UCL Centre for Blockchain Technologies

Discussion Paper Series

Q3 2022

Foreword

I am glad to present this issue of the CBT Discussion Paper Series. We have a content-rich edition with four contributions spanning across some of the main 'hot' topics in the blockchain and decentralized finance domain: the governance in Defi; central bank digital cash; stablecoins and a mechanism for blockchain consensus based on identity. Advances in any of these topics will have dramatic repercussions on the entire digital economy domain.

I find a sign of good health that researchers are attacking these challenges from all perspectives, and I am pleased to read convergences and divergences in the proposed approaches. We live in a complex world and technology is not simplifying it. Rather, it is empowering the system's dynamics enhancing such complexity to higher levels. For a long time, we have witnessed academics and technologists producing self-emphatic essays on the magnificent and progressive effects of new technologies on society. There was indeed a need for this, to promote change. Now we start witnessing the production of more mature reflections on fundamental unsolved problems that technology has exasperated. Indeed, we still need to master the delicate balance between authoritarian and democratic governance in all systems and – even more – in Defi. Stablecoins are great tools, but what about the inescapable mechanism underneath supply, demand, and price discovery? And eventually, who protects our personal data from human greed and madness?

I am sure readers will find this issue insightful and thought-provoking.

Tomaso Aste
UCL CBT Scientific Director & Chairman of the Editorial

July 2022

Acknowledgement

The editorial board wishes to thank Matilde Faro for her high-quality work and strong perseverance, without which this discussion paper series would have never been possible.

Discussion Paper Series Contents

CBT

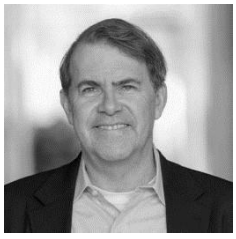
Editorial Board

Tomaso Aste

Chairman of the Editorial Board
Professor, Complexity Science, UCL

Quinn DuPont

Assistant Professor, University College Dublin



Daniel Heller

Honorary Professor, UCL

Seongbae Lim

Professor, St. Mary's University

Ralf Wandmacher

Professor, Accadis University

Andy Yee

Public Policy Director at Visa

Discussion Paper 1

Governance in
Decentralized Autonomous Organizations

Alexander Braun, University of St. Gallen

Niklas Haeusle, University of St. Gallen

Stephan Karpischek, Etherisc

Abstract

We develop a model of decentralized autonomous organizations and examine when these new institutional arrangements are both incentive compatible and collusion proof. Incentive compatibility can be achieved through a staking mechanism. Participants arerequired to post stakes, a digital form of collateral that carries voting rights. Additional effort must be spurred by higher stakes, which increase the risk of collusion. To be inherently collusion proof, DAOs either require a high revenue or a sufficient degree of decentralization in combination with a large network size. If neither one ofthese conditions is fulfilled, collusion can be precluded by governing the DAO with stochastic voting and having voting power concentrated among a few key token holders.

Keywords

DAO, Staking, Incentive Compatibility, Collusion Proofness, Game Theory.

CBT

Governance in Decentralized Autonomous Organizations

Alexander Braun, Niklas Haeusle, Stephan Karpischek[*]

June 1, 2022

Abstract

We develop a model of decentralized autonomous organizations and examine when these new institutional arrangements are both incentive compatible and collusion proof. Incentive compatibility can be achieved through a staking mechanism. Participants are required to post stakes, a digital form of collateral that carries voting rights. Additional effort must be spurred by higher stakes, which increase the risk of collusion. To be inherently collusion proof, DAOs either require a high revenue or a sufficient degree of decentralization in combination with a large network size. If neither one of these conditions is fulfilled, collusion can be precluded by governing the DAO with stochastic voting and having voting power concentrated among a few key token holders.

[*]Alexander Braun is Associate Professor of Insurance and Capital Markets as well as Director of the Institute of Insurance Economics at the University of St. Gallen, Tannenstrasse 19, CH-9000 St. Gallen, Switzerland. Niklas Haeusle is a Ph.D. student in Finance at the University of St. Gallen. Stephan Karpischek is with the Decentralized Insurance Foundation. The authors can be contacted via E-mail: alexander.braun@unisg.ch, niklas.haeusle@unisg.ch, and kpi@kpi.at. We thank Stefan Buehler, Samuel Haefner, Winfried Koeniger, Larry Samuelson, Rakesh Vohra as well as the participants of the European Economic Association (EEA) Annual Meeting (2020), the American Risk and Insurance Association (ARIA) Annual Meeting (2021), the European Group of Risk and Insurance Economists (EGRIE 2021) annual meeting and the household finance research seminar at the University of St. Gallen (2021) for their helpful comments and suggestions.

1 Introduction

In his original Ethereum whitepaper, Buterin (2013) discusses the concept of an idealized decentralized autonomous organization (DAO) that operates under a fully transparent and tamper-proof set of business rules recorded in specific pieces of computer code. The latter are called smart contracts and maintained by a blockchain network.[1] Workers remain anonymous and can join and leave a DAO without permission. Furthermore, there is no hierarchy or human management, meaning that decisions are based on consensus of the network participants, who identify themselves with cryptographic tokens native to the project. All tasks involved in the production of goods or services are completed on an entirely transactional basis. In line with these considerations, we define a DAO as an institutional arrangement with a decentralized value creation process and a decentralized decision making process.

DAOs have seen unprecedented growth in recent years in the blockchain space. In 2022, blockchain sources listed over 1'000 DAOs, 65 of which exhibited an asset base in excess of USD 100 million and more than 1'000 active members.[2] Together, these projects cover a wide range of applications: developer collectives such as BadgerDAO build infrastructure for the Web 3.0, decentralized finance (DeFi) organizations such as MakerDAO offer lending and borrowing services, and decentralized insurance platforms such as Etherisc offer parametric insurance, e.g., to protect smallholder farmers against natural catastrophes in emerging markets.

Decentralization creates the need for specific governance mechanisms that align different network participants along the common goals of the DAO. In traditional firms, governance typically relies on managerial oversight. This is not possible in a decentralized setting, since executives and reporting lines do not exist. Workers may participate on a one-off basis so that their liability is strictly limited. Hence, they cannot be held accountable for shirking after the transaction has been completed. Moreover, attempts to directly sue individual workers are likely associated with prohibitively high costs, even if the net-work is relatively small. The reason is that participants typically remain anonymous andmay be distributed across various jurisdictions around the globe.[3] Consequently, DAOs are at odds with a traditional enforcement of claims.

Against this background, we analyze how the governance of DAOs needs to be designed so that incentive and collusion problems can be resolved. A sufficient share of the revenue induces individuals to participate. A staking mechanism can incentivize indi-

[1] Blockchains are a form of distributed ledger technology that allows for record keeping by a network of nodes, rather than a trusted intermediary (see, e.g., World Economic Forum, 2018).

[2] For more information see https://deepdao.io/organizations.

[3] The incompleteness of contracts governing international transactions is known to be a limiting factor for traditional firms (see, e.g., Antràs, 2005a,b). A decentralized setting, however, exponentiates the problem.

viduals to exert effort. Stakes are a digital form of collateral that DAO workers deposit into a trusted smart contract, where it remains locked until the task is completed and the result becomes visible.[4] Unfortunately, staking mechanisms give rise to another governance problem in a DAO: collusion. Higher stakes are needed to instigate larger efforts. However, the higher the stake posted by a worker, the stronger his temptation to collude with others by substituting effort costs with bribes to peers. Malicious workers can thus manipulate the decision making process of a DAO in order to be acquitted and avoid the confiscation of their staked tokens. Against this background, it is all but trivial to design a DAO that is both incentive-compatible *and* collusion proof.

We develop a microeconomic model of a DAO and describe the constrained optimization problem faced by its planner, including target function, incentive-compatibility condition, participation condition, resource condition and non-collusion condition. We then use our model to scrutinize three key design choices for the DAO: i) the revenue sharing scheme, ii) the staking mechanism and iii) the voting system required for decentralized consensus. The staking mechanism itself consists of the size of demanded stakes and the usage of confiscated stakes (burning or redistribution). In a first step, we determinethe optimal revenue sharing scheme. Next, we derive the optimal incentive-compatible stake and assess whether redistribution or burning of confiscated stakes is the best way to reduce the risk of collusion. Finally, a substantial part of the paper focuses on the iden- tification of a collusion-proof consensus mechanism. To this end, we analyze the impact of concentrated voting power, the choice of different voting systems, and the existence of additional governance token holders.

Our work aims to establish a strand of the literature, focusing on economic issues in DAOs. Within the wider blockchain literature, we provide insights on how this technol-ogy helps to overcome architectural and governance challenges in digital platforms and infrastructures, which Constantinides et al. (2018) highlighted as a major area for future research. We also contribute to the discussion on decentralized consensus (Saleh, 2021). Apart from that, we add to the existing research on virtual teams (Jarvenpaa et al., 2004; Wakefield et al., 2008; Peng et al., 2019). Virtual teams are IT-enabled relationships with a decentralized production process, but a centralized decision making process. DAOs are blockchain-enabled relationships where both production and decision making are decen- tralized. Finally, in a wider economics context, our contribution is also related to the liter- atures on team production in general (Guillen et al., 2015), online labor markets (Huanget al., 2020), governance and platform design (Tiwana and Konsynski, 2010; Tiwana et al., 2010), and voting systems (Dal Bó, 2007; Dekel et al., 2008).

[4]Generally, collateral schemes are also possible in traditional organizations. Firms may demand performance bonds to be placed in a trust fund to deter workers from shirking. In Section 5 we will point out differences to staking in the blockchain context.

This paper proceeds as follows. In the next section, we briefly recap the characteristics of DAOs and introduce some recent examples. In the third section, we then provide a detailed explanation of the economic problem. In the fourth section, we lay out a model that captures the operating principles of an idealized DAO. We employ this model to identify the characteristics and governance choices that determine whether a DAO will be incentive compatible and collusion proof. In the penultimate section, we discuss DAOs in a broader context and answer some of the questions outlined in Constantinides et al. (2018) and Hendershott et al. (2021), regarding blockchain technology and governance. Finally, in the sixth section, we draw our conclusion.

2 Decentralized Autonomous Organizations

2.1 General Structure of a DAO

Figure 1 shows the generic structure of a DAO. Multiple workers coordinate themselves through a blockchain network to generate a product or service, which customers purchase with fiat currency (or stable tokens).[5] The DAO is built around a native token with several functions. First, it secures early stage funding through an initial coin offering (ICO). Second, it may be used in a staking mechanism with the goal of incentivizing workers to exert effort. Staked tokens act as a collateral to prevent malfeasance. Workers who want to participate in the DAO need to purchase these tokens and send them to a smart contract where they will remain locked until the task has been fulfilled. Third, the native token can carry voting rights and thus play a crucial role in reaching decentralized consensus.

Apart from workers and customers, a DAO usually also comprise passive token holders. They either purchase the token during the ICO or later on in the secondary market. By buying and selling their tokens on exchanges, they generate liquidity. As we will see later on in the course of our formal analysis, such token holders may also be an important line of defense against collusion.

[5] It should be noted that workers can be individuals as well as (small) firms.

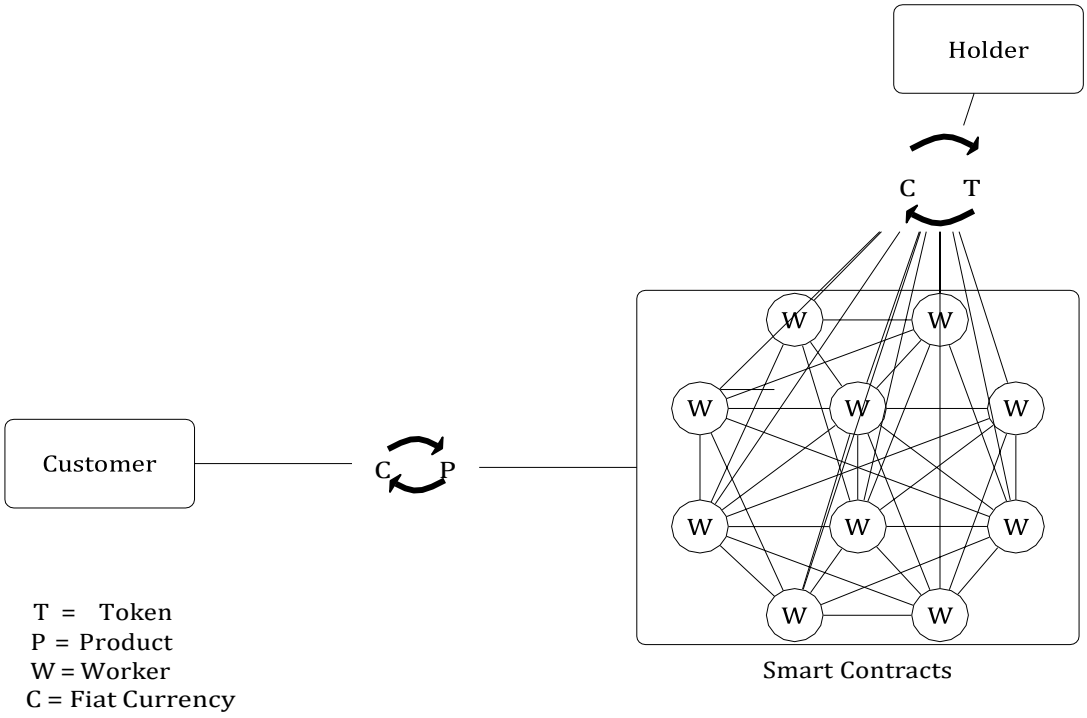


Figure 1: Generic Structure of a DAO

This figure illustrates the structure of an archetypical DAO. Workers denoted W contract on a transac- tional basis via a peer-to-peer network to establish a decentralized value creation process. Customers purchase the goods or services produced by the DAO with fiat currency (or stable coins). Apart from the workers, the DAO might also have passive members, calld token holders. They do not accomplish tasks but provide early-stage funding to the project by investing in the network-specific tokens issued during an ICO. Later on, they buy and sell tokens in the secondary market.

2.2 The Case of MakerDAO

An important application of the DAO concept is the DeFi space. In the following, we briefly introduce the example of MakerDAO, a platform created on the Ethereum blockchain to allow participants to lend and borrow cryptocurrencies. MakerDAO is the first DeFi application to see significant adoption, with a market capitalization of roughly USD 200 million at the time of writing.[6] The Maker Protocol is maintained by people around the world, who hold its governance token, the MKR. Decisions are made through a system of voting and governance polling. One MKR token locked into a voting contract equals one vote (MakerDAO, 2019).

In addition to its smart contract infrastructure, the Maker Protocol involves groups of actors (workers) to facilitate operations. Keepers maintain the target price and provide liquidity to the system. Price oracles provide real-time information about the market

[6]https://coinmarketcap.com/currencies/dao-maker/

price of the collateral assets so that liquidations can be triggered in time. Global settlers (emergency oracles) are a line of defense for attacks on oracles. Finally, core units (MakerDAO teams) provide specific services to the network in different roles (risk team members, governance facilitators etc.). A shirking problem exists for several worker groups. The core units are best suited to illustrate it. Consider, e.g., the so-called risk team members. They support the MakerDAO governance with financial risk research (MakerDAO, 2019). To generate accurate and reliable research, they need to exert effort. This effort can be avoided by creating low quality research.

3 The Economic Problem

Consider a decentralized production process with n necessary participants. Denote the set of all participants N . Revenue R is probabilistic and binary, i.e., either the full revenue occurs or there is no revenue at all. The probability that revenue is generated depends on the joint effort decision of the workers that form the DAO. The more workers exert effort, the higher the likelihood that revenue is generated. Every worker i decides for himself, if he exerts effort at costs e_i or shirks at costs 0. Hence, exerting effort is a binary discrete variable ($\{0, e_i\}$). If every single worker exerts his effort, revenue is generated with a probability of 1. If one individual shirks, this probability drops to p_0. The probability for failure is proportional to the number of individuals that shirk. More specifically, for $\tilde{n}$ malicious workers, the success probability equals $1 - \tilde{n}(1 - p_0)$. Any revenue generated by the DAO is distributed among all participants. This means that there is an interdependence between workers: a shirking worker imposes a negative externality on his peers. We are interested in equilibria where all workers exert effort.

A staking mechanism with worker-specific stakes S_i, which act as a collateral, can be used to solve this freerider problem. Under this mechanism, workers need to acquire a predetermined amount of network-specific tokens and deposit them into a trusted smart contract. The funds are then locked until the task is completed and the result becomes visible. We assume that posting a stake is associated with collateral costs rS_i.[7] The main problem of such a staking mechanism is to determine when stakes should be confiscated. For ease of exposition, we consider the case in which the effort is observable for humans, but shirking cannot be verified electronically.[8] Thus, it is not possible to automate pun-

[7] These costs can be any form of opportunity costs or capital costs (Bester, 1985; Wang et al., 2022).

[8] This assumption can be made without loss of generality. Extending the model by uncertainty about the identity of the malicious worker would complicate the exposition, but not change our key results. We reflect on the effect of a noisy signal for unobservable effort in Section 5.3.

ishment through a smart contract.[9] Instead, a decision by the DAO is required. Without a central decision making authority, however, every decision in a DAO must be made collectively by its participants. The decisions that determine whether a specific worker will be found guilty of shirking are economically interesting, because they imply a major conflict of interest. A shirker may use his own voting power in the DAO to prevent the confiscation of his stake. Furthermore, he might bribe others to influence the outcome of the decision process.

Figure 2 is an overview of the underlying game. In the first step, each worker de- cides whether or not to shirk. If all workers exert effort, the game ends directly at this point: revenue is collected and distributed among the participants. If a worker shirks, the game may still end through the generation of revenue, albeit with a lower probability(p_0). With probability ($1 - p_0$), the missing effort causes the DAO to fail and lose its rev- enue. In this situation, the punishment process built into the staking mechanism begins.First the shirker must be identified and accused. Given effort is assumed to be observable information, this step is straightforward. Next, the accused worker can offer bribes to other network participants to turn the voting process in his favor. Finally, voting takes place. Each participant votes for conviction or acquittal of the defendant. In the absence of bribes, the outcome will be conviction. If the shirker bribes others, however, he might get acquitted. After the voting decision, any promised bribes are paid and, in case of conviction, the stake of the culprit is confiscated. We assume that the acquittal of a malicious worker causes a loss of trust in the DAO and, in turn, a drop in the value of its network-specific token. Our focus lies on subgame perfect nash equilibria, in which everybody exerts effort. On the off-equilibrium path (where somebody does not exert effort), we are interested in the nash equilibrium without bribing by the malicious worker.

[9]For example, the DAO might need a data provider for its operations, with expected costs of 100 monetary units. A smart contract can be written which compensates the data provider automatically after delivery. However, the smart contract cannot recognize if the data provider just sent made-up numbers to save costs. Humans on the other hand can usually detect such malicious behavior fairly quickly.

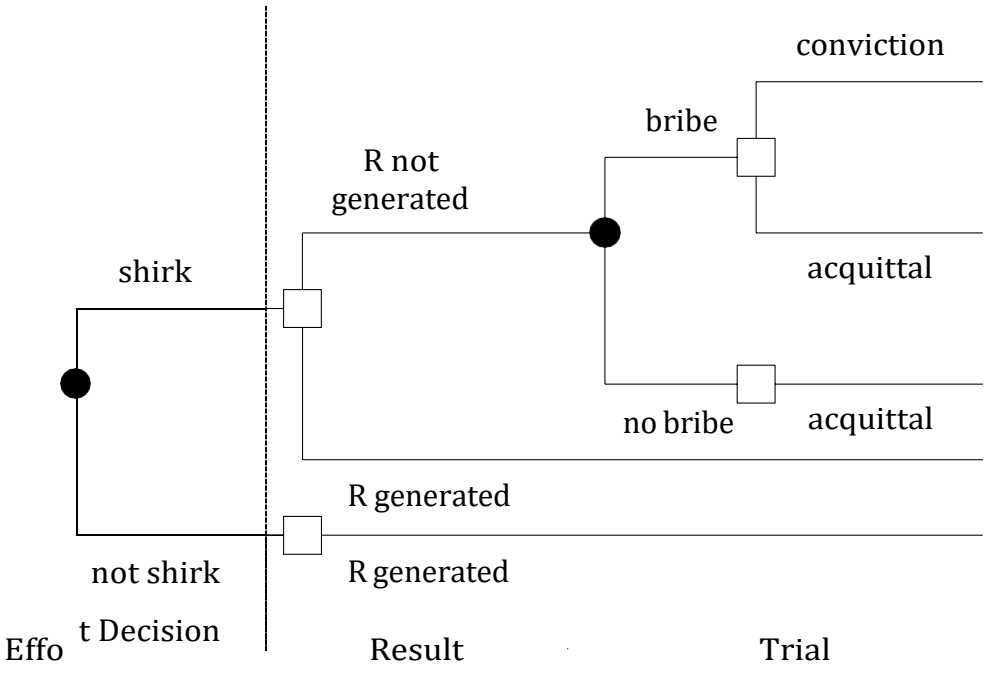


Figure 2: Overview of the Nash Game in a DAO

This figure illustrates the economic problem addressed in this paper. DAO workers play a multi-stage nash game. In the first step, they decide whether or not to expend effort. In the second step, the random revenue generation of the DAO takes place. Here the probability for a successful outcome decreases in the number of shirkers. The third step only occurs in case the DAO failed to generate revenue. The shirker is then identified and a voting to confiscate his stake is arranged. Before this voting is carried out, however, the shirker may decide to bribe other network participants for their collusion so that he will be acquitted.

To sum up, our main problem is the shirking problem, the solution is staking. However, compared to a traditional organization we cannot rely on a central decision making authority. Hence, the danger of collusion between network participants is prevalent. As it will turn out, this problem can be mitigated through certain key characteristics of a DAO: the voting scheme (masked voting, majority voting and stochastic voting), the assignment of voting rights (to staked tokens only or to all tokenholders), and the joint impact of decentralization (of voting rights) and network size (number of participants).

4 Modeling the DAO

Formally, let K denote the probability space of the voting outcome and $k \in K$ a specific realization therein.[10] C, A, and Q represent degenerate versions of the probability space. C subsumes all outcomes which lead to conviction, A subsumes all outcomes which lead to acquittal and Q subsumes all outcomes where an individual voting decision is pivotal,

[10]This space includes any possible combination of votes, e.g. individuals *i* and *j* voted for conviction and individuals *k*,*l*, and *m* voted for acquittal.

i.e., a switch of one voter changes the outcome from C to A or vice versa. Furthermore, $B_{j,k}$ stands for the bribe paid to person j in state k and m denotes the number of individuals to which a positive bribe was offered. While being costly, collusion with other blockchain network members is not impossible.[11] However, due to the anonymity and international dispersion of participants in a DAO, the identification of and outreach to peers becomes time consuming and resource intensive relatively quickly. Thus, the briber incurs search costs $c(m)$, that grow in the number of contacted network participants m. Finally, D is the relative loss in token value caused by the acquittal of a guilty worker, implying that the stake of every participant is then worth $(1 - D)S < S$ in real value terms.

Recall from the previous section that S_i denotes the stake, rS_i the staking (collateral) costs, and e_i the effort of worker i. Moreover, p_0 and $(1 - p_0)$ are the probability of success and the probability of failure for the DAO, if exactly one worker shirks. In the following, bold letters will be used to distinguish vectors from scalars. In particular, z is the vector of individual revenue shares z_i and B is the vector of individual bribe amounts $B_{j,k}$ paid to network participants.

We begin our formal analysis with the overall problem:

Problem 1

$$\underset{\mathbf{z},S}{\operatorname{argmax}} \quad R - \left(\sum_{i=1}^{n} e_i + r \sum_{i=1}^{n} S_i \right)$$

$$s.t.$$

$$z_i R - e_i \geq z_i R p_0 - S_i(1 - p_0) \qquad \forall i \qquad (IC)$$

$$z_i R - e_i - rS_i \geq 0 \qquad \forall i \qquad (PC)$$

$$\sum_{i=1}^{n} z_i R \leq R \qquad (RC)$$

$$S_i \leq \sum_{k \in K} P(k) \left(\sum_{j=1}^{m} B_{j,k} \right) + P(C|\mathbf{B}) S_i + c(m) \qquad \forall i. \qquad (BC)$$

When starting a DAO, the planner needs to select a revenue allocation scheme z and a staking scheme S. The overall staking mechanism then consists of i) the staking scheme, i.e., the size of the stake S_i to be demanded from each individual, and ii) the usage of confiscated stakes, i.e., the decision of whether those stakes should be redistributed among

[11]A well-known example for collusion in a decentralized setting are the pump and dump schemes in the crypto asset space (Xu and Livshits, 2019).

the honest workers or burnt.[12] Furthermore, the planner can select a voting scheme for the a DAO: masked voting, majority voting and stochastic voting. He makes these design choices with the goal of maximizing the surplus of the DAO, consisting of the accumulated revenue minus the accumulated effort and staking costs.

For his optimization, the planner has to take the following constraints into consideration:

- Incentive compatibility constraint (IC): it needs to be preferable for a individual to exert effort when all other workers in the DAO do so.
- Participation constraint (PC) constraint: workers will only be prepared to join the DAO, if they do not expect to suffer a loss.
- Resource constraint (RC): the sum of all distributions cannot exceed the revenue.
- Bribing constraint (BC): a solution without bribing can only exist if it is more expensive to bribe than to forfeit the stake.

The last constraint warrants some additional explanation. While (IC), (PC) and (RC) are also found in economic analyses of traditional organizations, (BC) is specific to the DAO. The first term on the right hand side of (BC) is the unconditional expected value of the bribes across the whole probability space K.[13] The second term is the expected stake loss in case of conviction, i.e., if the bribing fails. The probability of conviction $P(\mathrm{C}|\mathrm{B})$ is contingent on the allocation of the bribes B to the voters. Finally, the third term represents the search costs $c(m)$ that arise from the decentralized setting and make bribing more expensive, the larger the number of workers that have to be identified and contacted.

If Problem 1 has a solution, the DAO is feasible. Otherwise, rational economic agents will not be prepared to form it in the first place. Figure 3 illustrates how the feasibility of the DAO depends on the two critical parameters R and m. In Area A the existence of the DAO is not socially optimal:

$$R < \sum_{i=1}^{n} e_i. \tag{1}$$

The obvious reason is that the revenue is insufficient to compensate the aggregate effort expended by the workers. It is relatively straightforward to see that a situation in which $\tilde{n}$ workers shirk and the rest $n - \tilde{n}$ exerts effort is also not optimal:

$$R[1 - \tilde{n}(1 - p_0)] < \sum_{i \in \mathrm{N}} e_i, \tag{2}$$

[12] In a blockchain context, burning means irrecoverable destruction of tokens.

[13]Note that the bribes could also be conditioned on states where the shirker gets acquitted ($k \in A$).

where N denotes the set of honest (non-shirking) workers.

Area B represents the case, in which the sum of the aggregate effort and the aggregate collateral costs associated with the optimal incentive-compatible stakes S^*_i exceed the revenue R:[14]

$$R < \sum_{i=1}^{n} e_i + r \sum_{i=1}^{n} S^*_i. \tag{3}$$

The associated violation of the participation constraint implies that the DAO is infeasible.

Furthermore, in Area D, where R is high enough (c.p.), the DAO is generally feasible:

Lemma 1 *If the revenue is high enough, the DAO is incentive compatible and collusion proof.*

Proof. A large R ceteribus paribus implies

$$z_i R - e_i > z_i R p_0 \quad \forall i, \tag{4}$$

such that the workers are incentivized by the revenue sharing scheme itself. In this case, our problem becomes a trivial one: staking is unnecessary.■

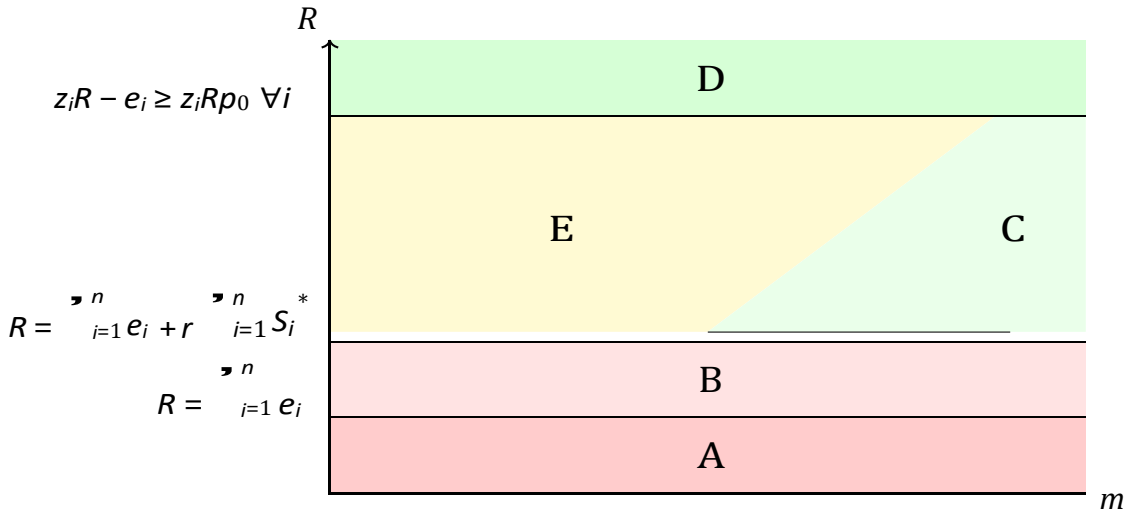


Figure 3: Economic Feasibility of a DAO

This figure shows the feasibility of a DAO, depending on its revenue R and the number of participants m that need to be contacted for bribing. Areas A and B are never feasible. Area D is always feasible. Areas C and E are only feasible, if the DAO can be made collusion-proof.

The freerider problem arises only in Areas C and E. In these cases, the DAO is generally feasible, but the revenue R is not high enough to incentivize the workers automatically. It

[14]We determine the optimal incentive-compatible stake in Proposition 1.

may thus be beneficial to shirk. To prevent shirking and the associated negative external- ity on all other participants, an efficient incentivization mechanism is required. The decentralized setting without a traditional decision making authority such as management, however, equips malicious workers with the additional possibility of bribing (collusion) to avoid a loss of their stake. Therefore, the planner needs to find an optimal[15] stake S^* that prevents shirking *and* he needs to consider whether or not that stake is also collusion proof. In the following subsections, it will turn out this implies a tradeoff: to make the DAO incentive compatible, the stake needs to be rather high. To make the DAO collusion proof, however, the stake needs to be rather low.

4.1 Optimal Revenue Sharing

As we will see below (refer to proposition 1), the optimal incentive-compatible stakes S_i^* cannot be derived independently from the revenue allocation mechanism of the DAO. We will therefore first identify the optimal revenue allocation scheme z and then use it to determine the optimal staking scheme S.

There are a variety of different revenue sharing schemes that maximize the surplus of the DAO. We focus on the solution, which additionally makes the DAO feasible for as many parametrizations as possible. Intuitively, the range of feasible revenue sharing schemes decreases in the revenue R. This is due to the fact that, for lower R, the partici- pation constraint (PC) becomes binding more quickly.

Lemma 2 *It is optimal to share the revenue proportionally to the effort.*

Proof. To see that it is optimal to distribute the revenue among workers proportional to their effort, first rearrange (IC) as follows.

$$\frac{e_i}{1 - \sum_i} - z_i R \leq S \tag{5}$$

The optimal (welfare-maximizing) stake is exactly equal to the left hand side of (5).[16] The collateral costs rS_i associated with this stake are:

$$\frac{\left(\frac{e_i}{1 - \sum_i} - z_i R\right) \cdot r}{1 - p_0} \tag{6}$$

[15]Optimal here means welfare (or surplus) maximizing. Since the stake enters the objective function negatively (see Problem 1), we would like it to be as small as possible, while still achieving incentive compatibility.

[16]Other stakes can also achieve incentive compatibility, but would be unnecessarily high.

Substituting these collateral costs into (PC) and rearranging[17] yields:

$$z_i R \geq e_i \gamma, \tag{7}$$

where

$$\gamma = 1 + \left(\frac{r}{1 - p_0} \right) \frac{1}{(1 + r)}. \tag{8}$$

Inequality (7) highlights that the individual worker i is only prepared to participate in the DAO, if the revenue share that he will receive is proportional to his effort: $z_i R \propto e_i$.

