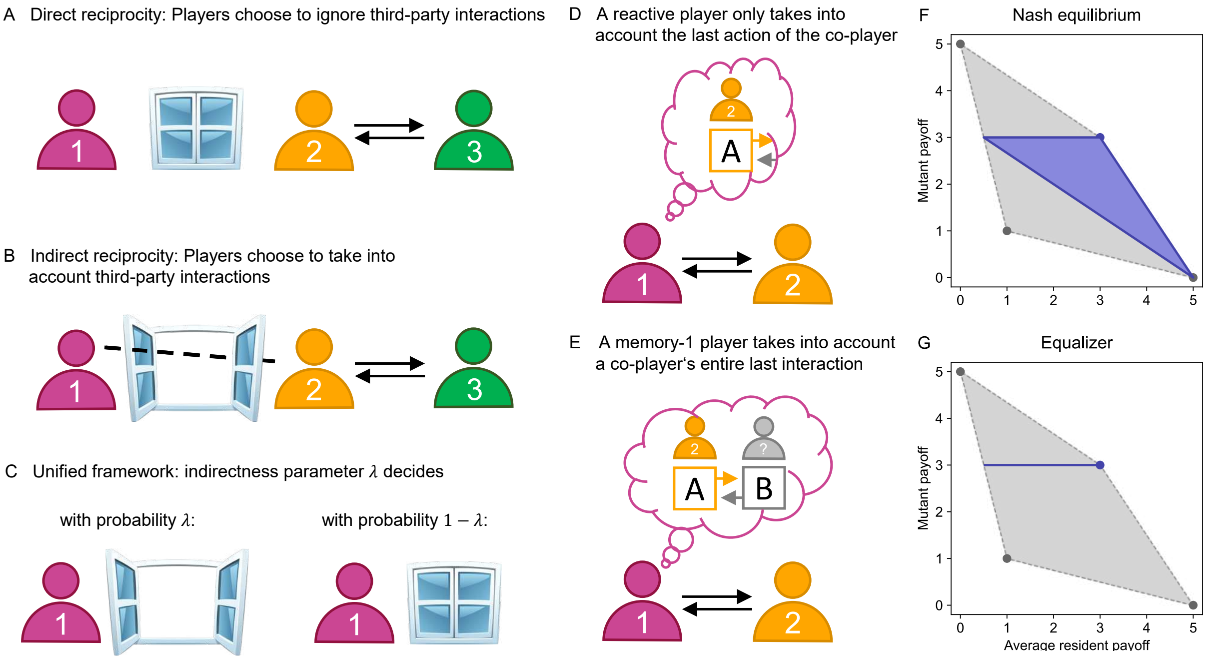

Figure 5.2: Strategies of direct and indirect reciprocity. Strategies of reciprocity differ in whether or not players (here, Player 1) take into account third-party interactions (here, between players 2 and 3). A, According to direct reciprocity, Player 1 ignores third-party interactions. B, According to indirect reciprocity, Player 1, takes such third-party interactions into account. C, Our framework also allows for intermediate cases, where Player 1 considers third-party interactions with some fixed probability λ. D, For our model, we consider strategies of different complexity. When using a reactive strategy, players condition their behaviour on the last observed action of the co-player. E, When using a memory-1 strategy, they condition their behaviour on the entire outcome of the co-player's last interaction. This also involves the action of the co-player's last opponent. F, A strategy is a Nash equilibrium if it is a best response to itself. The effect of such strategies can be represented graphically. The grey area represents all feasible payoffs in the respective game. The blue area represents the payoffs that are still feasible if every player adopts the same fixed resident strategy, except for one deviating mutant. In this example, the resident strategy yields a payoff of three against itself (indicated by the upper right dot). According to the blue area, no mutant strategy yields a higher payoff. Hence, the given resident strategy is a Nash equilibrium. G, An equalizer strategy is a special case of a Nash equilibrium. Here, the mutant's payoff is always the same, regardless of the mutant's strategy.

of strategies. For example, Schmid et al. [SCHN21] consider strategies of the form $\sigma = (p_0, p_A, p_B, \lambda)$. Here, the first parameter p_0 is a player's cooperation probability against an unknown co-player. The next two parameters p_A (p_B) give the player's cooperation probability against a co-player who cooperated (defected) in their last relevant interaction. Finally, the parameter λ determines which previous interactions of the co-player are deemed relevant. When $\lambda = 0$, only direct interactions matter. For example, if Bob previously defected against Alice (played B), but then cooperated with Charlie (played A), Alice would use cooperation probability p_B against Bob. That is, Alice implements a strategy of direct reciprocity (Fig. 5.2A). In contrast, when $\lambda = 1$, players take into account all their co-players' interactions equally, even interactions with third parties. As a result, such players base their decision on the very last action of the co-player, independently of whether or not they were personally involved. In the above example, if Bob defected against Alice (played B), but then cooperated with Charlie (played A), Alice's cooperation probability against Bob is p_A. Now, Alice uses a strategy of indirect reciprocity (Fig. 5.2B). The model also allows for intermediate values of $\lambda \in (0,1)$. In that case, Alice takes into account third-party interactions with probability λ (Fig. 5.2C). As the above strategies merely respond to the *co-player's* previous behaviour, they are called *reactive* [Sig10] (Fig. 5.2D). The set of reactive strategies includes Always Cooperate $\sigma = (1,1,1,\lambda)$, Tit-for-Tat $\sigma = (1,1,0,0)$, and its indirect reciprocity analogue Simple Scoring [NS98a] $\sigma = (1,1,0,1)$, among others.

In the context of direct reciprocity, it is also common to consider a slightly more general strategy set, called *memory-1 strategies*. Here, a player does not only take into account the co-player's last action. Rather the player takes into account the entire context of the co-player's previous interaction, including the action of the co-player's opponent (Fig. 5.2E). Memory-1 strategies take the form $\sigma = (p_0, p_{AA}, p_{AB}, p_{BA}, p_{BB}, \lambda)$. The interpretation of the entries p_0 and λ is the same as before. However, now p_{xy} is a player's cooperation probability given that in the co-player's last relevant interaction, the co-player used action y whereas the co-player's opponent used action x. Within the set of memory-1 strategies we can represent reactive strategies as those strategies for which $p_{AA} = p_{BA}$ and $p_{AB} = p_{BB}$. Here, only the co-player's last relevant action matters. A well-known example of a non-reactive memory-1 strategy is Win-Stay Lose-Shift [NS93], $\sigma = (1,1,0,0,1,0)$. Here, a player would only cooperate with a co-player if in their previous joint interaction either both cooperated, or no one did [NS93].

Partner strategies. In the following, we are interested in whether mutual cooperation can be sustained by either direct or indirect reciprocity. To this end, we study strategies σ with two properties. First, the strategy ought to be *nice* [Axe81]. That is, if σ is adopted by everyone, the entire population cooperates indefinitely in the absence of errors. Second, the respective strategy ought to be a Nash equilibrium: if adopted by everyone, no single player has an incentive to deviate (Fig. 5.2F). In the context of direct reciprocity, strategies that satisfy both properties are called *partners* [HTS15]. The answer to the question whether partners exist turns out to be trivial in the stag hunt game or in the so-called harmony game [MVCS12]. In those games, payoffs satisfy $R > T$. Therefore mutual cooperation is a Nash equilibrium even if the game is only played once. It trivially follows that mutual cooperation can also be sustained within our repeated setup – players merely need to use the strategy Always Cooperate. In the following, we will thus focus more on the other two game classes, the prisoner's dilemma and the hawk-dove game.

To show existence of partner strategies in those games, we characterise a particular subset of Nash equilibria, those based on equalizer strategies. Such strategies do not only ensure that no player can unilaterally improve their payoff; they ensure every deviating player gets the same payoff (Fig. 5.2G). For direct reciprocity ($\lambda = 0$), the existence of equalizers has been shown by Boerljist *et al* [BNS97] and Press & Dyson [PD12]. Their result applies to the infinitely repeated prisoner's dilemma without errors ($d = 1$ and $\varepsilon = 0$). For indirect reciprocity ($\lambda = 1$), the existence of equalizers follows from the work of Schmid *et al* [SCHN21], but only for the restrictive case of donation games (but arbitrary d and ε). Instead, here we characterise equalizers for all social dilemmas, for direct and indirect reciprocity, and for all continuation probabilities and error rates. All details and proofs are in the Supplementary Information. Below we summarise the respec-

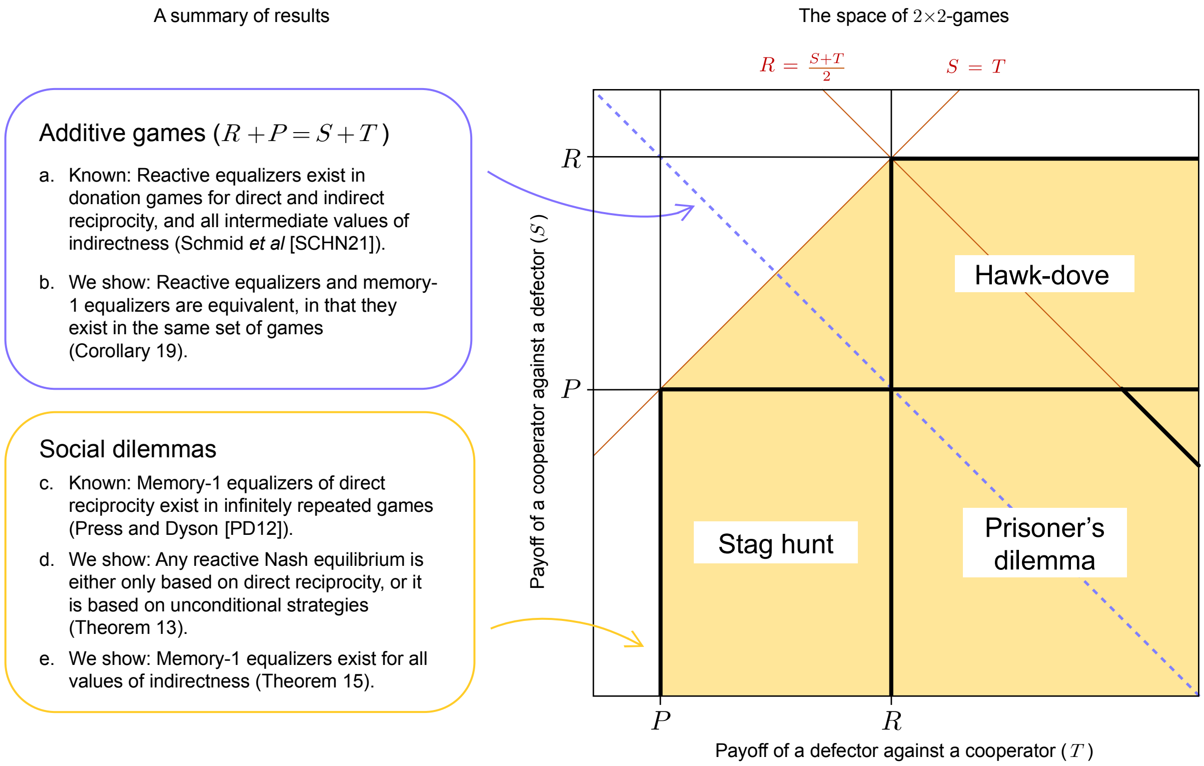

Figure 5.3: Summary of our results. Herein, we derive results on the existence of 'equalizer strategies' in pairwise social dilemmas. The right half of the figure represents the space of all such dilemmas graphically. Here, we keep the two payoff parameters R and P fixed (with $R>P$). We vary the remaining payoffs T and S. The orange region indicates all games that satisfy the conditions of a social dilemma. The blue dashed line indicates the subspace of 'additive games', which includes the donation game. All previous models that combine direct and indirect reciprocity focus on this blue subspace [Rob07, SCHN21, SUOY24, PHG24]. For the general space of social dilemmas, the existence of equalizers has only been established for direct reciprocity [PD12], but not for indirect reciprocity or any mixtures.

tive results. For a visual representation of previous work and our contribution, see Fig. 5.3.

Reactive strategies in the donation game. To better motivate our contribution, let us first recapitulate the results of Schmid *et al* [SCHN21]. They considered a similar setup as ours, but restricted to the donation game and to players with reactive strategies. Payoffs of the donation game are given by $R=b-c$, $S=-c$, $T=b$, and $P=0$, where b and c are the benefit and cost of cooperation. Schmid *et al* show that for full cooperation to be sustainable, the pairwise continuation probability δ needs to be sufficiently large (this is the probability that a given pair of players will interact again given it just interacted; this probability is directly related to the population-wide probability d, see Supplementary Information). The exact threshold for δ depends on the players' indirectness parameter λ. For direct ($\lambda=0$) and indirect reciprocity ($\lambda=1$), the respective thresholds are

$$\delta_0=\frac{c}{b} \quad \text{and} \quad \delta_1=\frac{c}{b+(n-2)\big((1-2\varepsilon)b-c\big)}. \tag{5.1}$$

In particular, indirect reciprocity makes it easier to sustain cooperation (compared to direct reciprocity) if the population is large and errors are rare. Either way, once the respective threshold is reached, mutual cooperation can be enforced with an equalizer strategy. In case of direct reciprocity, the respective equalizer is Generous Tit-for-Tat [NS92]. In case of indirect reciprocity, it is Generous Scoring [SCHN21].

To derive the above results, both the restriction to donation games and to reactive strategies turns out to be crucial. Because the donation game satisfies the additivity property $R+P=S+T$, the payoff of each player can be decomposed into a sum of two terms. The first term only depends on on the statistical distribution of the co-players' actions (affecting whether or not I receive a benefit b). The other term only depends on the statistical distribution of the own action (affecting whether or not I pay the cost c). That is, the players' actions affect payoffs independently. Furthermore, for reactive strategies, the distribution of a player's own actions affects the distribution of the co-players' actions by a linear relationship. Based on these two observations, one can derive a simple linear recursion for the players' likelihood to cooperate with each other in any given round. With this recursion, it becomes straightforward to compute payoffs. Unfortunately, once either the game is non-additive, or players use more complex strategies, the above approach is no longer viable. Hence, for the results below we rely on proof techniques that do not require us to compute the players' payoffs explicitly.

Memory-1 strategies in the donation game. To make progress, we first explore whether cooperation in the donation game is easier to sustain when players are allowed to use memory-1 strategies. More specifically, we ask whether there are memory-1 equalizers that can sustain full cooperation even when the respective condition in (5.1) is violated. The answer is negative. We find that for all game parameters and any indirectness λ, memory-1 equalizers exist if and only if reactive equalizers exist (Supplementary Information Corollary 5). This result resonates with earlier work on direct reciprocity. For the infinitely repeated donation game, it was shown that reactive strategies can enforce all linear payoff relationships that are theoretically possible [HNT13]. Thus at least in the donation game, allowing for more complex strategies does not provide any additional advantage with respect to implementing equalizer strategies (Fig. 5.3B).

Reactive strategies in general social dilemmas. Given the strong properties of reactive strategies in donation games, we ask whether they can also sustain full cooperation in other social dilemmas. Surprisingly, the answer is negative. To describe this result more formally, we introduce the notion of a degenerate strategy. A reactive strategy is degenerate if it exclusively relies on direct reciprocity ($\lambda=0$), or if it acts unconditionally ($p_A=p_B$). That is, degenerate strategies completely ignore any third-party interactions. Similarly, we say an equilibrium is degenerate if it requires players to use degenerate strategies. Using this notion, we can formulate the main result of this section as follows: In any non-additive game with more than two players ($n>2$) and positive error rates ($\varepsilon>0$), any Nash equilibrium in reactive strategies is degenerate. In other words, if players are in a Nash equilibrium that entails at least some indirect reciprocity ($\lambda>0$), players must be using unconditional strategies such as Always Defect. (The above result also implies that for $\lambda>0$, there are usually no equalizer strategies, because unconditional strategies are in general not equalizers [HNT13].) This finding suggests that earlier results on the donation game [SCHN21] are sensitive to the exact payoff values. Once the payoff matrix is slightly perturbed, reactive partner strategies that entail some indirect reciprocity cease to exist (Fig. 5.3D).

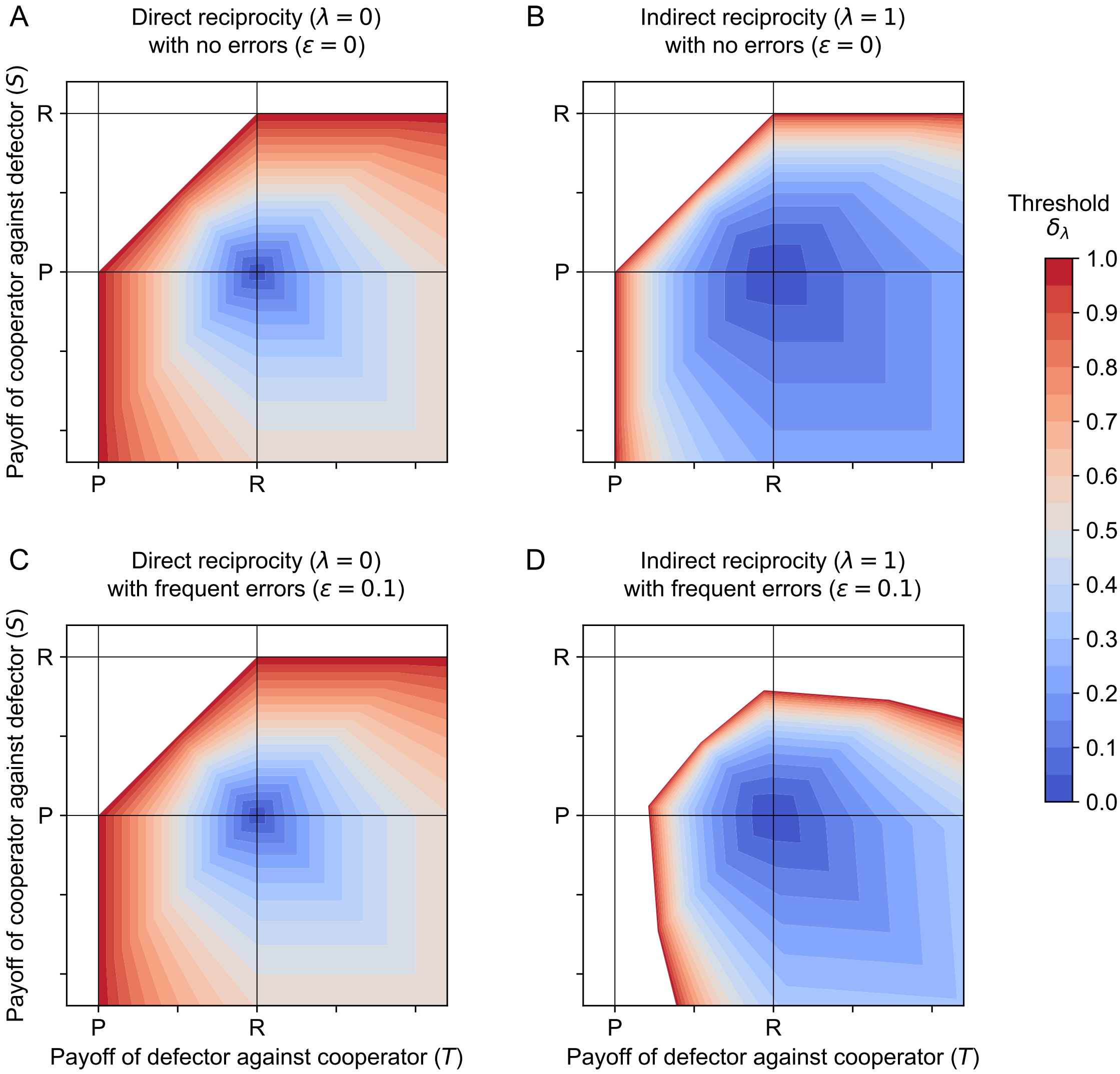

Figure 5.4: Feasibility of equalizer strategies across all social dilemmas. We graphically represent whether or not equalizer strategies exist. To this end, we consider four different cases. The cases depend on whether individuals use direct (left) or indirect reciprocity (right), and on whether or not there are errors (top vs. bottom). In each case, we consider the space of all social dilemmas (as in Fig. 5.3). For each possible game, we depict how large the pairwise continuation probability δ needs to be for equalizers to exist. Low values of δ (blue) indicate that the conditions for equalizers are easy to satisfy. Higher values of δ (red) suggest that equalizer strategies only exist for rather high continuation probabilities. A,B, The figure shows that without errors, indirect reciprocity is more favourable to the existence of equalizers. C,D, Once third-party observations are subject to errors, there are regions in which direct reciprocity allows for equalizers whereas indirect reciprocity does not.

The proof of the above result is constructive: We show that for any such reactive resident strategy, one can construct a deviating strategy that gets a strictly higher payoff. Interestingly, the deviating strategy is not reactive. Rather, it is a higher-memory strategy that takes into account the joint distribution of previous actions across different pairs of players. We show that such strategies have a payoff advantage when the process involves at least some randomness (e.g., when there are errors). We provide a description of the deviation strategies in the Methods, and a proof of their superiority in the Supplementary Information.

Memory-1 strategies in general social dilemmas. The above result raises the question whether beyond the simple donation game, non-degenerate equalizers exist at all. To explore that question, we search the space of memory-1 strategies. There, we find that the answer is positive. For any social dilemma and any indirectness λ, there exist equalizer strategies for sufficiently large continuation probabilities and sufficiently small error rates (Fig. 5.3E, see Supplementary Information). Similar to (5.1) for the donation game, the minimum continuation probability can be computed explicitly. For example, assuming $\varepsilon = 0$ and $T > R$, we find that fully cooperative equalizers with indirectness λ exist if and only if the pairwise continuation probability exceeds the threshold [5.2],

$$\delta_\lambda = \left(1 + \Big(1+(n-2)\lambda\Big)\frac{\min\{T,R\}-\max\{P,S\}}{\max\{|T-R|,\ |P-S|\}}\right)^{-1}. \tag{5.2}$$

A few remarks are in order. First, for social dilemmas, this threshold is strictly smaller than one. Hence the condition can be satisfied for sufficiently large δ. Second, for any population size $n > 2$, it is easy to verify that threshold (5.2) is strictly lower for indirect reciprocity than for direct reciprocity. This is a consequence of our assumption that there are no errors. Once the error rate becomes positive, direct reciprocity may become the more favourable mechanism for full cooperation (Fig. 5.4). Third, in the special case of the donation game, the threshold simplifies to the following values for direct ($\lambda = 0$) and indirect reciprocity ($\lambda = 1$):

$$\delta_0 = \frac{c}{b} \qquad \text{and} \qquad \delta_1 = \frac{c}{b+(n-2)(b-c)}. \tag{5.3}$$

That is, we recover the earlier conditions in (5.1) for reactive strategies for $\varepsilon = 0$.

In the more general case of an arbitrary prisoner's dilemma and of the hawk-dove game (with $T > R$), we show that once the condition (5.2) is satisfied, one can always find equalizer strategies that enforce the mutual cooperation payoff R (see Supplementary Information, Proposition 5). Again, our proof is constructive. In the Methods, we provide an algorithm that produces an optimal equalizer strategy for all social dilemmas (even for positive error rates). For games with $T > R$, this algorithm produces nice strategies (i.e. $p_0 = p_{AA} = 1$). Together with our earlier observation that cooperation is trivial to sustain in the stag hunt and the harmony game (with $T < R$), we conclude that stable cooperation can always be achieved with memory-1 strategies, based on direct or indirect reciprocity, or any arbitrary mixture of the two.

Overall, the above results represent a considerable generalisation of previous work. We recover the seminal results of Press & Dyson [PD12], when we restrict our framework to direct reciprocity in infinitely repeated games ($\lambda = 0$ and $d = 1$). Similarly, we recover the results of Schmid et al. [SCHN21], when we restrict our framework to reactive strategies, and to donation games only (see Methods for details).

Simulation results. To further illustrate the above results, we have explored the game dynamics when $n-1$ players act according to a given equalizer strategy (produced by Algorithm 1 in the Methods section). For the remaining player, we have sampled $N = 100$

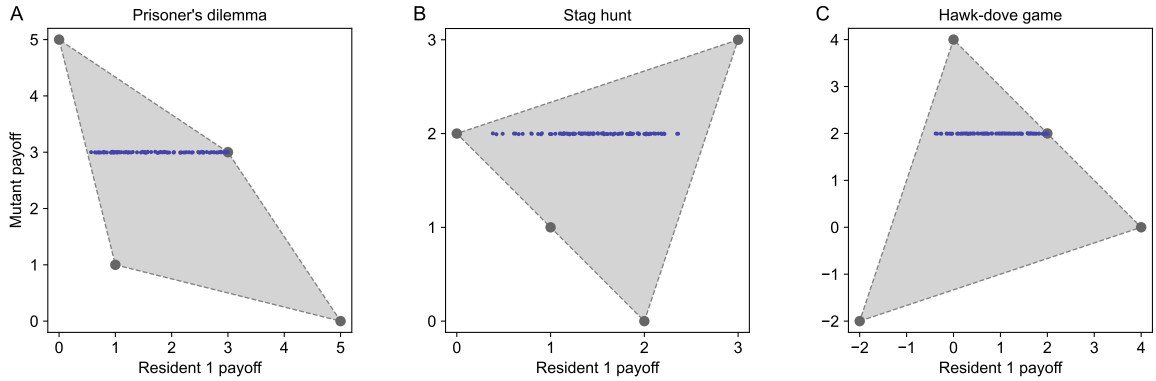

Figure 5.5: Simulation of equalizer strategies. We consider three social dilemmas in which $n-1$ residents use a fixed equalizer strategy. For the remaining player, we randomly sample $N=100$ mutant strategies. For each mutant strategy, we simulate the resulting game dynamics. Based on these simulations, we compute the average payoff in pairwise interactions between the deviating mutant player and a given resident player ('Resident 1', see Methods for details). These pairwise payoffs are depicted as blue dots. As expected from our analytical results, all mutant strategies yield the same average payoff (whereas the payoff of the resident may vary). In two of the three cases, the respective payoff is optimal (A, C). Only in the stag hunt game, equalizers cannot enforce the socially optimal payoff (B). But also in those games (with $T<R$), the socially optimal payoff of mutual cooperation can still be achieved in equilibrium. Players merely need to use the strategy of Always Cooperating instead. For reference, the grey area shows the set of feasible payoffs for the respective game.

random 'mutant' strategies. To approximate the players' resulting payoffs, we simulated many independent instances of the game dynamics, separately for each mutant strategy (in contrast to previous work on reactive strategies in donation games [SCHN21], there is no known formula to compute the players' payoffs explicitly). Fig. 5.5 shows the results. We depict the residents' and the mutant's average payoff for three different social dilemmas (the prisoner's dilemma, the stag hunt game, and the hawk-dove game). In each case, we find that the simulated payoffs indeed form a straight horizontal line, the characteristic property of an equalizer strategy (Fig. 5.2G). In particular, in each case the produced resident strategy is a Nash equilibrium: Once adopted by everyone, no mutant strategy has a selective advantage.

5.2 Discussion

Direct and indirect reciprocity are important determinants of human behaviour in social dilemmas [Now06b]. They are arguably among the key mechanisms to explain our exceptionally high cooperation rates [RN13]. Yet despite the many similarities between the two kinds of reciprocity, they are typically studied independently. Even worse, respective models often differ substantially. Models of direct reciprocity tend to study the prisoner's dilemma [Axe81, HCN18, GK21, RH23], whereas indirect reciprocity models are often based on the narrower class of donation games [NS98b, NS05, Sig12, Oka20]. Similarly, individuals in direct reciprocity models are typically assumed to adopt reactive or memory-1 strategies [Sig10, SHCN22]. In contrast, models of indirect reciprocity focus on subsets of 'third-order social norms' [OI04], which do not map easily onto either class of direct reciprocity strategies. These differences make it difficult to compare the two mechanisms directly. Moreover, they make it difficult to generalise insights from one field to the other. To address these problems, we join recent efforts to study a unified framework [Rob07, SCHN21, SUOY24, PHG24], in which individuals themselves choose which kind of reciprocity they use.

Previous models that combine direct and indirect reciprocity are based on the smallest common denominator of the two literatures, the donation game. This game is the simplest metaphor of cooperation: individuals pay some cost to provide a benefit to someone else. This simplicity promotes a mathematical analysis, and it permits an intuitive interpretation of the results. However, this game rules out possible interdependencies between the individuals' actions. Neither is it particularly beneficial if individuals cooperate at the same time, nor is it particularly damaging if they all defect simultaneously. This assumption makes it impossible to study games in which mutual cooperation yields synergistic benefits, as in the stag hunt game. Similarly, it rules out interactions in which individual actions are strategic substitutes, as the volunteer's dilemma [Die85] or the snowdrift game [DH05]. Even among all prisoner's dilemmas, donation games only represent a negligible subset of measure zero. These considerations highlight a need to study models that allow for more general types of social dilemmas. We present such a model herein.

We use this model to characterise strategies that can sustain full cooperation. Our results show that these strategies do not need to be overly complex. Instead, it suffices that individuals take into account the last interaction of the respective group member (e.g., to consider memory-1 strategies). Our results also show that simpler strategies (reactive strategies) in general do not suffice to support cooperation. In fact, the only domain in which these strategies suffice are the donation games considered earlier [Rob07, SCHN21, SUOY24, PHG24]. Together, these two observations represent a nice characterisation of the complexity of strategies that is necessary and sufficient to ensure stable cooperation for all social dilemmas.

While the basic setup we consider is thus similar to earlier unified frameworks of reciprocity [SCHN21, SUOY24, PHG24], the mathematical tools we apply are vastly different. Earlier work exploited the advantage that payoffs of reactive players in the donation game can be computed explicitly. Instead, for our results we focus more on the notion of equalizer strategies, as introduced by Boerljist *et al* [BNS97] and Press & Dyson [PD12]. These strategies have the remarkable property that they can unilaterally control the co-player's payoff, independent of the co-player's strategy (Fig. 5.2G). This makes them extremely useful tools to construct Nash equilibria. Once every population member adopts an equalizer strategy, no single player has an incentive to deviate. Interestingly, however, deviators suffer no harm either. In particular, Nash equilibria based on equalizer strategies do not satisfy the stronger notion of evolutionary stability [MSP73].

