Decentralized Markets for Public Goods: 
Solving Collective Action Problems Using Blockchains

By
Abhinav Sriram
Bachelor of Science (Sc.B.), Brown University, 2023

A Thesis Presented to The Department of Computer Science
In Partial Fulfillment of the Requirements for the Degree of
Master of Science (Sc.M.) in Computer Science

Brown University
Providence, Rhode Island
May 2023

Advisor: Maurice Herlihy
Abstract

Public goods are important, because their provision is, to varying degrees, essential to the smooth functioning of society - economically, politically, and culturally. The non-excludable and non-rivalrous nature of public goods results in it generally being insufficiently profitable to be provisioned by the private sector, and so, in the absence of government provision, would not be provisioned at all. However, due to unremunerated externalities and the free rider problem, there is under provision, overutilization, and degradation of most public goods that do end up being provisioned by governments. The collective action problem prevents society from being able to work together to address both the lack of provision of public goods and abuse of provisioned public goods. This paper presents a framework for a solution to that problem. Markets are extremely powerful but fail to provision adequate levels of public goods due to certain market failures. These market failures may be addressed by designing a mechanism that eliminates as many of the incentives as possible that cause them. A systematic review of the literature on the economic, biological, psychological, and sociological theories that affect collective human behavior is used to design a system around the aforementioned mechanism that encourages endogenous community formation to collectively organize around issues of public interest, primarily that of provisioning and maintaining public goods. The system proposed in this paper will provide a standardized framework via infrastructure that is decentralized, so as to establish trust via complete transparency, aid in establishing self-governance, and to eliminate as many of the remaining perverse incentives as possible that contribute to the collective action problem.
## Contents

Introduction ........................................................................................................................................ 1

1 Collective Action Problems ........................................................................................................ 2

1.1 Historical Context ..................................................................................................................... 3

1.1.1 Early Ideas ............................................................................................................................ 3

1.1.2 Modern Ideas ........................................................................................................................ 3

1.2 Theoretical Explanations .......................................................................................................... 4

1.2.1 Economic Theories .............................................................................................................. 4

1.2.2 Biological Theories ............................................................................................................. 5

1.2.3 Psychological Theories ....................................................................................................... 6

1.2.4 Sociological Theories ......................................................................................................... 7

1.3 Summary of Observations ....................................................................................................... 8

2 Public Goods .................................................................................................................................. 9

2.1 Semantics .................................................................................................................................. 9

2.2 Provisioning ............................................................................................................................... 10

2.2.1 Governments and Nonprofits ............................................................................................ 10

    Aside: Concentrated Benefits and Diffused Costs .................................................................... 13

    Aside: Coase Theorem .............................................................................................................. 14

2.2.2 Free Markets ....................................................................................................................... 16

    Aside: Efficiency ...................................................................................................................... 20

    Aside: Time Value of Money ..................................................................................................... 21

2.2.3 Market Failures .................................................................................................................... 22

    2.2.3.1 Valuation ...................................................................................................................... 22

    2.2.3.2 Externalities ................................................................................................................ 23

    2.2.3.3 Free Rider Problem .................................................................................................... 27

2.3 Local vs Global ......................................................................................................................... 29

2.4 Ownership ................................................................................................................................ 29
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.3.1 &quot;For&quot;</td>
<td>54</td>
</tr>
<tr>
<td>6.3.1 &quot;Against&quot;</td>
<td>56</td>
</tr>
<tr>
<td>6.4 Membership</td>
<td>58</td>
</tr>
<tr>
<td>6.4.1 Authentication</td>
<td>58</td>
</tr>
<tr>
<td>6.4.2 Wallets</td>
<td>59</td>
</tr>
<tr>
<td>6.5 Proposals</td>
<td>61</td>
</tr>
<tr>
<td>6.5.1 Approval</td>
<td>61</td>
</tr>
<tr>
<td>6.5.2 Tokens</td>
<td>61</td>
</tr>
<tr>
<td>6.5.3 Metrics</td>
<td>62</td>
</tr>
<tr>
<td>6.5.4 Bidding</td>
<td>64</td>
</tr>
<tr>
<td>6.5.5 Conflict Resolution</td>
<td>65</td>
</tr>
<tr>
<td>6.6 Assurance Contract</td>
<td>67</td>
</tr>
<tr>
<td>6.7 Philanthropy</td>
<td>69</td>
</tr>
<tr>
<td>6.7.1 Charities</td>
<td>70</td>
</tr>
<tr>
<td>6.7.2 Tax Benefits</td>
<td>71</td>
</tr>
<tr>
<td>6.8 Governance</td>
<td>72</td>
</tr>
<tr>
<td>6.8.1 Liquid Democracy</td>
<td>72</td>
</tr>
<tr>
<td>6.8.1.1 Decentralization</td>
<td>72</td>
</tr>
<tr>
<td>6.8.1.2 Historical Governance</td>
<td>73</td>
</tr>
<tr>
<td>6.8.1.3 Delegating Votes</td>
<td>73</td>
</tr>
<tr>
<td>6.8.1.4 Expert Committees</td>
<td>74</td>
</tr>
<tr>
<td>6.8.1.5 Arbitration Committee</td>
<td>75</td>
</tr>
<tr>
<td>6.8.2 Token-Based Governance</td>
<td>76</td>
</tr>
<tr>
<td>6.8.2.1 Incentives</td>
<td>76</td>
</tr>
<tr>
<td>6.8.2.2 Attacks</td>
<td>76</td>
</tr>
<tr>
<td>6.8.3 Reputation-Based Governance</td>
<td>79</td>
</tr>
<tr>
<td>6.8.3.1 Limited Governance</td>
<td>80</td>
</tr>
<tr>
<td>6.8.3.2 Vote-Buying</td>
<td>81</td>
</tr>
</tbody>
</table>
6.8.3.3 Bribing

6.8.3.4 Asymmetric Consequences

6.8.4 DAOs

6.8.4.1 The DAO

6.8.4.1.1 Responsibilities

6.8.4.1.2 Features

6.8.4.1.3 Structure

6.8.4.2 Contributors’ “DAO”

6.8.5 Voting Mechanisms

6.9 Communication

6.10 Reputation

6.10.1 Users

6.10.1.1 Money and Activity

6.10.1.2 Local and Global Time

6.10.1.3 Example Scores

6.10.1.4 Simple Heuristics

6.10.1.5 Fraud, Collusion and Stalling

6.10.1.6 Certification

6.10.1.7 Special Users

6.10.2 Bidders

6.11 Attacks

6.11.1 Fraud

6.11.2 Collusion

6.11.3 Stalling

Conclusion
Introduction

There is a small town with 20,000 residents. It is a burgeoning town with lots of young children and families. However, their quality of life is rapidly deteriorating as a neighboring coal power plant pollutes the air. It would cost the power plant $1 million to install a carbon scrubber that stops this from happening. The town makes a plea to the operators of the plant to do this because they fear their kids are growing up in an extremely unhealthy environment. However, as a corporation with the interests of their shareholders at heart, they choose not to, since that $1 million would have to come straight out of their pocket. The town has no recourse since it is a small town with no powerful governance structures and the corporation running the power plant has already received the necessary regulatory clearance from the state. What can the residents of this small town now do? The carbon scrubber is a public good. Public goods are often provisioned by governments via taxation and voting. The other prevalent method to provision such goods is to “privatize” a public good by making it exclusionary somehow so that it is sufficiently profitable to be provisioned by private companies. There are also those institutions that lean on moral, cultural, religious, or social motives to induce individuals to contribute to provisioning public goods. However, in the absence of such institutions, as is the case here, the rational, if not obvious, solution would be to attempt to raise the money for the carbon scrubber themselves. If the people in the town cared enough to make a plea to the power plant to install the scrubber, it means they value clean air for their children (and themselves) at a certain price. If that value, collectively across everyone in the town, is greater than $1 million, then they should raise the money to pay for the scrubber themselves. However, this is not as easy as it seems. Getting everyone to coordinate to volunteer money for the scrubber, to agree on what costs they believe they are incurring as a result of the polluted air, the costs of ensuring that the scrubber is actually installed once the money is collected, the issues that arise due to some people wanting to deliberately free ride etc. are being ignored. The town will be unable to install the scrubber themselves, even if the economics are in their favor, due to the collective action problem.

This is discussed in Section 1 alongside an overview of the literature on the economic, biological, psychological, and sociological theories that affect collective human behavior leading to this problem. A lot of the results from research in those fields will be used to design the system that solves those very problems in Section 6 and these ideas are built upon in Section 4. The collective action problem is particularly pronounced in the case of the provision and maintenance of public goods. Section 2 discusses this in more detail. It discusses how public goods are provisioned today - primarily through governments using taxation, and the inefficiencies of such an approach are elucidated in detail. The natural alternative of using free markets is then discussed. The power of markets is demonstrated but so are the market failures (primarily, the lack of valuation mechanisms, the presence of unremunerated externalities, and the free rider problem) that prevent it from provisioning the economically efficient level of public goods.
Section 3 summarizes the arguments for and against markets as a solution to the problem at hand; making the case that markets are the best solution to organizing people around solving collective action problems, and especially, for the provision of efficient quantities of public goods. Section 4 then provides an overview of three fields of study that aim to address the many problems with markets discussed thus far, namely mechanism design, system design, and governance design. Mechanism design aims to modify institutional arrangements to incentivize intended behavior in participants; but the conclusion arrived at is that it is impossible to do this in a mathematically provable manner for the successful provision of public goods by free markets. However, there is still hope as empirical data does not support these proofs. System design builds upon ideas discussed in Section 1 regarding economic, biological, psychological, and sociological theories that affect collective human behavior, and how these can be taken advantage of to encourage cooperation to solve collective action problems. This includes solutions ranging from smaller group sizes during community organization efforts to help discourage free riding, to reputation systems that reward and punish people as appropriate. Lastly, governance design provides an overview of political science theory behind various kinds of voting systems and the ones needed to allow communities to organize around issues of interest. The voting mechanism for this system will need to be strategy proof and hence uphold independence of irrelevant alternatives.

Section 5 summarizes everything learnt thus far from work that has been done on these topics into 10 desirable properties that must be satisfied by the system being built. Section 6 uses these properties as its north star to define a fully decentralized market for public goods that addresses all the problems free markets have today by means of mechanism design. It is implemented on a blockchain to provide trust and transparency that no other technology primitive can provide. This also aids with defining a robust governance structure that is pivotal to ensuring not just the success of provisioning public goods but also its long-term sustainability. The system designed possesses these features, and in essence, allows for people to make contributions to proposals that aim to provision public goods that they believe are personally beneficial to them or worthy of their hard-earned money. Any individual or organization is free to bid on such proposals. If there is a match between a bidder and a proposal, then it is pursued. Everyone who contributed to a proposal is given voting rights to govern the public good being provisioned and empowered with the tools they need to sustainably organize around the problem in a manner that encourages endogenous community formation.

1 Collective Action Problems

A collective action problem, also called a social dilemma [4][5][7][8], is a situation in which all participating parties would be better off cooperating with each other but fail to do so because of conflicting interests between individual parties. [1][2][3] When group members choose to pursue individual interests at the expense of the group, which most humans often tend to do, then collective action problems arise. It can be best understood through a game theoretical analysis and the free-rider problem resulting from the provision of public goods; and can be used to explain resource depletion, low voter turnout, overpopulation, and numerous other public policy concerns faced today.
1.1 Historical Context

Collective action problems can take many forms and have been studied across various disciplines such as psychology, economics, and political science throughout history.

1.1.1 Early Ideas

“If any two men desire the same thing, which nevertheless they cannot both enjoy, they become enemies.”
- Thomas Hobbes, Leviathan, 1651

Hobbes never explicitly used the words “collective action problem” but the phenomenon he was describing was the very same.

“Two neighbours may agree to drain a meadow, which they possess in common; because it is easy for them to know each others mind; and each must perceive, that the immediate consequence of his failing in his part, is, the abandoning the whole project. But it is very difficult, and indeed impossible, that a thousand persons should agree in any such action; it being difficult for them to concert so complicated a design, and still more difficult for them to execute it; while each seeks a pretext to free himself of the trouble and expence, and would lay the whole burden on others.”
- David Hume, A Treatise of Human Nature, 1738

Hume also never explicitly used the words “collective action problem” but he gets closer to the more modern definition. He claims that if a thousand people are expected to work together to achieve a common goal, individuals will be likely to “free ride” (Section 2.2.3.3), as they assume that each of the other members involved will do enough to achieve said goal. Whereas, in smaller groups, each individual plays a much greater role and thus is less inclined to feel like they can free ride, both because they feel like they are pivotal to the effort at hand, but also due to social pressure and risk of social ostracization.

| In larger groups, the natural tendency for individuals is to free ride but when group sizes are smaller, the incentive to free ride is lower as each individual feels they are pivotal to the effort at hand and are subject to social pressure and ostracization. |

1.1.2 Modern Ideas

The definition of a collective action problem today derives most closely from Mancur Olson’s 1965 book The Logic of Collective Action. [6] The primary idea that Olson furthered was that when it came to public goods i.e., goods that are non-rival and non-excludable (Section 2.1), when individuals pursued their own economic interests (i.e., those interests beneficial to themselves) then it was not necessarily beneficial to the overall public. This is in stark contrast to Adam Smith’s “invisible hand” which stated that when individuals pursued their own
economic interests it would result in the collective well-being of the majority i.e., the principle behind free markets. [6]

1.2 Theoretical Explanations

There are various theories from economics, biology, psychology, and sociology that attempt to explain why humans behave in a way that facilitates the existence of collective action problems.

1.2.1 Economic Theories

Game Theory (i.e., rational choice theory, expected utility theory) is a principal component of economic theory and assumes that individuals are rational actors motivated to maximize their utilities where utility is often narrowly defined in terms of people’s economic self-interest. It serves as a great framework for explaining how individuals allocate scarce resources and how scarcity drives human interaction and trade.

The prisoner’s dilemma [11] is one of the most common, simple, and famous examples of game theory and also happens to be crucial to understanding collective action problems [12] because it perfectly illustrates Olson’s claim that individual interests conflict with group interests. The premise of the game is that two players are accused of a crime and are in separate interrogation rooms unable to communicate with each other. If player A testifies against player B, then player A goes home free while player B receives a hefty sentence and vice versa. If both players testify against each other then they both receive hefty sentences but if they both keep quiet, then they will both serve a much shorter sentence. It seems obvious that both players should just choose to keep quiet. However, when it comes down to it, players who are unable to communicate with each other will choose to testify against the other player because they have an individual incentive to do so and get away with a commuted sentence, but this results in both receiving a hefty sentence instead.

<table>
<thead>
<tr>
<th>Player A</th>
<th>Cooperate</th>
<th>Defect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperate</td>
<td>Player A → 1 Year&lt;br&gt;Player B → 1 Year</td>
<td>Player A → 5 Years&lt;br&gt;Player B → No Sentence</td>
</tr>
<tr>
<td>Defect</td>
<td>Player A → No Sentence&lt;br&gt;Player B → 5 Years</td>
<td>Player A → 3 Years&lt;br&gt;Player B → 3 Years</td>
</tr>
</tbody>
</table>

This model is purely for illustrative purposes. This problem can be easily solved had the players been able to communicate or play the game repeatedly but in more complex real-world situations with hundreds, if not thousands of players, the collective action problem spirals out of control preventing groups of those sizes from making decisions that are of collective economic interest.
The unscrupulous diner’s dilemma is an $n$-person prisoner’s dilemma and illustrates concretely how even groups of single digit sizes end up creating unfavorable outcomes. The situation here involves one where a group of individuals go out to eat and before ordering, decide to split the bill evenly. Each individual must now choose to pick the cheaper or more expensive dish for themselves (where the perception of a more expensive dish is one that is better than cheaper ones but not by enough that would justify paying the difference for themselves). They all reason that since the incremental cost to their own bill will be small when they order the more expensive dish and because they believe that others will pick more expensive dishes, or at the very least, suspect that they could, then they too should order the more expensive dish. Since everyone reasons similarly, the bill, even when split, rises by the same amount had the individual been alone and chosen to order the more expensive dish i.e., everyone ends up paying more even though they would not have otherwise. This was tested experimentally [98] and actually showed that despite elimination of incentives such as social awkwardness or prior established relationships, the set up proved that participants did not consider the welfare of others involved, which is in contrast to other laboratory experiments that asked similar questions in contexts that were hypothetical or more abstract.

| The collective action problem spirals out of control with larger group sizes preventing them from making decisions of collective economic interest. |

1.2.2 Biological Theories

Thinking about biological and evolutionary explanations to decision making that often lead to the collective action problem, provide some useful, or at the very least, interesting, complementary insights.

The selfish gene theory [13][14][15] proposes that individuals may pursue a seemingly irrational strategy to cooperate if it benefits the survival of their genes. The theory of inclusive fitness is closely related and claims that family members tend to cooperate because doing so facilitates the survival of their genes. Reciprocity theory provides a slightly different explanation claiming that in repeated encounters with the same participants, the threat of punishment encourages cooperation.

| Selfish Gene Theory and Reciprocity Theory show that smaller groups with close-knit connections tend to be more successful at working together. |

Theories of indirect reciprocity and costly signaling may be useful in explaining larger scale cooperation because when individuals can selectively choose partners, it pays off to develop a cooperative reputation. This is further supported by the fact that cooperation communicates kindness and generosity which combine to make individuals an attractive and desirable group member.
1.2.3 Psychological Theories

The game theoretical explanation to collective action problems presupposed that the definition of an individual’s utility as being narrowly confined to their own economic self-interest was valid. Psychological theories challenge that assumption.