■

Based on Lemma 2 it is relatively straightforward to recognize that the proceeds can only be shared proportionally to the effort, if the DAO manages to generate a critical minimal level of revenue. Since we cannot distribute more than the full revenue, the weights z_i must sum up to 1:

$$1 = \sum_{i=1}^{n} z_i \quad \rightarrow \quad 1 = \sum_{i=1}^{n} \frac{e_i \gamma}{R}. \tag{9}$$

Given a specific e_i and γ, now consider a revenue level $\bar{R}$, for which (9) is exactly fulfilled:

$$\bar{R} = \sum_{i=1}^{n} z_i R = \sum_{i=1}^{n} e_i \gamma \tag{10}$$

Larger revenues $R^+ > \bar{R}$ provide the flexibility of deviating from a proportional scheme. Smaller revenues $R^- < \bar{R}$, on the other hand, violate (PC) and make the DAO infeasible.

Recall from Section 3 that the effort in our setting is observable for humans, but not for machines. Nevertheless, a smart contract can automatically distribute revenue among workers proportionally to their effort by scheduling ex ante targets. The ex post realization of the target effort, however, can only be observed by humans.

4.2 Optimal Incentive-Compatible Stake

Given the optimal revenue allocation scheme, we can determine the optimal incentive-compatible stake.

[17] We assume that the collateral cost rate r is strictly positive.

Proposition 1 *The optimal incentive-compatible stake for worker i is*

$$S_i^* = \frac{e_i}{1 - p_0} - \frac{e_i}{\sum_{i=1}^{n} e_i} R. \tag{11}$$

Proof. See Appendix. ■

The first component of the optimal incentive-compatible stake represents the individual's effort costs, scaled with the probability that no revenue is generated. Thus, the smaller the probability of failure, the higher the optimal incentive-compatible stake. Intuitively, for low probabilities of failure, the stake needs to be multiple times the effort to make the expected costs of losing it sufficiently large. The second component reduces the optimal stake that is required to make the DAO incentive compatible. It equals the revenue share that the worker receives from the DAO in case of success. Economically, this is similar to a bonus: the stake can be lowered accordingly without violating incentive compatibility. Notice that stake and revenue are both proportional to the effort. Also notice that if revenue is large enough, this equation becomes negative. However, in such a case, we are automatically in area D of Figure 3.

4.3 Redistribution vs. Burning of the Stake

As underlined in Section 3, our main interest lies in the question of how collusion can be avoided altogether. Nevertheless, we also consider the off-equilibrium path and answer the question of what should happen to the stake of a shirker after it has been confiscated.

In this case, the DAO has two options. It could either redistribute the stake to the honest workers as a compensation for the lost revenue, or it could simply burn the corresponding tokens, i.e., irrevocably remove them from the system.[18] With regard to incentive compatibility, it only matters that the stake will be lost to the malicious worker.

With regard to collusion resilience, however, the question of redistribution versus burning of stakes is a relevant governance decision for a DAO. In this regard, we present the following result.

Proposition 2 *A staking mechanism that redistributes stakes instead of burning them increases the collusion resilience of a DAO.*

Proof. See Appendix. ■

[18] In line with Lemma 2, in case of redistribution, the stake should be allocated proportionally to the effort of the honest workers.

Intuitively, stake redistribution eases the participation constraint by providing the honest participants implicit insurance against lost revenue due to the shirking of others.

4.4 Decentralization and Size

We now take a closer look at area C of Figure 3, in which the revenue R is not sufficient to make the DAO incentive compatible and collusion proof, but the bribing costs $c(m)$ for the shirker are high enough to ensure the feasibility of the optimal incentive-compatible stake. This is because the number of peers m that the shirker needs to identify and contact to offer bribes is large. First, we need to develop a more detailed understanding of the drivers of m. We draw on the Herfindahl index[19] H_V as a measure for the concentration of the voting power:

$$H_V = \sum_{i=1}^{n} \left(\frac{S_i}{\sum_{i=1}^{n} S_i} \right)^2 . \tag{12}$$

H_V is a function of i) the network size n and ii) the distribution of the stakes among the participants. It ranges from $\frac{1}{n}$ (perfect decentralization) to one (perfect concentration). When all participants in the DAO have equal stakes, $H_V = n \left(\frac{S}{nS} \right)^2 = \frac{1}{n}$.

Given H_V, we can model how m is related to the degree of decentralization of the DAO network. Intuitively, the less concentrated the voting power (i.e., the closer H_V is to $\frac{1}{n}$), the higher the number of network participants that the malicious worker needs to contact with his bribing offer.

As a corollary, the DAO will be collusion proof, if the network is both sufficiently decentralized *and* large. To see this, reconsider the bribing constraint (BC):

$$S_i \leq \sum_{k \in \mathrm{K}} P(k) \left(\sum_{j=1}^{m} B_{j,k} \right) + P(\mathrm{C}|\mathrm{B}) S_i + c(m) \qquad \forall i. \tag{13}$$

Clearly, (BC) places an upper bound on the feasible stake. For stakes that exceed this limit, the shirker has an incentive to bribe other members of the DAO in order to avoid conviction. If the optimal incentive-compatible stake exceeds the upper bound implied by (BC), it is not possible to make the DAO both incentive compatible *and* collusion proof. Obviously, any increase of the right hand side of (BC) reduces this problem, because it increases the set of feasible stakes.

Deriving an explicit expression for m, given a particular network size (n) as well as a specific degree of concentration of the voting power among the participants (H_V), and

[19]The Herfindahl index is widely used to measure concentration, for example by the FTC to measure market concentration.

inserting it into (BC) allows us to derive the following result:

Proposition 3 *A sufficient condition for a collusion-proof DAO is*

$$S^*_i \leq c \left(\frac{n}{2} - \frac{1}{2} \right) \sqrt{n^2 - \frac{1}{H_V}} \quad \forall i. \tag{14}$$

Proof. See Appendix. ■

Intuitively, for an equal distribution of the stakes $H_V = \frac{1}{n}$, the Herfindahl index is correlated with the number of participants in the DAO and a large n will lead to a low concentration of the voting power. This implies a large m and, in turn, prohibitively high costs of collusion. However, if the stakes are very unequally distributed, voting power can be highly concentrated, even if n is large ($H_V \gg 1$). The shirker then only has to contact and bribe a small number of influential DAO members that can provide him the decisive edge in the voting process.

Proposition 3 will be fulfilled if the DAO exhibits both a sufficient degree of decentralization *and* a large network size. However, this is not a design choice. Size is exogenous and intertwined with a chicken and egg problem: The DAO gets collusion proof, if its sufficiently large and decentralized, but nobody wants to participate in a DAO where collusion inhibits the decentralized value creation process. Furthermore, empirical facts point to the possibility that size and decentralization may be difficult to achieve at the same time: in Proof-of-Work as well as in Proof of Stake blockchain networks, size is often correlated with a tendency to recentralize (He et al., 2020).[20] Against this background, the next sections focus on area E of Figure 3, where collusion must be avoided through governance instruments.

4.5 Voting Systems for the Consensus Process

A key characteristic of a DAOs is that its participants must achieve decentralized consensus, i.e., without the intervention of a central authority as in traditional principal-agent relationships. While consensus is require for every decision of the DAO, we specifically scrutinize the problems, which are associated with the verification of shirking. To this end, we compare three potential voting mechanisms: i) masked voting, ii) majority voting, and iii) stochastic voting.

[20] A well-known example is the highly concentrated hashing power among Bitcoin mining pools.

4.5.1 Masked Voting

Consider the case where the voting rights of every worker are proportional to his stake. In case of failure and lost revenue, a consensus must be reached regarding the treatment of the shirker. If at least 50 percent of the votes are cast for conviction, the DAO confiscates the stake of the malicious worker. After the voting, the decision and the number of votes that carried it become public knowledge. However, under masked voting, the individual voting behavior of the network participants is not publicly disclosed. Therefore, the shirker does not find out who voted for acquittal and who for conviction. This means that bribes cannot only be conditioned on the total outcome, but not on individual votes. Under such a voting scheme the following holds true:

Proposition 4 *Masked voting increases the collusion resilience of a DAO.*

Proof. See Appendix. ■

It should be highlighted that masked voting makes a DAO more *collusion resilient*, but not necessarily *collusion proof*. If the token value response to a false acquittal is very small or the voting power in the network is extremely concentrated, collusion cannot be prevented with certainty. For a large range of reasonable parameter values, however, masked voting will yield a collusion-proof outcome. Interestingly, masked voting is hardly seen in existing DAOs.[21] This is likely due to the fact that a general characteristic of blockchain technology is transparency. DAOs with masked voting are at odds with this notion.

4.5.2 Majority Voting

Some DAOs rely on classic majority voting. From a technical perspective, network participants can cast their vote by sending a small amount of tokens to a smart contract. The individual votes are then weighted with the voting power of the senders. If a 50 per- cent majority is reached, the faulty worker will be convicted. After the voting, all details become fully transparent. Our analysis of majority voting delivers the following result:

Proposition 5 *Majority voting does not prevent collusion in a DAO.*

Proof. See Appendix. ■

The intuition behind the proof is as follows. First we find the most efficient bribing strategy. To do this we solve the following subproblem:

[21] A notable exception is Aragon Court.

Problem 2

$$\underset{B}{\operatorname{argmin}} \sum_{j=1}^{m} \sum_{k \in \mathrm{K}} B_{j,k} P(k) + SP(\mathrm{C}|B),$$

i.e., the briber picks the vector of bribes B that minimizes his bribing costs. We find that a unique nash equilibrium exists, in which accepting bribes is a strictly dominant strategy. This leads to acquittal of the malicious worker at virtually no costs.

4.5.3 Stochastic Voting

Under a weighted stochastic voting scheme, there is no need to achieve a 50 percent majority for the conviction of shirkers. Instead, the decision is made randomly. First, everybody places a vote. Then, a single decisive vote is drawn from the overall pool of votes with a probability that is proportional to the stake size.[22]

Proposition 6 *Stochastic voting does not prevent collusion in a DAO.*

Proof. See Appendix. ■

In the next section, we will show that, in the presence of passive token holders, stochastic voting becomes collusion proof. A majority voting scheme, in contrast, will still be susceptible to bribing.

4.6 Governance Token Holders

Until now, we assumed that the native token of the DAO resembles preferred stock. Put differently, tokens that are not staked only serve the purpose of providing liquidity – they do not carry voting power. In this section, we relax this assumption and explore another design choice the DAO can make: extending voting power to passive governance token holders, i.e., network participants that do not actively complete tasks for the DAO as workers. Unfortunately, the sheer presence of token holders alone is not enough to correct the incentives:

Corollary 1 *If a DAO is governed through majority voting, the presence of governance token holders does not prevent collusion.*

[22] This is similar to the determination of the next block writer in a Proof-of-Stake (PoS) consensus mechanism.

Proof. See Appendix. ■

The economic intuition behind this result is relatively straightforward. Under majority voting, token holders are virtually no different from workers. Their presence ceteribus paribus only increases the number of network participants. Accordingly, the dominant bribing strategy that we introduced above remains intact.

Things turn out differently, however, if token holders are present in a DAO with stochastic voting and voting power is concentrated:

Proposition 7 *A DAO with stochastic voting and a concentration of voting power among governance token holders is collusion-proof.*

Proof. See Appendix. ■

The economic intuition behind this result is that, in the presence of stochastic voting, token holders with a large concentration of voting power are non-linearly incentivized to refuse bribes. There are three reasons for this:

i The stochastic voting mechanism fixes the probability of being pivotal, making it independent of the voting behavior of others.

ii The probability of being pivotal is much higher for a large token holder compared toa small token holder.

iii Given a large token holder is pivotal, he has much more to lose, since any value decrease of the network tokens has a more substantial impact on his wealth position.

To be fair, it should be noted that stochastic voting will always be associated with a positive probability probability of acquittal, because a malicious worker can avert conviction, if he turns out to be the randomly chosen pivotal voter. However, this outcome refers to the off-equilibrium path and is therefore a minor issue. Stochastic voting in the presence of large token holders will discourage shirking in the first place, because the formerly dominant bribing strategy does no longer work *in expectation*.

Also note that Proposition 7 provides a theoretical explanation for the empirical observation that DAOs (and other blockchain-based projects) often exhibit so-called "whales": network participants that own a large fraction of the overall token supply.[23] However, while this has positive governance effects when the DAO is small, it seems to be difficult to ensure a proper decentralization when the DAO grows larger and the token concentration is no longer needed for governance reasons.

[23]See, e.g., Aragon, MakerDAO, BadgerDAO, Etherisc etc.

The mechanics behind Proposition 7 are illustrated in Figure 4. Moreover, Table 1 summarizes our main results on the governance of DAOs. In the next section, we will round out our analysis with a number of additional considerations beyond our model framework.

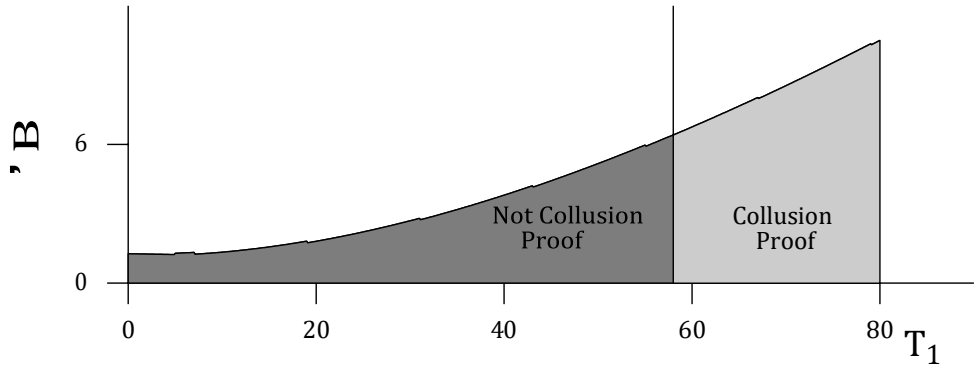


Figure 4: Illustration of the Importance of Concentrated Governance Token Holders

The graph shows the overall bribing costs for the shirker (B'), depending on the number of tokens held by one of two governance token holders (T_1). In this numeric example, the DAO has 15 workers, each with an optimal stake of six, and two additional governance token holders. One of the two possesses 85 tokens and the other one an amount T_1. The bribing costs from the perspective of a shirker increase with T_1. Once a critical concentration of tokens is reached, the DAO becomes collusion proof.

DAO Characteristic	Result	Description	Implication
Revenue	Lemma 1	Revenue is large	DAO collusion proof
Revenue allocation	Lemma 2	Opt. revenue sharing scheme	proportional to effort
Size of stakes	Proposition 1	Opt. incentive-compatible stake	proportional to effort
Confiscated stakes	Proposition 2	Stake redistribution (no burning)	DAO more collusion resilient
Decentralization / size	Proposition 3	Decentralized + large network	DAO collusion proof
Voting system	Proposition 4	Masked voting	DAO collusion proof*
	Proposition 5	Majority voting	DAO not collusion proof
	Proposition 6	Stochastic voting	DAO not collusion proof
Voting rights	Proposition 7	Governance token holders + stochastic voting + concentrated voting power	DAO collusion proof

Table 1: Summary of Main Results on DAO Governance

This table summarizes our main results. The different governance instruments and network characteristics determine whether a DAO is both incentive compatible and collusion proof. The result highlighted with an asterisk (*) is not true in general, but for a large range of reasonable parameter values.

5 Additional Considerations

5.1 Existence and Growth of DAOs

Our analysis focuses on the governance of DAOs, which has not been sufficiently scrutinized to date. However, existing economic theory can explain why such virtual enterprises based on blockchain networks exist and how they transform existing value-creating interactions (Constantinides et al., 2018). When Coase (1937) showed that production is organized in hierarchies, because their transaction costs in the context of imperfect information are lower than those of a pure market exchange, the view of the firm was still monolithic. This changed with the work of Jensen and Meckling (1976), through which we have come to consider the organization as a nexus of contracts. In this spirit, DAOs simply represent a radical unbundling and outsourcing of corporate functions.

To understand when and why this could be desirable, we need to consider the efficiency of institutional arrangements as measured by their transaction costs. Williamson (1973, 2010) predicts that the pure market exchange is favored over a traditional firm, if asset specificity and uncertainty are low. This closely reflects the situation in blockchain networks, which rely on generic digital production factors and remove information asymmetries through their transparent distributed ledgers. Thus, given the rapid dissemination of blockchain technology, transaction cost economics can rationalize the strong growth in both the number and size of DAOs highlighted in the introductory section.

5.2 The Importance of Staking

The staking mechanism is a crucial design choice for a DAO, because other incentive schemes are much less viable. Evidently, the absence of managers and hierarchies preclude the feasibility of monitoring approaches. In addition, signalling is hampered by the anonymity (pseudonymity) of the DAO participants and the known shortcomings of reputation-based rating systems. Another problem is the dominance of one-off assignments in online labour markets (Stanton and Thomas, 2020). As individuals are employed on a transactional basis and not through long-term contracts, they cannot be penalized for malfeasance after the task has been completed (see, e.g., Martimort et al., 2017).

5.3 Unobservable Effort

Our model framework is based on the assumption that effort is observable by humans, but not by machines.[24] This assumption can be relaxed in several directions. If effort is also

[24]This is generally reasonable. In the case of the MakerDAO risk team members (see Section 2.2), e.g., a software can only check the timely submission of research reports, but not their quality.

verifiable electronically, there will be no need for a voting process. The DAO can simply be made collusion proof by coding the punishment for malfeasance into a smart contract, which would automatically execute once a lack of effort is detected. Furthermore, assume effort is unobservable for both humans and machines, but humans receive a noisy signal. In this case, our results remain valid as long as the informative component of the signal amounts to at least 50 percent.[25] The shirker can then still be identified (and convicted) in expectation.[26] Finally, if effort is completely unobservable (even by humans), the DAO will not be feasible. This is because nobody can identify the shirker and instigate a voting for the confiscation and redistribution of his stake.

5.4 Staking vs. Performance Bonds

Economically, cryptographic stakes are similar to performance bonds, a form of collateral that is placed in a trust fund to deter workers from shirking (see, e.g., Becker and Stigler, 1974; Lazear, 1981; Akerlof and Katz, 1989). The latter, however, are known to suffer from material implementation hurdles such as double moral hazard problems and legal constraints (Ritter and Taylor, 1994). In a classical setting, double moral hazard arises, because the firm may be tempted to wrongfully accuse a worker to confiscate the collateral. This is much less of an issue in a DAO: the decision process is decentralized and the stake is paid in network-specific tokens. Accordingly, rational individuals should seelittle benefit in wrongfully convicting an honest peer as long as that harms the reputa- tion of the DAO and decreases the value of its tokens. Another major difference between performance bonds and stakes is that the latter carry voting rights. It is this specific characteristic that gives rise to our core economic problem in the first place: establishing the governance of the DAO in a way to make it incentive compatible and collusion proof.

5.5 Stake Redistribution

In Proposition 2, we showed that distributing the confiscated stake of a malicious worker among all honest peers is strictly more collusion resilient than burning it. There are additional considerations associated with this outcome. First of all, burning and stake redistribution might be equivalent under certain circumstances. If the supply of tokens is fixed and tokens are irreversibly removed from the network through burning, then the

[25] A simple extension of this case is a signal with less than 50 percent information content, but the possibility to improve it at a cost. If these costs are too high, then the DAO would be feasible, but socially undesirable (refer to Area A of Figure 3).

[26] Such a noisy signal would lower the bribing costs, because, given their imperfect information, some honest participants will vote for acquittal anyway.

value of the remaining tokens should rise for any given value of the DAO.[27] Thus, given an efficient secondary market for tokens, stake redistribution can also be done implicitly.

Secondly, explicit stake redistribution might be still be preferable to implicit stake redistribution, because it can be implemented contingent on voting. Nexus Mutual and Aragon (court) even use a form of strong vote contingent redistribution. This means that participants only get a reward if they voted with the consensus (Cuende and Izquierdo, 2017; Karp and Melbardis, 2017). While, in our model, we assumed that every person either votes for conviction or for acquittal, in reality individuals may also abstain from voting altogether. By redistributing stakes contingent on voting, participants are incentivized not to abstain. Despite this evident benefit, however, there is a risk that an undesired equilibrium will emerge, where an innocent worker instead of the faulty worker is convicted and loses his stake.

5.6 The Role of Blockchain Technology

There is a crucial role of the underlying blockchain technology for the feasibility of a DAO. In particular, DAOs need a permissionless (public) blockchain to come to fruition. If the governance principles and the records of the DAO were stored centrally or maintained by a handful of administrator nodes as in the permissioned blockchain space, then very few individuals or entities could control them and simply change any parameter to their advantage. In this scenario, all of our results become void: a decentralized value creation process can never be collusion proof without decentralized record keeping.

6 Conclusion

We develop a microeconomic model for blockchain-based DAOs and demonstrate that incentive compatibility and collusion proofness are challenging to reach at the same time. Incentive compatibility can be achieved through a staking mechanism. However, more effort must be instigated by higher stakes, which increase the temptation of workers to shirk and then bribe their peers for acquittal. To be inherently collusion proof, DAOs either require a high revenue or a sufficient degree of decentralization in combination witha large network size. While these are natural long term targets, most DAOs will need to start small and grow organically. Thus, in the early stages, an adequate set of governance instruments is imperative. Masked voting can help to mitigate the problem. Yet, it is not compatible with full transparency, as advocated in the original DAO idea. In addition, it

[27] Consider the following example: if a DAO with a token supply of 1'000 is worth USD 1 million, then the slashing of 500 tokens will increase of the value of the remaining tokens by 50%. Economically, this isthe same as proportionally redistributing the 500 tokens.

will still not be collusion proof in some situations, albeit rather theoretical ones. Therefore, the most promising approach to preclude collusion is to opt for stochastic voting and have the voting power concentrated among a few key governance token holders.

Owing to the unprecedented growth in the number and size of DAOs throughout the last years, this new organizational form has become a critical part of the Web3.0. Furthermore, their dominance in the DeFi space is likely to make DAOs a cornerstone of the future financial services infrastructure. Therefore, our work exhibits a high practical relevance, as it equips developers with an effective governance for their DAOs. Moreover, our findings may also relevant for other forms of decentralized value creation that have come to the fore in recent years. One case is the gig economy, an online labor market that enables principals to task freelance workers with jobs on a one-off basis. Another related concept are business ecosystems. The latter connect companies from different geographies through a digital platform to provide a single seamless solution for the customer. Both of these arrangements can be viewed as precursors of the DAO and have the potential to gradually evolve towards a pure transactional setting based on smart contracts.

Notwithstanding these considerations, there are limitations to our work and at least two directions for future research. First, while our analysis is based on a normative framework, there are behavioral reasons that could dampen both the incentive and the collusion problem. Network participants with a preference for honesty or a strong belief in the purpose of the DAO, e.g., might be less likely to act maliciously. The relevance of such factors could be determined by experimental research. Second, potential conflicts of interest between workers and token holders have not yet been considered. While the former should be interested in increasing the compensation for their tasks, the latter will want to increase the overall network value. The extent to which such issues could threaten the feasibility of a DAO warrants further scholarly attention.

References

Akerlof, G. A. and Katz, L. F. (1989). Workers' trust funds and the logic of wage profiles. *The Quarterly Journal of Economics*, 104(3):525–536.

Antràs, P. (2005a). Incomplete contracts and the product cycle. *American Economic Review*, 95(4):1054–1073.

Antràs, P. (2005b). Property rights and the international organization of production. *American Economic Review*, 95(2):25–32.

Becker, G. S. and Stigler, G. J. (1974). Law enforcement, malfeasance, and compensation of enforcers. *The Journal of Legal Studies*, 3(1):1–18.

Bester, H. (1985). Screening vs. rationing in credit markets with imperfect information. *American Economic Review*, 75(4):850–855.

Buterin, V. (2013). A next-generation smart contract and decentralized application platform. *Ethereum White Paper*.

Coase, R. H. (1937). The nature of the firm. *Economica*, 4(16):386–405.

Constantinides, P., Henfridsson, O., and Parker, G. G. (2018). Platforms and infrastructures in the digital age. *Information Systems Research*, 29(2):381–400.

Cuende, L. and Izquierdo, J. (2017). Aragon network: A decentralized infrastructure for value exchange. *Whitepaper*.

Dal Bó, E. (2007). Bribing voters. *American Journal of Political Science*, 51(4):789–803.

Dekel, E., Jackson, M. O., and Wolinsky, A. (2008). Vote buying: general elections. *Journal of Political Economy*, 116(2):351–380.

Guillen, P., Merrett, D., and Slonim, R. (2015). A new solution for the moral hazard problem in team production. *Management Science*, 61(7):1514–1530.

He, P., Tang, D., and Wang, J. (2020). Staking pool centralization in proof-of-stake blockchain network. *SSRN Electronic Journal*.

Hendershott, T., Zhang, X., Leon Zhao, J., and Zheng, Z. (2021). Fintech as a game changer: Overview of research frontiers. *Information Systems Research*, 32(1):1–17.

Huang, N., Burtch, G., Hong, Y., and Pavlou, P. A. (2020). Unemployment and worker participation in the gig economy: Evidence from an online labor market. *Information Systems Research*, 31(2):431–448.

Jarvenpaa, S. L., Shaw, T. R., and Staples, D. S. (2004). Toward contextualized theories of trust: The role of trust in global virtual teams. *Information Systems Research*, 15(3):250–267.

Jensen, M. C. and Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4):305–360.

Karp, H. and Melbardis, R. (2017). Nexus mutual whitepaper. *Whitepaper*.

Lazear, E. P. (1981). Agency, Earnings Profiles, Productivity, and Hours Restrictions. *American Economic Review*, 71(4):606–620.

MakerDAO (2019). The maker protocol: Makerdao's multi-collateral dai (mcd) system. *Whitepaper*.

Martimort, D., Semenov, A., and Stole, L. (2017). A theory of contracts with limited enforcement. *The Review of Economic Studies*, 84(2):816–852.

Peng, C. H., Lurie, N. H., and Slaughter, S. A. (2019). Using technology to persuade: Visual representation technologies and consensus seeking in virtual teams. *Information Systems Research*, 30(3):948–962.

Ritter, J. A. and Taylor, L. J. (1994). Workers as creditors: Performance bonds and efficiency wages. *American Economic Review*, 84(3):694–704.

Saleh, F. (2021). Blockchain without waste: Proof-of-stake. *The Review of Financial Studies*, 34(3):1156–1190.

Stanton, C. T. and Thomas, C. (2020). The gig economy beyond local services and transportation. *CESifo Forum*, 21(3):21–26.

Tiwana, A. and Konsynski, B. (2010). Complementarities between organizational IT architecture and governance structure. *Information Systems Research*, 21(2):288–304.

Tiwana, A., Konsynski, B., and Bush, A. A. (2010). Platform evolution: Coevolution of platform architecture, governance, and environmental dynamics. *Information Systems Research*, 21(4):675–687.

Wakefield, R. L., Leidner, D. E., and Garrison, G. (2008). A model of conflict, leadership, and performance in virtual teams. *Information Systems Research*, 19(4):434–455.

Wang, J. J., Capponi, A., and Zhang, H. (2022). A theory of collateral requirements for central counterparties. *Management Science*.

Williamson, O. E. (1973). Markets and hierarchies: Some elementary considerations. *American Economic Review*, 63(2):316–325.

Williamson, O. E. (2010). Transaction cost economics: The natural progression. *American Economic Review*, 100(3):673–690.

World Economic Forum (2018). Blockchain beyond the hype: A practical framework for business leaders.

Xu, J. and Livshits, B. (2019). The anatomy of a cryptocurrency pump-and-dump scheme. *28th USENIX Security Symposium*.

A Proofs

A.1 Proposition 1

Proof. Notice that (IC) puts a lower bound and (BC) an upper-bound on the stake. The DAO is feasible, if we can find a unique optimal stake that fulfills both (IC) and (BC).

First, we show that the optimal incentive-compatible stake is independent of the bribing constraint. If the bribing constraint is non-binding, we may select the lowest stake, which fulfills all other constraints. In contrast, if the bribing constraint is binding, then no solution exists and the DAO is not feasible.

First, reconsider (IC)

$$z_i R - e_i \geq z_i R p_0 - S_i(1 - p_0), \tag{A.1}$$

then solve for the incentive-compatible stake:

$$S_i \geq \frac{e_i}{1 - \ _i} - z\ R. \tag{A.2}$$

The stake enters the objective function value negatively. Hence, the *optimal* incentive-compatible stake is the smallest value of S_i, which still fulfills (A.2):

$$S_i = \frac{e_i}{1 - \ _i} - z\ R \tag{A.3}$$

In combination with the finding that the optimal revenue shares z_i are proportional to the effort (see Lemma 2), we can pin down the optimal stake as a function of total revenue R and individual effort e_i.:

$$S_i^* = \frac{e_i}{1 - p_0} - \ , \frac{e_i}{\sum_{i=1}^{n} e_i} R. \tag{A.4}$$

■

A.2 Proposition 2

Proof. To compare the DAO's options after the stake of a shirker has been confiscated, we consider the difference between the expected value of voting for conviction and the expected value of accepting a bribe and voting for acquittal. Denote this difference with Ψ:

$$\Psi = \mathrm{E}_i(i = \mathit{Conviction}) - \mathrm{E}_i(i = \mathit{Acquittal}). \tag{A.5}$$

For any giving bribing strategy targeted at person *i*, we aim to maximize Ψ, by selecting either burning or redistribution of confiscated stakes.

Burning

If the DAO decides in favor of burning the confiscated tokens, the expected payoff of worker *i*, given that he votes for conviction, equals:

$$E_i(i = Conviction) = 0P(C \setminus Q) + 0P(Q) - DS_iP(A \setminus Q). \quad (A.6)$$

The first two terms imply that there is nothing to gain by voting for conviction, both in the pivotal and in the nonpivotal case. The third term represents the absolute loss in token value caused by an acquittal of the guilty worker and the associated undermining of trust in the DAO. The probability of this scenario equals that of all realizations, which lead to acquittal, except the ones where the individual under consideration is the pivotal voter.

Moreover, the expected payoff of worker *i*, given he votes for acquittal is:

$$E_i(i = Acquittal) = 0P(C \setminus Q) - DS_iP(Q) - DS_iP(A \setminus Q) + \sum_{k \in K} B_{i,k}P(k) \quad (A.7)$$

Accepting the bribe, yields a non-negative payoff, depending on the bribers strategy and the realized voting outcome *k*. This is reflected by the last term in (A.7). The other difference compared to (A.6) comes from the second term. This is the loss in token value that voter *i* causes by his own pivotal decision to acquit the guilty worker.

Hence, we have

$$\Psi_{Burning} = DS_iP(Q) - \sum_{k \in K} B_{i,k}P(k). \quad (A.8)$$

Stake Redistribution

Let S_R denote the stake of the shirker, to differentiate it from the stake S_i of the person targeted with a bribe. In case of stake redistribution, the stake of the shirker is allocated to the honest workers, proportionally to their own stakes. To simplify notation, denote the shirker by *s* and the subset of workers, excluding the shirker by $M \subset N$ with $\{s \in N \mid s \$ M\}$. The expected payoff of worker *i*, given he votes for conviction then equals:

$$E_i(i = Conviction) = S_R, \frac{S_i}{\sum_{j \in M} S_j} P(C \setminus Q) + S_R, \frac{S_i}{\sum_{j \in M} S_j} P(Q) - DS_iP(A \setminus Q) \qquad i \neq j, \quad (A.9)$$

where we count to $n - 1$, because the convicted worker does not receive a part of the re-distributed stake ($i \setminus j$). If the voting outcome is *conviction*, then worker i receives his share of the confiscated stake S_R, not matter whether his voting decision was pivotal or not (first two terms). If the voting outcome is *acquittal*, he faces an absolute loss in token value (DS_i).