Evolutionary stability is generally difficult to achieve in repeated games. In fact, for the standard case of infinitely repeated games without errors, no strategy is evolutionarily stable [BL87, Lor94, GV18]: one can always identify mutant strategies that may invade by neutral drift. But even identifying strategies that satisfy the weaker condition of being a Nash equilibrium has been difficult in the field of indirect reciprocity. These difficulties are particularly apparent in models with 'private information' [Oka20]. In such models, individuals may hold different views on which reputation they assign to others. These disagreements may accumulate over time, which makes cooperation difficult to sustain [Uch10, HST+18, MH24]. To make analytical progress, the concept of equalizer strategies is particularly convenient. These strategies allow us to give a proof of principle: We rigorously show the existence of cooperative Nash equilibria, for any pairwise social dilemma, for all sufficiently large continuation probabilities, for direct and indirect reciprocity – even under private information.

Interestingly, however, even the most favorable equalizer strategies do not necessarily produce the socially optimal outcome. One counterexample is the stag hunt game (Fig. 5.5B). Here, equalizers exist, but they cannot ensure the optimal payoff of R. This insufficiency, however, does not diminish our results, nor is it a surprise. Because the stag hunt game's payoffs satisfy both $R>T$ and $R>P$, a co-player can always avoid an average payoff of R by defecting in all rounds. Hence, unilaterally imposing a guaranteed payoff of R on the co-player is clearly infeasible. Nevertheless, our more general result, that the game allows for full cooperation in equilibrium, holds. In this game, players simply need to adopt the strategy of always cooperating, instead of adopting an equalizer strategy.

To sum up, social dilemmas are at the core of many collective action problems. To resolve them, people frequently respond to an opponent's previous behaviour. Prior to making a decision, they form opinions about their opponent, either based on direct experiences, an opponent's third-party interactions, or both [MBE13, MSA20]. In our work, we mathematically characterise strategies people can use to sustain cooperation, independently of the kind of reciprocity they adopt, and independently of the specific social dilemma at hand.

5.3 Methods

Game dynamics and resulting payoffs. We consider a population of n players. Each round, two players are selected uniformly at random. They each play action A or action B and receive a payoff determined by the corresponding entries of the payoff matrix. The $n-2$ players who were not selected for that round receive a payoff of zero. Let $\pi_i(t)$ denote Player i's resulting expected payoff in round t. We define a player's total payoff as the sum of these one-round payoffs times a normalisation factor of $(1-d)n/2$, so that the total expected payoff is

$$\pi_i = (1-d)\frac{n}{2}\sum_{t=0}^{\infty} d^t \pi_i(t). \tag{5.4}$$

The normalisation factor ensures that the resulting values are in the same range as the game's one-round payoffs. For example, when all players use action A in all their interactions, the above formula guarantees that each player's expected total payoff (across all interactions and rounds) is R. Alternatively, one may also interpret the above payoff formula to represent a scenario in which individuals discount future payoffs by a constant rate d [FM86b, Abr88, NS95, MH16b, IM18]. In contrast, Press & Dyson [PD12], among many other works, consider the asymptotic behaviour of a game without discounting. This enables them to compute payoffs by calculating the stationary distribution of the Markov Chain defined by the game process. In the limit of $d \to 1$, our payoff definition approaches theirs.

Except for the case of a donation game among players with reactive strategies [SCHN21], there is no known closed-form solution to compute a player's expected payoff $\pi_i(t)$ in a given round t. Hence, for practical purposes, the value of (5.4) needs to be approximated numerically with simulations. For example, for a given population composition, we may independently simulate the above game dynamics k times. For each run, we sum up each player's payoff across all rounds. Then we sum up these total payoffs across all simulation runs, divide by k, and multiply with the normalisation constant $(1-d)n/2$.

In Fig. 5.5, we have run $k = 5 \cdot 10^5$ independent simulations for each mutant strategy. For graphical purposes, there we only report payoffs from interactions between the mutant and one given resident, Resident 1 (while ignoring the payoffs from all other interactions). In this case, the relevant normalisation constant is $(1-d)n(n-1)/2$. For the mutant, this procedure gives the same result as (5.4). However, for the resident, the result is different from (5.4), because we neglect the resident's payoff against other residents. For each of the three panels, the resident strategy is determined by Algorithm 1. All have $\lambda = 0.5$ and $p_0 = 1$. The other entries $(p_{AA}, p_{AB}, p_{BA}, p_{BB})$ are $(1.000, 0.331, 1.000, 0.666)$ for the prisoner's dilemma, $(0.498, 1.000, 1.000, 0.498)$ for stag hunt, and $(1.000, 0.498, 0.498, 1.000)$ for the hawk-dove game.

Instability of non-degenerate reactive strategies. In the results section, we have argued that in general social dilemmas, only degenerate reactive strategies can be stable. Here, we outline the respective proof (all details are in the Supplementary Information). To this end, consider a resident population of $n > 2$ players, who all adopt the same reactive strategy $\sigma = (p_0, p_A, p_B, \lambda)$. Assume the strategy is non-degenerate, $p_A \neq p_B$, $\lambda < 1$, and that errors are possible, $\varepsilon > 0$. For the proof, we construct a set of four (non-reactive) strategies. Then we show that at least one of them can invade the resident population.

First we construct events E_A and E_B. Both E_A and E_B completely define which players are selected and what actions they play in the first $n+3$ rounds, and do so identically apart from the action of Player 1 in round 2. The below table summarises these first $n+3$ rounds. The mutant player is Player 1. In round 2, x is action A in E_A and action B in E_B, whereas for y, any consistent choice is permissible. In round 3, $\overline{y}$ is the action that is not y. The dashes indicate actions that are defined by the event, but not specified explicitly in our proof. We show in Supplementary Information Proposition 10 that we can make these choices in such a way that E_A and E_B occur with positive probability.

Round	0	1	2	3	4	5	6	…	$n+2$
Players	2 3	1 3	1 2	2 3	1 3	1 3	1 4		1 n
Actions	– –	– –	x y	$\overline{y}$ –	A –	B –	– –		– –

The invader strategy σ' normally plays in the same way as strategy σ. Only when event E_x has occurred does σ' deviate from σ. It does so by, in one case of $x \in \{A, B\}$, slightly increasing its probability to play action A towards Player 2 next time they are selected to play together, and slightly decreasing it by an identical amount in the other case. Other than that, σ' continues to play exactly like σ.

Construction of equalizer strategies. In the following, we outline how equalizer strategies can be constructed within the space of memory-1 strategies. In the Supplementary Information, we show that for a strategy with cooperation probabilities $p = (p_{AA}, p_{AB}, p_{BA}, p_{BB})^\intercal$ and indirectness λ to be a generic equalizer, it needs to have the form

$$\begin{pmatrix} p_{AA} \\ p_{AB} \\ p_{BA} \\ p_{BB} \end{pmatrix} = (\delta^{-1} + \lambda(n-2))(I+\lambda(n-2)M_\varepsilon)^{-1} \begin{pmatrix} 1-\alpha R-\beta \\ 1-\alpha T-\beta \\ -\alpha S-\beta \\ -\alpha P-\beta \end{pmatrix}. \tag{5.5}$$

Here, α and β are constants, I denotes the identity matrix and M_ε is the error matrix

$$M_\varepsilon = \begin{pmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{pmatrix}^{\otimes 2} = \begin{pmatrix} (1-\varepsilon)^2 & (1-\varepsilon)\varepsilon & \varepsilon(1-\varepsilon) & \varepsilon^2 \\ (1-\varepsilon)\varepsilon & (1-\varepsilon)^2 & \varepsilon^2 & \varepsilon(1-\varepsilon) \\ \varepsilon(1-\varepsilon) & \varepsilon^2 & (1-\varepsilon)^2 & (1-\varepsilon)\varepsilon \\ \varepsilon^2 & \varepsilon(1-\varepsilon) & (1-\varepsilon)\varepsilon & (1-\varepsilon)^2 \end{pmatrix}.$$

The payoff that this equalizer strategy enforces is

$$\pi = \alpha^{-1} \left(\frac{1-\delta}{1+(n-2)\delta\lambda} p_0 - \beta \right).$$

Note that whether or not a strategy is an equalizer does not depend on p_0. This general result captures the results of several previous studies as special cases.

1. In Schmid *et al* [SCHN21], the authors consider the special case of reactive strategies in additive games (with $T \neq P$). Their Supplementary Information Eq. (13) states that a reactive strategy (p_0, p_A, p_B, λ) is an equalizer if and only if

   $$p_A - p_B = \frac{1+(n-2)\delta\lambda}{1+(n-2)(1-2\varepsilon)\lambda} \cdot \frac{P-S}{\delta(T-P)}. \tag{5.6}$$

   We can recover this result from our (5.5). For reactive strategies, it takes the form

   $$\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \lambda(n-2) \begin{pmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{pmatrix} \right) \begin{pmatrix} p_A \\ p_B \end{pmatrix} = \left(\delta^{-1} + \lambda(n-2) \right) \cdot \begin{pmatrix} -\alpha S - \beta \\ -\alpha P - \beta \end{pmatrix} \tag{5.7}$$

   where $\alpha = (T-P)^{-1}$. This is satisfiable (with exactly one β) if and only if condition [5.6] holds.

2. According to Press & Dyson [PD12], a memory-1 strategy is an equalizer in an infinitely repeated ($d=1$) two-player game ($n=2$) if and only if there are constants β (not identical with our β) and γ such that

   $$\begin{pmatrix} -1+p_{AA} \\ -1+p_{AB} \\ p_{BA} \\ p_{BB} \end{pmatrix} = \beta \begin{pmatrix} R \\ S \\ T \\ P \end{pmatrix} + \gamma \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \tag{5.8}$$

   This exactly corresponds to our (5.5) with $\delta=1$ and $n=2$ (or alternatively $\lambda=0$).

Based on (5.5), we can also provide an algorithm that takes the game parameters and an indirectness value λ as an input to compute an equalizer strategy that enforces the highest

Algorithm 1 Constructing a generic equalizer strategy with given indirectness λ that enforces the highest possible payoff.

function ComputeEqualizer(R, S, T, P, n, ε, λ)

$p_0 \leftarrow 1$

$\eta \leftarrow \frac{\delta(1+(n-2)\lambda)}{1+(n-2)\delta\lambda}$

$\zeta \leftarrow \frac{1+\lambda(n-2)}{1+\lambda(n-2)(1-2\varepsilon)}$

$M_\varepsilon \leftarrow \textsc{ErrorMatrix}(\varepsilon)$

$H_\varepsilon \leftarrow (1+\lambda(n-2))^{-1}(I+\lambda(n-2)M_\varepsilon)$

$(R_\varepsilon, S_\varepsilon, T_\varepsilon, P_\varepsilon)^\intercal \leftarrow H_\varepsilon^{-1}(R,S,T,P)^\intercal$

if $P_\varepsilon \neq R_\varepsilon$ **then**

  return null

$$\omega \leftarrow \max \left\{ \begin{array}{c} \max\{|R_\varepsilon - T_\varepsilon|, |S_\varepsilon - P_\varepsilon|\}/\eta, \\ (\max\{R_\varepsilon, T_\varepsilon\} - \min\{S_\varepsilon, P_\varepsilon\})/(\eta+\zeta), \\ (\min\{R_\varepsilon, T_\varepsilon\} - \min\{S_\varepsilon, P_\varepsilon\})/\zeta \end{array} \right\}$$

if $\omega > (\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\})/(\zeta-\eta)$ **then**

  return null

$\alpha \leftarrow \omega^{-1}$

$p_{AA} \leftarrow 1 - \eta^{-1}\alpha(R_\varepsilon - \min\{R_\varepsilon, T_\varepsilon\})$

$p_{AB} \leftarrow 1 - \eta^{-1}\alpha(T_\varepsilon - \min\{R_\varepsilon, T_\varepsilon\})$

$p_{BA} \leftarrow 1 - \eta^{-1}\zeta - \eta^{-1}\alpha(S_\varepsilon - \min\{R_\varepsilon, T_\varepsilon\})$

$p_{BB} \leftarrow 1 - \eta^{-1}\zeta - \eta^{-1}\alpha(P_\varepsilon - \min\{R_\varepsilon, T_\varepsilon\})$

$\pi \leftarrow \min\{R_\varepsilon, T_\varepsilon\} - (\zeta-1)/(2\alpha)$

return p_0, p_{AA}, p_{AB}, p_{BA}, p_{BB}, λ, π

payoff (among all equalizers with indirectness λ). It returns the strategy parameters as well as the payoff π that this strategy enforces, or **null** if no such equalizer strategy exists. The correctness of this Algorithm 1 is shown in Supplementary Information Theorem 6.

Figure parameters. Fig. 5.5: Each of $N = 100$ points represents the payoffs of a randomly generated mutant strategy against $n-1$ identical equalizer residents in a game with $n = 50$ players. Prisoner's dilemma: $R = 3$, $S = 0$, $T = 5$, $P = 1$. Stag hunt: $R = 3$, $S = 0$, $T = 2$, $P = 1$. Hawk-dove: $R = 2$, $S = 0$, $T = 4$, $P = -2$. $\varepsilon = 10^{-3}$, $\delta = 0.99$. Each point is an average over $5 \cdot 10^5$ samples with the same mutant. Equalizer strategies were generated with an implementation of Algorithm 1.

5.4 Supplementary Information

5.4.1 Model

The model we use is a generalisation of the model of Schmid et al. [SCHN21] from the donation game to general 2×2 matrix games. Like the model of [SCHN21], it is a unified framework of direct and indirect reciprocity. This lets us meaningfully compare the two mechanisms and also allows us to study strategies that combine both direct and indirect reciprocity.

Game model

Game matrix. We study a game with n players that is played in a sequence of rounds. In each round, a disjoint pair (i, j) of players is drawn uniformly at random. These two players then play a 2×2 matrix game G against each other, where G is a fixed matrix

$$G = \begin{pmatrix} R & S \\ T & P \end{pmatrix}.$$

We require $R \geq P$ throughout without loss of generality.

We use A and B as the indices for the rows and columns of G, where $\mathcal{A} = \{A, B\}$ is the set of actions that each player can choose from in a round in which they are choosen to play.

We also use the entries of G arranged in a 4-vector, in accordance with the below definitions.

Definition 13 (Payoff vector)**.** *Given a game G, we define the payoff vector g as*

$$g = \begin{pmatrix} G_{AA} \\ G_{AB} \\ G_{BA} \\ G_{BB} \end{pmatrix} = \begin{pmatrix} R \\ S \\ T \\ P \end{pmatrix}.$$

Definition 14 (Opponent's payoff vector)**.** *The opponent's payoff vector g_{op} is defined as*

$$g_{\text{op}} = \begin{pmatrix} G_{AA} \\ G_{BA} \\ G_{AB} \\ G_{BB} \end{pmatrix} = \begin{pmatrix} R \\ T \\ S \\ P \end{pmatrix}.$$

One-round payoff. We denote by e_A and e_B the standard basis vectors of $\mathbb{R}^{\mathcal{A}}$, i.e., $(1, 0)^\intercal$ and $(0, 1)^\intercal$, respectively. Let i and j be the players chosen to play in a given round t. Let $x \in \mathcal{A}$ be Player i's action and $y \in \mathcal{A}$ be Player j's action. Then Player i's payoff in that round, which we denote as $\Pi_i(t)$, is

$$\Pi_i(t) = e_x^\intercal G e_y.$$

Player j's payoff is $\Pi_j(t) = e_y^\intercal G e_x$. Players who are not drawn receive a payoff of 0 in that round.

Game continuation. After each round, it is randomly decided whether another round will be played, with a fixed probability d for the game to continue for at least one more round. With probability $1 - d$, the game ends. The probability d must satisfy $0 < d < 1$. Round 0 is always played. So the probability for all $t \geq 0$ that round t is played is d^t.

We define

$$\delta = \frac{2d}{2d + (n-1)n(1-d)}. \tag{5.9}$$

Then at any given point in time, for any pair (i, j) of players, δ is the probability that these two players will interact at least once in the future. The inverse of the relationship between δ and d can be expressed as

$$d = \frac{\delta n(n-1)}{2(1-\delta) + \delta n(n-1)}. \tag{5.10}$$

Total payoff. Player i's total payoff for the whole game is defined as

$$\Pi_i = (1-d)\frac{n}{2}\sum_{t=0}^{\infty} \Pi_i(t).$$

Because we allow player's strategies to be stochastic, $\Pi_i(t)$ is a stochastic quantity. We denote the expected payoff of Player i in round t, given that the game is played at least until round t, by $\pi_i(t) = \mathbb{E}\left[\Pi_i(t)\right]$. Then the expected total payoff of Player i is

$$\pi_i = \mathbb{E}\left[\Pi_i\right] = (1-d)\frac{n}{2}\sum_{t=0}^{\infty} d^t \pi_i(t). \tag{5.11}$$

The normalisation factor $(1-d)n/2$ is chosen so that the expected total payoff takes values in the same range as the one-round payoffs.

However, to avoid having to uneccessarily deal with the stochastic uncertainty of whether or not a given round t will be played when analysing $\pi_i(t)$, we henceforth consider a completely equivalent alternative setup where an infinite sequence of rounds is always played, and the payoffs from round t contribute to the total payoff with a weight of d^t. Naturally, Eq. (5.11) retains its validity.

Interaction and error probability. Players observe perfectly who is chosen in each round, and, when they are chosen to play, the action of their opponent. However, in the rounds that they are not involved in, each observed action is subject to errors at a rate of ε, where $0 \leq \varepsilon < 1/2$, independently of the other action.

For example, if two other players interacted and played the action pair (B, A), a third player will correctly observe (B, A) with probability $(1-\varepsilon)^2$, but make a wrong observation of (A, A) with probability $(1-\varepsilon)\varepsilon$, make a wrong observation of (B, B) also with probability $(1-\varepsilon)\varepsilon$, and wrongly observe both actions, i.e. make an observation of (A, B), with probability ε^2.

Together, the four parameters n, G, d, and ε specify a game in our model.

Game classification

We introduce some well-studied subclasses of the 2×2 matrix games.

(a) Games with $T > R > P > S$ and $2R > T + S$ are called prisoner's dilemmas.

(b) Prisoner's dilemmas that additionally satisfy $R + P = S + T$ are called donation games. W.l.o.g., these can also be expressed in terms of a benefit b and cost c of cooperation, where $b > c > 0$, as

$$G = \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix}.$$

(c) Games with $T > R > S > P$ are called hawk-dove games.

(d) Games with $R > T > P > S$ are called stag hunt games.

All of the above are examples of social dilemmas, which are defined as follows:

Definition 15 (Social dilemma)**.** *A game is a social dilemma if $R > S$, $R > P$, $T > S$ and $T > P$, i.e., if*

$$\min\{R, T\} > \max\{S, P\}.$$

Social dilemmas are of particular interest to us, because they feature a conflict between the collective and private preference.

The condition that distinguishes donation games from the more general case of prisoner's dilemmas can be stated generally as follows:

Definition 16 (Additive games)**.** *A game is additive if it satisfies $R + P = S + T$.*

With that definition, donation games are exactly the additive prisoner's dilemmas.

Given stochastic vectors $v, w \in [0, 1]^{\mathcal{A}}$ representing the probability distribution functions for the actions of two players who are chosen to play against each other in a given round, the payoff of the first player is given by the bilinear form $(v, w) \mapsto v^\intercal G w$. In the special case that $R + P = S + T$, it can be separated into two summands of which each only depends on one of the players' actions:

$$v^\intercal G w = v^\intercal \begin{pmatrix} S - P/2 \\ P/2 \end{pmatrix} + w^\intercal \begin{pmatrix} T - P/2 \\ P/2 \end{pmatrix}.$$

That is why we call the above property "additivity". An additive game allows for easier analysis, since we can exploit the linearity in (v, w) of the payoff function: A player's total payoff is a function of the totality of actions taken and the totality of actions received. In a non-additive game, where the payoff function is merely bilinear, we have to calculate payoffs per round. Figure 5.3 classifies the symmetric 2×2 matrix games. As is apparent from the figure, among the the three main game classes of interest, only prisoner's dilemmas can be additive.

Strategy spaces

We consider reciprocal strategies, where players remember past actions of their opponents and base their actions on the observations they remember. In our framework, a player playing a strategy of reciprocity remembers exactly one past action or interaction of each opponent. The past action or interaction that they remember is called the reputation that they assign to the given opponent. This reputation is private and not a priori known to the other players.

Reactive and Memory-1 strategies. We consider two classes of reciprocal strategies:

- Reactive strategies, which assign to each opponent a reputation consising of one past action of that opponent. In reactive strategies, reputations are thus elements of $\mathcal{A}$.

- Memory-1 strategies, which assign to each opponent a reputation consising of one past interaction of the opponent, meaning an action of the opponent together with the action of the player the opponent was interacting with. In memory-1 strategies, reputations are thus elements of $\mathcal{A}^2$.

We represent memory-1 reputations as tuples $(x, y) \in \mathcal{A}^2$. The second component, y, is always the action of the player whom the reputation is assigned to, whereas the first component, x, is the action of the player that the opponent was interacting with in the remembered interaction, which could be the focal memory-1 player themselves.

The strategy parameters of both reactive and memory-1 strategies include one probability value for each possible reputation (that is, two in the case of reactive strategies and four in the case of memory-1 strategies). When a player is chosen to interact with an opponent, the probability value in their strategy corresponding to the opponent's current reputation is the probability that they will play A.

Throughout the game, players update the reputations they assign to the other players based on their observations. When they are themselves chosen to play, we say that they directly observe the interaction. In such a round, they always update the reputation that they assign to their coplayer. When they are not chosen to play, we say that they make an indirect observation. In this case, they may or may not update the reputations they assign to the two involved players. A further strategy parameter, the indirectness λ, determines the probability for each of the two reputations to be updated. Traditionally, the extreme cases of direct and indirect reciprocity have been studied:

- A strategy of direct reciprocity is one that only takes into account directly observed actions. This corresponds to the case $\lambda = 0$ in our model.
- A strategy of indirect reciprocity equally takes into account direct and indirect observations. This corresponds to the case $\lambda = 1$ in our model.

What follows are a formal definition for reactive and memory-1 strategies.

Definition 17 (Reactive strategy)**.** *A reactive strategy is defined by the following parameters:*

1. *the initial action probability* $p_0 \in [0, 1]$
2. *the continuation vector* $p = (p_A, p_B) \in [0, 1]^{\mathcal{A}}$
3. *the indirectness probability* $\lambda \in [0, 1]$

A player employing the reactive strategy (p_0, p, λ) *acts as follows: The player privately assigns a reputation to each other player, whose value is updated throughout the game. Initially, no value is set for the reputation. Once it is set, the possible values for a reputation are* A *and* B*, which are the two possible actions in the one-round matrix game. Whenever a player* j *plays against the reactive player, the reputation of Player* j *will be updated to the action that Player* j *used. Whenever a player* j *plays against a third player* k*, with a probability of* λ *the reactive player will update* j*'s reputation to the observed action of* j *towards* k *(which might be observed wrongly due to errors, as described in the* *model), and with a probability of* $1-\lambda$ *will not update* j*'s reputation. The reactive player* i*, when chosen to play against a player* j*, will use the value of their register for Player* j *to choose their action. If the register is still empty,* i *plays* A *with probability* p_0*. If the register holds the reputation value* $x \in \mathcal{A}$*, then* i *will play* A *with probability* p_x*.*

Definition 18 (Memory-1 strategy)**.** *A memory-1 strategy is specified by the following parameters:*

1. *the initial action probability* $p_0 \in [0,1]$
2. *the continuation vector* $p = (p_{AA}, p_{AB}, p_{BA}, p_{BB}) \in [0,1]^{\mathcal{A}^2}$
3. *the indirectness probability* $\lambda \in [0,1]$

In the same way as with reactive strategies, the parameter λ *determines when registers are updated. Unlike with reactive strategies, the reputations are values in* $\mathcal{A}^2$*. So when the reputation assigned by Player* i *to Player* j *is* $(x,y) \in \mathcal{A}^2$*, this means that when the reputation was last updated, it was because Player* i *observed Player* j *play action* y *towards some player (which may have been Player* i *themselves) who simultaneously played action* x*. When a memory-1 player is choosen to play, the continuation vector* $p \in [0,1]^{\mathcal{A}^2}$ *determines their probability of playing* A *in the sense that if* $(x,y) \in \mathcal{A}^2$ *is the reputation assigned to the given opponent, then the memory-1 player plays* A *with probability* p_{xy}*.*

Reactive strategies are a special case of memory-1 strategies, where $p_{AA} = p_{BA}$ and $p_{AB} = p_{BB}$.

In the main text, instead of listing the parameters of a reactive or memory-1 strategy as (p_0, p, λ), we unwrap the vector p and list the parameters as (p_0, p_A, p_B, λ) for reactive strategies, and as $(p_0, p_{AA}, p_{AB}, p_{BA}, p_{BB}, \lambda)$ for memory-1 strategies.

Nash equilibria and equalizer strategies

We say a player i acts rationally if they play in such a way that they maximise their expected total payoff π_i. A Nash equilibrium strategy is one that is apt to be adopted by an entire population of rational players, in accordance with the following definition.

Definition 19 (Nash equilibrium [Nas50])**.** *A strategy* σ *is a Nash equilibrium strategy if, when all players employ* σ *as their strategy, no player can increase their expected payoff by unilaterally changing to a different strategy.*

The equalizer strategies, defined below, are a subclass of the Nash equilibrium strategies, as is easy to see from the definition. By constructing equalizer strategies, we can show the existence of Nash equilibria for a certain scenario.

Definition 20 (Equalizer strategy [PD12])**.** *A strategy* σ *is an equalizer strategy if, when all players except for one (called the mutant) employ* σ *as their strategy, then the mutant's expected payoff is independent of the mutant's strategy.*

5.4.2 Existing results

Two existing results from the literature are particularly relevant to our work.

The first is Press and Dyson's [PD12] result regarding the existence of equalizer strategies in the prisoner's dilemma. The model used in [PD12] is slighly different from ours, because they consider an infinitely repeated game (intuitively equivalent to $d = 1$ in our model, which is a case that we exclude). Other than that, the models are compatible. However, since their result concerns only two-player games, there are no indirect observations and thus also no observation errors. Thus, there is no need for the game parameter ε as it features in our model, nor for the indirectness λ as a strategy parameter. Because the game is infinitely repeated, payoffs are defined in terms of the limiting behaviour of the game, which is why the strategy parameter p_0 is also not needed. So in the model of [PD12], a game is given purely by the payoff matrix G, and a strategy purely by the continuation vector p.

Theorem 11 (Press and Dyson [PD12])**.** *Every infinitely repeated prisoner's dilemma among two players contains a memory-1 equalizer strategy of direct reciprocity.*

Regarding infinitely repeated games that are not prisoner's dilemmas, Press and Dyson also show that equalizer strategies exist in some cases.

Secondly, Schmid et al. [SCHN21] show the existence of equalizer strategies in donation games in a model that is exactly equivalent to a special case of ours, namely, its restriction to the donation game.

Theorem 12 (Schmid et al. [SCHN21])**.** *In all donation games, for all sufficiently large d, there exists a reactive equalizer strategy. More precisely, a reactive equalizer strategy with indirectness λ exists if and only if*

$$\frac{1 + (n-2)\delta\lambda}{1 + (n-2)(1-\varepsilon)\lambda} \cdot \frac{c}{\delta b} < 1. \tag{5.12}$$

Theorem 12 easily generalises from donation games to general additive games.

The results of [PD12] are about equalizer strategies in the context of direct reciprocity for a general payoff matrix, whereas the results of [SCHN21] are about Nash equilibria and equalizer strategies for a unified framework of direct and indirect reciprocity, but for the special class of donation games only. We present results for a general payoff matrix in the unified framework.

5.4.3 Analysis of reactive strategies

In this section, we aim to analyse the space of reactive Nash equilibria as comprehensively as possible, before later moving on to the more powerful memory-1 strategies in the subsequent section. We present separate results for additive and non-additive games. In additive games, we offer a complete description of all reactive Nash equilibria, which is a trivial generalisation from Schmid et al. [SCHN21]. In non-additive games, we show by a novel proof that, excluding edge cases of games and strategies, reactive Nash equilibria do not exist.

Reactive Nash equilibria in additive games

If all players (the residents) except for one (the mutant) play a reactive strategy, then the rate at which the mutant receives action A from the residents depends linearly on the mutant's own rate of playing action A towards the residents. Since the payoff equation is also linear, one may express the mutant's payoff as a linear function of the mutant's rate of playing action A (the mutant's A-rate), which is done in Lemma 4 below.

This lemma is a generalisation of Lemma 3 of [SCHN21] that is trivially obtained by dropping the mostly inconsequential assumption of $T > R > P$, which restricts additive games to the donation game. Compare also Eq. (71) in the proof of Lemma 3 of [SCHN21] for the derivation of our (5.13).

Lemma 4 (Schmid et al. [SCHN21])**.** *Consider a population where players but one, the mutant, apply the reactive strategy* (p_0, p, λ)*. Let* x *be the mutant's rate of playing action* A*. Then the mutant's payoff is equal to*

$$\pi = C_1((p_A - p_B)(T - P) - c_\lambda(P - S))x + C_2(T - P) + P, \tag{5.13}$$

where C_1*,* $C_2 > 0$ *are positive constants that depend only on the resident strategy, and*

$$c_\lambda = \delta^{-1}\frac{1 + (n-2)\delta\lambda}{1 + (n-2)(1-2\varepsilon)\lambda}. \tag{5.14}$$

Given that a mutant's payoff only depends on the mutant's A-rate, and given that it does so by a linear relationship, there are only three possible cases of what the mutant's best response can be, depending on the coefficient of x in Eq. 5.13. Either the only best response is to always play A, or to always play B, or, if the coefficient is 0, any strategy is a best response for the mutant. The latter is the case exactly when the resident strategy is an equalizer strategy.

We can use this to characterise all reactive Nash equilibria in additive games. The below Proposition 7 is analogous to Theorem 1 of [SCHN21].