Interdependence Theory claims that people transform the standard pay-off matrix to an effective matrix based on who the players in the game are. A prisoner’s dilemma with close kin changes the pay-off matrix into one where cooperation is rational.

Attribution models support the claim of such transformations of the pay-off matrix because the rational choice will vary based on whether the player believes people are naturally greedy or cooperative. Goal-Expectation theory supports this with the claim that if players have a cooperative goal and believe that other players will cooperate then they will do so as well. [9]

The Appropriateness model claims that people do not actually compute their payoffs in real time like the game theoretical argument assumes but rather use far simpler heuristics to make their decision (the theory of bounded rationality also makes a similar claim). The equality rule is a very common one in which people will simply ask themselves the question, "what does a person like me (identity) do (rules/heuristics) in a situation like this (recognition) given this culture (group)?" and the answer to that question is what influences their decision to cooperate. [17][18]

Reciprocity Theory and theories of Indirect Reciprocity and Costly Signaling show that reputation mechanisms incentivize cooperation in larger groups.

Interdependence Theory shows that people’s utility functions change to account for interpersonal relationships when there are close-knit connections.

Attribution Models and Goal-Expectation Theory show that cooperation increases when individuals believe others will cooperate.

The Appropriateness Model supports the provision of heuristics that easily indicate to individuals the degree to which they can trust others.
1.2.4 Sociological Theories

Most economic models tend to assume that rational people only care about themselves (and their close kin) and specifically, only care about their own economic outcomes, and thus would only care about things that are personally benefiting them. However, people do care about those around them. They care about doing good and helping others. Donations to charities and political campaigns are a testament to this fact. One way of looking at this that would be consistent with an economic model would be to assume that the act of giving and feeling like one cares about the welfare of others itself provides additional utility beyond any utility one may receive otherwise. [68]

People care about those around them and want to actively (monetarily) help them i.e., people derive utility from helping those around them.

Social norms also play a big role. When individuals meet repeatedly and get to know each other, the incentive to lie decreases and may eventually even be lost depending on cultural norms. Even if this were not the case, individuals will eventually come to realize that while lying may instinctively be a protective strategy with those one cannot trust, that it is in fact self-destructive once trust is established leading to a culture/norm of just telling the truth. [69][70] From a game theory perspective, this makes sense as it has long been well known that non-cooperative strategies are not optimal in repeated games with the same players (including, as previously mentioned, in the prisoner’s dilemma).

As people get to know each other and develop close-knit connections, telling the truth and being cooperative is the game theoretically dominant strategy.

There are also other fairness and equality social norms such as “everyone should do their bit” that may contribute to people willing to help others and contribute to community efforts. [71] This is especially true for those in close-knit communities as people tend to feel far more altruistic toward those they know, live next to and are similar to in other ways. [72] Even outside of smaller communities, there is data [99][100] showing positive correlation between trust and economic efficiency at the macroeconomic level i.e., trust need not just come from interpersonal relationships but rather can also be facilitated by institutions.

There is positive correlation between trust and economic efficiency in societies and this trust can be facilitated by institutions and not just interpersonal relationships.
1.3 Summary of Observations

In larger groups, the natural tendency for individuals is to free ride but when group sizes are smaller, the incentive to free ride is lower as each individual feels they are pivotal to the effort at hand and are subject to social pressure and ostracization.

The collective action problem spirals out of control with larger group sizes preventing them from making decisions of collective economic interest.

Selfish Gene Theory and Reciprocity Theory show that smaller groups with close-knit connections tend to be more successful at working together.

Reciprocity Theory and theories of Indirect Reciprocity and Costly Signaling show that reputation mechanisms incentivize cooperation in larger groups.

Interdependence Theory shows that people's utility functions change to account for interpersonal relationships when there are close-knit connections.

Attribution Models and Goal-Expectation Theory show that cooperation increases when individuals believe others will cooperate.

The Appropriateness Model supports the provision of heuristics that easily indicate to individuals the degree to which they can trust others.

People care about those around them and want to actively (monetarily) help them i.e., people derive utility from helping those around them.

As people get to know each other and develop close-knit connections, telling the truth and being cooperative is the game theoretically dominant strategy.

There is positive correlation between trust and economic efficiency in societies and this trust can be facilitated by institutions and not just interpersonal relationships.
In these observations, there are two common properties:

Property 1: The tendency to cooperate and not free ride is much greater in smaller groups than in larger groups; and in smaller groups this goes up when those in the group have close-knit interpersonal relationships and/or care about each other. This can be fostered by encouraging communication but also by establishing means to record past tendency to cooperate i.e., institutions and infrastructure can also enable trust.

Property 2: Reputation mechanisms that have deterrents (i.e., some punishment mechanism) with simple heuristics that easily indicate to others the degree of trustworthiness/cooperativeness of those they are working with helps increase the likelihood of cooperation in larger groups. This is especially true when individuals are pseudonymous and may not have previously interacted with each other.

2 Public Goods

A public good may be defined as one that is both non-excludable and non-rivalrous. [19] Non-excludable goods are goods that users cannot be barred from accessing by any means. Non-rivalrous goods are goods where the use of the good by one person does not take away from the ability of another person to use that same good. [20] A caveat with this definition is that a public good must be valuable to more than one person, if not, the fact that it is non-excludable and non-rivalrous would be economically irrelevant.

2.1 Semantics

While a public good is formally defined as above, it is hard to find goods that display these properties universally. They often do so to a certain extent or only at certain times and under specific conditions. This is primarily because of the strictness of the above definition. Copyrights and Patents only further complicate this definition as they often provide temporary monopolies over non-rivalrous goods. Things like law enforcement, streets, libraries, museums, parks etc. are often called quasi-public goods since some excludability is possible but they generally display the traits of public goods. [30][31] Lighthouses are another weird example. It is difficult to exclude a ship from using a lighthouse’s services and one ship’s use of it does not detract from another’s, but at the same time, the benefits are limited to ships docking at a specific port and these ships often pay port fees (that generally have some cost to maintain the lighthouse bundled within them).

In addition to such generally ambiguous situations there are those where the non-excludability and non-rivalrous nature of the good is determined by assignment of property rights (Section 2.2.1). For example, one can mark a piece of land as their own and indicate that it is intended only for private use by building a fence around it. However, any fence can be overcome. What makes private land excludable is that property rights are defined such that the landowner has the right to stop trespassers and to claim the land as their own - and the legal system enforces this (as in the US). If property rights were defined differently, then land too could become a non-
excludable good (as in the UK, where the public have right of way on private land) [97]. A sillier example would be that not everyone enjoys the sight of wrinkled clothing, or that of those wearing socks with sandals but it is well within an individual’s rights to do so and thus those who are bothered by this sight must bear with it. This holds true in societies with a culture that values freedom of expression. So, property rights determine the non-excludability and non-rivalrous nature of goods, and these can change with time and even across cultures.

There are a few more things that make the lines defining public goods even blurrier. There are limitations to consumers’ consumption decisions regarding the public good once it is provisioned. For example, different people may have different preferences regarding the level of national defense they would like, but once provisioned, they will take what they get. The same goes for clean air; people live with the air they find around them. This suggests that colloquially, public goods are called as such because they are “public” in nature. At the same time, it has been suggested that the “public” nature of a public good may actually be an effect of its provision by the public rather than its cause. [67] Does public investment in a private good make it public? Even if some of these notions are mistaken, they illustrate the fact that the boundary between public and private goods is a porous one.

In general, for the purposes of this paper, relying strictly on the formal definition of public goods is not helpful, but rather a good faith interpretation of the term “public goods”, whenever it is used, is. Despite context indicating that this stance is what is most ideal for this paper, formal clarification was needed to avoid any confusion. Steven Shavell’s way of thinking about this is an effective summary for the stance taken here. He says “when professional economists talk about public goods they do not mean that there are a general category of goods that share the same economic characteristics, manifest the same dysfunctions, and that may thus benefit from pretty similar corrective solutions...there is merely an infinite series of particular problems (some of overproduction, some of underproduction, and so on), each with a particular solution that cannot be deduced from the theory, but that instead would depend on local empirical factors.” [32]

2.2 Provisioning

2.2.1 Governments and Nonprofits

The non-excludable and non-rivalrous nature of public goods results in it generally being insufficiently profitable to be provisioned by the private sector, and so, in the absence of government provision, would not be provisioned at all. Thus, most public goods are in fact provisioned by the government using taxpayer funds. Taxpayers enjoy the benefits of these public goods in return for their tax contribution and for being law abiding citizens. [26][27] There are of course free riders, those who do not pay taxes, but the hope is that law enforcement acts as a sufficient deterrent to encourage the majority of people to pay taxes and hence not be free riders.
People have a natural tendency to free ride and generally require a strong deterrent (such as law enforcement) to discourage them from doing so.

An obvious issue with using taxpayer money is that taxes are generally charged as a flat rate (for those in a given income bracket, neighborhood etc.), however everyone in a given “group” (when grouped per tax brackets they fall within) may not value the public good provisioned with their tax money the same way. A proposed solution to this problem, that is quite hard to implement in practice, is the Lindahl tax. This was proposed by Swedish economist Erik Lindahl in 1919 and his idea was to tax individuals for the provision of public goods proportional to the marginal benefit they receive. Measuring the marginal benefit received by individuals is the hard part and so his proposal hinged on the idea that people would be voluntarily willing to pay for things that benefit them, and the more they benefit from something, the more willing they would be to pay for it. For obvious reasons, this did not work well in practice.

Despite the Lindahl tax itself not working, taxes continue to be used to fund most public goods and so naturally the primary responsibility to provision public goods continues to fall on the government. However, the government is in the business of pleasing the median voter, of doing just enough to continue to remain in power and nothing more. [28] As a result of which, public goods are almost always under provisioned, overused and/or degraded and that is where nonprofit organizations step in and fill the gap. These organizations are financed by those who want to increase the provision of public goods (whether this is for personal benefit or simply because they are altruistic). [29] Nonprofit organizations are also not necessarily all that efficient, exploring this in detail is beyond the scope of this paper, but will be addressed briefly in Section 6.2.2.

In addition to governments not provisioning the economically efficient level of public goods, doing so via taxation is highly inefficient.

![Demand Curve](Figure 2.2.1(a))

This is a demand curve (Figure 2.2.1(a)). The y-axis represents the price of a good and the x-axis represents the quantity of the good. As price goes up, demand for the good goes down and vice versa ($P_A > P_B$ but $Q_A < Q_B$).
Figure 2.2.1(b)

This is a supply curve (Figure 2.2.1(b)). The y-axis represents the price of a good and the x-axis represents the quantity of the good. As price goes up, the quantity supplied goes up, and as price goes down, the quantity supplied also goes down ($P_A < P_B$ and $Q_A < Q_B$).

Figure 2.2.1(c)

Equilibrium (Figure 2.2.1(c)) is special because in a competitive free market, it is the only point at which there is only one price in the market i.e., sellers get exactly what consumers pay = $P$ (excluding sales tax, credit card transaction fees etc.)
However, when a tax (of value $T$) is implemented (Figure 2.2.1(d)), there are now two prices instead of a single equilibrium price. There is the price that consumers pay and the price that sellers receive. Since supply is more elastic than demand and due to “Concentrated Benefits and Diffused Costs” (more below), producers bear the smaller burden of the tax ($p_s + T = p_c$) where $p_c = $9 and $p_s = $4 and consumers bear the larger burden (disproportionately). There is also massive deadweight loss ($DWL$) due to the tax. Deadweight loss is (economic) surplus that is just lost due to taxation.

In the above example, a tax of $5 is levied on a good costing $6 at equilibrium. Once the tax is levied, consumer surplus ($CS$) and producer surplus ($PS$) decrease drastically because the quantity of the good produced drops as consumers now pay $9 and sellers only receive $4 (i.e., demand drops due to higher price and supply drops due to lower profit margins). A huge chunk of $CS$ and $PS$ is converted to tax revenue and the rest is lost to $DWL$. It is pretty clear that taxation, especially taxes such as sales tax, income tax etc. that are applied across the board indiscriminately are an incredibly inefficient way to raise money to provision public goods.

---

**Aside: Concentrated Benefits and Diffused Costs**

With almost all government intervention (taxation, cap and trade, price ceiling/floor etc.) in free markets, there is some deadweight loss and some amount of consumer and producer surplus that is captured by the intervention but that captured surplus almost always just gets redirected to producers. Why is that? It is simply because there are far fewer producers than there are consumers and so the benefit to each individual producer is so high that they are willing to spend money lobbying for that captured surplus to be redirected to them. The benefits to producers are extremely concentrated while those for individual consumers are so diffused/low that it is not worth the effort for consumers to lobby for that captured surplus to be redirected to them. This is a general truth in political economies which is that whenever there is a law that benefits a few but makes society worse off as a whole, then the few have a strong incentive to lobby for it and such a law will most likely get passed due to concentrated benefits and diffused costs.
However, there will always be those that argue that the only reason governments must tax so heavily and use tax money to provision public goods is because private markets have failed to provision these goods adequately and that government interventions are necessary to correct market failures (Section 2.2.3). Taking this argument to be valid at face value, one can assume that taxation is the best tool amongst a set of bad tools. However, in developing countries where nonprofits do not operate as widely, governments are not necessarily the most honest, and funding for public goods is scarce (whether that is due to lack of taxpayers or other reasons), the responsibility to provision public goods falls on communities. [25] The ability of most communities to successfully provision public goods depends on the size of the group, the power and influence of group members, tastes and preferences of group members, and the distribution of benefits among group members (i.e., factors previously discussed in Section 1.2). As expected, when these factors are not in favor of the group i.e., the group is too large, the benefits do not feel tangible to group members, influential group members do not value it as highly, they lack the financial prowess needed to provision a good etc. the collective action problem results in a lack of cooperation and thus provision of the public good.

**When not provisioned by the government or by non-profits, the responsibility to provision public goods falls on communities who lack the infrastructure and funding to do so.**

Aside: Coase Theorem

The above observation is actually more profound than it seems. In fact, the Coase Theorem proves that “public goods” are often not provisioned due to “transaction costs”.

Consider the following scenario. A rancher and farmer own side by side lands. The farmer faces a negative externality because the rancher’s cattle comes onto their land and eats their crops. The Pigouvian solution is for government intervention via taxation (Section 2.2.3.2) i.e., to tax the rancher based on the value of crops their cattle ate. However, practically, this need not happen.

- **Scenario A:** If the cost of putting up a fence ($x$) to keep the cattle out of the farmer’s land is less than taxation ($y$), the rancher will put up the fence. If the cost of the fence ($x$) is higher than the tax ($y$) then the rancher will just pay the tax instead.
- **Scenario B:** Lack of a local government means taxation is not possible. The farmer will put up the fence if the cost of putting up a fence ($x$) to prevent the cattle from eating their crops is less than the value of crops ($y$) lost. If the cost of the fence ($x$) is higher than the value of crops ($y$) lost, it is better to lose the crops and so the fence will not be put up.
Scenario C: Property rights are well defined for crops so the farmer can sue the rancher for damages equaling the cost of the damaged crops ($y). If the cost of putting up a fence ($x) to prevent the cattle from eating the farmer’s crops is less than the damages ($y) then the rancher will build the fence, else they will get sued and pay damages.

In all cases, whether or not building the fence is efficient comes down to simple math, the fence will be built when $x < $y i.e., the cost of the fence is less than the alternative cost incurred. Who pays for it is determined by property rights. So, property rights not only define a public good but also determine who may pay for it.

Coase Theorem Element 1: The assignment of (property) rights determines who benefits but does not affect reaching the efficient outcome.

Now, consider another scenario. The very one that was discussed in the introduction. A coal power plant pollutes a neighboring town causing $y worth of damage. A carbon scrubber could be installed eliminating that damage and the scrubber costs $x. Once again, the Pigouvian solution is for government intervention via taxation (Section 2.2.3.2) i.e., to tax the power plant based on the value of damages. However, once again, practically, this need not happen.

Scenario A: If the cost of installing a scrubber ($x) to clean the air is less than taxation ($y), the power plant will install the scrubber. If the cost of the scrubber ($x) is higher than the tax ($y) then the power plant will just pay the tax instead.

Scenario B: Lack of a local government means taxation is not possible. The town will install the scrubber themselves if the cost of the scrubber ($x) is less than the damages ($y). If the cost of the scrubber ($x) is higher than the damages ($y), it is better to incur the damages and hence the scrubber will not be installed.

Scenario C: Property rights are well defined so the town can sue the power plant for damages ($y). If the cost of installing the scrubber ($x) is less than the damages ($y) then the power plant will install the scrubber, else they will get sued and pay damages.

The “damages” here is referring to the value that the town places on clean air, or rather the price they think they are paying because of unclean air. Once again, whether or not installing the scrubber is efficient comes down to simple math, the scrubber will be installed when $x < $y but in a situation like this, transaction costs are large. The costs associated with getting everyone to coordinate to volunteer money for the scrubber, to agree on what costs they believe they are incurring as a result of the polluted air, the costs of ensuring that the scrubber is actually installed once the money is collected, the issues that arise due to some people wanting to deliberately free ride etc. are being ignored.

