Decentralisation Conscious Players And System Reliability

Sarah Azouvi[1] and Alexander Hicks[2]

[1] Protocol Labs
sarah.azouvi@protocol.ai
[2] University College London
alexander.hicks@ucl.ac.uk

Abstract. We propose a game-theoretic model of the reliability of decentralised systems based on Varian's model of system reliability [27], to which we add a new normalized total effort case that models *decentralisation conscious players* that prioritize decentralisation.

We derive the Nash equilibria in the normalized total effort game. In these equilibria, either one or two values are played by players that do not free ride. The speed at which players can adjust their contributions can determine how an equilibrium is reached and equilibrium values. The behaviour of decentralisation conscious players is robust to deviations by other players.

Our results highlight the role that decentralisation conscious players can play in maintaining decentralisation. They also highlight, however, that by supporting an equilibrium that requires an important contribution they cannot be expected to increase decentralisation as contributing the equilibrium value may still imply a loss for many players. We also discuss practical constraints on decentralisation in the context of our model.

Keywords: decentralisation, public goods, free-riding, reliability

1 Introduction

The reliability of a system captures the likelihood that it performs as intended. For a decentralised system, there are two important components to consider, the number of participants and the distribution of power between them [26]. Even if there is a high number of contributors, if one of them has significantly more control over the system, there will be no meaningful level of decentralisation. This presents a problem that has been hard to solve in practice. How can the effort put into a system grow while maintaining an acceptable level of decentralisation?

Participation rewards can incentivise an increase in the effort invested in a system but a greater total effort can also be more centralised. Certain protocol considerations may alleviate this effect, e.g., at the consensus level [7]. It is also sometimes assumed that a portion of players will behave altruistically, following protocol guidelines even when an a priori more profitable strategies exist.

An alternative assumption, which we consider here, is that players have an incentive to maintain decentralisation. Short-term profits may be outweighed by

the possible long-term profits associated with maintaining a reliable system. For example, the value of a cryptocurrency that is vulnerable to hostile takeovers may decrease so miners have an incentive to maintain decentralisation and preserve the value of the tokens they hold and continue to receive.

Three observations support this assumption. First, the market price of a cryptocurrency is linked to its security [9]. Second, numerous flaws have been identified in the incentive structure of cryptocurrencies [11,21], yet attacks based on these have scarcely been observed [23]. Third, a mining pool has previously acted to avoid controlling more than half of Bitcoin's hash rate [20].

To further understand the rationality of maintaining decentralisation, this paper studies a game-theoretic model of decentralisation conscious players who prioritise decentralisation. With this model, we can analyse how such players will behave to ensure that a system remains decentralised, what effort they may contribute, and under which circumstances they will free-ride.

Our contributions The main contribution of this paper is the introduction and analysis of the normalised total effort game with decentralisation conscious players that extends Varian's system reliability model to decentralised systems.

We introduce our model based on the normalized total effort (NTE) function in the context of Varian's system reliability model [27] in Section 2. In Section 3, we derive the two types of Nash equilibria between decentralisation conscious players in which players contribute the same amount or two distinct amounts while others free ride. We also consider the social optimum, in which players contribute the same effort while minimizing their costs to maximize decentralisation.

To understand how decentralisation conscious players will behave in real systems alongside selfish and Byzantine players, we study in Section 4 the robustness of the previously derived equilibria when (i) the number of players change, which does not always affect the equilibrium; (ii) players deviate from the equilibrium, which can lead to a new equilibrium where players (possibly fewer) contribute a greater effort. Non-myopic players may, therefore, be incentivised to deviate from an equilibrium to reach a new equilibrium with fewer contributing players and a greater share of rewards.

Finally, we discuss in Section 6 some practical constraints on decentralisation in relation to our model.

2   Modelling System Reliability And Normalized Total Effort

Varian's original model of system reliability (treated as a public good) considers three cases based on how the individual efforts x_i of players are factored in [27]. The weakest link case considers the minimal effort exerted by any one of the players i.e, $F(x_1, \ldots, x_n) = \min_i(x_i)$. The total effort case considers the sum of every player's efforts i.e., $F(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i$. The best shot case considers the maximal effort exerted by any one of the players i.e., $F(x_1, \ldots, x_n) = \max_i(x_i)$.

Reliability will usually depend on a combination of these cases. For example, in the case of software security, a program's correctness can depend on the weakest link (the developer that introduces bugs), vulnerability testing depends on the total effort of all the testers, and the contributions of a system architect maps to the best shot case [4].

For each case, the Nash equilibria can be computed with the expected pay-off u_i for a player i expressed as in Equation 1, as can be the social optimum based on social pay-off SP expressed as in Equation 2. The likelihood that the system operates successfully is captured by $P(F(x_1, \ldots, x_n))$, which is assumed to be differentiable, increasing, and concave. The parameter v_i is the value derived by player i of the system operating successfully, and $c_i x_i$ is the cost to player i where c_i is a constant. The choice of a linear cost function of the form $c_i x_i$ implicitly ignores more complex forms of cost and any fixed costs. This is a limitation but it is realistic in relevant cases e.g., the energy required to operate a computer may be valued at a fixed price per kilowatt-hours.

$$u_i = P(F(x_1, \ldots, x_n))v_i - c_i x_i \tag{1}$$

$$SP = P(F(x_1, \ldots, x_n))(v_1 + \ldots + v_n) - \sum_{i=1}^{n} c_i x_i \tag{2}$$

The equilibria can be used to determine when free-riding can be expected to occur based on the form of F. For example, in the total effort case, the equilibrium is for players to free ride on the player who has the highest benefit-cost ratio $\frac{v_i}{c_i}$. The social optimum, obtained by maximizing the social pay-off rather than the player's utility functions, can also reveal how selfish behaviour from the players will lead to an outcome that is different from the social optimum. This is the case in the total effort case used as an example. Players free ride on the player with the highest benefit-cost ratio, which amounts to less total effort than in the social optimum, and the "wrong" players (those with the smallest benefit-cost ratio) can be found to contribute that effort.

The takeaway from Varian's results is that centralisation emerges even in the total effort case that involves everyone's contributions, and that rational behaviour can conflict with the social optimum i.e., selfish behaviour can lead to a weaker system – a concept known as *the price of anarchy* [24]. If decentralisation is desired, this means that an alternative model that produces individual and social outcomes that support a decentralised and stronger system is required.

To model decentralisation, the *relative* contribution of every player in the system must be taken into account because while the total effort should be as high as possible, the effort must also be as evenly distributed as possible. In practice, however, there are trade-offs between maximizing total effort and distributing effort evenly. It is unlikely that every player will have the same capacity to contribute, so maximizing the total effort is likely to come at the cost of a uniform distribution of effort, and vice versa.

With this in mind, we define in Equation 3 the normalized total effort (NTE) function based on the total effort and the maximal contribution. If the total effort

is high but the maximal effort is also high then the NTE may not be as high as
when the total effort is high but the maximal effort is low.

$$F(x_1, \ldots, x_n) = \frac{\sum_{i=1}^n x_i}{\max_i(x_i)} \tag{3}$$

The normalised total effort function is scale invariant i.e., $F(\alpha x_1, \ldots, \alpha x_n) = F(x_1, \ldots, x_n)$ for any α. This is because we are modelling players who care about
decentralisation over total effort. The goal is to capture the fact that in systems
that are designed to be decentralised, it is not only the total effort (studied by
Varian) that matters but the distribution of effort and, in particular, how much
the maximal contribution by a single player is as a portion of the total effort,
which our measure captures. Finding a measure that captures both this and the
benefits of a higher total effort is an open problem, and measures similar to ours
(e.g., the work of Kwon et al. [22]) suffer from the same limitation.

We show in Section 4 that contributions can still be expected to increase
given that other players who prioritize maximising their share of rewards exist.
Thus, much like in software security, a decentralised system's reliability depends
on nodes that are primarily concerned with decentralisation (normalised total
effort) and nodes that are primarily concerned with higher contributions (and
higher rewards) that increase the best shot and total effort. Because the best
shot and total effort case have already been studied by Varian, our focus in this
paper is the normalised total effort case.

3   Equilibria Between Decentralisation Conscious Players

We begin by studying the Nash equilibria of the NTE game defined below.