Similarly, the expected payoff of worker i, given he votes for acquittal, can be written as follows:

$$E_i(i = Acquittal) = S_R \frac{S_i}{\sum_{j \in M} S_j} P(C \setminus Q) - DS_i P(A \setminus Q) - DS_i P(Q) + \sum_{k \in K} B_{i,k} P(k) \qquad i \setminus j. \tag{A.10}$$

The difference to (A.9) lies in the last two terms. If the voting outcome is *acquittal* and the vote of worker i was pivotal, he caused himself a loss in token value of size DS_i. At the same time, however, in voting for acquittal, he expects to collect the non-negative cash flow associated with the bribe.

Combining (A.10) and (A.9) leads to

$$\Psi_{Redistribution} = S_R \frac{S_i}{\sum_{j \in M} S_j} P(Q) + DS_i P(Q) - \sum_{k \in K} B_{i,k} P(k) \qquad i \setminus j. \tag{A.11}$$

Through a comparison of (A.8) and (A.11), it is easy to see that

$$\Psi_{Redistribution} \geq \Psi_{Burning}. \tag{A.12}$$

■

A.3 Proposition 3

Proof. We are interested in a sufficient condition for a collusion-proof DAO, so we can ignore the first two summands on the right-hand side of (BC). Since

$$\frac{\partial c(m)}{\partial m} > 0, \tag{A.13}$$

we are essentially looking for a lower bound on m. First, we need to understand how m is linked to the concentration of the voting power:

Lemma 3 *The number of DAO members that a shirker needs to contact to achieve collusion is inversely related to the degree to which voting power in the DAO is concentrated.*

Proof. We need to show that

$$m = f(H_V), \qquad \frac{\partial m(H_V)}{\partial H_V} < 0. \tag{A.14}$$

The shirker decides on a bribing strategy,[28] given the degree of concentration of the voting power in the DAO as reflected by H_V. In a DAO with an absolute-majority decision rule, the minimum number of individuals (including the shirker) required to achieve a majority vote is:

$$m = \min_{\kappa} \left\{ \kappa \,\middle|\, \frac{\sum_{j=1}^{\kappa} S_j}{\sum_{i=1}^{n} S_i} \geq \frac{1}{2} \right\} \tag{A.15}$$

The malfeasant worker needs to secure at least half of the voting power. It is rational for him to choose a strategy that minimizes the amount of people that he needs to contact to achieve this goal.

Now, rearrange (12) as follows:

$$\sum_{i=1}^{n} S_i = \sqrt{\frac{1}{H_V} \sum_{i=1}^{n} S_i^2} \tag{A.16}$$

and insert into (A.15) to obtain:

$$m = \min_{\kappa} \left\{ \kappa \,\middle|\, \frac{\sum_{j=1}^{\kappa} S_j}{\sqrt{\sum_{i=1}^{n} S_i^2}} \geq \sqrt{\frac{1}{H_V}} \cdot \frac{1}{2} \right\}. \tag{A.17}$$

Evidently, m will be higher, the lower H_V, i.e., the less voting power is concentrated.

■

Next, we determine how the voting power must be allocated to minimize H_V for a fixed $m = \bar{m}$:

Lemma 4 *The lowest possible H_V for a given $m = \bar{m}$ is reached for a uniform distribution of the voting power among the smallest group of network participants that controls least 50 percent of the overall voting power.*

Proof. Consider the following situation: At least 50 percent of the voting power is uniformly distributed among $\bar{m}$ participants. These are the network participants, which the shirker aims to contact. Call a participant from this set a K_1 participant. The rest of

[28] A bribing strategy comprises the size of the individual bribes ($B_{j,k}$), the number of DAO members thatare contacted with the offer (m), and the condition under which bribes will be paid ($k \in$ K or $k \in$ A).

the voting power is uniformly distributed among the remaining $(n - \bar{m})$ participants with $(n - \bar{m}) \geq \bar{m}$. Call a participant from this set a K_2 participant. For ease of exposition and w.l.o.g., now consider the special case where the two groups K_1 and K_2 hold exactly $\frac{1}{2}$ of the voting power:

$$H_V = \bar{m}\left(\frac{1}{2\bar{m}}\right)^2 + (n - \bar{m})\left(\frac{1}{2(n - \bar{m})}\right)^2. \tag{A.18}$$

Notice how any deviation from the status quo distribution increases H_V. If, e.g., we increase the voting power of one of the K_1 participants by c at the expense of another K_1 participant, we obtain the following new value H_V^J:

$$H_V^J = \left(\frac{1+c}{2\bar{m}}\right)^2 + \left(\frac{1-c}{2\bar{m}}\right)^2 + (\bar{m} - 2)\left(\frac{1}{2\bar{m}}\right)^2 + (n - \bar{m})\left(\frac{1}{2(n - \bar{m})}\right)^2. \tag{A.19}$$

The net change is:

$$H_V^J - H_V = \left(\frac{1+c}{2\bar{m}}\right)^2 + \left(\frac{1-c}{2\bar{m}}\right)^2 - 2\left(\frac{1}{2\bar{m}}\right)^2, \tag{A.20}$$

which can be reduced to

$$H_V^J - H_V = \frac{2c^2}{(2\bar{m})^2} > 0. \tag{A.21}$$

The same argument can be made for any change of the voting power of K_2 participants.[29] Hence, the outlined situation with uniformly distributed voting rights in each of the two groups is associated with the minimal Herfindahl index H_V^{min} for a given network size n (and fixed $\bar{m}$). [30] ■

Finally, we express $\bar{m}$ as a function of a given network size $\bar{n}$ and concentration of the voting power $\bar{H}_V$. From (A.18) we get:

$$\bar{H}_V = \frac{1}{4\bar{m}} + \frac{1}{4(\bar{n} - \bar{m})} = \frac{\bar{n}}{4\bar{m}(\bar{n} - \bar{m})} \tag{A.22}$$

Rearranging delivers a quadratic equation:

$$\bar{m}^2 - \bar{n}\bar{m} + \frac{\bar{n}}{4\bar{H}_V} = 0. \tag{A.23}$$

[29] Moreover, this logic also applies to those cases in which the distribution of voting power between the K_1 participants and the K_2 participants is not 50/50. Note that if a single participant owns more that 50 percent of the voting rights, the DAO is generally unfeasible.

[30] Similarly, the maximum Herfindahl index H_V^{max} obtains, if the 50 percent voting power of the K_1 partic- ipants

is maximally concentrated inside the group, which is substantively no different from the case where $\bar{m} = 1$.

We are interested in the following root of (A.23):

$$\bar{m} = \frac{\bar{n}}{2} - \frac{1}{2}\sqrt{\bar{n}^2 - \frac{\bar{n}}{\bar{H}_V}}. \tag{A.24}$$

For the specific the network size $\bar{n}$ and degree of concentration $\bar{H}_V$, the shirker needs to contact and bribe at least $\bar{m}$ participants to be acquitted. Therefore, (A.24) describes a lower bound for m:[31]

$$m \geq \frac{n}{2} - \frac{1}{2}\sqrt{n^2 - \frac{n}{H_V}}. \tag{A.25}$$

■

A.4 Proposition 4

Proof.

Under a masked voting scheme, bribes cannot be conditioned on individual votes. This heavily restricts the probability space K, implying a substantial reduction of the set of bribing strategies available to the shirker. We therefore scrutinize strategies contingent on the voting power (stake share) of network participants, which is transparent even when the actual votes are masked.[32] To begin with, consider the following feasible strategy. Offering

$$\frac{1}{2}\left(S_R + D\sum_{i=1}^{n} S_i\right) + c \tag{A.26}$$

(with $c > 0$) to the majority of the voting power overcompensates their aggregate payoff in the conviction state and thus makes acquittal the dominant strategy.

A briber might be tempted to reduce the bribing costs associated with this strategy. He could do so by:

i offering an individual voter with voting power $\frac{S_i}{\sum_{i=1}^{n} S_i}$ less than $\frac{S_i}{\sum_{j\in M} S_j} S_R + DS_i$ $(i \in j)$. However, this will induce him to vote for conviction.

ii Offering the strategy to participants which together hold less than half of the voting power. However, this will never lead to $P(A) > 0$.

iii Turning to pivotal strategies.

[31] Evidently, the lower bound is inversely related to H_V. From Lemma 4 we know that, for a given numberof network participants n, the smallest H_V ($= H^{min}$) and therefore the largest lower bound for m is associated with a uniform distribution of voting rights (in the each of the two groups K_1 and K_2).

[32] Note that this proof is an application of an argument first made by Dal Bó (2007).

Due to masked voting, it is not possible to condition the bribing strategy on pivotal individuals. However, the briber may condition payoffs on specific aggregate voting outcomes. Our particular interest lies in pivotal results one of which is, e.g., given if the briber is acquitted with a razor thin margin. In this case, the briber could offer a payoff to more participants, not differentiating between those who voted for acquittal and those who voted for conviction, but only pay conditional on a specific vote share. Thus, to complete the proof, we need to show that such a pivotal bribing strategies does not work.

Under the aforementioned strategy, the bribe B offered to each individual participant, contingent on a pivotal aggregate voting outcome, must fulfill the following condition:

$$B \geq S_R \frac{S_i}{\sum_{j \in M} S_j} + DS_i \qquad i \neq j. \tag{A.27}$$

Furthermore note that the payment has to be atleast as large for any outcome where more participants vote for acquittal. If not the participant would vote for conviction to increase the likelihood of a pivotal outcome. However, this means that more than the majority voting power has to be compensated. Therefore any such strategy is bounded below by equation A.26.

$$B = \frac{1}{2}\left(S_R + D\sum_{i=1}^{n} S_i + c\right). \tag{A.28}$$

Now turn to the perspective of the briber. For the DAO to be collusion proof, the cheapest possible bribing strategy (A.26) must be more expensive than the stake of the briber that will be confiscated in case of conviction:

$$S_R \leq \frac{1}{2}\left(S_R + D\sum_{i=1}^{n} S_i + c\right) \tag{A.29}$$

In a nutshell, the optimal stake can be implemented and the system is collusion proof, except for a few extreme cases. Those are given, if ceteris paribus,

i the token value response is negligible,

ii the sum of all stakes in the network ($\sum_{i=1}^{n} S_i$) is small,

iii the largest stake commanded by any of the participants is dominant relative to the sum of all stakes, i.e., if the network is highly centralized.

■

A.5 Proposition 5

Proof.

We focus on strictly dominant bribing strategies. Under such strategies, a unique equilibrium emerges where bribing is nearly costless and the DAO is not collusion proof. To prove that this is true, we switch into the decision making perspective of the bribed. Recall from (A.11) that, in case of stake redistribution, the difference in expected payoffs between voting for conviction and voting for acquittal is:[33]

$$\Psi_i = 0P(\mathrm{C} \setminus \mathrm{Q}) + \left(S_R \frac{S_i}{\sum_{j \in M} S_j} + DS_i \right) P(\mathrm{Q}) + 0P(\mathrm{A} \setminus \mathrm{Q}) - \sum_{k \in \mathrm{K}} B_{i,k} P(k) \qquad i \neq j. \tag{A.30}$$

Individual i will only vote for acquittal, if $\Psi_i < 0$. To explicitly specify the bribe term, now consider the following dominant nash strategy from the perspective of the shirker: offer individuals that vote for acquittal a contingent bribe that exactly match the payoff of conviction in every state and additionally comprises an $c \geq 0$. Formally, such a bribe can be described as follows:

$$B_{i,k} = \begin{cases} c & \text{if } k \in \mathrm{C} \setminus \mathrm{Q} \\ S_R \frac{S_i}{\sum_{j \in M} S_j} + DS_i + c & i \neq j \quad \text{if } k \in \mathrm{Q} \\ c & \text{if } k \in \mathrm{A} \setminus \mathrm{Q}. \end{cases} \tag{A.31}$$

In words: if the outcome is conviction, the individual who voted for acquittal was not pivotal ($k \in \mathrm{C} \setminus \mathrm{Q}$) and will receive a compensation for the lost share in the redistributed stake plus c. If the outcome is acquittal and the individual was pivotal ($k \in \mathrm{Q}$), his payoffwill additionally include a compensation for the loss in token value (which is assumed tooccur in case of a false acquittal). Finally, if the outcome is acquittal and the individual was not pivotal ($k \in \mathrm{A} \setminus \mathrm{Q}$), the payoff will be c.

With this bribing scheme, it is a dominant nash-strategy for all participants to vote acquittal, implying $P(\mathrm{A} \setminus \mathrm{Q}) = 1$. In addition, the realized costs for the briber will be nc and therefore arbitrarily close to zero. To see this, simply compare the expected payoffs to the voters:

$$\mathrm{E}_i(i = \textit{Conviction}) = S_R \frac{S_i}{\sum_{j \in M} S_j} P(\mathrm{C} \setminus \mathrm{Q}) + S_R \frac{S_i}{\sum_{j \in M} S_j} P(\mathrm{Q}) - DS_i P(\mathrm{A} \setminus \mathrm{Q}) \qquad i \neq j \tag{A.32}$$

[33] For $\mathrm{E}_i(i = \textit{Conviction})$ and $\mathrm{E}_i(i = \textit{Acquittal})$ refer to equations (A.9) and (A.10), respectively. We count to $n - 1$, since the convicted worker does not count when determining the weights for the redistribution.

$$\mathrm{E}_i(i = \textit{Acquittal}) = \left(S_R, \frac{S_i}{\sum_{j\in M} S_j} + c\right) P(\mathrm{C} \setminus \mathrm{Q}) + \left(S_R, \frac{S_i}{\sum_{j\in M} S_j} + c\right) P(\mathrm{Q}) + (-DS_i + c)P(\mathrm{A} \setminus \mathrm{Q}) \qquad i \neq j. \tag{A.33}$$

Evidently, the shirker only needs to offer an infinitesimal amount c to ensure $\Psi_i = \mathrm{E}_i(i = \textit{Conviction}) - \mathrm{E}_i(i = \textit{Acquittal}) < 0$. This means that an acquittal decision can be ensured for a cost of $(n-1)c$, which is close to zero.

■

A.6 Proposition 6

Stochastic voting implies that voter i is pivotal with a fixed probability

$$P(Q) = \frac{S_i}{\sum_{i=1}^{n} S_i}. \tag{A.34}$$

Now, consider the same bribing strategy which led to collusion under majority voting: the shirker offers those individuals that vote for acquittal a contingent bribe, which exactly matches the payoff of conviction in every state and comprises an additional payment $c \geq 0$. Formally, such a bribe looks exactly as in (A.31).

The payoffs from the perspective of the bribed (without inserting (A.31) for B_i) are:

$$\mathrm{E}_i(i = \textit{Conviction}) = S_R \frac{S_i}{\sum_{j\in M} S_j} P(\mathrm{C} \setminus \mathrm{Q}) + S_R \frac{S_i}{\sum_{j\in M} S_j} \frac{S_i}{\sum_{i=1}^{n} S_i} - DS_i P(\mathrm{A} \setminus \mathrm{Q}) \qquad i \neq j, \tag{A.35}$$

and

$$\mathrm{E}_i(i = \textit{Acquittal}) = \left(S_R \frac{S_i}{\sum_{j\in M} S_j} + B_i\right) P(\mathrm{C} \setminus \mathrm{Q}) + (-DS_i + B_i) \frac{S_i}{\sum_{i=1}^{n} S_i} + (-DS_i + B_i)P(\mathrm{A} \setminus \mathrm{Q}) \qquad i \neq j. \tag{A.36}$$

This leads to

$$\Psi_i = \mathrm{E}_i(i = \textit{Conviction}) - \mathrm{E}_i(i = \textit{Acquittal}) = S_R \frac{S_i}{\sum_{j\in M} S_j} + DS_i \frac{S_i}{\sum_{i=1}^{n} S_i} - B_i \qquad i \neq j. \tag{A.37}$$

Hence, if

$$B_i > S_R \frac{S_i}{\sum_{j\in M} S_j} + DS_i \frac{S_i}{\sum_{i=1}^{n} S_i} \qquad i \neq j, \tag{A.38}$$

$j \quad M \qquad\qquad i=1$

then $\Psi_i < 0$ and person i votes for acquittal.

Now assume for a moment that every participant is homogeneous w.r.t. effort and, in turn, stake. In this case, we can drop the subscript and rewrite (A.38) as follows:

$$B = S\frac{1}{n-1} + DS\frac{1}{n}. \tag{A.39}$$

The bribe needs to be paid to half of the network participants. This implies:

$$B = \frac{1}{n-1} + D\frac{Sn}{n2}, \tag{A.40}$$

which can be reduced to:

$$B = \frac{S}{2}\frac{1}{n-1} + D\ . \tag{A.41}$$

Obviously, the total bribing cost will always be smaller than the stake of the shirker:

$$B = \frac{S}{2}\frac{1}{n-1} + D < S. \tag{A.42}$$

Therefore, stochastic voting is not collusion proof.

Now resort back to heterogeneity w.r.t. effort, which leads to heterogeneity in the optimal stake. Any such heterogeneity will make bribing more appealing to the owner of the largest stake. While the bribing costs (A.41) increase slightly, that person's willingness to pay (right hand side of A.42) will be much higher.

A.7 Corollary 1

Once more, we focus on strictly dominant strategies. For any token holder i, owning network-specific tokens worth T_i, the difference in expected payoffs between voting for conviction and voting for acquittal is:[34]

$$\Psi_i = 0P(\mathrm{C} \setminus \mathrm{Q}) + DT_iP(\mathrm{Q}) + 0P(\mathrm{A} \setminus \mathrm{Q}) - \sum_{k\in\mathrm{K}} B_{i,k}P(k). \tag{A.43}$$

The token holder would vote for acquittal, if $\Psi_i < 0$. Now, consider the same dominant nash strategy as in Proposition 5 and Proposition 6: the shirker offers state-contingent bribes that exactly match the payoff of conviction and additionally comprise a marginal amount c in each case. Formally, this strategy now looks as follows:

[34]Token holders do not receive any part of the redistributed stake. Hence, (A.43) exhibits the same structure as (A.8).

$$B_{i,k} = \begin{cases} \vdots & \text{if } k \in C \setminus Q \\ DS_i + c & \text{if } k \in Q \\ c & \text{if } k \in A \setminus Q. \end{cases} \tag{A.44}$$

Consequently, despite the presence of token holders, the total costs of bribing in equilibrium remain at just $(n-1)c$.

A.8 Proposition 7

In addition to the n workers, we now have t governance token holders. The voting power of any token holder i is based on the number T_i of network-specific tokens in his portfolio. Stochastic voting implies that voter i is pivotal with a fixed probability

$$P(Q) = \frac{V_i}{\sum_{i=1}^{n} S + \sum_{i=1}^{t} T_i}, \tag{A.45}$$

where $V_i \in \{S_i, T_i\}$ denotes the stake or the token inventory, in case the voter is a worker or a token holder, respectively. Taking into account that the stake of the malicious worker is exclusively redistributed to workers to compensate them for their loss of revenue, token holder i faces the following difference in expected payoffs between voting for conviction and voting for acquittal (accepting a bribe):[35]

$$\Psi_i = DT_i \frac{T_i}{\sum_{i=1}^{n} S_i + \sum_{i=1}^{t} T_i} - B_i. \tag{A.46}$$

Given that we need $\Psi < 0$ to convince token holder i, the following is the lower bound for the bribe B_i:

$$B_i > D \frac{T_i^2}{\sum_{i=1}^{n} S_i + \sum_{i=1}^{t} T_i}. \tag{A.47}$$

Evidently, the minimum bribe is a squared function of the tokens. Large governance token holders exhibit a higher probability to be the pivotal voter and, in case they do turn out to be pivotal, they have much more to lose.

The shirker pursues the cheapest bribing strategy that secures him the votes of half

[35] Notice the analogy to (A.37)

the network. The corresponding optimization problem can be described as follows.

$$\underset{1^B}{\operatorname{argmin}} \sum_{i=1}^{n} B_{S_i}(1^B_{S_i}) + \sum_{i=1}^{n} B_{T_i}(1^B_{T_i}) \tag{A.48}$$

s.t.

$$\sum_{i=1}^{n} 1^B_{S_i} S_i + \sum_{i=1}^{n} 1^B_{T_i} T_i \geq \frac{1}{2}\left(\sum_{i=1}^{n} S_i + \sum_{i=1}^{n} T_i\right) \tag{A.49}$$

$$B_{S_i}(1^B_{S_i}) = \left(S_R, \frac{S_i}{\sum_{j \in M} S_j} + DS_i\right) \cdot \frac{S_i}{\sum_{i=1}^{n} S_i + \sum_{i=1}^{t} T_i}, \quad i \neq j \tag{A.50}$$

$$B_{T_i}(1^B_{T_i}) = DT_i \cdot \frac{T_i}{\sum_{i=1}^{n} S_i + \sum_{i=1}^{t} T_i}, \tag{A.51}$$

where $1^B_{S_i}$ and $1^B_{T_i}$ indicate the workers and token holders that are bribed.

The shirker can either bribe governance token holders or workers. His objective is to identify and select those token holders and workers that minimize his bribing cost: (A.48). To be successful, at least half of the votes in the DAO need to be secured: (A.49).A token holder with the exact same token amount as a worker is a little bit cheaper to bribe, because the shirker does not have to compensate the former for the loss of the redistributed stake: (A.50) and (A.51). For larger token holders, however, this effect gets quickly outweighed by the bribing cost for the pivot case.

If the total bribing cost faced by the shirker for any strategy exceeds his stake, i.e., if:

$$\sum_{i=1}^{n} B_{S_i}(1^B_{S_i}) + \sum_{i=1}^{n} B_{T_i}(1^B_{T_i}) \geq S_R, \tag{A.52}$$

then the DAO is collusion proof. This implies that, to preclude bribing, at least 50 percent of the votes need to be concentrated among a few relatively wealthy token holders.

Discussion Paper 2

Central Bank Digital Cash

A Credible Commitment to Privacy

Ian Grigg, P4P Foundation

Abstract

Central Banks are committing themselves to issue digital cash to retail or end-user customers. In so doing, they face a number of contradictions in their goals. Imposing AML, promoting financial inclusion, and not upsetting term transformation pose challenges. Riding above these challenges is the paradox of privacy: users will only respect digital cash if it is private and they feel safe in their use. This same privacy ensures that the bounty of transaction data will be a prize of great value, to be enjoyed in the breach by a host of enemies. If Central Banks do not preserve the privacy of users' transactions, they will not adopt. As there is no easy or stable balance in privacy in such a dynamic and complex system, this is an undecidable problem, and so Central Banks must take the users' side. To ensure success, I propose that Central Banks must provide a credible commitment to privacy, else see their designs rejected by a sceptical public.

Keywords

CBDC, Privacy Paradox, Financial Inclusion.

CBT

Central Bank Digital Cash
A Credible Commitment to Privacy

Ian Grigg 2021-2022 [1]

Abstract: Central Banks are committing themselves to issue digital cash to retail or end-user customers. In so doing, they face a number of contradictions in their goals. Imposing AML, promoting financial inclusion, and not upsetting term transformation pose challenges. Riding above these challenges is the paradox of privacy: users will only respect digital cash if it is private and they feel safe in their use. This same privacy ensures that the bounty of transaction data will be a prize of great value, to be enjoyed in the breach by a host of enemies. If Central Banks do not preserve the privacy of users' transactions, they will not adopt. As there is no easyor stable balance in privacy in such a dynamic and complex system, this is an undecidable problem, and so Central Banks must take the users' side. To ensure success, I propose that Central Banks must provide *a credible commitment to privacy*, else see their designs rejectedby a sceptical public.

Introduction

A Central Bank Digital Cash (or CBDC) is a new digital payments system in competition with cash, and with banks. Technically, it is an accounting system moving fiat value from one device to another, where that device might be cards, phones or computers. Legally speaking, a CBDC is a contract of value issued by the Central Bank offering both guaranteed redemption for fiat and the right of transfer.

Much research is being done on this topic by Central Banks (CBs), and it will take a long time. For the most part, this delay is because (a) the CBs have not had to issue a new currency, or a new payments technology in a long time, and must recover the knowledge to do this; (b) a CBDC necessarily involves the retail public, and CBs are disjoint from that user community; and (c) there are a number of stark contradictions that they will face.

Let's be clear on one thing. We know how to build the technology to do this, and have been able to do this since the 1990s. Several systems were fielded in the 90s that were capable of reaching this approximate goal across a range of technologies. To name but a few: Mondex, Chipper and Chipknip put money on smart cards across many countries, and they worked offline. David Chaum's eCash (Chaum 1982) and my own Ricardo (Grigg 2000) are software money; eCash was blinded money (untraceable but known person) whereas Ricardo used pseudonyms (traceable but unknown person), the method that was later adopted by Bitcoin.

[1] This paper received useful and critical commentary from George Papageorgiuo and Konstantinos Sgantzos.

When compared to the issuance of a CBDC, Bitcoin brought little that was new or relevant to this equation, as blockchain's decentralised model does not bear directly on centralised issuance, and the smart contract model is clumsy when it comes to describing simple contractual digital cash. And, speaking as someone who issued fiat cash several times in the 90s and 2000s, Tether & friends are still behind the state of the art in terms of governance or operations. But, commercially, crypto has broken ground, and as a competitive powerhouse it has suggested that CBs must compete.

By this caveat, I stress, although the narrative might bubble with excitement about the technology, the hard questions at hand are not about the technology. Or, as Izabella Kaminska put it (Kaminska 2021):

> *But framing the conversation as a technological challenge is nonsensical. All it does is detract from the very real downsides of overly centralised systems and the true nature of the competition at hand, which is a function of interest-rate arbitrage, the sort of privacy users value more (privacy from the state or from data-mining merchants and other private sector entities), and the question of whether deplatforming is effective at all.*

Rather, all of the unknowns are political choices, which need to be turned into policy decisions, before being handed to the IT guys to implement. Of course, these are critical to get right. But they are also problematic - who makes these choices, and in whose interests?

Cui bono? Of these difficult choices, there are several that stand out as competing, and potentially problematic to achieve them all:

1. Anti-Money Laundering (AML) regulations
2. "Financial Inclusion"
3. Zero impact on banking
4. Privacy

I suggest that these choices present hard contradictions, each of which alone could sink the CBDC project. Combined, they are much more problematic, and thus demand a much more serious analysis from Central Banks than has hitherto been seen.

Of these four, one could argue that the first three are business as usual. But the fourth, privacy, is setting up society for a dramatic shift at a fundamental level, one which will move us from a world of private transactions to a world of mass financial surveillance. Such a change in the very fabric of how our society works demands more than Central Bank policy discussions - unwinding the financial privacy that has characterised society for its entire existence is an all-of-society question which will have consequences that few can predict.

For these motives – to examine the all-of-society question of the potential end of financial privacy – let's work briefly through the first three challenges in turn, and then dive more deeply into the fourth.

Challenge 1 - AML regulations

Our first choice, **Anti-Money Laundering,** is a non-choice. In the mid-2000s, Central Banks signed up to the agenda of a secretive and undemocratic organisation called the Financial Action Task Force (FATF), and have since become cheerleaders. The CBs consider it non-negotiable that a digital cash must deliver "financial integrity," a phrase that confusingly covers KYC (loosely derived from *know your customer*), AML (*anti-money laundering*), CFT (*counter-funding of terrorism*), and something to do with taxation.[2] From the perspective of insiders and the CBs themselves, the only thing that matters here are the details of implementation.

Unfortunately, these terms all speak to mass financial surveillance. As Darbha & Arora put it (Darbha & Arora 2020):

> "Maintaining privacy and **complying** with regulations (the latter which requires disclosure of information) present a dichotomy for a CBDC. This is further complicated by **the need for proactive disclosure** to prevent fraud."

The basic assumption of AML is that the authorities can see what you are up to, and can stop you, if you are in some sense on their naughty list.

Controls that kick in at a certain limit

Typically, authorities talk about setting a limit - above some number X they can see, and "follow the money," while below X they can't see.[3] From a security point of view, this is unimpressive, as we don't have anything in the computer science or cryptography toolkit that clearly and obviously fixes X as a number that the people can rely upon. Nor, from a political point of view, will a fixed X be acceptable.

To underscore the intent of occasionally changing X to suit local politics, in recent times, the US tax department tried to move one such reporting X down to $600 *annually* (Davison 2021) and the EU voted to set its equivalent X, also aggregated over a year, to €1000. Meanwhile the President of Kenya instructed that its X be lifted to above $10,000 for a single transaction (Indeje 2021).

[2] In this paper, for brevity, I prefer the term AML, and use it broadly to include KYC, CFT, and any other applicable terms such as "financial integrity," which latter is best avoided as a misnomer, as the financial system retains its integrity no matter how much money laundering it suffers.

[3] We can make a similar argument about blocking transactions above X, but that has primarily an economic effect not a privacy effect.

By way of comparison, Kenya's limit would be above their annual average salary. For Americans, this would equate to moving the reporting threshold *for a single transaction* above say $50,000, suggesting a variance in reporting of 100 to 1, without even considering the aggregate demand of the IRS. Not all is in Kenyans' favour as Kenya's tax department went to court to get the right to see all transactions (Juma 2020), suggesting they want to set their X at zero.

Hence any control will necessarily be a dial that can move X up and down at the behest of authorities, at will, and thus X will be arbitrary; in essence, the people are being offered a bait & switch by their Central Banks as they will likely promote a high X in rollout and then find themselves lowering it according to the politics of the day.

Privacy on the other hand says it's none of other people's business, and dangerous to the individual if this information is shared. Legally, privacy is well established in principle with for example US 4th Amendment; *"The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated…"*; Also see EU's Charter of Fundamental Rights, Articles 7, 8, and the UN's Universal Declaration, Article 12.

All of these legal rights suggest that spying on people's transactions on a routine basis is illegal, and only court-supervised access should be permitted.

Thus, Darbha & Arora's point above is not only a dichotomy, it is a pressing legal conflict. The people should be justly scared of any Central Bank that introduces mass financial surveillance over their privave transactions. And the people would have a valid cause to reject the project. To underscore the arbitrariness of the CB's acceptance of the AML agenda without question, these prescriptions from the FATF are on record as having delivered approximately zero benefit after 3 decades of trying, for massive costs to the people (Pol, 2020).[4]

Hence our first contradiction - the CBs want financial surveillance to stop bad people doing bad things. The people do not want financial surveillance, to stop bad people doing bad things to them.

Challenge 2 - Financial Inclusion

The next challenge taken on by Central Banks is **Financial Inclusion**, a euphemism that can be best expressed simply as the desire to ensure that all people have access to means ofpayment. Such access is provided naturally by paper and coin money. Beyond any particular

[4] Dr Pol suggests that the benefit is almost zero, in that it is measured at below 1% and above 0%. As a scientist I would suggest that (a) such a measurement is 'below the noise level' e.g. is approximately zero, and (b) considering the normal error bounds for such a measurement in social sciences, it is as likely to be below 0% as above 0%. That is, we need to investigate the hypothesis that AML causes ML, as much as we need to confirm the FATF hypothesis that AML reduces ML. That polemic noted, I suggest that 'approximately zero benefits' captures the quibble with brevity.

technology, payment is a human right: access to means of payment is a necessary consequence of people's right to be paid for work, and the merchant's right to be paid for their goods. The alternative is dark and stark - people who cannot be paid for their work are slaves, people who cannot pay for their food must rob to eat.

Financial Inclusion originally had a simpler, illusory meaning of access to bank accounts - this was a view championed by international organisations such as the World Bank, IMF, and the NGO charity or aid sector. Hence the arisal of an aphorism - let's bank the unbanked.

But, the term has always suffered. For one, it has become a favoured catch-word in contexts that make little sense. The international aid or NGO world has championed it for 3 decades since the work of Mohammad Yunus and Hernando de Sota, running trial after trial, without everseemingly tying financial inclusion to the actual problems on the ground. More recently, the Central Bank of Bahamas rolled out their 'sand dollar' for motives of financial inclusion in a context that does not suggest it is a big issue (data artist, 2020).

For another, the term is backwards, and it is better to think of Financial Exclusion than inclusion. What is excluding people? In the Global South, people do not have access to bank accounts for several good reasons: they were too expensive, banks were too untrustworthy, or accounts delivered no feature that people want or need. This might be a surprise, but consider that in Africa, the cost of running a bank account is about the same as in the west, around $5 per month, including in areas that boast monthly income around $60. In Africa, bank accounts are luxury items! These people are 'unbanked' in the same sense that people in Kibera don't buy Prada or Rolexes, but nobody talks about Rolexing the UnRolexed.

