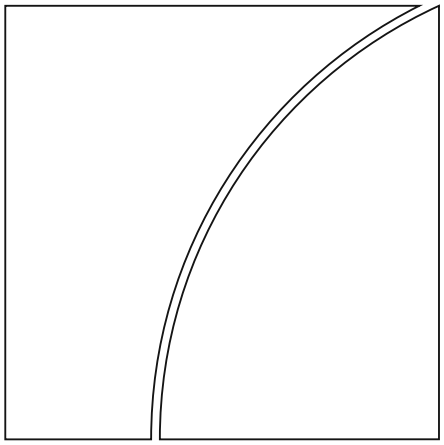

# BIS Working Papers

No 1117

# An impossibility theorem on truth-telling in fully decentralized systems

by Rodney Garratt and Cyril Monnet

Monetary and Economic Department

August 2023

JEL classification: C72, D72, D86, O33.

Keywords: decentralized systems, smart contracts, truth-telling, oracle problem.

BIS Working Papers are written by members of the Monetary and Economic Department of the Bank for International Settlements, and from time to time by other economists, and are published by the Bank. The papers are on subjects of topical interest and are technical in character. The views expressed in them are those of their authors and not necessarily the views of the BIS.

This publication is available on the BIS website (www.bis.org).



ISSN 1020-0959 (print)
ISSN 1682-7678 (online)

# An Impossibility Theorem on Truth-Telling in Fully Decentralized Systems*

Rodney Garratt† Cyril Monnet‡

August 2023

**Abstract**

We show that truthful reporting about the realization of a publicly observed event cannot be implemented as a unique equilibrium in a completely decentralized environment. Our work provides a theoretical underpinning of the need for oracles and the related "oracle problem."

**Keywords:** decentralized systems, smart contracts, truth-telling, oracle problem

**JEL Classification Numbers:** C72, D72, D86, O33

*The truth? What's that? Don't you know that the day has come when the truth is what we care to make it?*

– Iain Crichton Smith, *Consider the Lilies*

*The views expressed are those of the authors and do not necessarily reflect those of the BIS. We are grateful to Bruno Biais, Gabriele Camera, Jonathan Chiu, Johannes Hoerner, Thor Koeppl, Jean-Charles Rochet, Hyun Song Shin, Jens Witkowski and conference and seminar participants at the 2022 Theory and Experiments in Monetary Economics conference at George Mason University, Tokenomics 2022, the 2023 ASSA Winter Meetings, the Frankfurt School of Finance and Management, the University of Bonn, and the Toulouse School of Economics for helpful comments.

†Bank for International Settlements. Email: rodney.garratt@bis.org.

‡University of Bern and Study Center Gerzensee. Email: cyril.monnet@unibe.ch

# 1 Introduction

Smart contracts are self-executing programmable contracts between two or more parties. Smart contracts do not require a vetting authority because their legitimacy relies on decentralized ledger technologies. However the implementation of many potentially useful smart contract applications depends upon verifying that some real-world event has taken place (think insurance contracts). This is a problem. Given their fully decentralized nature, how does a smart contract select what the true state of the world is? More generally, how do fully decentralized systems function when their operation depends on the existence of a single, mutually accepted, record of the truth, but there is no single authority that can provide this record?

This paper considers a situation where multiple individuals seek to enter into agreements based on the outcome of a real-world event, but there is no trusted party (i.e., contractible source) that can be used to determine payoffs. In this case, payoffs must be based on some form of collective agreement on the true state of the world. By appealing to three basic properties, anonymity, neutrality, and monotonicity, we can restrict attention to majority voting. In this environment, agents report the truth that is most beneficial for them. Incentives to reward consensus do not necessarily make things better. Rather, they lead to a situation that is akin to a beauty contest à la Keynes (1936), in which agents report what they think the majority of other agents will report. With or without reporting incentives, the report of the true state that results from majority voting does not depend upon the true state.

Our main analysis focuses on simultaneous voting. However, we also briefly explore the possibility of sequential play, by allowing individuals to vote in a random order and allowing each individual to see the previous votes. In a three-agent model with sequential voting individuals unanimously agree on what the state of the world is, but this may not be the true state. The result suggests that sequential voting is not a solution to the truth-telling problem.