Coase Theorem Element 2: Externalities are frequently a problem of transaction costs. Property rights assignment can affect efficiency through the mechanism of affecting transaction costs.
2.2.2 Free Markets

If the onus of provisioning public goods falls solely on governments, then public goods will not just be provisioned inadequately but will be done in an incredibly inefficient manner - taxation (as previously discussed). However, that discussion presupposed that markets when left to their own devices operate in an efficient manner. This fact should seem obvious from the prevalence of markets but an illustration of the manner in which it achieves this can be useful to drive home the point. The economic framework, at its core, is about making the world a better place by having more “good stuff” and less “bad stuff”. “Good stuff” is stuff that a person is willing to sacrifice to have. “Bad stuff” is stuff that a person is willing to sacrifice to avoid. Assuming the amount of sacrifice someone is willing to make can be measured, if people are willing to sacrifice more to have something than other people are willing to sacrifice to avoid it, then it is “good stuff” and if people are willing to sacrifice more to avoid something than other people are willing to sacrifice to have it then it is “bad stuff”.

Most things are good for some and bad for others and this can be measured based on the amount of sacrifice a person is willing to make. The idea of being able to measure and compare sacrifices presupposes that sacrifices can be valued objectively. It also assumes that the system for aggregating sacrifices over the entire population does so in a manner that accurately values each individual’s perception of what their sacrifice is worth. This idea generally draws further criticism when the amount of sacrifice is measured in dollar terms.

There are two very common objections to the claims made above, one of which is valid and the other not so much.

● The first objection is that there are just certain things that cannot have a price tag slapped on them. Those objecting, fail to see that dollars are just a unit of measurement. Any unit of measurement including “number of happy days” would work equally well but dollars are convenient, generally universally understood and act as a sufficiently good proxy for whatever utility mechanism most people have. For example, people love to praise the “completely free” British healthcare system but a huge part of how (and why) it works is that they put a dollar (or more accurately, pound) value on “a quality adjusted extra year of life” (QALY) [59]. This is how they decide where to allocate resources. If Alice’s QALY is higher than Bob’s, then Alice will receive treatment over Bob. The truth is that dollar values are assigned to things all the time and even though it may not be “nice”, it is necessary.

● The second objection is that monetary sacrifices vary greatly based on net worth and income of the individuals under consideration. This is absolutely true. Someone whose annual income is $100,000 is going to be far more reluctant about spending $10,000 (and hence if they do, it is worth much more to them) than someone whose income is $100,000,000. Finland has a very interesting system of dynamically priced fines based on income levels. [101] But at the same time, $10,000 is $10,000 no matter where the
money comes from and those $10,000 will still do the same amount of good. So, there is an objective value to what the money is capable of doing despite it likely having far higher value for some than others. This is a valid concern, and it does mean that a dollar-based valuation system is not perfect, but a compromise must be made somewhere and if that means accepting that the valuation system is imperfect rather than having no valuation system at all, it seems like a reasonable compromise.

With that out of the way, returning to “good stuff” and “bad stuff”. If people sacrifice exactly the amount they are willing to, in exchange for good stuff, then the bad stuff “cancels out” the good stuff and the world is not better off. The world is better off when someone gets good stuff for less than the maximum amount they are willing to sacrifice. The sum of the total sacrifice people made to get good stuff subtracted from the sum of the total sacrifice people were willing to make to get good stuff is social surplus. The higher the social surplus, the better; and markets can be thought of as machines that maximize social surplus.

![Figure 2.2.2(a)](image)

As previously mentioned, this is a demand curve (*Figure 2.2.2(a)*). Another interesting way of thinking about a demand curve is that it is a marginal benefit (MB) curve i.e., every point on the curve represents the maximum price someone is willing to pay (or sacrifice) for that good. $A$ is the maximum price that the Quantity ($A$)th ($Q_A$) person (assuming each person buys exactly one unit) is willing to pay for the good, and the same goes for point $B$. 

As previously mentioned, this is a supply curve (Figure 2.2.2(b)). Another interesting way of thinking about a supply curve is that it is a marginal cost (MC) curve i.e., every point on the curve represents the lowest price someone is willing to accept (or sacrifice) to provide a unit of that good. Points A and B are the equivalent marginal cost analogues to those discussed in the demand curve.

Combining the curves previously discussed, results in Figure 2.2.2(c) (that was briefly mentioned earlier when discussing taxation and deadweight loss) that illustrates exactly what markets do. They reach equilibrium, which generally happens to be economically efficient (or close enough) [61] and that is the point at which social surplus is maximized. Why is this the case? This is because at all points lower than the equilibrium point, marginal benefit is greater than marginal cost so production is increased until the point where they are equal and beyond that point marginal cost is greater than marginal benefit, so it does not make sense to produce any more. It is important to acknowledge that equilibrium is when social surplus is maximized but consumer or producer surplus can be further maximized through market interventions.
For example, establishing a price ceiling (Figure 2.2.2(d)), a type of market intervention, results in consumer surplus increasing \((A + B \rightarrow A + C)\) but producer surplus decreasing \((C + D + E \rightarrow E)\) and deadweight loss i.e., surplus that either consumers \((B)\) or producers \((D)\) could have had but now cannot due to the market intervention. Similar to deadweight loss in taxation, this is economic surplus that is lost due to interfering with a naturally efficient market since \(Q_{\text{new}}\) is being produced now instead of \(Q\).

This shows how markets are like machines that help maximize social surplus when they reach equilibrium (and as previously mentioned, per Adam Smith’s invisible hand, they will reach equilibrium without market intervention). Without getting into mathematical details, the crux of it is that markets do this by taking extremely complex problems and boiling them down to multiple parts, each of which is quite simple and independent of the others; and each stakeholder in the process only needs to worry about their own part and making that process as efficient as possible, but they get to share in the profits from all parts working well together. “The Toaster Project” is a great book that illustrates the power of markets through a simple and practical example of attempting to make a toaster from scratch that ends up costing 100x more than buying a toaster from a store (and also does not work, very well, if at all).

So, it has been established that markets are very powerful. But what has really been established is that markets are powerful at being very efficient. Is that the best metric to determine if a market is powerful? What about equity? In economics, academically or otherwise, efficiency (maximizing total amount of social surplus) and equity (distributing social surplus among groups in society in the best way) are both important but efficiency always gets prioritized. Why is that? It most certainly is not an evil plot by the rich to help them get richer but rather that measuring, and thus achieving, efficiency is far easier than equity. It is as simple as, more social surplus is better and less social surplus is worse, or in the words of Jeremy Bentham himself, “it is the greatest happiness of the greatest number that is the measure of right and wrong”. But equity asks the question of what is the best (not most fair) distribution of social surplus? This is a much harder and far more nuanced question. One without a great answer; and so, focusing on maximizing social surplus first and then redistributing if the resulting distribution is not optimal is the compromise made.
Aside: Efficiency

When thinking about maximizing social surplus, there are two kinds of improvements that can be made. The first is called a Pareto improvement. This is a change that makes at least one person better off and harms no one i.e., unambiguous “good stuff”. In other words, Pareto efficiency/optimality is a situation where no action or allocation is available that makes one individual better off without making another worse off. [34] Competitive free markets are Pareto efficient but for goods that do not have such markets, Pareto improvements are quite unrealistic and rare, especially if the improvements need to be applied at scale. This is the case for public goods.

Specifically, the Pareto optimal provision of a public good is defined as a situation in which the sum of the marginal value of the public good (taken across all individuals that will use that public good) is equal to the marginal cost of providing that public good (Samuelson condition). Formally, the marginal value of the public good is calculated by the marginal rate of substitution (MRS) relative to a reference private good (trivially, this would be money) i.e., how much of the private good is someone willing to exchange for the public good while maintaining the same level of utility; and the marginal cost of the public good is given by the margin rate of transformation (MRT) relative to the same reference private good i.e., how much of that private good it costs to produce an incremental unit of the public good. The Pareto optimality condition for a public good is different from that of a private good for which it states that the consumer’s valuation of the private good should be equal to its marginal cost of production. [35][36]

The second improvement is called a Kaldor-Hicks improvement, also called a potential Pareto improvement. This is a change that makes some better off and some worse off but the gains to the gainers are greater than the losses to the losers. It is called a potential Pareto improvement because Kaldor-Hicks improvements can be made Pareto improvements by offsetting the losses to the losers. This is, however, quite hard to do in practice because it relies on accurate quantification of those losses.

A policy/the world is Pareto efficient, if and only if, no member of society could be made better off by an alternative policy without making at least one person worse off. A policy/the world is Kaldor-Hicks efficient if it would be Pareto efficient provided that the losers are compensated for their losses, but the actual compensation need not take place, it just needs to be possible. The world must become more Kaldor-Hicks efficient, and eventually, more Pareto efficient.

It is important to note that these efficiencies are defined with the assumption that they are happening under idealized conditions of perfect information. [37] This implies that Pareto (or Kaldor-Hicks) efficiency for public goods is virtually impossible to achieve in practice. [35]
Aside: Time Value of Money

Any rational actor will prefer $100 right now to $100 in a year but there is a certain amount of money such that the same rational actor will prefer $x today instead of $100 in a year (where $x < 100). This is known as the time value of money. If the market risk-free interest rate (generally the rate of return on default-free U.S. Treasury Bonds) is $r$ then the present value of $x$ in $n$ years is $\frac{x}{(1+r)^n}$. This is primarily because money grows when kept in financial instruments and financial markets work on the basic principle of those who are patient, lending to the impatient, in exchange for a fee (interest).

In the prior example, the following formula provides the value of $x$ such that a rational actor would prefer $x$ today instead of $100 \frac{x}{(1+r)^n} = \frac{100}{(1+r)^n}$ and assuming that $r$ is 3% then $x$ is approximately $97$. This calculation seems simple enough, but what about the present value of $xa/year for the next $n$ years? Why would this be useful information? Depending on the way a person views life, their circumstances, the kind of public good in question etc. they may find it hard to put a value on something. For example, imagine a family in a neighborhood raising money to build a park for the children in the neighborhood to use. The family knows that their only kid would outgrow the park in three years. This means that if they could intuitively put a dollar value on what the park is worth to them/their kid each year then they could accurately value what it is worth to them today given they would only use it for three years.

Generalizing the formula used above yields the following formula $\sum_{i=1}^{n} \frac{x}{(1+r)^i} = \frac{x}{r} \left(1 - \frac{1}{(1+r)^n}\right)$ which gives the present value of $xa/year for the next $n$ years. So, if someone wants to find out what $xa/year for the next $n$ years is worth today, they can use this formula. If the park was worth $100/year to the family and the risk-free rate ($r$) was 3% then the present value would be approximately $282. Earlier, it was established that any “rational” actor would have this preference, however, all people are not rational. In addition to which, even if they were all rational, it would be incredibly rare for them to perform this kind of “rigorous” mathematical analysis on what something is worth to them (the theory of bounded rationality) [102]. However, this is the only way to arrive at a person’s true valuation for a public good.

The value $r$ which is the market risk-free interest rate is also known as the discounting rate. The choice of the discounting rate is extremely important. Choosing to spend money on something now is equivalent to making an investment and the return on that investment is important. The return could be monetary (when investing in the stock market) but it also need not be (when investing in public goods). An example of the importance of the discounting rate is that if an asteroid would destroy Florida in two centuries and it was going to be prevented then society would spend $3 million to do so at a 7% discount rate but would spend $43 billion at a 2% discount rate.
So, a slightly more nuanced definition of what markets are doing is that they are maximizing social surplus but are doing so when defining social surplus as the difference between total present discounted benefits and total present discounted costs.

Despite having established that markets maximize social surplus and are very powerful at doing that, there are some valid critiques of this model. The first is that discounting privileges current generations. The second is that simply maximizing social surplus disregards moral responsibility. The third is that equity is being ignored both during valuation and during distribution. The fourth is that people are not rational and may not be the best judges of their own well-being (or rather, show bounded rationality) [102]. The fifth is that people’s tastes are not fixed but are rather learned and malleable. These are valid critiques with the system but do not affect its effectiveness. However, there are a class of problems that do affect the system’s effectiveness, and those are discussed in the following section.

2.2.3 Market Failures

So, if markets are so great (despite the valid critiques stated above), then why can their power not be harnessed to provision public goods? Why do markets not do a great job of providing public goods? The reason they do not is because it is very hard to create a competitive free market for public goods due to certain kinds of market failures that prevent markets from reaching efficient equilibriums without any market intervention; and when there is a market intervention, as previously discussed, there is almost always deadweight loss.

2.2.3.1 Valuation

For markets to be able to function efficiently public goods need to be priced. This is hard.

The Related Goods Method is one way of doing so. The value of national parks can be estimated based on how much someone is willing to spend to get there and on accommodation, hiking gear etc. The value of pollution-free air can be estimated by comparing housing prices in polluted vs non-polluted areas adjusting for other differences (such as schools, safety, proximity to grocery stores etc.) using hedonic pricing adjustments. The biggest advantage of this method is that it performs valuations based on actual actions, however someone needs to do the work to calculate the value of a public good (this will vary based on the specific circumstances in which the public good is being provisioned), and this method cannot be applied to everything.

The Stated Preference Method is a more straightforward way of valuing public goods. It involves just asking people what they value things at and their willingness to pay for things. The biggest advantage here is that it can be done for anything but there is no actual transaction taking place i.e., there are no real stakes and people do not always do a great job of verbalizing their actual choices especially when there is fear of public judgment. They
tend to overestimate/underestimate what things are worth to them and then do not follow through on their estimations when it comes to actually paying up. [35][60]

People will underestimate or overestimate the value of a public good depending on the circumstances in which they are asked to provide a valuation.

While the lack of a valuation technique that works for public goods may not be formally classified as a market failure in economics literature, it is one of the primary reasons competitive free markets for public goods do not exist.

2.2.3.2 Externalities

An externality occurs when the actions of some actor (individual or firm) have a direct, unintentional, and uncompensated effect on the well-being of other individuals or the profits of other firms.

Consider the market for gasoline (Figure 2.2.3(a)). The demand curve is represented by the marginal private benefit (MPB) curve which is the benefit received by individual participants in the market. In the case of gasoline, it also happens to be the marginal social benefit (MSB) curve which is the benefit received by society as a whole. The supply curve is represented by the marginal private costs (MPC) curve which is the marginal cost for individual participants in the market. However, there is also a marginal social cost curve (MSC) which is the marginal cost for society as a whole. The MSC and MPC curves differ because the use of gasoline produces harmful pollutants and greenhouse gasses that cause global warming, asthma, allergies etc. i.e., the cost incurred by society as a whole is higher than that incurred by individuals in isolation because of this negative externality. Graphically, the negative externality is highlighted by the portion between the two curves.
The equilibrium market outcome $Q_m$ is, of course, given by the intersection of the $MPC$ and $MPB$ curves, however, the truly efficient quantity $Q_e$ is given by the intersection of the $MSC$ and $MSB$ curves. This results in deadweight loss ($DWL$) (Figure 2.2.3(b)). The efficient quantity is not zero because people get a lot of benefits from the use of gasoline and up to a certain point, those benefits outweigh the negative side-effects (externalities) of using gasoline.

This kind of $DWL$ (Figure 2.2.3(b)) is different from that which comes with a market intervention like a price ceiling (Figure 2.2.3(c)). That $DWL$ is due to underproduction but the $DWL$ here is due to overproduction. People do get benefits from using that additional gasoline but the marginal benefit of those additional units of gasoline is lower than the marginal cost of using it, and that difference is precisely the negative externality.
The same goes for positive externalities (Figure 2.2.3(d)). Consider the market for bees. The demand curve is represented by the marginal private benefit (MPB) curve which is the benefit received by individual participants in the market. However, there is also a marginal social benefit (MSB) curve which is the marginal benefit received by society as a whole. The supply curve is represented by the marginal private costs (MPC) curve which is the marginal cost for individual participants in the market. In the case of bees, it also happens to be the marginal social cost curve (MSC) which is the marginal cost incurred by society as a whole. The MSB and MPB curves differ because bees not only produce honey but also pollinate flowers (allowing fruits and vegetables to grow) i.e., the private benefit is the honey, but the pollination of flowers is a positive externality. Graphically, the positive externality is highlighted by the portion between the two curves.

The equilibrium market outcome = $Q_m$ is, of course, given by the intersection of the MPC and MPB curves, however, the truly efficient quantity = $Q_e$ is given by the intersection of the MSC and MSB curves. This results in deadweight loss (DWL) (Figure 2.2.3(e)). The efficient quantity is more than what is produced by a free-market equilibrium because bees have a positive externality.
The presence of externalities that are not remunerated is a market failure.

Figure 2.2.3(f)

To make this crystal clear, as this is one of the biggest reasons why markets do not efficiently provision public goods, the gasoline example is worth revisiting. In Figure 2.2.3(f) (which is just an alternative depiction of Figure 2.2.3(b)) consumer surplus is given by $A + B + C$ and producer surplus is given by $D + E + F$, the deadweight loss is $G$ and the cost of the externality is $B + E + C + F + G$ but if the efficient quantity $Q_e$ were produced instead, $C$ would be lost from consumer surplus and $F$ from producer surplus; but there would be no deadweight loss and the cost of the externality actually decreases by that amount. The “lost” producer and consumer surplus is just the gains of the winners being transferred to the losers (i.e., those affected by the externality), so there is no loss, just an elimination of $DWL$. Formally, externalities can be internalized by taxation using Pigouvian taxes.