An observation. When I lived in my first non-Western country, I noticed that many of the houses were built with flat concrete roofs, but unfinished - they had rebar (metal spikes to reinforce concrete) and plastic water pipes sticking upwards to the next missing floor, awaiting further building.

It took me a while to find out why - a local explained that their banks are unsafe as they frequently collapse through insider theft and bad management. People had learnt over time to put their savings into building not into banks. Savers start out with land, and build up the supplies to work on each successive floor; this also explained the piles of concrete blocks and sand sitting beside unfinished houses for many years.

Since then, I've lived in and visited many nominally poor countries, and there's a clear correlation between people saving by building, and the weakness of the banking sector, which you can spot by walking around and looking at the rebar.

Banks in poor countries are also typically unsafe, as it is an easy scam for insiders (tellers) to raid fat accounts, and they collapse more frequently. Perhaps unsurprisingly, banks see someone on a low income as a poor account and loan risk. In summary, Financial Inclusion

might have been an enticing narrative in the West, but it just did not work, *implicitly*, for the economics.

But in the 21st century, two new contrasting developments emerged to change the approach to Financial Inclusion. Firstly, the invention and rapid deployment of *mPesa* in Kenya showed that banking and unbanking wasn't what it was about. Instead, a simple non-bank payment product from a mobile phone telecommunications company ("telco") did the trick, and as millions of micro-businesses could suddenly and safely pay remotely for goods, Kenya experienced what iswidely recognised as a financial miracle. This experience directly validates the Central Banks' desire to consider Financial Inclusion as a plausible goal, and to issue a separate money as digital cash to meet that goal.

Secondly, a new blockage emerged, as Financial Inclusion changed character with the introduction of deliberate and aggressive moves to exclude people. The global rollout of AML, accompanied by the motive of *suspicion*, created explicit financial exclusion as policy; it increased the expense of payments accounts, and made providers more untrustworthy as payments were blocked and reversed, and accounts were frozen or closed for arbitrary motivesdelivered by bank compliance personnel struggling to apply suspicion over sparse transaction data (or, by artificial intelligence engines badly trained to exclude at a much lower cost).

Central Banks will be unable or unlikely to improve reliability for inclusive payments with digital cash because they have already decided that AML is required, and it works very well to exclude in a digital system. CBs will not be able to assist in financial inclusion because they are on the side of financial exclusion. This is a shame, because the human rights aspect of giving people access to payments clearly trumps that of fighting an ineffective and unwinnable war against money laundering - a choice that, perhaps uniquely in the world, the Central Bank of Kenya (CBK) made correctly in supporting mPesa for all.

An anecdote. Back in Kenya in the early 2010s, we were talking to CBK about how our proposals for social savings were to work (Grigg 2021). Once they understood what we were trying to do, they were all smiles and they opened up with chatter and anecdotes.

Mpesa was the big thing in Kenya, so all conversations turned to this. One such story stuck with me: they shared with us their observations of how the big money was moving using mPesa. In essence what was happening was this - local merchants were loading up SIM cards to full capacity, which if I recall correctly was around $10k worth, more than an annual average salary, and they were trading the SIMs directly, *as if money*.

That is, in some big deal, a payer would hand over to his payee a plastic baggie of SIMs with the PIN number written on each. The payee would insert each in turn into their own phone, and check that Safaricom would recognise the SIM as being fully loaded to the limit of $10k. And in this way, the bag of cards was recognised for its entire value.

Now, at one level, it was a laugh to hear how limits were bypassed by the savvy street traders.

But far more important was how the CBK was dealing with it. They were watching! They weren't fussed about it in the slightest. They had intelligence! It struck me that these guys were wise; much wiser than that steady stream of idiot white visitors, called the Wazungu, from the various and many lettered international agencies such as FATF, IMF, WB, UN and so forth, those that came blind and cloth-eared, aiming only to promote their one-size-fits-all WEIRD prescriptions[5].

The CBK was conducting actual real risk analysis, a thing that the Wazungu did not have, because they could not see the wood for the trees.

How then did Kenya manage the miracle, when so many others failed? There were many factors but a few highlights include (Omwansa 2012):

- When mPesa was rolled out in 2007, the AML measures that cause financial exclusion were not as strong as they are today, and CBK was able to slow and dampen the impact of exclusionary measures until mPesa had achieved full reach across the economy.
- The issuer, Safaricom, was a non-bank, specifically a telco, and therefore had no conflict of interest with protecting its other payments, of which, more in the next section.
- mPesa was the first, and thus banks were taken by surprise.
- Once they woke up to the danger, CBK ran interference to stop the banks from destroying mPesa.
- The political system was aligned more with Safaricom than with the banks; the President's family was a big shareholder in the privatised national telco.

This is by way of recognising that these factors, critical for mPesa's success in bringing an entire country's worth of poor into digital payments, are hard to replicate. Choosing financial inclusion, or fighting financial exclusion, is in contradiction to the place that today's Central Banks find themselves in.

Challenge 3 - Zero impact on the banks

Next, **impact on the banks**. If the CB's digital cash works, people will hold balances in this digital unit; indeed one can suggest that the leading metric of success for the project will be howmuch balances the people hold, in comparison to other holdings, especially cash at bank and cash in pocket. In effect, any reduction of bank deposits and of cash holdings will be useful evidence of the success of the CBDC.

In classical digital cash thinking, CBDC is cash, and is for example uninsured and losable.[6] It would not be considered as equivalent to cash held in bank accounts, and thus would not count

[5] WEIRD stands for Western, Educated, Industrialised, Rich, Democratic - a comment on policies and practices invented for those countries but applied without thought to others, generally with bad outcomes. Wazungu translates roughly as aimless or blind wanderer.

[6] By "classical thinking" I mean the theory and practice developed in the 1990s, but note that some re-thinking is to be expected.

as *deposits* to banks. Indeed, we could suggest that if digital cash works as expected, people won't need bank accounts for much of their activity, as routine daily and monthly payments consume the vast majority of most people's balances. The retail individual banking customer will only want bank accounts for *credit* functions such as cards, overdrafts, mortgages, car leases, emergencies, and some savings if interest rates ever rise again. Perhaps that's it?

Yet, historically, banks have reacted aggressively to competition in payments (Dowd 1996). This aggression is in parts because (a) to banks, deposits are subsidised by government and industry structure, and are thus ultra-cheap and safe loans to the bank, (b) a base of depositors allows banks to upsell profitable credit and other services, (c) they earn fee revenue in payments flow, which has become a serious component under today's era of low interest rates, and (d) the rise of deep data processing in the digital pattern of payments.

At least for (a) above, the banking world has frequently granted itself a quasi-monopoly on payment instruments - because, their narrative says, without very strong protective barriers, banks would lose deposits and then be less able to issue the loans that drive the expansive sideof the economy. Whether we like banks or not, it is a fact that collecting all of the public's deposits *on demand* and turning them into loans to the public *at term* is a very important pillar ofthe operation of the economy. Some would even say that this function, called *term transformation*, is the very definition of banking, and the very reason why Central Banks are essential. Central Banks are of course very aware of this issue, and supportive of this power.

Then, for all of these reasons we can expect banks to fight aggressively against the CBDC. This is to be expected, even in the Global South where banking products are not competitive, simply because of the way banks think – in Kenya we knew a senior banking analyst whose unofficial mission was to kill mPesa, even as bank accounts were unsuitable for the larger population. Indeed, it was this reaction that killed the rollout of mPesa across wider Africa, as banks in other countries made mPesa impossible to roll out by hook or by crook. Banks are also behind the ongoing war against cash, as they see every cash transaction as a stolen opportunity. In a recent win, banks were able to scare people away from using cash in the recent COVID19 epidemic.

To address these concerns, the CBs are building compromises into the CBDC product.

One compromise is to outsource the customer-facing part of the business to the commercial banks. Initially, this was considered sensible because the Central Banks have no capability in retail operations, and, as rolling out a new customer-facing infrastructure would be a heavy lift, it's possibly wiser to use the available customer-facing resources and networks that already exist in the market. This then results in a two tier architecture where banks will own their portionof customers, to some extent alleviating the competition between CBDCs and deposit accounts.

One might reasonably critique this arrangement as suggesting that the CBs are not the right agents to get involved in a retail product, but the counter to that is that the network effects of a money are so strong that one single issuer is likely to dominate. One can also point to the

obvious flaw in the argument that banks are the right commercial customer-facing agents; telcos are actually a lot more adept at this because they already have the secure platform, theyalready deal in low levels of money, and the mPesa experience is proof that it works.

A second compromise is to lean more heavily on AML than is warranted. The unfortunate logic is that if banks employ hard AML on digital cash, they can reduce the effectiveness of digital cash, and thus preserve the rightful territory of payments to themselves. Indeed, we could reasonably ask if the motive of "protecting the banks" is a bigger factor than the more customary"financial integrity" excuse; we can certainly expect the banks to promote AML for their own mission of reducing competition. Which reveals a form of hubris: AML fails at effectiveness inits headline tasks (Pol, 2020), but hands to its employers a powerful weapon: *carte blanche* to discriminate against whosoever is not wanted.

And so the fundamental contradiction is revealed between the banks' competitive hatred of alternate payment systems, and the Central Banks' promise to do no harm to the banks. CBs would like to have their cake and eat it too, but the outcome of compromise is a simple weakening of the digital cash product, one that imperils the entire project.

Which reveals the paucity of the CB's mission thinking. What was it they were trying to do in the first place? What was the particular benefit that made CBs want to do this? If banks can so easily manage to corral the CBDC back into their custody, was the mission really that important?

Challenge 4 - Privacy

The above gives a brief description of three of the factors that make the CBDC an inordinately difficult goal, brief because of space, and only because we want to show something of the map where it impacts privacy.

Now turning to **privacy**. The ECB revealed that the #1 requirement of the users they surveyed was privacy (Arnold, 2021). That matches my experience - in past efforts, consumers may not have been vocal about this issue, but when they detected that privacy is inadequate, they quietly avoided the product. And especially, note that (a) early adopters are more privacy conscious than those who follow, (b) they have an outsized say in who follows, and (c) that anymoney is extremely dependent on network effects. Money needs mass adoption in order to be money.

Another anecdote. While I was doing a stint working on the über-private Chipper smart card money system of the late 1990s, in the Netherlands, a country known then for its obsessive approach to privacy, I happened to be at a party. With lots of drugs, which is normal in Amsterdam, not that I took the opportunity.

And at this party was a chap who was tagged as a local dealer of recreational product. Now, I'd long since learnt that if you want to do anything in the security world, you have to know both your customer and your customer's enemy. This sounded like an ideal chance to kill two birds with one stone, so I struck up a conversation about dealing, and we got around to talking about digital cash.

Over beers, he was very forthcoming and helpful on his trade, so I reciprocated and told him about this new form of super privacy money coming out. And I asked, "would you use this for your deals?" He said, absolutely not. He was unequivocal on this point.

Surprised, I asked why not, and he replied, words to effect, "it won't be private. No matter what they say or what people think, the government will ensure it won't be private." That put a new perspective on things for me, because, firstly I'd just spent some time telling him how private it was, and secondly, I knew the security team. They were good, strong, honest people. Dutch people, obsessed with privacy! They'd worked for 2 years on the security model, and *it was private*. Dammit!

Which left me thinking, a digital cash would surely gain a mark of privacy quality about it, if we could only get the dealers to use it. Or, any criminals, really. But I wasn't clear on how to sell this marketing proposal to anyone else...

Given the attention of early adopters within the *product adoption cycle*, it is critical, crucially critical to get privacy right up front, in order to create a positive climate for roll-out. And, by the by, we could probably measure the success of the privacy goal by checking how the bad guys perceive and use the system, as shown above with CBK.

This is just the dynamics of how payment systems work. In essence you can't fudge the privacy aspect, you can't leave it until later or lean on vague marketing words. It's got to be solid. Rock solid.

Diving Deep in Privacy

That turns out to be a pretty big challenge, as there will be sceptics. Let's dive deeper into this thing called privacy.

In the security world, we talk about threats – often using the phrase "what's your threat model?" ("WYTM?").

We can start that discussion by asking, *who* is the threat here? For privacy, there is a long list: thieves, neighbours, family, employees & employers, data miners, BigTech, police, tax, spooks, social security, random other government departments… Indeed, it is fairer to say that, by default, *everyone is a threat to your privacy*.

Therefore, in privacy engineering, we generally start from a principled position of "it's none of your business" and allow nobody to access a person's data. This is actually something that people work on, and that search has sparked such innovations as public key cryptography, the original eCash design of blinded signatures (of which there is a long and studied literature) and disappearing messages. Privacy engineering is a thing.

However, Central Banks are authorities, and authorities are not independent, rather they are beholden to many different interests. For example, one of their interests is having detailed data on what people are doing with money, for monetary policy and other reasons. Hence, CBs will not adopt the notion that "it's none of their business" because they believe that what you are doing with your money is indeed their business.

> *"Anonymised/aggregate data on the use of the digital euro should be available to the Eurosystem under any privacy option for statistical, research, supervisory and oversight purposes, including to fight fraud/illicit activities."* (ECB 2022)

Further, limitations on the privacy of money is not without precedent; even cash for example has technologies built in to breach the privacy of payments. A century or so ago, serial numbers were added onto notes following a rash of kidnapping cases in the USA, in which the kidnappers could cash out safely as the notes were really, actually untraceable. Adding serial numbers to all notes gave a way to trace where the ransom was being spent, and gave a lead toinvestigators.

If it happened to cash - the holiest of holies in the privacy world - then it will happen to digital cash. We've already stated that AML will be an article of faith in CBDCs, so we can assume that the starting point is that all digital cash will be traceable and identifiable as to the persons inany transactions. The ECB calls this its baseline position:

> *"Transparent to intermediary*
> *Checks during onboarding*
> *Data transparent to intermediary for AML/CFT purposes"* (ECB 2022)

The BIS has also put a heavy stake in the ground - not only will digital cash carry solid KYC, they are going to expand from there, solve the 'identity problem' and put all sorts of other personal data into the system (BIS 2021). Space does not permit us to examine this flight of fantasy (see for example Grigg 2021), but suffice to say that this is currently the BIS thinking, which we can assume is advice that all CBs will be comfortable adopting, even as they will find it difficult to implement in a free and democratic society.

Then, following from the above discussion, we can assume that the data will exist. The next question then is, who gets to enjoy it?

Who has access to the data?

Let's start with no-one. Not a soul. As a thought experiment, because some of us technologists know it can be done, and therefore it should be done! [7]

Unfortunately, the claim of *no-one* will be subject to a number of exceptions. Let's count them.

The Intelligence Community

The first exception is the spooks. The intelligence community (IC) are basically those who are sanctioned (permitted) to conduct crimes on behalf of the state. Spying, and other similar crimes such as stealing intellectual property and military secrets, kidnappings, regime changes, fraud, renditions, torture and assassinations, are allowed by the IC: they are "legal" when initiated within an initiating country against a target country, but are basically and obviously totally illegal inside every country they are done to. Indeed, being caught spying in somebody else's country is probably the most illegal thing that can be done, it gets a poorer reception than murder or robbing a bank - in war time, being caught spying on the battlefield gets you summaryexecution, which means death *without trial*.

One outcome of living in the IC is that people within it are very comfortable with breaking the law, and they only get slightly uncomfortable when it's done in their jurisdiction, to their own people, and then, only if they get caught. The spooks are masters of not getting caught! For them, laws might apply in principle, but need not apply in practice.

Which leads us to our first answer: the spooks will be given access to the entire database, because if not, ***they'll steal it***. Doesn't matter how they do it, and space does not permit us to dwell on this, but suffice to say it is 100% sure that they'll steal it.

Not that this is a serious consideration because it is also entirely sure that the IC will be in the room and will steer the technical design of any digital cash so that it's easy for them to access. How do we know? Because they are in the room for every other similar system; the IC is interested in wherever cryptography standards are being created or deployed at scale, such as mobile phones. Every major corporation that ships cryptography will do it with the IC's oversight, which means backdoors, and standards that use cryptography are nudged by the ICaway from our defensive mission and towards their offensive mission. It gets worse - now they have expanded their purview to wherever data is touched including data protection (EC 2021)

[7] Many technologists will rush forward and say that we can solve this problem of privacy with what we call exotic cryptography - blinded money, zero-knowledge proofs, ZK-SNARKs and the like. But, firstly, these things are not mainstream, and are scary for a reason. Until they've stood the test of time, and especially proven they can be simply treated as black boxes with solid characteristics, they are more likely to be too brittle to be trusted in the very important goal of protecting users' privacy against persistent attackers. Secondly, this misses the entire point ofthis discussion, which is about whether we choose private money, not how to build it. And thirdly, combining these two points, if some exotic cryptography is deployed, *because privacy*, how are we to tell whether it is real, or it is backdoored somehow?

and digital services (Fanta 2022). Every major system that tries to hide important data has somewhere some person who represents the needs of the IC.

It's just the way they work. You might not spot their influence, but that's only because they live in the shadows. Once you know the tricks of their trade, they can be spotted.

In conclusion, the spooks will have access. To everything. Supporters will rush to claim civil society, anti-terrorism, misinformation and so forth, but it's not our role today to judge their today excuses (a fascinating subject that it is). Let's just assume the spooks have it.

We have our first exception. Who's next?

The Police

How about the police? We could for example say that the police will have access if they get a particularised court order from a judge, pursuant to probable cause. This will be a strong protection as judges know what that means and won't grant an order without seeing some evidence.

This will work, but only for a time, because the spooks will eventually get sick of the crimes going on and will leak the information to the police. For example, in the USA, the NSA was caught sharing intelligence information on crimes with 19 different domestic agencies, and evengoing so far as to teach the police how to lie to the judge - by the creation of a false train of evidence to convince the judge that it was based on some other random source such as a tip- off, an insider or a lucky traffic stop.

They called it *parallel construction*, I call it perjury, obstruction of justice, contempt of court, conspiracy, deception and probably some more crimes thrown in. One can understand the frustration of the NSA and other spying agencies. They've often got the taps on all the criminal gangs, they hear who gets knocked off and who is about to get hit. Why not share it with the police?

The point here is to not judge their actions but to point to the power of human nature. The information will flow, if it is useful. For example, consider the various and many pandemic track & trace systems (Venkataramakrishnan 2021):

> *"In response to the coronavirus pandemic, the Singaporean government set up a contact tracing system reliant on a central database, allowing staff to rapidly locate and contact those who may have been exposed to coronavirus. Such a trade-off in personal privacy for the public good may have been deemed acceptable during a crisis, but a change in policy now allows the police to access this data."*

This process of human nature will happen with the people inside the CBs - the ones who have to manage the secret tracking for the spooks. They will eventually set up access to the police,

or if they show some moral spine, they will be replaced with people who are comfortable with doing it, which is called a *secret cell* in the trade.

That's our second exception. Perhaps we are comfortable with the police having some access? After all, we've nothing to hide because we've done nothing wrong.

The Banks

If the list of exceptions could be particularised and made strong, then this would be perhaps a credible commitment. But it cannot.

What about the banks? We already killed that one as the BIS has already stated that an individual can only have one CBDC account, so, as a consequence, each bank has to have access to all known CBDC accounts to at least find out if yours is already set up. And they'll find a way to encourage you to open up the transaction records so they can be more holistic intheir financial surveillance. But it's even easier than that.

Another anecdote. While (still) working on the über-private Chipper solution, I found a manager who was trying to whip his card out of the reader at exactly the right time to crash the protocol.

Let me explain: in a classical digital cash protocol on smart cards, 1990s style, each card would have a little spot inside it with a value for the money. My spot might have 100, yours 200. When I pay you, we now have to do a "dual database transaction", instantaneously, on both databases. As a fundamental of computer science, instantaneous between 2 different places never works, and nobody's built a quantum smart card yet, so it has to be serialised - one after the other.

If say, your spot increases by 50 to 250, and then mine decreases by 50, we have an interesting problem. My manager mate can whip his card out just after the first action, and then he's got 50 extra, and I've still got my 50! Woot! Free money… which in the trade was known as the *evergreen card*.

Or we could do it the other way around: decrease mine and only then increase yours afterwards. It's just a protocol choice, and this second way is generally how smart card money was done. But it still leaves the problem of what happens when you whip out your card at the precise wrong point, because now we - or I! - have the *lost money problem*.

Recalling that knowing your customers was critical to this business of security, I asked him why he wanted to lose his money. Because, he said, then he'd go up to the support office, and ask them to get it back!

My ears were buzzing at this point. This was a good question, and I was somewhat self-disgusted at not asking it myself. OK, *how are they going to get it back?*

"They'll look in *the database*, confirm it happened, then restart the transaction!"

I didn't hear the last part, because by now my ears had been blown off by the alarm bells. At this point I was very carefully going back in my mind to read the first part above, over and over again. "They'll look in the *database…*" What *database*?

And so the whole house of cards collapsed. Talking to everyone I was able to figure out what had happened.

1. The Central Bank of the Netherlands ("De Nederlandsche Bank" or DNB) had ordered the Chipper team to store all the transactions into a database so that in the event of a *meltdown* (another term of art, meaning that somehow, crooks found a way to really spank the system and make it completely compromised), then the damage could be debugged andrepaired.
2. The security team, and the whole company, were dead set on privacy, so the compromise was to keep the database entirely secret, and only 2 senior managers were trusted with special access.
3. Then the "lost money" problem turned up, and the 2 trusted people had to dive in, check the database, and proceed to recover the money.
4. Then, more lost money, and more and more… so the 2 special people handed access to the support team so they could deal with it more routinely.

And all of this happened while everyone (else) in the company still believed the system was private. Massive institutional cognitive dissonance!

And of course, I was now pondering my own beliefs, as the drug dealer had been right, and I'd been a muppet to believe the narrative.

Considering the above anecdote: you're happily doing your digital cash thing, and then a problem occurs. Doesn't matter what problem, as it's software, and problems always occur. So you pick up the phone to your local support agency, which is almost certainly your sponsoring bank (see Challenge 3 above) and ask for help.

You describe the problem… What is the bank going to do?

Of course, the bank must solve the problem, otherwise, happy customer becomes angry customer. Today, if anything, systems need more support because consumers have become more demanding, less tolerant and less willing to blame themselves. And, not to forget, it's *money*; whilst banks might be annoyed by a minor complaint, for the customers this is at minimum a red flag, and at maximum a loss of their earnings.

The banks will need to engage seriously on this, and of course, as their customer service agents, the banks will need to turn angry customers into happy customers. The help desk operator will need to see all your activity. Oops! The customer won't be able to just provide it on the spot, for various security and other reasons - the help desk operator will need an independent path of access so that they can discover what really happened, and choose the appropriate mitigation.[8]

And so we have our 3rd exception - the help desk will have access, else *no help*, no growth, and no success.

The Taxman

Let's move right along. Consider taxation. The tax department in your country will look at that treasure trove of payment information and quite fairly, in its humble opinion, decide it needs it. For example, the US Treasury recently proposed a rule that any recipient should report any annual aggregate inflows/outflows exceeding $600 of crypto directly to the IRS (Davison 2021).

As we the people will assume that the tax department doesn't have it, because *privacy*, then we will all use CBDC for all our dodgy deals. The tax department will of course hear about this, and appeal to the courts for a blanket order to get access, or to the parliament for a blanket law. It will probably be rejected, and that will be expected, so the taxman will try again. The more we the people are using it for private transactions, the more the money works, the more the tax department will want access. So they'll try again and again.

Eventually a sympathetic judge or parliament will let the taxman have it - and order whoever has the data to open it up.

This reveals a paradox of privacy - the more the privacy, the more we trust and use the money; yet the more the privacy, the more our trust, the greater the value of breaking it, and the more the tax department will fight to get it. Perversely, a corollary to the paradox is that once that privacy is broken, the more those who want privacy will flee to other methods; will those that are left with nothing to hide be enough to sustain the system?

It is without a doubt that eventually the tax department will have access. Once that happens, we now have civil authorities involved outside the initial core set of actors. Which means that any similar government department that has a plausible story can do so as well - this is the nature of precedents.

Pandemic? Sure, health authorities need to know who is buying banned drugs in pharmacies. Immigration? It is unfortunate that undocumented workers can access digital cash, but we can

[8] We can supercharge this example by considering a major robbery of digital cash from a major merchant. Happy merchant discovers that their day's takings have disappeared! Of course, the fast return of funds will be demanded, else, happy merchant becomes angry merchant, and they will turn off the terminals. Big oops!

track and trace them, and show how they're being paid for work they are not permitted to do. Dirtbag fathers skipping out on child support? We can't have that. Social security will need access because of the amount of benefits fraud.

Our 4th exception then is every "official" agency that has a good case to make. Which is a lot of them.

Your Enemies

And, once a half dozen or more "trusted" authorities get access, that means anyone can have access - for a fee. All private investigators will have their friends inside the authorities who for a little consideration will do searches. For example, if you end up in that unfortunate story known as divorce, any secret expenses you paid for with super private digital cash with your new dating partner might be presented in court. Oops!

Same for criminals, who are even more adept at acquiring special sources. Same for hackers who love the challenge of picking up entire databases of privacy information.

And there's our 5th exception: every enemy of yours who has a little money to bribe someone who has access to a support facility of some form. Enough enemies will have access to the Central Bank's private records that your CBDC will be only private from the ordinary, honest people.

What's the good of that?

Consequences

What are the outcomes from this cascading series of privacy breaches?

Everyone will know that the Central Bank's Digital Cash cannot be used for private things. And people will intend to not use it for private things. Which might ordinarily be OK as private things might only be 10% of our daily purchases.

Yet, it isn't that simple:

1. Nobody can predict what comes back to bite them in the future. That is, our private things might only be 10%, but we don't know which 10% they are. Hence, the desire fora blanket privacy commitment is strong and broad, rather than being particularised or narrow - I can't be attacked by you about something you don't know about; as Cardinal Richelieu said in the French Revolution, *"If you give me six lines written by the hand of the most honest of men, I will find something in them which will hang him."*
2. Out in the public space, there is already a lot of privacy scrutiny on the CBDCs. Already, outside the narrow insider space of CBs and partners, there has arisen broad and deep scepticism, which will only grow as weaknesses are scrutinised by researchers of

greater or lesser expertise. As the system gets closer to the public eye, it is to those adversarial opinions that the public will turn, rather than the marketing blather of CBs.

3. Early adopters are savvy, and privacy is something they typically know more about than the mass market. They will not be happy with something that clearly sets them up for problems in the future.
4. In contrast, because of network effects, CBDCs will be very dependent on mass adoption.
5. This system is dynamic rather than static. Whatever technology the CBs utilise, it will be tunable; it will come with many switches to flip and knobs to turn in both technical and governance terms. As new threats emerge in the minds of the noisy or political classes, such threats have to be stopped, right? This is the slippery slope that turned Paypal from a new libertarian money into a church cake-baking donation system.

In such an environment, a real risk - a clear and present danger - is that the people reject the CBDC.

Not vocally, not by marching on the Central Bank building and burning it down, but passively, by not adopting it. And if they don't adopt it, then neither will the merchants. And if the merchants don't adopt it, it fails - the iron rule of the double sided market is you need a guaranteed win on one side, and you also need a guaranteed win on the other side.

CBDCs are set up for a guaranteed Fail on both sides. As the Director General of GCHQ puts it, about China's DC/EP (Khalaf and Warrell, 2021):

> *"If wrongly implemented, it gives a hostile state the ability to surveil transactions," [Sir Jeremy Fleming] said. "It gives them the ability . . . to be able to exercise control over what is conducted on those digital currencies."*

Obviously, once they have breached privacy, the CBs are in the driver's seat to exercise control over actions - freezing, blocking, closing, seizing - which outcomes also follow naturally from AML. Fleming is talking about China because that's how China stated it will be done, but he could equally well be talking about any country's digital cash, as Western CBs may not be able to do it any other way.

Which is to suggest that the two most likely outcomes are, (i) a mass financial surveillance system, which has been forced on society, or (ii) the users will reject the system.

The users' goal of privacy is not some polite nicety, not just another feature in a list. Privacy is fundamental to our society, our ways and our life. Our evolution through history is based on privacy of money, and it is only in the last 50 years that digital transactions have offered a way to breach. Privacy of money is not a dial to fiddle with because of policy, it is rather part of our very economy and society.

In order to overcome this barrier, I propose that Central Bankers are going to need *a credible commitment to privacy*.

Central banks will need to *commit* because if it isn't a cast-iron promise, CBs will back out. Such a commitment will need to be *credible* because the public is already sceptical and alreadycritical, and simply won't believe words like "we're more friendly to privacy than Facebook." Users will expect CBs to back out, and will walk away before they give CBs the chance.

"Don't give me problems, give me solutions!"

I do not in this paper aim to provide a full solution. Instead, I fall short of "the solution," at the point of just laying out the difficulties, because I do not believe Central Banks have understood what they are heading into. It is my hope that CBs will take on the issue of privacy more seriously than they so far have, if they understand how severely failure of privacy imperils theirproject – privacy isn't a feature, it's a fundamental right. The users know it and demand it, and woe betide any society that walks blindly into a transparent world.

Having said that – "Don't bring me problems, give me solutions!" rings in my ears. Some discussion of potential solutions could assist. One strategy I propose is:

a) Give the spooks access (they'll steal it anyway).
b) Make digital cash through a 2nd tier bank count as *deposits*. That is, if a bank introduces a customer to digital cash, through a tied app, then digital cash inside that app or wallet, owned by the customer, counts as deposits for the purposes of the bank'sbalance sheet and lending equations.
c) Declare access by anyone (other than bona fide national security business, i.e. the spooks) subject to a strict court barrier: particularised and probable cause. No fishing expeditions, no speculative requests, no tricky contracts that conveniently slip in permission in the fine print.
d) Make all other access a criminal offence with mandatory jail time. Strict liability.

This won't stop the most egregious offences, but it will give pause to the thousands of police, bureaucrats, tellers and supporting personnel who might otherwise desire a look, because reasons, because they're the good guys. Selling off access or pushing for loopholes can rebound and jail time is a strong incentive. Find another way.

This solution has drawbacks.

- It is only a partial solution, as we cannot solve the support problem this way. That's unfortunate because support is a big issue, including being critical to adoption. And I'd say, if we can only support the product by opening up a privacy loophole then that may be too high a price to pay. More thinking is required here.
- Central banks have already handed over their independence mandate to the AML compliance juggernaut, and stopping CBs from sharing the data won't be easy. Could there be another exception? Easily, but every exception opens up the privacy to

criminals and others - compliance departments are big and full of bureaucrats trying to make monthly rent (AML *as* corruption is another fascinating topic, no time today).

- It is also going to be an unpopular solution in the political world, as dozens of agencies in each country are already thinking about the potential bounty of all that data. To backtrack on that is going to spark the fiercest pushback from the powers that be, the insiders. But note the paradox of privacy, that bounty is only worth something if the users think it safe from spying, which incentivises all those agencies to breach it.
- Central Bank employees will be surprised that working with statistical data exposes them to strict liability if that data can be de-anonymised, but that's a surprise we the people need them to have. Indeed, nobody in government agencies with access will want (d) above, being mandatory jail time with strict liability. Fair enough, one might say, we didn't expect our bureaucrats to take on liability like that. Or did we? If we the people are to hand over *our very financial lives* to a host of inscrutable and unpunishable agencies, isn't liability, and strong liability, a fair price to ask? Because we the people are going to be facing liabilities if and when our data gets shared, and no government agency will pay the tab for their mistakes. It will all be on us, the unprotected users. If there isn't a balance of liabilities, just why would we do it?

Another approach is proposed in the ECASH bill being prepared for US Congress (Lynch 2022). For privacy, its draft wording requires that it be:

> *"(9) classified and regulated in a manner similar to physical currency for the purposes of anti-money laundering, know-your-customer, counter-terrorism, and transaction reporting laws, and thus not subject to third-party exemptions to a reasonable expectation of privacy;"*

This proposal accepts that the name of each person holding a payment card is known to the authorities, but the transactions between payment cards are treated the same as cash - presumably untraceable. Technically, this proposal expects that a card can do the balance transfer to another card, without each card recording the owner of the other. While providing privacy on paper, we can note that (a) it is only a proposal, and it will have to fight the dozens of agencies who will lobby in Congress committees for their special needs; (b) due to the overall complexity and opacity of a digital cash system, the users will still have difficulty knowing that it is indeed private, and there is no backdoor or bait & switch lurking under the plastic; (c) there is no answer to the meltdown and support issues; and (d) it breaches the Travel Rule which is already in place.