Our general result, which applies to simultaneous voting games, is that the only way that individuals are willing to vote according to the true state is if they are completely indifferent as to what the true state should be. That is, their payoffs cannot depend on their actions or their individual reports. This general result suggests that absent additional motivation (e.g., deontological preferences toward truth-telling, see Bergstrom et al. (2019)) that links individual payoffs to the truth, there is no way to implement contracts that pay out based on an observed state without a trusted source.

Our analysis has aspects in common with peer prediction games (PPG); see, for example Gao et al. (2016) and Kong and Schoenbeck (2019) and the references

therein. In these settings, there is typically one object that has an unknown characteristic and several agents can exert effort to obtain a signal about this characteristic. These signals can be correlated (i.e, come from the same distribution), but there is typically no known "ground truth." The literature studies peer prediction mechanisms that are used to make sure agents exert effort and that they communicate their signals truthfully.

An aspect that makes both our game and PPG similar is that in both cases agents have to send signals about what they observe. Also neither the peer prediction mechanism nor ours can rely on the ground truth (even if it exists) to discipline agents. There are, however, important differences between PPGs and our game. Our game is one of common knowledge (everybody knows the true state, and everybody knows that everybody knows, etc.), while there is no common knowledge in PPG. This is in part because there may be no objective ground truth, rather the truth may be subjective. In addition, in PPGs, agents do not know the signals of others. Trivially, our agents do not need to exert effort to coordinate on what they commonly and freely observe. Our agents have an ex-ante stake in the "truth," while in PPG the only reward is the one agents obtain by "getting it right" relative to other agents. Finally, in our game, the reward is idiosyncratic and depends on "types".

The importance of the ability to write contracts based on observed states has long been recognized in the economics literature. Radner (1968) formalized the requirement for state verification in classic competitive analysis by limiting contracts to events that could be verified by all parties. Scenarios where parties are asymmetrically informed were first analyzed by Townsend (1979) and Gale and Hellwig (1998), who use the implied limitations of asymmetric information to explain the prevalence of simple debt contracts. This literature, including the many papers that followed these seminal works, recognizes the issues that arise when state verification is costly, but typically assumes one party (who may determine the state through private actions) knows the true state and that this is not easily verified by the other party. The problem we consider is quite different in that there is no informational asymmetry about the realization of the true state and the true state is determined exogenously. It is also distinct from the literature on incomplete contracts (e.g., Grossman and Hart (1986), Hart and Moore (1999), Tirole (1999)), which has more to do with the inability to specify all contingencies, or the costs associated with doing so.

This work is relevant to decentralized finance applications in which contracts are written with payoffs that depend on the outcomes of real world events. Our works provides a theoretical underpinning of the need for oracles and the related "oracle problem" (Caldarelli (2020)).

# 2 A coin toss example

We start with an example in which an odd number $N$, of agents bet on the outcome of a public coin toss. The modelling begins after the bets are made. These bets are determined by exogenous factors that are outside the model. In fact, it is fine to assume that these bets are determined by agents' own independent coin tosses. The purpose is to illustrate the difficulty of eliciting the truthful outcome of the **publicly observable** coin toss from the population of bettors.

## 2.1 Simultaneous Voting

Each agent $i = 1, ..., N$'s initial bet is denoted by $b_i \in \{H, T\}$, where $H$ stands for heads and $T$ stands for tails. Let $\boldsymbol{b} = \{b_1, ..., b_N\}$ be the betting profile. The fraction of agents who bet on heads is denoted by $\beta = \sum_i \mathbb{I}\{b_i = H\}/N$, where $\mathbb{I}\{x\}$ is the indicator function. Following their bet, agents learn the aggregate betting record $\beta$, but no individual bets. Then nature tosses the coin with realization $\theta \in \{H, T\}$. This realization is publicly observable, but not directly contractible. Rather, the outcome of the coin toss, for the purposes of paying off bets, must be determined by some voting procedure. So agent $i$'s vote is, at least in principle, a function $v_i(b_i, \beta, \theta)$ that maps agent $i$'s initial bet, the aggregate betting record, and the public realization of the coin toss into $\{H, T\}$. We assume the voting rule is majority voting. This is justified by appealing to May's Theorem (May (1952)):

**May's Theorem.** *Consider any 2-candidate election with an odd number of voters. If the voting system satisfies (1) equal treatment of voters (anonymity), (2) equal treatment of candidates (neutrality), and (3) monotonicity (if A wins with a votes, then A wins with a+1 votes), then the voting system is simple majority.*

Given a complete set of votes $\boldsymbol{v} = \{v_1(b_1, \beta, \theta), ..., v_N(b_N, \beta, \theta)\}$ the recorded realization of the public coin toss is

$$r(\boldsymbol{v}) \equiv \begin{cases} H \text{ if } \sum_i \mathbb{I}\{v_i = H\} > \frac{N}{2} \\ T \text{ otherwise} \end{cases}$$

In words, the recorded realization is $H$ if the majority of voters says it is $H$, and it is $T$ otherwise. This recorded realization is all that matters for the purpose of paying off bets. Agent $i$ is a winner whenever $b_i = r(\boldsymbol{v})$, otherwise agent $i$ is not a winner.