Figure 2.2.3(g)

A Pigouvian tax (Figure 2.2.3(g)) is a tax equal to the per unit cost of the externality at the efficient quantity; thereby internalizing the externality and resulting in the efficient quantity being produced in a free market. It is
the only tax with no *DWL* as the region that is traditionally *DWL* is actually just a portion of the externality that is now being eliminated.

### 2.2.3.3 Free Rider Problem

Public goods are also subject to the “free rider problem” because they are non-excludable. [22] This results in public goods being underproduced, overused and/or degraded. [21] The individual incentive to free ride conflicts with the collective benefit of the group if everyone participated honestly and fairly. So, the lack of provision of public goods as a result of the “free rider problem” due to their non-excludable nature is a collective action problem. The free rider problem is a known example of a market failure i.e., a situation in which individual gain seeking behavior does not produce economically efficient results, and this is because public goods have positive externalities that are not remunerated.

![Figure 2.2.3(g)](image)

Consider the following scenario (*Figure 2.2.3(g)*). Alice and Bob are roommates and share a flower garden. The garden is a public good because they each get non-rival benefits from it i.e., they can both enjoy the garden without taking away from each other’s enjoyment; and neither can exclude the other from enjoying it since it is a common garden i.e., it is non-excludable.

The benefit that Alice gets from the garden is given by *MB_A* and the benefit that Bob gets from the garden is given by *MB_B*. Adding their marginal benefit curves yields *MB_{A+B} = MSB*. Thus, the socially efficient quantity is *Q* (intersection of *MSB* and *MC* curves) but neither Alice nor Bob is individually incentivized to take care of the garden enough to get to that quantity. What ends up happening is that if Alice were the only person taking care of the garden, then Alice would put in *q_A* amount of effort; and if Bob were the only person taking care of the garden, then Bob would put in *q_B* amount of effort. Thus, when both are taking care of the garden, Bob will put in *q_B* effort and Alice gets the extra benefit *q_B* – *q_A* that is not remunerated i.e., Alice experiences a positive externality.
<table>
<thead>
<tr>
<th></th>
<th>Alice</th>
<th>Bob</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Contribute</td>
<td>Defect</td>
</tr>
<tr>
<td>Contribute</td>
<td>Alice → $400 - $300 = $100</td>
<td>Alice → $200 - $300 = -$100</td>
</tr>
<tr>
<td></td>
<td>Bob → $400 - $300 = $100</td>
<td>Bob → $200</td>
</tr>
<tr>
<td>Defect</td>
<td>Alice → $200</td>
<td>Alice → $0</td>
</tr>
<tr>
<td></td>
<td>Bob → $200 - $300 = -$100</td>
<td>Bob → $0</td>
</tr>
</tbody>
</table>

What is interesting is that when this plays out in real life, even Bob does not provision \( q_B \). Suppose that each person who works in the garden increases its benefits by $200 but they incur a cost of $300 to put in that work. Then the Pareto efficient and Kaldor-Hicks efficient outcome would be for both to contribute, resulting in the benefit totaling $400 and the cost to each being $300 thus resulting in a $100 net benefit to each. However, as previously discussed in the prisoner’s dilemma payoff matrix, something similar takes place here and the Nash equilibrium (situation in which neither player can improve their payoff by only changing their action) ends up being the situation wherein neither one contributes to taking care of the garden.

The free rider problem with public goods (especially at scale) can be more generally defined as a public goods dilemma. This is a situation in which the whole group can benefit if a certain number of members contribute but individuals will benefit from “free riding” i.e., not contributing but simply benefiting, as long as enough members contribute. [24] The economic theory approach of explaining this phenomenon is called the “free rider problem” and is broadly applicable but this is also explained in social psychology literature as social loafing. [23] Social loafing specifically refers to the tendency for people to exert less effort when in a group than when working alone.

The canonical example of the free rider problem at scale is when a town is considering putting on a fireworks show. The town has 500 residents and the show costs $5000. If every resident is asked to contribute $10 to fund the show, then it will almost always never take place (i.e., a voluntary “tax”). This is because everyone can view the show for free from their houses, so many would choose to not pay the $10 even if it is worth that much to them, choosing to free ride instead. There is also the matter of households for whom a fireworks show may be worth $0 (i.e., they are indifferent) and they will also choose to not contribute. It could be worth less than $10 to them, but as long as it is worth more than $0, they would still be free riding if they did not contribute. It could even be inflicting a negative externality on them. At the same time, there are households, say those with young kids, to whom the fireworks show may have been worth $20 or even $40, but because of the way any kind of taxation is set up, they cannot contribute more to it even if they wanted to. In fact, the best way to imagine this situation is that it is one with multiple payoff matrices within the community, many of which result in Nash equilibriums wherein people end up not paying and thus free riding. This results in free riders preventing the show from happening, but a more subtle problem is the inability to collectively organize around this problem and
solve it easily (i.e., the fact that even if the fireworks show were worth more than $5000 to the town collectively it would not happen).

**Communities lack the infrastructure to collectively organizing around problems they care about and solve them.**

The free rider problem is further complicated by the demand revelation problem i.e., when non-excludability of a good results in a failure to pay the true marginal value, it simultaneously results in failure to generate proper income levels since households would not be willing to give up valuable leisure if they cannot individually increment a good. [33] This implies that if a public good does not have strong special interest support, then it will more than likely be under provisioned since cost-benefit analysis for that good would be conducted at incorrect income levels and all income would have been spent on the public good, apart from general equilibrium considerations.

**Public goods without strong organized special interest support generally never get provisioned due to the demand revelation problem.**

### 2.3 Local vs Global

Global public goods may be defined as those from which anyone in the world can derive value at any time. This includes things like knowledge, national defense, clean air, property rights etc. Local public goods are those for which the benefits accrue more to a specific group of people (in a neighborhood) than the general public. Public goods have been modeled differently in light of this information. Specifically, it has been noticed that these neighborhoods tend to overlap and the effects of this are not negligible and thus groups of overlapping neighborhoods have been modeled as networks. [38] Individual neighborhoods continue to be modeled via the Tiebout model (i.e., competitive pressures in local governments could result in provision of optimal levels of public goods). This distinction is primarily relevant to thinking about under provision of public goods and selective steps governments can take to combat that, however, it is not relevant to the system that will be proposed in this paper (at least in the early stages) and is simply mentioned briefly for the sake of comprehensiveness. A future addendum to the proposed system could account for networks, but at this time, for the sake of brevity, it will not be addressed.

### 2.4 Ownership

With private goods, it is understood that the party with the better investment technology should be the owner. [39] However, with public goods, it is not so clear. Some argue that the party with the larger valuation of the public good must be the owner and that the investment technology does not matter. [40] However, there are also those that argue that the investment technology should continue to remain an important determinant of the
optimal ownership structure of the public good. [41] The system that will be proposed in this paper will certainly make ownership an interesting point of contention and it is one that shall be revisited (Section 6.5.4).

2.5 Summary of Observations

People have a natural tendency to free ride and generally require a strong deterrent (such as law enforcement) to discourage them from doing so.

When not provisioned by the government or by non-profits, the responsibility to provision public goods falls on communities who lack the infrastructure and funding to do so.

Transaction costs are generally sufficiently large to prevent public goods from being provisioned but can be reduced through appropriate assignment of property rights.

It is hard for even the most honest person to accurately value what something like a public good is worth to them.

People will underestimate or overestimate the value of a public good depending on the circumstances in which they are asked to provide a valuation.

Communities lack the infrastructure to collectively organizing around problems they care about and solve them.

Public goods without strong organized special interest support generally never get provisioned due to the demand revelation problem.

In these observations, there are two common properties, adding onto the two properties previously observed:

Property 3: The responsibility to provision public goods often falls on communities who lack the infrastructure and funding to provision public goods themselves. Without infrastructure that helps them collectively organize around issues they care about, the transaction costs involved with provisioning public goods to address those issues is too high and cannot be met without special interest support.
Property 4: People will underestimate or overestimate what a public good is worth to them when asked to put a dollar value on it, and may do so intentionally, but even the most honest people will end up doing so unintentionally. Most people may be rational, but their rationality is bounded. They need to “put their money where their mouth is”, and there is simply no substitute for doing so.

3 Markets

3.1 “For”

Despite having talked about markets in such detail, a consensus on whether or not they are the best tool for the problem at hand i.e., solving collective action problems to provision adequate (i.e., the economically efficient) level of public goods, has not been reached. In essence, there are two primary ways to get people to coordinate. Force i.e., “do this or I will hurt you” and Markets i.e., “do this because I will pay you”. There is also a third way of doing so and that is Legitimacy i.e., “do this because it is right, do this for god, king, country, people, status etc.”

These three ways are not independent of each other but rather work together. For example, people are told, follow the law because they are good and following the law is good (Legitimacy), this lets them work and make money (Market) but if they do not follow the law then they will pay a fine or be sent to prison (Force). There are drawbacks to each of these techniques. Force is extremely resource intensive. Significant resources are required to back one’s threats. It is not exactly pleasant to think about and is the least noble of the three (a slippery slope toward tyranny and dictatorships). Legitimacy is extremely time intensive. To change people’s minds and appeal to the goodness in their hearts and their humanity is a long-term project, it takes time. It is the most noble, but it also lacks “teeth”, simply asking people to do something (a favor). Markets are hard to control, can be a little sloppy and can fail, however they are in the middle when it comes to nobility, people are not exactly being forced to do anything but at the same time, they do not like putting prices on everything.

Despite this, of the three, Markets are probably the best and most powerful solution because they are the best at getting people to coordinate since the incentive (money) is very strong. They are particularly powerful because they act as an information creation and coordinating function of the price system i.e., prices generate information about preferences and scarcities (which is otherwise very hard to gather). [82] This mechanism fails to operate when there is a government intervention and it provides the good instead, as is the case with public goods today. When the government intervenes with free markets, and once a public good is provisioned by the government, it is hard to tell how much people really value it (due to previously discussed limitations to consumers’ consumption decisions in Section 2.1), how much to invest in it each year, whether to keep it running, and/or to invest in R&D for newer and more advanced forms of the good. Markets have also proven to be widely successful throughout history, and with a little bit of Legitimacy and Force thrown into the mix, it could just produce a winning solution.
3.1 “Against”

By this point, it should be clear that Markets are the best solution, but just to play Devil’s advocate, evaluating some of the most popular arguments against the use of markets to privately provision public goods can be insightful. At its core, the argument that governments are justified in provisioning public goods via taxation is rooted in private markets’ inability to provision sufficient levels of public goods. As previously hinted at, this is likely because markets have never been given a chance. Regardless, with that out of the way, exploring some of the more nuanced takes on why public goods should not be provisioned by markets can be insightful.

Some argue that the government should provision certain public goods, regardless of whether private markets are capable of provisioning adequate quantities of those same goods. Anderson argues that the sphere of the market should be limited because market norms do not embody certain important values and market exchange may undermine ideals and interests legitimately protected by the state. [83] According to Anderson, market norms have the following five features: they are impersonal (independent of the relationship and of the other party’s ends), egoistic (the goal is to satisfy one’s personal interests), exclusive (people can be excluded), want-regarding (as opposed to responsive to objective needs), and oriented to “exit” rather than “voice” (rather than voicing one’s complaint, one seeks an alternative). [83] An economic good is one whose production, distribution, and enjoyment are properly governed by market norms. Public goods actually fall into Anderson’s category of “political goods”. They are characterized by three norms that oppose the market norms. First, one’s freedom is exercised through voice rather than exit. Second, goods are distributed according to public principles rather than subjective wants. Third, the goods are provided on a non-exclusive basis. She says “Some goods can be secured only through a form of democratic provision that is nonexclusive, principle- and need-regarding, and regulated primarily through voice. To attempt to provide these goods through market mechanisms is to undermine our capacity to value and realize ourselves as fraternal democratic citizens.”

While that may have sounded a tad dramatic, in a similar vein, Kallholff argues that some public goods have externalities that are essential to the functioning of a democracy and hence need to be provisioned by the government. She says “Some public goods constitute visible expressions of solidarity and social justice among citizens (“central goods”), some support connectivity and serve as representations of shared interests (“connectivity goods”), and some serve as visible representations of a shared sense of citizenship (“identification goods”). The absence of these goods would undermine the ability of the citizenry to engage in public deliberation and develop a sense of self-determination.” [84] More practically, a lot of contemporary economic theory has emphasized the importance of increasing returns, especially through investment in innovation, knowledge, and development [103]. A lot of modern urban development and theory that supports it, comes from literature on agglomeration and spatial economics that emphasize that increasing returns are what allow cities that created the idea of the “middle class citizen” to exist and prosper today. [104] However, despite its importance, modern capitalistic societies are not well suited to provide public goods to the extent one would define as optimal due to the previously discussed market failures.
To clarify, no kind of market/system for privately funding public goods (including and especially the one that will be proposed in the paper) would act as a replacement to government funding (at least not for the foreseeable future); but rather it will help fill the obvious gap and could help solve collective action problems in a non-resource intensive manner.

4 Solutions

4.1 Mechanism Design

In Section 2.2.3.1, it was established that the lack of a valuation mechanism is one of the primary reasons free markets cannot be used to provision the efficient quantity of public goods. Mechanism Design is the art of designing institutional arrangements that deliver the required/optimal outcomes (i.e., for the system that will be proposed in this paper, mechanisms that encourage individuals to reveal their true valuations of public goods) instead of attempting to take the arrangement as a given and incentivize players’ behavior to deliver the required/optimal outcome. The goal is to design a mechanism that holds the following properties (such a mechanism is defined as “optimal” henceforth).

Property 1: The mechanism is efficient i.e., it always delivers the Pareto efficient outcome.
Property 2: The mechanism is incentive-compatible i.e., all players have an incentive to be truthful.
Property 3: The mechanism is budget-balanced i.e., it never runs a deficit and raises the required capital.
Property 4: The incentive-compatible mechanism is direct i.e., lying is weakly dominated by telling the truth.

It is also known that if there exists a mechanism that delivers a certain outcome, then there also exists a direct incentive-compatible mechanism that delivers the same outcome. This is known as the revelation principle. The resulting implication is that the optimal mechanism will be found in the set of all direct incentive-compatible budget-balanced mechanisms that are efficient (i.e., those that hold all four properties listed above).

4.1.1 The Vickrey-Clarke-Groves Mechanism

The Vickrey-Clarke-Groves Mechanism is the only mechanism that gets close to optimality as defined earlier (i.e., it is one that is direct, incentive-compatible, and efficient but not budget-balanced). How does it achieve this? Suppose a public good that costs $x to be provisioned needs to be funded i.e., the quantity of the public good provisioned will be proportional to $x. Let the value, for a given individual i, of the public good being provisioned be \( v_i(x) \) where \( v_i \) is a quasi-linear utility function (i.e., it is concave, differentiable, and monotonically increasing) such that \( v_i(0) = 0 \). This condition for the utility function is required because the impossibility results of Gibbard [85] and Satterthwaite [86], show that if every possible preference is admissible, and if the mechanism must always have a dominant strategy, then the outcomes chosen must be exactly those chosen by simply assigning one agent to be a dictator; quasi-linear utility functions address this concern. The efficient level of
funding for the public good is $\alpha = [\sum_i v_i(x)] - x$ that maximizes social surplus = $\alpha$. However, due to the free rider problem, people have no incentive to be honest about their $v_i(x)$. So, to address this, the VCG Mechanism asks everyone for their $v_i(x)$ and then finds the $\alpha$ that maximizes $\alpha = [\sum_i v_i(x)] - x$. Then, every individual is given $[\sum_{j \neq i} v_j(\alpha)] - \alpha$. This value could be positive or negative (if negative, it acts like a tax). When individuals add their own utility $v_i(x)$ to the transfer $[\sum_{j \neq i} v_j(\alpha)] - \alpha$ then their personal payoff is the “same” as society’s. Hence individuals are incentivized to tell the truth since they want $\alpha$ to be computed correctly. This shows that the VCG mechanism is always incentive compatible and direct. [105] These side-payments or transfers (i.e., $[\sum_{j \neq i} v_j(\alpha)] - \alpha$) are very expensive. However, they can be made much cheaper by eliminating constants from the transfers i.e., by eliminating any function independent of $v_i$. When the ideal function $(\max_x [\sum_{j \neq i} v_j(\alpha)] - \alpha)$ is chosen then the amount raised from each individual is $(\max_x [\sum_{j \neq i} v_j(x)] - x) \cdot (([\sum_{j \neq i} v_j(\alpha)] - \alpha) > 0$ i.e., each individual needs to contribute at least as much as their side-payment (which is inefficient); but at the same time, there is no incentive compatible mechanism that raises more total funds. [87] Thus, it is efficient as well. [105] However, there are several issues with VCG. The first is that it requires too much information. Every individual is required to submit their exact valuation for each level of total funding for every public good (i.e., $v_i(x)$ needs to be defined for all $x$ or at least all $x$ in some finite set that is the range of all possible values of $\alpha$). This is simply too much work. The second is that it fails if individuals can collude. With anonymity, the ways to collude and cheat in VCG are generally NP-hard to compute so this issue is somewhat overblown. The third is that it can end up raising less than $\alpha$ and would require some kind of subsidization/taxation to compensate for the funding gap (the proof of this fact is beyond the scope of this paper though one could work out a simple example where this holds). This last issue is particularly interesting and profound because it has been proven that there is no mechanism that can raise more funds than VCG. [87] This implies that there is no direct incentive-compatible efficient mechanism to raise funds for public goods that never runs a deficit.