In conclusion, in launching this adventurous mission to deliver private digital cash, Central Banks have truly placed themselves on the horns of a dilemma. How much privacy is needed? A little or a lot? This is an undecidable problem, but if Central Banks don't take on the side of users in this mission, they imperil the project. Central Banks need *a credible commitment to privacy* to bring on the users. Get privacy wrong, and they will set back the field for a decade.

I'm happy to be proven wrong.

References

Martin Arnold, “Europeans raise privacy concerns over digital currency” Financial Times 14th April 2021.
Also see European Central Bank (ECB), “ECB digital euro consultation ends with record level of public feedback” ECB Press release, 13 Jan 2021
https://www.ecb.europa.eu/press/pr/date/2021/html/ecb.pr210113~ec9929f446.en.html

Bank of International Settlements (BIS), “III. CBDCs: an opportunity for the monetary system,” *BIS Annual Economic Report*, BIS 23 June 2021

David Chaum, "Achieving Electronic Privacy," Scientific American, August 1992, p. 96-101
Also see David Chaum. "Blind signatures for untraceable payments". *Advances in Cryptology Proceedings of Crypto 82*. 1983

(Wikipedia) “Chipknip,” accessed 24 Oct 2021
https://en.wikipedia.org/wiki/Chipknip

Sriram Darbha, Rakesh Arora, *“Privacy in CBDC technology”* 2020, Bank of Canada
https://www.bankofcanada.ca/2020/06/staff-analytical-note-2020-9/

The Data Artist, “How many Cbankers does it take to do meaningful CBDC research in the Bahamas?” The Blind Spot 2022
https://the-blindspot.com/how-many-cbankers-does-it-take-to-do-meaningful-cbdc-research-in-the-bahamas/

Laura Davison, “Treasury Rips ‘Misinformation’ on Tax Plan’s Bank-Data Provision”, 14th October 2021
https://www.bloomberg.com/news/articles/2021-10-14/treasury-blasts-misinformation-on-tax-plan-that-faces-hurdles

Kevin Dowd, *Laissez Faire Banking*, Routledge 1996 ISBN 0415137322.
see Chapter 1 “The Evolution of a Free Banking System”
https://iang.org/free_banking/dowd_lfb_intro.html

European Central Bank (ECB), “Digital privacy options,” 2022
https://www.ecb.europa.eu/paym/digital_euro/investigation/profuse/shared/files/dedocs/ecb.dedocs220404.en.pdf

European Commission, “Data Protection: Commission sends a reasoned opinion to BELGIUM for lack of independence of its Data Protection Authority,” in *October infringements package: key decisions*, 12th November 2021

Alexander Fanta, “NATO center wants to be allowed to do research with Facebook data,” Netzpolitik.org, 27th January 2022 (in German).

Ian Grigg, “Financial Cryptography in 7 layers,” Financial Cryptography 2000
https://iang.org/papers/fc7.html

Ian Grigg, *Identity Cycle*, 2021
https://iang.org/identity_cycle/

David Indeje, “Kenya’s Treasury to Revise Upwards Reporting on Large Cash Transactions,” 2021 Khusoko
https://khusoko.com/2021/10/20/kenyas-treasury-to-revise-upwards-reporting-on-large-cash-transactions/

Juma, “Court Allows KRA To Access Your M-Pesa Records And Phone Data”, SokoDirectory 2020
https://sokodirectory.com/2020/02/court-allows-kra-to-access-your-m-pesa-records-and-phone-data/

Izabella Kaminska, “Is the central bank panic about the PBOC coin justified?,” Financial Times, 19th April 2021
https://amp.ft.com/content/76e450be-e8b3-40f3-a452-b20284e0bd63

Roula Khalaf and Helen Warrell, “UK spy chief raises fears over China’s digital renminbi,” Financial Times, 10th December 2021

Rep. Stephen Lynch (MA-08), Chair of the House Committee on Financial Services' Task Force on Financial Technology, “H.R. 7231 - The Electronic Currency and Secure Hardware (ECASH) Act,” 28th March 2022
https://ecashact.us/
https://lynch.house.gov/_cache/files/1/7/17e47e5a-2bff-42c1-b509-eb8a50931fa6/BDDFF183213326821B5C88B7F326EABB.ecashact-lynch.pdf

(Wikipedia) “Mondex,” accessed 24 Oct 2021
https://en.wikipedia.org/wiki/Mondex

Tonny Omwansa, *Money, Real Quick*, Balloon View Ltd 2012 ISBN 101907798455

Ronald F. Pol (2020), “Anti-money laundering: The world's least effective policy experiment? Together, we can fix it, Policy Design and Practice,” DOI: 10.1080/25741292.2020.1725366
https://doi.org/10.1080/25741292.2020.1725366

Siddharth Venkataramakrishnan, “Online privacy: a fraught philosophical debate,” Financial Times

Discussion Paper 3

Stablecoin Regulation: EU, UK and US Perspectives

Pierre Ostercamp

Abstract

Recent EU, UK and US proposals for stablecoin regulation are all based on existing banking, e-money and payments regulations to varying degrees. However, there are limits to applying existing financial regulation to stablecoins due to their unique technological characteristics. As such, this paper argues that the regulation of stablecoins should follow the principle of 'same risk, same regulatory outcome', where stablecoin activities posing the same risks as currently regulated activities could be subject to amended regulatory requirements as long as the same regulatory outcomes are achieved. This paper outlines and compares the EU, UK and US proposals, highlights the limits of applying existing financial regulation to stablecoins, and proposes arguments for how stablecoins should be regulated. This paper argues that centralised stablecoin issuers should be subject to strict requirements on the investment and management of reserve assets in addition to liquidity and capital buffers to ensure redemptions can always be met. Under these and other requirements such as segregation and safeguarding requirements, bank-like deposit protection would not be needed. This paper disagrees with the approach proposed inthe US of regulating stablecoins like bank deposits given that equivalent protections to deposit insurance can be achieved through other means, and stablecoins are generally treated as a means of payment but not a means of saving.

Keywords

Stablecoins, Regulation, Law, Crypto, Cryptoassets.

CBT

Stablecoin Regulation: EU, UK and US Perspectives

Pierre Ostercamp – ostercamppierre@gmail.com

Abstract

Recent EU, UK and US proposals for stablecoin regulation are all based on existing banking, e-money and payments regulations to varying degrees. However, there are limits to applying existing financial regulation to stablecoins due to their unique technological characteristics. As such, this paper argues that the regulation of stablecoins should follow the principle of 'same risk, same regulatory outcome', where stablecoin activities posing the same risks as currently regulated activities could be subject to amended regulatory requirements as long as the same regulatory outcomes are achieved. This paper outlines and compares the EU, UK and USproposals, highlights the limits of applying existing financial regulation to stablecoins, andproposes arguments for how stablecoins should be regulated. This paper argues that centralised stablecoin issuers should be subject to strict requirements on the investment and management of reserve assets in addition to liquidity and capital buffers to ensure redemptions can always be met. Under these and other requirements such as segregation and safeguarding requirements, bank-like deposit protection would not be needed. This paper disagrees with the approachproposed in the US of regulating stablecoins like bank deposits given that equivalent protections to deposit insurance can be achieved through other means, and stablecoins are generally treated as a means of payment but not a means of saving.

Contents

Introduction

Stablecoins attracted international regulatory scrutiny in 2019 when Facebook announced Libra, a global stablecoin project aimed at fostering financial inclusion and innovation by providing a lower cost alternative to traditional payment systems.[1] Regulators around the world were quick to voice concerns that Libra or other global stablecoins could become systemically important payment systems posing systemic risks, and many argued that these risks should be addressed by strict regulatory requirements.[2] Facebook later published a revised proposal for Libra to better account for regulatory concerns, before rebranding Libra as 'Diem'.[3]

Stablecoins are privately-issued digital assets that seek to maintain a stable value relative to the value of one or more fiat currencies, assets and/or commodities. While stablecoin designs vary, they most commonly reference the value of a single fiat currency, in which case they have functional similarities to other payment instruments such as electronic money (e-money) or commercial bank money. Equally, the infrastructure supporting the use of a stablecoin ('stablecoin arrangement') is functionally similar to traditional payment systems and financial market infrastructures. Given these similarities, jurisdictions including the EU, the UK and the US have proposed to apply regulatory requirements to stablecoins and stablecoin arrangements which are largely or entirely based on existing banking, e-money, payments and payment system regulations. Some requirements such as anti-money laundering regulations already apply to stablecoins and stablecoin arrangements in these jurisdictions, but further regulation is needed to mitigate the various risks that stablecoins pose. Notably, these include the risk that a stablecoin issuer would be unable to meet all redemptions of stablecoins for their underlying assets, resulting in users incurring losses.

Following the regulatory approach to cryptoasset-based financial activities proposed by Chiu[4], this paper will 1) identify the risks of stablecoins and relevant regulatory objectives; 2) assess the justifications for introducing stablecoin regulation; and 3) determine how stablecoin regulation should be designed. The first section of this paper will discuss the design and risks of stablecoins (Section I) before evaluating the EU, UK and US proposals for stablecoin regulation

[1] Dirk Zetsche, Ross Buckley and Douglas Arner, 'Regulating Libra' (2020) 41 Oxford Journal of Legal Studies 80.
[2] BIS, 'Investigating the impact of global stablecoins' (*BIS*, 2019) <https://www.bis.org/cpmi/publ/d187.pdf> accessed 19 February 2022.
[3] Diem, 'Economics and the Libra Reserve' (*Diem*, 2020) <https://wp.diem.com/en-US/wp-content/uploads/sites/23/2020/12/EconomicsAndTheReserve_DD_April2020.pdf> accessed 19 February 2022.
[4] Iris H-Y Chiu, *Regulating the Crypto Economy* (Bloomsbury 2021) 291.

(Section II). Next, this paper will discuss further normative arguments and regulatory challenges (Section III), and draw conclusions on how stablecoins should be regulated.

This paper broadly agrees with the regulatory approaches proposed by the EU and the UK, but emphasises that regulators should remain open to different forms of regulation as longas they are likely to achieve the same regulatory outcomes. This is the principle of 'same risk, same regulatory outcome' adopted by the UK government and regulators. Additionally, this paper disagrees with the approach proposed in the US of regulating stablecoins like bankdeposits as stablecoins are generally treated as a means of payment but not a means of saving, and their issuance should not be restricted to banks as bank-like protections can be ensured without deposit insurance. Nonetheless, a key risk of stablecoins is that they will be vulnerable toa run similar to a bank run, so bank-like prudential regulation is justified to mitigate this risk. However, where stablecoin issuers back stablecoins with low-risk liquid assets and do notengage in lending, they should not be subject to prudential requirements as strict as those appliedto relatively higher-risk fractional reserve banks. Further, this paper broadly supports the EU and UK proposals for enhanced regulation and supervision for systemic stablecoins and stablecoin arrangements given the higher risks they pose, but emphasises the importance of clearly defined and limited regulatory discretion.

I. Stablecoins

a. Stablecoin Design

The most common form of a stablecoin is a cryptoasset that references the value of a single fiat currency. As defined under the UK's amended 2017 Money Laundering Regulations (MLRs), a cryptoasset is 'a cryptographically secured digital representation of value or contractual rights that uses a form of distributed ledger technology and can be transferred, stored or traded electronically'.[5] Although some definitions of cryptoassets are technology-neutral, the key underlying technology of a cryptoasset is typically a blockchain: a chronological chain of data blocks which are verified and cryptographically linked to each other such that they cannot be amended.[6] Unless otherwise stated, references to stablecoins in this paper are to blockchain- based stablecoins. This paper proposes that stablecoins can broadly be classified into four

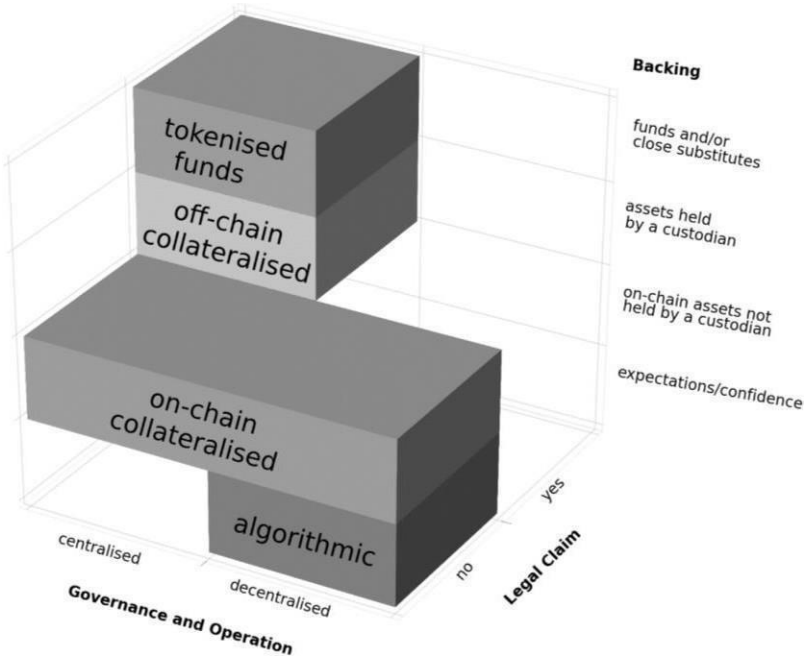
[5] MLRs Regulation 14(3)(a).
[6] Andreas Antonopoulos, *Mastering Bitcoin* (O'Reilly 2017) 211.

categories according to three key characteristics, similar to the classification proposed by the European Central Bank.[7] These characteristics are:

1) type of backing;
2) whether stablecoin holders have a legal claim on the issuer or reserve assets; and
3) (de)centralisation of governance and operation.

Depending on these characteristics, as shown below on Figure 1 (based on the figure proposed by the ECB[8]), stablecoins can be broadly classified as (i) tokenised funds; (ii) off-chain collateralised; (iii) on-chain collateralised; or (iv) algorithmic stablecoins (where on-chain and off-chain refer to, respectively, on-blockchain and off-blockchain). Tokenised funds are pegged to the value of a single fiat currency, most commonly the US dollar, while other stablecoins can in principle be pegged to any value. As tokenised funds are the most prevalent stablecoins and the focus of regulatory scrutiny, this paper will focus on tokenised funds stablecoins.

Figure 1 – Stablecoin Classification



[7] ECB, 'In search for stability in crypto-assets: are stablecoins the solution?' (*ECB*, 2019) <https://op.europa.eu/en/publication-detail/-/publication/61cb4b79-b040-11ea-bb7a-01aa75ed71a1/language-en> accessed 19 February 2022; ECB, 'Stablecoins – no coins, but are they stable?' (*ECB*, 2019) <https://www.ecb.europa.eu/paym/intro/publications/pdf/ecb.mipinfocus191128.en.pdf?adf9157e3e022dceaadac7b24b54ed0b> accessed 19 February 2022.

[8] ibid.

As this is a broad classification framework there may be some exceptions, and some stablecoins are 'hybrid', combining multiple design features such as partial or full on-chain collateralisation with algorithmic stabilisation mechanisms. Algorithmic stablecoins are generally the only type of stablecoin not backed by a reserve of assets, therefore, they are often less stable and perceived as higher risk. As the stability of algorithmic stablecoins depends on confidence rather than backing assets, they are vulnerable to crises of confidence and price collapses even during unstressed market conditions, and so in their current form are unlikely to be viable options for large-scale activities requiring sustained price stability.[9] Nonetheless, algorithmic stablecoins should remain under consideration for future regulation.[10]

Except for some unbacked algorithmic stablecoins, most stablecoins are issued upon receipt of funds or assets, as visualised by the ECB.[11] For tokenised funds, which are the most prevalent form of stablecoins, users typically transfer funds or close substitutes to a stablecoin issuer or their agent or custodian, before the issuer or agent instructs a smart contract on the blockchain to issue stablecoins to the user. Users can generally redeem stablecoins for the equivalent value in funds, where stablecoins are then destroyed by the smart contract before funds are sent to the user. In the context of blockchains, smart contracts are computer programs on blockchains that run automatically when pre-defined conditions are met.[12] Although tokenisedfunds may use smart contracts to issue and redeem stablecoins, they still rely on a centralised issuer, agent and/or custodian to conduct issuance and redemption. References in this paper to smart contracts issuing or redeeming stablecoins refer to decentralised stablecoins where users can interact with the smart contract directly to issue and redeem stablecoins.

While all stablecoins pose risks, many maintain a degree of price stability in relation to their reference assets by maintaining reserve assets backing the stablecoin, as well as other stabilisation mechanisms. A primary risk is of a run on the stablecoin, similar to a run on a bank, where holders lose confidence in the stablecoin's value or the issuer's ability to fulfil redemptions and then all may try to sell or redeem their stablecoins at once. Confidence in a

[9] Christian Catalini, Alonso de Gortari and Nihar Shah, 'Some Simple Economics of Stablecoins' (2021) 13-14 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3985699> accessed 19 February 2022.
[10] UK Finance, 'HMT: UK regulatory approach to cryptoassets and stablecoins' (*UK Finance*, 2021) <https://www.ukfinance.org.uk/system/files/HMT%20Stablecoin%20UKF%20Response%20v1.0.pdf> accessed 19 February 2022.
[11] cf ECB (n 7).
[12] Andreas Antonopoulos and Gavin Wood, *Mastering Ethereum* (O'Reilly 2018) 127.

stablecoin can be maintained by issuers consistently fulfilling redemptions[13], but as reserve assets are not risk-free, there is always a risk that issuers cannot meet all redemptions at par value (Li and Mayer 2020).[14] This risk is elevated in the case of a run as large liquidations of reserve assets over a short time may trigger a fire sale with declining prices, which may force issuers to sell reserve assets at a discount, resulting in insufficient funds being available to meet redemptions. Some issuers aim to protect themselves against this risk by reserving the right to delay redemption and apply early redemption fees to slow a run on the reserve.[15]

However, consistent, fast and frictionless redemption is key as it facilitates arbitrage: another key stabilisation mechanism of almost all stablecoins where market participants 1) buy and then redeem stablecoins; or 2) mint and then sell stablecoins; in order to profit from price discrepancies when the market price is, respectively, below or above the target price. With algorithmic stablecoins, or for participants without redemption rights, this can be done over time based on confidence that the market price will return to the target price. Arbitrage helps minimise fluctuations as significant price deviations can invoke counterbalancing market incentives.[16] Even with effective arbitrage, however, fluctuation would always be possible withina spread, at least because redemption-based arbitrage would only be effective when the stablecoin price deviates by a margin that is greater than the redemption fee plus any other fees.

Stablecoins also differ in the presence and nature of a legal claim on the issuer or reserve assets. Where stablecoins offer a legal claim on an 'identifiable entity', the ECB argues that existing regulation should apply with additional requirements as necessary[17], which will be the case under EU proposals. For decentralised stablecoins, in their current form, there cannot be a legal claim as there is no centralised identifiable entity and smart contracts are generally not recognised as legal contracts. Commentators argue that it still remains necessary for the legal

[13] Tether, 'Response to FSB Stablecoin consultation' (*Tether*, 2020) <https://www.fsb.org/wp-content/uploads/Tether.pdf> accessed 19 February 2022.
[14] Ye Li and Simon Mayer, 'Money Creation in Decentralized Finance: A Dynamic Model of Stablecoin and Crypto Shadow Banking' (2020) Fisher College of Business Working Paper; Timothy Massad, 'Regulating stablecoins isn't just about avoiding systemic risk' (*Brookings*, 2021) <https://www.brookings.edu/research/regulating-stablecoins-isnt-just-about-avoiding-systemic-risk/> accessed 19 February 2022.
[15] cf Diem (n 3).
[16] Amani Moin, Emin Gün Sirer and Kevin Sekniqi, 'A Classification framework for stablecoin designs' (2019) International Conference on Financial Cryptography and Data Security 174.
[17] ECB, 'Stablecoins: Implications for monetary policy, financial stability, market infrastructure and payments, and banking supervision in the euro area' (*ECB*, 2020) <https://www.ecb.europa.eu/pub/pdf/scpops/ecb.op247~fe3df92991.en.pdf?b85631de8b2fdfa5395c2a4c87de05e1 > accessed 19 February 2022.

system to uphold claims in cases of dispute including smart contract failures.[18] Nonetheless, where a claim on the reserve assets is coded in a smart contract and upheld by the consensus rules of the underlying blockchain, that claim might be perceived as similarly robust to a claim upheld by a legal system. Moreover, some smart contracts may soon be recognised as legal contracts. In November 2021, the Law Commission concluded that 'smart legal contracts' can beaccommodated under the law of England and Wales.[19]

Stablecoin governance and operation involves setting parameters such as fees, delays and interest rates, and conducting activities such as issuance and redemption.[20] Governance and operation can broadly be categorised as decentralised or centralised, though levels of (de)centralisation may vary significantly within these categories due to combinations of centralised and decentralised elements. Decentralised governance and operation appeals to many cryptoasset users who value decentralisation and censorship resistance. However, decentralisation can involve co-ordination difficulties and uncertainty due to the absence of trusted intermediaries which have accountability for conducting specified activities.[21] The absence of identifiable intermediaries, which are generally the subject of traditional financial regulation, impedes the application of existing regulations and may explain the apparent omission of decentralised stablecoins in the regulatory proposals discussed in this paper.

All types of stablecoins may charge fees on issuance, redemption and transfer. For centralised stablecoins, fees tend to be imposed to offset operational costs, whereas for decentralised stablecoins, fees often contribute to a secondary price stability mechanism.[22] For example, fees collected by a decentralised stablecoin protocol can be used by a smart contract to buy or sell the stablecoin when its market price is, respectively, significantly below or above the target price, which in turn drives the market price closer to the target price. Stablecoins also vary in the openness of issuance and redemptions. Whereas decentralised stablecoins typically allow any cryptoasset user to request issuance or redemption of stablecoins, tokenised funds and

[18] BIS, 'Stablecoins: risks, potential and regulation' (*BIS*, 2020) <https://www.bis.org/publ/work905.htm> accessed 19 February 2022; cf Zetsche et al. (n 1).
[19] Law Commission, 'Smart legal contracts Advice to Government' (*Law Commission*, 2021) <https://s3-eu-west-2.amazonaws.com/lawcom-prod-storage-11jsxou24uy7q/uploads/2021/11/Smart-legal-contracts-accessible.pdf> accessed 19 February 2022.
[20] Ariah Klages-Mundt, Dominik Harz, Lewis Gudgeon, Jun-You Liu and Andreea Minca, 'Stablecoins 2.0: Economic Foundations and Risk-based Models' (2020) Proceedings of the 2nd ACM Conference on Advances in Financial Technologies 59.
[21] cf Chiu (n 4) 74; Michèle Finck, 'Blockchains: Regulating the Unknown' (2018) 19 German Law Journal 665.
[22] cf ECB (n 7).

centralised on-chain collateralised stablecoins often restrict issuance and redemption to authorised intermediaries and individuals. In this case, the legal claim on the issuer or reserve assets is not universal.

Stablecoins and related service providers also differ in their systemic importance and global reach, with systemic stablecoins and global stablecoins attracting the most regulatory scrutiny. Though inter-related, global reach and systemic importance are distinct characteristics, and global or widespread use does not necessarily imply increased systemic importance or risk.[23]Global stablecoins are stablecoins which have 'potential reach and adoption across multiple jurisdictions and the potential to achieve substantial volume', and could become systemically important.[24] Systemic stablecoins and stablecoin arrangements are those deemed systemically important according to specific criteria or by designation by a regulatory authority, including criteria of size and interconnectedness with the financial system. A good starting point for measuring systemic importance is the guidance under the Principles for Financial Market Infrastructures (PFMI).[25] The PFMI apply to the transfer function of stablecoin arrangements, and to the entire stablecoin arrangement if systemic.[26] Given that systemic stablecoins and arrangements may pose systemic risk, where 'failure or distress of that entity could adversely affect financial stability and the real economy'[27], they warrant stricter regulation according to a risk-based approach.[28]

Use of stablecoins is facilitated by the functions of a broader 'stablecoin arrangement', including 1) issuance, redemption and stabilisation; 2) exchange, transfer and payment; and 3) user services, including custody and network access.[29] Key participants in a stablecoin

[23] cf Tether (n 13) 3.
[24] FSB, 'Regulation, Supervision and Oversight of "Global Stablecoin" Arrangements' (*FSB*, 2020) 5 <https://www.fsb.org/wp-content/uploads/P131020-3.pdf> accessed 19 February 2022.
[25] BIS, 'Principles for Financial Market Infrastructures' (*BIS*, 2012) <https://www.bis.org/cpmi/publ/d101a.pdf> accessed 19 February 2022.
[26] BIS, 'Application of the Principles for Financial Market Infrastructures to stablecoin arrangements' (*BIS*, 2021) 4 <https://www.bis.org/cpmi/publ/d198.pdf> accessed 19 February 2022.
[27] US Treasury, 'President's Working Group on Financial Markets Releases Report and Recommendations on Stablecoins' (*US Treasury*, 2021) 14 <https://home.treasury.gov/news/press-releases/jy0454> accessed 19 February 2021.
[28] Bank of England, 'New forms of digital money' (*BoE*, 2021) 11 <https://www.bankofengland.co.uk/paper/2021/new-forms-of-digital-money> accessed 19 February 2022.
[29] FSB, 'Addressing the regulatory, supervisory and oversight challenges raised by "global stablecoin" arrangements: Consultative document' (*FSB*, 2020) <https://www.fsb.org/2020/04/addressing-the-regulatory-supervisory-and-oversight-challenges-raised-by-global-stablecoin-arrangements-consultative-document/> accessed 19 February 2022.

arrangement include 1) issuers; 2) system operators; 3) cryptoasset exchanges; and 4) cryptoasset wallets.[30] The core composition of a typical stablecoin arrangement is outlined below in Table 1.

Table 1 – Typical Stablecoin Arrangement

Function	Activity	Entities likely to conduct activity
Governance	Setting governance rules for the stablecoin arrangement	Issuers or governance token holders
	Blockchain protocol operation and upgrades	Software developers, nodes and miners/stakers
Management of the stablecoin supply and reserve assets	Issuance, redemption, creation and destruction of stablecoins	Issuers, users or the stablecoin protocol
	Management of reserve assets to ensure stability and redeemability	Issuers or stablecoin protocols
	Custody/trust of reserve assets	Issuers, custodians or stablecoin protocols
Exchange, transfer and payment	Transaction validation	Nodes and miners/stakers
	Stablecoin exchange	Issuers, wallets and exchanges
	Stablecoin transfer	Wallets or payment systems
	Stablecoin payment	Issuers, wallets, exchanges or payment systems
User services	Custody and administration of stablecoins and cryptoassets	Custodial wallets and exchanges
	User access to blockchain networks and applications	Software developers providing user interfaces

b. Stablecoin Risks

As they are a disruptive technological innovation, there are both potentially valuable opportunities and considerable risks associated with stablecoins, which must be carefully balanced by policymakers. Stablecoins could offer improvements in financial services by makingtransactions faster, cheaper, more secure and more efficient, and by enhancing financial inclusion, cross-border payments and payment system resilience. Further, stablecoins couldstrengthen monetary policy by enhancing transmission and decreasing the lower bound.[31]

However, stablecoins may pose various risks depending on their design, systemic importance and interconnectedness with the financial system. Chiu proposes that regulators should identify the risks of cryptoasset-based financial activities and the associated regulatory

[30] HM Treasury, 'UK regulatory approach to cryptoassets and stablecoins: Consultation and call for evidence' (*HMT*, 2021) <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/950206/HM_Treasury_Cryptoasset_and_Stablecoin_consultation.pdf> accessed 19 February 2022.

[31] Bank of England, 'New forms of digital money' (*BoE*, 2021) <https://www.bankofengland.co.uk/paper/2021/new-forms-of-digital-money> accessed 19 February 2022.

objectives before deciding whether regulation is needed to achieve those objectives and how regulation should be designed[32]; this approach will be followed in this paper. A key concern is that stablecoins may pose risks to financial stability, market integrity, consumer protection and effective market competition. Protecting against these risks is often the foundation of a financial regulator's objectives and mandate, therefore, financial regulation aimed at mitigating these risks would be justified. For example, where stablecoins are backed by assets subject to financial risks such as liquidity risk, but there are no user protections such as deposit insurance, prudential regulation could mitigate risks of financial instability and payments disruption.[33]

Where stablecoins are widely used, there would be risks of runs leading to fire sales of reserve assets which may have adverse contagion effects on the broader financial system.[34] A run on a global stablecoin reserve could have systemic negative spill-over effects on global financial markets as the issuer liquidates reserve assets to meet stablecoin redemptions.[35] While issuers could restrict redemptions to mitigate this issue, this could also have negative consequences, including detriment to confidence in financial systems.

Systemic stablecoins and stablecoin arrangements may pose greater risks and even systemic risks due to their size and impact, including risks to financial stability, monetary policy sovereignty and market competition.[36] Without adequate regulation, systemic stablecoins 'could undermine confidence in money and payments and in the financial system as a whole' due to, for example, failures in the security and reliability of stablecoins as payment instruments.[37] Key stakeholders including regulators, governments, central banks and standard-setting bodies have emphasised that regulation, supervision and oversight of stablecoins and related arrangements is necessary to mitigate the associated risks, particularly around systemic stablecoins.[38] Some of the primary risks associated with stablecoins are outlined below in Table 2, which draws on the risk

[32] cf Chiu (n 4) 291.
[33] Bank of England, 'Financial Stability Report August 2020' (*BoE*, 2020) 12 <https://www.bankofengland.co.uk/-/media/boe/files/financial-stability-report/2020/august-2020.pdf> accessed 19 February 2022.
[34] cf BIS (n 18); cf ECB (n 17); ECB, 'A regulatory and financial stability perspective on global stablecoins' (*ECB*, 2020) <https://www.ecb.europa.eu/pub/financial-stability/macroprudential-bulletin/html/ecb.mpbu202005_1~3e9ac10eb1.en.html> accessed 19 February 2022.
[35] ibid.
[36] cf BIS (n 2).
[37] cf BoE (n 31) 45.
[38] ibid; cf ECB (n 17); cf FSB (n 24); cf HMT (n 30); US Treasury, 'G7 Finance Ministers and Central Bank Governors' Statement on Digital Payments' (*US Treasury*, 2020) <https://home.treasury.gov/news/press-releases/sm1152> accessed 19 February 2022.

mapping framework proposed by the WEF[39] and includes the systemic risks identified by the G7 Working Group on Stablecoins.[40]

Table 2 – Key Stablecoin Risks

Risk category	Related risks
Financial risk Risks to the stability and reliability of the stablecoin, particularly around reserve assets	Liquidity risk
	Credit risk
	Market risk
	Counterparty risk
Legal compliance risk Breach or evasion of laws and regulations by stablecoin users or service providers	Financial crime and market integrity
	Data protection and privacy
	Regulatory arbitrage
	Legal uncertainty
Operational risk Failures in the operation and co-ordination of key systems and individuals	Custody risk
	Governance risk
	Operational and cyber resilience
	Blockchain protocol upgrade risk
Technical risk Risks to the security of the stablecoin due to software failures or exploits	Smart contract risk
	Transaction validation and execution risk
	Miner and staker risk
	Oracle risk
Systemic risk Risks to financial stability and the economy posed by systemically important stablecoins and stablecoin arrangements	Financial stability
	Monetary policy transmission
	Monetary sovereignty
	Fair competition and anti-trust policy

c. Regulatory Design

Financial regulation aims to correct market failures.[41] Many of the above risks can be categorised as potential market failures, which the unregulated market may not prevent or sufficientlymitigate due to limitations of market-based governance and self-regulation. Moreover, financial regulators have objectives and legislative mandates to tackle many of these risks. As such, regulation of stablecoins and related arrangements appears necessary and justified, with the remaining question being how stablecoins should be regulated. Regulatory approaches to stablecoins include 1) regulating stablecoins under existing legislation; 2) creating a bespoke stablecoins regime; and 3) imposing restrictions on stablecoin use. The EC considered these options before proposing the Markets in Crypto-assets Regulation (MiCA), which opted for a

[39] World Economic Forum, 'Decentralized Finance (DeFi) Policy-Maker Toolkit' (*WEF*, 2021) <https://www3.weforum.org/docs/WEF_DeFi_Policy_Maker_Toolkit_2021.pdf> accessed 19 February 2022.
[40] cf BIS (n 2).
[41] Steven Schwarcz, 'Regulating Digital Currencies: Towards an Analytical Framework' (2021) 34 Duke Law School Public Law & Legal Theory Series.

combination of a bespoke regime based on existing legislation. The EU, UK and US proposals for stablecoin regulation are all based on existing legislation to varying degrees, drawing in particular on banking, e-money and payments regulations.