Let $W(\boldsymbol{b}, \boldsymbol{v}) = \sum_i \mathbb{I}\{b_i = r(\boldsymbol{v})\}$ denote the number of winners. Moreover, let $\gamma$ denote that fraction of bets that are paid to the winners. The fraction $1 - \gamma$ is retained to incentivize reporting. This amount is analogous to what betting houses call the "vigorish", or "vig" for short. Winners each get

$$w(W) = \frac{\gamma N}{W(\boldsymbol{b}, \boldsymbol{v})}$$

with $\gamma \leq 1$. If $\gamma < 1$, then agents also get a reward for reporting in line with the majority. Let $M(\boldsymbol{b}, \boldsymbol{v}) = \sum_i \mathbb{I}\{v_i = r(\boldsymbol{v})\}$ denote the number of agents that vote with the majority. Then, only agents that vote with the majority get a reward

$$\pi(M) = \frac{(1-\gamma)N}{M(\boldsymbol{b}, \boldsymbol{v})}.$$

Combining everything, the payoff of agent $i$ is

$$V_i = w\mathbb{I}\{b_i = r(\boldsymbol{v})\} + \pi\mathbb{I}\{v_i = r(\boldsymbol{v})\}.$$

Our first result assumes voting is done simultaneously.[1]

**Conjecture 1** *If $\gamma = 1$, then there is a Bayesian Nash equilibrium in weakly dominant strategies where $v_i(b_i, \beta, \theta) = b_i$ and hence $r(\boldsymbol{v}) = H$ if $\beta > \frac{1}{2}$ and $T$ otherwise. In this case, voting strategies do not depend on the publicly observed outcome of the coin toss. If $\gamma < 1$, then there are two Bayesian Nash equilibria, one in which everyone votes $H$ and one in which everyone votes $T$. In this case, voting strategies do not depend on betting positions or the publicly observed outcome of the coin toss.*

**Proof.** If $\gamma = 1$, then agents' payoffs depend entirely on whether their initial bets match $r(\boldsymbol{v})$, the state that is decided upon by majority voting. Consider an agent $i$ who bet $b_i$, and who considers voting $H$ or $T$. The only time the agent's vote matters for their own payoff is when it is pivotal and beneficial; that is it changes the outcome from $r(\boldsymbol{v}) \neq b_i$ to $r(\boldsymbol{v}) = b_i$. In such a circumstance the agent strictly prefers to vote the same way that they bet, and in all other circumstances the agent's vote does not matter. Hence $v_i(b_i, \beta, \theta) = b_i$ for all $i$ and all triples $(b_i, \beta, \theta)$ is a weakly dominant strategy. Given that agents vote their own type, the unique equilibrium will have $r(\boldsymbol{v}) = H$ when $\beta > 1/2$ and $r(\boldsymbol{v}) = T$ otherwise, as stated in the conjecture.

[1] The definition of Bayesian Nash Equilibrium for this game is standard. A formal definition is provided using more general notation in Section 3.

If $\gamma < 1$, then agent's can earn a positive payoff by voting with the majority, even if their bet does not match the majority report. If an agent placed the bet $b_i = T$, then they would strictly prefer to vote for $T$ whenever they expect the number of others voting for $T$ to be greater than or equal to $\frac{N-1}{2}$. Otherwise, the type $b_i = T$ agent would strictly prefer to vote for $H$. Likewise, if an agent placed the bet $b_i = H$, then they would strictly prefer to vote for $H$ whenever they expect the number of others voting for $H$ to be greater than or equal to $\frac{N-1}{2}$. Otherwise, the type $b_i = H$ agent would strictly prefer to vote for $T$. Given that voting now depends on expectations regarding how others are voting there are two equilibria: $v_i(b_i, \beta, \theta) = H$ for all $i$ and all triples $(b_i, \beta, \theta)$ and $v_i(b_i, \beta, \theta) = T$ for all $i$ and all triples $(b_i, \beta, \theta)$. ■

Intuitively, the result for $\gamma = 1$ holds because voting according to how they bet, can never lower an agent's payoff and sometimes it might help them win. If everyone votes according to how they bet, then the record will match the bet of the majority. If $\gamma < 1$, then agents no longer have a weakly dominant strategy to vote according to their own bet. This can be costly if the majority votes differently.