Proposition 7. *In an additive game, the reactive Nash equilibria are exactly the following strategies, with* c_λ *defined as in (5.14).*

- *The equalizer strategies. They are*
  - *all strategies such that*
    $$p_A - p_B = c_\lambda\frac{P - S}{T - P}$$
    when $T \neq P$*,*
  - *and all strategies when* $R = S = T = P$*.*
- *Those Nash equilibria that are not equalizer strategies and always play action* A*. They are*
  - *all strategies with* $p_0 = p_A = p_B = 1$ *when* $P < S$
  - *and all strategies with* $p_0 = p_A = 1$ *and* $(1 - p_B)(T - P) > c_\lambda(P - S)$ *and either of* $\lambda = 0$ *or* $n = 2$ *or* $\varepsilon = 0$*.*

- *Those Nash equilibria that are not equalizer strategies and always play action B. They are*
  - *all strategies with $p_0 = p_A = p_B = 0$ when $P > S$*
  - *and all strategies with $p_0 = p_B = 0$ and $p_A(T-P) < c_\lambda(P-S)$ and either of $\lambda = 0$ or $n = 2$ or $\varepsilon = 0$.*

Proof. We consider the sign of the coefficient of x in (5.13). It is the sign of

$$(p_A - p_B)(T-P) - c_\lambda(P-S). \tag{5.15}$$

If the value of this expression is positive, the best-response mutant strategies are exactly those with $x = 1$, so the resident strategy is a Nash equilibrium exactly if it always plays action A against itself in a homogenous population. Similarly, if (5.15) is negative, the strategy is a Nash equilibrium exactly if it always plays action B against itself. The strategy is an equalizer strategy exactly if the value of (5.15) is 0.

It is shown in Lemma 2 of [SCHN21] that a strategy (p_0, p, λ) always plays action A against itself exactly if $p_0 = p_A = p_B = 1$ or $p_0 = p_A = 1$ and either $\lambda = 0$ or $n = 2$ or $\varepsilon = 0$. Similarly, it always plays action B against itself exactly if $p_0 = p_A = p_B = 0$ or $p_0 = p_B = 0$ and either $\lambda = 0$ or $n = 2$ or $\varepsilon = 0$.

The rest follows simply from examining the sign of (5.15) in the various cases. □

In Section 5.4.4, we analyse memory-1 strategies. Due to their greater complexity, it is not possible to characterise all Nash equilibria, but we present existence and non-existence results for equalizer strategies, which are an important subclass. In order to compare those results to the case of reactive strategies, we now derive existence criteria for reactive equaliser strategies from Proposition 7.

Corollary 15. *In an additive game, reactive equalizer strategies with indirectness λ exist exactly if (a) $T \neq P$ and*

$$\delta^{-1} \frac{1 + (n-2)\delta\lambda}{1 + (n-2)(1-2\varepsilon)\lambda} \cdot \frac{P-S}{T-P} \in [-1, 1],$$

or (b) $R = S = T = P$.

Corollary 16. *In an additive game, reactive equalizer strategies exist exactly if (a) $T \neq P$ and*

$$\delta^{-1} \frac{|S-P|}{|T-P|} \leq 1 \qquad \text{or} \qquad \delta^{-1} \frac{1 + (n-2)\delta}{1 + (n-2)(1-2\varepsilon)} \cdot \frac{|S-P|}{|T-P|} \leq 1,$$

or (b) $R = S = T = P$.

Corollary 17. *In an additive game, reactive equalizer strategies exist for sufficiently large δ exactly if the game is either a social dilemma or satisfies $R = S = T = P$.*

Proof. We use the characterisation of equalizer strategies given in Corollary 16.

Firstly, we treat the case $T = P$. Reactive equalizer strategies exist exactly if $R = S = T = P$.

Secondly, we treat the case $T \neq P$. Assume that a reactive equalizer strategy exists. One of the two inequalities in Corollary 16 must hold. Both of them imply $|S-P| < |T-P|$, which is therefore true. The general assumption $P \leq R$ is equivalent to $0 \leq (T-P) + (S-P)$ in an additive game. Taking these two inequalities together, we get $0 < T - P$ and $0 < (T-P) + (S-P)$, or alternatively $P < T$ and $P < R$. This is sufficient for an additive game to be a social dilemma.

Now assume instead that the game is a social dilemma. Then $S < T$ and $P < R$ respectively imply that $S - P < T - P$ and $P - S < T - P$. Using $T - P > 0$, we can write this as $|S - P| < |T - P|$. Again using $T - P > 0$, this means that $\frac{|S-P|}{|T-P|} \leq \delta$ for some sufficiently large $\delta < 1$, and that $T \neq P$. So by Corollary 16, the reactive equalizer strategies exist for sufficiently large δ.

We have thus shown the equivalence of the game being a social dilemma and the existence of reactive equalizer strategies for sufficiently large δ in the case of $T \neq P$. □

Reactive Nash equilibria in non-additive games

In contrast to the case of additive games, Nash equilibria cannot exist in non-additive games other than in special cases where they are not affected by errors.

The most obvious way in which a Nash equilibrium can be unaffected by errors is if the game has an error rate of zero ($\varepsilon = 0$). Another way is for the game to only have two players ($n = 2$), since in a two-player game, all observations are direct, and errors by definition only occur in indirect observations. The following definition captures this distinction between games where errors can and cannot occur.

Definition 21 (Game with errors). *We say a game given by the parameters n, G, and ε is a game with errors if $n > 2$ and $\varepsilon > 0$.*

But even in a game with errors, a particular reactive strategy can still be unaffected by errors. This is certainly the case if the strategy is a strategy of direct reciprocity ($\lambda = 0$), since errors only affect indirect observations, which this strategy disregards. It is also the case if the strategy is memoryless ($p_A = p_B$), in which case it disregards both direct and indirect observations. The following definition expresses the concept of a reactive strategy that can be affected by errors.

Definition 22 (Error-prone reactive strategy). *We call a reactive strategy (p_0, p, λ) error-prone if it satisfies $p_A \neq p_B$ and $\lambda > 0$.*

With these definitions, we can formally state our result about the non-existence of Nash equilibria that are affected by errors.

Theorem 13. *In non-additive games with errors, no error-prone reactive strategy is a Nash equilibrium strategy.*

The proof of this theorem is found in Appendix Section 5.4.5.

As we can see from the proof, it is not the distorting effect of errors per se that prevents the existence of Nash equilibria. Rather, the function of errors in the proof is merely to ensure that the game does not follow along a predetermined path, but that instead a sufficient variety of situations arises with positive probability. Once that is the case, non-reactive strategies can outperform any reactive resident strategy in an additive game by using their larger memory. We therefore expect that future research will be able to further restrict the conditions under which reactive Nash equilibria can exist in non-additive games.

5.4.4 Analysis of memory-1 equalizer strategies

We have seen with Theorem 13 that reactive strategies only contain a very restricted set of Nash equilibria other than in additive games. In this section, we analyse the more general class of memory-1 strategies and find that they do contain a rich set of Nash equilibria also for non-additive games. We show this by exactly characterising the set of memory-1 equalizer strategies. We see that they exist in all classes of social dilemmas, at least when δ is sufficiently large and ε is sufficiently small, but not necessarily 0. In many social dilemmas, including all where unconditional cooperation is not an equilibrium, memory-1 equalizers can achieve full cooperation.

Definitions

We first define some concepts and notation that we will use throughout Section 5.4.4.

Firstly, we define some quantities that may depend on the game parameters but not on a specific strategy. Recall the definitions of the payoff vector $g = (R, S, T, P)^\intercal \in \mathbb{R}^{\mathcal{A}^2}$ and the opponent's payoff vector $g_{\text{op}} = (R, T, S, P)^\intercal \in \mathbb{R}^{\mathcal{A}^2}$. We write $1_{\mathcal{A}^2} = (1, 1, 1, 1)^\intercal$ for the vector of ones in $\mathbb{R}^{\mathcal{A}^2}$. Furthermore, we write e_{A*} for the vector

$$e_{A*} \coloneqq e_{AA} + e_{AB} = (1, 1, 0, 0)^\intercal.$$

We write $I_{\mathcal{A}^2} \in \mathcal{M}_{\mathcal{A}^2}(\mathbb{R})$ for the identity matrix of $\mathbb{R}^{\mathcal{A}^2}$.

We write $M_\varepsilon \in \mathcal{M}_{\mathcal{A}^2}(\mathbb{R})$ for the error matrix

$$M_\varepsilon \coloneqq \begin{pmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{pmatrix}^{\otimes 2} = \begin{pmatrix} (1-\varepsilon)^2 & (1-\varepsilon)\varepsilon & \varepsilon(1-\varepsilon) & \varepsilon^2 \\ (1-\varepsilon)\varepsilon & (1-\varepsilon)^2 & \varepsilon^2 & \varepsilon(1-\varepsilon) \\ \varepsilon(1-\varepsilon) & \varepsilon^2 & (1-\varepsilon)^2 & (1-\varepsilon)\varepsilon \\ \varepsilon^2 & \varepsilon(1-\varepsilon) & (1-\varepsilon)\varepsilon & (1-\varepsilon)^2 \end{pmatrix}$$

for observation vectors. If the stochastic vector $x \in \mathbb{R}^{\mathcal{A}^2}$ is the probability distribution over the outcome of an interaction between a player pair, then $M_\varepsilon x$ is the distribution over the error-prone observation of this interaction by a third player.

The following quantities depend on the indirectness λ of a given memory-1 strategy (p_0, p, λ). For a given strategy, we define the auxiliary quantities $\eta \in \mathbb{R}$ and $\zeta \in \mathbb{R}$ as

$$\eta \coloneqq \delta \frac{1 + \lambda(n-2)}{1 + \delta\lambda(n-2)} \tag{5.16}$$

and

$$\zeta \coloneqq \frac{1 + \lambda(n-2)}{1 + (1-2\varepsilon)\lambda(n-2)}. \tag{5.17}$$

We always have $\delta < \eta < 1$ and thus $\eta \uparrow 1$ as $\delta \to 1$. We always have $\zeta \geq 1$, with equality exactly if $n = 2$ or $\varepsilon = 0$ or $\lambda = 0$. That is, $\zeta = 1$ exactly if the strategy (p_0, p, λ) cannot make errors with the updating of reputations.

We define the error-adjustment matrix $H_\varepsilon \in \mathcal{M}_{\mathcal{A}^2}(\mathbb{R})$ as

$$H_\varepsilon := (1 + \lambda(n-2))^{-1}(I_{\mathcal{A}^2} + \lambda(n-2)M_\varepsilon). \tag{5.18}$$

We further define the error-adjusted payoff values $R_\varepsilon, S_\varepsilon, T_\varepsilon, P_\varepsilon \in \mathbb{R}$ by

$$(R_\varepsilon, S_\varepsilon, T_\varepsilon, P_\varepsilon)^\intercal := H_\varepsilon^{-1}(R, S, T, P)^\intercal. \tag{5.19}$$

If $\varepsilon = 0$, then $H_\varepsilon = I_{\mathcal{A}^2}$ and $R_\varepsilon = R$, $S_\varepsilon = S$, $T_\varepsilon = T$, $P_\varepsilon = P$.

Finally, we will use the following central definition:

Definition 23 (Press-Dyson vector)**.** *Given a strategy (p_0, p, λ), we define its Press-Dyson vector $\tilde{p}$ as*

$$\tilde{p} = e_{A*} - \eta H_\varepsilon p. \tag{5.20}$$

Note that $\tilde{p}$ is independent of p_0.

Characterisation of memory-1 equalizer strategies

In this section, we derive a complete characterisation of the memory-1 equalizer strategies for all games.

In other models, a useful result for the construction of equalizer strategies is Akin's Lemma [Aki16]. We also show a version of Akin's Lemma for this model, and use it to show that all strategies included in our characterisation are indeed equalizer strategies. The other direction will be shown in a different way.

Lemma 5 (Akin's Lemma)**.** *Consider a game of $n-1$ residents playing the memory-1 strategy (p_0, p, λ) and one mutant playing an arbitrary strategy.*

Let for all $t \geq 0$ the vector $v(t) \in \mathbb{R}^{\mathcal{A}^2}$ be the mutant's outcome of round t, but from the view of the mutant's opponents, in the sense that for all $x, y \in \mathcal{A}$, the value of v_{xy} is 1 if the mutant player was selected in round t, played y, and their opponent played x, and is 0 otherwise. Let

$$v := (1-d)\sum_{t=0}^{\infty} d^t v(t).$$

Then we have

$$\tilde{p}^\intercal \mathbb{E}v = (1-\eta)\frac{2}{n}p_0. \tag{5.21}$$

Proof. Let π be the overall expected payoff of the mutant. Then

$$\pi = g_{\text{op}}^\intercal \mathbb{E}v = (1-d)g_{\text{op}}^\intercal \sum_{t=0}^{\infty} d^t \mathbb{E}v(t).$$

Write $x(t) \in \mathcal{P}(\mathcal{A}^2)$ for the probability distribution over the action pairs in round t under the condition that the mutant player was selected, again with the mutant's action in the second component. Then

$$\mathbb{E}v(t) = \frac{2}{n}x(t),$$

so

$$\pi = (1-d)\frac{2}{n} g_{\text{op}}^{\intercal} \sum_{t=0}^{\infty} d^t x(t).$$

We also write

$$x := (1-d) \sum_{t=0}^{\infty} d^t x(t).$$

Let r be the probability that in a given round, a given opponent does not update the reputation they assign to the mutant. This can be the case either because the mutant did not play in that round, or because the mutant played against a third player and the given opponent did not observe the interaction. So

$$r = \frac{n-2}{n} + (1-\lambda)\frac{2(n-2)}{n(n-1)} = \frac{(n-2)(n+1-2\lambda)}{n(n-1)}.$$

For any $\tau \geq 1$, let $r_{\text{dir}}(\tau)$ be the probability that at a given time t, where $t > \tau$, under the condition that the mutant and a given opponent are playing in round t, the reputation that the opponent assigns to the mutant stems from a directly observed interaction (i.e. an interaction between the same two players) in round $t-\tau$. This is equivalent to saying that in round $t-\tau$, the two players were selected, and in every subsequent round the opponent did not update reputation they assign to the mutant. These are τ independent events, and so

$$r_{\text{dir}}(\tau) = \frac{2}{n(n-1)} r^{\tau-1}.$$

Similarly, for any $\tau \geq 1$, let $r_{\text{ind}}(\tau)$ be the probability that, again under the condition that the mutant and a given opponent are playing in round $t > \tau$, the reputation that the opponent assigns to the mutant stems from an indirectly observed interaction in round $t-\tau$. Then

$$r_{\text{ind}}(\tau) = \lambda \frac{2(n-2)}{n(n-1)} r^{\tau-1}.$$

Finally, for any $t \geq 0$, let $r_{\text{init}}(t)$ be the probability that, at time t, a given opponent does not yet have a reputation assigned to the mutant. We have

$$r_{\text{init}}(t) = r^t.$$

Using these definitions, we can write

$$e_{A*}^{\intercal} x(t) = p^{\intercal} \left(\sum_{\tau=1}^{t} r_{\text{dir}}(\tau) x(t-\tau) + \sum_{\tau=1}^{t} r_{\text{ind}}(\tau) M_{\varepsilon} x(t-\tau) \right) + r^t p_0,$$

Applying $(1-d)\sum_{t=0}^{\infty} d^t$, we get

$$\begin{aligned}
e_{A*}^{\intercal} x &= (1-d)\left(p^{\intercal} \sum_{t=0}^{\infty} d^t \sum_{\tau=1}^{t} r_{\text{dir}}(\tau) x(t-\tau) + p^{\intercal} \sum_{t=0}^{\infty} d^t \sum_{\tau=1}^{t} r_{\text{ind}}(\tau) M_{\varepsilon} x(t-\tau) + \sum_{t=0}^{\infty} d^t r^t p_0 \right) \\
&= (1-d)\left(p^{\intercal} \sum_{t=0}^{\infty} d^t x(t) \sum_{\tau=1}^{\infty} d^{\tau} r_{\text{dir}}(\tau) + p^{\intercal} M_{\varepsilon} \sum_{t=0}^{\infty} d^t x(t) \sum_{\tau=1}^{\infty} d^{\tau} r_{\text{ind}}(\tau) + \sum_{t=0}^{\infty} d^t r^t p_0 \right) \\
&= p^{\intercal} x \sum_{\tau=1}^{\infty} d^{\tau} r_{\text{dir}}(\tau) + p^{\intercal} M_{\varepsilon} x \sum_{\tau=1}^{\infty} d^{\tau} r_{\text{ind}}(\tau) + (1-d) p_0 \sum_{t=0}^{\infty} d^t r^t \\
&= \left(p^{\intercal} x \cdot d \frac{2}{n(n-1)} + p^{\intercal} M_{\varepsilon} x \cdot d\lambda \frac{2(n-2)}{n(n-1)} + (1-d) p_0 \right) \sum_{t=0}^{\infty} d^t r^t.
\end{aligned}$$

We calculate

$$\sum_{t=0}^{\infty} d^t r^t = \sum_{t=0}^{\infty} \left(d \frac{(n-2)(n+1-2\lambda)}{n(n-1)} \right)^t = \frac{n(n-1)}{n(n-1) - d(n-2)(n+1-2\lambda)}.$$

So

$$\begin{aligned} e_{A*}^{\intercal} x = {} & \frac{n(n-1)}{n(n-1) - d(n-2)(n+1-2\lambda)} p^{\intercal} \left(d \frac{2}{n(n-1)} I_{\mathcal{A}^2} + d\lambda \frac{2(n-2)}{n(n-1)} M_{\varepsilon} \right) x \\ & + (1-d) \frac{n(n-1)}{n(n-1) - d(n-2)(n+1-2\lambda)} p_0 \\ = {} & \frac{2d}{n(n-1) - d(n-2)(n+1-2\lambda)} p^{\intercal} \left(I_{\mathcal{A}^2} + \lambda(n-2) M_{\varepsilon} \right) x \\ & + (1-d) \frac{n(n-1)}{n(n-1) - d(n-2)(n+1-2\lambda)} p_0. \end{aligned}$$

Using (5.10) to substitute for d an expression in terms of δ, we get

$$\frac{2d}{n(n-1) - d(n-2)(n+1-2\lambda)} = \delta(1 + \delta\lambda(n-2))^{-1}.$$

Using the same substitution, as well as the definition of η in (5.16), we also get

$$(1-d) \frac{n(n-1)}{n(n-1) - d(n-2)(n+1-2\lambda)} = 1 - \eta.$$

Taken together, we can write

$$e_{A*}^{\intercal} x = \delta(1 + \delta\lambda(n-2))^{-1} p^{\intercal} \left(I_{\mathcal{A}^2} + \lambda(n-2) M_{\varepsilon} \right) x + (1-\eta) p_0.$$

By (5.16) and (5.18), this is equivalent to

$$e_{A*}^{\intercal} x = \eta p^{\intercal} H_{\varepsilon} x + (1-\eta) p_0.$$

We rearrange this to

$$(e_{A*}^{\intercal} - \eta p^{\intercal} H_{\varepsilon}) x = (1-\eta) p_0.$$

Since H_{ε} is symmetric, we have $e_{A*}^{\intercal} - \eta p^{\intercal} H_{\varepsilon} = (e_{A*} - \eta H_{\varepsilon} p)^{\intercal} = \tilde{p}^{\intercal}$. So

$$\tilde{p}^{\intercal} x = (1-\eta) p_0.$$

This is equivalent to

$$\tilde{p}^{\intercal} \mathbb{E} v = (1-\eta) \frac{2}{n} p_0. \tag{5.22}$$

□

With Lemma 5, it is simple to construct some memory-1 equalizer strategies. That is what we do in the following proposition. We call these equalizer strategies the generic equalizers. Others also exist, but those are edge cases that can be dealt with directly.

Proposition 8. *If a memory-1 strategy* (p_0, p, λ) *is such that* $\tilde{p} = \alpha g_{\text{op}} + \beta 1_{\mathcal{A}^2}$ *for some* $\alpha, \beta \in \mathbb{R}$*, then it is an equalizer strategy. Necessarily* $\alpha > 0$*. The payoff that the equalizer strategy enforces is*

$$\pi = \alpha^{-1}((1-\eta)p_0 - \beta). \tag{5.23}$$

Proof. Take (p_0, p, λ) such that $\tilde{p} = \alpha g_{\text{op}} + \beta 1_{\mathcal{A}^2}$ for some $\alpha, \beta \in \mathbb{R}$. We can check that indeed $\alpha > 0$, since otherwise p could not be a vector of probabilities. With v defined as in Lemma 5, we have

$$\tilde{p}^\intercal \mathbb{E}v = \alpha g_{\text{op}}^\intercal \mathbb{E}v + \beta 1_{\mathcal{A}^2}^\intercal \mathbb{E}v = \alpha \frac{2}{n}\pi + \beta \frac{2}{n}.$$

So, using Lemma 5,

$$\alpha \frac{2}{n}\pi + \beta \frac{2}{n} = (1-\eta)\frac{2}{n}p_0,$$

which is equivalent to

$$\pi = \alpha^{-1}\left((1-\eta)p_0 - \beta\right).$$

So (p_0, p, λ) is indeed an equalizer strategy that enforces a payoff of π. □

Next, we will show the converse direction: if a strategy is an equalizer strategy, then it is of the given form. But for now, we exclude some edge cases, which we will deal with separately later. We call these edge cases the effectively unconditional strategies.

Definition 24 (Effectively unconditional strategy)**.** *For a given strategy* σ*, let* $n-1$ *residents play strategy* σ *against one mutant with an arbitrary strategy. For some set of events that are directly observable to the mutant before some time* T*, by which we mean the selection of player pairs in rounds* $t < T$ *as well as, in rounds where the mutant is playing, the actions of the two players, consider the probability that, given that the mutant is drawn to play in round* T*, the mutant's opponent plays action* A*. If this probability is independent of the mutant strategy, of* T*, and of the set of directly observable events that we conditioned on, then we say that* σ *is an effectively unconditional strategy.*

In Proposition 9, we provide an algebraic characterisation of the effectively unconditional strategies.

Proposition 9. *The effectively unconditional strategies exactly consist of those strategies that satisfy* $p_0 = p_{AA} = p_{AB} = p_{BA} = p_{BB}$*, and, when one of* $n = 2$ *or* $\varepsilon = 0$ *or* $\lambda = 0$ *holds, additionally also those that satisfy* $p_0 = p_{AA} = p_{AB} = 1$ *or* $p_0 = p_{BA} = p_{BB} = 0$*.*

Proof. Clearly, any strategy that satisfies $p_0 = p_{AA} = p_{AB} = p_{BA} = p_{BB}$ is effectively unconditional.

It is also easy to see that when the resident strategy satisfies $p_0 = p_{AA} = p_{AB} = 1$, and only direct observations matter, i.e. $n = 2$ or $\varepsilon = 0$ or $\lambda = 0$, then all residents will always assign a reputation of either AA or AB to the mutant (or none). So this strategy will always play A and is thus effectively unconditional. Similarly, when $n = 2$ or $\varepsilon = 0$ or $\lambda = 0$, any strategy with $p_0 = p_{BA} = p_{BB} = 0$ is also effectively unconditional.

It remains to show the converse direction. Assume $p_0 = 1$. Then the strategy is almost unconditional exactly if it always plays action A. Consider the case that the same resident is selected to play against the mutant in both round 0 and round 1, and the mutant played action X in round 0. Because $p_0 = 1$, the resident played action A in round 0, so the reputation assigned by the resident to the mutant at time 1 is AX. Since the resident is almost unconditional, we must have $p_{AX} = 1$. Now additionally assume that $n > 2$ and $\varepsilon > 0$ and $\lambda > 0$, and instead let a different resident be selected to play against the mutant in round 1. With probability $1 - \lambda$, the resident does not yet assign a reputation to the mutant and will thus play action A. Write $\overline{X}$ for the action that is not X. With probability $\lambda(1-\varepsilon)$, the resident assigns reputation AX or $A\overline{X}$ to the mutant and will thus also play action A in this case. But with probability $\lambda\varepsilon(1-\varepsilon)$, the reputation is BX, and with probability $\lambda\varepsilon^2$, it is $B\overline{X}$. In those cases, the resident plays action A with probability p_{BX} and $p_{B\overline{X}}$, respectively. So we must have $p_{BX} = p_{B\overline{X}} = 1$ as well. We can similarly deal with the case $p_0 = 0$.

What remains is the converse direction for the case $0 < p_0 < 1$. Again consider the scenario that the same resident plays against the mutant in the first two rounds. Here, all reputations in $\mathcal{A}^2$ are possible in round 0. In each of those cases, the mutant needs to play action A with probability p_0 in round 1, so we need $p_0 = p_{AA} = p_{AB} = p_{BA} = p_{BB}$. □

We are now ready to show that if a strategy is an equalizer strategy, excluding almost unconditional strategies, then it is of the given form. In the following Proposition 10, we use the functional definition of almost unconditional strategies, whereas in the proof of Theorem 14, we will use the algebraic characterisation provided by Proposition 9.

Proposition 10. *Take any memory-1 equalizer strategy (p_0, p, λ) that is not effectively unconditional. Then either $\tilde{p}$ is of the form $\tilde{p} = \alpha g_{\mathrm{op}} + \beta 1_{\mathcal{A}^2}$ for some $\alpha, \beta \in \mathbb{R}$, or $R = S = T = P$.*

Proof. Let $\sigma = (p_0, p, \lambda)$ be a memory-1 equalizer strategy that is not effectively unconditional. Consider the case of one mutant with an arbitrary strategy playing against a resident population of σ. Let $a(t)$ be the probability that, under the condition that the mutant player is drawn in round t, the mutant's opponent plays action A.

For all $t \geq 0$, we can express $a(t+1)$ in terms of $a(t)$ by distinguishing multiple cases:

Case 1. The mutant did not play in round t. This occurs with a probability of $\frac{n-2}{n}$. In that case, the reputations that the residents assign to the mutant are unchanged compared to round t, and hence the probability that the mutant's opponent plays A is $a(t+1) = a(t)$.

Case 2. The mutant played in round t. This occurs with a probability of $\frac{2}{n}$. We write $v(t) \in \mathbb{R}^{\mathcal{A}^2}$ for (the probability mass function of) the distribution over the outcome in round t under the condition that the mutant played in that round, from the view of the mutant's opponent. So for all $x, y \in \mathcal{A}$, the entry $v_{xy}(t)$ is the probability that the mutant plays action y and the opponent plays action x.

Case 2.1. In round t, the mutant played against the same opponent as in round $t+1$. This occurs with a probability of $\frac{1}{n-1}$ under the condition of Case 2. The opponent's action is determined by the continuation vector p of the resident strategy: The opponent will play A with probability $p^\intercal v(t)$.

Case 2.2. The mutant played against a different opponent as in round $t+1$. This occurs with a probability of $\frac{n-2}{n-1}$ under the condition of Case 2.

Case 2.2.1. The mutant's opponent in round $t+1$ updated the reputation they assign to the mutant in reaction to observing the interaction between the mutant and the other opponent in round t. This occurs with a probability of λ under the condition of Case 2.2. The reputation that the opponent assigns to the mutant at time $t+1$ is distributed as $M_\varepsilon v(t)$. Thus, the opponent's probability to play action A in round $t+1$ is $p^\intercal M_\varepsilon v(t)$.

Case 2.2.2. The mutant's opponent in round $t+1$ did not update the reputation they assign to the mutant in reaction to observing the interaction between the mutant and the other opponent in round t. This occurs with a probability of $1-\lambda$ under the condition of Case 2.2. The reputation assigned by the opponent to the mutant follows the same distribution as that of a randomly drawn mutant in round t, since the events of drawing players and of observation or non-observation occur independently from all past events. Therefore, the opponent's probability to play action A is $a(t+1) = a(t)$.

By the law of total probability, we can write

$$a(t+1) = \frac{n-2}{n}a(t) + \frac{2}{n}\left(\frac{1}{n-1}p^\intercal v(t) + \frac{n-2}{n-1}\left(\lambda p^\intercal M_\varepsilon v(t) + (1-\lambda)a(t)\right)\right) \tag{5.24}$$

$$= \frac{n-2}{n}a(t) + \frac{2}{n(n-1)}p^\intercal\left(I_{\mathcal{A}^2} + \lambda(n-2)M_\varepsilon\right)v(t) + \frac{2(n-2)}{n(n-1)}(1-\lambda)a(t) \tag{5.25}$$

for all $t \geq 0$.