There is no direct incentive-compatible efficient (i.e., first-best) mechanism to raise funds for public goods that never runs a deficit.

So, it has just been proven that the first-best mechanism (i.e., the one that satisfies all required conditions without being bound to any constraints) does not exist. In fact, in addition to not satisfying the required conditions, when tested practically, only certain variants of VCG (that had even weaker guarantees and sometimes dropped one of the three guarantees entirely) succeeded, [106] and it was later shown that any variant of VCG (even weakened variants) holding all properties was generally impractical and could not be implemented in practice. [107] The only second-best mechanism (i.e., the one that satisfies all required conditions with some constraints but cannot be outperformed by another mechanism with the same constraints) is a fixed-split mechanism (i.e., taxation). The proof of this fact is once again beyond the scope of this paper.

The only second-best mechanism to raise funds for public goods is a fixed-split mechanism (i.e., taxation).
4.1.2 Quadratic Finance

After VCG, a few methods [108][109] were proposed in response to the above results being proven that suggested a quadratic mechanism for determining the efficient level of public goods using near-optimal collective decision-making. However, these methods required a centralized iterative process or required a strong assumption of complete information, and they do not, in general, satisfy individual rationality. There has not been much further work in the field of mechanism design since the 70s where the aforementioned results were proven.

The primary insight of quadratic pricing of collective choices re-emerged recently via the idea of “Quadratic Vote Buying” [110][111] and showed that with certain fairly relaxed assumptions it led to approximately optimal decisions on public goods. A recent paper in this space co-authored by Vitalik Buterin further built on this idea. He is especially passionate about the use of blockchains as a means of decentralized funding and endogenous community formation to provision public goods and talks about an idea known as Quadratic Finance (QF) [88]. QF solves the first issue with VCG by allocating resources according to $\alpha = \left(\sum_i c_i / \sqrt{c_i}\right)^2$ where $c_i$ is the contribution made by each individual. This implies that participants no longer need to submit their exact valuation for each level of total funding for every public good (i.e., $v_i(x)$ does not need to be defined for all $x$).

It can be shown that QF results in the same efficient allocation $\alpha$ that the VCG mechanism ends up at (this is proved in the aforementioned paper [88]). However, QF is different from VCG in that it does not attempt to maximize $\sum_i c_i$ but instead aims to find the socially efficient allocation $\alpha$. Thus, QF will always have a deficit given by $D_QF = \alpha - \sum_i c_i = \left(\sum_i \sqrt{c_i}\right)^2 - \sum_i c_i$ which will always be positive as long as there is more than one contributor (Jensen’s inequality [89]) i.e., outside funding is always required for QF. The authors do not mind this as they assert that raising funds exclusively privately for public goods will always result in inadequate funding (due to the collective action problem).

QF also continues to be plagued by the other two problems of VCG. The first is that it fails if individuals can collude, and it is orders of magnitude easier to collude in QF than VCG. Simply consider the scenario where someone wants to contribute $c_i$ and they have friends who do not want to contribute anything; they can simply split their payments amongst their friends and have them contribute parts of their payment which will quadratically inflate their contribution (which could also then just be funneled back to themselves via dishonest projects “needing” to raise money). QF also always has a deficit that will be larger than VCG i.e., it will require a tax/grant to be completely funded.

In addition to these problems, QF also breaks down when contributions are in the form of single lump sums from foundations or philanthropists, specifically in the case where the value of the grant $G$ is larger than the deficit $D_{QF}$. The authors acknowledge this and propose an approximate solution called Capital-Constrained Quadratic Finance (CQF) where $\alpha = \beta \left(\sum_i \sqrt{c_i}\right)^2 + (1 - \beta) \sum_i c_i$ where $0 \leq \beta \leq 1$. When $\beta = 0$, CQF is the same as exclusively privately funding a good and when $\beta = 1$, CQF is the same as QF. The deficit of CQF is $\beta D_{QF}$ and the idea is that $\beta$ can be changed as needed depending upon how much of the funding can be externally sourced. However, it is important to note that CQF does not yield the efficient allocation of the good unless $\beta = 1$. 
All that being said, QF is extremely attractive in cases where the deficit is funded by a large pool of money provided by philanthropic donors who want a method to determine where to allocate their dollars. There are several issues with a simple voting system (or even more complicated ones like quadratic voting) where voters have nothing at stake, but QF addresses that by allowing voters to put as little as a dollar toward a project and have their donations matched (in a manner that scales quadratically with support for the project) from a larger pool of money. It has been implemented by FundOSS, Gitcoin, Downtown Stimulus, clr.fund, and various Ethereum projects. When implemented in practice, neither QF nor CQF are actually used but rather another variant of QF where \( \alpha = (\sum_i c_i^p) + G \left( \frac{\sum_i \sqrt{c_i^p}}{\sum_p \sum_i \sqrt{c_i^p}} \right)^2 \) is used. \( G \) is the size of the large pool of money/grant from philanthropic donors and \( p \) denotes one public good from a group of public goods, all of which want to raise money. This variant of QF while seemingly related to QF/CQF is actually quite different. It does not guarantee an efficient allocation. It also gives any project with at least one contributor some portion of \( G \) but this is not the case with QF, CQF, and even VCG. This means that the system is susceptible to attacks that fraudulently profit off \( G \). The other drawback is that choosing to support/donate to one project actually takes away from another, since all projects share a common pool of money \( G \). This approach to public goods funding is an interesting one but due to its inherent complexity, the need for external sources of funding, and its approach involving projects competing with each other for the same pool of money it does not fit the criteria for the system that will be proposed in this paper.

### 4.1.3 Empirical Results

For the sake of comprehensiveness, VCG, the classical mechanism for eliciting perfectly honest preferences and QF, a more recently proposed attempt to create a flexible framework for philanthropic or publicly funded seeding to allow (near) optimal provision of a decentralized, self-organizing ecosystem of public goods via endogenous community formation were discussed. The goal of QF is very much in line with the system that will be proposed in this paper.

Despite VCG and QF being “state of the art” mechanism designs that elicit truthful valuations, the “best” mechanism is the one that allows for the successful provision of the most allocation problems amongst a given set of allocation problems. This set must be presented to equivalent groups of individuals with a real vested interest in the allocation problems at hand and then it must be empirically determined as to which mechanism design (or variants that do not necessarily have the strongest theoretical guarantees but have practical ones) end up performing the “best” (where best is defined as above). The reason this has not been done is because conducting such an experiment at a scale that would make the findings valuable is extremely hard and impractical.

So, the work-around to this problem is laboratory testing, where cash payments are used to induce known preferences in real human subjects, who then participate in the mechanisms of interest. Here, the laboratory serves as a wind tunnel or “testbed” to examine mechanism properties in the face of real behavior with real incentives, but on a manageable scale with more control and observability. However, as previously discussed, such results are
not reliable because there is not really anything at stake for those involved, they are playing a game, one in which they can only act as they believe they would in real life. [90] They would likely hold themselves to the very highest moral standards without regard for the implications on themselves and their loved ones (because those implications are purely hypothetical when in a laboratory setting).

This reliance on laboratory testing often results in premature publication of theories that eventually fail to hold in reality. A survey of the many cases [91] in which this has happened are beyond the scope of this paper. However, there are two consistent lessons learned across all such studies. [92][93][94][95][96] The first conclusion is one previously discussed, which is that actual behavior converges to the Nash equilibrium in repeated games with a stable efficient equilibrium. Thus, any mechanism must have stability as part of the required desiderata if the mechanism is to be repeated successfully to reach efficient equilibrium. The second is that the truth-telling dominant strategy that mechanisms such as VCG guarantee is only a weak guarantee. Empirically it is found that certain individuals tend to be indifferent between telling the truth and lying i.e., they experience no change in utility despite the knowledge that lying would result in a worse outcome for them. If they are inclined to lie, then they will do so i.e., as previously mentioned, individuals are often not rational.

4.2 System Design

So far, a collection of observations has been made (Sections 1.5 and 2.5) and has been summarized as four properties to be satisfied in the system being designed. There are other ways to further enhance the system design to provide the greatest likelihood of success when attempting to solve collective action problems. These techniques build upon research results discussed in Section 1.2, and are discussed below.

4.2.1 Motivational Solutions

Social Value Orientations show that people have stable preferences for how much they value outcomes for themselves vs others. People generally fall in one of the following three categories: 1. Individualism (maximizing their own outcome regardless of others); 2. Competition (maximizing their own outcome relative to others); 3. Cooperation (maximizing joint outcomes). The first two are called proself orientations while the third is called a prosocial orientation. Prosocial individuals tend to be far more concerned about the moral and societal implications of their decisions than Proself individuals. [46]
Prosocial individuals are far more concerned about the moral and societal implications of their decisions and care more about those around them than proself individuals.

There are a variety of factors including but not limited to family history, age, culture, gender, level of education, even university majors etc. that affect whether someone is prosocial or proself but the underlying mechanisms for this behavior are not sufficiently well understood to intervene and influence them one way or the other, nor is it known if the behavior is binary/discrete or continuous and what its distribution is like in the population. However, something very interesting is that communication affects the tendency of individuals to exhibit prosocial or proself behavior. Cooperation and prosocial tendencies increase when people are given a chance to talk to each other and are in contact with others. It could be that communication reinforces a sense of group identity. [47] It is also possible that communication allows individuals to make promises and get explicit commitments in exchange for their behavior and compliance. However, it is not clear if people stick to their commitments; prior relationships between individuals as well as time spent in communication tends to affect this. [48] On the flipside, communication means an awareness of the willingness of others in the group to cooperate which could end up encouraging free riding.

Prosocial tendencies increase when individuals are allowed to communicate, establish connections, and build trust but too much communication encourages free riding.

4.2.2 Strategic Solutions

Generous tit-for-tat (GTFT) [49] is a strategy in which cooperation is reciprocated with cooperation and defection is reciprocated with defection; however, occasionally, at random, GTFT will forgive defection by the other player and cooperate in response. This strategy is often a Nash equilibrium and is evolutionarily stable in repeated interactions. The more beneficial cooperation is, the more forgiving GTFT can be while still resulting in positive outcomes. Tit-for-tat (TFT) is a stricter strategy but still works because it is “nice, but firm”, as a result of which, most real-world policies like marriage contracts, rental agreements, and international trade policies use TFT.

Some degree of forgiveness tends to work better in games with repeated interactions where actions of individuals are public or observed.

The important caveat here is that TFT or GTFT works well in repeated interactions as well as those in which choices of individuals are public and/or monitored by others. Cooperators tend to create better opportunities for themselves than non-cooperators and public acts of altruism and cooperation like charitable donations, philanthropy etc. are manifestations of reputation-based cooperation.
4.2.3 Structural Solutions

Solutions in this category aim to solve the collective action problem by modifying the rules of the game or getting rid of the problem altogether. Cooperation in collective action problems is based on a mix of the following factors: whether or not individuals have the ability to monitor the situation, to punish or "sanction" defectors, if they are legitimized by external political structures to cooperate and self-organize, can communicate with one another and share information, know one another, have effective arenas for conflict resolution, and are managing social and ecological systems that have well-defined boundaries or are easily monitorable. [50][51]

However, reward and punishment systems are hard to implement and can be problematic for several reasons.

First, there are costs associated with implementing and running such systems that can often be quite high for the marginal benefit they provide. The infrastructure needed to continuously monitor cooperators and non-cooperators and then fairly but selectively distribute rewards and penalties is not simple and is an expensive endeavor.

Second, such a system itself is a public good because individuals can enjoy the benefits of such a system without contributing to it just like how citizens who are not taxpayers enjoy the benefits of law enforcement. This raises the obvious question of how such a reputation system would be financed and if people would be willing to pay something equivalent to a tax to run such a system. Research suggests that particularly low trust individuals would be willing to finance such a system but also that a considerable number of people are willing to invest money into punishing non-cooperators even if they do not personally profit in any way and this form of altruistic punishment is believed to be an evolved mechanism for human cooperation. [52]
Third, these systems might undermine people’s willingness to voluntarily cooperate because some people simply feel good about doing so, and providing selective incentives for cooperation, as well as punishments for non-cooperation, might crowd out their voluntary willingness to cooperate and the mere existence of such systems could have the effect of decreasing the trust that people have in others. [53][54]

Reputation systems, especially when too strict and unforgiving, tend to have opposite the intended effect i.e., they decrease trust and willingness to cooperate.

Experiments show that there is a strong preference for democratically elected leaders with limited power (as defined by the group that elects the leader) to oversee the provisioning and upkeep of common resources and public goods. [55] An interesting observation is that the preference for a leader with limited power is primarily when ties within the group are strong and group sizes are small but in larger groups or those in which ties are weaker, there is a preference for a leader with stronger powers. Other research shows that an external authority is not what is needed (or wanted) but rather the ability to self-organize and collectively govern and care for the resource/good is and may even be pivotal to the sustainability of the provisioned good. [56]

Larger groups prefer democratically elected leaders with more power while smaller groups prefer democratically elected leaders with less power.

The ability to self-organize and collectively govern a provisioned public good is crucial to its long-term sustainability.

Yet another structural solution is to reduce group size (for obvious reasons). Cooperation declines with increasing group size due to the previously discussed phenomenon of individuals feeling like their contribution does not matter with increasing group sizes resulting in more free riding. The most common and effective solution to this problem is to manually reduce the scale of the problem by dividing it into smaller parts each of which is “assigned” to a subset of the original group. However, it is important to note that group size has a curvilinear effect. When group sizes are too small then they do not have sufficient bandwidth to research, manage, govern, and administer the issue at hand. [56]

Smaller groups encourage cooperation, but group size has a curvilinear effect, when groups are too small, they are unable to self-govern and organize around a problem.

Privatization is also a common means of restructuring collective action problems as it restructures incentives and essentially eliminates the temptation to place individual needs above that of the group however it also might end
up eroding people’s intrinsic voluntary cooperative tendencies by externalizing the locus of control and thus it is important to strike the right balance between privatization and local self-governance.

**While privatization is necessary to provision economically efficient levels of public goods, too much privatization takes the locus of control away from communities.**

Nearly all environmental problems can be classified as collective action problems. [44] This is simply because these problems are a result of humans around the world behaving a certain way and so they require a collective behavior change in order to be addressed. However, without government regulation and the threat of disobedience being punished by law enforcement, individuals are unlikely to contribute to the collective behavior change as they are naturally incentivized to choose the cheaper/easier alternative which almost always differs from what the collective needs. Thus, it makes perfect sense that over half of all Americans believe that government regulation does more harm than good but at the same time say that it is necessary when it affects their quality of life (food and water quality, FDA regulation etc.) [45]

**People are far more likely to cooperate, and maybe even support government intervention, when they believe it will positively affect their lives and that of those around them.**

There are several other structural solutions that can be used to modify collective action problems and many of these solutions can be used in conjunction with each other, but at the end of the day, the likelihood of successfully co-managing a shared resource, successfully organizing to self-govern, or successfully cooperating to solve a collective action problem depends on many things, from the nature of the resource/good, to the nature of the social system the actors are a part of, to the political position of external authorities, to the ability to communicate effectively, to the rules-in-place regarding the management of the resource/good and the infrastructure that exists to support these activities. [57] However, sub-optimal or failed results in a collective action problem tend to occur "when resource users do not know who all is involved, do not have a foundation of trust and reciprocity, cannot communicate, have no established rules, and lack effective monitoring and sanctioning mechanisms." [58]

**The primary reasons communities cannot provision public goods are: when they do not know who is involved, do not have a foundation of trust, cannot communicate, have no established rules, and lack effective monitoring and sanctioning mechanisms.**

### 4.2.5 Field Studies and Empirical Evidence

#### 4.2.5.1 Public Goods Games

A public goods game is essentially an n-person Prisoner’s Dilemma (like the unscrupulous diner’s dilemma previously discussed). The specifics of the game can take any form and there are several variants used in various
studies (the specifics seem to not have an impact on the outcomes of these studies, the themes observed are consistent). Surprisingly, the Nash-equilibrium strategy is actually only played by some players, some tend to be far more optimistic and altruistic, while others play to “win”. Some interesting observations include the fact that in one-shot trials and initial stages of repeated games, subjects generally provide contributions halfway between the Pareto efficient level and free riding level, contributions actually decline with repetition (as players learn to game the system) and communication improves the rate of contribution. [73] This shows that people free ride less often than advocates of strong government intervention claim, but provisioning of public goods would still be at suboptimal levels if exclusively privately provisioned (this can also be shown mathematically). Jon Elster explains that there are six types of cooperative behavior and associated personalities and for the greatest chances of cooperative success, one should attempt to create the optimal mix of these six personalities. [74] However, this is largely irrelevant for the system being proposed in this paper since it would be impossible to determine which “type” an individual is, at scale, and then to optimize for that. Contributions increase with increasing thresholds at which the public good is provided while the probability that the threshold will be reached goes down. [75] Interestingly, contrary to economists’ expectations, group size can have a positive effect on contributions despite it diluting the effect of marginal returns. [76]

Contributions increase with increasing thresholds at which the public good is provided while the probability that the threshold will be reached goes down.