*Remark 1.* The evidence in Bergstrom et al. (2019)) among others, point to the fact that some agents may have a preference or compulsion for telling the truth. Would our results stand if a fraction $\tau$ of agents from a [0,1]-continuum always tell the truth? If $\tau$ is common knowledge, we show that they indeed hold as long as $\tau$ is not too large relative to the majority of types. Without loss of generality, suppose there is a fraction $\beta > 1/2$ of agents who bet on heads. If the outcome of the coin flip is is tails, a measure $\tau\beta$ of "heads" agents will cast a tail vote because these agents always tell the truth. Hence, the rest of the agents that bet on heads (who do not necessarily feel compelled to tell the truth) will vote for heads whenever: $\beta(1-\tau) > (1-\beta) + \tau\beta$ or $\tau < 1 - 1/(2\beta)$, and they will vote tails otherwise. If $\beta$ is large, the measure of truth-tellers has to be large to reverse the incentives of other agents. In particular, if $\beta$ is arbitrarily close to 1/2, we obtain truth-telling for all $\tau > 0$, because in that case the truth-tellers become pivotal.

## 2.2 Sequential Voting

So far we have analyzed a game where agents vote simultaneously on the observed outcome of the coin toss. In that context, the outcome of the game is indeterminate because it depends on agents' beliefs. We now attempt to solve this indeterminacy by having agents vote sequentially. Intuitively, with sequential voting, agents can

update their beliefs based on the votes that have already been made. We show that while sequential voting may break the indeterminacy, agents still cannot coordinate on the publicly observed outcome.

Suppose that nature randomly orders agents to sequentially cast a vote. The agent in place $s$ votes $v_s(b_s, \beta, \theta, c^{s-1}) \in \{H, T\}$, after observing $\theta$, $\beta$, and the count $c^{s-1} = \sum_{t=1}^{s-1} \mathbb{I}\{v_t = H\}$ of number of $H$ votes from agents voting before her, where $s = 1$ means the agent is the first to vote and $s = N$ means the agent is the last one to vote, and $c^0 = \emptyset$.

In the simultaneous voting game of the previous section, each individual agent's private information about how they bet was irrelevant. In the sequential setting the private information about individual bets matters since agents must draw inferences about whether agents are following particular type-contingent strategies. These inferences determine the probabilities agents assign to being at any particular information set in the game.

We now consider a game in which an agent's strategy specifies their action at each information set, and agents have beliefs about which node they are at within each information set. The solution concept we use is perfect Bayesian equilibrium (PBE).

**Definition 1** *A perfect Bayesian equilibrium of the sequential voting game is a strategy profile* $\{v_s(b_s, \beta, \theta, c^{s-1})\}_{s=1,..,N}$ *and agent beliefs such that (1) at each information set, each agent's strategy is a best response to the strategies of the other agents given their beliefs and, and (2) given the strategy profile, the beliefs are consistent with Bayes' rule whenever possible.*

**Conjecture 2** *In a three-agent model with sequential voting and* $\frac{1}{4} < \gamma < 1$ *there is a unique perfect Bayesian equilibrium in which individuals unanimously agree that the outcome of the coin toss matches the bet placed by the majority of bettors.*

Conjecture 2 says that we can obtain a unique equilibrium, but it will only involve truth-telling when the bet placed by the majority or bettors happens to be correct. To prove this conjecture, we show that there is a unique PBE where the recorded realization is the bet placed by the majority of bettors independently of the true realization $\theta$ of the coin toss, and the vote is unanimous. For this result to hold, we require that the incentives to vote with the majority, as determined by the value $1 - \gamma$, are non zero, but sufficiently small.

**Proof.** Suppose $\gamma < 1$ so there are rewards for voting like the majority does. Suppose there are two agents whose initial bet is $T$ and one agent whose initial bet is $H$. Since

aggregate bets are known, the agent whose bet is $H$ must assign probability 1 to each other agent being a type $T$ bettor regardless of how they vote.