Since $I_{\mathcal{A}^2} + \lambda(n-2)M_\varepsilon$ is symmetric,

$$p^\intercal\left(I_{\mathcal{A}^2} + \lambda(n-2)M_\varepsilon\right)v(t) = v(t)^\intercal\left(I_{\mathcal{A}^2} + \lambda(n-2)M_\varepsilon\right)p.$$

By definition of $\tilde{p}$,

$$\left(I_{\mathcal{A}^2} + \lambda(n-2)M_\varepsilon\right)p = \delta^{-1}(1+\delta\lambda(n-2))(e_{A*} - \tilde{p}).$$

Inserting into (5.25) gives

$$a(t+1) = \frac{n-2}{n}a(t) + \frac{2\delta^{-1}(1+\delta\lambda(n-2))}{n(n-1)}v(t)^\intercal(e_{A*} - \tilde{p}) + \frac{2(n-2)}{n(n-1)}(1-\lambda)a(t). \tag{5.26}$$

Fix some action $X \in \mathcal{A}$ and consider the case that the mutant plays the strategy All$X$, that is, the mutant unconditionally plays action X in every round. Then $v(t) = a(t)e_{AX} + (1-a(t))e_{BX}$ for every $t \geq 0$. With that, we obtain from (5.26) the linear recurrence relation in $a(t)$

$$a(t+1) = \frac{n-2}{n}a(t) + \frac{2\delta^{-1}(1+\delta\lambda(n-2))}{n(n-1)}\left(a(t) - (a(t)\tilde{p}_{AX} + (1-a(t))\tilde{p}_{BX})\right) + \frac{2(n-2)}{n(n-1)}(1-\lambda)a(t) \tag{5.27}$$

$$= \left(\frac{n-2}{n} + \frac{2\delta^{-1}(1+\delta\lambda(n-2))}{n(n-1)}(1-(\tilde{p}_{AX}-\tilde{p}_{BX})) + \frac{2(n-2)}{n(n-1)}(1-\lambda)\right)a(t) - \frac{2\delta^{-1}(1+\delta\lambda(n-2))}{n(n-1)}\tilde{p}_{BX}. \tag{5.28}$$

Let

$$b_X = \frac{n-2}{n} + \frac{2\delta^{-1}(1+\delta\lambda(n-2))}{n(n-1)}(1-(\tilde{p}_{AX}-\tilde{p}_{BX})) + \frac{2(n-2)}{n(n-1)}(1-\lambda) \tag{5.29}$$

and

$$c_X = -\frac{2\delta^{-1}(1+\delta\lambda(n-2))}{n(n-1)}\tilde{p}_{BX}, \tag{5.30}$$

so we can express (5.28) as

$$a(t+1) = b_X a(t) + c_X. \tag{5.31}$$

Noting that $|b_X| < 1$, we can solve that recurrence relation with the initial condition $a(0) = p_0$ as

$$a(t) = \left(p_0 - \frac{c_X}{1-b_X}\right) {b_X}^t + \frac{c_X}{1-b_X}. \tag{5.32}$$

So the overall discount-weighted rate a_X of the resident playing A against an All$X$ mutant is

$$a_X = (1-d)\frac{n}{2}\sum_{t=0}^{\infty}\frac{2}{n}d^t a(t) = \left(p_0 - \frac{c_X}{1-b_X}\right)\frac{1-d}{1-db_X} + \frac{c_X}{1-b_X}. \tag{5.33}$$

The payoff of the All$X$ mutant is then given by

$$\begin{aligned}\pi_X &= G_{XA}a_X + G_{XB}(1-a_X)\\ &= G_{XB} + (G_{XA}-G_{XB})a_X\\ &= G_{XB} + (G_{XA}-G_{XB})\left(\left(p_0 - \frac{c_X}{1-b_X}\right)\frac{1-d}{1-db_X} + \frac{c_X}{1-b_X}\right).\end{aligned} \tag{5.34}$$

By the equalizer property, we must have

$$\pi_A = \pi_B,$$

which is equal to

$$\begin{aligned}S + (R-S)&\left(\left(p_0 - \frac{c_A}{1-b_A}\right)\frac{1-d}{1-db_A} + \frac{c_A}{1-b_A}\right)\\ &= P + (T-P)\left(\left(p_0 - \frac{c_B}{1-b_B}\right)\frac{1-d}{1-db_B} + \frac{c_B}{1-b_B}\right).\end{aligned} \tag{5.35}$$

Since the strategy σ is not effectively unconditional, we can find a mutant strategy σ' and a time T and an event E that is directly observable by the mutant before time T such that E occurs with positive probability and when conditioning on E and on the mutant being selected in round T, the probability that the mutant's opponent plays action A in round t is not equal to $a(0) = p_0$. Fix such σ', T and E, and let that probability be p_0'. For either value of $X \in \mathcal{A}$, we construct the mutant strategy AllX' as follows: The strategy AllX' plays indentically to σ', except if the event E occured. In that case, starting from round T, the strategy AllX' unconditionally always plays action X.

Of course, the recurrence relation (5.31) holds identically in this case for $t \geq T$, now with the initial condition $a(T) = p_0'$. By the equalizer property for σ, the mutant strategies

AllA' and AllB' must have the same continuation payoff at time T under the condition E. By a change of variables $t \mapsto t - T$, we obtain the same equality (5.35) for p_0' instead of p_0. (Note that b_A, b_B, c_A, and c_B are independent of p_0, as we can check from (5.29) and (5.30).) By construction, $p_0' \neq p_0$. So (5.35) is a linear equation in p_0 with at least two solutions. Equality must thus hold for both coefficients:

$$S + (R-S)\frac{c_A}{1-b_A}\left(1 - \frac{1-d}{1-db_A}\right) = P + (T-P)\frac{c_B}{1-b_B}\left(1 - \frac{1-d}{1-db_B}\right) \tag{5.36}$$

$$(R-S)\frac{1-d}{1-db_A} = (T-P)\frac{1-d}{1-db_B} \tag{5.37}$$

We analyse the expression $1-db_X$. Using (5.10) to substitute for d, as well as the definition of b_X, we obtain

$$\begin{aligned}
&1 - db_X \\
&= 1 - \frac{\delta n(n-1)}{2(1-\delta)+\delta n(n-1)}\Bigg(\frac{n-2}{n} + \frac{2\delta^{-1}(1+\delta\lambda(n-2))}{n(n-1)}(1-(\tilde{p}_{AX}-\tilde{p}_{BX})) \\
&\qquad\qquad + \frac{2(n-2)}{n(n-1)}(1-\lambda)\Bigg) \\
&= \frac{1}{2(1-\delta)+\delta n(n-1)}\Bigg((2(1-\delta)+\delta n(n-1)) \\
&\quad - \delta n(n-1)\left(\frac{n-2}{n} + \frac{2\delta^{-1}(1+\delta\lambda(n-2))}{n(n-1)}(1-(\tilde{p}_{AX}-\tilde{p}_{BX})) + \frac{2(n-2)}{n(n-1)}(1-\lambda)\right)\Bigg) \\
&= \frac{1}{2(1-\delta)+\delta n(n-1)}\Big(2(1-\delta)+\delta n(n-1) \\
&\qquad - \delta(n-1)(n-2) - 2(1+\delta\lambda(n-2))(1-(\tilde{p}_{AX}-\tilde{p}_{BX})) - 2\delta(n-2)(1-\lambda)\Big) \\
&= \frac{2(1+\delta\lambda(n-2))}{2(1-\delta)+\delta n(n-1)}(\tilde{p}_{AX}-\tilde{p}_{BX}).
\end{aligned} \tag{5.38}$$

So we can express (5.37) as

$$(R-S)(\tilde{p}_{AB}-\tilde{p}_{BB}) = (T-P)(\tilde{p}_{AA}-\tilde{p}_{BA}). \tag{5.39}$$

Equation (5.36) can be rewritten as

$$S + (R-S)\frac{dc_A}{1-db_A} = P + (T-P)\frac{dc_B}{1-db_B}.$$

Using (5.37), it follows that

$$(S-P)(1-db_A) = -d(R-S)(c_A-c_B).$$

We insert from (5.38) for $T-P$, from (5.10) for d, and from (5.30) for c_A and c_B to get

$$\begin{aligned}
&(S-P)\frac{2(1+\delta\lambda(n-2))}{2(1-\delta)+\delta n(n-1)}(\tilde{p}_{AA}-\tilde{p}_{BA}) \\
&\qquad = \frac{\delta n(n-1)}{2(1-\delta)+\delta n(n-1)}(R-S)\frac{2\delta^{-1}(1+\delta\lambda(n-2))}{n(n-1)}(\tilde{p}_{BA}-\tilde{p}_{BB})
\end{aligned}$$

and thus, after simplification,

$$(S-P)(\tilde{p}_{AA}-\tilde{p}_{BA}) = (R-S)(\tilde{p}_{BA}-\tilde{p}_{BB}). \tag{5.40}$$

Similarly, we may obtain

$$(S-P)(\tilde{p}_{AB}-\tilde{p}_{BB}) = (T-P)(\tilde{p}_{BA}-\tilde{p}_{BB}) \tag{5.41}$$

by using (5.38) for $R-S$ instead of $T-P$.

We distinguish three exhaustive cases:

1. $R \neq S$. Choose $\alpha, \beta \in \mathbb{R}$ such that $\tilde{p}_{AA} = \alpha R + \beta$ and $\tilde{p}_{BA} = \alpha S + \beta$. Note that $\alpha \neq 0$. Then from (5.40) we get that $\tilde{p}_{BB} = \alpha P + \beta$ and consequently from (5.39) that $\tilde{p}_{AB} = \alpha T + \beta$. So $\tilde{p} = \alpha g_{\mathrm{op}} + \beta 1_{\mathcal{A}^2}$.
2. $T \neq P$. Similarly to the previous case, this time using (5.41) and (5.39), we conclude that $\tilde{p} = \alpha g_{\mathrm{op}} + \beta 1_{\mathcal{A}^2}$.
3. $R = S$ and $T = P$. In this case, the mutant's payoff only depends on the mutant's own actions. So σ being an equalizer requires $R = S = T = P$.

So we have shown the statement for all cases. $\square$

Theorem 14. *A strategy $\sigma = (p_0, p, \lambda)$ is an equalizer strategy exactly if one of the following hold:*

1. *The Press-Dyson vector $\tilde{p}$ is of the form $\tilde{p} = \alpha g_{\mathrm{op}} + \beta 1_{\mathcal{A}^2}$ for some $\alpha, \beta \in \mathbb{R}$.*

   Necessarily $\alpha > 0$. This strategy enforces payoff

   $$\pi = \alpha^{-1}\left((1-\eta)p_0 - \beta\right). \tag{5.42}$$

   We call the equalizers of this form the generic equalizers.

2. *The strategy satisfies*

   $$p_0 = p_{AA} = p_{AB} = p_{BA} = p_{BB} = \frac{P-S}{R-S-T+P}.$$

   This strategy enforces payoff

   $$\pi = \frac{RP-ST}{R-S-T+P}.$$

3. *$R = T$, $p_0 = p_{AA} = p_{AB} = 1$, and additionally either of $p_{BA} = p_{BB} = 1$ or $n = 2$ or $\lambda = 0$ or $\varepsilon = 0$.*

   This strategy enforces payoff $\pi = R$.

4. *$S = P$, $p_0 = p_{BA} = p_{BB} = 0$, and additionally either of $p_{AA} = p_{AB} = 0$ or $n = 2$ or $\lambda = 0$ or $\varepsilon = 0$.*

   This strategy enforces payoff $\pi = S$.

5. *$R = S = T = P$.*

   This strategy enforces payoff $\pi = R$.

Proof. First, we consider the case that (p_0, p, λ) is not effectively unconditional. Proposition 10 shows that if (p_0, p, λ) is an equalizer strategy, then either $\tilde{p} = \alpha g_{\mathrm{op}} + \beta 1_{\mathcal{A}^2}$ or $R = S = T = P$. Proposition 8 shows that if $\tilde{p} = \alpha g_{\mathrm{op}} + \beta 1_{\mathcal{A}^2}$, then (p_0, p, λ) is an equalizer strategy. It is obvious that all strategies are equalizer strategies when $R = S = T = P$. So we have equivalence in this case.

It remains to consider the case that σ is effectively unconditional. Proposition 9 gives an algebraic characterisation of the effectively unconditional strategies with given p_0. It remains to determine for which payoff matrices an effectively unconditional strategy with given p_0 is an equalizer strategy. The mutant's payoff in one round where the mutant is selected is $p_0 R + (1 - p_0) S$ if the mutant plays action A and $p_0 T + (1 - p_0) P$ if the mutant plays action B. So the condition for the effectively unconditional resident strategy to be an equalizer strategy is simply

$$p_0 R + (1 - p_0) S = p_0 T + (1 - p_0) P. \tag{5.43}$$

It is easy to check that this is equivalent to the conditions given in the statement of the theorem.

So we have shown the statement of the theorem in all cases. □

Existence of memory-1 equalizer strategies

Theorem 14 gives a complete characterisation of the algebraic forms of memory-1 equalizer strategies. But with case 1 (the generic equalizers) of the theorem, and to a lesser degree also case 2, it is not clear a priori when the given conditions can be satisfied, i.e. in which cases memory-1 equalizer strategies exist and in which cases they do not. So with Theorem 15, we derive the conditions for the existence of memory-1 equalizer strategies from Theorem 14.

Theorem 15. *Generic memory-1 equalizer strategies with indirectness λ exist exactly if $P \neq R$ and*

$$\max\left\{|R_\varepsilon - T_\varepsilon|, |S_\varepsilon - P_\varepsilon|\right\} \leq \frac{\eta}{\zeta - \eta}\left(\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\}\right). \tag{5.44}$$

Additionally, for any λ, the conditions for non-generic memory-1 equalizer strategies to exist, and the payoffs they enforce, are:

1. *If either (a) $R > T$ and $S > P$ or (b) $R < T$ and $S < P$, non-generic memory-1 equalizer strategies with indirectness λ exist and enforce payoff*

   $$\pi = \frac{RP - ST}{R - S - T + P}.$$

2. *If $R = T$, non-generic memory-1 equalizer strategies with indirectness λ exist and enforce payoff $\pi = R$.*

3. *If $S = P$, non-generic memory-1 equalizer strategies with indirectness λ exist and enforce payoff $\pi = S$.*

4. *If either (a) $R > T$ and $S < P$ or (b) $R < T$ and $S > P$, non-generic memory-1 equalizer strategies do not exist.*

Proof. It is easy to derive the conditions for non-generic memory-1 equalizer strategies, so we omit that part. Below, we derive the condition for generic memory-1 equalizer strategies. That is, we show that strategies of the form $\tilde{p} = \alpha g_{\text{op}} + \beta 1_{\mathcal{A}^2}$ (by Theorem 14) exist exactly if $P \neq R$ and (5.43) holds:

$$\max\{|R_\varepsilon - T_\varepsilon|, |S_\varepsilon - P_\varepsilon|\} \leq \frac{\eta}{\zeta - \eta}\left(\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\}\right)$$

Recall the definition of $\tilde{p}$ as $\tilde{p} = e_{A*} - H_\varepsilon p$. We want to find conditions determining whether there exists p such that $\tilde{p}$ is of the form $\tilde{p} = \alpha g_{\text{op}} + \beta 1_{\mathcal{A}^2}$, or equivalently, whether we can choose p, α and β such that

$$e_{A*} - \eta H_\varepsilon p = \alpha g_{\text{op}} + \beta 1_{\mathcal{A}^2}. \tag{5.45}$$

Since H_ε is invertible, we can rearrange (5.45) to

$$\eta p = H_\varepsilon^{-1}(e_{A*} - \alpha g_{\text{op}} - \beta 1_{\mathcal{A}^2}), \tag{5.46}$$

Recalling the definition of ζ in (5.17), we may verify that

$$H_\varepsilon^{-1} 1_{\mathcal{A}^2} = 1_{\mathcal{A}^2}$$

and

$$H_\varepsilon^{-1} e_{A*} = \zeta e_{A*} + \frac{1-\zeta}{2} 1_{\mathcal{A}^2}.$$

So (5.46) becomes

$$\eta p = \zeta e_{A*} + \frac{1-\zeta}{2} 1_{\mathcal{A}^2} - \alpha H_\varepsilon^{-1} g_{\text{op}} - \beta 1_{\mathcal{A}^2}. \tag{5.47}$$

We want to determine when it is possible to find $p \in [0,1]^{\mathcal{A}^2}$ and $\alpha, \beta \in \mathbb{R}$ such that (5.47) holds. Recalling our notation $(R_\varepsilon, T_\varepsilon, S_\varepsilon, P_\varepsilon)^\intercal := H_\varepsilon^{-1} g_{\text{op}}$, that is the case exactly if we can choose $\alpha, \beta \in \mathbb{R}$ such that

$$0 \leq \frac{1+\zeta}{2} - \alpha R_\varepsilon - \beta \leq \eta \tag{5.48}$$

$$0 \leq \frac{1+\zeta}{2} - \alpha T_\varepsilon - \beta \leq \eta \tag{5.49}$$

$$0 \leq \frac{1-\zeta}{2} - \alpha S_\varepsilon - \beta \leq \eta \tag{5.50}$$

$$0 \leq \frac{1-\zeta}{2} - \alpha P_\varepsilon - \beta \leq \eta. \tag{5.51}$$

Since $P \leq R$, we also have $P_\varepsilon \leq R_\varepsilon$. Furthermore, $\eta < 1$ and $\zeta \geq 1$. So we need $\alpha > 0$ for the above inequalities to be satisfied. The conditions are thus equivalent to the existence of $\alpha \in \mathbb{R}_{>0}$ and $\beta \in \mathbb{R}$ such that

$$\beta \leq \frac{1+\zeta}{2} - \alpha \max\{R_\varepsilon, T_\varepsilon\} \tag{5.52}$$

$$\beta \leq \frac{1-\zeta}{2} - \alpha \max\{S_\varepsilon, P_\varepsilon\} \tag{5.53}$$

$$\frac{1+\zeta}{2} - \alpha \min\{R_\varepsilon, T_\varepsilon\} - \eta \leq \beta \tag{5.54}$$

$$\frac{1-\zeta}{2} - \alpha \min\{S_\varepsilon, P_\varepsilon\} - \eta \leq \beta. \tag{5.55}$$

This is equivalent to the existence of $\alpha \in \mathbb{R}_{>0}$ such that

$$\frac{1+\zeta}{2} - \alpha \min\{R_\varepsilon, T_\varepsilon\} - \eta \leq \frac{1+\zeta}{2} - \alpha \max\{R_\varepsilon, T_\varepsilon\} \tag{5.56}$$

$$\frac{1-\zeta}{2} - \alpha \min\{S_\varepsilon, P_\varepsilon\} - \eta \leq \frac{1+\zeta}{2} - \alpha \max\{R_\varepsilon, T_\varepsilon\} \tag{5.57}$$

$$\frac{1+\zeta}{2} - \alpha \min\{R_\varepsilon, T_\varepsilon\} - \eta \leq \frac{1-\zeta}{2} - \alpha \max\{S_\varepsilon, P_\varepsilon\} \tag{5.58}$$

$$\frac{1-\zeta}{2} - \alpha \min\{S_\varepsilon, P_\varepsilon\} - \eta \leq \frac{1-\zeta}{2} - \alpha \max\{S_\varepsilon, P_\varepsilon\} \tag{5.59}$$

or equivalently, such that

$$\alpha |R_\varepsilon - T_\varepsilon| \leq \eta \tag{5.60}$$

$$\alpha \max\{R_\varepsilon, T_\varepsilon\} - \alpha \min\{S_\varepsilon, P_\varepsilon\} \leq \eta + \zeta \tag{5.61}$$

$$\alpha \max\{S_\varepsilon, P_\varepsilon\} - \alpha \min\{R_\varepsilon, T_\varepsilon\} \leq \eta - \zeta \tag{5.62}$$

$$\alpha |S_\varepsilon - P_\varepsilon| \leq \eta. \tag{5.63}$$

Noting that $\zeta - \eta > 0$, this is equivalent to the existence of $\alpha \in \mathbb{R}_{>0}$ such that

$$\frac{\max\{|R_\varepsilon - T_\varepsilon|, |S_\varepsilon - P_\varepsilon|\}}{\eta} \leq \alpha^{-1} \tag{5.64}$$

$$\frac{\max\{R_\varepsilon, T_\varepsilon\} - \min\{S_\varepsilon, P_\varepsilon\}}{\eta + \zeta} \leq \alpha^{-1} \tag{5.65}$$

$$\alpha^{-1} \leq \frac{\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\}}{\zeta - \eta}. \tag{5.66}$$

That is the case exactly if $\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\} > 0$ and

$$(\zeta - \eta) \max\{|R_\varepsilon - T_\varepsilon|, |S_\varepsilon - P_\varepsilon|\} \leq \eta \left(\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\}\right) \tag{5.67}$$

$$(\zeta - \eta) \left(\max\{R_\varepsilon, T_\varepsilon\} - \min\{S_\varepsilon, P_\varepsilon\}\right) \leq (\eta + \zeta) \left(\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\}\right). \tag{5.68}$$

The condition $\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\} > 0$ requires $P_\varepsilon < R_\varepsilon$. But if $P_\varepsilon < R_\varepsilon$ is given then also $\max\{R_\varepsilon, T_\varepsilon\} - \min\{S_\varepsilon, P_\varepsilon\} > 0$, so by (5.68), we have $\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\} > 0$ as well. We can check with the definition of H_ε that $P_\varepsilon < R_\varepsilon$ is equivalent to $P < R$, and thus to $P \neq R$. Thus, we can replace the condition $\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\} > 0$ by $P \neq R$.

Of the two inequalities (5.67) and (5.68), we can also drop (5.68), because it follows from (5.67): From (5.67), we get the weaker

$$(\zeta - \eta) \left(|R_\varepsilon - T_\varepsilon| + |S_\varepsilon - P_\varepsilon|\right) \leq 2\eta \left(\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\}\right) \tag{5.69}$$

and thus

$$(\zeta-\eta)\left(\max\{R_\varepsilon,T_\varepsilon\}-\min\{R_\varepsilon,T_\varepsilon\}+\max\{S_\varepsilon,P_\varepsilon\}-\min\{S_\varepsilon,P_\varepsilon\}\right)$$
$$\leq 2\eta\left(\min\{R_\varepsilon,T_\varepsilon\}-\max\{S_\varepsilon,P_\varepsilon\}\right), \quad (5.70)$$

which is equivalent to (5.68).

So generic memory-1 equalizer strategies exist exactly if $P \neq R$ and (5.67) holds. □

We obtain several corollaries from this result by restricting to the special case of additive games, or letting δ go to 1.

Corollary 18. *In an additive game, memory-1 equalizer strategies with indirectness λ exist exactly if either (a) $P \neq R$ and*

$$|S-P| \leq \zeta^{-1}\eta(T-P) \quad (5.71)$$

or (b) $R = T$.

Proof. We adapt the statement of Theorem 15 for the case of additive games.

In an additive game, $R = T$ and $S = P$ are equivalent. The conjunction of $R > T$ and $S > P$ is impossible, same as $R < T$ and $S < P$. So non-generic memory-1 equalizer strategies exist exactly if $R = T$.

It remains to simplify the condition for generic equalizer strategies. Consider Eq. (5.44):

$$\max\{|R_\varepsilon-T_\varepsilon|,|S_\varepsilon-P_\varepsilon|\} \leq \frac{\eta}{\zeta-\eta}\left(\min\{R_\varepsilon,T_\varepsilon\}-\max\{S_\varepsilon,P_\varepsilon\}\right)$$

Note that $R+P=S+T \Rightarrow R_\varepsilon+P_\varepsilon=S_\varepsilon+T_\varepsilon$. So we have $|R_\varepsilon-T_\varepsilon|=|S_\varepsilon-P_\varepsilon|$. Furthermore,

$$\min\{R_\varepsilon,T_\varepsilon\}=T_\varepsilon+\min\{(S_\varepsilon-P_\varepsilon),0\}$$

and

$$\max\{S_\varepsilon,P_\varepsilon\}=P_\varepsilon+\max\{(S_\varepsilon-P_\varepsilon),0\}.$$

So (5.44) becomes

$$|S_\varepsilon-P_\varepsilon| \leq \frac{\eta}{\zeta-\eta}\left((T_\varepsilon-P_\varepsilon)-|S_\varepsilon-P_\varepsilon|\right)$$

which is equivalent to

$$|S_\varepsilon-P_\varepsilon| \leq \zeta^{-1}\eta(T_\varepsilon-P_\varepsilon). \quad (5.72)$$

We can check that in our case of $R+P=S+T$, we have $S_\varepsilon-P_\varepsilon=S-P$ and $T_\varepsilon-P_\varepsilon=T-P$. □

When we compare this Corollary 18, which states when memory-1 equalizers exist in additive games, to Corollary 15 from Section 5.4.3, which states when reactive equalizers exist in additive games, we see that the conditions are equivalent:

Corollary 19. *In an additive game, for given λ, reactive equalizer strategies with indirectness λ exist exactly if memory-1 equalizer strategies with indirectness λ exist.*

Proof. Assume that memory-1 equalizer strategies with indirectness for some given λ exist. Consider first the case that $P \neq R$. We have $\zeta^{-1}\eta < 1$. So (5.71) in Corollary 18 implies that either $|S-P| < |T-P|$ or $S = T = P$. But $P \neq R$ is equivalent to $S \neq T$ in an additive game. So we must have $|S-P| < |T-P|$ and thus $T \neq P$. We can therefore express the condition (5.71) by

$$\delta^{-1} \frac{1+(n-2)\delta\lambda}{1+(n-2)(1-2\varepsilon)\lambda} \cdot \frac{P-S}{T-P} \in [-1, 1]$$

as in Corollary 15.

Now instead consider the case $P = R$. By Corollary 18, we also have $R = T$. So $R = S = T = P$.

Of course, reactive strategies are a subset of memory-1 strategies, so the existence of reactive equalizers trivially implies the existence of memory-1 equalizers. □

Finally, we ask when memory-1 equalizer strategies (with arbitrary λ) exist for sufficiently large δ. The conditions for non-generic memory-1 equalizer strategies to exist are obvious from Theorem 15, since they are independent of δ and λ. The below corollary gives the condition for generic memory-1 equalizers.

Corollary 20. *Generic memory-1 equalizer strategies exist for sufficiently large δ if and only if G is a social dilemma.*

Proof. By Theorem 15, generic memory-1 equalizer strategies exist exactly if (5.44) holds:

$$\max\{|R_\varepsilon - T_\varepsilon|, |S_\varepsilon - P_\varepsilon|\} \leq \frac{\eta}{\zeta - \eta}\left(\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\}\right)$$

As $\delta \to 1$, this inequality becomes

$$(\zeta - 1)\max\{|R_\varepsilon - T_\varepsilon|, |S_\varepsilon - P_\varepsilon|\} \leq \min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\}. \quad (5.73)$$

Assume that generic memory-1 equalizers exist. Then by (5.73), we have $\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\} > 0$. Thus also $\min\{R, T\} - \max\{S, P\} > 0$. So the game is a social dilemma.

Now assume instead that the game is a social dilemma. For $\lambda = 0$, (5.73) takes the form

$$0 \leq \min\{R, T\} - \max\{S, P\}.$$

This is given, so generic memory-1 equalizers exist.

We have thus shown equivalence. □

Within the social dilemmas, we now know that generic memory-1 equalizer strategies exist in all of them, while non-generic memory-1 equalizers additionally exist in the stag hunt and hawk-dove games. Of these, the non-generic memory-1 equalizer strategies are in general unique up to the irrelevant choice of λ, where they exist. In contrast, where generic memory-1 equalizer strategies exist, in general there exists a whole family of them, parameterised by p_0, α, β and λ. It is not surprising then that when we look for equalizer strategies in social dilemmas that have the desirable property of being cooperative, which we do in Section 5.4.4 below, that we find them among the generic equalizers.

Optimisation over memory-1 equalizer strategies

Theorem 16. *Given game parameters and an indirectness value λ as input, Algorithm 1 on page 92 correctly outputs the generic equalizer with the highest payoff if such an equalizer exists, and* **null** *otherwise. In the case $\zeta \neq 1$, if equalizers exist, the equalizer output by the algorithm is unique in having the highest payoff.*

Proof. A generic equalizer strategy is determined by $p_0 \in [0,1]$, $\alpha \in \mathbb{R}_{>0}$, $\beta \in \mathbb{R}$ and $\lambda \in [0,1]$. It enforces a payoff of

$$\pi = \alpha^{-1}((1-\eta)p_0 - \beta). \tag{5.74}$$

Following the proof of Theorem 15, we see that the conditions for a choice of p_0, α, β and λ to produce a valid equalizer are equivalent to:

1. $P_\varepsilon \neq R_\varepsilon$,
2. $\alpha^{-1} \geq \omega_{\min}$, where $\omega_{\min} := \max\{\max\{|R_\varepsilon - T_\varepsilon|, |S_\varepsilon - P_\varepsilon|\}/\eta, (\max\{R_\varepsilon, T_\varepsilon\} - \min\{S_\varepsilon, P_\varepsilon\})/(\eta+\zeta)\}$
3. $\alpha^{-1} \leq \omega_{\max}$, where $\omega_{\max} := (\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\})/(\zeta - \eta)$,
4. $\beta \geq \beta_{\min}$, where $\beta_{\min} := \max\{(1+\zeta)/2 - \alpha\min\{R_\varepsilon, T_\varepsilon\}, (1-\zeta)/2 - \alpha\min\{S_\varepsilon, P_\varepsilon\}\} - \eta$,
5. and $\beta \leq \beta_{\max}$, where $\beta_{\max} := \min\{(1+\zeta)/2 - \alpha\max\{R_\varepsilon, T_\varepsilon\}, (1-\zeta)/2 - \alpha\max\{S_\varepsilon, P_\varepsilon\}\}$.

The parameter λ is given as input. The algorithm needs to compute p_0, α, and β that maximise π. The conditions above are independent of p_0. From (5.74), we see that for fixed α and β, the optimal choice of p_0 is always $p_0 = 1$. So we can start by setting $p_0 = 1$.

Secondly, we check that the first condition $P_\varepsilon \neq R_\varepsilon$, which does not depend on α and β, is met, and output **null** otherwise.

We know that given this first condition, and given any $\alpha \in \mathbb{R}_{>0}$, it is exactly when α satisfies $\omega_{\min} \leq \alpha^{-1} \leq \omega_{\max}$ that there exists β such that $\beta_{\min} \leq \beta \leq \beta_{\max}$. So assume we have a fixed choice of α that satisfies $\omega_{\min} \leq \alpha^{-1} \leq \omega_{\max}$. Then clearly from (5.74), the value of β that conditionally maximises π is $\beta = \beta_{\min}$. So the maximal payoff with this α is

$$\pi(\alpha) = \alpha^{-1}\left((1-\eta) - \left(\max\left\{\frac{1+\zeta}{2} - \alpha\min\{R_\varepsilon, T_\varepsilon\}, \frac{1-\zeta}{2} - \alpha\min\{S_\varepsilon, P_\varepsilon\}\right\} - \eta\right)\right) \tag{5.75}$$

$$= \min\left\{\frac{1-\zeta}{2\alpha} + \min\{R_\varepsilon, T_\varepsilon\}, \frac{1+\zeta}{2\alpha} + \min\{S_\varepsilon, P_\varepsilon\}\right\}. \tag{5.76}$$

We write

$$\omega_{\text{break}} = (\min\{R_\varepsilon, T_\varepsilon\} - \min\{S_\varepsilon, P_\varepsilon\})/\zeta.$$

For convenience, we write $\omega = \alpha^{-1}$ and let the algorithm choose α by choosing a value for ω. Then we have

$$\pi(\omega) = \begin{cases} \omega\frac{1+\zeta}{2} + \min\{S_\varepsilon, P_\varepsilon\} & \text{if } \omega \leq \omega_{\text{break}} \\ \omega\frac{1-\zeta}{2} + \min\{R_\varepsilon, T_\varepsilon\} & \text{if } \omega \geq \omega_{\text{break}} \end{cases} \tag{5.77}$$

for all $\omega \in \mathbb{R}_{>0}$.

If the interval $[\omega_{\min}, \omega_{\max}]$ is non-empty, the algorithm must now choose $\omega \in [\omega_{\min}, \omega_{\max}]$ that maximises π. If the interval is empty, the algorithm must output **null**. Practically, it sets $\omega = \max\{\omega_{\min}, \omega_{\text{break}}\}$ and returns **null** if $\omega > \omega_{\max}$. We show below that that is equivalent.