The key determinant of the applicability of existing regulation is the legal classification of stablecoins, which for tokenised funds and off-chain collateralised stablecoins is most frequently as 1) e-money; 2) payment systems; or 3) money market funds.[42] Equally, under the EU framework, a stablecoin could be deemed equivalent to a deposit under banking regulation, and a stablecoin arrangement's asset management function could be regulated as an investment fund.[43] As the taxonomy of stablecoins remains under debate, several classifications and existing regulations could apply, though stablecoins are most commonly classified as e-money.[44] As defined under Article 2(2) of the EU's Second E-Money Directive (EMD2), electronic money (e-money) is 'electronically, including magnetically, stored monetary value as represented by a claim on the issuer which is issued on receipt of funds for the purpose of making payment transactions…', and is accepted by persons other than the issuer.

Contrary to commercial bank deposits, e-money is designed to be purely a payment instrument, and 'involves the purchase of a means of payment rather than the creation of a debtor-creditor relationship'.[45] Nonetheless, e-money has functional similarities to deposits, though e-money is not protected by deposit insurance. To prevent users from treating e-money like deposits by using them as a means of saving, most countries prohibit interest payments on e- money holdings.[46] Whether stablecoins are functionally similar to e-money and can thus be regulated similarly depends on their different characteristics, including the legal claim (if any), redemption conditions and stabilisation mechanism.[47] Indeed, MiCA acknowledges that single- fiat stablecoins 'are close to the definition of e-money under the Electronic Money Directive'.[48]

[42] Banque de France, 'Stablecoins: A Brave New World' (*Banque de France*, 2020) <https://publications.banque-france.fr/sites/default/files/medias/documents/wp757.pdf> accessed 19 February 2022.
[43] cf ECB (n 34); Oliver Read and Stefan Schäfer, 'Libra Project: Regulators Act on Global Stablecoins' (2020) 55 Intereconomics 392.
[44] BIS, 'Fintech and payments: regulating digital payment services and e-money' (*BIS*, 2021) <https://www.bis.org/fsi/publ/insights33.htm> accessed 19 February 2022; Agata Ferreira, 'The Curious Case of Stablecoins—Balancing Risks and Rewards?' (2021) 4 Journal of International Economic Law 755.
[45] Mykyta Sokolov, 'Are Libra, Tether, MakerDAO and Paxos issuing e-money? Analysis of 9 stablecoin types under the EU and UK e-money frameworks' (2020) 26 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3746250> accessed 19 February 2022.
[46] cf BIS (n 44) 17.
[47] ibid.
[48] MiCA page 8.

However, of nine major stablecoins analysed under the UK and EU e-money frameworks, Sokolov found that only USDT could be considered e-money because the rest were either not issued on receipt of funds, had variable redemption value or did not grant rights to token holders.[49] On this basis, it is submitted that amendments would be necessary to apply existing e-money regulations to stablecoins, and the approaches for doing so proposed by the EU and UK will be discussed in sections II.a and II.b. Equally, there are challenges in regulating stablecoins as deposits under existing banking regulation, which will be discussed in the context of the US regulatory proposals in section II.c.

II. Regulatory Landscape

a. European Union (EU)

In response to the risks of regulatory fragmentation and regulatory arbitrage due to some cryptoassets falling outside of the scope of EU financial services law, MiCA provides a comprehensive regulatory framework based on existing financial regulation, applying broadly to persons that issue cryptoassets or provide related services within the EU. Like its requirements, MiCA's objectives are inspired by those of existing financial regulation: MiCA aims to achieve legal certainty, promote innovation, protect consumers, investors and market integrity, ensure financial stability, and support the proper functioning of cryptoasset markets.

Stablecoins within scope of MiCA are classified as either 1) 'e-money tokens' (EMTs): stablecoins pegged to a single fiat currency; or 2) 'asset-referenced tokens' (ARTs): stablecoins pegged to multiple fiat currencies, one or more assets, or a combination of the two. As MiCA applies to persons, decentralised on-chain collateralised stablecoins and algorithmic stablecoins appear to fall outside MiCA's scope as they are generally issued by smart contracts rather than natural or legal persons. While Recital 26 suggests that some algorithmic stablecoins could be considered ARTs, their issuance could only be regulated under MiCA if it is centralised and conducted by a person, which is generally not the case. Where EMTs reference an EU currency, they are within scope of MiCA even if they are issued outside the EU.

Stablecoins which meet criteria measuring their systemic importance are deemed 'significant' stablecoins and their issuers are subject to enhanced requirements such as maintaining a higher amount of own funds and a liquidity management policy and procedures.

[49] cf Sokolov (n 45).

Further, significant stablecoin issuers are subject to enhanced supervision at the EU level by the European Banking Authority (EBA) and other members of a broad supervisory college. This is characteristic of the risk-based approach to regulation and supervision, where activities posing greater risks are subject to enhanced regulatory requirements and supervision. The risk-based approach is effective in optimising the use of limited regulatory resources, but regulatory discretion in the application of enhanced requirements and supervision should be clearly defined and limited to ensure legal certainty.[50] Moreover, while co-operation between regulatory authorities at the domestic and international level is key in tackling the risks of stablecoins[51], it remains to be seen how effectively the broad supervisory college outlined under MiCA could co- operate to address these novel risks.

Regulation of stablecoin issuers and related cryptoasset service providers (CASPs) under MiCA includes a swathe of prudential, conduct and organisational requirements based on existing EU banking, payments and e-money regulations. Nonetheless, MiCA is a bespokecryptoassets regime, which indicates that policymakers agreed it was inappropriate to subject cryptoassets to existing financial regulation without modifications given the unique features of cryptoassets compared to similar financial products and services.[52] These unique features of cryptoassets give rise to different and sometimes greater risks, and MiCA noted that EMD2 and the Second Payment Services Directive (PSD2) may be insufficient to protect consumers from significant risks around cryptoassets.[53] Notably, 'significant' stablecoins could pose significant or systemic risks, hence the application of enhanced requirements and supervision to their issuers and related service providers.

Under MiCA, EMT and ART issuers must be authorised by the competent authority in their home state, be an EU-based legal entity, and publish a 'white paper' including various disclosures around the stablecoin and issuer - similar to the prospectus required for securities issuers under the EU Prospectus Regulation. As under the Prospectus Regulation, MiCA ensures full regulatory harmonisation across the EU as the Regulation is directly applicable in each Member State. In contrast, EU Directives require national implementation that can differ across

[50] cf Ferreira (n 44).
[51] cf FSB (n 24).
[52] cf Chiu (n 4) 104.
[53] MiCA page 8.

the Union and result in regulatory arbitrage and barriers to entry into each Member State's market.

Transparency and disclosure requirements for issuers are a key aspect of MiCA supporting confidence in stablecoins by providing assurances to the market, most importantly assuring that a stablecoin's reserve assets are sufficient to meet redemptions. Indeed, an investigation by the New York Attorney General found that Tether Holdings Limited (Tether), the issuer of the USDT stablecoin, had misrepresented the status of its reserve assets by falsely claiming that USDT was 100% backed by US dollars.[54] The settlement agreement comprised a fine of $18.5 million and temporary disclosure requirements imposed on Tether, and Tether has since published independent assurances of its reserves detailing their composition and attesting that their value is greater than the value of outstanding USDT.[55] However, many industry participants call for greater transparency, and some even accuse Tether of large-scale fraud. This paper argues that standardised transparency and disclosure requirements such as those under MiCA will be key in addressing market participants' reservations around the reserves held by stablecoin issuers.

Additionally, to maintain sufficient collateral and liquidity to fulfil redemptions, MiCA requires issuers to ensure that their outstanding issuance of stablecoins is always 100% collateralised by reserve assets or funds which are segregated and safeguarded and can only be invested in liquid low-risk assets, and issuers must maintain an additional capital buffer of their own funds. Further, similar to e-money issuers, EMT issuers must provide holders with a claim on the issuer and redeem EMTs on demand at par value. While ART issuers are not generally required to provide holders with a claim or redemption rights, MiCA provides that holders must have the right to redeem ARTs directly from the issuer if the ART's market value differs significantly from the value of the reference assets or reserve assets. These measures aim to ensure that stablecoins can always be redeemed at their nominal value, even where this differs from their market value or where there are crises of confidence or market sell-offs.

[54] New York Attorney General, 'Attorney General James Ends Virtual Currency Trading Platform Bitfinex's Illegal Activities in New York' (*NY Attorney General*, 2021) <https://ag.ny.gov/press-release/2021/attorney-general-james-ends-virtual-currency-trading-platform-bitfinexs-illegal> accessed 19 February 2022.
[55] Tether, 'Tether Assurance Consolidated Reserves Report' (*Tether*, 2021) <https://assets.ctfassets.net/vyse88cgwfbl/01lZdtaNYx7jZ4jU5xmlYO/90aa0d5b1e3559c393ff135f987ddbd0/tether-assurance-sept-30-2021.pdf> accessed 19 February 2022.

Stablecoins are often compared to money market funds (MMFs), which are mutual funds invested in liquid low-risk assets and aiming to maintain a stable value and enable par-value on-demand redemption, and which were subject to runs in 2008 and 2020.[56] Regulatory proposals in the EU, the UK and the US clearly aim to prevent a stablecoin run, where issuers are unable to meet large-scale redemptions over a short time, particularly on a global stablecoin which could have systemic contagion effects. The primary risk of a run is that where there is large-scale liquidation of reserve assets in order to fulfil redemptions, this can cause a 'fire sale' of assets driving prices downwards to the point where the value of reserve assets falls below that of the outstanding stablecoins, so not all stablecoins could be redeemed at their nominal value.

Issuers can mitigate this risk by reserving the right to delay or apply fees to redemptions, though MiCA requires the time period for redemptions to be stated in the white paper and be no greater than 30 days. To mitigate the risk of a bank run, banks around the world are subject to prudential requirements including minimum capital and liquidity requirements under national frameworks and their implementations of the Basel framework (international standards for prudential regulation of banks set by the Basel Committee on Banking Supervision (BCBS)). Maintaining minimum amounts of capital and liquidity ensures that banks can meet withdrawals even in stressed conditions, either by drawing on liquidity stock or by liquidating capital.Similarly, it is submitted that by imposing minimum capital and liquidity requirements onstablecoin issuers, MiCA reduces the risk of issuers being unable to meet redemptions of stablecoins even in stressed conditions.

Table 3 below provides a non-exhaustive list of the types of requirements MiCA applies to stablecoin issuers and CASPs, and those applicable to EMT issuers under EMD2. Article references are to provisions under MiCA unless otherwise stated.

[56] Gary Gorton and Jeffery Zhang, 'Taming Wildcat Stablecoins' (2021) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3888752> accessed 19 February 2022.

Table 3 – MiCA Requirements

Issuers of ARTs	Issuers of EMTs	Crypto-asset service providers
Authorisation (Articles 15-16)	Authorisation (Article 43)	Authorisation (Articles 53 and 54)
White paper and attached liability (Articles 17 and 22)	White paper and attached liability (Articles 46 and 47)	Prudential safeguards (Article 60)
Conduct requirements (Article 23)	Issuance and redeemability (Article 44)	Conduct requirements (Article 59)
Marketing communications (Article 25)	Marketing communications (Article 48)	Organisational requirements (Article 61)
Ongoing disclosures (Article 26)	Prohibition of interest (Article 45)	Safeguarding clients' crypto-assets and funds (Article 63)
Complaint handling procedure (Article 27)	Investment of funds (Article 49)	Complaint handling procedure (Article 64)
Conflicts of interest (Article 28)	Compliance with EMD2 Titles II and III unless otherwise stated (Article 43)	Conflicts of interest (Article 65)
Notification of management changes (Article 29)	Prohibition from issuing e-money except for EMIs (EMD2 Article 10)	Outsourcing (Article 66)
Governance arrangements (Article 30)	Permitted activities (EMD2 Article 6)	
Own funds (Article 31)	Own funds (EMD2 Article 5)	
Reserve assets (Article 32)	Initial capital (EMD2 Article 4)	
Custody of reserve assets (Article 33)	Safeguarding (EMD2 Article 7)	
Investment of reserve assets (Article 34)	General prudential rules (EMD2 Article 3)	
Prohibition of interest (Article 36)	Prohibition of interest (EMD2 Article 12)	
Disclosure of holders' claim or rights (Article 35)		
Orderly wind-down procedure (Article 42)		
Additional requirements for significant ART issuers (Article 41)	**Additional requirements for significant EMT issuers (Article 52)**	**Additional requirements for specific cryptoasset services**
Remuneration policy (Article 41)	Requirements on custody and investment of reserve assets (Articles 33 and 34) instead of EMD2 Article 7	Custody and administration (Article 67): - Custody policy - Liability for loss
Interoperability (Article 41)	Remuneration policy, interoperability, liquidity management policy and procedures (Article 41 paras 1-3)	Exchange of cryptoassets (Article 69): - Commercial policy - Publication of order and transaction details
Liquidity management policy and procedures (Article 41)	Orderly wind-down procedure (Article 42)	Execution of cryptoasset orders (Article 70): - Best execution
Higher own funds requirement (Article 41)	Higher own funds requirement (Article 41(4)) instead of EMD2 Article 5	Reception and transmission of orders (Article 72): - Procedures

MiCA is one of four legislative proposals in the EU's digital finance package, an initiative aimed at enhancing digital financial services in the EU and supporting responsible innovation.[57] Of the other proposals, noteworthy provisions include requirements under theproposed regulation on digital operational resilience that all EMT and ART issuers and CASPs must implement an ICT risk management framework which ensures robust operational resilience. This includes implementation of 'strong authentication mechanisms … to prevent access to cryptographic keys' (Article 8(4)(d)), which is critical to preventing theft of cryptoassets. Also, under the proposed pilot regime regulation (COM(2020) 594), financialmarket infrastructures based on distributed ledger technology would be permitted to settle payments in e-money tokens, which may facilitate innovation and institutional adoption ofstablecoins.

In addition to the digital finance package, a new Eurosystem oversight framework for electronic payment instruments, schemes and arrangements ('PISA framework') was recently approved. The PISA framework implements most of the PFMI (with some exceptions and differences) and applies to stablecoins and stablecoin arrangements.[58] As the PFMI are key principles supporting the resilience of global financial market infrastructures, the PISA framework is likely to reduce risks around stablecoins and stablecoin arrangements, particularly operational risks and financial risks.

b. United Kingdom (UK)

Perhaps spurred by the EU's 2020 MiCA proposal, the UK published similar proposals in 2021 for subjecting most stablecoins and stablecoin arrangements to financial services regulation. The UK proposals draw on existing regulation including the E-Money Regulations 2011 (EMRs), the Payment Services Regulations 2017 (PSRs), the Banking Act 2009, and the Financial Services (Banking Reform) Act 2013, with amendments where necessary to account for the idiosyncrasies of stablecoins. Currently, the UK's regulatory approach to cryptoassets is based on 3 classifications of cryptoassets outlined in the FCA's guidance (2019):

[57] European Commission, 'Digital finance package' (*EC*, 2020) <https://ec.europa.eu/info/publications/200924-digital-finance-proposals_en> accessed 19 February 2022.
[58] ECB, 'Eurosystem publishes new framework for overseeing electronic payments' (*ECB*, 2021) <https://www.ecb.europa.eu/press/pr/date/2021/html/ecb.pr211122~381857cdfe.en.html> accessed 19 February 2022.

1) security tokens (cryptoassets similar to specified investments under the Regulated Activities Order 2001 (RAO));
2) e-money tokens (cryptoassets that qualify as e-money under the EMRs); and
3) unregulated tokens (all other cryptoassets including exchange tokens such as Bitcoin and utility tokens such as Link).

As these classifications clearly define the regulatory perimeter, this approach helps to avoid over-regulation and the stifling of innovation.[59] However, in regulating only two categories of cryptoassets, this approach has led to under-regulation given that stablecoins usually fall outside of UK financial regulation despite involving similar activities and risks to traditional financial services. Stablecoins are most likely to qualify as either (regulated) e-money tokens or (unregulated) exchange tokens according to the UK Treasury[60], but since most major stablecoins would not qualify as e-money under the EMRs[61], they likely would instead be unregulated exchange tokens.

Nonetheless, various regulations apply to those who provide services related to unregulated cryptoassets. For example, cryptoasset exchanges and custodian wallet providers are regulated under the amended 2017 Money Laundering Regulations, and regulations around data protection and advertising are applicable to activities involving 'unregulated tokens'. Also,payment services relating to funds may be regulated under the PSRs even if they are facilitated by unregulated cryptoassets.[62] However, these regulations cannot sufficiently mitigate the range of risks posed by stablecoins and stablecoin arrangements.

To bring most stablecoins and related activities within the regulatory perimeter andaddress associated risks, including risks to financial stability, market integrity, consumers and competition, the UK government proposed to create a fourth classification of cryptoassets called 'stable tokens'. Stable tokens are defined as tokens that aim to maintain a stable value by referencing a fiat currency or other assets, and as this definition applies regardless of theunderlying technology, it may remain applicable to future technological developments.[63] As the government is seeking to regulate 'tokens which could be reliably used for retail or wholesale

[59] cf Chiu (n 4) 91.
[60] cf HMT (n 30) 5.
[61] cf Sokolov (n 45).
[62] ibid 18; FCA, 'Guidance on Cryptoassets' (*FCA*, 2019) 39 <https://www.fca.org.uk/publication/policy/ps19-22. pdf> accessed 19 February 2022.
[63] cf HMT (n 30) 6-7.

transactions', only 'single-fiat stable tokens' (linked to a single fiat currency) and 'other asset-linked stable tokens' (those linked to other assets or multiple currencies) are within scope of the proposals.

Algorithmic stablecoins are outside of scope as they are deemed most similar to unbacked (and unregulated) exchange tokens. Decentralised on-chain collateralised stablecoins are not explicitly mentioned, but are likely outside of scope where issuance is decentralised asthe primary target of regulation (the issuer) likely cannot be subject to regulatory requirements, particularly if issuance is conducted by a smart contract. Exemption of such decentralised stablecoins is also supported by the fact that '[t]he government does not currently propose to bring specific DeFi [Decentralised Finance] activities into the scope of regulation'.[64]

The proposals give a high-level overview of the requirements likely to apply to different activities and propose an approach to cryptoasset regulation where requirements would be determined by the relevant regulators themselves according to the objectives and principles set by HMT and the government. These relevant regulators are the Financial Conduct Authority (FCA), the Bank of England's Prudential Regulation Authority (PRA) and the Payment Systems Regulator (PSR). Similar to the objectives of the FCA and PRA, the proposed objectives are to ensure financial stability, market integrity and consumer protection, and promote competition, innovation and UK competitiveness. Notably, the proposed principles for the regulatory approach include applying the 'same risk, same regulatory outcome' principle, following a risk- based and proportionate approach, and ensuring a flexible and agile approach with international regulatory harmonisation where possible.

Under the proposals, issuers of single-fiat or asset-linked stable tokens would be subject to an FCA authorisation regime and various operational, prudential and conduct requirements.[65] Single-fiat token issuers would be subject to the EMRs and PSRs with some modifications and additions as necessary, similar to the treatment of e-money tokens under MiCA. Businesses providing stable token custody and administration, execution or exchange services would be subject to FCA authorisation and all the requirements applicable to asset-linked stablecoinissuers except for requirements on operational resilience and maintenance and management of reserve assets.

[64] ibid 34.
[65] ibid 26-29.

Where systemic, i.e. where disruption of service could pose risks to financial stability, all listed activities in relation to in-scope stable tokens would be subject to the PFMI with modifications and additions as necessary. Further, where systemic, stable token service providers and payment systems would be subject to regulation by the Bank of England (BoE) under the Banking Act 2009, and by the PSR under the Financial Services (Banking Reform) Act 2013. Indeed, the BoE recently published a discussion paper on potential regulatory approaches to new forms of digital money, focused on the context of a GBP-denominated, retail-focused stablecoin which is systemic or could quickly become systemic.[66] The BoE proposes four potential regulatory models for systemic stablecoins[67]:

1) Bank model: stablecoin operating as a bank and backed by a broad range of assets;
2) HQLA model: stablecoin backed by a range of assets including high-quality liquid assets (HQLAs) and reserves;
3) Central Bank Liability (CBL) model: stablecoin backed by central bank reserves, often referred to as a synthetic central bank digital currency (CBDC); and
4) Deposit Backed (DB) model: stablecoin backed by commercial bank deposits.

The paper examines how the proposed models would meet the BoE's Financial Policy Committee (FPC) expectations regarding systemic stablecoin regulation, which 'aim to ensure the same level of public confidence in stablecoins as commercial bank money', confidence which underpins the BoE's objectives of achieving monetary and financial stability.[68] The FPC expectations are twofold[69]:

1) 'Payment chains that use stablecoins should be regulated to standards equivalent to those applied to traditional payment chains. Firms in stablecoin-based systemic payment chains that are critical to their functioning should be regulated accordingly'.
2) 'Where stablecoins are used in systemic payment chains as money-like instruments they should meet standards equivalent to those expected of commercial bank money in relation to stability of value, robustness of legal claim and the ability to redeem at par in fiat'.

[66] cf BoE (n 31).
[67] ibid.
[68] ibid.
[69] Bank of England, 'Financial Stability Report December 2019' (*BoE*, 2019) 82 <https://www.bankofengland.co.uk/-/media/boe/files/financial-stability-report/2019/december-2019.pdf?la=en&hash=4A650CF0FB871B5094C614C99689D9AD930CAA01> accessed 19 February 2022.

Key to both expectations is the notion of equivalent standards rather than identical standards, which underlies the overarching regulatory principle of 'same risk, same regulatory outcome' adopted by the UK government and regulators in their approach to cryptoasset regulation, as discussed further in section III.a. In stark contrast to the approach of 'same risk, same rules' taken by the EU[70] and the US[71], the 'same risk, same regulatory outcome' principle allows different regulation to be applied to cryptoassets posing the same or similar risks as comparable products and services, provided that the same regulatory outcomes are achieved. In any case, the BoE argue that the existing e-money regime applied to systemic single-fiat stablecoins would be insufficient to meet the FPC expectations, so enhancements would be necessary.[72] Nevertheless, the UK's regulatory proposals are still largely based on existing financial regulation.

Notably, the BoE argue that the FPC's second expectation would require stablecoin regulation to include core features of banking regulation: a legal claim, capital requirements, liquidity requirements and support, and a backstop for depositors.[73] The BoE note that the primary differences in their four proposed models is in their backing assets, which must always cover 100% of the value of outstanding coins unless the stablecoin issuer is operating as a bank. However, the BoE notes that under the CBL model 'the backing assets are lower risk, so the backstop elements could reflect this', seeming to suggest that lower risks could justify relaxed requirements in other areas.[74]

Such a risk-based approach would be desirable in better accounting for the idiosyncrasies of different business models or regulatory models, allowing greater flexibility in regulation as long as the same outcomes are achieved, true to the principle of 'same risk, same regulatory outcome'. The regulatory approach to cryptoassets should, however, reflect this principle at an even higher level. For example, instead of designing regulation of systemic stablecoins inaccordance with the FPC expectations, the focus should be on achieving the underlying regulatory objective of the expectations: ensuring the same level of public confidence in stablecoins as commercial bank money. The BoE's proposed regulatory models would likely meet the FPC expectations, but it is questionable whether all the core features of banking

[70] cf European Commission (n 57).
[71] US Treasury, 'President's Working Group on Financial Markets Releases Statement on Key Regulatory and Supervisory Issues Relevant to Certain Stablecoins' (*US Treasury*, 2020) <https://home.treasury.gov/news/press-releases/sm1223> accessed 19 February 2022.
[72] cf BoE (n 31) 61.
[73] ibid 57.
[74] ibid 60.

regulation would be necessary to achieve the objectives of stablecoin regulation, both of public confidence and broader objectives. In particular, if some aspects of a stablecoin business model or regulatory model pose demonstrably lower risks than others, they may warrant the application of commensurably less restrictive regulation in other areas, provided that the regulatory model asa whole achieves the intended outcomes.

A key reason for public confidence in commercial bank money is that demand deposits are usually protected by deposit insurance, which in the UK is up to £85,000 per individual per institution under the Financial Services Compensation Scheme (FSCS). The BoE asserts that a backstop of measures, which could include the FSCS deposit guarantee scheme, is necessary to ensure stablecoin holders do not lose their funds. While FSCS protection of stablecoins could contribute to similar public confidence in stablecoins as in commercial bank money, this may be difficult to implement. FSCS protection of bank deposits is funded by industry levies, which may be an infeasible model given that there are currently only a small number of large stablecoins, whereas including stablecoins in the same class as banks for FSCS protection may be unfair for banks and unaffordable for stablecoin issuers.[75] The same issue applies in the EU and US where deposit guarantees are also industry-funded.

Equally, a publicly-funded backstop for stablecoins may be difficult to justify for at least two reasons. Firstly, financial regulation following the 2008 financial crisis aims to prevent state-funded bail-outs of financial sector firms.[76] Secondly, although public backstops include other measures such as access to central bank liquidity, these are justified by the key role banks play in financial intermediation.[77] Liquidity support could be justified for issuers operating as banks and engaging in financial intermediation under the Bank model, but under the HQLA or DB model this may be more difficult to justify in absence of financial intermediation. Therefore, given that the BoE maintains that systemic stablecoin regulation must be supported by central bank liquidity for issuers, it appears the Bank or CBL model may be more easily justified according to the BoE's criteria. However, the equivalent outcomes of issuers having sufficient liquidity could be ensured without access to central bank liquidity. For example, as discussed in the context of US proposals below, an issuer holding reserve assets exclusively in low-risk liquid assets as

[75] ibid 69.
[76] ibid 69.
[77] Vincenzo Bavoso, 'Financial Intermediation in the Age of FinTech: P2P Lending and the Reinvention of Banking' (2021) Oxford Journal of Legal Studies.

under the BoE's proposed HQLA model would always meet the minimum liquidity coverage ratio under the Basel framework, and a capital buffer could provide an additional safeguard to ensure sufficient liquidity to meet redemptions.

Given the practical difficulties in implementing deposit protection for stablecoins, other backstops are likely to be necessary. One approach could be strict safeguarding requirements, perhaps including custodian liability for losses as under MiCA. While the FCA has recently emphasised the difference between e-money safeguarding and FSCS coverage[78], policymakers should consider whether strict safeguarding regulation can achieve equivalent outcomes to deposit protection, either alone or in combination with other requirements such as custodian liability for losses as under MiCA.

Of the four models proposed by the BoE, this paper favours the HQLA model as it would impose proportionate requirements allowing some flexibility in the stablecoin's backing assets while still ensuring that the issuer has sufficient capital and liquidity to meet redemptions. While the Bank model would allow even greater flexibility in the backing assets, with respect to stablecoin issuers that do not engage in lending, applying the full range of bank regulation to issuers would be disproportionate. The deposit-backed model would be feasible, but would likely increase interconnectedness in the financial system which may increase financial stability risks.[79]

Finally, the central bank liability model is workable, but would closely resemble a central bank digital currency (CBDC): a digital currency that is a liability of the central bank. As such, stablecoins backed by a central bank liability are often referred to as 'synthetic CBDCs', with the key difference being that they represent a liability of the issuer rather than of the central bank.[80] While synthetic CBDCs may well be an attractive option for central banks to leverage private sector innovation, such a regulatory model is not currently an appropriate generalised approach to regulation of stablecoins. This model has a complex range of implications including risks of disintermediating the commercial banking system and impeding monetary policy transmission. These implications and the merits of issuing CBDCs and synthetic CBDCs warrant detailed individual consideration before such a model can be implemented in practice.

[78] FCA, 'Business Plan 2021/22' (*FCA*, 2021) 29 <https://www.fca.org.uk/publications/business-plans/2021-22> accessed 19 February 2022.
[79] cf BoE (n 31) 11.
[80] ibid 60.

c. United States (US)

In the US, the future regulatory approach to stablecoins remains under debate, though some issuers and service providers already comply with federal regulation under the Bank Secrecy Act (BSA) and state regulation under money transmission laws. Much of the debate focuses around the extent to which stablecoins can be treated as deposits and issuers should be regulated as banks. In December 2020, lawmakers proposed the Stablecoin Tethering and Bank Licensing Enforcement (STABLE) Act, which would treat stablecoins as deposits and limit stablecoin issuance to insured depository institutions. Issuers would be required to obtain a banking charter and comply with applicable banking regulation. Stablecoins are defined broadly under the STABLE act as a 'cryptocurrency or other privately-issued digital financial instrument … with a fixed nominal redemption value', 'denominated in or pegged to' a fiat currency, and which eitheraims or succeeds at 'creating a reasonable expectation or belief' that the nominal redemption value will remain 'so stable as to render the nominal redemption value effectively fixed'. Stability according to 'reasonable expectation or belief' under the definition is subjective and cannot be measured.[81] The definition is, therefore, broad and likely applicable to all areasintended, but lacks legal certainty. Also, off-chain collateralised stablecoins are outside of scope, despite the fact that they are centralised and could be regulated, which would be warranted given the potential associated risks. While not explicitly mentioned, decentralised stablecoins would appear to be outside of scope as their reserve assets (or lack thereof) could not be held asdeposits.

Only weeks after the STABLE Act was proposed, the President's Working Group on Financial Markets (PWG) released a statement on key issues around US stablecoin regulation, which suggested that stablecoins could constitute securities, commodities or derivatives subject to existing federal laws.[82] The statement recommended that systemic stablecoin arrangements facilitating US retail payments should be designed in accordance with key principles that encompass objectives of financial regulation. Several recommendations for meeting these principles are provided, including that stablecoin arrangements should maintain 100% collateralisation in high quality assets in addition to a capital buffer.

[81] cf Ferreira (n 44).
[82] cf US Treasury (n 71).

While reserve asset requirements for stablecoins are likely necessary to ensure stability and redeemability, not all of the statement's recommendations would be necessary to achieve key regulatory objectives, and some appear overly strict. For example, the suggestion thatstablecoin arrangements should be able to verify the identities of all transacting parties is unrealistic, particularly given the practical difficulties in verifying ownership of private cryptoasset wallets. Any requirement to this effect would be overly burdensome or even impossible to comply with, and stretch beyond the recommendations of the FATF, an international body that sets globally-recognised policy standards for combating money laundering, terrorist financing and proliferation financing.[83] Specific requirements for stablecoinsare yet to be agreed, but many US policymakers, regulators and businesses agree that stablecoin regulation should include requirements around reserve assets, audit and transparency, illicitfinancing and sanctions, prohibitions of liquidity and maturity transformation, and operational risk.[84]

In November 2021, a report issued by the PWG, Federal Deposit Insurance Corporation (FDIC) and Office of the Comptroller of the Currency (OCC) recommended that Congress legislate to apply prudential requirements to stablecoin arrangements facilitating payments, essentially treating stablecoins like deposits.[85] Citing various risks around stablecoins used for payments, the agencies recommended that legislation impose requirements that issuers be insured depository institutions, apply federal oversight to custodial wallets and apply risk management standards for critical services, and restrict commercial affiliation with stablecoin issuers and custodial wallets.[86]

The functional similarities of stablecoins and issuers compared to deposits and banks are clear, hence the calls for equal regulatory treatment and the argument that 'issuers of stablecoins are essentially unregulated banks'.[87] However, stablecoin issuers are not banks and are not unregulated. Awrey notes that the definition of a bank under US federal banking law is circular,

[83] FATF, 'The FATF Recommendations' (*FATF*, 2021) <https://www.fatf-gafi.org/publications/fatfrecommendations/documents/fatf-recommendations.html> accessed 19 February 2022.
[84] Jai Massari (2021) Written Testimony for the Hearing on "Stablecoins: How Do They Work, How Are They Used, and What Are Their Risks?" Before the U.S. Senate Committee on Banking, Housing, and Urban Affairs <https://www.banking.senate.gov/imo/media/doc/Massari%20Testimony%2012-14-211.pdf> accessed 19 February 2022.
[85] cf US Treasury (n 27).
[86] ibid.
[87] cf Gorton and Zhang (n 56) 6.

and proposes a functional definition of a bank: 'any financial institution that combines lending with the creation of monetary liabilities'.[88] According to Awrey's definition, stablecoin issuers generally are not banks as they do not lend. Additionally, stablecoin issuers such as Circle are already regulated under state money transmission laws including permissible investment rules restricting investment of reserve assets, and US stablecoin issuers are subject to federal requirements under the BSA.