We are looking for an equilibrium where, given that $\beta = \frac{1}{3}$, $v_s(b_s, \beta, \theta, c^{s-1}) = T$ for all $b_s$ and $\theta$; that is everyone votes $T$. We assume off path beliefs are defined as follows: an agent of type $b_s = T$ who sees someone vote $H$ assumes that person's initial bet is $H$ with probability 1. (Note that there are no other off-path beliefs to define since an agent who bet $H$ knows how the other agents bet regardless of how they vote.)

We can think in terms of the extensive form game that starts with nature selecting one of three paths for the order of sequential voting ($HTT$ when the agent who votes first bet $H$, $THT$ when the agent who votes second bet $H$, and $TTH$ when the agent who votes last bet $H$), all with equal probability.

Consider the third voter. On the equilibrium path this agent sees two $T$ votes. If she bet $H$, she knows which information set she is at and votes $T$. This earns her a payoff of $\pi(3)$. If she bet $T$, she assigns equal probability to two possible voting paths but still votes $T$ since this yields the highest possible payoff of $w(2) + \pi(3)$. Off the equilibrium path, the third voter sees at least one $H$ vote and the agent who bet $H$ and votes last will vote $H$ (since there is already at least one other $H$ vote). An agent who bet $T$ and votes last will vote $T$ unless the off path voting history is $(H, H)$ in which case she will vote $H$ as well and earn $\pi(3)$.

Next, consider the voting decision of the second voter. On the equilibrium path she sees a vote of $T$ by the first voter. If she bet $T$, she assigns equal weight to the two voting paths (which depend on the unknown bettor type of the first voter) and votes $T$, as this guarantees her the highest possible payoff of $w(2) + \pi(3)$. If she bet $H$, then she knows the path, and she knows the third voter will vote $T$, so she also votes $T$ to earn a payoff of $\pi(3)$. Off the equilibrium path, if she sees the first voter vote $H$ and she bet $T$ then she believes with probability 1 that the first voter bet $H$. Hence, it is best to vote $T$ since she believes the third voter must have bet $T$. If the second voter bet $H$ and she sees the first voter vote $H$, then she votes $H$ to earn the maximum payoff $w(2) + \pi(3)$.

Finally, consider the voting decision of the first voter. If the first voter bet $H$, then she knows the next two voters bet $T$ and that they will vote $T$, regardless of how she votes, so the first voter votes $T$ to earn the payoff $\pi(3)$. If the first voter bet $T$, she votes $T$ expecting the two remaining voters to follow through with the equilibrium strategies.

That establishes the existence of a PBE in which everyone votes $T$. Since there are only two BE in the simultaneous game (Conjecture 1), and since PBE is a refinement

of BE, uniqueness is established if we can show there is no PBE in which everyone votes $H$.

Suppose $v_s(b_s, \beta, \theta, c^{s-1}) = H$ for all $\theta$ and $s$ and that on path beliefs are given by Bayes rule and off path beliefs are free to be chosen.

Consider the last voter. On the equilibrium path this agent votes $H$ regardless of how she bet. Off the equilibrium path, she votes $T$ if she bet $T$ to gain maximum payoff. If the final voter bet $H$, then she votes $H$ if at least one of the previous voter voted $H$ (so in response to the sequence of votes $(T, H)$ or $(H, T)$) and $T$ otherwise(regardless of her beliefs).

Consider now the second voter. On the equilibrium path if the second voter bet $H$, she will vote $H$ and guarantee the maximum payoff. However if she bet $T$, she has to compute the expected payoff of each option. Since she believes the first voter (and hence the last voter) is equally likely to have placed each type of bet she votes $H$ if $\pi(3) \geq .5(\pi(2) + w(2))$. We can rule this out by assuming $\gamma$ is large enough so that $\pi(3) < .5(\pi(2) + w(2))$.

The implication is that even if the first voter goes along with the proposed equilibrium where everyone votes $H$, the second voter, regardless of her type, will not follow.

■

Interestingly, introducing sequential voting to the case with $\gamma < 1$ induces agents to condition their votes on their individual types (i.e., exogenously determined initial bets). That is, sequential voting allows agents to coordinate on one of the two equilibria that are present in the simultaneous game. Naturally, they coordinate on the equilibria that is preferred by most.