First, assume that $[\omega_{\min}, \omega_{\max}]$ is non-empty, i.e. $\omega_{\min} \leq \omega_{\max}$. Then we can show that $\omega_{\text{break}} \leq \omega_{\max}$ as well: By definition,

$$|S_\varepsilon - P_\varepsilon|/\eta \leq \omega_{\min}.$$

Combining with $\omega_{\min} \leq \omega_{\max}$ gives

$$|S_\varepsilon - P_\varepsilon|/\eta \leq (\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\})/(\zeta - \eta).$$

We multiply with the denominators on both sides and get

$$(\zeta - \eta)|S_\varepsilon - P_\varepsilon| \leq \eta(\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\}),$$

which we can also write as

$$\begin{aligned} \zeta\max\{S_\varepsilon, P_\varepsilon\} - \zeta\min\{S_\varepsilon, P_\varepsilon\} + \eta\min\{S_\varepsilon, P_\varepsilon\} - \eta\max\{S_\varepsilon, P_\varepsilon\} \\ \leq \eta\min\{R_\varepsilon, T_\varepsilon\} - \eta\max\{S_\varepsilon, P_\varepsilon\}. \end{aligned}$$

This is equivalent to

$$\begin{aligned} \zeta\min\{R_\varepsilon, T_\varepsilon\} - \eta\min\{R_\varepsilon, T_\varepsilon\} - \zeta\min\{S_\varepsilon, P_\varepsilon\} + \eta\min\{S_\varepsilon, P_\varepsilon\} \\ \leq \zeta\min\{R_\varepsilon, T_\varepsilon\} - \zeta\max\{S_\varepsilon, P_\varepsilon\}. \end{aligned}$$

Dividing by $\zeta - \eta$ and ζ gives

$$(\min\{R_\varepsilon, T_\varepsilon\} - \min\{S_\varepsilon, P_\varepsilon\})/\zeta \leq (\min\{R_\varepsilon, T_\varepsilon\} - \max\{S_\varepsilon, P_\varepsilon\})/(\zeta - \eta).$$

By definition, that is $\omega_{\text{break}} \leq \omega_{\max}$.

Consider the derivative of $\pi(\omega)$ with respect to ω. We obtain from (5.77) that

$$\frac{d}{d\omega}\pi(\omega) = \begin{cases} \frac{1+\zeta}{2} & \text{if } \omega < \omega_{\text{break}} \\ \frac{1-\zeta}{2} & \text{if } \omega > \omega_{\text{break}} \end{cases}. \tag{5.78}$$

We know that $\zeta \geq 1$. So the derivative is negative on $\omega > \omega_{\text{break}}$, zero on $\omega < \omega_{\text{break}}$ if $\zeta = 1$, and positive on $\omega < \omega_{\text{break}}$ if $\zeta > 1$. The case $\zeta = 1$ occurs exactly if $n = 2$ or $\lambda = 0$ or $\varepsilon = 0$.

We can distinguish three cases as to which values $\omega \in [\omega_{\min}, \omega_{\max}]$ maximise $\pi(\omega)$. The first case is that $\omega_{\min} \leq \omega_{\text{break}}$ and $\zeta > 1$. In this case, $\omega = \omega_{\text{break}}$ is the only optimum. The second case is that $\omega_{\min} \leq \omega_{\text{break}}$ and $\zeta = 1$. In this case, the optimal values are exactly the set $[\omega_{\min}, \omega_{\text{break}}]$. The third case is that $\omega_{\text{break}} < \omega_{\min}$. In this case, $\omega = \omega_{\min}$ is the only optimum. So the choice we made, which is $\omega = \min\{\omega_{\min}, \omega_{\text{break}}\}$, is indeed optimal. The condition $\omega > \omega_{\max}$ evaulates as false, since both $\omega_{\text{break}} \leq \omega_{\max}$ and $\omega_{\min} \leq \omega_{\max}$. So the algorithm does not return **null** at this point, which is correct.

It remains to consider the case that $[\omega_{\min}, \omega_{\max}]$ is empty, i.e. that $\omega_{\min} > \omega_{\max}$. Here, since $\omega \geq \omega_{\min}$, the condition $\omega > \omega_{\max}$ evalates as true and the algorithm correctly returns **null**.

Because ω satisfies $\omega \geq \omega_{\text{break}}$, the expression for the optimal β, which is $\beta_{\min}$, simplifies to

$$\beta = \frac{1-\zeta}{2} - \alpha \min\{R_\varepsilon, T_\varepsilon\} - \eta.$$

The enforced payoff, by (5.77), can be written as

$$\pi = \frac{1-\zeta}{2\alpha} + \min\{R_\varepsilon, T_\varepsilon\}. \tag{5.79}$$

According to (5.46), the continuation vector p of the equalizer strategy is given by

$$p = \eta^{-1} H_\varepsilon^{-1}(e_{A*} - \alpha g_{\text{op}} - \beta 1_{\mathcal{A}^2}).$$

We may check manually or with a computer algebra system that

$$H_\varepsilon \left(\frac{1+\zeta}{2}, \frac{1+\zeta}{2}, \frac{1-\zeta}{2}, \frac{1-\zeta}{2}\right)^\intercal = (1,1,0,0)^\intercal = e_{A*}.$$

Of course, $H_\varepsilon^{-1} 1_{\mathcal{A}^2} = 1_{\mathcal{A}^2}$. So indeed we have

$$p = 1_{\mathcal{A}^2} - \eta^{-1}\zeta e_{B*} - \eta^{-1}\alpha(H_\varepsilon^{-1} g_{\text{op}} - \min\{R_\varepsilon, T_\varepsilon\}).$$

By definition of $R_\varepsilon, S_\varepsilon, T_\varepsilon, P_\varepsilon$, and by H_ε being symmetric,

$$H_\varepsilon^{-1} g_{\text{op}} = (R_\varepsilon, T_\varepsilon, S_\varepsilon, P_\varepsilon)^\intercal,$$

and we can also write p as in Algorithm 1 as

$$\begin{pmatrix} p_{AA} \\ p_{AB} \\ p_{BA} \\ p_{BB} \end{pmatrix} = \begin{pmatrix} 1 - \eta^{-1}\alpha(R_\varepsilon - \min\{R_\varepsilon, T_\varepsilon\}) \\ 1 - \eta^{-1}\alpha(T_\varepsilon - \min\{R_\varepsilon, T_\varepsilon\}) \\ 1 - \eta^{-1}\zeta - \eta^{-1}\alpha(S_\varepsilon - \min\{R_\varepsilon, T_\varepsilon\}) \\ 1 - \eta^{-1}\zeta - \eta^{-1}\alpha(P_\varepsilon - \min\{R_\varepsilon, T_\varepsilon\}) \end{pmatrix}. \tag{5.80}$$

□

This algorithm allows us to obtain some results about the highest payoff that is enforceable with generic equalizer strategies, if any, and about the existence of those particularly desirable equalizers in social dilemmas that always play action A, unless they made an observation error. It is already clear from Theorem 15 that non-generic equalizers do not have this property unless $R = T$, even in those social dilemmas where they do exist.

Corollary 21. *If a generic equalizer enforces payoff* π*, then*

$$\max\{S_\varepsilon, P_\varepsilon\} \leq \pi \leq \min\{R_\varepsilon, T_\varepsilon\}.$$

Both of these bounds are attained if and only if $n = 2$ *or* $\lambda = 0$ *or* $\varepsilon = 0$*, as long as equalizers exist at all.*

Proof. From Theorem 16, we know that the maximal enforceable payoff $\pi_{\max}$ is of the form

$$\pi = \min\{R_\varepsilon, T_\varepsilon\} - \alpha^{-1}\frac{\zeta - 1}{2}.$$

The upper bound follows immediately, knowing that $\zeta \geq 1$. The condition $\zeta = 1$ is equivalent to $n = 2$ or $\lambda = 0$ or $\varepsilon = 0$.

The statement about the lower bound is reduced to the statement about the upper bound by a transformation of the payoff matrix as $(R, S, T, P) \mapsto (-P, -T, -S, -R)$. □

Definition 25 (Action-A-playing)**.** *A strategy is action-A-playing if it always plays action A at least until the first error occurs.*

In social dilemmas, this desirable property is usually referred to as being cooperative.

Proposition 11. *An action-A-playing memory-1 equalizer strategy with indirectness* λ *exists exactly if any memory-1 equalizer strategy with indirectness* λ *exists and*

$$R_\varepsilon \leq T_\varepsilon. \tag{5.81}$$

The condition (5.81) is equivalent to

$$(T - R)(1 + \lambda(n-2)(1-2\varepsilon)(1-\varepsilon)) + (S - P)\lambda(n-2)(1-2\varepsilon)\varepsilon \geq 0. \tag{5.82}$$

In error-free games, it simplifies to $R \leq T$*.*

Proof. Clearly a memory-1 strategy is action-A-playing exactly if $p_0 = p_{AA} = 1$. The condition $p_0 = 1$ is simple. The condition $p_{AA} = 1$ corresponds a choice of β such that inequality (5.48) in the proof of Theorem 15 is tight. We see from inequalities (5.48–5.51) and (5.54–5.55) that we can choose such a β, for given α, exactly if

$$\frac{1+\zeta}{2} - \alpha T_\varepsilon \leq \frac{1+\zeta}{2} - \alpha R_\varepsilon \tag{5.83}$$

and

$$\frac{1+\zeta}{2} - \alpha \min\{R_\varepsilon, T_\varepsilon\} - \eta \leq \frac{1-\zeta}{2} - \alpha \min\{S_\varepsilon, P_\varepsilon\} - \eta. \tag{5.84}$$

Inequality (5.84) is equivalent to $\alpha^{-1} \leq \omega_{\text{break}}$ where $\omega_{\text{break}} = (\min\{R_\varepsilon, T_\varepsilon\} - \min\{S_\varepsilon, P_\varepsilon\})/\zeta$ like in the proof of Theorem 16. Since $\omega_{\text{break}} \leq \omega_{\max}$ is always given as long as $\omega_{\min} \leq \omega_{\max}$ (i.e. as long as generic equalizer strategies with indirectness λ exist), the condition (5.84) is void in that case.

The other condition, (5.83), is equivalent to $R_\varepsilon \leq T_\varepsilon$. This shows the statement in the first form.

Now, $R_\varepsilon \le T_\varepsilon$ is equivalent to

$$(-1,0,1,0)(R_\varepsilon,S_\varepsilon,T_\varepsilon,P_\varepsilon)^\intercal \ge 0. \tag{5.85}$$

By definition, $(R_\varepsilon,S_\varepsilon,T_\varepsilon,P_\varepsilon)^\intercal = (1+\lambda(n-2))(I_{\mathcal{A}^2}+\lambda(n-2)M_\varepsilon)^{-1}(R,S,T,P)^\intercal$, so this is equivalent to

$$(-1,0,1,0)(I_{\mathcal{A}^2}+\lambda(n-2)M_\varepsilon)^{-1}(R,S,T,P)^\intercal \ge 0, \tag{5.86}$$

or, since $I_{\mathcal{A}^2}+\lambda(n-2)M_\varepsilon$ is symmetric,

$$(R,S,T,P)(I_{\mathcal{A}^2}+\lambda(n-2)M_\varepsilon)^{-1}(-1,0,1,0)^\intercal \ge 0. \tag{5.87}$$

We can verify by hand that

$$(-1,0,1,0)^\intercal = \frac{1}{2}(I_{\mathcal{A}^2}+\lambda(n-2)M_\varepsilon)\left(\frac{(-1,-1,1,1)^\intercal}{1+\lambda(n-2)(1-2\varepsilon)} + \frac{(-1,1,1,-1)^\intercal}{1+\lambda(n-2)(1-2\varepsilon)^2}\right). \tag{5.88}$$

So (5.87) is equivalent to

$$\frac{-R-S+T+P}{1+\lambda(n-2)(1-2\varepsilon)} + \frac{-R+S+T-P}{1+\lambda(n-2)(1-2\varepsilon)^2} \ge 0, \tag{5.89}$$

which in turn can be simplified to

$$(T-R)(1+\lambda(n-2)(1-2\varepsilon)(1-\varepsilon)) + (S-P)\lambda(n-2)(1-2\varepsilon)\varepsilon \ge 0. \tag{5.90}$$

□

Corollary 22. *Assume that equalizer strategies exist.*

If $R_\varepsilon \le T_\varepsilon$, then all equalizer strategies that enforce the maximal possible payoff are action-A-playing. In particular, they have $p_{AA} = 1$.

If $T_\varepsilon \le R_\varepsilon$, then all equalizer strategies that enforce the maximal possible payoff have $p_{AB} = 1$.

Proof. Even when there are multiple optimal equalizer strategies, we see from the proof of Theorem 16 that all of them satisfy (5.80) for some value of α.

If $R_\varepsilon \le T_\varepsilon$, we may insert $\min\{R_\varepsilon, T_\varepsilon\} = R_\varepsilon$ into the first component of (5.80) to obtain $p_{AA} = 1$. Since also p_0, these strategies are action-A-playing.

If $T_\varepsilon \le R_\varepsilon$, we similarly obtain $p_{AB} = 1$. □

Corollary 23. *If any action-A-playing equalizer strategy exists, then there is also an action-A-playing equalizer strategy that enforces the maximal payoff among all equalizer strategies.*

In the special case that $\zeta = 1$, every action-A-playing equalizer strategy enforces this maximal payoff.

Proof. By Proposition 11, if an action-A-playing equalizer strategy exists, then $R_\varepsilon \le T_\varepsilon$. By Theorem 16, there is an equalizer strategy that enforces a maximal payoff. Take one of those. By Corollary 22, it is action-A-playing.

If $\zeta = 1$, then observation errors never occur, because either $n = 2$ or $\lambda = 0$ or $\varepsilon = 0$. So in this case, an action-A-playing strategy by definition always plays action A and thus obtains payoff R in a homogeneous population. If it is an equalizer, the payoff it enforces is therefore also $\pi = R$. In particular, all action-A-playing equalizer strategies enforce the same payoff, which is the maximal payoff. □

5.4.5 Proof of Theorem 13

We consider a population of n players who are all playing the same reactive strategy $\sigma = (p_0, p, \lambda)$. We analyse the stochastic process of these players' game, which will allow us to show that an alternative strategy can achieve a higher payoff than σ against an otherwise homogeneous σ population in the cases stated in Theorem 13.

First, we define some notation: We write $\mathcal{R} = \{(i,j) : i,j \in [n], i \neq j\}$ for the set of disjoint player pairs. In each round, a pair $(i,j) \in \mathcal{R}$ is selected uniformly at random. We write $\mathcal{A} = \{A, B\}$ for the set of actions. Where X is a random variable taking values in set S, we write $p[X] \in \mathbb{R}^S$ for the probability distribution function of X.

Intents

For mathematical convenience, we stipulate that whenever players update the reputation they assign to another player, they immediately form a secrent intent for how they will act towards that player next time that the are selected to interact with them, given that they don't update their reputation again before that time. This is formally expressed by the following definition.

Definition 26 (Intent)**.** *Player i's intent towards Player j at time t, denoted as $Y_{(i,j)}(t)$, is a random variable taking values in $\mathcal{A}$. It depends only on the history of the game before time t, and it does so in the following way:*

1. *If $t = 0$, then $Y_{(i,j)}(t) = A$ with probability p_0 and $Y_{(i,j)}(t) = B$ with probability $1 - p_0$.*
2. *If $t > 0$ and Player i did not update the reputation they assign to Player j in round $t-1$, then $Y_{(i,j)}(t) = Y_{(i,j)}(t-1)$.*
3. *If $t > 0$ and Player i updated the reputation they assign to Player j in round $t-1$ to value $X \in \mathcal{A}$, then $Y_{(i,j)}(t) = A$ with probability p_X and $Y_{(i,j)}(t) = B$ with probability $1 - p_X$.*

If players i and j are selected to play, Player i's action is given by $Y_{(i,j)}(t)$.

With the above definition, we not only constructed a set of random variables called intents, but also redefined actions in relation to intents rather than reputations, as they were previously defined. Naturally, we would have considered the reputations players assign to each other as the primary object representing the state of the game. But by studying the stochastic process in terms of intents rather than reputations, we avoid having to distinguish between the case when a player already assigns a reputation to another player and when they do not. Clearly, this is an equivalent formulation of the game.

Definition 27 (Intent vector)**.** *The intent vector $Y(t)$ at time t is the collection of all intents at time t:*

$$Y(t) \coloneqq (Y_{(i,j)}(t) : (i,j) \in \mathcal{R})$$

The state space of the intent vector $Y(t)$ is the set of all possible configurations of intents, $\mathcal{A}^{\mathcal{R}}$. It is easy to see that the sequence $(Y(t))_t$ satisfies the time-independent Markov property. That is, for all $t > 0$, $Y(t+1)$ only depends on $Y(t)$ and not

$Y(0), \ldots, Y(t-1)$, and there is a transition function f, which is independent of t, such that $p\,[Y(t+1) \mid Y(t) = y(t)] = f(y(t))$ for all t. Given that, we write P for the transition matrix of $(Y(t))_t$, meaning that if f is the transition function, then

$$\mathbb{P}\,[Y(t+1) = y' | Y(t) = y] = P_{y'y} = f_{y'}(y) \text{ for all } y,\, y'.$$

Note that technically, $\mathbb{P}[Y(t) = y]$ can be zero for some y, in which case P and f are ill-defined. But we may simply alternatively define P and f by saying that for all y, if a group of n players with start in a state in which they have intents y, then after one round of following strategy σ, their intents are distributed as $f(y)$.

We can derive Player i's expected payoff in round t from the intent vector $Y(t)$ as

$$\pi_i(t) = \frac{2}{n(n-1)} \sum_{j \in [n] \setminus \{i\}} g^\intercal p \left[(Y_{(i,j)}(t), Y_{(j,i)}(t)) \right]. \tag{5.91}$$

So, in theory, understanding the Markov chain $(Y(t))_t$ could allow us to analyse the equilibria of the game. Unfortunately, the complexity of its transition function makes this difficult, so we will first reduce the system to a simpler one. Even before that, we need to introduce an additional concept, which is intent origins.

Intent origins

In the original formulation of the game process, there are four sources of randomness: (a) the selection of a player pair, (b) the selection of actions by the two players, and finally, with those players who were not involved, (c) the occurrence of observation errors and (d) the occurrence of reputation updates. We replaced the random selection of actions with the random selection of intents, which means that actions, given a selected player pair, are now a deterministic function of intents. So now, the four sources of randomness, again in temporal order, are (a) intents, (b) player pair selection, (c) errors, and (d) reputation updates. We observe that (c) and (d) only depend on (b) of the same round, and (b) depends on nothing at all. (Note: with (d), we are referring to the random event of a player deciding whether or not to update a reputation, not to the actual value that the reputation is updated with, which does depend on (a).) Meanwhile, intents can depend on all events of the previous round.

We call the the player pair selection, the occurrence of errors, and the occurrence of reputation updates the circumstantial events. We write $C(t)$ for the random variable representing the circumstantial events at time t. The random variables $C(0), C(1), \ldots$ are independent and identically distributed.

Knowing that $(Y(t))_t$ satisfies the time-independent Markov property, we may analyse its transition matrix P without reference to a particular time t: For a intents given by the random variable Y, let Y' be the random variable representing the intents after one round of gameplay. Write C for the random variable representing the circumstantial events of that one round. Write $\mathcal{C}$ for the set of its possible values. By definition of f, the conditional distribution of Y' given Y is $f(Y)$.

We condition on the entire process up to time t, as well as on the circumstantial events at time $t+1$, and analyse the conditional distribution of $Y(t+1)$. To that end, consider a general intent $Y_{(i,j)}(t+1)$ at time $t+1$. We distinguish two cases, depending on the value of C:

Definition 28 (Intent origin). *Conditional on $C = c$, and for a given player pair $(i,j) \in \mathcal{R}$, we define the intent origin of (i,j), which is also a player pair, and which we denote by $a_c(i,j) \in \mathcal{R}$. We define $a_c(i,j)$ by distinguishing three different cases.*

- *If Player j was not among those two players who were chosen to play according to c, then $a_c(i,j) = (i,j)$.*
- *If Player j was selected to play, but not against Player i, and Player i did not observe the event (and thus also did not update their intent towards Player j), then $a_c(i,j) = (i,j)$.*
- *If Player j was selected to play against some Player k (who may or may not be identical to i), and Player i updated their intent towards Player j in response to observing Player j's action, then $a_c(i,j) = (j,k)$.*

In each of these cases, the intent $Y'_{(i,j)}$ of the final state somehow traces back to the intent $Y_{a_c(i,j)}$ of the original state: In the first two cases, this is simply because the intent was not updated. In the third case, this is because i's final intent towards j, that is $Y'_{(i,j)}$, was formed because of j's action towards another player, and that action was determined by the original intent $Y_{a_c(i,j)}$.

In fact, when conditioning on $C = c$, the intent $Y'_{(i,j)}$ only depends on $Y_{a_c(i,j)}$, and not of any of the other intents in Y. This is formally expressed by the following proposition.

Proposition 12. *Fix a player pair $(i,j) \in \mathcal{R}$. Given any outcome c of the circumstantial events of one round, there exists a (necessarily linear) function $Q_c(i,j) : \mathbb{R}^{\mathcal{A}} \to \mathbb{R}^{\mathcal{A}}$ such that*

$$p\left[Y'_{(i,j)} \mid C = c\right] = Q_c(i,j)p\left[Y_{a_c(i,j)}\right]. \tag{5.92}$$

for all y.

Proof. We make a similar case distinction as in Definition 28 based on the value of c.

If c is such that Player j did not play, or they did, but Player i did not update the reputation they assign to and thus their intent towards Player j, then $a_c(i,j) = (i,j)$. Therefore, we have $Y'_{(i,j)} = Y_{(i,j)}$ with probability 1, i.e.

$$p\left[Y'_{(i,j)} \mid C = c\right] = p\left[Y_{a_c(i,j)}\right].$$

If c is such that Player j played against Player i, then $a_c(i,j) = (j,i)$. The action of Player j was equal to $Y_{(j,i)}(t)$, which is distributed as $y_{(j,i)}(t)$. The action was observed by Player i, who then updated their intent towards Player j. Player 1 reacts to the observation based on their continuation vector p, which is a component of the reactive strategy $\sigma = (p_0, p, \lambda)$:

$$p\left[Y'_{(i,j)} \mid C = c\right] = \begin{pmatrix} p_A & p_B \\ 1-p_A & 1-p_B \end{pmatrix} p\left[Y_{a_c(i,j)}\right].$$

Finally, if c is such that Player j played against some Player k, who is not identical with i, then we have $a_c(i,j) = (j,k)$ for an analogous reason as in the previous case. The equation contains an error matrix as an additional factor, because observation errors may occur:

$$p\left[Y'_{(i,j)} \mid C = c\right] = \begin{pmatrix} p_A & p_B \\ 1-p_A & 1-p_B \end{pmatrix} \begin{pmatrix} 1-\varepsilon & \varepsilon \\ \varepsilon & 1-\varepsilon \end{pmatrix} p\left[Y_{a_c(i,j)}\right].$$

The relevant point is only that the conditional distribution of $Y'_{(i,j)}$ is purely a function of $p\left[Y_{a_c(i,j)}\right]$. We have observed that this is true for all possible values of c. □

Example.

We illustrate the concept of intent origins with an example of $n = 3$ players. In this example, there are $\binom{n}{2} = 6$ intents.

Say that in some round t, Player 1 played against Player 2. Player 3 observed this and randomly decided to update their intent towards Player 1, but not that towards Player 2. The intent origins are as follows.

- First of all, $a_c((1,3)) = (1,3)$ and $a_c((2,3)) = (2,3)$. Since Player 3 was not selected to play, other players did not update their intents towards Player 3. These intents are thus unchanged and are their own parents.
- Similarly, $a_c((3,2)) = (3,2)$. Player 2 was selected to play, but Player 3 nonetheless did not update the intent $(3,2)$.
- But Player 3 did update the intent $(3,1)$. The new intent is based on Player 1's action towards Player 2, which in turn is based on Player 1's intent of Player 2, that is $(1,2)$. So $a_c((3,1)) = (1,2)$.
- The selected Players, 1 and 2, certainly updated their intents towards each other. So the new value of the intent $(2,1)$ also stochastically depends on $(1,2)$ in the previous round, and vice versa. We have $a_c((1,2)) = (2,1)$ and $a_c((2,1)) = (1,2)$.

Marginal distributions

We write $w_{(i,j)} \coloneqq p\left[Y_{(i,j)}\right]$ for the marginal distribution of $Y_{(i,j)}$, and similarly $w'_{(i,j)} \coloneqq p\left[Y'_{(i,j)}\right]$. We write

$$w = (w_{(i,j)} : (i,j) \in \mathcal{R}) = \left(p\left[Y_{(i,j)}\right] : (i,j) \in \mathcal{R}\right)$$

for the collection of these marginal distributions, and similarly w' for the collection of the $w'_{(i,j)}$.

We interpret w as an element of the vector space $\mathbb{R}^{\mathcal{R}} \otimes \mathbb{R}^{\mathcal{A}}$, where $\mathbb{R}^{\mathcal{R}} \otimes \mathbb{R}^{\mathcal{A}}$ denotes the tensor product of $\mathbb{R}^{\mathcal{R}}$ and $\mathbb{R}^{\mathcal{A}}$. For a general introduction to tensor products, see e.g. Lang [Lan02]. Here, we just mention some basic properties that we will use. Firstly, the vector space $\mathbb{R}^{\mathcal{R}} \otimes \mathbb{R}^{\mathcal{A}}$ is isomorphic to $\mathbb{R}^{\mathcal{R}\times\mathcal{A}}$. Given elements $v \in \mathbb{R}^{\mathcal{R}}$ and $w \in \mathbb{R}^{\mathcal{A}}$, the element $a \in \mathbb{R}^{\mathcal{R}} \otimes \mathbb{R}^{\mathcal{A}}$ defined by $a_{(i,j),x} = v_{(i,j)} w_x$ for all $(i,j) \in \mathcal{R}$ and $x \in \mathcal{A}$ is denoted by $a = v \otimes w$. Not all elements of $\mathbb{R}^{\mathcal{R}\times\mathcal{A}}$ can be written in this form, but a basis of $\mathbb{R}^{\mathcal{R}} \otimes \mathbb{R}^{\mathcal{A}}$ is given by $\{e_{(i,j),x} = e_{(i,j)} \otimes e_x : (i,j) \in \mathcal{R}, x \in \mathcal{A}\}$.

When we multiply e.g. a row $\mathcal{R}$-vector with a vector from from $\mathbb{R}^{\mathcal{R}} \otimes \mathbb{R}^{\mathcal{A}}$, the result is an $\mathcal{A}$-vector. In particular, if $(i,j) \in \mathcal{R}$ and $a \in \mathbb{R}^{\mathcal{A}} \otimes \mathbb{R}^{\mathcal{A}}$ is given by the general form $a = \sum_{(i,j)\in\mathcal{R},x\in\mathcal{A}} a_{(i,j),x} e_{(i,j),x}$ where $a_{(i,j),x} \in \mathbb{R}$ for all $(i,j) \in \mathcal{R}$ and $x \in \mathcal{A}$, then $e_{(i,j)}^{\intercal} a = \sum_{x\in\mathcal{A}} a_{(i,j)x} e_x \in \mathbb{R}^{\mathcal{A}}$.

Fundamentally, the vector $w \in \mathbb{R}^{\mathcal{R}} \otimes \mathbb{R}^{\mathcal{A}}$ assigns a real number to each pair of a player pair $(i,j) \in \mathcal{R}$ and an action $x \in \mathcal{A}$, which is the probability that Player i's intent towards Player j is x.

The following is a direct consequence of Proposition 12:

Corollary 24. *There exists a matrix $Q \in \mathcal{M}_{\mathcal{R}}(\mathbb{R}) \otimes \mathcal{M}_{\mathcal{A}}(\mathbb{R})$ such that for all w and corresponding w', we have $w' = Qw$.*

Proof. Fix a player pair $(i,j) \in \mathcal{R}$. By the law of total probability, summing (5.92) over all possible values of c,

$$\begin{aligned} w'_{(i,j)} = p\left[Y'_{(i,j)}\right] = \sum_{c\in\mathcal{C}} \mathbb{P}[C=c] Q_c(i,j) p\left[Y_{a_c(i,j)}\right] &= \sum_{c\in\mathcal{C}} \mathbb{P}[C=c] Q_c(i,j) w_{a_c(i,j)} \\ &= \sum_{c\in\mathcal{C}} \mathbb{P}[C=c] Q_c(i,j) (e_{a_c(i,j)}^{\intercal} w) \\ &= \sum_{c\in\mathcal{C}} \mathbb{P}[C=c] (e_{a_c(i,j)}^{\intercal} \otimes Q_c(i,j)) w. \end{aligned}$$

Define the matrix $Q \in \mathcal{M}_{\mathcal{R}}(\mathbb{R}) \otimes \mathcal{M}_{\mathcal{A}}(\mathbb{R})$ by

$$Q = \sum_{(i,j)\in\mathcal{R}} e_{(i,j)} \sum_{c\in\mathcal{C}} \mathbb{P}[C=c] e_{a_c(i,j)}^{\intercal} \otimes Q_c(i,j). \tag{5.93}$$

We note that Q is defined independently of w and w'. In particular, the probabilities $\mathbb{P}[C=c]$ are independent of w. Then we have

$$w' = Qw. \tag{5.94}$$

□

Writing $w(t) = (p\left[Y_{(i,j)}(t)\right] : (i,j) \in \mathcal{R})$ for the collection of marginal distributions of intents at time t, we can now express $w(t)$ simply as $w(t) = Q^t w(0)$. We say that $(w(t))_t$ has the time-independent marginal Markov property, in accordance with the following definition.

Definition 29 (Marginal Markov property)**.** *We are given a sequence of random variables $(X(t))_t$ whose elements take values in $\bigtimes_{i\in I} S_i$ for some index set I. In other words, values of $X(t)$ are of the form $(x_i : i \in I)$, where $x_i \in S_i$ for all $i \in I$.*

The sequence $(X(t))_t$ satisfies the marginal Markov property if for each t and each $i \in I$, the marginal distribution of $X_i(t+1)$ only depends on the marginal distributions $(p\left[X_j(t)\right] : j \in I)$.