While stablecoin issuers that do not lend are not banks in the traditional sense, they resemble 'narrow banks': banks which back their demand deposits with 100% of the equivalent value in reserves in liquid low-risk assets and do not engage in lending.[89] In contrast, traditional fractional reserve banks engage in lending and back deposits with a high proportion of illiquid, long-term and higher-risk assets such as loans. To mitigate the risks of financial transformation and bank runs, banks across the world are subject to strict prudential regulation and oversight, including capital requirements such as minimum leverage ratios. Under the Basel framework, the leverage ratio is the ratio of tier 1 capital to total exposure, where banks must always maintain a minimum 3% leverage ratio.[90]

Given the lower risks relative to fractional reserve banking, stablecoin issuers should not be subject to all of the same regulatory requirements as fractional reserve banks which engage in lending. Where issuers invest reserve assets exclusively in liquid low-risk assets such as short-term US treasuries, and do not engage in lending, issuers should not be subject to a minimum leverage ratio designed to apply to more complex and higher risk financial institutions.[91] Moreover, stablecoin issuance should not be limited to insured depository institutions as the risk of runs can be mitigated by requirements of 100% collateralisation in short-term, low-risk liquid assets and a capital buffer, without the need for deposit insurance.[92]

However, for stablecoin issuers to qualify for relaxed capital requirements compared to banks, they should be subject to strict requirements on their reserve assets, which could include requiring 100% collateralisation in specific short-term, low-risk liquid assets and a capital buffer.

[88] Dan Awrey, 'Unbundling Banking, Money, and Payments' (2021) 7 ECGI Working Paper Series in Law <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3776739> accessed 19 February 2022.
[89] Bert Ely, 'The Narrow Bank' (1991) Cato Review of Business and Government.
[90] BIS, 'Basel III leverage ratio framework' (*BIS*, 2017) <https://www.bis.org/fsi/fsisummaries/b3_lrf.htm> accessed 19 February 2022.
[91] Christian Catalini and Nihar Shah, 'Setting Standards for Stablecoin Reserves' (2021) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3970885> accessed 19 February 2022.
[92] cf Massari (n 84).

Additionally, issuers should be required to maintain a minimum 100% liquidity coverage ratio under the Basel framework to ensure a sufficient stock of HQLAs that can be liquidated to meet redemptions.[93] This ratio calculates the value of a firm's HQLAs against the value of net cash outflows during the next 30 days, and would be at least 100% where issuers invest exclusively in HQLAs.

Further, stablecoin issuers and service providers should still be subject to other aspects of financial regulation applicable to banks and e-money institutions, including prudential, conduct and operational requirements. This proposed regulatory approach would be similar to that adopted in the EU and the UK where stablecoins are treated like e-money or investment productsrather than deposits. While there are similarities between stablecoins and deposits and some regulators fear users would treat them as equivalent, regulatory requirements such as prohibition of interest could help distinguish between stablecoins and deposits, as under UK and EU e- money regulation.

As proposed by Catalini and Massari as a 'true stablecoin' model, issuers could ensure price stability and redeemability by committing to redeem stablecoins at par value, holding ringfenced reserves equivalent to 100% of the issuance value in HQLA, and holding a capital buffer.[94] This paper broadly supports this regulatory model, though the importance of an issuer's commitment to redemption of stablecoins is questioned below in section II.b, which argues that a legal claim on the issuer should not be a general requirement provided that measures are in place to ensure users can reliably redeem stablecoins.

Commentators have suggested that the risks of systemic stablecoins in the US can be mitigated if the FSOC designates their issuers as a systemically important financial institution which would be subject to prudential regulatory supervision by the Federal Reserve, even if the issuer is a nonbank financial institution.[95] Alternatively, the FSOC could designate the issuance of stablecoins as a systemic payment activity under the Dodd-Frank Act, which would empower

[93] cf Catalini and Shah (n 91); BIS, 'Liquidity Coverage Ratio (LCR)' (*BIS*, 2018) <https://www.bis.org/fsi/fsisummaries/lcr.htm> accessed 19 February 2022.
[94] Christian Catalini and Jai Massari, 'Stablecoins and the Future of Money' (*Harvard Business Review*, 2021) <https://hbr.org/2021/08/stablecoins-and-the-future-of-money> accessed 19 February 2022.
[95] cf Massad (n 14); Hilary Allen (2021) Written Testimony for the Hearing on "Stablecoins: How Do They Work, How Are They Used, and What Are Their Risks?" Before the U.S. Senate Committee on Banking, Housing, and Urban Affairs <https://www.banking.senate.gov/imo/media/doc/Allen%20Testimony%2012-14-211.pdf> accessed 19 February 2022; Steven Schwarcz, 'Regulating Global Stablecoins: A Model-Law Strategy' (2021) Duke Law School Public Law & Legal Theory Series.

the Federal Reserve to regulate stablecoin issuance of financial institutions.[96] As both approaches are only applicable to financial institutions, the broader regulatory model for stablecoin issuance would need to require issuers to be financial institutions, thereby restricting competition. FSOC designation of issuers may be a successful approach, but both the criteria for designation and the requirements applicable should be clearly defined and limited to prevent abuse of regulatory discretion. Designating all stablecoin issuance of financial institutions as a systemic payment activity, on the other hand, would be an overly inclusive and thus undesirable approach, as even the issuance of stablecoins with low market capitalisation and posing low risk would be deemed a systemic payment activity.

III. Normative Arguments

a. Regulatory Approaches

In evaluating how regulation should be designed, it is important to consider what the overarching regulatory approaches and principles should be. When discussing how stablecoins should be regulated, the 'same risk, same rules' principle has been proposed by authorities in the EU[97] and the US[98], as well as by international regulatory bodies such as the Bank for International Settlements[99] and the Financial Stability Board.[100] However, despite the functional similarities between stablecoins and payment instruments or bank deposits, and between stablecoin arrangements and payment systems, the technological idiosyncrasies of stablecoins mean some of the risks they pose are not the same.

As such, applying the principle of 'same risk, same rules' to stablecoin regulation is challenging and likely to be somewhat ineffective.[101] Instead, stablecoin regulation should be based on the principle of 'same risk, same regulatory outcome' as proposed by the UK government and regulators[102], as the flexibility of this principle would better support technological innovation around stablecoins while ensuring regulatory outcomes are achieved. Under this principle, regulation different to existing frameworks could be applied, which may include exemptions or looser requirements in certain areas, so long as the same regulatory

[96] cf Gorton and Zhang (n 56).
[97] cf European Commission (n 57).
[98] cf US Treasury (n 71).
[99] cf BIS (n 26).
[100] cf FSB (n 29).
[101] cf BoE (n 31); cf Tether (n 13) 3.
[102] cf BoE (n 31); cf HMT (n 30).

outcomes are achieved. For example, as discussed in section II.c, if stablecoin issuers are subject to strict requirements on reserve assets and do not engage in lending, they should not be subject to all of the same rules including the minimum leverage ratio applicable to relatively higher risk businesses such as fractional reserve banks.

Additionally, adopting a functional approach to regulation may help account for the unique technological characteristics of stablecoins and ensure regulation can adapt to future innovations.[103] Under the functional approach, regulation is designed and applied based on the function and purpose of the regulated activity. In contrast, current payment services regulations often apply based on rigid definitions outlining the scope of regulation, according to a definitional approach.[104] In supporting the functional approach to cryptoasset regulation, Chiu highlights that collateralised stablecoins have two key functions of 1) payments, which should beregulated similar to existing payment instruments and payment systems; and 2) reserve asset or investment management, which could be subject to existing financial regulation.[105] Collateralisedstablecoins should be regulated according to both their asset and payment functions, in contrast to the EU's approach under MiCA of categorising and regulating stablecoins based on either their asset or payment function.[106]

However, the approaches and principles that regulators may adopt are limited by factors including the existing legal and regulatory frameworks and the practical challenges in implementation. For example, the US financial regulatory framework is highly fragmented between sectors and regulatory agencies[107] under the US's sectoral approach where different agencies are responsible for regulating different sectors of the financial industry. This sectoral approach 'does not fully capture innovative characteristics, such as the multifunctional nature of stablecoins as both payment instrument and asset'[108] and thus reduces the feasibility of the functional approach to stablecoin regulation. Equally, a functional approach would stipulate that decentralised stablecoins and arrangements that serve payment or asset management functions should be regulated according to these functions. However, due to the practical challenges in

[103] cf Chiu (n 4) 208.
[104] James Burnie, Andrew Henderson and Andrew Burnie, 'Blockchain, Cryptocurrencies and How They Fit Within Current Payments Regulation' in Susanne Chishti et al. (eds), *The Paytech Book* (Wiley 2020) 114.
[105] cf Chiu (n 4) 212.
[106] ibid 205.
[107] cf Gorton and Zhang (n 56).
[108] cf Chiu (n 4) 99.

imposing regulatory requirements in this context (as discussed in section III.c), there are few viable options for regulating decentralised stablecoins and stablecoin arrangements.

b. Backing Assets

Regardless of the regulatory approach adopted, some of the key features that stablecoin regulation should include are requirements on the nature, quantity and management of reserve assets. In order to prevent the primary risk of a stablecoin issuer being unable to meet redemptions on demand and at par value, this paper proposes that centralised stablecoin issuers should be subject to restrictions on the types of reserve assets that can be held as well as minimum requirements on the amount of capital and liquidity that must be held. For tokenised funds, the reserve assets should be restricted to cash, cash equivalents, HQLAs and other low- risk liquid assets, in addition to a liquidity and capital buffer. Instead of regulating stablecoin issuers as banks, the same regulatory outcomes of protecting the price stability and redeemability of stablecoins could be achieved by requiring issuers to comply with the Basel framework (except for certain unsuitable requirements such as a minimum leverage ratio).[109]

Also, the effect of redeemability can be protected without requiring issuers to provide holders with a legal claim on the issuer or reserve assets. Whereas e-money is typically only used by holders who have a contractual relationship with the issuer, stablecoins are generally freely traded and used by many users who have no contractual relationship with the issuer. Rather than redeem stablecoins, users often simply trade them for cryptoassets, or even for fiat currency on a cryptoasset exchange, making redemption often unnecessary in realising approximately the par value of the stablecoin. At the same time, on-demand redemption at par value remains important in facilitating arbitrage and ensuring users can recover the nominal value of stablecoins. However, this paper argues that instead of requiring issuers to provide all holders with a legal claim, redeemability can be protected by regulatory requirements that issuers must have contractual arrangements with third parties who agree to facilitate redemption (as under MiCAfor ART issuers). Regulation should nonetheless require that holders have a legal claim on the issuer or reserve assets in exceptional circumstances where they are otherwise unable to recover the nominal value of their stablecoins, such as where the third-party contractors no longer fulfil redemption.

[109] cf Catalini and Shah (n 91).

c. Decentralised Stablecoins

Decentralised stablecoins, i.e. algorithmic stablecoins and decentralised on-chain collateralised stablecoins, appear to fall outside the scope of the EU, UK and US proposals for stablecoin regulation. In their current form, to the extent that these stablecoins are fully or mostly decentralised, this paper agrees that they should be outside the scope of current regulatory proposals for at least two reasons. Firstly, these three jurisdictions are yet to propose regulation for DeFi (decentralised financial services based on smart contracts), which should ideally be implemented in tandem with any approach to regulating decentralised stablecoins given their interconnectedness. Secondly, there are many practical challenges in applying traditional financial regulation to decentralised systems, so regulation of decentralised stablecoins and DeFi would likely require a bespoke regime separate to that proposed for centralised stablecoins.

Traditional financial regulation focuses on regulation of centralised intermediaries. There would be challenges in applying this regulatory approach to decentralised stablecoin arrangements due to the relative lack of intermediaries compared to in the traditional financial industry. While there may remain some centralised components of otherwise decentralised stablecoin arrangements, such as centralised exchanges and custodian wallet providers, these entities are already subject to anti-money laundering regulations in the EU, the UK and the US. Where the key functions of stablecoin issuance and governance are, respectively, carried out by an automated smart contract and a decentralised group of governance token holders, the applicability of regulatory requirements is limited.

Instead of targeting intermediaries, regulatory requirements could be imposed on software developers with respect to the smart contract code they create for stablecoin arrangements, but such an approach is difficult to justify as it could stifle innovation and even be criticised as restricting free speech.[110] Equally, regulation of a decentralised population of governance token holders would be burdensome on users and practically impossible to enforce. A successful approach to regulating decentralised stablecoins could involve co-regulation which fosters co-operation between regulatory authorities and the cryptoasset industry, with the aim of both supporting innovation and achieving regulatory objectives.[111] However, such a novel approach would likely need to be integrated in a bespoke regime following public consultations

[110] cf WEF (n 39) 22.
[111] cf Chiu (n 4); Michèle Finck, *Blockchain Regulation And Governance In Europe* (Cambridge University Press 2019).

and engagement, and the present regulatory focus should first be on regulating the centralised stablecoins and arrangements which currently dominate the market and may become systemic. Moreover, even where unregulated, the risks of decentralised stablecoins can be mitigated through measures such as consumer warnings and awareness campaigns.

Conclusion

Stablecoins may enable considerable improvements to financial services and payment systems, but they also pose risks which should be mitigated through regulation of the entities within stablecoin arrangements, particularly where they are systemically important. Recent EU, UK and US proposals for stablecoin regulation are all based on existing banking, e-money and payments regulations to varying degrees, which is justified given the functional similarities of stablecoins and stablecoin arrangements to existing components of the regulated financial sector.

However, there are limits to applying existing financial regulation to stablecoins due to their unique technological characteristics. As such, this paper argues that regulation of stablecoins should follow the principle of 'same risk, same regulatory outcome', where stablecoin activities posing the same risks as currently regulated activities could be subject to amended regulatory requirements as long as the same regulatory outcomes are achieved. This principle better supports both innovation and effective regulation by allowing a degree of flexibility in the regulatory approach which can accommodate the idiosyncrasies of stablecoins while ensuring the desired regulatory outcomes are achieved.

Notably, regulatory outcomes that stablecoin regulation should achieve include ensuring that issuers or their contractors can always meet redemptions of stablecoins on demand and at par value, even in stressed conditions. This paper broadly agrees with several of the EU and UK proposals for achieving this outcome via the application of existing regulations with modifications. Centralised stablecoin issuers should be required to fully collateralise their outstanding stablecoin issuance with specified liquid low-risk assets or the assets which the stablecoin references, and hold additional liquidity and capital buffers. Reserve assets and funds should be safeguarded and segregated (in a bankruptcy-remote manner).

Where issuers and relevant entities in stablecoin arrangements are subject to these requirements, other aspects of existing financial regulation may not need to be applied to achieve regulatory outcomes such as ensuring reliable redemption of stablecoins. For example, deposit

protection for stablecoins may not be necessary where strict safeguarding regulation applies, particularly if combined with other requirements such as custodian liability for losses as proposed under MiCA. As such, stablecoin issuance should not be limited to banks as was proposed by US lawmakers.[112] Moreover, stablecoin issuers that do not engage in lending shouldnot be subject to the same bank regulation as relatively higher-risk businesses such as fractional reserve banks that engage in lending.

With respect to systemic stablecoins and stablecoin arrangements, the enhanced regulation and supervision proposed by the EU and UK will likely mitigate the higher associated risks and is consistent with the pragmatic risk-based approach to regulation.[113] Indeed, the PFMI apply to systemic stablecoin arrangements[114], though modifications will likely be necessary to ensure their applicability. In regulating systemic stablecoins and stablecoin arrangements, regulators should clearly outline applicable requirements and criteria for designation as systemically important to ensure legal certainty and prevent abuse of regulatory discretion.

[112] cf Massari (n 84).
[113] cf BoE (n 31) 11.
[114] cf BIS (n 26) 4.

Discussion Paper 4

The proof of identity

Andrea Dalla Val, Human Income Coin - Proof of Identity Network

Abstract

Despite the recent development of a number of decentralised applications, blockchain technology is still lacking of proper fundamentals as currently used consensus protocols are poorly achieving the original goals at the base of its inception. In fact, Proof of Work is dependent on high amounts of electricity supply thus easily censorable, and Proof of Stake is undemocratic as one does not count one, and thus does not have a great appeal to small holders. The inspiration of this white paper comes from Prof. Bryan Ford and the EPFL, and their idea of proof of personhood which places individuals as the base of the minting system and creates the idea of pseudonymous parties; however, here we introduce the proof of identity which eliminates the need of multiple attendances to pseudonymous parties and creates a permanent proof. Another notable paper introducing the use of biometric identification, was proposed by Mohammad Javad Hajialikhani, Mohammad Mahdi Jahanara June 2018, Cornell University, however the proposed design relies on Ethereum blockchain, and P2P identification, in this paper instead, the methodology is based on identification parties, and additional verification parties randomly organised by a decentralised AI engine.
The Proof of Identity network is truly decentralised and democratic because based on a multitude of individuals, it is structured as income generator as opposed to an investment scheme, is non-quantitative, i.e. able to validate all blocks of all blockchains without significant additional efforts. The protocol is based on humanity parties that are fully managed by an AI engine stored in the blockchain (cohort leader servers). Humanity parties are complying with standards that ensure: “everyone or noone is cheating” i.e. there cannot be cheaters and honest participants attending the same party; parties are automatically formed and managed, no authority or entity hasany power or discretion over them, parties venues are locked and tamper sealed; “everyone or no one is cheating” is the foundation of a pyramid validation structure which ultimately requires to collude globally to avoid being discovered. The Proof of Identity network is permanently censorship resilient because based on the principles that attackers cannot possibly fake identities, established minters are unaffected by governmental bans, it is the democratic choice of a multitude of identified individuals, all prerogatives missed by the currently used proofs. Under these premises, any open source blockchain can be effortlessly validated and even forcibly (i.e. without any involvement of miners) hard forked into a duplicate network run under PolD. Imagine the slogan : “Bitcoin is hard forking into a 0 energy consumption network!”

Keywords

Blockchain Consensus Protocol, Proof of Identity, Blockchain Technology.

CBT

THEPROOF OF IDENTITY

ACONSENSUSPROTOCOLBASED ON PERMANENT PROOF & UNIFYING ALLBLOCKCHAINS

THE NETWORK RUN BY INDIVIDUALS NOT BY MACHINES

Written by: Andrea Dalla Val

3rd June 2021

Rev.C2

Table of Contents

PROOF OF IDENTITY: THE IDENTITY COIN.

1. EXECUTIVE SUMMARY

Despite the recent development of a number of decentralised applications, blockchain technology is still lacking of proper fundamentals as currently used consensus protocols are poorly achieving the original goals at the base of its inception.
In fact, Proof of Work is dependent on high amounts of electricity supply thus easily censorable, and Proof of Stake is undemocratic as one does not count one, and thus does not have a great appeal to small holders.

The inspiration of this white paper comes from Prof. Bryan Ford and the EPFL, and their idea of proof of personhood which places individuals as the base of the minting system and creates the idea of pseudonymous parties; however, here we introduce the proof of identity which eliminates the need of multiple attendances to pseudonymous parties and creates a permanent proof.

Another notable paper introducing the use of biometric identification, was proposed by MohammadJavad Hajialikhani, MohammadMahdi Jahanara June 2018, Cornell University, however the proposed design relies on Ethereum blockchain, and P2P identification, in this paper instead, the methodology is based on identification parties, and additional verification parties randomly organised by a decentralised AI engine.

The Proof of Identity network is truly decentralised and democratic because based on a multitude of individuals, it is structured as income generator as opposed to an investment scheme, is non-quantitative, i.e. able to validate all blocks of all blockchains without significant additional efforts.

As many countries (Sweden, Norway etc.) are developing reliable and public electronic ID, identifications may be carried out entirely online, but in general, the protocol is based on humanity parties that are fully managed by an AI engine stored in the blockchain (cohort leader servers).
Humanity parties are complying with standards that ensure:
- ''everyone or none is cheating'' i.e. there cannot be cheaters and honest participants attending the same party

- parties are automatically formed and managed, no authority or entity has any power or discretion over them, parties venues are locked and tamper sealed
- ''everyone or none is cheating'' is the foundation of a pyramid validation structure which ultimately requires cheaters to collude globally to avoid being discovered.

The Proof of Identity network is permanently censorship resilient because based on the principles that attackers cannot possibly fake identities, established minters are unaffected by governmental bans, it is the democratic choice of a multitude of identified individuals, all prerogatives missed by the currently used proofs.
Under these premises, any open source blockchain can be effortlessly validated and even forcibly (i.e. without any involvement of miners) hard forked into a duplicate network run under the PoID.
Imagine the slogan: ''Bitcoin is hard forking into a 0 energy consumption network''!

2. PROOF OF IDENTITY, INTRODUCTION

Cryptocurrencies have been so far conceived with the aim to achieve fully distributed and censorship resilient networks, however, it has recently become clear to many that these goals are not achievable under the Proof of Work protocol, simply because the amount of wasted energy required to protect the network is just too high for it to come from non-governmental providers.

Bitcoin is the clear example of that, at a mere market capitalisation of USD 1T, it needs the same amount of energy as Argentina to secure the network against Sybil attacks, a waste that is not only incompatible with the resources of the globe but also it depends on governmental agencies approval, i.e. easily censorable.
The necessary (but not sufficient) condition for a PoW network to be secure is:

gain from double-spending < 51% of the block hash rate cost

otherwise, a 51% attack is economically attractive.

Being network security tied to energy consumption, PoW cannot achieve at the same time the 3 basic requirements of a globally distributed cryptocurrency: significant market capitalisation, network security, censorship resilience (Fig. 1).

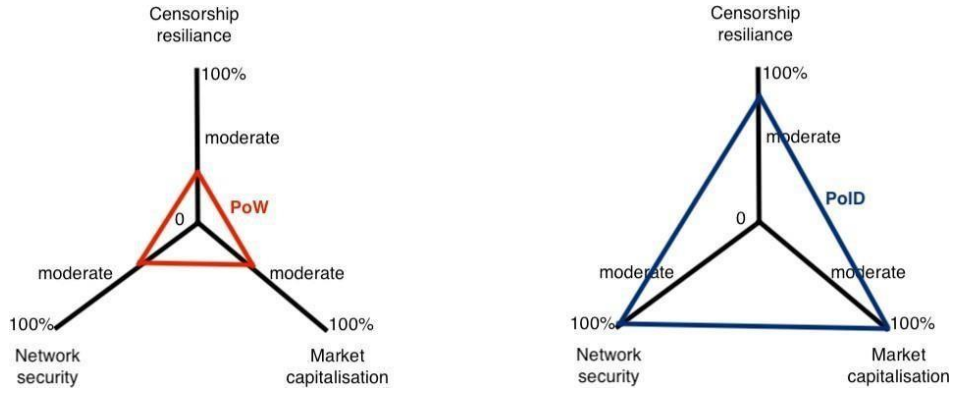


Fig. 1: The 3 fundamentals of a cryptocurrency in PoW and PoID.

Another flop of Proof of Work blockchains is that unlike what is commonly believed, PoW fails to achieve its first reason to exist: the goal of decentralisation.
Bitcoin network was initially thought of as householders connecting from their homes,in reality, 10 mining pools cover 90% of the PoW network (Fig.2), they are often forced to relocate their facilities in the quest of a bigger amount of energy at low cost;

eventually, the intrinsic dependency of PoW networks from governments will disprove the idea they are decentralised, distributed and censorship resilient.

Proof of Stake networks boast more security and independence from government agencies as they do not waste electricity, but unfortunately, they propose to individuals a bad deal.

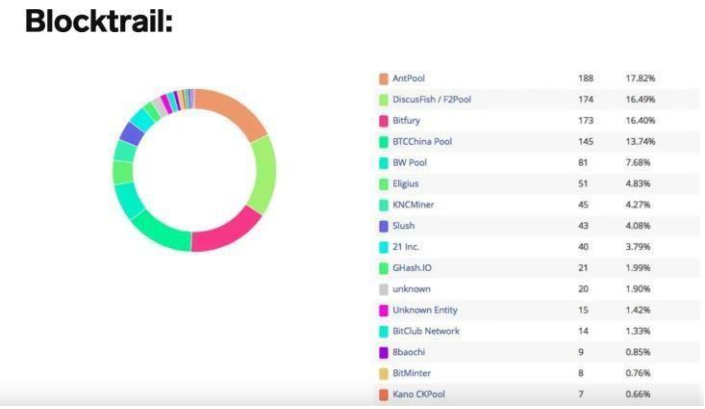


Fig.2: Mining farms worldwide, Source Blocktrail

The Filecoin is a good example of that: here, nodes of the network are storing computers, and anyone can join the network with his own computer and offer storage to the global co-shared iCloud; yet it is easy to see big iCloud corporations having great advantages when both running machines on a higher scale and acquiring larger network governance.

As the number of participants increases, the benefits to individuals running nodes in a homemade fashion, shrink to little or nothing, meanwhile big holders prosper from the increased size of the consolidated market and the gained control over the newly created cryptocurrency, benefits that mostly came from the participation of the single individuals. Ultimately PoS networks are unattractive to individuals because there is little incentive for smallholders.

Fig.3: PoS, the bigger benefits from the smaller

Because the problems with PoW and PoS are inherent to their design and hence unsolvable, a different type of proof is required to attain in full those goals inspired at the inception of blockchain technology and clearly the only possible option is a proof based on individuals.

In this white paper we propose a network made of identified individuals, and create the Proof of Identity network (PoID).

A network made of identified individuals is truly decentralised and un-censorable, it can be scaled up to any market capitalisation while providing high transaction output (Fig. 1), it does not waste electricity, it distributes wealth democratically to individuals, not to capital holders or machines owners/energy providers.

Most interestingly, PoW and PoS are quantitive proofs, i.e. the used amount of work or stake is a function of the total capitalisation of the connected blockchain and its security; there would be no synergies in the aggregation of PoW or PoS networks.

Fig. 4: Both PoW and PoS are quantitative proofs, the bigger is the size of the network, the more is the proof required to secure the network.

Conversely, the Proof of Identity network (PoID) is non-quantitative proof, its security is based on the impossibility for an attacker to fake identities, in reflection to that, a relatively small number of nodes/network participants makes up already a safe network.

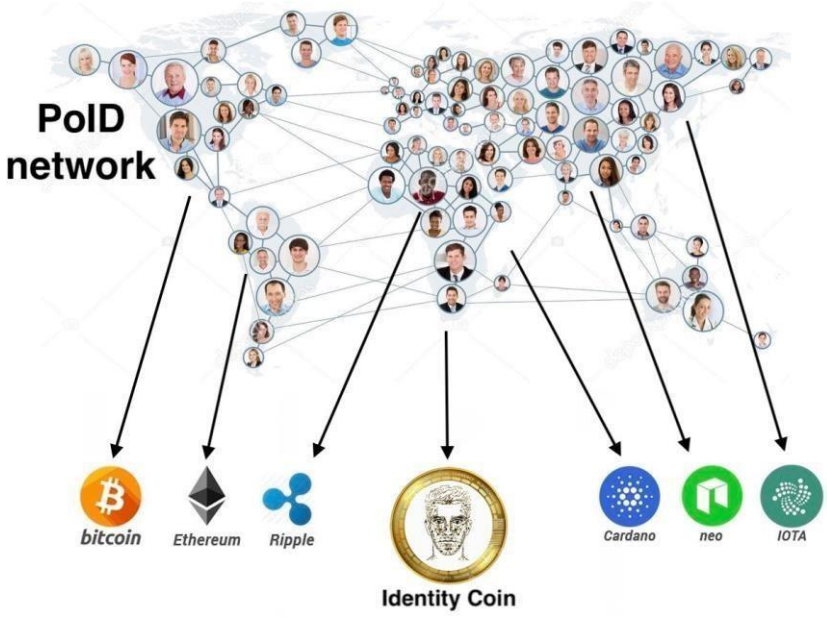


Fig. 5: PoID is non-quantitative proof, it can validate blocks of any blockchain with no additional effort required.

Once PoID achieves some global diffusion, the number of nodes/cohorts becomes by far more than what is needed to validate all blockchains at the same time, potentially

offering a unified ecosystem in which all blockchains can focus on their own features while having their blocks validated by the PoID (Fig.5).

This capacity suggests the interesting idea of cloning other blockchains (like Bitcoin for example) and globally replace PoW with PoID with forced hard-forks and without the involvement of miners (see chap. 10).

3. PROOF OF IDENTITY BACKGROUND

In the paper: Proof-of-Personhood: Redemocratizing Permissionless Cryptocurrencies (https://bford.info/pub/dec/pop.pdf)
the idea of Proof of Personhood has been introduced by Professor Bryan Ford and the EPFL, and this white paper is crediting the idea that proof of genuine participation to a permission-less network has to be based on individuals, not on machines or stakes.

This white paper embraces in full the technical architecture designed by the EPFL, the use of the technologies RandHound, ByzCoin, Linkable Ring Signatures, Collective Signing with the only difference being that whilst the Proof of Personhood envisages individuals receiving PoP tokens by attending recurrent pseudonym parties, we believe that individuals should receive their everlasting PoID minting token through identification.

Proof of Personhood has certainly great merits, however, is a difficult and discouraging process to implement.
Pseudonym parties have to be repeated many times and simultaneously organised in many places to cover a good size of the population and prevent individuals from getting multiple minting tokens.
Synchronisation and coordination are very challenging in the PoP protocol: after a first successful round of pseudonymous parties, a new blockchain to adopt the PoP protocol would have to wait until all other blockchains repeat their pseudonymous parties to synchronise with them and avoid the chance of multiple attendances.

It is hard to figure out the logistics of the parties in big urban areas like Paris or London where the traffic is already congested, and not many places can contain hundreds of people at the same time.

If attendance is free, an adversary may ruffle the estimated size of the party, if it is payable, more difficulty is added to the PoP process. The possibility for an attacker to create a multitude of pseudonymous parties with big number of attendees may be an unsolved concern for the PoP network which does not seem to prevent the case of a party fully attended by faked personhood.

As the network grows, recurrent pseudonymous parties become more visible to authorities and thus more exposed to censorship, it is easy to imagine how a governmental crackdown can terminate a PoP network very quickly.

Furthermore, the need of repeating parties within uncertainty may frustrate participants from the idea of building a system stable in time: every epoch can be better or worse than the previous one, PoP networks do not guarantee any consistent growth, it is ultimately unclear whether a permanent value is actually constructed.

In this white paper, we believe that the problem of the Proof of Personhood network can be solved with the Proof of Identity, a protocol in which parties are not pseudonymous and therefore do not require individuals to repeatedly attend them.

Another important proposal in the field is the proof of UniqueID, published in June 2018 by MohammadJavad Hajialikhani, MohammadMahdi Jahanara (https://arxiv.org/abs/1806.07583) Cornell University. The protocol introduces the use of biometric identification performed P2P in person with some verifiers. The method relies on Ethereum blockchain for the execution of a set of smart contracts, the system also include the use of Captchas solving, and stores biometric data in a trusted setup.

The main disadvantage of this design is its exposure to collusions among verifiers, particularly in remote areas of the planet not well connected, and the lack of any supervision over possible cheats.
Moreover, the reliance on the Ethereum blockchain ties down the security of the system to the security of the Ethereum network.

The proof of identity network introduces a new design which significantly improves the problems with the proof of UniqueID and creates a network collusion proofed and self detecting any possible cheat.