# 3 Impossibility result

The number of agents, $N$, is large and odd. Agent $i \in \{1, ..., N\}$ has type (previously called their bet) $b_i \in \{0, 1\}$. Let $\boldsymbol{b}$ be the randomly determined type-profile for this economy. Suppose nature (randomly) selects the state (previously the outcome of the coin toss) $\theta \in \{0, 1\}$, that is commonly known to be observed by all. Each of the $N$ agents have to announce what the state is. As before, the fraction of agents whose type is 1 is denoted by $\beta = \sum_i b_i/N$.

A decentralized mechanism for determining the agreed upon state is a message space $\mathcal{M}$, a state function $g : \mathcal{M}^N \to \{0, 1\}$ that maps the message profile $\boldsymbol{m} = (m_1, ..., m_N) \in \mathcal{M}^N$ into a state, and an outcome function $f : \{0, 1\}^N \times \{0, 1\} \to \mathbb{R}^N$ that maps type profile and the outcome of the messaging function into a balanced

vector of payments to the $N$ agents. The utility agent $i$ receives from the mechanism is written as $u\left(f_i\left(\boldsymbol{b}, g(\boldsymbol{m})\right)\right)$ where $u(\cdot)$ is increasing and concave in $f$, $u(0) = 0$. The mechanism is balanced whenever $\sum_i f_i(\boldsymbol{b}, g(\boldsymbol{m})) = 0$. The mechanism uniquely implements the true state $\theta$ whenever $g(\boldsymbol{m}) = \theta$ for all $\theta \in \{0, 1\}$ and $\beta \in [0, 1]$.

A strategy for agent $i$ is a message $m_i(b_i, \beta, \theta) \in \mathcal{M}$. Note that strategies can be type dependent and types are private information. The solution concept is therefore, Bayesian equilibrium (BE).

**Definition 2** *Given a commonly observed type statistic $\beta \in [0, 1]$ and state $\theta \in \{0, 1\}$, a Bayesian Nash equilibrium of the simultaneous voting game is a strategy profile $\{m_i(b_i, \beta, \theta)\}_{i=1,\ldots,N} \in \mathcal{M}^N$ such that for each agent $i$ we have that $m_i(b_i, \beta, \theta)$ maximizes the expected utility of agent $i$ with type $b_i$ given the strategies of the other agents and the probabilities of their type profiles.*

Note that even though individuals only know their own type and the summary statistic $\beta$, we assume that the mechanism can condition payoffs on individual types. This reflects only an assumption on how information that is received by the mechanism is transmitted to participants before voting. We consider this to be a more realistic assumption, consistent with, for example, how sports books report odds but not the bets of each individual.

We consider only symmetric mechanisms whereby all agents who are of the same type and who vote the same are treated the same. We also assume agents are individually rational in the sense that they are willing to play under the assumption that types will be randomly assigned (independently and with equal probability) and willing to vote after learning their own type $b_i$, the type statistic $\beta$ and $\theta$.

Here we argue the following:

**Theorem 1** *When there are two states, there is no individually rational and budget balanced mechanism that implements truthful reporting as a unique Bayesian Nash equilibrium in a completely decentralized environment.*

**Proof.** We show that the only outcome function for which agents are willing to report the truth is one where $f \equiv 0$ for all state realizations and type profiles. We call this the "no stake" outcome function. Since, in this situation, any report is equally good for all agents regardless of what other agents do, any strategy profile is a Nash equilibrium. Therefore, truthful reporting cannot be implemented as a unique equilibrium.

Without loss of generality we consider the case where $\mathcal{M} = \{0, 1\}$. We make the assumption that $g(\boldsymbol{m})$ has to satisfy (1) equal treatment of voters (anonymity), (2)

equal treatment of states (neutrality), and (3) monotonicity (if A wins with $a$ votes, then A wins with $a+1$ votes). According to May's theorem, $g$ has to satisfy simple majority, that is

$$g(\boldsymbol{m}) = \mathbb{I}\left\{\sum_i m_i \geq \frac{N}{2}\right\}.$$

By considering different realizations of the majority type and true state we can arrive at the desired result.

Suppose that a majority of agents have the type $b_i = 0$, i.e., $\beta < \frac{1}{2}$ (the alternative case would follow the same argument), and, for starters, that the realized state is $\theta = 1$. If an agent of type $b_i = 0$ prefers to send message $m = 0$, then the state $\theta = 1$ cannot be implemented. Therefore we need

$$u\left(f\left(\boldsymbol{b}, 1; 0\right)\right) \geq u\left(f\left(\boldsymbol{b}, 0; 0\right)\right),$$

where we use the notation $f(\boldsymbol{b}, g(\boldsymbol{m}); b_i)$ to identify agent $i$'s type in the profile $\boldsymbol{b}$.