If in addition there exists a (necessarily linear) function f such that

$$(p\left[X_i(t+1)\right] : i \in I) = f((p\left[X_i(t)\right] : i \in I)),$$

then the sequence satisfies the time-independent marginal Markov property.

Corollary 24, which is stated with the knowledge that $(Y(t))_t$ satisfies the time-independent Markov property, shows that $(Y(t))_t$ also satisfies the time-independent marginal Markov property. However, neither is in general a consequence of the other.

The vectors $w(t)$ are elements of $\mathbb{R}^{\mathcal{R}} \otimes \mathbb{R}^{\mathcal{A}} \cong \mathbb{R}^{\mathcal{R}\times\mathcal{A}}$, whereas the distributions of the $Y(t)$ are elements of the vastly much larger space $\mathbb{R}^{\mathcal{A}^{\mathcal{R}}}$, since $Y(t)$ takes values in $\mathcal{A}^{\mathcal{R}}$. And indeed, analysing Q is much simpler than analysing the transition matrix of the Markov chain $(Y(t))_t$. However, $(w(t))_t$ unfortunately does not contain all the information we need to compute payoffs: From (5.91), we see that the payoffs depend on pairwise joint distributions of intents. It is only in additive games that the following holds:

$$\pi_i(t) = \frac{2}{n(n-1)} \sum_{j\in[n]\setminus\{i\}} p\left[Y_{(i,j)}(t)\right]^{\intercal} Gp\left[Y_{(j,i)}(t)\right]. \tag{5.95}$$

So $(w(t))_t$ is an oversimplification of the system.

However, in the same way as we derived the marginal Markov property for $(Y(t))_t$ from Proposition 12, we can do the same for the collection of pairwise intent distributions.

We define the pairwise intent $\ddot{Y}_{(i_1,j_1),(i_2,j_2)}(t)$ for all t and all $(i_1, j_1) \in \mathcal{R}$, $(i_2, j_2) \in \mathcal{R}$ by

$$\ddot{Y}_{(i_1,j_1),(i_2,j_2)}(t) := (Y_{(i_1,j_1)}(t), Y_{(i_2,j_2)}(t))$$

as a random variable taking values in $\mathcal{A}^2$. We write $\ddot{Y}(t)$ for the collection of all pairwise intents at time t:

$$\ddot{Y}(t) := (\ddot{Y}_\rho(t) : \rho \in \mathcal{R}^2).$$

Again like with $Y(t)$, we denote the componentwise marginal distributions as follows:

$$\ddot{w}(t) = (\ddot{w}_\rho(t) : \rho \in \mathcal{R}^2) = (p\left[\ddot{Y}_\rho(t)\right] : \rho \in \mathcal{R}^2).$$

These objects are somewhat internally redundant, since for $(i_1, j_1) \neq (i_2, j_2)$, the two distributions $\ddot{w}_{(i_1,j_1),(i_2,j_2)}(t)$ and $\ddot{w}_{(i_2,j_2),(i_1,j_1)}(t)$ represent the same information. However, this is notationally convenient.

We define the vector $z \in \mathbb{R}^{\mathcal{A}}$ as

$$z = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{5.96}$$

and the coordination vector $\ddot{z} \in \mathbb{R}^{\mathcal{A}^2}$ as

$$\ddot{z} = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} = z^{\otimes 2}. \tag{5.97}$$

Note: In the above equation (5.97), the symbol "$\otimes$" denotes the Kronecker product and not the tensor product, which simply means that we interpret $\ddot{z}$ as a vector in $\mathbb{R}^{\mathcal{A}^2}$ and not the naturally isomorphic $\mathbb{R}^{\mathcal{A}} \otimes \mathbb{R}^{\mathcal{A}}$.

Proposition 13. *Fix two player pairs $(i_1, j_1) \in \mathcal{R}$ and $(i_2, j_2) \in \mathcal{R}$. Given any outcome c of the circumstantial events of one round, there exists a matrix $\ddot{Q}_c((i_1, j_1), (i_2, j_2)) \in \mathcal{M}_{\mathcal{R}^2}(\mathbb{R}) \otimes \mathcal{M}_{\mathcal{A}^2}(\mathbb{R})$ such that*

$$p\left[\ddot{Y}'_{(i_1,j_1),(i_2,j_2)} \mid C = c\right] = \ddot{Q}_c((i_1, j_1), (i_2, j_2))p\left[\ddot{Y}_{a_c(i_1,j_1),a_c(i_2,j_2)}\right], \tag{5.98}$$

for all $c \in \mathcal{C}$ and such that $\ddot{z}$ is an eigenvector of $\ddot{Q}_c((i_1, j_1), (i_2, j_2))$ for some eigenvalue μ satisfying $|\mu| \leq 1$.

Proof. We show this by distinguishing the cases $(i_1, j_1) = (i_2, j_2)$ and $(i_1, j_1) \neq (i_2, j_2)$, both of which we will reduce to Proposition 12. Fix some $c \in \mathcal{C}$.

First, if $(i_1, j_1) = (i_2, j_2)$, then of course also $a_c(i_1, j_1) = a_c(i_2, j_2)$. By Proposition 12, take $Q_c(i_1, j_2)$ such that

$$p\left[Y'_{(i_1,j_1)} \mid C = c\right] = Q_c(i_1, j_2) p\left[Y_{a_c(i_1,j_1)}\right].$$

Let

$$Q_c(i_1, j_1) = \begin{pmatrix} q_{AA} & q_{BA} \\ q_{AB} & q_{BB} \end{pmatrix}.$$

Then we can define

$$\ddot{Q}_c((i_1, j_1), (i_2, j_2)) = \begin{pmatrix} q_{AA} & q_{AA} & q_{BA} & q_{BA} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ q_{AB} & q_{AB} & q_{BB} & q_{BB} \end{pmatrix}.$$

(This is one of multiple possible choices.) Clearly, this satisfies the conditions, with $\mu = 0$.

Secondly, consider the case that $(i_1, j_1) \neq (i_2, j_2)$. When we condition on $Y_{a_c(i_1,j_1)}$ and $Y_{a_c(i_2,j_2)}$ in addition to C, then the random variables $Y_{(i_1,j_1)}$ and $Y_{(i_2,j_2)}$ are independent, since the formation of intents, given the observations they are based on, is independent. We can express this as

$$\begin{aligned} p&\left[(Y'_{(i_1,j_1)}, Y'_{(i_2,j_2)}) \mid C = c \wedge Y_{a_c(i_1,j_1)} = y_1 \wedge Y_{a_c(i_2,j_2)} = y_2\right] \\ &= p\left[Y'_{(i_1,j_1)} Y_{a_c(i_1,j_1)} = y_1 \wedge Y_{a_c(i_2,j_2)} = y_2\right] \\ &\qquad \otimes p\left[Y'_{(i_2,j_2)} Y_{a_c(i_1,j_1)} = y_1 \wedge Y_{a_c(i_2,j_2)} = y_2\right] \end{aligned}$$

for all $y_1, y_2 \in \mathcal{A}$. (Again, like in (5.97), "$\otimes$" denotes the Kronecker product here.) Take $Q_c(i_1, j_1)$ and $Q_c(i_2, j_2)$ as in Proposition 12. Like all stochastic matrices of order 2, they have z as an eigenvector. Let $\mu_1, \mu_2 \in [-1, 1]$ be the respectivly corresponding eigenvalues. By construction of $Q_c(i_1, j_1)$ and $Q_c(i_2, j_2)$, we have

$$\begin{aligned} p&\left[(Y'_{(i_1,j_1)}, Y'_{(i_2,j_2)}) \mid C = c \wedge Y_{a_c(i_1,j_1)} = y_1 \wedge Y_{a_c(i_2,j_2)} = y_2\right] \\ &= Q_c(i_1, j_1) p\left[Y_{a_c(i_1,j_1)} Y_{a_c(i_1,j_1)} = y_1 \wedge Y_{a_c(i_2,j_2)} = y_2\right] \\ &\qquad \otimes Q_c(i_2, j_2) p\left[Y_{a_c(i_2,j_2)} Y_{a_c(i_1,j_1)} = y_1 \wedge Y_{a_c(i_2,j_2)} = y_2\right] \end{aligned}$$

for all $y_1, y_2 \in \mathcal{A}$. Summing over y_1 and y_2, we get

$$p\left[(Y'_{(i_1,j_1)}, Y'_{(i_2,j_2)}) \mid C = c\right] = Q_c(i_1, j_1) p\left[Y_{a_c(i_1,j_1)}\right] \otimes Q_c(i_2, j_2) p\left[Y_{a_c(i_2,j_2)}\right]$$

or equivalently

$$p\left[(Y'_{(i_1,j_1)}, Y'_{(i_2,j_2)}) \mid C = c\right] = Q_c(i_1, j_1) w_{a_c(i_1,j_1)} \otimes Q_c(i_2, j_2) w_{a_c(i_2,j_2)}. \tag{5.99}$$

Defining $\ddot{Q}_c((i_1, j_1), (i_2, j_2))$ by

$$\ddot{Q}_c((i_1, j_1), (i_2, j_2)) = Q_c(i_1, j_1) \otimes Q_c(i_2, j_2),$$

we may express (5.99) as

$$p\left[(Y'_{(i_1,j_1)}, Y'_{(i_2,j_2)}) \mid C = c\right] = \ddot{Q}_c((i_1,j_1),(i_2,j_2))\ddot{w}_{a_c(i_1,j_1),a_c(i_2,j_2)}.$$

Furthermore, we have $\ddot{Q}_c((i_1,j_1),(i_2,j_2))\ddot{z} = \mu_1\mu_2\ddot{z}$, and $|\mu_1\mu_2| \leq 1$. So all conditions are satisfied. □

The following statement, also a consequence of Proposition 12, shows that $(\ddot{Y}(t))_t$ satisfies the time-independent marginal Markov property.

Proposition 14. *There exist matrices $\ddot{Q} \in \mathcal{M}_{\mathcal{R}^2}(\mathbb{R}) \otimes \mathcal{M}_{\mathcal{A}^2}(\mathbb{R})$ and $A \in \mathcal{M}_{\mathcal{R}^2}(\mathbb{R})$ with the following properties. For all $\ddot{w}$ and corresponding $\ddot{w}'$, we have $\ddot{w}' = \ddot{Q}\ddot{w}$. The matrix A satisfies $\|Ax\|_\infty \leq \|x\|_\infty$ for all $x \in \mathbb{R}^{\mathcal{R}^2}$. Finally, $\ddot{Q}\ddot{z} = A \otimes \ddot{z}$.*

Note: The expression $\ddot{Q}\ddot{z} = A \otimes \ddot{z}$ may appear unfamiliar. It is equivalent to saying that $\ddot{Q}(x \otimes \ddot{z}) = (Ax) \otimes \ddot{z}$ for all $x \in \mathbb{R}^{\mathcal{R}^2}$.

Proof. Fix two player pairs $(i_1,j_1) \in \mathcal{R}$ and $(i_2,j_2) \in \mathcal{R}$.

By the law of total probability, summing (5.98) over all possible values of c,

$$\ddot{w}'_{(i_1,j_1),(i_2,j_2)} = \sum_{c\in\mathcal{C}} \mathbb{P}[C=c]\ddot{Q}_c((i_1,j_1),(i_2,j_2))e^{\intercal}_{a_c(i_1,j_1),a_c(i_2,j_2)}\ddot{w}$$

Define the matrix $\ddot{Q} \in \mathcal{M}_{\mathcal{R}^2}(\mathbb{R}) \otimes \mathcal{M}_{\mathcal{A}^2}(\mathbb{R})$ by

$$\ddot{Q} = \sum_{(i_1,j_1),(i_2,j_2)\in\mathcal{R}} e_{(i_1,j_1),(i_2,j_2)} \sum_{c\in\mathcal{C}} \mathbb{P}[C=c]e^{\intercal}_{a_c(i_1,j_1),a_c(i_2,j_2)} \otimes \ddot{Q}_c((i_1,j_1),(i_2,j_2)). \tag{5.100}$$

Then

$$\ddot{w}' = \ddot{Q}\ddot{w}. \tag{5.101}$$

Consider the value of $\ddot{Q}\ddot{z}$, and let $\mu_c((i_1,j_1),(i_2,j_2))$ be the eigenvalue of $\ddot{Q}_c((i_1,j_1),(i_2,j_2))$ corresponding to the eigenvector $\ddot{z}$. We have

$$\begin{aligned}\ddot{Q}\ddot{z} &= \sum_{(i_1,j_1),(i_2,j_2)\in\mathcal{R}} e_{(i_1,j_1),(i_2,j_2)} \sum_{c\in\mathcal{C}} \mathbb{P}[C=c](e^{\intercal}_{a_c(i_1,j_1),a_c(i_2,j_2)} \otimes \ddot{Q}_c((i_1,j_1),(i_2,j_2)))\ddot{z} \\ &= \left(\sum_{(i_1,j_1),(i_2,j_2)\in\mathcal{R}} e_{(i_1,j_1),(i_2,j_2)} \sum_{c\in\mathcal{C}} \mathbb{P}[C=c]\mu_c((i_1,j_1),(i_2,j_2))e^{\intercal}_{a_c(i_1,j_1),a_c(i_2,j_2)}\right) \otimes \ddot{z}.\end{aligned}$$

Define

$$A = \sum_{(i_1,j_1),(i_2,j_2)\in\mathcal{R}} \sum_{c\in\mathcal{C}} \mathbb{P}[C=c]\mu_c((i_1,j_1),(i_2,j_2))e_{(i_1,j_1),(i_2,j_2)}e^{\intercal}_{a_c(i_1,j_1),a_c(i_2,j_2)}.$$

By construction, A has the property $\ddot{Q}\ddot{z} = A \otimes \ddot{z}$.

It remains to show that $\|Ax\|_\infty \leq \|x\|_\infty$ for all $x \in \mathbb{R}^{\mathcal{R}^2}$. Take any $x \in \mathbb{R}^{\mathcal{R}^2}$, and consider the absolute value of the $((i_1, j_1), (i_2, j_2))$-th component of Ax:

$$\begin{aligned}
\left|(Ax)_{(i_1,j_1),(i_2,j_2)}\right| &= \left|\sum_{c\in\mathcal{C}} \mathbb{P}[C=c]\mu_c((i_1,j_1),(i_2,j_2))e^{\intercal}_{a_c(i_1,j_1),a_c(i_2,j_2)}x\right| \\
&= \left|\sum_{c\in\mathcal{C}} \mathbb{P}[C=c]\mu_c((i_1,j_1),(i_2,j_2))x_{a_c(i_1,j_1),a_c(i_2,j_2)}\right| \\
&\leq \sum_{c\in\mathcal{C}} \mathbb{P}[C=c]\left|\mu_c((i_1,j_1),(i_2,j_2))\right|\left|x_{a_c(i_1,j_1),a_c(i_2,j_2)}\right| \\
&\leq \|x\|_\infty \sum_{c\in\mathcal{C}} \mathbb{P}[C=c]\left|\mu_c((i_1,j_1),(i_2,j_2))\right| \\
&\leq \|x\|_\infty
\end{aligned}$$

This is true for all $((i_1, j_1), (i_2, j_2)) \in \mathcal{R}^2$. So indeed $\|Ax\|_\infty \leq \|x\|_\infty$. □

The effect of one unit of coordination

The double intent $\ddot{Y}_{((i,j),(j,i))}(t)$ determines payoffs in round t if players i and j are selected to play. Double intents that are not formed of a matching pair (i.e. where one pair is (i, j) but the other one is not (j, i)), are not needed to derive players' payoffs, but the double intents in subsequent rounds, including the payoff-relevant matching double intents, do depend on them. We denote the set of matching double player pairs that contain Player 1 by $\mathcal{I}$:

$$\mathcal{I} := \{((1,i),(i,1)) : i \neq 1\} \cup \{((i,1),(1,i)) : i \neq 1\} \subseteq \mathcal{R}^2$$

In a given round t, for any given player $i \neq 1$, Player 1 has a probability of $2/(n(n-1))$ of being drawn together with that player. The set $\mathcal{I}$ actually contains two elements corresponding to the unordered pair $\{1, i\}$, but we may assign a probability of $1/(n(n-1))$ to each of them. Over all of $\mathcal{I}$, these sum to $2/n$, which is Player 1's probability of being one of the selected players.

Given then that the players $(1, i)$ are drawn to interact in round t, Player 1's expected payoff is given by

$$\pi_1(t) = g^{\intercal}\ddot{w}_{((1,i),(i,1))}(t).$$

If we don't know which player pair will be drawn, the expected payoff is

$$\pi_1(t) = \frac{1}{n(n-1)}\sum_{\rho\in\mathcal{I}}(e_\rho \otimes g)^{\intercal}\ddot{w}(t). \tag{5.102}$$

For convenience, we define the vector $s \in \mathbb{R}^{\mathcal{R}^2}$ as

$$s = \sum_{\rho\in\mathcal{I}} e_\rho,$$

so that we can write

$$\pi_1(t) = \frac{1}{n(n-1)}(s \otimes g)^{\intercal}\ddot{w}(t). \tag{5.103}$$

While not relevant for the subsequent proof, we may observe that the following set of vectors forms an orthogonal basis of $\mathbb{R}^{\mathcal{A}^2}$:

$$\begin{aligned} b_1 &= e_{AA} + e_{AB} + e_{BA} + e_{BB} \\ b_S &= e_{AA} + e_{AB} - e_{BA} - e_{BB} \\ b_O &= e_{AA} - e_{AB} + e_{BA} - e_{BB} \\ \ddot{z} = b_C &= e_{AA} - e_{AB} - e_{BA} + e_{BB}. \end{aligned}$$

Given some $\rho \in \mathcal{I}$, we can analyse $\ddot{w}_\rho$ in this basis:

- We always have $b_1^\intercal \ddot{w}_\rho = 1$, since $\ddot{w}_\rho$ is a stochastic vector.
- We can interpret $b_S^\intercal \ddot{w}_\rho$ as encoding Player 1's A-rate, that is, Player 1's probability of playing action A. Specifically, if Player 1's A-rate is p, then $b_S^\intercal \ddot{w}_\rho = 2p - 1$.
- Similarly, we can interpret $b_O^\intercal \ddot{w}_\rho$ as encoding the A-rate of Player 1's opponent.
- Finally, $b_C^\intercal \ddot{w}_\rho$ can be interpreted as encoding the coordination rate between the two players. If p is the probability that the players both play the same action, then $b_C^\intercal \ddot{w}_\rho = 2p - 1$.

Player 1's total (i.e. summed over all rounds) A-rate, the opponents' total A-rate, and Player 1's total coordination rate with their opponents together determine Player 1's payoff.

We know from [SCHN21] that if the other players all use reactive strategies, then their A-rates at time t only depend on the A-rate of Player 1 in the preceding rounds. Furthermore, Proposition 14 shows that coordination only affects coordination, in the sense that if at any point in the game we intervene by increasing the probability that Player 1 and Player 2 will both play A in the next round, while keeping their individual probabilities of playing A constant, then all future A-rates will be also unchanged. This is a consequence of the fact that reactive strategies only take into account individual actions, not pairwise interactions.

Our aim is to show that within the conditions of Theorem 13, reactive Nash equilibria do not exist. We know that reactive Nash equilibria do exist in many games that are additive, but otherwise satisfy the criteria of Theorem 13. So the requirement that the game be non-additive is required for the theorem to hold. It is the defining feature of an additive game that one-round payoffs only depend on the two players' A-rates, but not on their coordination rate. Analogously to the reactive equalizer strategies of additive games, we could also construct reactive strategies for non-additive games that neutralise the effect of a mutant's A-rate on the mutant's payoff. In order to outperform such strategies, Player 1's only option is to manipulate the coordination rate.

As we will ultimately conclude, it is possible to profitably do so against all reactive strategies within the conditions of Theorem 13. By Proposition 14, the mutant can be sure that such a manipulation of the coordination rate will only affect the total coordination rate, but not their own or their opponents' A-rates, against any reactive opponent. Depending on the payoff matrix, the mutant may either wish to increase or to decrease the total coordination rate.

Intuitively though, while Player 1 has full control over their own A-rate, they only partially control their coordination rate with the other players. It is not immediately clear how they would willfully increase or decrease the latter. Even if they are to have a way of positively or negatively affecting their coordination rate in a given round t, this may (and indeed will) have an effect on the coordination rate in the subsequent rounds, and a priori we cannot exclude the possibility that that subsequent effect either reverses or exactly cancels out the effect in round t.

But the following proposition states that this is never the case.

Proposition 15.

$$s^\intercal \sum_{\tau=0}^{\infty} d^\tau A^\tau (e_{((1,2),(2,1))} + e_{((2,1),(1,2))}) > 0. \tag{5.104}$$

As constructed from Proposition 14, the matrix $A \in \mathcal{M}_{\mathcal{R}^2}(\mathbb{R})$ quantifies the effects that the current coordination rates of all the pairs of player pairs $((i_1, j_1), (i_2, j_2)) \in \mathcal{R}^2$ have on the coordination rates in the next round.

Because we use a redundant representation for notational convenience, it is the sum of the two basis vectors,

$$e_{((1,2),(2,1))} + e_{((2,1),(1,2))},$$

that represents one unit of Player 1's coordination rate with Player 2. We use the notation

$$u := e_{((1,2),(2,1))} + e_{((2,1),(1,2))}. \tag{5.105}$$

Proposition 15 states that this unit of coordination with Player 2 in any one round has a positive effect on Player 1's total coordination rate with all players. So if Player 1 can increase their coordination rate with Player 2 in round t compared to some reference, then they have thereby also increased their total coordination rate, and thus, in an additive game, positively or negatively affected their own payoff.

Proof. By symmetry,

$$s^\intercal \sum_{\tau=0}^{\infty} d^\tau A^\tau (e_{((1,i),(i,1))} + e_{((i,1),(1,i))}) = s^\intercal \sum_{\tau=0}^{\infty} d^\tau A^\tau (e_{((1,2),(2,1))} + e_{((2,1),(1,2))}) \tag{5.106}$$

for all $i \neq 1$, so (5.104) is equivalent to

$$s^\intercal \sum_{\tau=0}^{\infty} d^\tau A^\tau s > 0. \tag{5.107}$$

(Technically, there can be more than one possible choice of $\ddot{Q}$ that satisfies Proposition 14, and we only know that we may choose $\ddot{Q}$ in such a way that A has these symmetry properties. But that is sufficient.)

We will first show

$$s^\intercal \left(\sum_{\tau=0}^{\infty} d^\tau A^\tau \right) s \neq 0. \tag{5.108}$$

and then infer (5.107) using the intermediate value theorem.

Define the matrix $Z \in \mathcal{M}_{\mathcal{R}^2}(\mathbb{R})$ as

$$Z = \sum_{\rho \notin \mathcal{I}} e_\rho \otimes e_\rho^\intercal.$$

So given a vector $x \in \mathbb{R}^{\mathcal{R}^2}$, the vector Zx is equal to x with all components in the standard basis indexed by some $\rho \in \mathcal{I}$ set to 0.

Consider the double-sequence $v(l, m)$ defined on $l, m \geq 0$ by

$$v(l,m) = \begin{cases} (ZA)^l s & \text{if } l \leq m \\ A^{l-m}(ZA)^m s & \text{if } l \geq m. \end{cases},$$

Observe that for all $l \geq 0$, we have

$$\lim_{m\to\infty} v(l,m) = (ZA)^l s. \tag{5.109}$$

We know that $\|Ax\|_\infty \leq \|x\|_\infty$ for all $x \in \mathbb{R}^{\mathcal{R}^2}$. It is clear from the definition of Z that also $\|Zx\|_\infty \leq \|x\|_\infty$ for all $x \in \mathbb{R}^{\mathcal{R}^2}$. So we have $||v(l,m)||_\infty \leq ||s||_\infty = 1$ for all l and m.

Define the double sequence $b(L, m)$ for $L, m \geq 0$ by

$$b(L,m) = s^\intercal \sum_{l=0}^{L} d^l v(l,m).$$

By the Weierstraß M-test, $b(L, m)$ converges uniformly in L as $L \to \infty$.

By (5.109), and using the fact that $s^\intercal Z = 0$, we have

$$\lim_{m\to\infty} b(L,m) = s^\intercal \left(\sum_{l=0}^{L} d^l (ZA)^l s \right) = s^\intercal s = 2(n-1)$$

for all L.

Since we are thus given uniform convergence in L as well as pointwise convergence in m, we know that both double limits exist and are equal:

$$\lim_{m\to\infty} \lim_{L\to\infty} b(L,m) = \lim_{L\to\infty} \lim_{m\to\infty} b(L,m) = 2(n-1) \tag{5.110}$$

For a proof of this basic result, see Theorem 7.11 in Rudin [Rud76].

We define the sequence $a(m)$ for $m \geq 0$ by

$$a(m) := \lim_{L\to\infty} b(L,m) = s^\intercal \sum_{l=0}^{\infty} d^l v(l,m).$$

Equation (5.110) states that $\lim_{m\to\infty} a(m) = 2(n-1)$.

Assume $a(0) = 0$. Now consider the expression $a(m+1) - a(m)$ for $m \geq 0$. We have

$$\begin{aligned} a(m+1) - a(m) &= s^\intercal \sum_{l=0}^{\infty} d^l v(l,m+1) - s^\intercal \sum_{l=0}^{\infty} d^l v(l,m) \\ &= s^\intercal \sum_{l=0}^{\infty} d^l \left(v(l,m+1) - v(l,m) \right). \end{aligned}$$

For $l \leq m$, we have

$$v(l, m+1) - v(l, m) = 0.$$

For $l > m \geq 0$, we have

$$\begin{aligned} v(l, m+1) - v(l, m) &= A^{l-m-1}(ZA)^m s - A^{l-m}(ZA)^{m-1}s \\ &= A^{l-m-1}(ZA)^{m-1}s - A^{l-m+1}(ZA)^{m-2}s \\ &= A^{l-m-1}(Z - I_{\mathcal{R}^2})A(ZA)^{m-1}s \\ &= -A^{l-m-1}\left(\sum_{\rho\in\mathcal{I}} e_\rho \otimes e_\rho^\intercal\right) A(ZA)^{m-1}s \end{aligned}$$

By symmetry, we have that $e_\rho^\intercal v(l, m) = e_{(1,2),(2,1)}^\intercal v(l, m)$ for all $\rho \in \mathcal{I}$ and all l, m. Consequently,

$$\begin{aligned} v(l, m+1) - v(l, m) &= -A^{l-m-1}\left(\sum_{\rho\in\mathcal{I}} e_\rho \otimes e_{(1,2),(2,1)}^\intercal\right) A(ZA)^{m-1}s \\ &= (-e_{(1,2),(2,1)}^\intercal A(ZA)^{m-1}s) \cdot A^{l-m-1}\sum_{\rho\in\mathcal{I}} e_\rho \\ &= (-e_{(1,2),(2,1)}^\intercal A(ZA)^{m-1}s) \cdot A^{l-m-1}s \\ &= (-e_{(1,2),(2,1)}^\intercal A(ZA)^{m-1}s) \cdot v(l-m-1, 0). \end{aligned}$$

So, for all $m \geq 0$,

$$\begin{aligned} a(m+1) - a(m) &= (-e_{(1,2),(2,1)}^\intercal A(ZA)^{m-1}s) \cdot s^\intercal \sum_{l=m+1}^{\infty} d^l v(l-m-1, 0) \\ &= (-e_{(1,2),(2,1)}^\intercal A(ZA)^{m-1}s) \cdot d^{m+1}s^\intercal \sum_{l=0}^{\infty} d^l v(l, 0) \\ &= (-e_{(1,2),(2,1)}^\intercal A(ZA)^{m-1}s) \cdot d^{m+1}a(0) \\ &= 0 \end{aligned}$$

by our assumption of $a(0) = 0$. So $a(m+1) = a(m)$ for all $m \geq 0$. By induction, $a(m) = a(0) = 0$ for all m. But $a(m) \to 2(n-1)$. This is a contradition. So our assumption of $a(0) = 0$ must be false; $a(0) \neq 0$.

Since

$$a(0) = s^\intercal \sum_{l=0}^{\infty} d^l v(l, 0) = s^\intercal \sum_{l=0}^{\infty} d^l A^l s,$$

$a(0) \neq 0$ is equivalent to (5.108).

So we know that

$$s^\intercal \left(\sum_{\tau=0}^{\infty} d^\tau A^\tau\right) s \neq 0. \tag{5.111}$$

This is true for all $d \in [0, 1)$. For $d = 0$, the left-hand side of (5.111) is equal to $2(n-1)$. By continuity in d on $[0, 1)$, it is positive for all $d \in [0, 1)$. □

A condition for the existence of profitable deviations

By now, we know that if a mutant Player 1 can effect a change in their coordination rate with Player 2 in any given round t, while not affecting any other aspect of the state vector $\ddot{w}(t)$, then they have as a either result increased or decreased their payoff. By linearity, either an increase or a decrease in the coordination rate will effect a payoff increase. The following lemma formally shows this.

So far, we have only considered the case of a population of n players who all use the reactive strategy σ. Starting with the below lemma, we need to evaluate σ against alternative, not necessarily reactive, strategies for a potentially deviating Player 1. For these arbitrary mutant strategies, a priori there is no concept of intents, and there is no natural way to define an analogue of the pairwise intent distribution vectors $(\ddot{w}(t))_t$ that we have in the case of a homogeneous reactive population. But given a sequence of vectors, say $(\ddot{w}'(t))_t$, we can define the concept of action compatibility:

Definition 30 (Action compatibility). *Given an arbitrary strategy profile $(\sigma_i)_{i\in[n]}$ (i.e., an assignment of one strategy to each player), and given a sequence of vectors $(\ddot{w}'(t))_t$ in $\mathbb{R}^{\mathcal{R}^2} \otimes \mathbb{R}^{\mathcal{A}^2}$, we say that $(\ddot{w}'(t))_t$ is action compatible with $(\sigma_i)_{i\in[n]}$ if for all t and all $(i,j) \in \mathcal{R}$, the probability distribution over the pair of actions of Player i and Player j in round t, given that they are selected to play, is equal to $e^{\intercal}_{(i,j),(j,i)}\ddot{w}'(t)$.*

In particular, if some $(\ddot{w}'(t))_t$ is action compatible with some $(\sigma_i)_{i\in[n]}$, then the expected payoff of Player i in round t is given by

$$\pi_i(t) = \frac{1}{n(n-1)} \sum_{j\in[n]\setminus\{i\}} ((e_{(i,j),(j,i)} + e_{(j,i),(i,j)}) \otimes g)^{\intercal} \ddot{w}'(t).$$

With that, we can state the following lemma.