Definition 1 (Normalized Total Effort Game). *We call the normalized total effort game (NTEG) the game consisting of n players with costs $(c_1, \ldots, c_n) \in (\mathbb{R}_+^*)^n$, valuations $(v_1, \ldots, v_n) \in (\mathbb{R}_+^*)^n$, contributions $(x_1, \ldots, x_n) \in (\mathbb{R}_+)^n$, utility functions defined by equations 3 and 1, benefit-cost ratios $\beta_i = \frac{v_i}{c_i}$ such that $\beta_1 < \ldots < \beta_n$, and where we assume a logarithmic reliability function $P(F(x_1, \ldots, x_n)) = \ln\left(\frac{\sum_{i=1}^n x_i}{\max_i(x_i)}\right)$ for $\max_i(x_i) > 0$. By convention we have $P(F(0, \ldots, 0)) = 0$ i.e., a system with no contributions does not function.*

$$F(x_1, \ldots, x_n) = \frac{\sum_{i=1}^n x_i}{\max_i(x_i)} \tag{3}$$

$$u_i = v_i P(F(x_1, \ldots, x_n)) - c_i x_i \tag{1}$$

Two-player case We start by considering the simple case of a two-player game
and the following theorem, which we prove in Appendix A.

Theorem 1. *In a two-player NTEG, the Nash equilibria are for both players to
contribute the same effort $x_1 = x_2 = x_{eq}$ such that $x_{eq} \leq \frac{1}{2}\min(\beta_1, \beta_2)$.*

Both players contribute the same effort when the equilibrium is played, which
is the only possible "decentralised" solution.

Multiplayer case For $n > 2$ players, we prove the following in Appendix B.

Theorem 2. *In a $n > 2$ player NTEG, there exist two types of equilibrium.*

1. *(1-value equilibria) players $i+1$ to n (for $1 \le i < n$) contribute x_{eq} subject to the constraint expressed by Inequality 4 and players 1 to i with the smallest benefit-cost ratio free ride on them.*

$$\frac{1}{n-i}\beta_i \le x_{eq} \le \frac{1}{n-i}\beta_{i+1} \tag{4}$$

2. *(2-value equilibria) player i contributes x_m , players $(i+1$ to $n)$ contribute x_M, where $x_m < x_M$, subject to the constraints in Inequality 5 and Equation 6 and players 1 to $i-1$ free ride, for $1 \le i \le n$ (with no players free riding if $i = 1$).*

$$\frac{1}{n-i+1}\beta_i < x_M < \frac{1}{n-i}\beta_i \tag{5}$$

$$x_m = \beta_i - (n-i)x_M \tag{6}$$

We highlight Lemma 1 (proven as part of the proof) that we will reuse later.

Lemma 1. *If there exist two contributing rational players whose contributions are strictly less than $\max_i(x_i)$ and who play their best strategy, then those players must have the same benefit-cost ratio.*

Unless specified otherwise, we denote by x_{eq} the value played by the players or bulk of players in the 1-value or 2-value equilibrium, respectively. For both types of equilibrium, the lower x_{eq} is the more decentralised the system is, as more players can contribute and the less free-riding there is.

The fact that one equilibrium is for all players to contribute the same amount of effort makes sense as the NTE function encodes the social goal of maximizing decentralisation. It also prevents the perverse effects of any feedback loops that enable some players to contribute increasingly more than other players.

The 2-value equilibrium is less expected. It shows that, even if some players cannot match the other players' contributions (due to their own costs or valuation), they may still be incentivised to contribute.

3.1 The impact of a reward

The equilibria we have derived above include the case where everyone contributes no effort. Adding a reward function $R_i(x_1, \ldots, x_n)$ to the utility function, as in Equation 7. (e.g., cryptocurrency mining rewards) is a way of explicitly incentivising non-zero contributions, particularly from new players.

$$u_i = P(F(x_1, \ldots, x_n))v_i - c_i x_i + R_i(x_1, \ldots, x_n) \tag{7}$$

A reward separate from the valuation v models the compensation for the effort invested in the system rather than the benefit derived from being able to use the system. In practice, it may be a constant R that can be won by players with a probability proportional to the effort they contribute. Under certain conditions, this is an optimal allocation rule [14] so we restrict ourselves to this case.

$$R_i(x_1, \ldots, x_n) = \begin{cases} R\frac{x_i}{\sum_{j=1}^n x_j}, & \text{if } \max(x_1, \ldots, x_n) > 0 \\ 0, & \text{if } \max(x_1, \ldots, x_n) = 0 \end{cases} \tag{8}$$

This removes the $x_{eq} = 0$ equilibrium without significantly affecting other equilibria. In the two-player case, an equilibrium still involves the two players contributing the same value x subject to different constraints and under the additional assumptions that $R < \min(v_1, v_2)$. This expresses the fact that the player's valuations of the system must be at least greater than the value of the reward – it would make little sense to gain a reward that is greater than the value of the system functioning. We prove the following theorem in Appendix C.

Theorem 3. *In a two player NTEG with reward $R < \min(v_1, v_2)$ there exist infinite Nash equilibria where both players contribute the same value x such that*

$$\begin{cases} \frac{c_1}{4R}((R\frac{1-\sqrt{\Delta_1'}}{2c_1})^2 - \beta_1^2) < x < \frac{c_1}{4R}((R\frac{1+\sqrt{\Delta_1'}}{2c_1})^2 - \beta_1^2) \\ \frac{c_2}{4R}((R\frac{1-\sqrt{\Delta_2'}}{2c_2})^2 - \beta_2^2) < x < \frac{c_2}{4R}((R\frac{1+\sqrt{\Delta_2'}}{2c_2})^2 - \beta_2^2) \end{cases} \tag{9}$$

with $\Delta_1' = 1 + 4\frac{c_1}{R}(\frac{v_1^2 c_1}{R} + \frac{v_1}{c_1})$ and $\Delta_2' = 1 + 4\frac{c_2}{R}(\frac{v_2^2 c_2}{R} + \frac{v_2}{c_2})$.

We leave the multiplayer analysis as future work.

3.2   Social optimum

An insight from Varian's work is that the equilibria and social optima are not necessarily the same e.g., the total effort social optimum involves players contributing much more than in the Nash equilibrium [27].

In the NTE case, the social optimum is for players to contribute the smallest non-zero amount possible as this maximizes the level of decentralisation while minimizing their costs. If all contributions are equal then in most cases it is also a Nash equilibrium. This convenient outcome is expected from our choice of NTE that reflects a desire to ensure that the social goal of decentralisation is met, so the NTE function is well defined in that sense. The only exception is when the benefit-cost ratio of some players is too low as they then free-ride.

Figuring out an acceptable minimal contribution can be straightforward when it is possible to impose a minimum contribution. Ethereum's implementation of proof-of-stake does this, but not all systems impose a minimum contribution.

4   Robustness Of Decentralisation Conscious Players to Variations By Others

In practice, players may leave or join the game, as well as increase or decrease their contributions because of selfish behaviour or, more generally, Byzantine faults. Thus, it is important to analyse how decentralisation conscious players tolerate variations in the actions of other players. We do this by studying how the equilibria for the NTEG change after such events.

In the analysis that follows we will be using a result derived in the proof of Theorem 2, which is that for each player j the best response to (fixed) contributions of other players is as follows.

1. if $\sum_{i \neq j} x_i < \beta_j$, contribute $\min(\max_{i \neq j}(x_i), \beta_j - \sum_{i \neq j} x_i)$
2. if $\sum_{i \neq j} x_i \geq \beta_j$, contribute zero.

Equivalently, player's j best response can be written as in Equation 10.

$$\max\{0, \min(\max_{i \neq j}(x_i), \beta_j - \sum_{i \neq j} x_i)\} \qquad (10)$$

In this section, we note n the number of contributing players.

4.1   Change in number of players

New player joining

1-value equilibrium We first consider the case where the players play the 1-value equilibrium described in Theorem 2. If one player joins the game, the robustness of the equilibrium depends both on x_{eq} and on the benefit-cost ratio β of the new player. From Condition 4, we have that $x_{eq} \leq \frac{1}{n}\beta_j, \ \forall 1 \leq j \leq n$.