Suppose now that a majority of agents have the type $b_i = 0$, but that the state is $\theta = 0$. By the same argument, if an agent of type $b_i = 0$ prefers to send the message $m = 1$, then the state $\theta = 0$ cannot be implemented. Therefore we need

$$u\left(f\left(\boldsymbol{b}, 0; 0\right)\right) \geq u\left(f\left(\boldsymbol{b}, 1; 0\right)\right)$$

Hence, we require

$$u\left(f\left(\boldsymbol{b}, 0; 0\right)\right) = u\left(f\left(\boldsymbol{b}, 1; 0\right)\right).$$

Hence $f\left(\boldsymbol{b}, 0; 0\right) = f\left(\boldsymbol{b}, 1; 0\right) = f(\boldsymbol{b})$, is the transfer to an agent of type $b = 0$ when they have the majority in type profile $\boldsymbol{b}$. The payoff $f(\boldsymbol{b})$ must be such that agents in the majority have an incentive to vote, that is $f(\boldsymbol{b}) \geq 0$. Since the mechanism is balanced,

$$f(\boldsymbol{b}, 0; 1) \;\; = f(\boldsymbol{b}, 1; 1) = \;\; -\frac{\sharp b_0}{\sharp b_1} f\left(\boldsymbol{b}\right).$$

where $\sharp b_0$ is the number of agents of type 0 in the type profile $\boldsymbol{b}$, and $\sharp b_1$ is the number of agents of type 1 in the same type profile.

Let $q(\boldsymbol{b})$ denote the probability of type profile $\boldsymbol{b}$. Then the participation constraint of agents at the start of the game (before they know their type) is

$$Prob(b(i)=0)\left\{\sum_{\{\boldsymbol{b}:\sharp b_0>\sharp b_1\}} q(\boldsymbol{b})u(f(\boldsymbol{b})) + \sum_{\{\boldsymbol{b}:\sharp b_0<\sharp b_1\}} q(\boldsymbol{b})u\left(-\frac{\sharp b_1}{\sharp b_0}f\left(\boldsymbol{b}\right)\right)\right\} +$$
$$Prob(b(i)=1)\left\{\sum_{\{\boldsymbol{b}:\sharp b_1>\sharp b_0\}} q(\boldsymbol{b})u(f(\boldsymbol{b})) + \sum_{\{\boldsymbol{b}:\sharp b_1<\sharp b_0\}} q(\boldsymbol{b})u\left(-\frac{\sharp b_0}{\sharp b_1}f\left(\boldsymbol{b}\right)\right)\right\} \geq 0$$

Arranging terms, and using $p = Prob(b(i)=0)$,

$$\sum_{\{\boldsymbol{b}:\sharp b_0>\sharp b_1\}} q(\boldsymbol{b})\left[pu(f(\boldsymbol{b})) + (1-p)u\left(-\frac{\sharp b_0}{\sharp b_1}f\left(\boldsymbol{b}\right)\right)\right] +$$
$$\sum_{\{\boldsymbol{b}:\sharp b_0<\sharp b_1\}} q(\boldsymbol{b})\left[(1-p)u(f(\boldsymbol{b})) + pu\left(-\frac{\sharp b_1}{\sharp b_0}f\left(\boldsymbol{b}\right)\right)\right] \geq 0$$

Then using Jensen's inequality on each terms in square-brackets, for instance for $\boldsymbol{b}:\sharp b_0>\sharp b_1$,

$$pu(f(\boldsymbol{b})) + (1-p)u\left(-\frac{\sharp b_0}{\sharp b_1}f\left(\boldsymbol{b}\right)\right) <$$
$$u\left(pf(\boldsymbol{b}) - (1-p)\frac{\sharp b_0}{\sharp b_1}f\left(\boldsymbol{b}\right)\right) = u\left(\left(p-(1-p)\frac{\sharp b_0}{\sharp b_1}\right)f\left(\boldsymbol{b}\right)\right)$$

Since $p=1/2$, $\boldsymbol{b}$ is such that $\sharp b_0 > \sharp b_1$, and $f\left(\boldsymbol{b}\right)\geq 0$, while $u(0)=0$, the last term in the utility function is negative. Therefore, the participation constraint can never be satisfied unless $f(\boldsymbol{b})=0$. As claimed, the only outcome function for which people are willing to report the truth is the no stake outcome function. ■

# 4 Conclusion

A completely decentralized system can only function well when agents have no stake in it. While we express this result as an impossibility theorem, it really points to the limit of decentralization: the decentralization of the truth has to be handed to a group of agents who have no other stake in what the truth is than the payment

they receive to do their jobs. This opens the door to a free rider problem and the so-called oracle problem. We leave the study of the best organization for determining what the true state of the world is to future research.