Lemma 6. *Let σ be a given reactive strategy. As always, we write $(\ddot{w}(t))_t$ for the pairwise intent distributions in the homogeneous σ population.*

Given a time $T \in \mathbb{N}$ and a real number γ satisfying $\gamma g^{\intercal}\ddot{z} > 0$, we define the sequence $(\ddot{w}'(t))_t$ in $\mathbb{R}^{\mathcal{R}^2} \otimes \mathbb{R}^{\mathcal{A}^2}$ by

- $\ddot{w}'(t) = \ddot{w}(t)$ *for all* $t < T$
- $\ddot{w}'(T) = \ddot{w}(T) + \gamma\, u \otimes \ddot{z}$
- $\ddot{w}'(t+1) = \ddot{Q}\ddot{w}'(t)$ *for all* $t \geq T$,

If there exist T, γ as above, and a (not necessarily reactive) mutant strategy σ' such that $(\ddot{w}'(t))_t$ is action compatible with the strategy profile in which a mutant Player 1 plays σ' and the residents play σ, then σ is not a Nash equilibrium strategy.

Note that the defining property of additive games can be expressed as $g^{\intercal}\ddot{z} = 0$. So it is already clear that Lemma 6 is only applicable to non-additive games.

Proof. Let σ be a reactive strategy that satisfies the conditions of the lemma. We want to show that σ is not a Nash equilibrium strategy.

Take T, γ, and σ' as in the statement of the lemma. We write $\pi'(t)$ for Player 1's expected payoff in round t when using strategy σ'. Then, by action compatibility of $(\ddot{w}'(t))_t$, we have

$$\pi'(t) = \frac{1}{n(n-1)}(s \otimes g)^\intercal \ddot{w}'(t). \tag{5.112}$$

where $\pi(t)$ is Player 1's expected payoff in round t when using strategy σ.

Writing π' of course for the total expected payoff of Player 1 with σ', we have

$$\begin{aligned}
\pi' &= (1-d)\sum_{t=0}^{\infty} d^t \pi'(t) && (5.113)\\
&= (1-d)\frac{1}{n(n-1)}(s \otimes g)^\intercal \sum_{t=0}^{\infty} d^t \ddot{w}'(t) && (5.114)\\
&= (1-d)\frac{1}{n(n-1)}(s \otimes g)^\intercal \left(\sum_{t=0}^{T-1} d^t \ddot{w}(t) + \sum_{t=T}^{\infty} d^t \ddot{Q}^{t-T} \ddot{w}'(T)\right) && (5.115)\\
&= (1-d)\frac{1}{n(n-1)}(s \otimes g)^\intercal \left(\sum_{t=0}^{T-1} d^t \ddot{w}(t) + \sum_{t=T}^{\infty} d^t \ddot{Q}^{t-T} (\ddot{w}(T) + \gamma\, u \otimes \ddot{z})\right) && (5.116)\\
&= (1-d)\frac{1}{n(n-1)}(s \otimes g)^\intercal \left(\sum_{t=0}^{\infty} d^t \ddot{w}(t) + \gamma \sum_{t=T}^{\infty} d^t \ddot{Q}^{t-T} (u \otimes \ddot{z})\right) && (5.117)\\
&= \pi + (1-d)\frac{1}{n(n-1)}\gamma (s \otimes g)^\intercal \sum_{t=T}^{\infty} d^t \ddot{Q}^{t-T} (u \otimes \ddot{z}). && (5.118)
\end{aligned}$$

By Proposition 14, we therefore have that

$$\begin{aligned}
\pi' - \pi &= (1-d)\frac{1}{n(n-1)}\gamma (s \otimes g)^\intercal \sum_{t=T}^{\infty} d^t \left(A^{t-T} u\right) \otimes \ddot{z} && (5.119)\\
&= (1-d)\frac{1}{n(n-1)} d^T \gamma \left(s^\intercal \sum_{t=0}^{\infty} d^t A^t u\right) \cdot g^\intercal \ddot{z}. && (5.120)
\end{aligned}$$

We denote the term in the big parentheses by b. Proposition 15 states that $b > 0$. We thus have

$$\pi' - \pi = (1-d)\frac{1}{n(n-1)} d^T b \gamma g^\intercal \ddot{z}. \tag{5.121}$$

But also $\gamma g^\intercal \ddot{z} > 0$ by assumption. So $\pi' > \pi$; the mutant strategy σ' achieves a better payoff than σ. So σ is not a Nash equilibrium strategy. □

Constructing profitable deviations

Now that we proved Lemma 6, what remains to show Theorem 13 is to construct appropriate T, γ, and σ' for all cases covered by the theorem. We do this separately for $\lambda < 1$ (in Proposition 16) and $\lambda = 1$ (in Proposition 17).

Proposition 16. *Let $\sigma = (p_0, p, \lambda)$ be a reactive strategy with $p_A \neq p_B$ and $0 < \lambda < 1$ in a game with errors. Then σ is not a Nash equilibrium strategy.*

Proof. It is enough to construct T, γ, and σ' that satisfy the conditions of Lemma 6.

Let $T = n + 3$. We construct σ' as follows.

First, we construct for each $x \in \mathcal{A}$ an event E_x. Both E_A and E_B completely specify the drawn player pairs and their actions in rounds 0 to $T - 1$. The player pairs are independent of x, and the actions other than the action of Player 1 in round 2 are also independent of x. The below table is a visual summary of E_x and is described in detail below. Cells marked with a dash indicate an action that is fixed by E_x, but not explicitly specified in our construction.

Round	0		1		2		3		4		5		6		...	$n+2$	
Players	2	3	1	3	1	2	2	3	1	3	1	3	1	4		1	n
Actions	–	–	–	–	x	y	$\overline{y}$	–	A	–	B	–	–	–		–	–

To construct the sequence of actions in E_A and E_B, we proceed as follows. First, for rounds 0 and 1, choose any actions that have positive probability in a homogeneous σ population. Then, due to the presence of errors, any pair of actions is possible in round 2. Choose some action for Player 2. Player 1 plays x, i.e., action A in E_A and action B in E_B. In round 3, choose any possible action for Player 3. It is possible to do so consistently for E_A and E_B, since the action of Player 1 towards Player 2 in round 2 is not relevant here. For Player 2, choose the action that Player 2 did not use in round 2. This is also certainly possible, since the intent $Y_{(3,2)}(3)$ is potentially ($\lambda > 0$) influenced by the action of Player 3 towards Player 1 in round 1, which might have been observed with errors. In round 4 again, due to the presence of errors, anything is possible. For Player 1, choose action A. For Player 3, choose any action that allows Player 1 to play action B in round 5. There must be at least one possible such choice, since $p_A \neq p_B$. In round 5, choose action B for Player 1 and any action for Player 3. For all i such that $4 \leq i \leq n$, the players $(1, i)$ are selected to play in round $i + 2$. We choose any actions that have positive probability.

Firstly, observe that we have constructed E_A and E_B such that both occur with positive probability when all players use strategy σ. Secondly, fix $x \in \mathcal{A}$ and consider the distribution of the intent of Player 1 towards Player 2 at time $n + 3$ under the condition that E_x occurred. The intent originates either from the action of Player 2 in round 2 or in round 3, which were different. Since $0 < \lambda < 1$, either of those is possible. Since also $p_A \neq p_B$, the intent $Y_{(1,2)}(n + 3)$ is neither certainly A nor certainly B. (Note that this is not due to Player 1's susceptibility to errors.) Write

$$\varphi_x = \mathbb{P}\left[Y_{(1,2)}(n+3) = A \mid E_x\right].$$

Choose some positive real number $\Delta_x > 0$ that is sufficiently small such that

$$\Delta_x \leq \varphi_x \leq 1 - \Delta_x.$$

Let $\Delta = \min_{x \in \mathcal{A}} \Delta_x$.

For both $x \in \mathcal{A}$, we define

$$\psi_x = \mathbb{P}\left[Y_{(2,1)}(n+3) = A \mid E_x\right].$$

Since $0 < \lambda < 1$, there is a positive probability that the intent $Y_{(2,1)}$ was last updated in round 2, when Player 2 directly observed Player 1's action. The probability for this is the same when conditioning on E_A and E_B. If it is the case that the intent was last updated at that time, then the reputations are distributed differently between E_A and E_B (since $\varepsilon < \frac{1}{2}$). If it is not the case, then they are distributed equally between E_A and E_B, since all other actions are the same in E_A and E_B. So overall the distribution at time $n+3$ is different between E_A and E_B, i.e., $\psi_A \neq \psi_B$.

Let $\iota_A \in \{+1, -1\}$ such that ι_A has the same sign as $(\psi_A - \psi_B)g^\intercal \ddot{z}$. Let $\iota_B = -\iota_A$.

Now, the strategy σ' for Player 1 operates as follows. Normally, it behaves exactly like σ. Only after round $n+2$, before round $n+3$ is played, if event E_x occurred for some $x \in \mathcal{A}$, then σ' modifies its intent towards Player 2. It does so by randomly choosing a new value for the intent, with $\varphi_x + \iota_x \Delta \mathbb{P}\left[E_{\bar{x}}\right]$ as the probability for it being A, where $\bar{x}$ is the action that is not x. So the probability of the intent being A, given that E_x occurred, changes from φ_x to $\varphi_x + \iota_x \Delta \mathbb{P}\left[E_{\bar{x}}\right]$.

Let $E = E_A \cup E_B$. We write $Y'_{(1,2)}(n+3)$ for the (potentially, because E need not occur) updated intent. When conditioning on E, the distribution of $Y'_{(1,2)}(n+3)$ is identical to that of $Y_{(1,2)}(n+3)$:

$$\begin{aligned}
\mathbb{P}\left[Y'_{(1,2)}(n+3) = A \mid E\right] &= \mathbb{P}\left[E_A\right]\mathbb{P}\left[Y'_{(1,2)}(n+3) = A \mid E_A\right] \\
&\qquad + \mathbb{P}\left[E_B\right]\mathbb{P}\left[Y'_{(1,2)}(n+3) = A \mid E_B\right] \\
&= \mathbb{P}\left[E_A\right]\left(\varphi_A + \iota_A \Delta \mathbb{P}\left[E_B\right]\right) + \mathbb{P}\left[E_B\right]\left(\varphi_B - \iota_A \Delta \mathbb{P}\left[E_A\right]\right) \\
&= \mathbb{P}\left[E_A\right]\varphi_A + \mathbb{P}\left[E_B\right]\varphi_B \\
&= \mathbb{P}\left[Y_{(1,2)}(n+3) = A \mid E\right]
\end{aligned}$$

We analyse the pairwise distributions $\ddot{Y}'_{(1,2),(i,j)}(n+3)$, where $(i,j) \in \mathcal{R}$, under the condition E. The same statements are true for their redundant analogues $\ddot{Y}'_{(i,j),(1,2)}(n+3)$.

First, consider the case $(i,j) = (1,2)$. It follows from the above paragraph that this pairwise distribution is unchanged compared to $\ddot{Y}$.

Secondly, consider the case that $(i,j) \neq (1,2)$ and $(i,j) \neq (2,1)$. The intent $Y'_{(1,2)}(2)$ is either in its initial state, or was last update in response to an action whose value is identically fixed by E_A and E_B. (The only action for which this is not the case is the action of Player 1 in round 2. But the last time that Player i's intent towards Player j was updated cannot have been in response to that action. If it were so, then $j = 2$ and thus $i \neq 1$ and $i \neq 2$. But then Player i had a direct interaction with Player 1 in round $i+2$.) So $Y'_{(i,j)}(n+3)$ is independent of E_A and E_B under the condition E, and thus also of $Y'_{(1,2)}(n+3)$. The same is true for the pair $Y_{(i,j)}(n+3)$ and $Y_{(1,2)}(n+3)$. So the pairwise distribution is unchanged.

Finally, we consider $(i,j) = (2,1)$. For each $x \in \mathcal{A}$, we have

$$\mathbb{P}\left[Y'_{(1,2)}(n+3) = A \mid E_x\right] = \varphi_x + \iota_x \Delta \mathbb{P}\left[E_{\bar{x}}\right]$$

and

$$\mathbb{P}\left[Y'_{(2,1)}(n+3) = A \mid E_x\right] = \psi_x.$$

Since with E_x we are conditioning on all observable events prior to round T, the intents are again independent. So we have

$$p\left[Y'_{((1,2),(2,1))}(n+3) \mid E_x\right] = \begin{pmatrix} (\varphi_x + \iota_x \Delta \mathbb{P}[E_{\bar{x}}])\psi_x \\ (\varphi_x + \iota_x \Delta \mathbb{P}[E_{\bar{x}}])(1-\psi_x) \\ (1-\varphi_x - \iota_x \Delta \mathbb{P}[E_{\bar{x}}])\psi_x \\ (1-\varphi_x - \iota_x \Delta \mathbb{P}[E_{\bar{x}}])(1-\psi_x) \end{pmatrix}. \tag{5.122}$$

Consequently, we have

$$p\left[Y'_{((1,2),(2,1))}(n+3) \mid E_x\right] - p\left[Y_{((1,2),(2,1))}(n+3) \mid E_x\right] = \iota_x \Delta \mathbb{P}[E_{\bar{x}}] \begin{pmatrix} \psi_x \\ (1-\psi_x) \\ -\psi_x \\ -(1-\psi_x) \end{pmatrix}. \tag{5.123}$$

We can thus state

$$\begin{aligned} & p\left[Y'_{((1,2),(2,1))}(n+3) \mid E\right] - p\left[Y_{((1,2),(2,1))}(n+3) \mid E\right] \\ &\quad = \mathbb{P}[E_A \mid E]\, \iota_A \Delta \mathbb{P}[E_B] \begin{pmatrix} \psi_A \\ (1-\psi_A) \\ -\psi_A \\ -(1-\psi_A) \end{pmatrix} + \mathbb{P}[E_B \mid E]\, \iota_B \Delta \mathbb{P}[E_A] \begin{pmatrix} \psi_B \\ (1-\psi_B) \\ -\psi_B \\ -(1-\psi_B) \end{pmatrix} \\ &\quad = \iota_A \Delta \frac{\mathbb{P}[E_A \mid E]\, \mathbb{P}[E_B \mid E]}{\mathbb{P}[E]} (\psi_A - \psi_B) \ddot{z}. \end{aligned} \tag{5.124}$$

Of course

$$p\left[Y'_\rho(n+3) \mid E^{\mathsf{c}}\right] = p\left[Y_\rho(n+3) \mid E^{\mathsf{c}}\right]$$

for all $\rho \in \mathcal{R}$, since σ' acts identically to σ when E does not occur.

Let

$$\gamma = \iota_A \Delta \mathbb{P}[E_A \mid E]\, \mathbb{P}[E_B \mid E]\, (\psi_A - \psi_B).$$

Then $\gamma g^{\intercal} \ddot{z} > 0$ by construction of ι_A, and

$$\left(p\left[\ddot{Y}'_\rho(n+3)\right] : \rho \in \mathcal{R}^2\right) = \ddot{w}(n+3) + \gamma\, u \otimes \ddot{z}.$$

So we have

$$\left(p\left[\ddot{Y}'_\rho(n+3)\right] : \rho \in \mathcal{R}^2\right) = \ddot{w}'(n+3),$$

for $(\ddot{w}'(t))_t$ as it is defined by T and γ.

Since after time T, σ' plays like σ given the potentially updated intent, we also have

$$\left(p\left[\ddot{Y}'_\rho(t)\right] : \rho \in \mathcal{R}^2\right) = \ddot{w}'(t)$$

for all $t > T$. For $t < T$, the same is trivially given.

So indeed, the strategy profile where Player 1 plays σ' and all other players play σ is action compatible with $(\ddot{w}'(t))_t$ as defined by T and γ. Thus, by Lemma 6, σ is not a Nash equilibrium. □

Proposition 17. *Let $\sigma = (p_0, p, \lambda)$ be a reactive strategy with $p_A \neq p_B$ and $\lambda = 1$ in a game with errors. Then σ is not a Nash equilibrium strategy.*

Proof. Again, it is enough to construct T, γ, and σ' that satisfy the conditions of Lemma 6.

Let this time $T = 3$. We construct σ' as follows.

First, we construct for each $x \in \mathcal{A}$ an event E_x. For both values of $x \in \mathcal{A}$, the event E_x completely specifies the drawn player pairs and their actions in rounds 0 to 2 apart from the action of Player 2 in round 2. The player pairs are independent of x, and the actions other than the action of Player 1 in round 2 are also independent of x. The below table is a visual summary of E_x and is described in detail below. Cells marked with a dash indicate an action that is fixed by E_x, but not explicitly specified in our construction. The cell marked with a question mark is an action that is not fixed by E_x.

Round	0		1		2	
Players	1	3	2	3	1	2
Actions	–	–	–	–	x	?

To construct the actions in E_A and E_B, we proceed as follows. For rounds 0 and 1, choose any actions that have positive probability, identically for E_A and E_B. In round 2, Player 1 plays A in E_A and B in E_B. The action of Player 2 in round 2 is not fixed by E_A and E_B.

Analogously to the proof of Proposition 16, we may observe that E_A and E_B occur with positive probability when all players use strategy σ. We analogously define

$$\varphi_x = \mathbb{P}\left[Y_{(1,2)}(3) = A \mid E_x\right]$$

and

$$\psi_x = \mathbb{P}\left[Y_{(2,1)}(3) = A \mid E_x\right].$$

In round 2, the selected players' intents towards each other certainly stem from indirect observations of the respective other's action towards Player 3 in the previous rounds. Due to the presence of errors, all combinations of reputations and thus, by $p_A \neq p_B$, all action pairs are possible in round 2. In particular, this means that $0 < \varphi_x < 1$ for all $x \in \mathcal{A}$. We have $\psi_A = p_A$ and $\psi_B = p_B$, so $\psi_A \neq \psi_B$ is given.

Like in the proof of Proposition 16, we can thus choose $\Delta > 0$ sufficiently small such that

$$\Delta \leq \varphi_x \leq 1 - \Delta$$

for all x.

We again choose $\iota_A = -\iota_B \in \{+1, -1\}$ such that ι_A has the same sign as $(\psi_A - \psi_B)g^\intercal \ddot{z}$.

In this case, the strategy σ' is defined to operate as follows. Normally, it behaves exactly like σ. However, if after round 2 it determines that event E_x occurred for some $x \in \mathcal{A}$, then it forms an alternative intent towards Player 2, which we denote by $Y^*_{(1,2)}(3)$. It does so by randomly choosing a value, with $\varphi_x + \iota_x \Delta \mathbb{P}\left[E_{\bar{x}}\right]$ as the probability for the alternative intent to be A. Now, the subsequent actions depend on which of two things happen first. If either Player 1 or Player 2 are selected to play with some third player before they are selected to play with each other, then Player 1 forgets the alternative intent and keeps the original intent, and thus continues to play like σ. However, if Player 1 and Player 2 are selected to play with each other in some round $t \geq 3$ before either of them played with someone else, then Player 1 forgets the original intent towards Player 2 and enacts the alternative intent towards Player 2 in that round.

Let R be the event that the two players interact with each other before either of them interacts with someone else. We have $\mathbb{P}[R] > 0$. We define

$$\gamma = \iota_A \Delta \mathbb{P}[E_A \mid E] \mathbb{P}[E_B \mid E] \mathbb{P}[R](\psi_A - \psi_B).$$

It remains to show that the strategy profile given by σ' for Player 1 and σ for the residents is action compatible with $(\ddot{w}'(t))_t$ as defined by T and γ.

We analyse the distribution of the alternative intent $Y^*_{(1,2)}(3)$ under the condition E, and the pairwise joint distributions with other intents. We know from Proposition 14 that the pairwise joint distributions of intents are collectively sufficient to infer the distribution of all future actions.

Firstly, by an analogous calculation as in the proof of Proposition 16, the alternative intent is distributed identically to the original intent. Intents of the form $Y_{(i,j)}(3)$ with $j \neq 1$ and $j \neq 2$ are independent of both Player 1's original intent $Y_{(1,2)}(3)$ and the alternative intent $Y^*_{(1,2)}(3)$, since we are conditioning with E on all observable events so far other than the actions of Player 1 and Player 2 in round 2.

The joint pairwise distributions of the alternative intent $Y^*_{(1,2)}(3)$ together with intents of the form $Y_{(i,j)}(3)$, where either $i \neq 1$ and $j = 2$, or $i \neq 2$ and $j = 1$, are irrelevant for all actions, because if R occurs, then such an intent $Y_{(i,j)}(3)$ will certainly ($\lambda = 1$) be replaced wihtout ever being enacted, and if R does not occur, then Player 1 uses the original intent and not the alternative. For the same reason, the joint distribution of the original intent with Player 2's intent towards Player 1, $Y_{(2,1)}(3)$, is also irrelevant.

So the actions of the players from round 3 onwards are as if Player 1 were a normal σ player whose intent towards Player 2, say $Y'_{(2,1)}(3)$, is distributed exactly as it would be with a homogeneous σ population, except that in the case that E occurs, the joint distribution between $Y'_{(1,2)}(3)$ and $Y'_{(2,1)}(3)$ is given by

$$p\left[Y'_{((1,2),(2,1))}(3) \mid E\right] = p\left[Y_{((1,2),(2,1))}(3) \mid E\right] + \iota_A \Delta \mathbb{P}[E_A \mid E] \mathbb{P}[E_B \mid E] \mathbb{P}[R](\psi_A - \psi_B)\ddot{z}.$$

Of course, σ' behaves identically to σ prior to round 3.

So indeed, the strategy profile where Player 1 plays σ' and all other players play σ is action compatible with $(\ddot{w}'(t))_t$ as defined by T and γ. Thus, by Lemma 6, σ is not a Nash equilibrium. □

The statement of Theorem 13 is simply the conjunction of Proposition 16 and Proposition 17.

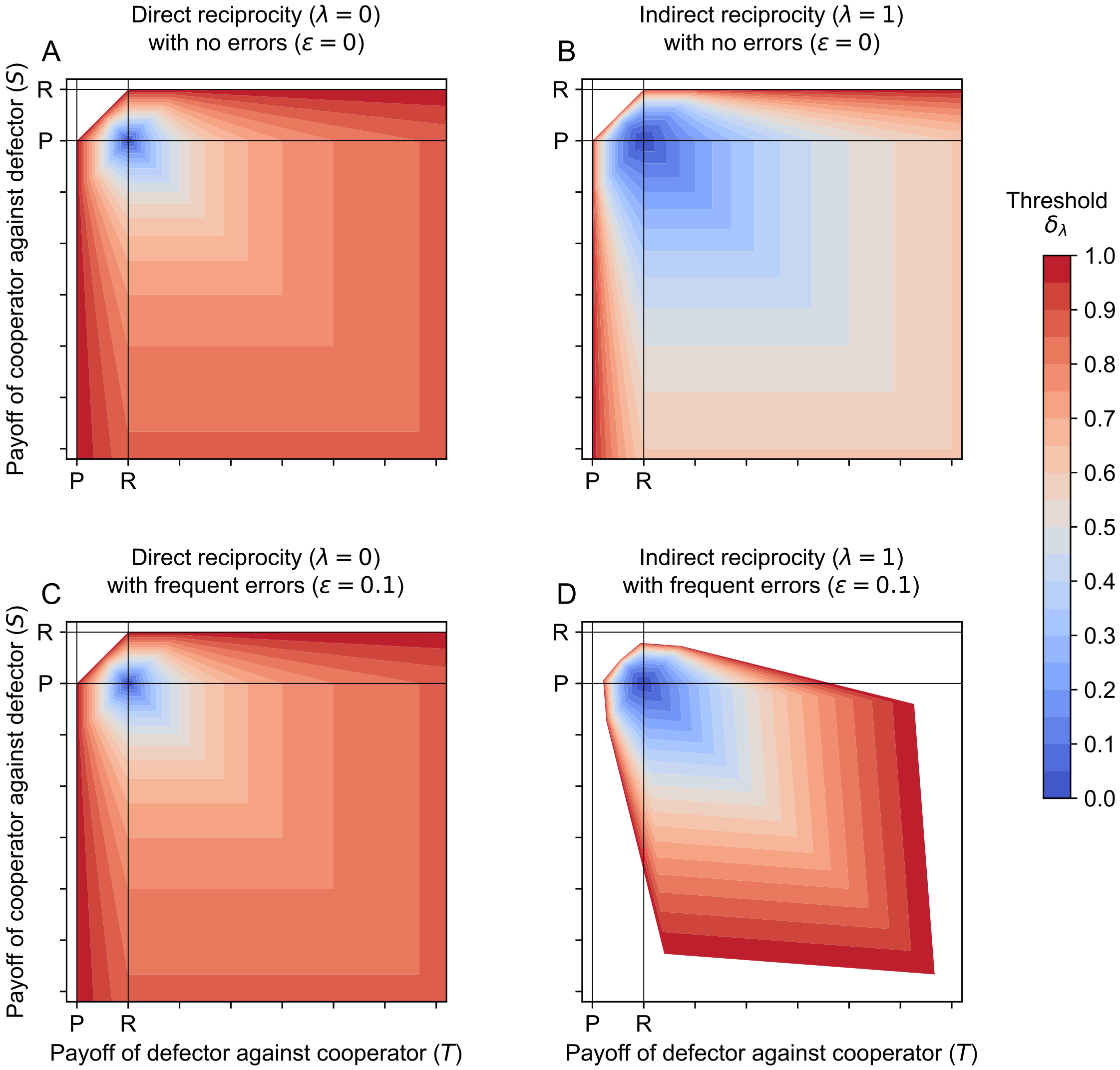

Figure 5.6: An expanded view on Fig. 5.4. For the four combinations of pure direct reciprocity ($\lambda = 0$) and pure indirect reciprocity ($\lambda = 1$) with no errors ($\varepsilon = 0$) and with frequent errors ($\varepsilon = 0.1$), we show the largest δ for which equalizers exist, depending on the payoff matrix G. The parameters are the same as in Fig. 5.4, but we show a wider section of the payoff matrix space. We see that in the case $\lambda = 1$ and $\varepsilon > 0$, for fixed P and R, the area in the T-S-plane where equalizers exist is bounded. But in the cases where $\lambda = 0$ or $\varepsilon = 0$, equalizer strategies exist in all social dilemmas for sufficiently large δ, meaning that the coloured area extends infinitely.

Data and software availability

The data shown in Figure 5.5 were generated in a computer simulation written in Rust (compiled with rustc 1.78) and Python 3.10. The computer code and the obtained simulation data are available in [H2̈4a].

Author contribution statement

All authors conceived and discussed the study; V.H. analyzed the model; C.H., V.H., and

K.C. wrote the main text manuscript; V.H. wrote the Supplementary Information text and conducted simulations; all authors discussed the results and edited both texts. C.H. and K.C. contributed equally to this work.

Bibliography

[Abr88] Dilip Abreu. On the theory of infinitely repeated games with discounting. *Econometrica*, 56(2):383–396, 1988.

[ACBS+18] Maria Abou Chakra, Silke Bumann, Hanna Schenk, Andreas Oschlies, and Arne Traulsen. Immediate action is the best strategy when facing uncertain climate change. *Nature Communications*, 9(1):2566, 2018.

[ACT14] Maria Abou Chakra and Arne Traulsen. Under high stakes and uncertainty the rich should lend the poor a helping hand. *Journal of Theoretical Biology*, 341:123–130, 2014.

[Ado99] Ralph Adolphs. Social cognition and the human brain. *Trends in Cognitive Sciences*, 3(12):469–479, December 1999.

[ADS94] Dilip Abreu, Prajit K. Dutta, and Lones Smith. The Folk Theorem for Repeated Games: A NEU Condition. *Econometrica*, 62(4):939–48, 1994.

[AH81] R. Axelrod and W. D. Hamilton. The evolution of cooperation. *Science*, 211(4489):1390–1396, 1981.

[Aki15] Ethan Akin. What you gotta know to play good in the iterated prisoner's dilemma. *Games*, 6(3):175–190, 2015.

[Aki16] Ethan Akin. *The iterated Prisoner's Dilemma: good strategies and their dynamics*, pages 77–107. De Gruyter, June 2016.

[Ale17] Richard Alexander. *The biology of moral systems*. Routledge, 2017.

[AS11] M. Archetti and I. Scheuring. Coexistence of cooperation and defection in public goods games. *Evolution: International Journal of Organic Evolution*, 65(4):1140–1148, 2011. Publisher: Wiley Online Library.

[Axe81] R. Axelrod. The emergence of cooperation among egoists. *American political science review*, 75(2):306–318, 1981.

[BBV86] Theodore Bergstrom, Lawrence Blume, and Hal Varian. On the private provision of public goods. *Journal of Public Economics*, 29(1):25–49, 1986.

[BC06] Edward Buckley and Rachel Croson. Income and wealth heterogeneity in the voluntary provision of linear public goods. *Journal of Public Economics*, 90(4-5):935–955, 2006.

[BDV+20] Wolfram Barfuss, Jonathan F Donges, Vítor V Vasconcelos, Jürgen Kurths, and Simon A Levin. Caring for the future can turn tragedy into comedy for long-term collective action under risk of collapse. *Proceedings of the National Academy of Sciences*, 117(23):12915–12922, 2020.

[Ben00] Roland Benabou. Unequal societies: Income distribution and the social contract. *American Economic Review*, 91(1):96–129, 2000.

[BG16] U. Berger and A. Grüne. On the stability of cooperation under indirect reciprocity with first-order information. *Games and Economic Behavior*, 98:19–33, 2016.

[BJHN16] Seung Ki Baek, Hyeong-Chai Jeong, Christian Hilbe, and Martin A Nowak. Comparing reactive and memory-one strategies of direct reciprocity. *Scientific Reports*, 6(1):1–13, 2016.

[BL87] Robert Boyd and Jeffrey P. Lorberbaum. No pure strategy is evolutionarily stable in the repeated prisoner's dilemma game. *Nature*, 327(6117):58–59, May 1987.