We proceed as follows. For different values of x_{eq} we study what would be the new player's best response x_{new} and whether they would join the game i.e., contribute a non-zero effort. We then look at whether the introduction of a new player playing x_{new} disrupts the equilibrium for the rest of the players i.e., whether having n players play x_{eq} and one player play x_{new} is still an equilibrium. We find that the original n players change their contributions if and only if $\beta > \beta_1$ and $\frac{1}{n+1}\beta_1 < x_{eq} < \frac{1}{n}\beta$. We prove this result in Appendix D.

Theorem 4. *In a NTEG that is in a state of 1-value equilibrium with n players contributing x_{eq}, the introduction of a new player with benefit-cost ratio β changes the value played by the other players if and only if $\beta > \beta_1$ and $\frac{1}{n+1}\beta_1 < x_{eq} < \frac{1}{n}\beta$.*

2-value equilibrium In the case where the players were initially in a 2-value equilibrium, we have the following theorem, which we prove in Appendix E.

Theorem 5. *In a NTEG that is in a state of 2-value equilibrium with player 1 playing x_1 and the other $n-1$ players playing x_{eq}, the introduction of a new player with benefit-cost ratio β does not change the value played by the other players unless $\sum_{i=1}^{n} x_i < \beta$ or $\beta_1 < \beta$.*

We now consider how the utility of each player changes following the introduction of a new player. If the new player contributes a strictly positive effort and the value played at equilibrium stays unchanged for the other players it is clear that the introduction of a new player increases everyone's utility as it increases the reliability of the system without changing anyone's cost. When the equilibrium is changed, if only one player (player 1) leaves the system, then this is simply a player replacement and the reliability of the system stays the same. Player 1 increases their utility in this case as the reliability of the system is the same as before but their cost is now zero.

However, from the proof of Theorem 4 we have that a new player could potentially incentivise more than one player to decrease their contribution. Lemma 1 tells us that this means that the players would potentially need many iterations before reaching a new equilibrium if they reach one, where only one or two values are played. Although it could be presumed that a new player joining should increase the reliability of the system, this result shows that if one or more players have to decrease their contributions then it is not clear that the final reliability of the system will be higher with $n+1$ player than with the original n players. We study simulations of equilibrium disruption in Section 5 and leave a rigorous study of the outcome of the new game as an open problem.

Player leaving the game In the case where a player leaves the game, we have the following theorem, which we prove in Appendix F.

Theorem 6. *In the NTEG, if the n players are playing a $1-value$ Nash equilibrium, the removal of a new player with benefit-cost ratio β_i does not change the value played by the other players.*

If other contributions stay unchanged, a player leaving the system decreases the reliability of the system as it renders it more centralised. The utilities of the remaining players will therefore always decrease in this case.

4.2   Deviation from an equilibrium

We now consider the case where one player (player k) deviates from the equilibrium and changes their contribution to x_{k_0}. We are concerned with the response of the $n-1$ other players and what new equilibrium is reached, regardless of whether it will be the best strategy for player k to keep their value x_{k_0} in the new equilibrium (i.e., player k may be irrational). We prove the following theorem in Appendix G

Theorem 7. *In the NTEG with n players contributing the same value x_{eq} at the equilibrium, the deviation of player k with benefit-cost ratio β_k to a new value x_{k_0} does not change the value played by the other players unless $x_{k_0} > x_{eq}$.*

In the 2-value equilibrium, the results are very similar. We prove the following theorem in Appendix H.

Theorem 8. *In the NTEG with n players playing a 2-value equilibrium where players 2 to n play the same value x_{eq} at the equilibrium, the deviation of player k with benefit-cost ratio β_k to a new value x_{k_0} changes the value played by the other players unless in the specific case where player 1 is deviating to a new value $x_{k_0} \neq x_1$ and for all $2 \leq j \leq n$ we have (1) $x_{k_0} < x_{eq}$ (2) $\beta_j > (n-2)x_{eq} + x_{k_0} + \max(x_{k_0}, x_{k_0})$ and (3) $(n-2)x_{k_0} + x_{k_0} < \beta_2$.*

In the case where the players do not change their equilibrium after an irrational player deviates (i.e., $x_{k_0} < x_{eq}$) the utility of players will decrease as reliability will be lower for the same costs and contributions.

In the other case, before the other players can adjust their contribution, their utility will also decrease, and in some realistic cases, players may not be able to change their contribution as we discuss in Section 6. This is an undesirable effect defined as *immunity* by Abraham et al. [2] in the context of distributed systems where one or more irrational players can negatively impact the utility of rational players. If players can change their contributions, the reliability functions could go up or down depending on the new value x_{eq} and the benefit-cost of other players (i.e., whether they will free ride).

After the deviation from player k, we have from Condition 10 that each player i such that $x_{k_0} \geq \frac{\beta_i - (n-2)x_{eq}}{2}$ changes their contribution to $x_{i,\text{new}} = \beta_i - x_{k_0} - (n-2)x_{eq}$ or zero if that value is negative, and each player i such that $x_{k_0} \leq \frac{\beta_i - (n-2)x_{eq}}{2}$ changes their contribution to x_{k_0}.

In Lemma 1, we showed that if there exists two contributing rational players whose contributions are strictly less than $\max_j(x_j)$, then those players must have the same benefit-cost ratio. This is true regardless of the existence of an irrational player. Since we assume that all the benefit-cost ratios are different, this means that there can be at most one rational player playing strictly less than the maximum value x_M. According to the strategy defined in Condition 10, no rational player is incentivised to play more than $\max_j(x_j)$. Thus, even after players adjust their contributions we will still have $\max_j(x_j) = x_{k_0}$ and, following the deviation, the bulk of the players will align with the deviating players or free ride, except for one rational player. By setting x_{k_0} high enough, the deviating player could ensure that many players switch to free-riding, which could pose a threat to the system if it facilitates one party taking control of the system (e.g., a 51% attack).

4.3   Non-myopic players

Motivated by Brünjes et al. [13], we consider *non-myopic players* deviating from the equilibrium. The utility function of such players accounts for the effects an

action will have on the other players, unlike Nash equilibria that consider the best response of players given that the other players' strategies are fixed.

In the previous section, we have seen that a player deviating from the equilibrium may disrupt the best response of the other players and lead to a new equilibrium. In a Nash equilibrium, assuming that other players keep their contribution unchanged, deviating means that one's utility is reduced, but this does not account for the possibility of a new equilibrium being reached. A new equilibrium (if reached) may be a better equilibrium for the deviating player if their utility is higher in the new equilibrium.

Does a non-myopic player have incentives to deviate from the equilibria we have derived? We have established that a condition to disrupt the equilibrium is to change one's contribution to a value $x_{k_0} > x_{eq}$. We have also observed that by setting this value high enough, the deviating player can cause some players to free ride. In a NTEG without a reward, the new equilibrium would, therefore, have fewer contributing players with greater contributions. In a 1-value equilibrium, this would mean we have $F(x_1, \ldots, x_n) = n_{\text{new}} < n$ where n_{new} is the new number of contributing players. However, because $x_{k_0} > x_{eq}$, the cost will be higher and this strategy is therefore not rational as the new equilibrium results in less utility for the deviating player and the other players.

In a NTEG with reward, however, fewer players implies a greater share of rewards. Thus a non-myopic player may be incentivised to deviate from an existing equilibrium to reach a new one with fewer contributing players.

This suggests that a fixed proportional reward may increase centralisation. Designing a protocol with a variable reward such that players would earn similar revenue regardless of the number of players is an open problem due to the pseudonymous nature of systems like cryptocurrencies. Another alternative is to rely on a fixed reward but design the system such that it is not possible to increase one's contribution, as in proof-of-personhood schemes [12].

4.4   Coalition-resistance

A group of miners may decide to form a coalition if this increases their expected gain, even if doing so centralises the system. In this case a coalition is equivalent to having one player contributing $X = \sum_{i \in [i_1, \ldots, i_c]} x_i$ for all the players $(i_1, \ldots, i_c)$ in the coalition instead of having each contributing separately. Because the sum of the efforts stay the same but the maximum effort potentially increases, $F(x_1, \ldots, x_{i_1}, \ldots, x_{i_c}, \ldots, x_n) \geq F(x_1, \ldots, X, \ldots, x_n)$. Thus, the utility of decentralisation conscious players decreases when they form a coalition i.e., they are not incentivised to create coalitions.