# References


Bergstrom T, Garratt R, Leo G. 2019. Let me, or let george? motives of competing altruists. *Games and Economic Behavior* **118**: 269–283.

Caldarelli G. 2020. Understanding the blockchain oracle problem: A call for action. *Information* **11**: 509.

Gale D, Hellwig M. 1998. Incentive-compatible debt contracts: The one-period problem. *Review of Economic Studies* **52**: 647–664.

Gao AX, Wright JR, Leyton-Brown K. 2016. Incentivizing evaluation via limited access to ground truth: Peer-prediction makes things worse. *arXiv preprint arXiv:1606.07042* .

Grossman SJ, Hart OD. 1986. The costs and benefits of ownership: A theory of lateral and vertical integration. *Journal of Political Economy* **94**: 691–719.

Hart O, Moore J. 1999. Foundations of incomplete contracts. *The Review of Economic Studies* **66**: 115–138.

Keynes JM. 1936. *The General Theory of Employment, Interest and Money*. New York: Harcourt Brace and Co.

Kong Y, Schoenbeck G. 2019. An information theoretic framework for designing information elicitation mechanisms that reward truth-telling. *ACM Transactions on Economics and Computation* **7(1)**: 2:1–2:33.

May KO. 1952. A set of independent necessary and sufficient conditions for simple majority decisions. *Econometrica* **20**: 680–6847.

Radner R. 1968. Competitive equilibrium under uncertainty. *Econometrica* **36**: 31–58.

Tirole J. 1999. Incomplete contracts: Where do we stand? *Econometrica* **67**: 741–781.

Townsend RM. 1979. Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory* **21**: 265–293.

## Previous volumes in this series

| | | |
|---|---|---|
| 1116<br>August 2023 | Absolute blockchain strength? Evidence from the ABS market in China | Jing Liu, Ilhyock Shim and Yanfeng Zheng |
| 1115<br>August 2023 | Sharks in the dark: quantifying HFT dark pool latency arbitrage | Matteo Aquilina, Sean Foley, Peter O'Neill and Thomas Ruf |
| 1114<br>August 2023 | The term structure of inflation forecasts disagreement and monetary policy transmission | Alessandro Barbera, Fan Dora Xia and Xingyu Sonya Zhu |
| 1113<br>August 2023 | To Lend or Not to Lend: the Bank of Japan's ETF purchase program and securities lending | Mitsuru Katagiri, Junnosuke Shino and Koji Takahashi |
| 1112<br>July 2023 | Trust bridges and money flows | Tobias Adrian, Rodney Garratt, Dong He, and Tommaso Mancini-Griffoli |
| 1111<br>July 2023 | How much do firms need to satisfy the employees? – Evidence from credit spreads and online employee reviews | Koji Takahashi and Sumiko Takaoka |
| 1110<br>July 2023 | Fiscal sources of inflation risk in EMDEs: the role of the external channel | Ryan Banerjee, Valerie Boctor, Aaron Mehrotra and Fabrizio Zampolli |
| 1109<br>July 2023 | Original sin redux: role of duration risk | Carol Bertaut, Valentina Bruno and Hyun Song Shin |
| 1108<br>July 2023 | Innovation convergence | Bryan Hardy and Can Sever |
| 1107<br>July 2023 | Financial heterogeneity and monetary union | Simon Gilchrist, Raphael Schoenle, Jae Sim and Egon Zakrajsek |
| 1106<br>July 2023 | Global public goods, fiscal policy coordination, and welfare in the world economy | Pierre-Richard Agénor and Luiz A Pereira da Silva |
| 1105<br>June 2023 | The demand for government debt | Egemen Eren, Andreas Schrimpf and Fan Dora Xia |
| 1104<br>June 2023 | The Crypto Multiplier | Rodney Garratt and Maarten R C van Oordt |
| 1103<br>June 2023 | Privacy regulation and fintech lending | Sebastian Doerr, Leonardo Gambacorta, Luigi Guiso and Marina Sanchez del Villar |

All volumes are available on our website www.bis.org.