[BN93] Abhijit V Banerjee and Andrew F Newman. Occupational choice and the process of development. *Journal of Political Economy*, 101(2):274–298, 1993.

[BNS97] M. C. Boerlijst, M. A. Nowak, and K. Sigmund. Equal pay for all prisoners. *American Mathematical Monthly*, 104:303–307, 1997.

[BP94] Abraham. Berman and Robert J. Plemmons. *Nonnegative matrices in the mathematical sciences*. Society for Industrial and Applied Mathematics, 1994.

[BR88a] Robert Boyd and Peter J. Richerson. *Culture and the Evolutionary Process*. University of Chicago Press, Chicago, IL, June 1988.

[BR88b] Robert Boyd and Peter J. Richerson. The evolution of reciprocity in sizable groups. *Journal of Theoretical Biology*, 132(3):337–356, June 1988.

[BR13] M. Broom and J. Rychtář. *Game-Theoretical Models in Biology*. Chapman and Hall / CRC Press, London, 2013.

[BRH11] Robert Boyd, Peter J. Richerson, and Joseph Henrich. The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, 108(supplement2):10918–10925, June 2011.

[BS04] H. Brandt and K. Sigmund. The logic of reprobation: assessment and action rules for indirect reciprocation. *J. Theor. Biol.*, 231, 2004.

[Cam03] Colin F Camerer. *Behavioral game theory*. The Roundtable Series in Behavioral Economics. Princeton University Press, Princeton, NJ, February 2003.

[CB01] Valerie Curtis and Adam Biran. Dirt, disgust, and disease: Is hygiene in our genes? *Perspectives in Biology and Medicine*, 44(1):17–31, December 2001.

[CFW20] Daniel Clark, Drew Fudenberg, and Alexander Wolitzky. Indirect reciprocity with simple records. *Proc. Natl. Acad. Sci. USA*, 117(21):11344–11349, 2020.

[CGH22] M. C. Couto, S. Giaimo, and C. Hilbe. Introspection dynamics: a simple model of counterfactual learning in asymmetric games. *New Journal of Physics*, 24(6):063010, June 2022. Publisher: IOP Publishing.

[CKS05] Todd L Cherry, Stephan Kroll, and Jason F Shogren. The impact of endowment heterogeneity and origin on public good contributions: evidence from the lab. *Journal of Economic Behavior & Organization*, 57(3):357–365, 2005.

[CMMM99] Kenneth S Chan, Stuart Mestelman, Robert Moir, and R Andrew Muller. Heterogeneity and the voluntary provision of public goods. *Experimental Economics*, 2(1):5–30, 1999.

[CP23] Marta Couto and Saptarshi Pal. Introspection Dynamics in Asymmetric Multiplayer Games. *Dynamic Games and Applications*, 2023. Publisher: Springer.

[CT13] Leda Cosmides and John Tooby. Unraveling the enigma of human intelligence: Evolutionary psychology and the multimodular mind. In *The evolution of intelligence*, pages 145–198. Psychology Press, 2013.

[Dar09] Charles Darwin. *The Origin of Species: By Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life.* Cambridge University Press, July 2009.

[DBF18] Pedro Dal Bó and Guillaume R Fréchette. On the determinants of cooperation in infinitely repeated games: A survey. *Journal of Economic Literature*, 56(1):60–114, 2018.

[DD22] Ryan J Dwyer and Elizabeth W Dunn. Wealth redistribution promotes happiness. *Proceedings of the National Academy of Sciences*, 119(46):e2211123119, 2022.

[DH05] Michael Doebeli and Christoph Hauert. Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game. *Ecology letters*, 8(7):748–766, 2005. Publisher: Wiley Online Library.

[Die85] A. Diekmann. Volunteer's dilemma. *Journal of Conflict Resolution*, 29:605–610, 1985.

[DLTZ15] Yali Dong, Cong Li, Yi Tao, and Boyu Zhang. Evolution of conformity in social dilemmas. *PLoS One*, 10(9):e0137435, 2015.

[DSS73] Partha Dasgupta, Amartya Sen, and David Starrett. Notes on the measurement of inequality. *Journal of Economic Theory*, 6(2):180–187, April 1973.

[DVS06] Frederique De Vignemont and Tania Singer. The empathic brain: how, when and why? *Trends in cognitive sciences*, 10(10):435–441, 2006.

[FAO24] Food and Agriculture Organization of the United Nations. Food and balance sheets 2010–2022 – global, regional and country trends. Technical report, FAOSTAT Analytical Brief Series, Rome, 2024.

[FGF01] U. Fischbacher, S. Gächter, and E. Fehr. Are people conditionally cooperative? evidence from a public goods experiment. *Econ. Lett.*, 71, 2001.

[Fis30] R. A. Fisher. *The genetical theory of natural selection: a complete variorum edition*. Oxford University Press, UK, 1930.

[FM86a] D. Fudenberg and E. Maskin. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54(3):533–554, 1986.

[FM86b] Drew Fudenberg and Eric Maskin. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54(3):533–554, 1986.

[Fra98] Steven A Frank. *Foundations of social evolution*, volume 19. Princeton University Press, 1998.

[Fre94] Marcus R Frean. The prisoner's dilemma without synchrony. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 257(1348):75–79, 1994.

[Fri71] J. Friedman. A non-cooperative equilibrium for supergames. *Review of Economic Studies*, 38:1–12, 1971.

[FT91] Drew Fudenberg and Jean Tirole. *Game Theory*. The MIT Press. MIT Press, London, England, August 1991.

[Fuk03] Francis Fukuyama. *Our posthuman future*. Profile Books, London, England, April 2003.

[FWR06] Kevin R Foster, Tom Wenseleers, and Francis LW Ratnieks. Kin selection is the key to altruism. *Trends in ecology & evolution*, 21(2):57–60, 2006.

[Gäc15] Simon Gächter. Visible inequality breeds more inequality. *Nature*, 526(7573):333–334, 2015.

[GANH24] N.E. Glynatsi, E. Akin, M. A. Nowak, and C. Hilbe. Conditional cooperation with longer memory. *Proceedings of the National Academy of Sciences USA*, 121:e2420125121, 2024.

[GFGGvV24] C. Graser, T. Fujiwara-Greve, J. García, and M. van Veelen. Repeated games with partner choice. *Tinbergen Institute Discussion Paper*, pages TI 2024–038/I, 2024.

[GK21] N.E. Glynatsi and V.A. Knight. A bibliometric study of research topics, collaboration and centrality in the field of the Iterated Prisoner's Dilemma. *Humanities and Social Sciences Communications*, 8:45, 2021.

[GMTV17] Simon Gächter, Friederike Mengel, Elias Tsakas, and Alexander Vostroknutov. Growth and inequality in public good provision. *Journal of Public Economics*, 150:1–13, 2017.

[GT12] Julián García and Arne Traulsen. The structure of mutations and the evolution of cooperation. *PloS one*, 7(4):e35287, 2012.

[GV18] J. García and M. Veelen. No strategy can win in the repeated prisoner's dilemma: linking game theory and computer simulations. *Front. Robot. AI*, 5, 2018.

[GVDDVD20] Jörg Gross, Sonja Veistola, Carsten KW De Dreu, and Eric Van Dijk. Self-reliance crowds out group cooperation and increases wealth inequality. *Nature Communications*, 11(1):5161, 2020.

[GWB04] Ashleigh S. Griffin, Stuart A. West, and Angus Buckling. Cooperation and competition in pathogenic bacteria. *Nature*, 430(7003):1024–1027, August 2004.

[GZ93] Oded Galor and Joseph Zeira. Income distribution and macroeconomics. *The Review of Economic Studies*, 60(1):35–52, 1993.

[H2̈4a] Valentin Hübner. Computer code and simulation data for "stable strategies of direct and indirect reciprocity across all social dilemmas", 2024.

[H2̈4b] Valentin Hübner. Computer code used to compute optimal contribution sequences and other data for the research article "time-dependent strategies in repeated asymmetric public goods games", 2024.

[Ham64] W. Hamilton. The genetical evolution of social behaviour. II. *Journal of Theoretical Biology*, 7(1):17–52, 1964.

[Har68] G. Hardin. The tragedy of the commons. *Science*, 162:1243–1248, 1968.

[HCN18] Christian Hilbe, Krishnendu Chatterjee, and Martin A Nowak. Partners and rivals in direct reciprocity. *Nature Human Behaviour*, 2(7):469–477, 2018.

[HG23] Ashley Harrell and Anna S Greenleaf. Resource asymmetry reduces generosity and paying forward generosity, among the resource-advantaged and disadvantaged. *Social Science Research*, 109:102786, 2023.

[HHCN19] Oliver P. Hauser, Christian Hilbe, Krishnendu Chatterjee, and Martin A. Nowak. Social dilemmas among unequals. *Nature 2019 572:7770*, 572(7770):524–527, 8 2019.

[HHS+25] Valentin Hübner, Christian Hilbe, Manuel Staab, Maria Kleshnina, and Krishnendu Chatterjee. Time-dependent strategies in repeated asymmetric public goods games. *Dynamic Games and Applications*, February 2025.

[HHV02] Shaun Hargreaves-Heap and Yanis Varoufakis. *Game Theory: A Critical Introduction.* Routledge, London, England, September 2002.

[HK24] Valentin Hübner and Maria Kleshnina. Computer code for "efficiency and resilience of cooperation in asymmetric social dilemmas", 2024.

[HL95] Karla Hoff and Andrew B Lyon. Non-leaky buckets: Optimal redistributive taxation and agency costs. *Journal of Public Economics*, 58(3):365–390, 1995.

[HMD+11] Peter Hajek, Katie Myers, Al-Rehan Dhanji, Oliver West, and Hayden McRobbie. Weight change during and after ramadan fasting. *Journal of Public Health*, 34(3):377–381, November 2011.

[HMS96] Elizabeth Hoffman, Kevin McCabe, and Vernon L. Smith. Social distance and other-regarding behavior in dictator games. *The American Economic Review*, 86(3):653–660, 1996.

[HNT13] C. Hilbe, M. A. Nowak, and A. Traulsen. Adaptive dynamics of exortion and compliance. *PLoS One*, 8:e77886, 2013.

[Hol73] Crawford S Holling. Resilience and stability of ecological systems. *Annual review of ecology and systematics*, 4(1):1–23, 1973.

[HRS16] Shaun P Hargreaves Heap, Abhijit Ramalingam, and Brock V Stoddard. Endowment inequality in public goods games: A re-examination. *Economics Letters*, 146:4–7, 2016.

[HS97] C. Hauert and H. G. Schuster. Effects of increasing the number of players and memory size in the iterated prisoner's dilemma: a numerical approach. *Proc. R. Soc. B*, 264, 1997.

[HSH+24] Valentin Hübner, Manuel Staab, Christian Hilbe, Krishnendu Chatterjee, and Maria Kleshnina. Efficiency and resilience of cooperation in asymmetric social dilemmas. *Proceedings of the National Academy of Sciences*, 121(10), February 2024.

[HSHC25] Valentin Hübner, Laura Schmid, Christian Hilbe, and Krishnendu Chatterjee. Stable strategies of direct and indirect reciprocity across all social dilemmas. *PNAS Nexus*, May 2025.

[HST+18] C. Hilbe, L. Schmid, J. Tkadlec, K. Chatterjee, and M. A. Nowak. Indirect reciprocity with private, noisy, and incomplete information. *Proc. Natl Acad. Sci. USA*, 115, 2018.

[HTN14] Oliver P Hauser, Arne Traulsen, and Martin A Nowak. Heterogeneity in background fitness acts as a suppressor of selection. *Journal of Theoretical Biology*, 343:178–185, 2014. Publisher: Elsevier.

[HTS15] C. Hilbe, A. Traulsen, and K. Sigmund. Partners or rivals? Strategies for the iterated prisoner's dilemma. *Games and Economic Behavior*, 92:41–52, 2015.

[Hum78] David Hume. *David Hume: A Treatise of Human Nature (Second Edition)*. Oxford University Press, November 1978.

[IM18] Genki Ichinose and Naoki Masuda. Zero-determinant strategies in finitely repeated games. *Journal of Theoretical Biology*, 438:61–77, 2018.

[IN10] Lorens A Imhof and Martin A Nowak. Stochastic evolutionary dynamics of direct reciprocity. *Proceedings of the Royal Society B: Biological Sciences*, 277(1680):463–468, 2010.

[JBRH16] Meike Janssen, Claudia Busch, Manika Rödiger, and Ulrich Hamm. Motives of consumers following a vegan diet and their attitudes towards animal agriculture. *Appetite*, 105:643–651, October 2016.

[KDK99] Timothy Killingback, Michael Doebeli, and Nancy Knowlton. Variable investment, the continuous prisoner's dilemma, and the origin of cooperation. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 266(1430):1723–1728, 1999.

[KGSF04] Benjamin Kerr, Peter Godfrey-Smith, and Marcus W Feldman. What is altruism? *Trends in Ecology & Evolution*, 19(3):135–140, 2004.

[KHv+23] M. Kleshnina, C. Hilbe, S. Šimsa, K. Chatterjee, and M.A. Nowak. The effect of environmental information on evolution of cooperation in stochastic games. *Nature Communications*, 14:4153, 2023.

[KI09] Shun Kurokawa and Yasuo Ihara. Emergence of cooperation in public goods games. *Proceedings of the Royal Society B: Biological Sciences*, 276(1660):1379–1384, January 2009.

[KKT86] Daniel Kahneman, Jack L. Knetsch, and Richard Thaler. Fairness as a constraint on profit seeking: Entitlements in the market. *The American Economic Review*, 76(4):728–741, 1986.

[KMS17] Claudia Keser, Andreas Markstädter, and Martin Schmidt. Mandatory minimum contributions, heterogeneous endowments and voluntary public-good provision. *Games and Economic Behavior*, 101:291–310, 2017.

[Kur16] Shun Kurokawa. Evolution of cooperation: The analysis of the case wherein a different player has a different benefit and a different cost. *Letters on Evolutionary Behavioral Science*, 7(2), November 2016.

[KWI10] Shun Kurokawa, Joe Yuichiro Wakano, and Yasuo Ihara. Generous cooperators can outperform non-generous cooperators when replacing a population of defectors. *Theoretical Population Biology*, 77(4):257–262, 2010.

[Lan02] Serge Lang. *Algebra*. Springer New York, 2002.

[LH01] O. Leimar and P. Hammerstein. Evolution of cooperation through indirect reciprocity. *Proc. R. Soc. B*, 268, 2001.

[LM23] Olof Leimar and John M McNamara. Game theory in biology: 50 years and onwards. *Philosophical Transactions of the Royal Society B*, 378(1876):20210509, 2023. Publisher: The Royal Society.

[Lor94] J. M. D. Lorberbaum. No strategy is evolutionary stable in the repeated Prisoner's Dilemma. *Journal of Theoretical Biology*, 168:117–130, 1994.

[Lou81] Glenn C Loury. Intergenerational transfers and the distribution of earnings. *Econometrica: Journal of the Econometric Society*, pages 843–867, 1981.

[MBB16] Gert Mensink, Clarissa Lage Barbosa, and Anna-Kristin Brettschneider. Prevalence of persons following a vegetarian diet in germany. In *Journal of Health Monitoring*, volume 1. Robert Koch-Institut, Epidemiologie und Gesundheitsberichterstattung, 2016.

[MBE13] L. Molleman, E. Broek, and M. Egas. Personal experience and reputation interact in human decisions to help reciprocally. *Proc. R. Soc. B*, 28, 2013.

[Mey00] C. D. Meyer. *Matrix analysis and applied linear algebra.* Society for Industrial and Applied Mathematics, Philadelphia, 2000.

[MH15] A. McAvoy and C. Hauert. Asymmetric evolutionary games. *PLoS Computational Biology*, 11(8):e1004349, 2015.

[MH16a] Alex McAvoy and Christoph Hauert. Autocratic strategies for iterated games with arbitrary action spaces. *Proceedings of the National Academy of Sciences*, 113(13):3573–3578, 2016.

[MH16b] Alex McAvoy and Christoph Hauert. Autocratic strategies for iterated games with arbitrary action spaces. *Proceedings of the National Academy of Sciences*, 113(13):3573–3578, 2016.

[MH23] Yohsuke Murase and Christian Hilbe. Indirect reciprocity with stochastic and dual reputation updates. *PLOS Computational Biology*, 19(7):e1011271, 2023.

[MH24] Y. Murase and C. Hilbe. Computational evolution of social norms in well-mixed and group-structured populations. *Proceedings of the National Academy of Sciences USA*, 121:e2406885121, 2024.

[MIF+16] Fabrizia Minandri, Francesco Imperi, Emanuela Frangipani, Carlo Bonchi, Daniela Visaggio, Marcella Facchini, Paolo Pasquali, Alessandra Bragonzi, and Paolo Visca. Role of iron uptake systems in pseudomonas aeruginosa virulence and airway infection. *Infection and Immunity*, 84(8):2324–2335, August 2016.

[MMKS+24] S. Michel-Mata, M. Kawakatsu, J. Sartini, T.A. Kessinger, J.B. Plotkin, and C. E. Tarnita. The evolution of private reputations in information-abundant landscapes. *Nature*, 634:883–889, 2024.

[Mol85] P. Molander. The optimal level of generosity in a selfish, uncertain environment. *Journal of Conflict Resolution*, 29:611–618, 1985.

[MPGFDL22] E. Montero-Porras, J. Grujić, E. Fernández Domingos, and T. Lenaerts. Inferring strategies from observations in long iterated prisoner's dilemma experiments. *Scientific Reports*, 12:7589, 2022.

[MRH21] A. McAvoy, A. Rao, and C. Hauert. Intriguing effects of selection intensity on the evolution of prosocial behaviors. *PLoS Computational Biology*, 17:e1009611, 2021.

[MS82] John Maynard Smith. *Evolution and the Theory of Games*. Cambridge University Press, October 1982.

[MSA20] D. Melamed, B. Simpson, and J. Abernathy. The robustness of reciprocity: Experimental evidence that each form of reciprocity is robust to the presence of other forms of reciprocity. *Science Advances*, 6(23):eaba0504, 2020.

[MSMP22] David Melamed, Brent Simpson, Bradley Montgomery, and Vedang Patel. Inequality and cooperation in social networks. *Scientific Reports*, 12(1):6789, 2022.

[MSMS21] Ramona Merhej, Fernando P Santos, Francisco S Melo, and Francisco C Santos. Cooperation between independent reinforcement learners under wealth inequality and collective risks. In *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*, pages 898–906, 2021.

[MSP73] J. Maynard Smith and G. R. Price. The logic of animal conflict. *Nature*, 246:15–18, 1973.

[MVCS12] Luis A. Martinez-Vaquero, José A. Cuesta, and Angel Sánchez. Generosity pays in the presence of direct reciprocity: A comprehensive study of 2×2 repeated games. *PLoS ONE*, 7(4):e35135, 2012.

[MW22] Lucía Macchia and Ashley V Whillans. The link between income, income inequality, and prosocial behavior around the world. *Social Psychology*, 2022.

[N$^+$12] Martin A Nowak et al. Evolving cooperation. *Journal of Theoretical Biology*, 299(0):1–8, 2012.

[NAKP20] Laila Nockur, Laetitia Arndt, Johannes Keller, and Stefan Pfattheicher. Collective choice fosters sustainable resource management in the presence of asymmetric opportunities. *Scientific Reports*, 10(1):10724, 2020.

[Nas50] John F. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49, January 1950.

[NO14] Mitsuhiro Nakamura and Hisashi Ohtsuki. Indirect reciprocity in three types of social dilemmas. *Journal of Theoretical Biology*, 355:117–127, 2014.

[Now06a] M. A. Nowak. *Evolutionary dynamics: exploring the equations of life*. The Belknap press of Harvard University press, UK, 2006.

[Now06b] M. A. Nowak. Five rules for the evolution of cooperation. *Science*, 314(5805):1560–1563, 2006.

[NS92] Martin A Nowak and Karl Sigmund. Tit for tat in heterogeneous populations. *Nature*, 355(6357):250–253, 1992.

[NS93] M. A. Nowak and K. Sigmund. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*, 364:56–58, 1993.

[NS95] Martin A. Nowak and Karl Sigmund. Invasion dynamics of the finitely repeated prisoner's dilemma. *Games and Economic Behavior*, 11(2):364–390, 1995.

[NS98a] M. A. Nowak and K. Sigmund. The dynamics of indirect reciprocity. *Journal of Theoretical Biology*, 194:561–574, 1998.

[NS98b] M. A. Nowak and K. Sigmund. Evolution of indirect reciprocity by image scoring. *Nature*, 393, 1998.

[NS05] M. A. Nowak and K. Sigmund. Evolution of indirect reciprocity. *Nature*, 437:1291–1298, 2005.

[NSRC15] Akihiro Nishi, Hirokazu Shirado, David G Rand, and Nicholas A Christakis. Inequality and visibility of wealth in experimental social networks. *Nature*, 526(7573):426–429, 2015.

[OI04] H. Ohtsuki and Y. Iwasa. How should we define goodness? – reputation dynamics in indirect reciprocity. *J. Theor. Biol.*, 231, 2004.

[Oka20] I. Okada. A review of theoretical studies on indirect reciprocity. *Games*, 11(3):27, 2020.

[Ols71] M. Olson. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press, 1971.

[OYS+18] I. Okada, H. Yamamoto, Y. Sato, S. Uchida, and T. Sasaki. Experimental evidence of selective inattention in reputation-based cooperation. *Sci. Rep.*, 8, 2018.

[PD12] W. H. Press and F. D. Dyson. Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proc. Natl Acad. Sci. USA*, 109, 2012.

[PdW02] Stephanie D. Preston and Frans B. M. de Waal. Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25(1):1–20, February 2002.

[Pew12] Pew Research Center. The world's muslims: Unity and diversity. Technical report, Washington, D.C., August 2012.

[PGGS+13] M. Perc, J. Gómez-Gardeñes, A. Szolnoki, L. M. Floría, and Y. Moreno. Evolutionary dynamics of group interactions on structured populations: A review. *Journal of The Royal Society Interface*, 10(80):20120997, 2013.

[PHG24] S. Pal, C. Hilbe, and N.E. Glynatsi. The co-evolution of direct, indirect, and generalized reciprocity. *arXiv*, 11:2411.03488, 2024.

[PHRZ15] Liming Pan, Dong Hao, Zhihai Rong, and Tao Zhou. Zero-determinant strategies in iterated public goods game. *Scientific Reports*, 5(1):1–10, 2015.

[Pik97] Thomas Piketty. The dynamics of the wealth distribution and the interest rate with credit rationing. *The Review of Economic Studies*, 64(2):173–189, 1997.

[Pin10] Steven Pinker. The cognitive niche: Coevolution of intelligence, sociality, and language. *Proceedings of the National Academy of Sciences*, 107(supplement_2):8993–8999, 2010.

[PSSS09] J. M. Pacheco, F. C. Santos, M. O. Souza, and B. Skyrms. Evolutionary dynamics of collective action in n-person stag hunt dilemmas. *Proceedings of the Royal Society B*, 276:315–321, 2009.

[RD00] Colin Ratledge and Lynn G Dover. Iron metabolism in pathogenic bacteria. *Annual Review of Microbiology*, 54(1):881–941, October 2000.

[RH23] C.S.L. Rossetti and C. Hilbe. Direct reciprocity among humans. *Ethology*, 130:e13407, 2023.

[RHR+18] Johannes G Reiter, Christian Hilbe, David G Rand, Krishnendu Chatterjee, and Martin A Nowak. Crosstalk in concurrent repeated games impedes direct reciprocity and requires stronger levels of forgiveness. *Nature Communications*, 9(1):555, 2018.

[Ric05] Peter J Richerson. *Not by genes alone*. University of Chicago Press, Chicago, IL, January 2005.

[Rid03] Matt Ridley. *Nature via nurture: Genes, experience, and what makes us human.* HarperCollins Publishers, 2003.

[RN13] David G. Rand and Martin A. Nowak. Human cooperation. *Trends in Cognitive Sciences*, 17(8):413–425, 8 2013.

[Rob07] G. Roberts. Evolution of direct and indirect reciprocity. *Proc. R. Soc. B*, 275, 2007.

[RR03] Paul B Rainey and Katrina Rainey. Evolution of cooperation and conflict in experimental bacterial populations. *Nature*, 425(6953):72–74, 2003.

[RSP19] A. L Radzvilavicius, A. J. Stewart, and J. B. Plotkin. Evolution of empathetic moral evaluation. *eLife*, 8:e44269, 2019.

[Rud76] Walter Rudin. *Principles of mathematical analysis (int'l ed).* McGraw-Hill Professional, New York, NY, 3 edition, September 1976.

[SBM+09] John P Spencer, Mark S Blumberg, Bob McMurray, Scott R Robinson, Larissa K Samuelson, and J Bruce Tomblin. Short arms and talking eggs. 2009.

[SCHN21] Laura Schmid, Krishnendu Chatterjee, Christian Hilbe, and Martin A. Nowak. A unified framework of direct and indirect reciprocity. *Nature Human Behaviour*, 5(10):1292–1302, 2021.

[Sel88] Reinhard Selten. *Reexamination of the perfectness concept for equilibrium points in extensive games*. Springer, 1988.

[SHCN22] Laura Schmid, Christian Hilbe, Krishnendu Chatterjee, and Martin A Nowak. Direct reciprocity between individuals that use different strategy spaces. *PLoS Computational Biology*, 18(6):e1010149, 2022.

[Sig10] K. Sigmund. *The Calculus of Selfishness*. Princeton Univ. Press, 2010.

[Sig12] K. Sigmund. Moral assessment in indirect reciprocity. *Journal of theoretical biology*, 299:25–30, 2012.

[Sky03] B. Skyrms. *The Stag-Hunt Game and the Evolution of Social Structure*. Cambridge University Press, Cambridge, 2003.

[Slo17] Michael Slote. *Empathy as an Instinct*, page 133–141. Cambridge University Press, December 2017.

[SMB20] Alexander J Stewart, Nolan McCarty, and Joanna J Bryson. Polarization under rising inequality and economic decline. *Science Advances*, 6(50):eabd4201, 2020.

[SN23] Joshua C Skewes and Laila Nockur. National inequality, social capital, and public goods decision-making. *Current Research in Ecological and Social Psychology*, page 100112, 2023.

[SP13] A. J. Stewart and J. B. Plotkin. From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proceedings of the National Academy of Sciences*, 110(38):15348–15353, 2013.

[SP14a] Alexander J Stewart and Joshua B Plotkin. Collapse of cooperation in evolving games. *Proceedings of the National Academy of Sciences*, 111(49):17558–17563, 2014.

[SP14b] A. Szolnoki and M. Perc. Evolution of extortion in structured populations. *Phys. Rev. E*, 89, 2014.

[SP14c] Attila Szolnoki and Matjaž Perc. Defection and extortion as unexpected catalysts of unconditional cooperation in structured populations. *Scientific Reports*, 4(1):1–6, 2014.

[SP15] Alexander J Stewart and Joshua B Plotkin. The evolvability of cooperation under local and non-local mutations. *Games*, 6(3):231–250, 2015.

[SP16] A. J. Stewart and J. B. Plotkin. Small groups and long memories promote cooperation. *Scientific Reports*, 6:26889, 2016.

[SSG+23] M. Stallen, L.L. Snijder, J. Gross, L.P. Hilbert, and C.K.W. De Dreu. Partner choice and cooperation in social dilemmas can increase resource inequality. *Nature Communications*, 14:6432, 2023.

[SSP18] F. P. Santos, F. C. Santos, and J. M. Pacheco. Social norm complexity and past reputations in the evolution of cooperation. *Nature*, 555, 2018.

[Ste12] Kim Sterelny. *The evolved apprentice.* MIT press, 2012.

[SUOY24] T. Sasaki, S. Uchida, I. Okada, and H. Yamamoto. The evolution of cooperation and diversity under integrated indirect reciprocity. *Games*, 15:15, 2024.

[TN06] Arne Traulsen and Martin A Nowak. Evolution of cooperation by multilevel selection. *Proceedings of the National Academy of Sciences*, 103(29):10952–10955, 2006.

[Tri71] R. Trivers. The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46(1):35–57, 1971.

[TRS14] Danielle FP Toupo, David G Rand, and Steven H Strogatz. Limit cycles sparked by mutation in the repeated prisoner's dilemma. *International Journal of Bifurcation and Chaos*, 24(12):1430035, 2014.

[Uch10] S. Uchida. Effect of private information on indirect reciprocity. *Phys. Rev. E*, 82, 2010.

[VGE08] Ajit Varki, Daniel H. Geschwind, and Evan E. Eichler. Human uniqueness: genome interactions with environment, behaviour and culture. *Nature Reviews Genetics*, 9(10):749–763, October 2008.

[VSPL14] Vítor V Vasconcelos, Francisco C Santos, Jorge M Pacheco, and Simon A Levin. Climate policies under wealth inequality. *Proceedings of the National Academy of Sciences*, 111(6):2212–2216, 2014.

[VSPLS12] Sven Van Segbroeck, Jorge M Pacheco, Tom Lenaerts, and Francisco C Santos. Emergence of fairness in repeated group interactions. *Physical Review Letters*, 108(15):158104, 2012.

[WCW+23] Xiaomin Wang, Marta C Couto, Nianyi Wang, Xinmiao An, Bin Chen, Yali Dong, Christian Hilbe, and Boyu Zhang. Cooperation and coordination in heterogeneous populations. *Philosophical Transactions of the Royal Society B*, 378(1876):20210504, 2023.

[WM00] C. Wedekind and M. Milinski. Cooperation through image scoring in humans. *Science*, 288, 2000.

[WPG02] Stuart A West, Ido Pen, and Ashleigh S Griffin. Cooperation and competition between relatives. *Science*, 296(5565):72–75, 2002.

[YWH22] Fengyuan Yu, Jianwei Wang, and Jialu He. Inequal dependence on members stabilizes cooperation in spatial public goods game. *Chaos, Solitons & Fractals*, 165:112755, 2022.

[Zel03] Jennifer Zelmer. Linear public goods experiments: A meta-analysis. *Experimental Economics*, 6:299–310, 2003.

[Zha18] Huanren Zhang. Errors can increase cooperation in finite populations. *Games and Economic Behavior*, 107:203–219, January 2018.