5   Dynamics Of Decentralisation Conscious Players

As we have shown, there are many possible equilibria, each corresponding to different equilibrium values. How an equilibrium is reached i.e., how quickly and how many players reach it, as well as which equilibrium value is reached could depend on several factors that we look at in this section.
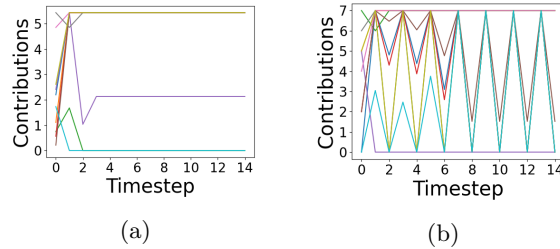
Fig. 1: Without constraints on contribution changes players can reach an equilibrium (Figure 1a) but may also oscillate indefinitely (Figure 1b).

Methodology Using a Python script, we simulate the NTEG where each player computes their best strategy at each time unit. By iterating over multiple time units we observe how players (simultaneously) re-evaluate their contributions based on the effort of other players in the previous time unit. The scenarios we simulate are not exhaustive but highlight interesting behaviour, the benefit-cost ratios were chosen randomly within a range.

Random initial values To observe how an equilibrium is reached, we initialize a NTEG with 10 players to which we assign random initial values (contributions, costs, benefits) and look at how they change their contributions until an equilibrium is reached. (The same initial contributions and benefit-cost ratios are used for every simulation.) According to the strategy defined by Equation 10, no decentralisation conscious player is incentivised to contribute more than other players hence the player that has the maximum contribution in step 1 of the game (set by nature's move) will be reducing their contribution in the next step. On the other hand, other players with a high enough benefit-cost ratio will be incentivised to increase their contributions to the maximum value in step two of the game.

Under ideal conditions i.e., when the maximum contribution $x_{\max}$ is such that $nx_{\max} < \min_i \beta_i$, the equilibrium is reached after a few steps. Players with the greatest benefit-cost ratios align their contributions to the maximum value (except perhaps for one of them, resulting in a 2-value equilibrium) while the remaining players free ride, as shown in Figure 1a.

In other cases, as shown in Figure 1b, some players may keep oscillating indefinitely. For these players, it must be the case that $\beta_j < nx_{eq}$, else playing x_{eq} at the same time as other players will be their best strategy and an equilibrium will be reached. Thus whenever everyone is playing x_{eq} at one time unit, they decrease their contributions to $\beta_j - nx_{eq}$ in the next step. However, after other oscillating players have also decreased their contributions, it is now the best strategy to go back to x_{eq}, and so on. This is due to players being myopic, not anticipating that other players will increase their contributions at the same time as them.

(a) $\Delta = 0.1$ (b) $\Delta = 0.3$ (c) $\Delta = 0.5$
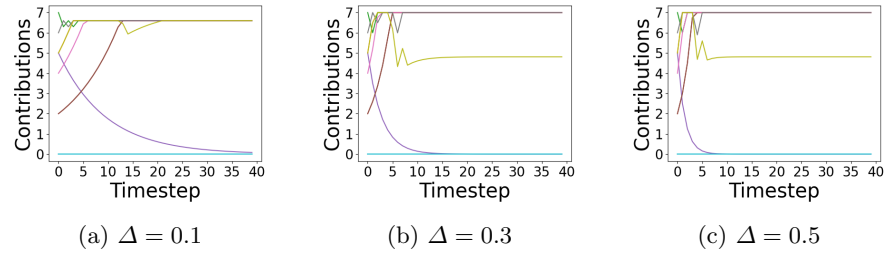
Fig. 2: Oscillations disappear with constraints on contribution changes. The speed at which equilibria are reached depends on the constraints (slower with $\Delta = 0.1$, faster with $\Delta = 0.5$), as do the type of equilibria (1-value with $\Delta = 0.1$, 2-value $\Delta = 0.3$ or 0.5) and equilibrium value (greater with $\Delta = 0.3$ or 0.5).

Constraints on the rate of change of contributions. To avoid the unrealistic case where players oscillate forever we constrain the change in each player's contribution from one time unit to another by a factor Δ. This dampens the oscillations and allows players to converge to an equilibrium.

Because Δ affects how quickly players can converge to an equilibrium, the equilibrium that is reached varies with Δ. For example, in the case where $\Delta = 0.1$ participants are allowed to change their contributions by at most 10% from one time unit to another and a 1-value equilibrium is reached, as shown in Figure 2a. When $\Delta = 0.3$ or $\Delta = 0.5$, a 2-value equilibrium is reached, as shown in Figures 2b and 2c. Keeping this in mind we will, however, stick to the $\Delta = 0.1$ case in most of the simulations that follow for simplicity as the overall player behaviours i.e., players increasing their contribution or free-riding are the same although the final equilibrium differs.

We have also computed the different values of the reliability in each case but did not observe any clear pattern. Whether there is a pattern that is not clearly observable is left as an open problem.

Not only is a constraint on the change in the effort of players useful for them to efficiently converge to an equilibrium, it is also realistic. Players in real life are likely to understand the adverse effects of over correcting and are also likely to have constraints on how much they can change their effort (at least upwards) due to the cost of doing so. We discuss this constraint further in the next section, in relation to resource scarcity.

Moreover, every player updating their contributions at the same time is not a realistic assumption either. Bounding the change of contribution of each player from one step to another also helps get closer to a continuous time model.

Constraints on total effort Another constraint that can be implemented is a limit on the overall change in the effort of all players i.e., the total effort. This models the constraint that the stock of resources used to contribute effort (e.g., new hardware) may be limited at any point in time. Figure 3a shows that in

(a) $\Delta = 0.1$ (b) $\Delta_+ = 0.05$, $\Delta_- = 0.4$ (c) $\Delta_+ = 0.4$, $\Delta_- = 0.05$
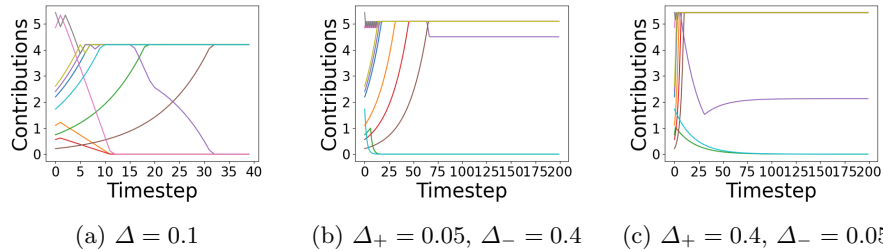
Fig. 3: Constraining the total effort can increase free-riding and reduce equilibrium values (Figure 2a), as can reductions in contributions being easier than increases (Figures 3b and 3c).

this case, some players may not be able to change their contribution enough to converge to the equilibrium and, therefore, switch to free riding.

Since it is usually easier to reduce one's contribution than to increase it, we also simulate the game with different constraints on the increase and the decrease of contributions from one step to another. We see in Figures 3b and 3c that a 2-value equilibrium is reached, although the relative constraints on increasing and decreasing contributions result in different equilibrium values. The value played by the bulk of the player x_{eq} is higher when there is a greater constraint on the increase of contributions than on the decrease. This is because players can more rapidly reach the new maximum value. As a consequence, the second value played at the equilibrium is smaller.
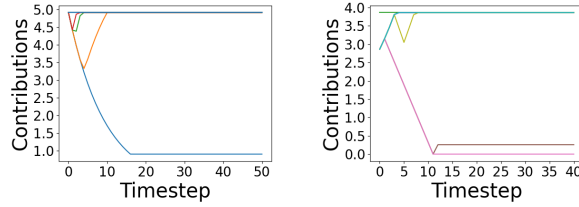
Disruptions to an equilibrium A new player joining the game when it is in a 1-value equilibrium (which happens according to the conditions defined in Theorem 4) can lead to a new equilibrium being reached after a few steps, as shown in Figure 4a in the case of a strong constraint.

When an equilibrium is disrupted by a player deviating from the equilibrium, players that increase their contribution to contribute more effort than the equilibrium value incentivise other decentralisation conscious players to free ride or increase their effort to reach a new equilibrium value if it is allowed by their benefit-cost ratio. This is shown in Figure 4b, in the case of a strong constraint.