Group size has a positive effect on contributions but naturally this ends up also diluting marginal returns.

4.2.5.2 Substitutes or Complements?

All field studies confirm that individuals make some voluntary contributions to public good funding but what is more interesting are the factors that affect the sizes of the contributions and the mechanisms used to encourage people to contribute. Specifically, a number of field studies have been performed to determine the relationship between an individual’s contribution to a public good and others’ contributions. There are two schools of thought. The first is that if an individual is primarily motivated by altruism, they will care about the consumption of others and therefore contribute less when others or the government already contributes enough. In this case, the two kinds of contributions would be substitutes. The second is when an individual is primarily motivated by social norms such as fairness or reciprocity, then they will contribute when others do their bit as well. In this case, the two kinds of contributions are complements.

Field experiments show support for the complementary hypothesis. [77] Specifically, they found that if social information was provided about a high contribution of one donor, then pledges increased and the likelihood that each individual donor contributed again also went up. The same was found in another study on voluntary contributions to an information good. [78] At the same time, observational studies have found that the notion
that the two kinds of contributions are perfect substitutes i.e., that government or other private contributions completely crowd out individual contributions has no empirical backing. Most studies found either no crowding out or very minimal crowding out. [79][80]

**Most people see others’ funding a good as complementary to their own funding of the good, not as substitutes; and seeing greater contributions also elicits contributions.**

### 4.4 Governance Design

Due to the emphasis on governance [10] from all the “observations” that have been made so far, further investigation is warranted.

#### 4.4.1 Rational Ignorance

The Paradox of Voting says that when a rational self-interested individual casts a vote, the probability of their individual vote influencing the outcome of an election is so low that the costs almost certainly outweigh the benefits. Some of the costs associated with casting a vote include the opportunity cost of doing something else at that time, the literal cost of making the effort to cast a vote, the cost of being an informed voter etc. This leads to the theory of Rational Ignorance which is that the low probability of influencing the outcome of an election leaves little to no incentive to learn about any of the issues at hand, especially when the payoffs from solving those issues for each individual tend to be quite low. In fact, empirical studies show that in the political realm, insufficient knowledge is the norm rather than the exception, even amongst active participants. [81]

The costs outweigh the benefits if someone is the only person making the effort to vote and/or if their candidate loses. However, if everyone adopted this view, or even a vast majority did, then democracies would not be able to function and once they hit a certain critical mass would break down as a result of this realization. This is a collective action problem since individuals are incentivized to stay at home since that avoids the “cost” of voting and they know that if they do, it does not matter, since their vote would not sway the outcome of the election anyway; but, what if everyone made the effort and their candidate won, then the benefits may outweigh the costs as they do have the power to sway an election and to bring into effect changes they would like to see.

So, it seems like, to solve the collective action problem associated with the provisioning of public goods, a system that involves voting into effect solutions to those collective action problems has instead been proposed. In essence, this posits that despite the inability to coordinate to provision public goods, the ability to coordinate to vote into effect proposals that will provision public goods is possible. If this were true, why not just coordinate to provision public goods in the first place? It seems like one collective action problem has been replaced with another.

Interestingly enough, despite high levels of political apathy in the United States and the fact that there is only a one in ten million chance that one vote could sway the outcome of a United States Presidential election even in a
battleground state [42], this does not turn out to be a problem simply because most Americans believe the power of their vote to be much higher than it actually is. [43] It appears that the mere belief that collective action will lead to individual benefits is sufficient to overcome the collective action problem (as discussed earlier). This holds true for nearly all democracies around the world and is one of the primary reasons this form of governance works as well as it does.

The mere belief that collective action will lead to individual benefits is sufficient to overcome the collective action problem.

4.4.2 Independence of Irrelevant Alternatives

Independence of Irrelevant Alternatives (IIA) states that if candidate A beats candidate B in a head-to-head matchup, introducing candidate C should not make candidate B a winner. If each voter can only vote for one candidate and the winner is decided by plurality, then IIA is violated. This happened during the 2000 U.S. Presidential Election in which Bush ended up winning by a small margin over Al Gore due to votes that would have otherwise been cast for Al Gore being cast for Ralph Nader (the “spoiler” candidate i.e., a loser whose presence in the race changed who the winner is) instead. [62]

In a head-to-head matchup, Al Gore would have beat Bush. However, the introduction of Nader (far left candidate) meant that some voters who planned to vote for Gore now voted for Nader instead, thereby letting Bush win. [63] Nader received 22,000 votes in New Hampshire, where Bush beat Gore by 7,000 votes and he also received 97,000 votes in Florida, where Bush beat Gore by just 537 votes. If even a small subset of Nader’s voters had voted for Gore in either of those states, Gore would have won. This whole fiasco and the controversy surrounding it could have been avoided if a ranked choice voting system was used.

<table>
<thead>
<tr>
<th>Presidential candidate</th>
<th>Party</th>
<th>Home state</th>
<th>Popular vote</th>
<th>Electoral vote</th>
<th>Running mate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Count</td>
<td>Percentage</td>
<td>Vice-presidential candidate</td>
</tr>
<tr>
<td>George Walker Bush</td>
<td>Republican</td>
<td>Texas</td>
<td>50,456,002</td>
<td>47.86%</td>
<td>Richard Bruce Cheney</td>
</tr>
<tr>
<td>Albert Arnold Gore, Jr.</td>
<td>Democratic</td>
<td>Tennessee</td>
<td>50,999,897</td>
<td>48.38%</td>
<td>Joseph Isadore Lieberman</td>
</tr>
<tr>
<td>Ralph Nader</td>
<td>Green</td>
<td>Connecticut</td>
<td>2,882,955</td>
<td>2.74%</td>
<td>Winona LaDuke</td>
</tr>
<tr>
<td>Pat Buchanan</td>
<td>Reform</td>
<td>Virginia</td>
<td>448,895</td>
<td>0.43%</td>
<td>Ezola B. Foster</td>
</tr>
<tr>
<td>Harry Browne</td>
<td>Libertarian</td>
<td>Tennessee</td>
<td>384,431</td>
<td>0.36%</td>
<td>Art Olivier</td>
</tr>
<tr>
<td>Howard Phillips</td>
<td>Constitution</td>
<td>Virginia</td>
<td>98,020</td>
<td>0.09%</td>
<td>Curtis Frazier</td>
</tr>
<tr>
<td>John Hagelin</td>
<td>Natural Law</td>
<td>Iowa</td>
<td>83,714</td>
<td>0.08%</td>
<td>Nat Goldhaber</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td>51,186</td>
<td>0.05%</td>
<td>Other</td>
</tr>
<tr>
<td>(abstention)¹³</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>(abstention)¹³</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td>105,421,423</td>
<td>100%</td>
<td>538</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Needed to win</th>
<th>270</th>
</tr>
</thead>
</table>
Instant Runoff Voting (IRV) is a type of ranked choice voting where multiple rounds are held until a candidate has a majority. After each round where no candidate has a majority, the candidate with the least votes gets eliminated and their votes are redistributed to the next ranked candidate for each voter. If IRV were used in the 2000 elections, then after all candidates with votes fewer than Nader were eliminated, Nader himself would have been eliminated as well. Upon redistribution of Nader’s votes to Gore and Bush, Gore would have won since far-left voters would have definitely preferred a candidate on the left than on the right. Even though IRV does not violate IIA in this scenario, it does in others. This happened in the 2009 Burlington Mayoral race.

<table>
<thead>
<tr>
<th>Candidates</th>
<th>Party</th>
<th>Votes</th>
<th>%</th>
<th>% Active</th>
<th>±</th>
<th>Votes</th>
<th>%</th>
<th>% Active</th>
<th>±</th>
<th>Votes</th>
<th>%</th>
<th>% Active</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bob Kiss</td>
<td>Progressive</td>
<td>2,585</td>
<td>28.8%</td>
<td>28.8%</td>
<td>+396</td>
<td>2,981</td>
<td>33.2%</td>
<td>33.8%</td>
<td>+1332</td>
<td>4,313</td>
<td>48.0%</td>
<td>51.5%</td>
</tr>
<tr>
<td>Kurt Wright</td>
<td>Republican</td>
<td>2,951</td>
<td>32.9%</td>
<td>32.9%</td>
<td>+343</td>
<td>3,294</td>
<td>36.7%</td>
<td>37.3%</td>
<td>+767</td>
<td>4,061</td>
<td>45.2%</td>
<td>48.5%</td>
</tr>
<tr>
<td>Andy Montroll</td>
<td>Democrat</td>
<td>2,063</td>
<td>23.0%</td>
<td>23.0%</td>
<td>+491</td>
<td>2,554</td>
<td>28.4%</td>
<td>28.9%</td>
<td>-2,554</td>
<td>0</td>
<td>0.0%</td>
<td></td>
</tr>
<tr>
<td>Dan Smith</td>
<td>Independent</td>
<td>1,306</td>
<td>14.5%</td>
<td>14.5%</td>
<td>-1,306</td>
<td>0</td>
<td>0.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>James Simpson</td>
<td>Green</td>
<td>35</td>
<td>0.4%</td>
<td>0.4%</td>
<td>-35</td>
<td>0</td>
<td>0.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Write-In</td>
<td></td>
<td>36</td>
<td>0.4%</td>
<td>0.4%</td>
<td>-36</td>
<td>0</td>
<td>0.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXHAUSTED PILE</td>
<td></td>
<td>4</td>
<td>0.0%</td>
<td>0.0%</td>
<td>+147</td>
<td>151</td>
<td>1.7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTALS</td>
<td></td>
<td>8980</td>
<td>100.0%</td>
<td>100.0%</td>
<td>8980</td>
<td>100.0%</td>
<td>8980</td>
<td>100.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Bob Kiss (the incumbent) ended up winning the IRV election but was neither the plurality winner (which was Kurt Right) [65] nor the Condorcet winner (which was Andy Montroll) [64]. The Condorcet winner is the candidate that beats any of their opponents in a head-to-head matchup.

Andy Montroll defeated both Bob Kiss and Kurt Wright in separate pairwise contests and was eliminated in the semifinal round of IRV due to vote-splitting with both candidates. Thus, IRV violated IIA. Kurt Wright became the “spoiler” candidate, splitting the vote against Bob Kiss; Wright received more first-choice votes (including promoted votes to first choice) than Montroll due to Kiss splitting the vote against Wright. [66]

At this point, it should seem like finding a good voting system is hard. Arrow’s impossibility theorem proves exactly that.

### 4.4.3 Arrow’s Impossibility Theorem

Arrow’s impossibility theorem states that when voters have three or more distinct choices, no ranked choice voting system can convert the ranked preferences of individuals into a community-wide (complete and transitive) ranking while also meeting the following set of criteria: *unrestricted domain/universality, non-dictatorship, unanimity/Pareto efficiency, and independence of irrelevant alternatives.*
Unrestricted domain/universality essentially means that for any set of voter preferences, the voting system must present a unique ranking of choices such that it is a complete ranking of preferences for society, and it must deterministically provide the same ranking every time voters’ preferences are presented the same way.

Non-dictatorship refers to the fact that the voting system must account for the wishes of multiple voters i.e., no single voter must have the power to determine the outcome of an election by themselves.

Unanimity/Pareto efficiency refers to the fact that if every voter prefers candidate A to candidate B, so should the resulting societal preference order.

These requirements seem sufficiently basic that given the faith society places in democratic systems, so there must exist a voting system that can satisfy this minimal set of criteria. Arrow’s impossibility theorem proves that there is not. The details of the proof are beyond the scope of this paper but given how striking this result is, there are myriad presentations of this proof that appeal to people with varying degrees of mathematical inclination that the reader can choose to refer to should they be interested in more details.

This shows that the results of any voting system are always flawed, and the “people’s will” really depends on whether they prioritize IIA or Unanimity – a given system can only have one, not both. When it comes to choosing between the two, given the nature of the problem being solved in this paper, IIA is preferred over Unanimity. The reason for this is because if a voting system allows IIA to be violated then it is not strategy proof. A voting system is strategy proof if no voter can get a better outcome in the election by lying about their preferences. There should be no incentive for anyone to lie about their preferences, if they do, the chances of building a competitive free market that works is slim, and so, IIA must be upheld by the governance mechanism.

A strategy-proof voting system must uphold Independence of Irrelevant Alternatives.

4.5 Summary of Observations

There is no direct incentive-compatible efficient (i.e., first-best) mechanism to raise funds for public goods that never runs a deficit.

The only second-best mechanism to raise funds for public goods is a fixed-split mechanism (i.e., taxation).

Behavior of individuals converges to the Nash equilibrium in repeated games with stable efficient equilibrium conditions.
Individuals often have no change in utility due to lying despite active knowledge that telling the truth would be unequivocally better for them.

Prosocial individuals are far more concerned about the moral and societal implications of their decisions and care more about those around them than proself individuals.

Prosocial tendencies increase when individuals are allowed to communicate, establish connections, and build trust but too much communication encourages free riding.

Some degree of forgiveness tends to work better in games with repeated interactions where actions of individuals are public or observed.

Public acts of altruism and cooperation like charitable donations, philanthropy etc. are manifestations of reputation-based cooperation.

The ability to monitor the situation (i.e., transparency), to punish or sanction defectors, legitimation from the government, the ability to communicate, and effective means of conflict resolution impact likelihood of cooperation.

Reputation systems themselves are public goods that have running costs and need to be funded by some mechanism or be made exclusionary.

The majority of people believe in altruistic punishment i.e., they are willing to dedicate significant (monetary) resources to punish defectors.

Reputation systems, especially when too strict and unforgiving, tend to have opposite the intended effect i.e., they decrease trust and willingness to cooperate.

Larger groups prefer democratically elected leaders with more power while smaller groups prefer democratically elected leaders with less power.
The ability to self-organize and collectively govern a provisioned public good is crucial to its long-term sustainability.

Smaller groups encourage cooperation, but group size has a curvilinear effect, when groups are too small, they are unable to self-govern and organize around a problem.

While privatization is necessary to provision economically efficient levels of public goods, too much privatization takes the locus of control away from communities.

People are far more likely to cooperate, and maybe even support government intervention, when they believe it will positively affect their lives and that of those around them.

The primary reasons communities cannot provision public goods are: when they do not know who is involved, do not have a foundation of trust, cannot communicate, have no established rules, and lack effective monitoring and sanctioning mechanisms.

Contributions increase with increasing thresholds at which the public good is provided while the probability that the threshold will be reached goes down.

Group size has a positive effect on contributions but naturally this ends up also diluting marginal returns.

Most people see others’ funding a good as complementary to their own funding of the good, not as substitutes; and seeing greater contributions also elicits contributions.

The mere belief that collective action will lead to individual benefits is sufficient to overcome the collective action problem.

A strategy-proof voting system must uphold Independence of Irrelevant Alternatives.

In these observations, there are six common properties, adding onto the four properties previously observed:
Property 5: There is no mathematically provable mechanism that is capable of raising adequate funds for public goods in an efficient, direct, and incentive-compatible manner. However, these results do not hold in practice, as individuals are not rational (bounded rationality), are often indifferent about lying, and have malleable preferences that change in repeated interactions with stable equilibrium.

Property 6: Prosocial individuals care more about those around them than proself individuals and thus tend to be far more altruistic and willing to cooperate; but these orientations are not fixed. Prosocial tendencies increase when individuals are allowed to communicate, have a means to build and establish trust, and have universally agreed upon conflict resolution schemes.

Property 7: Reputation systems encourage cooperation and altruism. These systems must allow for the punishment of defectors; but must strike a balance between doing so in a manner sufficiently strict that it deters defectors yet has some room for forgiveness. These systems cost money to run but people are willing to pay for them (altruistic punishment).

Property 8: The ability to “see everything” and to monitor situations with complete transparency is vital to the success of an effort to solve a collective action problem as it is pivotal to building and establishing trust amongst pseudonymous identities who previously had none. When privatization is involved, transparency is key to making sure that communities do not feel like the locus of control is being taken away from them.

Property 9: The ability to collectively govern a public good once provisioned is pivotal to its long-term sustainability. Larger groups prefer elected leaders with more power than smaller groups do. Any voting system being used to elect leaders or make decisions must be strategy-proof. The governance system must instill confidence and the feeling that everyone’s vote matters (the mere belief that it does is sufficient).

Property 10: Larger group sizes have a positive effect on contributions (frequency and magnitude), and seeing greater contributions only elicits more, but they have a negative effect on the tendency to not free ride and cooperate. At the same time, groups that are too small lack the bandwidth to effectively self-organize and govern.

5 The 10 Desired Properties

Property 1: The tendency to cooperate and not free ride is much greater in smaller groups than in larger groups; and in smaller groups this goes up when those in the group have close-knit interpersonal relationships and/or care about each other. This can be fostered by encouraging communication but also by establishing means to record past tendency to cooperate i.e., institutions and infrastructure can also enable trust.