4. POID: STANDARD IDENTIFICATION PROCEDURE

In the PoID network, identification means negative identification, the evidence establishing the 2 conditions:

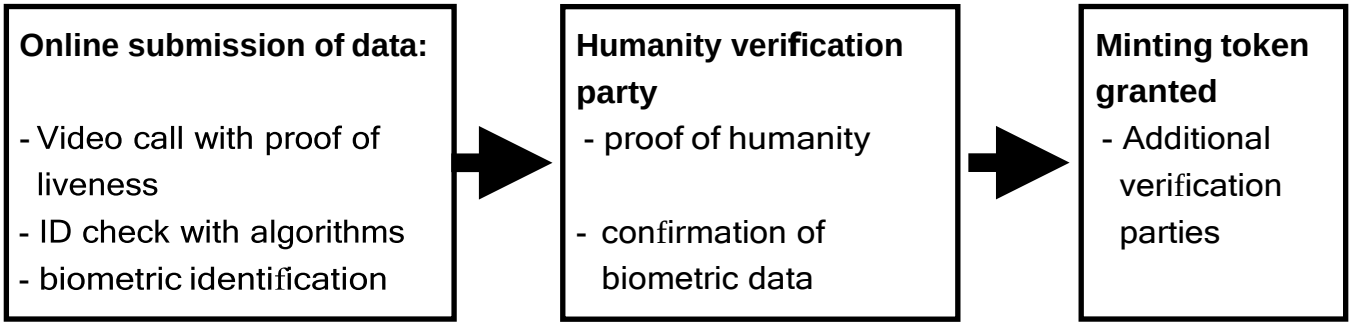
1. **the applicant is a human**

and

2. **is not among the group of people already known to the system.**

Once an individual has been identified, he receives a PoID minting token which enables his mobile app to join a cohort of minters and take part in the voting system of the PoID network / Identity Coin and therefore to receive identity coins without any other effort required.

The identification procedure is summarised in 3 phases:



4.1 ONLINE SUBMISSION OF DATA

In this phase face biometric data of the individual are taken and stored in the PGO central server, the application is also recorded in the blockchain.

Biometric identification is performed using face recognition technologies 3D mapping.

According to software leading companies, facial identification 1:N is secure over 99.999% against each single spoof attempt, whereas the word ''single'' means the method can not be breached in serial mode.

The technology is based on the concept that some

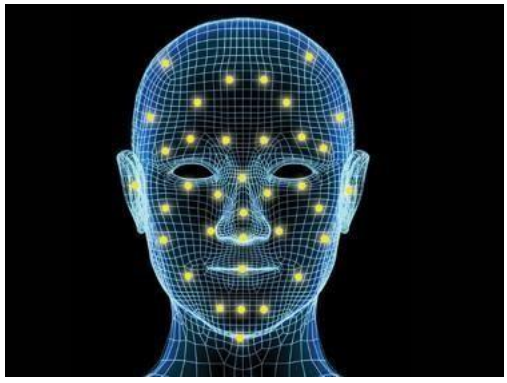


Fig. 6: examples of biometrics technology, with 3D face recognition

characteristics of the human face such as the distance between eyes, the geometry of the bones, do not change in time and cannot be changed by surgeries.
Algorithms are used to compute the probability two records belong to the same individual.

Proof of liveness and ID check with algorithms are also included in the procedure as strong discouragements to cheaters: knowing that humanity verification will disprove a video spoof, it would be pointless even to try the cheat, which is something requiring significant investment in AI video spoofing technologies.

Proof of liveness video-based technology is currently considered 99.6% secure (if vocal verification is added the odds increase even further), however here the odds are against a serial breaching of the system therefore it should not be used to validate condition 1 of the identification.

Proof of liveness algorithms are hosted in our central server, they are performed in the background and therefore the attacker would have to carry out several attempts only to work out the basic functioning of it.
Proof of liveness technology may include several checks assessing the face moves of a person against his selfie, the lighting, the movement of the eyes when following a dot on the screen (Fig. 7), the responsiveness of certain areas of the face while talking etc.

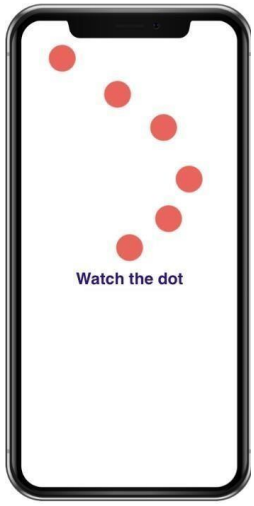


Fig. 7: an example of proof of liveness

Most notably, we have included the check of the ID of the individual against our in-server algorithms, as another strong deterrent for cheaters, however, the recording of the names is not strictly necessary for the purposes of our network; once an individual is proven a human and unique in our system, no further proof is needed.

Once the procedure is completed, a bar code is issued to the user to attend a humanity verification party.

4.2 HUMANITY IDENTIFICATION PARTIES

Humanity parties are carried out to validate the humanity of individuals and to confirm their biometric data; they are performed automatically without the need of

supervision/control, as long as premises and equipment are tamper sealed and in compliance with the protocol, and party participants are randomly chosen.

The simultaneous presence of several individuals randomly chosen by the system, makes collusions very difficult because the attacker is unaware of the composition of the party.
Reciprocal recordings of biometric data among attendees is a strong deterrent to using video spoofing because the plot will have to be synchronised with others.

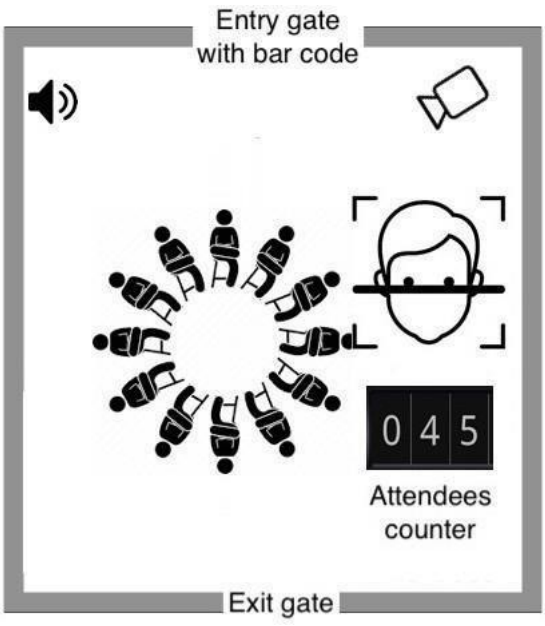


Fig. 8: the scheme of the premises of a humanity party

The presence of a randomly chosen verifier coming from a remote location additionally strengthens the protocol.

The scheme of the venue of a humanity party (Fig. 8) is designed to achieve the following procedure:
- the party is fully video recorded and internal speakers page participants.
- the bar code received at 4.1 is scanned to enter the premises
- individuals gather in the room by the scheduled time, if the number of attendees is inferior to the minimum required, the party is delayed or canceled
- participants confirm the count of the attendees of the party
- each attendee performs a face recognition of all the others to confirm their biometric data using the face recognition feature of the minting app. Vocal identification can be added to obtain more refined data.
- one or more verifiers may attend the party in incognito.
- at the end of the party all participants are also requested to feedback on the tamper seal of the venue and the regularity of the party.

Party transcripts are posted for public consultation in the blockchain, minting tokens are automatically issued to identified individuals once biometric data and party feedbacks are validated.

If one venue is tampered or vandalised, or somehow proved to be suspicious because of a high rate of canceled parties/absences, the AI engine increases scrutiny on its identified individuals, and/or premises are subjected to new inspection and wiring.

4.3 ADDITIONAL VERIFICATIONS PARTIES

To ensure the health of the ecosystem against any possible cheating, additional verification parties are carried out.
The aim is to create an anti-collusion pyramid structure that is based on the principle ''everyone or none is cheating'' as described in chapter 6.

Individuals are contacted via the minting app and email to attend an additional verification party organised automatically by the AI engine which is stored in the blockchain. Each participant is selected from a different humanity party, absences increase the scrutiny on unverified parties.

Additional verification parties play out the same way as human identification parties and use the same locations.
Because verification parties cover a wider catchment area, individuals should be encouraged to travel and participate with awards/coins.

The bigger is the average size of the party the more secure and verifiable is the system.

In a country with 1 million identified individuals, and an average size of humanity verification parties equal to 20 individuals, and 5 additional verification parties performed every day, even a very small intake of 0.01% of forgeries has good chances to be spotted after just a few days. The following table shows how long it takes to discover a cheater over different levels of forgery intake (fig. 9) when the principle ''everyone or none is cheating'' is attained.

100 additional verifications/day	100 forgeries / 1 million	500 forgeries / 1 million	1000 forgeries / 1 milion
Chances to spot a cheater within 1 day	18%	63%	86%
Chances to spot a cheater within 5 days	63%	99%	~100%
Chances to spot a cheater within 10 days	87%	~100%	~100%

Fig. 9: The odds to spot cheaters while performing 100 additional verifications/day (5 parties) in a country with 1 million individuals identified, with a humanity verification party average size of 20 individuals.

4.4 THEORETICAL CHEATS

Cheat 1: The party is entirely made of spoofed identities.
Because of the random formation of the parties, an attacker would have to produce several spoofed applications which will create a high rate of absences to the surrounding parties, and reasonably that will increase the scrutiny of the AI engine. The attacker will also have to:
- pass the online identification
- tamper the identification venue, which is video recorded and allows only single entries with bar code
- run spoofing videos on all devices
- collude with the verifier who comes from a remote location

In this cheat the principle ''everyone or none is cheating'' applies, and the AI engine will discovered the cheat very quickly by running additional verifications (see fig. 9).

Cheat 2: The party is attended by honest individuals but one is a spoofed identity.
This cheat is reasonably impossible when the state of art biometric technology is inplace. Furthermore, colluding with other attendees in the attempt to run a spoofing video, is also very unlikely.

However, the AI engine running additional verification parties, will discover the cheat outside the principle ''everyone none is cheating'' in just a longer amount of time (see fig. 10).

4.5 FALSE POSITIVE AND DISPUTES

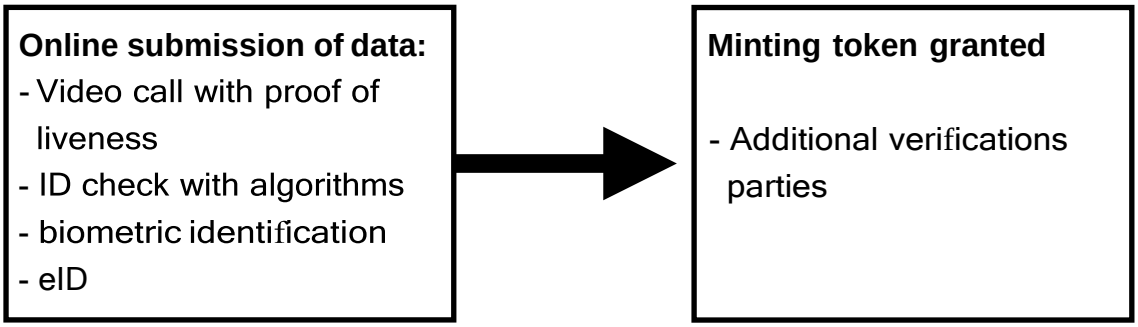
In the event of identification denial (positive identification), individuals may appeal/ request rectification directly via app. In most of the cases the appeal will result into either a denial or the request to attend a verification ruled directly by the AI engine.

In exceptional cases, individuals may request the rule of the PGO. This does introduce some discretion on the entire system, however, the number of those cases should be kept very low so that the ecosystem is enriched by this option and not devalued by.

5. ALTERNATIVE IDENTIFICATION PROCEDURE WITH EID

Electronic identification, eID, is developing in many countries, and available to the public, this suggests the opportunity to run an alternative method without the need for humanity identification parties (additional verification parties will still be in place).

With eID the identification procedure is carried out entirely online:



Because eID supposedly guarantees both conditions 1 and 2, an attack is always classified as government cheating or eID inefficiency, and in order to detect that, the system runs additional verification parties.

In a country with 1 million identified individuals and 100 additional verifications performed every day (5 parties with 20 attendees), even a very small intake of 0.01% of forgeries has good chances to be spotted after just a few months (fig. 10).

If government cheating/inefficiency is discovered, the entire national eID system is disqualified and the country is moved to the standard identification procedure.

100 back checks /day	100 forgeries / 1 million	500 forgeries / 1 million	1000 forgeries / 1 milion
Chances to spot a cheater within 10 days	10%	39%	63%
Chances to spot a cheater within 50 days	39%	92%	99%
Chances to spot a cheater within 100 days	63%	99%	~100%
Chances to spot a cheater wthin 150 days	78%	~100%	~100%

Fig. 10: The odds to spot cheaters while performing 5 verification parties/day in a country with 1 million individuals identified, outside the principle ''everyone or none is cheating''.

National eIDs ease a quick diffusion of the PoID network, and attract small entrepreneurs to invest into identification venues (see chap. 9).

In Europe, many countries are offering public eID services: UK, Sweden, Germany, France, Italy, and many more; in fact, the European Community is pushing all members to pursue the program (https://digital-strategy.ec.europa.eu/en/policies/electronic-identification), in the US a similar program is also kicking off.

5.1 THEORETICALCHEAT

Cheating this procedure requires to:
- implement an AI spoofing video tool able to cheat our proof of liveness algorithms
- video counterfeit documents/IDs to cheat our ID video check algorithms
- tamper the birth register or pass the eID process on behalf of different existing identities

However, the AI engine running verification parties, will discover the cheat outside the principle ‘’everyone none is cheating’’ as shown in fig. 10.

6. VERIFICATION PYRAMID STRUCTURE - VERIFICATION PARTIES

As PoW and PoS are battles to acquire electrical energy and accumulate capital, PoID is a massive effort to identify individuals on a global scale, based on an AI engine which organises verification parties and continuously computes the chances of cheating at individual/area/country/region level.

Negative feedbacks increase the level of scrutiny on the specific issue, e.g. an abnormal rate of parties absences would trigger a higher rate of verification parties or even the disqualification of the venues; negative feedbacks on a venue tamper seal would trigger further inspections or re-wiring.

Additional verification parties are designed to form a verification pyramid system based on the following principles:

- **Randomness**: parties participants are randomly assigned by the AI engine to one identification venue of the area, no entity can manipulate the list of the attendees, whose data are publicly stored in the blockchain.
This condition makes it very difficult for a cheater to spoof an entire party because he would have to apply for several identifications before gaining control of one party.

- **Everyone or none is cheating**: parties cannot be attended by both honest humans and video spoofs, therefore all individuals of a party are either spoofs or humans.
In reflection to this, it is fair to assume the verification of one single individual of a party verifies the humanity of all the participants of the party.

- **Pyramid structure**: because of the above condition, spoofing parties are detected as soon as one spoofed identity is grouped into a level up party attended by humans.
In other words, one cheater must spoof an entire humanity party to cheat the system at the first level; however, at the above level, the party will have again to be entirely made of spoofed identities to not be detected, and so on.
The higher is the level of the verification parties and the interoperability of the verifiers, the more extensive the collusion has to be to not be detected (fig. 11).

The figure below shows a massive collusion involving the PGO, the majority of the verifiers of one region, and producing spoofed identities at all parties of one country (the country has no honest human applicants).

However, the collusion would be discovered at the level of intra-countries-verifications, once a party would group individuals coming from another country.

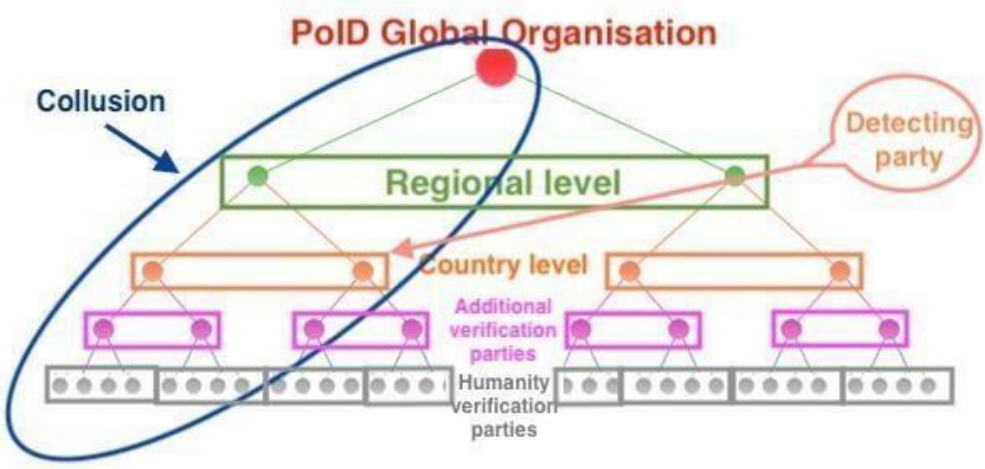


Fig. 11: An extensive collusion carried out across all countries of a region and one entire country made of spoofed identities, is detected as soon as one party groups individuals coming from human and non-human sub-level parties.

7. DATA STORAGE

Data in the PoID are stored into two different levels:

Level 1: data are stored in a centralised server managed by the PGO:

- Online data submission
- Parties video recordings

Level 2: data are stored in the blockchain:

- Biometric data of the minters
- AI engine
- Parties formations, feedbacks, venues breaks in transcripts

Fig. 12: data storage in the PoID.

It is important to notice that the role of the PoID Global Organisation (PGO), is limited to facilitate the early stages of the network, PGO's activities and data storage are not essential for the system to run. Once the network is globally adopted it is reasonable to expect the PGO involvement to be limited to only disputes resolution.

Biometric data of the minters are stored partially encrypted to preserve the privacy of the minters. The amount of data publicly stored in the blockchain is enough to avoid a good number of collisions with new participants, yet it is not enough to securely match one complete record to its collision in the blockchain.
Once a collision occurs, the AI engine requests the minter to temporarily disclose his encryption key for biometric verification; new encryption keys are then issued to the two participants.

Fig. 13: biometric data of minters are partially encrypted.

The AI engine and the biometric technologies are initially made by the PGO, but the aim is to kick off in time an open source library, in which participation of coders is incentivised with awards/coins.

8. ONGOING PROOF VS PERMANENT PROOF

A big fault in the design of PoW networks is that the mining is ongoing, the dependency from the energy is ongoing and so is the liability to censorship. PoW protocol requires work at every block.

A similar condition is found in the Proof of Personhood network: PoP tokens are temporary/expiring, in general, every time new individuals request to join the network, all minters previously personhood'ed have to attend again to a new pseudonymous party. This suggests that either the process will be endless or it will eventually leave out those individuals incapable to attend the last party.

A better setup is offered by Proof of Stake networks, here the ongoing proof of stake does not require continuous efforts, yet it is not permanent either, it never gets to a stable and secure status, oligarchies can be formed anytime.

In the PoID network instead, identifications take place only once and last the individual's life (or a good part of it); each identification adds a permanent and stable- in-time value to the network.

Once an individual joins the PoID network, he assigns his signature to a cohort of minters who will take part in the validation of the blocks on his behalf.
A cohort is a node with the delegated voting power of its participants, the consensus of the network is reached according to the majority of the voting power and so is the block award distribution.
The system works as a PoS protocol, but being based on individuals and not on stakes, it is much more democratic, distributed, and most importantly permanently stable: there is never an incentive for a minter to cheat the system as that will end up in one against all others.

As soon as the network is reasonably dispersed i.e. minters cannot possibly contact each other in big numbers, cannot infiltrate the network in big hostile groups, the formation of oligarchies becomes permanently impossible. In fact, a very low number of nodes is enough to prevent the formation of oligarchies and outnumber the infiltration of hostile groups.

Because the PoID network builds on permanent value, it is also resilient to censorship. In the PoID, governments can restrict access to eID and even outlaw verification parties, but that would not affect the minters who joined the network earlier on, they would continue to run the network and benefit from the minting, this is what we call permanent censorship resilience.

In regard to hypothetical IPs ban, as there are no specific minimum or maximum constraints over cohort size, PoID network can potentially use a great number of IPs on the internet if globally adopted (as opposed to PoW which only uses 5-600.000), and therefore not less than the shut down of the entire internet is required to censor the PoID network.

Arguably, the PoID network puts pressure on governments (see Fig. 14), as opposed to being liable to censorship; in a hypothetical situation in which the network has grown to the 50% of the population of a democratic country, a government ban would look very authoritarian and unfair to the point to be publicly questioned.

PoID network is made of real individuals and therefore is the democratic expression of the will of the people, another prerogative that other networks do not have.
The fact that minters are real people and easily reachable via minting app, also suggests the opportunity to offer to people an instrument of democratic participation, and run petitions and surveys on politics or other matters.

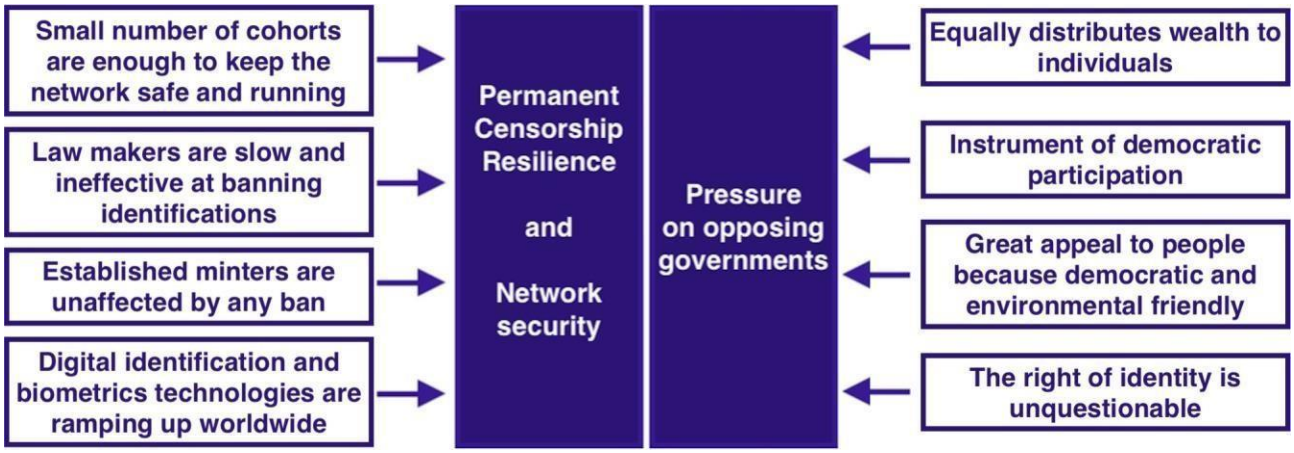


Fig. 14: The Identity Coin (PoID coin) is censorship resilient and secure as soon as the network is formed by a relatively small number of cohorts, also being the expression of the people will, it puts pressure on opposing agencies

9. THE ORGANISATION OF THE POID

The PoID network generates and distributes wealth through the Identity Coin minting process. The network is fully controlled by its minters and cohorts leaders.

The identification process runs without the control or involvement of humans: parties are formed by the system, internal doors are automatically controlled and all data are recorded in the blockchain.

The creation of new identification venues is also done automatically by the blockchain: once a good number of applications is received in an area/country, the system offers opportunities to identification investors, following a criteria which maximises coverage.

The PoID network is made by the following roles:

Minters: individuals that once identified, choose a cohort to join the network voting system and receive in return Identity Coins.
Minters can move from one cohort to another, effectively awarding those cohorts that have high validation efficiency and discouraging those operating poorly.
Minters also periodically vote for the president of the PGO, their vote is performed via the minting app.

Cohort leaders (C.L.): individuals running the minting nodes of the network to validate blocks.
C.L. may run fast and highly efficient nodes if high transaction output is needed, within an incentive mechanism also balancing the need of a high number of nodes.
If the PoID network is joined by several blockchains as shown in Fig. 5, cohorts leaders can choose the blockchain to validate according to their interest and output capacity.

Identification investors: individuals who invest in identification equipment and venues; they are responsible for the tamper seal of the premises and liable to fraud charges, their license is revokable.

Verifiers: individuals who have become minters are randomly offered the opportunity to attend parties as verifying guests and receive coins in return for their contribution.
The idea behind the role of a verifier is:
- verifiers are often people coming from abroad, their jobs are randomly assigned,

hence collusions are very difficult to form.
- once a verifier confirms his interest to attend a party, his attendance is very likely and even compulsory to confirm the party; this condition strengthens the impossibility of a cheater to gain control of a spoofing party, or to create collusions in a party.

PGO headquarters: The PoID Global Organisation is a profitable company that manage the installation of identification equipment worldwide.
Its role is merely technical and has no involvement in the assignment of the licenses: once a new venue is granted and equipment has been installed, a technical verifier is randomly selected to check the tamper seal of the venue.
This task is reasonably important but not critical for the health of the ecosystem because all venues and parties are constantly feedback'ed by individuals and verifiers, therefore any non-compliance would automatically trigger a new inspection or disqualification.
According to the spirit of the PoID, PGO governance is elected by the entire network of minters to avoid collusions and to keep a high turnover of people in the system.

10. HARD-FORKING POW BLOCKCHAINS INTO THE POID NETWORK

The capacity of the PoID network to securely validate blocks of any blockchain without particular additional efforts opens up interesting scenarios.
Here we examine a possible hard-fork of Bitcoin Core into Bitcoin PoID:

- A warm-up period is started and the Bitcoin ledger is copied and mirrored at each block into a duplicate ledger/blockchain;
- wallets owners are invited to gain access to their mirrored wallets in the duplicate d blockchain: they open a depository box in our system which returns an application code; the application code is attached to an OP message as a transfer of 1 Satoshi is executed from the wallet in the Bitcoin network to a service wallet; once the transfer is received, the depository box shows the secret key of the new wallet in the duplicate blockchain
- during the warm-up period, wallets in the duplicate blockchain are completely inactive and cannot send or received coins, new wallets cannot be created, unless that is done in the original blockchain
- once some interest has been shown by wallet owners, the warm-up period ends and the new blockchain stops mirroring the original blockchain, transactions are now independently performed under the PoID protocol, the hard fork is completed.
- wallets owners will still be able to gain access to their duplicate wallets in the PoID blockchain if those exist at the last block prior to the hard fork.

The duplication of an existing PoW blockchain into PoID, is nothing more than a marketing/advertising exercise (imagine the slogan: *"Bitcoin is Hard-Forking into an electricity-free blockchain"*), but it shows case the benefits, it arouses people interest and debate around the PoID.

However, it is interesting to note that the PoID network has a great potential to change the entire crypto industry alluring or even forcing other blockchains to join its network. All blockchains with open source code can be hard-forked/duplicated into newly created blockchains running on the PoID without any involvement of the current miners.
It will be possible to create new coins which initially have no history in the market, but being identical as per features and wallets balances, clearly are worth more than the originals simply because of the network superior features: no energy consumption, faster transactions with fewer network fees, no security issues (no 51% attack, etc.).

11. THE OPPORTUNITIES INVOLVED IN THE POID AND IDENTITY COIN

The goal of identifying individuals on a global scale is certainly a grandiose and ambitious project, but the opportunities carried along are also fascinating, among those, we envisage the following:

1. The opportunity to unify all blockchains into a single blocks validation network.

2. The opportunity to create a global income for individuals, based on the undeniable value of their individual identity.

3. The opportunity to run a global organisation aiming to facilitate the identification of billions of people worldwide, distributing jobs to thousands of people, and prospering from the earnings coming from identification services.

4. As minters are real and identified individuals, the opportunity to build on the minting app a democratic tool of participation to national and global matters, (e.g. Brexit re-run, Bielorussia election validity, etc).

It is easy to foresee a viral diffusion of the Identity Coin as it spreads wealth equally to everyone, it does not consume energy, it does not allow the formation of oligarchies or favour the richer.

12. SUMMARY TABLES

Table 1: main proofs comparison.

	Network Decentralisation	Network Distribution	Energy consumption	Censorship resistance	Capacity to build permanent value	Negatives
PoW theoretically BITCOIN	Full: it was expected to be ran by householders	Full: it was expected to be ran by householders	High	Full: it was expected to be ran by householders	None: proof of work is needed at each block	Inefficient - poor scalability - storing value only
PoW actually BITCOIN	Poor: it is actually ran by mining corporations	Poor: it is actually ran by mining corporations	High	Poor: governments can actually restrict access to electricity	None: proof of work is needed at each block	All the above + unsustainable, easy to censor
PoS FILECOIN	Full	Full	None	Full	Fair: proof of stake is effortless but it is not stable/permanent	Bad deal for individuals, the bigger benefits from the participation of the smaller
PoP POP-COIN	Full	Fair: pseudonymous parties can be carried out in most of the countries	None	Poor, governments can disrupt pseudonymous parties	Poor: each epoch can be better or worse than the previous one	Pseudonymous parties are easy censorable and need many repetitions
PoID IDENTITY COIN	Full	Fair: identifications can be carried out in most of the countries	None	High: identifications can be outlawed but existing minters stay unaffected	Minters are forever and there is never an incentive to cheating	Relies on humanity parties

Table 2: proofs based on individuals comparison.

	Proof of Personhood	Proof of UniqueID	Proof of Identity
Organization of the parties	By everyone	Initiated by a community of enthusiastic people	Randomly assigned by the AI engine
Participation method	Attendance at pseudonymous party	Biometric identification, P2P with some verifiers	Biometric identification, everyone to all, during identification parties
Supervision of possible cheats	None	None	Computed by the AI engine, detected by additional verification parties
Privacy of participants	Individuals can attend parties in disguise	Data are encrypted on a trusted setup	Data are partially encrypted on a decentralized setup
Builds on permanent value	No, every epoch can be better or worse than the previous	Yes	Yes
Use cases	Universal basic income	- Universal basic income - Consensus protocol for all blockchains - Voting system	- Universal basic income - Consensus protocol for all blockchains - Voting system
Negatives	Endless process or leaving out those unable to attend last round or parties	- Relying on Ethereum blockchain - Sybil attacks are unlikely but theoretically possible	- Lots of data to store on the blockchain - the protocol is eased by a global organization, which may be seen as a center of authority

About UCL CBT

The UCL CBT is the first centre globally to actively focus on blockchain-related research onthe adoption and integration of Blockchain and Distributed Ledger Technologies into our socio-economic system.

The unique characteristics of the CBT at UCL provides a cross-sectoral platform connecting expertise and drawing knowledge from eight UCL departments centrally in one place. The CBT is a centre of excellence fostering open dialogue between industry players and sharing expertise and resources. It is a neutral think tank providing consultancy services to industry members, dedicated knowledge-transfer activities and cutting-edge in-house solutions.

For engagement outside of the academic world, the CBT's activities have been tailored to industry and policymakers' needs. The UCL CBT draws on its world-leading academic expertise to produce blockchain solutions for industry, start-ups and regulators. With a community of over 247 Research & Industry Associates and Industry Partners, it is the largest Academic Blockchain Centre in the world.

Notable Work

- The CBT released a report on the current adoption of DLT in global physical supply chains. The report featured an analysis of over 100 different projects taking place all over the world in the Grocery, Pharmaceutical and Fashion industries. Access the report here.
- The CBT is a founding member of the Covid Task Force alongside The International Association for Trusted Blockchain Applications (INATBA) and the European Commission. The task force is convening key players in the global blockchain ecosystem to identify deployable technology solutions that address governmental, social, and commercial challenges caused by COVID. As well as identifying solutions, the Task Force will work to expedite their deployment.
- The CBT successfully funded nine research proposals that investigated topics including stable coin policy, smart contract innovation, blockchain economics and blockchain governance models. Research teams who were funded were made up of individuals from a variety of academic and industry organisations. Learn more aboutthe projects here.
- The CBT launched the Block-Sprint hackathon to promote DLT innovation in the financial services sector. Over 100 individuals took part in both the 2019 and the 2020 edition forming teams made up of industry practitioners, academics, and students. Learn about the winners and innovate ideas that were generated in the hackathon here.

About the Discussion Paper Series

The *UCL CBT Discussion Paper* *is* published on a quarterly basis featuring the latest developments in the blockchain and DLT space. The aim of the discussion paper series is to share recent developments and state-of-the-art solutions on blockchain and DLT of researchers from an interdisciplinary background with the CBT community. All accepted submissions are available in the CBT paper database.

The submissions are circulated among the members of the UCL CBT Editorial Board, led by the Scientific Director so that the results of the research receive prompt and thorough professional scrutiny.

If you are interested in submitting a paper to be included in forthcoming editions, please visit our website here to see what the latest theme and criteria for submission are.

UCL Centre for Blockchain Technologies

http://blockchain.cs.ucl.ac.uk/

UCL Computer Science
Malet Place
London WC1E 6BT
United Kingdom