6 Discussion

6.1 The role of decentralisation conscious players

Our model and choice of NTE function shows that decentralisation conscious players can help maintain a decentralised system. However, as Theorems 7 and 8 show, decentralisation conscious players only ever increase their effort in response to another player increasing their contribution at the cost of decentralisation. They maintain decentralisation within the constraints of their benefit-cost ratio

(a) $\Delta = 0.1$, a new player joins

(b) $\Delta = 0.1$, a player deviates

Fig. 4: Disruptions to an equilibrium due to a new player joining or a player deviating lead to new equilibriums.

but ignore players that free ride after their benefit-cost ratio no longer allows them to contribute.

Because decentralisation conscious players can only maintain a pre-existing level of decentralisation and can be leveraged by selfish players to implement a minimum benefit-cost ratio that acts as a form of gate-keeping against players with lower benefit-cost ratios, there is a distinction between decentralisation conscious players and *altruistic* players that operate regardless of their benefit-cost ratio. This suggests that new mechanisms dictating how effort is contributed or rewarded may be needed for players to have rational ways of increasing decentralisation outside of purely altruistic behaviour.

6.2   Modelling constraints

Resource scarcity Players contribute based on their benefit-cost ratios and, as we have seen in Section 5, equilibria depend on the rate of change of contributions. An implicit assumption made by our model is that a player can contribute more (at a cost) should they wish to do so but this may not be possible. For example, cryptocurrency mining hardware has suffered from shortages that forced buyers to obtain hardware at significant premiums and logistical difficulties [28]. When resources are unobtainable, it can become impossible to contribute more or continue contributing the same amount (if resources must be replaced), causing involuntary deviations from otherwise rational strategies.

If it is impossible to acquire the resources to contribute, the system will rely on players having a high valuation of the system. Contributors to systems like Tor [25] operating nodes at a loss may demonstrate this but in the case of cryptocurrencies new miners are less likely to have a high valuation of the system because they are unlikely to have a stake in it, unlike miners that have accumulated rewards. Miners in cryptocurrencies that are more centralised due to the high practical costs of mining can, therefore, form an effective oligopoly [16,5].

Can we avoid issues of resource scarcity? One way of avoiding the problematic reliance on resources with variable stock (e.g., stake, hardware) is to opt for

mechanisms like proof-of-personhood [12], which is equally distributed ("1 person = 1 vote") and maximizes the NTE, although this has other issues to overcome.

Geographical and political decentralisation Because players contribute based on their benefit-cost ratios so the geographical distribution of players will matter if costs vary with location. For example, cryptocurrency mining is concentrated in the few areas where mining is most profitable.

Markets are also affected by political power and changes in regulations. China controlled 65% of Bitcoin's hashpower in 2019 [17] but following new Chinese regulations [8] the share of hashpower in the US has grown due to political stability with respect to Bitcoin mining [3]. The impact of markets and political power on decentralisation adds complexity and uncertainty in models, which may motivate decentralised systems less reliant on other markets e.g., proof-of-stake (based on the cryptocurrency's native tokens) or proof-of-personhood may be easier to reason about than proof-of-work (energy and hardware markets).

Incomplete and unequal information Our model has assumed perfect information at each step with players changing their contributions based on this information, but players could hide information such as the stock of unused resources they have at their disposal. Attacks such as selfish mining in proof-of-work cryptocurrencies [15] are based on abusing information asymmetry, as are hostile takeovers which use previously unused but available mining capacity [10]. There is also an inherent delay in information propagating through a network. This may result in different equilibria as players adapt their contributions based on the information they receive at a point where it may no longer be accurate.

How much this matters is hard to determine. Attacks such as selfish mining have seldom been observed, and effort rarely varies across short time periods. (See the Bitcoin hashrate distribution over short time periods, even if the larger trend is growth [1].) This may be due to issues like acquiring the additional resources needed to contribute more effort, but it may also be to maintain a level of decentralisation as our model suggests miners might do.

6.3 Related Work

There is an important literature on modelling incentives in cryptocurrencies through refinements of Nash Equilibria that has been systematized [6]. Although the types of players and games considered vary across papers, none of the papers surveyed (except Varian's paper [27]) consider the reliability of the system.

Varian's system reliability paper [27] has previously been extended by Grossklags et al. [18] in the context of investments in security and insurance. Grossklags et al. [19] have also applied Varian's model to study the difference between expert and naive players in security games to quantify the impact of information. In this work, we have instead focused on decentralisation and introduced the NTEG, which extends Varian's model in another direction.

7   Conclusion

We have proposed a model for decentralisation conscious players based on the NTE function we have introduced. The Nash equilibria show what could be expected from such players. Using simulations we have also considered how players may reach an equilibrium, including after disruptions. There is a variety of possibilities for future work and opportunities to apply our model to specific cases. This includes cases with valuations of the system which are hard to precisely define e.g., ideological commitment, as well as cases with very explicit valuations and dependencies on rewards but complex financial optimization such as cryptocurrencies. Protocol designers who wish to incorporate rational players, as opposed to honest players, but also wish to incorporate the reliability of the system in addition to short-term rewards could use the NTE function.

Acknowledgments

References

1. pools-timeseries, https://www.blockchain.com/charts/pools-timeseries
2. Abraham, I., Dolev, D., Gonen, R., Halpern, J.: Distributed computing meets game theory: robust mechanisms for rational secret sharing and multiparty computation. In: Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing. pp. 53–62 (2006)
3. Allison, I.: Long in China's shadow, the US is becoming a Bitcoin mining power again (November 2020), https://www.coindesk.com/us-becoming-bitcoin-mining-power-again
4. Anderson, R.: Security engineering: a guide to building dependable distributed systems. John Wiley & Sons (2020)
5. Arnosti, N., Weinberg, S.M.: Bitcoin: A natural oligopoly. arXiv preprint arXiv:1811.08572 (2018)
6. Azouvi, S., Hicks, A.: Sok: Tools for game theoretic models of security for cryptocurrencies. arXiv preprint arXiv:1905.08595 (2019)
7. Bano, S., Sonnino, A., Al-Bassam, M., Azouvi, S., McCorry, P., Meiklejohn, S., Danezis, G.: Sok: Consensus in the age of blockchains. In: Proceedings of the 1st ACM Conference on Advances in Financial Technologies. pp. 183–198 (2019)
8. Baydakova, A.: China's crypto miners struggle to pay power bills as regulators clamp down on OTC desks (November 2020), https://www.coindesk.com/chinese-miners-struggle-to-pay-for-electricity
9. Bissias, G., Böhme, R., Thibodeau, D., Levine, B.N.: Pricing security in proof-of-work systems. arXiv preprint ar Xiv:2012.03706 (2020)
10. Bonneau, J.: Hostile blockchain takeovers (short paper). In: International Conference on Financial Cryptography and Data Security. pp. 92–100. Springer (2018)
11. Bonneau, J., Miller, A., Clark, J., Narayanan, A., Kroll, J.A., Felten, E.W.: Sok: Research perspectives and challenges for Bitcoin and cryptocurrencies. In: 2015 IEEE symposium on security and privacy. pp. 104–121. IEEE (2015)

12. Borge, M., Kokoris-Kogias, E., Jovanovic, P., Gasser, L., Gailly, N., Ford, B.: Proof-of-personhood: Redemocratizing permissionless cryptocurrencies. In: 2017 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW). pp. 23–26. IEEE (2017)

13. Brünjes, L., Kiayias, A., Koutsoupias, E., Stouka, A.P.: Reward sharing schemes for stake pools. arXiv preprint arXiv:1807.11218 (2018)

14. Chen, X., Papadimitriou, C., Roughgarden, T.: An axiomatic approach to block rewards. In: Proceedings of the 1st ACM Conference on Advances in Financial Technologies. pp. 124–131 (2019)

15. Eyal, I., Sirer, E.G.: Majority is not enough: Bitcoin mining is vulnerable. In: International conference on financial cryptography and data security. pp. 436–454. Springer (2014)

16. Gencer, A.E., Basu, S., Eyal, I., Van Renesse, R., Sirer, E.G.: Decentralization in Bitcoin and ethereum networks. In: International Conference on Financial Cryptography and Data Security. pp. 439–457. Springer (2018)

17. Godbole, O.: Highest in 2 years: 65% of Bitcoin hash power is in China, report finds (December 2019), https://www.coindesk.com/highest-in-2-years-65-of-bitcoin-hash-power-is-in-china-report-finds