Property 2: Reputation mechanisms that have deterrents (i.e., some punishment mechanism) with simple heuristics that easily indicate to others the degree of trustworthiness/cooperativeness of those they are working
with helps increase the likelihood of cooperation in larger groups. This is especially true when individuals are pseudonymous and may not have previously interacted with each other.

Property 3: The responsibility to provision public goods often falls on communities who lack the infrastructure and funding to provision public goods themselves. Without infrastructure that helps them collectively organize around issues they care about, the transaction costs involved with provisioning public goods to address those issues is too high and cannot be met without special interest support.

Property 4: People will underestimate or overestimate what a public good is worth to them when asked to put a dollar value on it, and may do so intentionally, but even the most honest people will end up doing so unintentionally. Most people may be rational, but their rationality is bounded. They need to “put their money where their mouth is”, and there is simply no substitute for doing so.

Property 5: There is no mathematically provable mechanism that is capable of raising adequate funds for public goods in an efficient, direct, and incentive-compatible manner. However, these results do not hold in practice, as individuals are not rational (bounded rationality), are often indifferent about lying, and have malleable preferences that change in repeated interactions with stable equilibrium.

Property 6: Prosocial individuals care more about those around them than prosel individuals and thus tend to be far more altruistic and willing to cooperate; but these orientations are not fixed. Prosocial tendencies increase when individuals are allowed to communicate, have a means to build and establish trust, and have universally agreed upon conflict resolution schemes.

Property 7: Reputation systems encourage cooperation and altruism. These systems must allow for the punishment of defectors; but must strike a balance between doing so in a manner sufficiently strict that it deters defectors yet has some room for forgiveness. These systems cost money to run but people are willing to pay for them (altruistic punishment).

Property 8: The ability to “see everything” and to monitor situations with complete transparency is vital to the success of an effort to solve a collective action problem as it is pivotal to building and establishing trust amongst pseudonymous identities who previously had none. When privatization is involved, transparency is key to making sure that communities do not feel like the locus of control is being taken away from them.

Property 9: The ability to collectively govern a public good once provisioned is pivotal to its long-term sustainability. Larger groups prefer elected leaders with more power than smaller groups do. Any voting system being used to elect leaders or make decisions must be strategy-proof. The governance system must instill confidence and the feeling that everyone’s vote matters (the mere belief that it does is sufficient).
Property 10: Larger group sizes have a positive effect on contributions (frequency and magnitude), and seeing greater contributions only elicits more, but they have a negative effect on the tendency to not free ride and cooperate. At the same time, groups that are too small lack the bandwidth to effectively self-organize and govern.

These 10 properties are not a comprehensive summary of everything touched on in prior sections but cover the most important results observed from existing economic, biological, psychological, and sociological literature on collective human behavior. A lot of these properties are overlapping to certain degrees and emphasize common themes that can be encapsulated in the following set of features the system being designed in this paper must address. Group Size, Governance, Reputation, Transparency, Pseudonymity, Communication, Valuation Mechanisms, and Conflict Resolution Schemes.

6 Blockchains

6.1 User Interface

The precise “look and feel” of the proposed system (henceforth referred to as “the system”) requires a lot of thought, effort, and research, and may even affect the formal rules and specifications of the system in some way but is beyond the scope of this paper. At times, it will be hard to describe rules and specifications for satisfying the 10 desired properties (Section 5) without bleeding into implementation details. Thus, while a best attempt is made at keeping details as abstractable as possible, there will be times where ideas will be presented via implementation (as in Section 6.10) from which abstracting out formal specifications is too verbose; but at the same time, it is quite easy to get a sense of how those parts of the system would work in practice from the sample implementations discussed.

However, these sample implementations are fairly arbitrary in their choice of details, UI elements etc. i.e., it is extremely unlikely that the sample implementations of various components discussed can work together cohesively as a single system without further concerted research. Any discussion of implementation details must merely be taken as a suggestion and nothing more. This will also be clarified every time such details are presented but should there be a spot that was missed, this general disclaimer stands. For the same reason, should anything be unclear/ambiguous or seem unattainable given other system design choices that will vary based on implementation details, defaulting to the 10 desired properties (Section 5) and making sure those are satisfied (to the greatest extent possible) will result in the design of the most optimal system.
6.2 The Two Arms

6.2.1 Roles

The system will formally have two “arms”. The first will be a “foundation” (henceforth referred to as “the foundation”) which is a non-profit organization that will build the protocols for which the specifications are described here, get the system up and running by assuming responsibility for various administrative tasks, and shepherd the system toward its goals in the initial days. The specifics of the functions of the foundation will be described in future sections. The second will be a DAO (henceforth referred to as “the DAO”) that will function like any other DAO i.e., a group of individuals that form a governing body and make decisions in interest of the system through a bottom-up approach involving voting on proposals and allocating funding. [110] DAOs are autopoietic, alegal, hyper scalable, executable, permissionless, aligned, co-owned, and mnemonic (Section 6.8.4.1.2). [155] The specifics of the functions of the DAO will also be described in future sections.

The foundation will establish rules on how and when all of their operational and administrative duties will be transferred over to the DAO. There are several blockchain projects such as Uniswap, Aave, The Arbitrum Foundation/Arbitrum DAO, The Dai Foundation/MakerDAO etc. that have successfully followed a two-armed governance structure involving transition of all powers to a DAO. [113][114][115][116] This structure is especially relevant here because the nature of the system being proposed is one that requires manual human intervention from time to time, which at the start, will be provided by the foundation. In addition to this, when the system is first launched, there are several functions that simply need to be performed by some kind of central entity. This entity must be able to act unanimously in the interest of the system, primarily to address security concerns (Section 6.11). Once again, the foundation shall fill this role. Once the system reaches a certain critical mass of users, most security concerns will be alleviated. At that point, powers can transition from the foundation to the DAO per some predefined protocol.

The specifics on the roles of the foundation and the DAO, as well as when and how powers transition from one to the other are being kept intentionally vague at this point. The former, because it will become clear as the system is specified in greater detail in forthcoming sections, and the latter, because it will nearly entirely be dictated by the implementation specifics of the system but also by the regions in which it is operating, the frequency with which there are security breaches, the degree to which the DAO is capable of assuming all necessary responsibilities etc. There are simply too many variables to even define a broad framework that would be realistically applicable for how and when this transition of power can take place. Thus, it is entirely left up to implementation. The aforementioned successful transitions of power from foundations to DAOs may be used to guide this effort. There are many valuable lessons in those prior attempts, however, exploring those attempts in detail is beyond the scope of this paper but it is worth investigating prior to attempting to set up this system. There will also be many other successful (and unsuccessful) attempts at blockchain governance and administration over time, all of which will also be valuable sources of learning and information.
6.2.2 Funding

It will become clear that the foundation, and eventually the DAO, will have a large role to play in administering, governing, and most importantly, moderating the system. At least initially, the foundation will need to be funded by some means to perform all these tasks. For now, one can consider the most naive mechanism to do so, which is that a certain percentage of the funds raised for every proposal (Section 6.5.2) will be used to fund the foundation. This percentage buffer will actively be reduced until it only accounts for the most essential tasks that need to be explicitly funded i.e., tasks that are likely to be outsourced (such as conversions from cryptocurrencies to fiat currencies etc.). This initial period in which the buffer pays for overhead fees is contrary to the core mission of this system which is to achieve economic efficiency in a market for public goods; it is a temporary but necessary evil at this stage of the system. At the same time, it is important to remember that charities, nonprofits, and NGOs have very significant overhead costs (Section 6.7.1), so while this initial period with overhead costs is not optimal, these overhead costs will likely still be far lower than that for charities, nonprofits, and NGOs; and at the very least, the way in which overhead expenses are spent will be transparent.

However, once the DAO takes over from the foundation, it is unclear if explicit funding will be needed. This is for two reasons. The first is that, traditionally, DAOs have been run by volunteers with no real direct financial incentive to do so. [201] However, most DAOs also eventually do have some financial incentive, generally in the form of airdropped tokens that reward volunteers. [202] The possibility of something similar is not entirely eliminated but is also not discussed here since it clouds the mission of the system to an uncomfortable degree and creates certain perverse incentives that can impede the system from succeeding at its core mission.

The second is that, as previously discussed (Section 4.2.3), reward and punishment mechanisms i.e., reputation mechanisms (Section 6.10) are in and of themselves public goods that traditionally would need to somehow be financed, and generally be made exclusionary (however, this system is not exclusionary). At the same time, research (Section 4.2.3) also shows that people are willing to pay (monetarily or otherwise) to run altruistic reputation mechanisms that punish defectors. If the latter were to hold true in practice, then the DAO (by way of the reputation mechanism) could even be explicitly funded with some kind of subscription.

There are myriad ways to fund overhead costs like these without directly eating into funding raised for proposals (Section 6.5.2), because doing so can be a slippery slope toward very high overhead margins and so should only be used as a last resort. In general, as long as the source of funding cannot directly influence the governance of the system (Section 6.8) in any way, then it may be considered a viable means of funding. The details on funding sources are left to implementation but is definitely an important area for consideration since no one knows how.
this will play out and what the true costs incurred would be. For instance, there will likely be various housekeeping smart contract calls that are executed periodically, and the gas fees for those contracts will need to be paid for by someone.

6.3 Decentralization

6.3.1 “For”

Before proceeding further, it is imperative to establish why blockchains are needed for this system. Specifically, at this stage, to clarify why governance and administration are being performed through this two-armed approach that necessitates the use of a blockchain. The reason is quite straightforward, it is simply because decentralization is important. Though there are several other reasons that also necessitate the use of a blockchain for this system (that will become clear in forthcoming sections), this is one of the most important ones. Some of the most common arguments for decentralization often include resisting government censorship, libertarian political views etc. While those may be valid to varying degrees depending on political views, where one lives, their culture etc. there is a more concrete reason why decentralization matters; and it has to do with centralized platforms.

All centralized platforms have the same playbook. At the start, often fueled by vast amounts of venture capital, they do everything they can to make their platform attractive - one where their users and/or their partners (developers, businesses, creators etc.) have the better side of an asymmetric deal - something that cannot be passed up. They do this to make their services more valuable as most services are valued (at least partially) as functions of their users and/or partners but also because most businesses are systems with multi-sided network effects. [111] As these centralized entities move up the adoption S-curve, what was once an asymmetric deal where they did not have the better side, quickly flips.

When they hit the top of that S-curve their relationship with network participants goes from positive-sum to zero-sum. The only way to continue growing (and make more money) is to extract data from users and to compete with anything that may be construed as a competitor until nothing remains (a certain company named after a famous tropical river is widely known for this strategy) i.e., to become monopolies. Some historical examples of
This includes Microsoft vs. Netscape, Google (Reviews) vs. Yelp, Facebook (Games) vs. Zynga, Twitter vs. 3rd party clients (like Tweetbot), etc. Operating systems like iOS/Android have had a slightly better track record though they still charge a 30%+ commission on all sales of and within apps [117], delist/reject apps for seemingly arbitrary reasons [119][120], and subsume functionality of 3rd party apps at will [121]. YouTube is another example where it has gradually, over the years, increased its own cut of creator ad revenue [118] in addition to adding more ads on every video, at increasing intervals [122], only to launch YouTube Premium as a solution.

This transition from cooperation to competition happens fairly gradually, yet it feels like a bait-and-switch for users and partners alike; but by that point, their dependency on those platforms is so great that neither group can protest in a meaningful way. Over time, the best entrepreneurs, developers, investors, and creators have become wary of depending on centralized platforms in a manner crucial to their interests because of decades of evidence that doing so only results in disappointment. In addition to which, users give up privacy, control of their data, and are vulnerable to the downstream impacts from potential data breaches when their data is stored in unknown ways by these centralized entities. These problems with centralized platforms are real and are being regulated heavily in the EU [123] but even that is not enough, especially when there is virtually no such regulation in the U.S. [124][125], which is where these companies are largely based, managed, and make a lot of their money.

Blockchains are a solution to this. They use decentralized consensus mechanisms to maintain and update state (instead of centralized servers); and associated cryptocurrencies incentivize consensus participants, users, and other network participants to run and maintain the system (instead of a centralized entity doing this). They combine the best of open-source technical specifications and protocols (www, VoIP, email etc.) from the early era of the internet with that of economic incentives to participate and build such networks offered by more recent centralized entities (Google, YouTube, Facebook etc.). They are able to balance this without too many of the negative side effects of either approach due to their open-source nature, transparency, and mechanisms of checks and balances via community governance (both on-chain and off-chain).

Ethereum is a general-purpose programming platform run on a blockchain with all of the properties that were just described, and for that reason, in this paper, it is the suggested network/platform to implement the proposed system. Whenever implementation specifics are described, it will involve references to specifics in Ethereum. However, this is only a suggestion. There are several blockchains that have similar properties, there may even be ones particularly well suited to this application, and one could well make an argument for building their own blockchain for this system, however, once again, these details are left to implementation.

Property 8: The ability to “see everything” and to monitor situations with complete transparency is vital to the success of an effort to solve a collective action problem as it is pivotal to building and establishing trust amongst pseudonymous identities who previously had none. When privatization is involved, transparency is key to making sure that communities do not feel like the locus of control is being taken away from them.
In addition to the decentralization that blockchains provide, that in and of itself makes the system being built more trustworthy than a centralized one, the completely public and transparent nature of all activities on the blockchain further enhances trust in a way that no other technology primitive can. In fact, every aspect described henceforth (be it governance mechanisms, reputation scores etc.) are enhanced by the inherent trust as a result of them being deployed on/via a blockchain. The literature discussed prior to this harped on the importance of building trust and showed that it is, by far, one of the most important factors that determine the success of efforts to solve collective action problems, and this happens to be specifically true when it comes to community organized efforts to provision public goods. Transparency, above all, enhances trust the most.

6.3.1 “Against”

It is also important to acknowledge that there is debate as to whether blockchains are the best tool to align network participants to work toward a common goal given its many limitations (scale and security), its inefficiency (performance), energy usage etc. but it is also the only one that works (for the most part), and so for the purposes of this paper, it will be used as the underlying technology primitive that makes the system being proposed a reality. The advantage to laying out the paper in this manner is that even if there were to eventually be a technology primitive derived from, or even completely unassociated to, blockchains that were a better fit for the system being proposed then one could still use everything learned up to Section 5 alongside the general approach taken in Section 6 to design specifications for a system that achieves the same goals using a different technology primitive.

Another question of particular relevance is the degree of decentralization required, the aspects of the system for which it is required, and whether the downsides to it are worth it. There has always been a lot of pessimism and critiques of highly decentralized systems [180][181], that DAOs do not really work in practice [182][183], and that DAO governance should just start to closely resemble that of traditional corporations in order to remain competitive. [184][185][186] The arguments for this are always fairly consistent and highlight a few themes: decentralization is inefficient, traditional corporate governance structures with boards that have indirect oversight over a CEO with domain expertise have evolved over hundreds of years of capitalism for a reason, the egalitarian ideals of decentralization are unrealistic and it is naive to believe that DAOs can do what corporations have tried to do but failed at achieving numerous times.

Refuting those lines of argument are both beyond the scope of this paper but also, to a certain extent, futile, because all those arguments have varying degrees of truth to them. In fact, some of those lessons will be incorporated in the systems defined in the paper. But even assuming the drawbacks of decentralized systems outweigh those of centralized systems highlighted in the prior section, there are still a few scenarios where decentralization is simply better at producing desirable outcomes; and all of these scenarios describe the very system being built in this paper. In short, decentralization is better for making decisions in concave environments, for censorship resistance, and as a means of credible fairness.
Decentralization is better for making decisions in concave environments. What exactly does that mean? Consider binary decisions. Before considering a choice between A and B, a higher-order question is worth considering, and that is whether a compromise between A and B or a coin flip between A and B is a better choice? Concave decisions are those where a compromise is better while convex decisions are those where a coin flip is better.

Some examples of concave decisions include judicial decisions, where the average between two independently chosen judgements is likely fairer than that from a single judgment (or a random choice between the two). Tax rates are another great example where due to deadweight loss (that scales quadratically with tax rates, as discussed in Section 2.2.1), a tax rate of $x$ is only about 25% as harmful as a tax rate of $2x$, and it is better to settle on a tax rate somewhere between the two rather than to just flip a coin and pick one randomly. These decisions are those where a compromise is better than an absolute choice and so they are better suited for decentralized decision-making processes where the diverse inputs of hundreds if not thousands of individuals go into the decision-making process.

On the other hand, some examples of convex decisions include the pandemic response, where a 100% ban on travel, eating out etc. would have stopped the spread of the virus while a 0% ban would not have, but at least would not have inconvenienced people; however, a 50% ban is the worst of both worlds, and did not stop the virus but also inconvenienced people. Convex decisions are those where a single answer is required, likely from a domain expert and decentralizing such decisions can often lead to poor outcomes. Military decisions are yet another example where it is worth considering whether attacking from the east or the west is the best strategy but splitting one’s forces across the two fronts equally is a recipe for disaster.

If this is not clear already, the majority of decisions made in this system will be of the concave nature, however, there are also a few decisions that are convex in nature and for those, decisions will continue to be made in a decentralized manner while offering the advantages of making such decisions in a centralized manner. This will be achieved by means of expert committees (Section 6.8.1.4) that are elected to power by the foundation. These committees will provide assessments of situations and of organizations based on their domain expertise and
effectively act as oracles that users can use to guide their voting decisions. Furthermore, if the results of any voting process deviate from the expected/recommended results from the expert committee by a certain margin, then those decisions will be flagged for audit and shall be reviewed by a specially elected committee or by a third party known as the arbitration committee (Section 6.8.1.5). These committees shall be elaborated upon in detail in their respective sections.