18. Grossklags, J., Christin, N., Chuang, J.: Secure or insure? a game-theoretic analysis of information security games. In: Proceedings of the 17th international conference on World Wide Web. pp. 209–218 (2008)

19. Grossklags, J., Johnson, B., Christin, N.: When information improves information security. In: International Conference on Financial Cryptography and Data Security. pp. 416–423. Springer (2010)

20. Hajdarbegovic, N.: Bitcoin miners ditch ghash.io pool over fears of 51% attack (Apr 2014), https://www.coindesk.com/bitcoin-miners-ditch-ghash-io-pool-51-attack

21. Judmayer, A., Stifter, N., Zamyatin, A., Tsabary, I., Eyal, I., Gazi, P., Meiklejohn, S., Weippl, E.R.: Pay-to-win: Incentive attacks on proof-of-work cryptocurrencies. IACR Cryptol. ePrint Arch. **2019**, 775 (2019)

22. Kwon, Y., Liu, J., Kim, M., Song, D., Kim, Y.: Impossibility of full decentralization in permissionless blockchains. In: Proceedings of the 1st ACM Conference on Advances in Financial Technologies. pp. 110–123. AFT '19, ACM, New York, NY, USA (2019). https://doi.org/10.1145/3318041.3355463, http://doi.acm.org/10.1145/3318041.3355463

23. Neudecker, T., Hartenstein, H.: Short paper: An empirical analysis of blockchain forks in Bitcoin. In: International Conference on Financial Cryptography and Data Security. pp. 84–92. Springer (2019)

24. Roughgarden, T.: Selfish routing and the price of anarchy, vol. 174. MIT press Cambridge (2005)

25. Syverson, P., Dingledine, R., Mathewson, N.: Tor: The second generation onion router. In: Usenix Security. pp. 303–320 (2004)

26. Troncoso, C., Isaakidis, M., Danezis, G., Halpin, H.: Systematizing decentralization and privacy: Lessons from 15 years of research and deployments. Proceedings on Privacy Enhancing Technologies **2017**(4), 404–426 (2017)

27. Varian, H.: System reliability and free riding. In: Economics of information security, pp. 1–15. Springer (2004)

28. Wong, J.I.: Ethereum miners are renting Boeing 747s to ship graphics cards and AMD shares are soaring (July 2017), https://qz.com/1039809/amd-shares-are-soaring-ethereum-miners-are-renting-boeing-747s-to-ship-graphics-cards-to-mines/

A   Proof of Theorem 1

To determine player 1's best strategy, we first study the continuous function $x_1 \mapsto u_1(x_1)$ for a fixed x_2 and try to find its global maximum. There are two cases to consider, depending on whether $x_1 \leq x_2$ or $x_1 > x_2$.

If $x_1 < x_2$, we have $u_1 = \ln(\frac{x_1+x_2}{x_2})v_1 - c_1 x_1$ and $\frac{du_1}{dx_1} = \frac{v_1}{x_1+x_2} - c_1$, so $\frac{du_1}{dx_1} \geq 0$ iff Condition 11 holds.

$$x_1 + x_2 \leq \beta_1 \tag{11}$$

If $x_1 > x_2$, we have $u_1 = \ln(\frac{x_1+x_2}{x_1})v_1 - c_1 x_1$ and $\frac{du_1}{dx_1} = v_1(\frac{1}{x_1+x_2} - \frac{1}{x_1}) - c_1$, so $\frac{du_1}{dx_1} < 0$.

We can, therefore, define player 1's strategy as follows.

1. If $\beta_1 < x_2$ then (from condition 11) u_1 is a decreasing function and player 1's best strategy is to play $x_1 = 0$.
2. If $x_2 \leq \beta_1$ then (from condition 11) u_1 is increasing up to $\min(x_2, \beta_1 - x_2)$ so there are two cases to consider.
   (a) If $x_2 \leq \beta_1 - x_2$ i.e., if $x_2 \leq \frac{1}{2}\beta_1$, then condition (11) is always satisfied for $x_1 \leq x_2$ and hence player 1 best strategy is to play $x_1 = x_2$. When $x_1 > x_2$ u_1 decreases so player 1 maximizes u_1 by playing x_2.
   (b) If $x_2 > \beta_1 - x_2$ or equivalently if $\frac{1}{2}\beta_1 < x_2 \leq \beta_1$: then player 1 best strategy is to play $x_1 = \beta_1 - x_2$ (after which u_1 starts decreasing according to condition (11)).

The same analysis can be repeated for u_2, giving the same results. Therefore, there exist infinite equilibria where $x_1 = x_2 = x_{eq}$ and $x_{eq} \leq \frac{1}{2}\min(\beta_1, \beta_2)$. We now show that these are the only equilibria of the game.

If one of the contributions, say x_2, is zero, then we are in case (2a) of the strategy for player 1 and their best response is also zero. Thus, we have $x_1 = x_2 = x_{eq}$ such that $x \leq \frac{1}{2}\min(\beta_1, \beta_2)$.

For the rest of this proof, we assume that $x_1, x_2 > 0$. Proceeding by contradiction, we assume that there exists a Nash equilibrium (x_{1_0}, x_{2_0}) for which $x_{2_0} \neq x_{1_0}$. Since $x_{2_0} \neq x_{1_0}$ and $x_{1_0} > 0$ by assumption, this means we are in case (2b) of player 1's strategy and player 1 maximizes u_1 when $x_1 = \beta_1 - x_{2_0} = x_{1_0}$.

Similarly, we have $\beta_2 - x_{1_0} = x_{2_0}$. Solving this system of two equations gives us $x_{1_0} + x_{2_0} = \beta_1 = \beta_2$. So unless $\beta_1 = \beta_2$, we have a contradiction.

If $\beta_1 = \beta_2 = v/c$, then we have $x_{1_0} + x_{2_0} = v/c$. Again, proceeding by contradiction, we assume that $x_{1_0} > \frac{1}{2}v/c$. This implies that $x_{2_0} < \frac{1}{2}v/c$, implying that we are in case (2a) of player's 1 strategy and thus that player 1's best response is $x_{1_0} = x_{2_0}$. The analysis is the same if $x_{1_0} < \frac{1}{2}v/c$, applied to player 2. This proves that players contribute the same at the equilibrium.

B   Proof of Theorem 2

The analysis of $x_1 \mapsto u_1(x_1)$ for $(x_2, \ldots, x_n)$ fixed is similar to the analysis done in the two player case.

If $x_1 < \max_{i=2}^n(x_i)$, we have $u_1 = \ln(\frac{x_1 + \sum_{i=2}^n x_i}{m})v_1 - c_1 x_1$ and $\frac{du_1}{dx_1} = \frac{v_1}{\sum_{i=1}^n x_i} - c_1$. Hence, $\frac{du_1}{dx_1} \geq 0$ iff Condition 12 holds.

$$\sum_{i=1}^n x_i \leq \beta_1 \tag{12}$$

If $x_1 > \max_{i=2}^n(x_i)$, we have $u_1 = \ln(\frac{x_1 + \sum_{i=2}^n x_i}{x_1})v_1 - c_1 x_1$ $\frac{du_1}{dx_1} = v_1(\frac{1}{\sum_{i=1}^n x_i} - \frac{1}{x_1}) - c_1$, so $\frac{du_1}{dx_1} < 0$.

As in the two player case, we can establish that player 1's best strategy for a fixed $(x_2, \ldots, x_n)$ is as follows.

1. If $\max_{i=2}^n(x_i) + \sum_{i=2}^n x_i \leq \beta_1$, contribute $\max_{i=2}^n(x_i)$. This is because we will always have $\sum_{i=1}^n x_i \leq \beta_1$ as long as $x_1 \leq \max_{i=2}^n(x_i)$, thus u_1 increases as x_1 increases up to $\max_{i=2}^n(x_i)$ and then decreases. The best response is thus $x_1 = \max_{i=2}^n(x_i)$.
2. If $\sum_{i=2}^n x_i < \beta_1 < \max_{i=2}^n(x_i) + \sum_{i=2}^n x_i$, contribute $\beta_1 - \sum_{i=2}^n x_i$. This is because (from Condition 12) u_1 is increasing up to that point, then decreasing.
3. If $\sum_{i=2}^n x_i \geq \beta_1$ then u_1 is decreasing and thus player 1's best response is to contribute nothing: $x_1 = 0$.

The best strategy of every other player is derived in the same way.

It is straightforward to show that for every contributing player, the following holds at the equilibrium.

$$\forall i \in [1, n]: \sum_{j=1}^n x_j \leq \beta_i \tag{13}$$

Additionally, from the strategy defined above, we have that the case where players $i+1$ to n (for any $1 \leq i \leq n-1$) contribute the same value x such that $\frac{1}{n-i}\beta_i \leq x \leq \frac{1}{n-i}\beta_{i+1}$ and others free-ride is a Nash equilibrium. All the players $i+1$ to n are in case (1) of their strategies and thus contribute x whereas all the players 1 to i are in case (3) and contribute zero. It is straightforward that these are the only type of equilibria that exist where all contributing players contribute the same value.

Now, assume that there exist at least two contributing players i_1 and i_2 who contribute two values x_{i_1} and x_{i_2} at the equilibrium such that $x_{i_1} \neq x_{i_2}$. Without loss of generality we assume $x_{i_1} < x_{i_2}$. We start by showing that for every other contributing player i, their equilibrium contribution x_i is equal to x_{i_2}. We show this by contradiction: we assume that $x_i \neq x_{i_2}$.

If $x_i < x_{i_2}$, then $x_i < \max_{j \neq i} x_j$ so we must be in case (2) of player i's strategy defined above ($x_i > 0$ by assumption). We conclude that $x_i = \beta_i - \sum_{j \neq i} x_j$, which tells us that $\beta_i = \sum_{j=1}^n x_j$. In a similar way, we have that $x_{i_1} = \beta_{i_1} - \sum_{j \neq i_1} x_j$ and thus $\beta_{i_1} = \sum_{j=1}^n x_j = \beta_i$, so players i_1 and i have the same benefit-cost ratio which contradicts our assumption (Definition 1). This also shows that there cannot exist two contributing players whose contributions are strictly less than

$\max_j(x_j)$ unless they have the same benefit-cost ratio. (This will later be used as Lemma 1.)

It must be that $x_i > x_{i_2}$, then $x_{i_2} < \max_j x_j$ and the exact same analysis as above can be applied to x_{i_2}. This leads us to $\beta_{i_2} = \beta_{i_1}$. This is a contradiction, meaning that $x_i \leq x_{i_2}$ and, therefore, $x_i = x_{i_2}$.

Thus at the equilibrium there can exist only two different non-zero contributions possible.

We assume that there exists 1 players contributing x_{i_1} and $M = n - 1$ contributing x_{i_2}. According to player i_1's' strategy (case (2)) we have the following.

$$x_{i_1} = \beta_{i_1} - Mx_{i_2} \tag{14}$$

$$\tag{15}$$

Because player i_1 is a contributing player, we have by assumption that $x_{i_1} > 0$, which implies $x_{i_2} < \frac{\beta_{i_1}}{M}$. We also have by assumption that $x_{i_1} < x_{i_2}$, which implies $\beta_{i_1} - Mx_{i_2} < x_{i_2}$, which in turn implies $\frac{\beta_{i_1}}{n} < x_{i_2}$.

According to condition 12, applied to player i_2, we have that $Mx_{i_2} + x_{i_1} \leq \beta_{i_2}$. Replacing the value of x_{i_1} in this inequality leads to $\beta_{i_1} \leq \beta_{i_2}$. We, therefore, have $\frac{\beta_{i_1}}{n} < x_{i_2} < \frac{\beta_{i_1}}{M} \leq \frac{\beta_{i_2}}{M}$.

It is straightforward to verify that having one player with the smallest benefit-cost ratio play x_{i_1} and the rest plays x_M is a Nash equilibrium.

C   Proof of Theorem 3

With the addition of the reward, the utility function for player 1 is the following.

$$u_1(x_1) = \begin{cases} \ln(\frac{x_1+x_2}{\max(x_1,x_2)})v_1 - c_1 x_1 + R\frac{x_1}{x_1+x_2} \text{ if } \max(x_1,x_2) > 0 \\ 0 \text{ if } \max(x_1,x_2) = 0 \end{cases} \tag{16}$$

If $x_1 < x_2$, $u_1 = \ln(\frac{x_1+x_2}{x_2})v_1 - c_1 x_1 + R\frac{x_1}{x_1+x_2}$ and $\frac{du_1}{dx_1} = \frac{v_1}{x_1+x_2} - c_1 + R\frac{x_2}{(x_1+x_2)^2}$.

We solve the inequality $\frac{du_1}{dx_1} > 0$, written in terms of $X = x_1 + x_2$.

$$\frac{du_1}{dx_1} > 0 \Leftrightarrow X^2 - \frac{v_1}{c_1}X - \frac{Rx_2}{c_1} < 0 \tag{17}$$

Consider $\Delta_1 = (\frac{v_1}{c_1})^2 + 4\frac{Rx_2}{c_1}$ and $X_\pm = \frac{1}{2}(\frac{v_1}{c_1} \pm \sqrt{\Delta_1})$, where $X_\pm$ are the roots of the quadratic equation $X^2 - \frac{v_1}{c_1}X - \frac{Rx_2}{c_1} = 0$. We can rewrite inequality 17 in terms of $X_\pm$ as $(X - X_-)(X - X_+) > 0$. Since $X_- < X_+$, the solution to this inequality is $X_- < X < X_+$. We also note that $\sqrt{\Delta_1} > \frac{v_1}{c_1}$ and hence $X_- < 0$. Since we also have $X \geq 0$, we can conclude the following for a fixed x_2.

$$\frac{du_1}{dx_1} > 0 \Leftrightarrow x_1 + x_2 < \frac{1}{2}(\frac{v_1}{c_1} + \sqrt{\Delta_1}) \tag{18}$$

If $x_1 > x_2$, $u_1 = \ln(\frac{x_1+x_2}{x_1})v_1 - c_1 x_1 + R\frac{x_1}{x_1+x_2}$ and $\frac{du_1}{dx_1} = v_1(\frac{1}{x_1+x_2} - \frac{1}{x_1}) - c_1 + R\frac{x_2}{(x_1+x_2)^2}$. In that case we have $\frac{du_1}{dx_1} > 0$ iff $\frac{x_2(x_1(R-v_1)-v_1 x_2)}{x_1(x_1+x_2)^2} - c_1 > 0$ and, therefore, $\frac{du_1}{dx_1} < 0$ if $R < v_1$.

Assume that (x_{1_0}, x_{2_0}) are the equilibrium values. If $x_{1_0} > \frac{1}{2}(\frac{v_2}{c_2} + \sqrt{\Delta_2}) = \frac{1}{2}(\frac{v_2}{c_2} + \sqrt{(\frac{v_2}{c_2})^2 + 4\frac{Rx_{1_0}}{c_2}})$ then $x_2 \mapsto u_2$ is decreasing (due to the result derived above and the assumption that $R < \min(v_1, v_2)$) and thus player 2 maximizes their utility by contributing $x_{2_0} = 0$. If $x_{2_0} = 0$, player 1 is better off contributing a very small amount to be sure to get the reward while minimizing their cost, so $x_{1_0} \approx 0$. This contradicts the condition $x_{1_0} > \frac{1}{2}(\frac{v_2}{c_2} + \sqrt{\Delta_2})$, which means that $x_{1_0} < \frac{1}{2}(\frac{v_2}{c_2} + \sqrt{\Delta_2})$. The exact same argument can be made to derive $x_{2_0} < \frac{1}{2}(\frac{v_1}{c_1} + \sqrt{\Delta_1})$.

Assume now that $x_{1_0} \neq x_{2_0}$ and, without loss of generality, that $\max(x_{1_0}, x_{2_0}) = x_{1_0}$. Since $R < v_1$ then $\frac{du_1}{dx_1} < 0$ for $x_1 \geq x_{2_0}$ and hence player 1 best strategy is to contribute $x_{1_0} \leq x_{2_0}$ which contradicts our assumption. We thus conclude that $x_{1_0} = x_{2_0}$.

Combining $x_{1_0} = x_{2_0} = x$ with the two inequalities derived in the first part of this proof, we have the following bounds on x.

$$\begin{cases} x < \frac{1}{4}(\frac{v_1}{c_1} + \sqrt{(\frac{v_1}{c_1})^2 + 4\frac{Rx}{c_1}}) \\ x < \frac{1}{4}(\frac{v_2}{c_2} + \sqrt{(\frac{v_2}{c_2})^2 + 4\frac{Rx}{c_2}}) \end{cases} \tag{19}$$

We now derive a closed-form solution for the constraint on x. To solve the Inequality System 19, we write $X = \sqrt{(\frac{v_1}{c_1})^2 + 4\frac{Rx}{c_1}}$ and, for ease of notation, write $c = c_1$ and $v = v_1$. Squaring and expanding X, one finds that $4x = \frac{c}{R}(X^2 - (\frac{v}{c})^2)$. Accordingly, rewriting the Inequality System 19 in terms of X gives us the following.

$$\frac{c}{R}X^2 - X - \frac{v^2}{Rc} - \frac{v}{c} < 0 \tag{20}$$

With $\Delta' = 1 + 4\frac{c}{R}(\frac{v^2}{Rc} + \frac{v}{c})$, we have that $\Delta' > 0$ so the inequality can be simplified as $(X - X'_-)(X - X'_+) < 0$ with $X'_\pm = R\frac{1\pm\sqrt{\Delta'}}{2c}$. Since $X'_- < X'_+$, the solution to Inequality 20 is $X'_- < X < X'_+$. As $X'_- > 0$, we have $(X'_-)^2 < X^2 < (X'_+)^2$ and, therefore, the solutions to the Inequality System 19 are the following, with $\Delta'_1 = 1 + 4\frac{c_1}{R}(\frac{v_1^2}{Rc_1} + \frac{v_1}{c_1})$ and $\Delta'_2 = 1 + 4\frac{c_2}{R}(\frac{v_2^2}{Rc_2} + \frac{v_2}{c_2})$.

$$\begin{cases} \frac{c_1}{4R}((R\frac{1-\sqrt{\Delta'_1}}{2c_1})^2 - (\frac{v_1}{c_1})^2) < x < \frac{c_1}{4R}((R\frac{1+\sqrt{\Delta'_1}}{2c_1})^2 - (\frac{v_1}{c_1})^2) \\ \frac{c_2}{4R}((R\frac{1-\sqrt{\Delta'_2}}{2c_2})^2 - (\frac{v_2}{c_2})^2) < x < \frac{c_2}{4R}((R\frac{1+\sqrt{\Delta'_2}}{2c_2})^2 - (\frac{v_2}{c_2})^2) \end{cases} \tag{21}$$

D Proof of Theorem 4

Consider the strategy of player $n + 1$ (the new player) as outlined in 10. In the 1-value equilibrium we have $\sum_{i=1}^{n} x_i = nx_{eq}$.

1. If $nx_{eq} \geq \beta$, player $n + 1$ is not incentivised to contribute anything so the equilibrium stays unchanged.
2. If $nx_{eq} < \beta$, there are two cases to consider.
   (a) If $x_{eq} < \beta - nx_{eq}$ then player $n + 1$ contributes x_{eq}. Since $nx_{eq} < \beta_j$ (from 4) this means that for every player $1 \leq j \leq n$ we are in case (1) of their strategy. They only change their contribution after the introduction of player $n + 1$ if $\beta_j - nx_{eq} < x_{eq}$ or equivalently $\beta_j < (n + 1)x_{eq}$. Since $\beta_1 \leq \beta_j$, there is at least one player who will change their contribution in this case if and only if $\beta_1 < (n + 1)x_{eq}$, in which case we also have $\beta_1 < \beta$ from $x_{eq} < \beta - nx_{eq}$.
   (b) If $x_{eq} \geq \beta - nx_{eq}$ then player $n + 1$ contributes $\beta - nx_{eq}$. As before, player j only changes their contribution after the introduction of player $n + 1$ if $\beta_j - (n - 1)x_{eq} - (\beta - nx_{eq}) < x_{eq}$ or equivalently if $\beta_j < \beta$. This happens only if $\beta_1 < \beta$. In this case, we also have $\beta_1 < \beta < (n + 1)x_{eq}$ from $x_{eq} \geq \beta - nx_{eq}$.

A necessary and sufficient condition to having one player changing their contribution is, therefore, $\beta > \beta_1$ and $\frac{1}{n+1}\beta_1 < x_{eq} < \frac{1}{n}\beta$.

E   Proof of Theorem 5

From Equation 6, we have $\sum_{i=1}^{n} x_i = \beta_1$. For any new player $n + 1$ playing $x_{n+1} > 0$ we will, therefore, have $\sum_{i=1}^{n+1} x_i > \beta_1$, so according to 10 the best strategy for player 1 would now be to free ride. According to 10, the new player contributes if and only if $\sum_{i=1}^{n} x_i < \beta$ or equivalently $\beta_1 < \beta$. Under this condition, a new player disrupts the equilibrium, which causes player 1 to free ride.

F   Proof of Theorem 6

If a player leaves the game by no longer contributing, we have the following. In the $1-$value equilibrium case, since $x_{eq} \leq \frac{1}{n-1}\beta_1 < \frac{1}{n-1}\beta_2 < \ldots < \frac{1}{n-1}\beta_n$ (implied by Condition 4), a player leaving the game does not change the equilibrium (as specified by 10). Hence a player leaving the game does not disrupt the equilibrium.

G   Proof of Theorem 7

By assumption, every player (other than player k) is playing x_{eq}. After the deviation happens, but before any other changes, players k is playing x_{k_0} and the $n-2$ other players are still contributing x_{eq}. Each player's new best response then becomes $x_{i,\text{new}} = \max\{0, \min(\max(x_{eq}, x_{k_0})), (\beta_i - x_{k_0}) - (n - 2)x_{eq}\}$ due to Condition 10.

Consider the following cases.

1. If $x_{k_0} < x_{eq}$ then $\max(x_{eq}, x_{k_0}) = x_{eq}$. For each player i we have $x_{k_0} + (n-1)x_{eq} \leq nx_{eq} \leq \beta_i$ and thus $(\beta_i - x_{k_0}) - (n-2)x_{eq} \geq x_{eq}$. $(\sum_{j \neq i} x_j < \beta_j$ still holds with $x_{k_0} < x_{eq}$.) This means each player's best response is not impacted. Therefore, x_{eq} is still the equilibrium.

2. If $x_{k_0} > x_{eq}$ then we have that $\max(x_{eq}, x_{k_0}) = x_{k_0}$. Therefore, Condition 10 becomes $\max\{0, \min(x_{k_0}, \beta_i - x_{k_0} - (n-2)x_{eq})\}$ for each player and for each player $i \neq k_0$ we have the following.

    (a) If $x_{k_0} \geq \beta_i - x_{k_0} - (n-2)x_{eq}$, or equivalently $x_{k_0} \geq \frac{\beta_i - (n-2)x_{eq}}{2}$, player i's new best response is to play $x_{i,\text{new}} = \beta_i - x_{k_0} - (n-2)x_{eq}$ or 0 if this value is negative. This disrupts the current equilibrium and all the rational players have to change their contribution accordingly.

    (b) If $x_{k_0} \leq \beta_i - x_{k_0} - (n-2)x_{eq}$, or equivalently $x_{k_0} \leq \frac{\beta_i - (n-2)x_{eq}}{2}$, player i's best response is to contribute x_{k_0}. Again, the equilibrium is disrupted.

To summarize, if a player deviates from the equilibrium and its new contribution is $x_{k_0} \leq x_{eq}$, the best responses of other players stay unchanged and the equilibrium is not disrupted. In the other case, if $x_{k_0} > x_{eq}$, the equilibrium changes.

H   Proof of Theorem 8

Assume that $k > 1$. In the case where players are in a $2-$value equilibrium, according to condition 10, player 1 should change its value to either $\beta_1 - (n-2)x_{eq} - x_{k_0}$ or to $\max(x_{eq}, x_{k_0})$ or to zero. Thus, player 1 changes their contribution.

In the case where player 1 is the deviating player (i.e., $k = 1$), then the other players change their contribution if and only if there exists a j such that $\beta_j - (n-2)x_{eq} - x_{k_0} < \max(x_{eq}, x_{k_0})$, or $\beta_j \leq (n-2)x_{eq} - x_{k_0}$ according to Condition 10.