Decentralization as a means of censorship resistance should be fairly straightforward to see from arguments made thus far - both how it achieves this and why it is important. It is also a heavily studied topic. [188] If this is not clear thus far, then it definitely will become clear in forthcoming sections. The last case where decentralization matters is that of credible fairness. This is of particular concern in situations where voters depend on subjective judgements to inform their voting process and these subjective judgements are beyond the skills of the average voter thus, they need to be performed by a “trusted” third party. This also applies to virtually any scenario where decisions cannot be made on-chain in a trustless manner. To address concerns around trust in these scenarios, there is an unavoidable need for very robust governance. The governance mechanisms must be resistant to attacks and credibly convince a large and untrusting public that it is robust, trustable, and empowers people to maintain control. The topic of decentralized governance and how it achieves this is addressed in detail in Section 6.8.

There are several arguments that can be made both for and against decentralization, however, the specific system being built in this paper is one that benefits from decentralization in more ways than it suffers as a result of it. Anyone may feel free to disagree and poke holes in the (not so comprehensive) arguments presented but it is the stance taken here.

6.4 Membership

6.4.1 Authentication

When anyone joins the system, they will be required to provide KYC/AML information. This is done for a few reasons. Though the system will never need to act as a “custodian of assets” per the SEC definition of the term [126], the SEC has been on a recent (as of 2023) rampage of suing, suspending, and issuing Wells notices to any company involved with anything to do with blockchains and cryptocurrencies, so much so that their activity has been construed as “regulation by enforcement” and is being probed by congress. [127][129][130] Their primary argument is that anything that promises some kind of return to “investors” and can be traded, is a security and must be registered with the SEC. [128] The public nature of the goods being provisioned by the system should alleviate their concerns, however, there is a chance that even the slightest private involvement (Section 6.5.4) could trigger the SEC to investigate and potentially sue the system claiming that voting rights to govern a public good provisioned privately (even though these rights are not tradable) is equivalent to a security that needs to be registered with the SEC since people expect to derive utility from the good being provisioned. So, it is best to remain as compliant as possible (by design).
In addition to the SEC, the Treasury has also stepped in and made a formal statement: “DeFi services engaged in covered activity under the Bank Secrecy Act have AML obligations regardless of whether the services claim that they currently are or plan to be decentralized”. [137] As expected, the activity covered under the Bank Secrecy Act is very broad and open to interpretation, but the essence of this statement is that it implies that a transaction is not considered peer-to-peer if smart contracts are involved (and in this system, they are, as will be described in Section 6.5.2). The Treasury expects fully decentralized organizations that use smart contracts to facilitate transactions to run AML compliance checks and register themselves as exchanges. This is particularly concerning for the system proposed in this paper, but the establishment of the foundation and the DAO as separate entities, should allow for the foundation to handle any regulatory concerns, beyond which, approvals to operate within the limits of any regulatory framework may be transferred from the foundation to the DAO. These regulatory concerns are those that are pressing (as of 2023), however, as regulation is formalized in the coming decade, these issues should become clearer. Thus, it is important to note that any future references to specifications with respect to regulatory concerns in this paper must be taken with a grain of salt and those specifications may be modified per the judgment and discretion of those implementing the system based on the latest regulatory guidelines.

Regardless of how the legislation ends up playing out, collecting KYC/AML information is one of the simplest ways to comply with regulatory authorities. These organizations only require that they be given access to certain kinds of information when it is warranted by a court of law, as long as the facility for doing so is built into the system, then it should be compliant by design. On the flip side, those opposed to collecting this information are generally those that have something to hide or believe it is antithetical to the anonymous/pseudonymous nature of blockchains. The former are those who should not be using the system in the first place and the latter is a fairly shallow argument.

In addition to providing regulatory benefits, it also provides security benefits, especially during the early days. Sybil attacks (i.e., when an attacker subverts a system’s security/reputation mechanisms by creating a large number of fake pseudonymous identities to gain a disproportionately large influence) [131] are one of the simplest yet largest concerns when user bases have not reached certain critical mass thresholds. KYC/AML helps address this easily [163], especially if the system has phased rollouts starting with regions with stricter identity verification norms such as the U.S. In the U.S., one could require federal ID to be verified (via an automated process offered by companies like ID.me) to become a member of the system. There are caveats to this that are outlined in Section 6.11.1. So, the specifics are left to implementation, but collecting KYC/AML information by some means that also aids against Sybil attacks is strongly recommended, if not a requirement. There are a few systems already attempting to do this such as ProofOfHumanity, BrightID, and Idenanetwork that may be used for inspiration. [164][165][166]

6.4.2 Wallets

Once a user is approved to become a member, their KYC/AML information must be stored in a central database hosted by the foundation. At this point, the user should also receive a pseudo-randomly generated wallet address.
(for Ethereum) that is tied to their KYC/AML information. This mapping must also be stored on the same centralized database hosted by the foundation. This mapping is required for reputation mechanisms (Section 6.10) and tax benefits (Section 6.7.2). When the foundation ceases to exist post transferring all operational and administrative responsibilities to the DAO, this data could be stored in a secure distributed hash table (like Tapestry) [132] where each of the nodes are “computers” hosted by users. This infrastructure can also be relatively primitive and can be developed by the foundation in collaboration with the DAO. For example, if the DAO agrees that this information can be continued to be stored centrally then control of the central database hosted by the foundation will be transferred to the DAO (i.e., the ownership of a central database by the DAO is considered a sufficient proxy for true decentralization). Either way, details are left to implementation.

In addition to the wallet address initially assigned to each user, which will be their “primary” wallet address, a new wallet address is generated for each proposal (Section 6.5) they are involved with, to further enhance their pseudonymity. However, there is only a certain degree to which their “identity” can be hidden because their reputation information (Section 6.10) will be publicly accessible by everyone in proposals to which they have contributed; and the reputation scores will be the same across all proposals (for obvious reasons), so those who really want to identify individuals they are “working” with across multiple proposals will be able to do so by cross-referencing their reputation records. Furthermore, mechanisms to enhance community and conversation (Section 6.9) between users contributing to a specific proposal will reveal the same pseudonymous wallet IDs and reputation records as identifiers for each user; and even make those details as prominent as possible to build trust. A possible solution to this conundrum is to provide a percentage buffer within which everyone’s reputation scores are randomized when displayed to others so as to prevent direct cross-referencing and matching of reputation records across proposals. However, it should be clear how one could build an “algorithm” that analyzes all publicly available data on the system to determine who they are working with if they really wanted to do so (even with randomized buffers for the reputation scores). The reason this is of such concern is due to the risk of collusion (Section 6.11.2).

On the flip side, the option of using zero-knowledge proofs/encryption to secure identity information was weighed against the specification proposed here but ultimately decided against due to risk of money laundering (Section 6.11.1) and other illegal activity that could stem from offering such degrees of privacy. Once again, due to regulatory risks of the foundation being shut down by the SEC, DOJ and/or Treasury due to being unable to provide sufficient transparency into the mechanics underlying the system, it is best to stay away from complete anonymity and to once again, err on the side of being as compliant as possible (by design). In addition to regulatory concerns, this is somewhat of a gray area where it is hard to tell exactly what would work best in practice without further large-scale testing to find the right balance between anonymity/pseudonymity, prevention of system misuse/abuse, and community formation. It will likely be some mix of these properties that results in a stable equilibrium so further details are left to implementation.
6.5 Proposals

Now that users have been validated and have joined the system (Section 6.4.1), and also have a means to interact with the system (Section 6.4.2), what exactly are they interacting with? The very thing at the heart of this system - proposals. A proposal is defined henceforth as a project (generally a public good, but exceptions will be evaluated on a case-by-case basis) that a user wants provisioned. This would be something they believe others would derive value from and hence would be willing to commit resources (monetary and otherwise) to. Any user can create a new proposal at any time.

6.5.1 Approval

When a user creates a proposal, they can only specify in words what they would like to achieve, they cannot submit any estimations of cost (Section 6.5.3). These proposals may also be required to be submitted with a certain degree of detail. The proposals that users submit must be related to the provision of public goods. Outside these restrictions, a proposal can likely contain just about anything. The process of reviewing and approving proposals would seem like a fairly manual and labor-intensive process. However, with the current advances in LLMs (large language models), any kind of manual review from someone in the foundation, should only happen outside of traditional scenarios that would have otherwise required manual review (that would now be handled by an LLM) such as duplicate proposals, inadequate information, prohibited information (such as cost estimations), proposals that clearly appear to be jokes/scams etc.

All proposals can also later be flagged by users for manual review should any of those conditions not be met but the proposal still manages to bypass the LLM review process. To prevent misuse of this system, the penalties for doing so are extremely steep (Section 6.10.1.5) and should act as a sufficient deterrent. One could go as far as to also specify formal models and provide the LLM with the ability to evaluate proposals as satisfiability problems to guarantee accuracy in its approvals. Regardless, the extent to which the system needs to be automated will also be determined somewhat experimentally (i.e., how often do users submit proposals, what percentage of them are legitimate etc.) so details are left to implementation.

6.5.2 Tokens

Once a proposal is approved, users can contribute any amount (in USD) to any proposal. When money is contributed to a proposal it is directly routed to an exchange that converts the USD into USDC/USDT or another USD backed stablecoin. The total money raised for a proposal is held in the form of stablecoins within a smart contract, so everyone knows where their money is. All users will be able to see open-source code for the smart contract in question (for transparency and confidence) but will not be able to see the specific contract instance that is deployed to perform this action until after the initial funding round closes. The reasons for this are elaborated upon in Section 6.5.3. As previously mentioned, the reason the system must never handle USD and hold all assets as stablecoins in a smart contract is to not become a “custodian of assets”. [126] This is, once again, to avoid SEC regulation and remain as compliant as possible (by design). To clarify, the premise behind
avoiding any potential reason for the SEC (or the DOJ/Treasury/any branch of government) interpreting this system as being in violation of its rules is not because it could be, but rather because proving compliance is yet another overhead that someone must pay for, and once again, part of the core mission is to eliminate such overhead that makes the system inefficient.

If users wish to withdraw their contributions at any point before the proposal times out (Section 6.5.3), they can do so. However, once a proposal times out, their money is effectively locked in until the ensuing stages (Section 6.5.4 and 6.5.5) play out in their entirety. The following mechanism is somewhat redundant as it could be replaced by a simple internal bookkeeping mechanism but the decentralized nature of the system as well as the need to inspire confidence in users’ notion that their money is not “locked up” necessitates this. For each USD that flows into a proposal, a unique token (implementing ERC-20, ERC-777 or ERC-1155) is minted. These tokens are stablecoins in and of themselves, each of which is backed by a USDC/USDT/some USD backed stablecoin. These tokens are effectively a cryptographically secure means by which users can be assured that their money has gone toward a specific proposal and that should they ever want to (prior to the proposal timing out), they can withdraw their contribution instantaneously. They can do this via a local exchange on the system which will be the only exchange on which these tokens can be traded (despite technically being able to be globally traded as they implement ERC-20, ERC-777 or ERC-1155).

| Property 4: People will underestimate or overestimate what a public good is worth to them when asked to put a dollar value on it, and may do so intentionally, but even the most honest people will end up doing so unintentionally. Most people may be rational, but their rationality is bounded. They need to “put their money where their mouth is”, and there is simply no substitute for doing so. |

The mechanism described in this section allows for people to “put their money where their mouth is” but also alleviates their concerns as it pertains to ever needing to retrieve that money for any reason. This is elaborated on in Section 6.6 and forms the crux of the assurance contract currently being constructed through careful mechanism design.

### 6.5.3 Metrics

As previously mentioned, proposals will eventually time out. All users will be able to see a countdown timer on how much time a proposal has left before it closes its initial funding round. The length of time a proposal has, in order to raise money, is predetermined and the same for every proposal. This timer starts as soon as a proposal is approved, whether that happens automatically or through the manual review process (Section 6.5.1). In addition to the countdown timer, users will also be given some (likely imperfect and delayed) indication of the total funding level of a given proposal as well as the “momentum” a given proposal has. This information should help users determine the marginal impact of their dollars in addition to what has been already contributed to the proposal. There should likely be some kind of visual aid or calculator that illustrates this to users. This also serves as an imperfect proxy for the likelihood of success of a given proposal. The idea is that access to this information
will help avoid fragmentation. The fact that the impact of proposals with more money from more contributors is likely going to be exponentially higher than that with less money and fewer contributors or even simply that with fewer contributors, should help prevent thousands of nearly identical proposals from proliferating and simultaneously discourage support for proposals with only a few large contributors. It is important to note that this information cannot be provided directly i.e., the amount of money raised thus far, the number of contributors, a record of when money was raised, by whom etc. cannot be publicly stated because doing so would enable market failures (Section 2.2.3). The fact that everything is happening on-chain means obscuring these details will involve some effort. If an L2 or completely new blockchain is built for this system, then provisions to obscure these details can be provided from the get-go. However, if everything is run on an existing L2 or L1 (like Ethereum), then everything posted to the blockchain may be done in the form of batched updates that are encrypted using a zero-knowledge scheme that allows users to verify that their vote was cast as intended but reveals nothing else. This is not hard, but it is additional infrastructure that needs to be engineered.

The reason these details need to be obscured and only made available through some imperfect proxies are to root out free riders. The idea is to not set a required funding goal for any proposal (though there is one implicitly based on how much a proposal of similar scope would cost), so users are encouraged to honestly evaluate and put forth whatever personal value they would assign to seeing a given proposal actualize. For instance, if a given user would like to see a specific proposal realized, and they could see that it had nearly reached its goal, they could very well conclude that their contribution is unnecessary for that proposal to be completed and therefore withhold their contribution. This could create an undesirable dynamic whereby users associate the probability of a proposal being achieved with the probability of them incurring a personal cost, therefore stemming funding for a proposal as it approaches its goal. This is why the total amount of money that has been raised and the money that needs to be raised are both obscured. There are several other game theoretical scenarios leading to perverse incentives that would prevent proposals from reaching their funding goals if these details were made available. Obscuring these details helps eliminate those scenarios, or at the very least, disincentivize those perverse choices. Furthermore, not having a preset funding threshold allows for true evaluation of the value that society places on a good.

Property 10: Larger group sizes have a positive effect on contributions (frequency and magnitude), and seeing greater contributions only elicits more, but they have a negative effect on the tendency to not free ride and cooperate. At the same time, groups that are too small lack the bandwidth to effectively self-organize and govern.

But at the same time, as previously hinted at, some proxy is needed for users to evaluate the “health” of a given proposal, since evaluating a proposal which is unlikely to be completed as evidenced by low community contribution or lack of recent activity could aptly be considered a waste of users’ time. On the flip side, seeing the “momentum” a given proposal has and the general engagement with a proposal, when positive, elicits more contributions from those previously on the fence. The elements that convey this crucial information to users hinge on the “look and feel” of the system, a factor previously discussed in Section 6.1; and so specification details are kept fairly sparse, only outlining the goal and rationale of each of these elements. The bare minimum that
must be conveyed to users is something that indicates the scale of the funding achieved thus far, but in a manner that does not stem further contributions (i.e., it should not should not show progress toward a goal of any kind), because contributions increase with increasing thresholds at which the public good is provided while the probability that the threshold will be reached goes down. [75] The other important piece of information to be displayed is the number of people involved (without directly stating the number) because interestingly, contrary to economists’ expectations, group size can have a positive effect on contributions despite diluting the effect of marginal returns. [76]

6.5.4 Bidding

A marketplace has two sides yet only one has been discussed thus far. The second side comes into play once a proposal times out. Once this happens, there will be a period (predetermined and constant for all proposals) during which any organization can place bids on how much they would charge to provision what was requested in the proposal. The organizations placing bids can ask clarifying questions about the proposal which can be answered by those who contributed to the proposal. The exact mechanism for this clarification system could be as sophisticated as voting into power a representative or Twitter/Reddit style forums with upvotes but could also be as simple as the original user who proposed the project answering the questions. Regardless of the format, these communiques must happen in public, observable by all contributors. In fact, outside of the requirement that these conversations take place publicly, the specifics could even be left up to the contributors to the proposal themselves, because by this point, they will have the means to communicate (Section 6.9), and this could serve as a great first step toward establishing rules and guidelines for their self-governance (Section 6.8.4.2).

Regardless, at this stage, bids placed by organizations are merely estimates and do not need to be exact so while there needs to be some mechanism for communication between the bidders and contributors, it does not have to be perfect. Eventually the bidder that is chosen to provision the good will be able to join the same communication platform (Section 6.9) all the contributors are a part of. The bidders will also have a system by which they can submit “proof” of their work to make a case for why they must win the bid and provision the public good in this proposal. The bidders will also have their own reputation scores (Section 6.10.2) if they have at least worked on a proposal once prior to this. Once bids have been placed, if there are multiple bidders, the names of the bidders (and any “proof” of competency they submit along with their reputation scores, if they exist) but not their actual bids themselves, are revealed to contributors who must then provide a ranked choice preference list of which bidders they want provisioning the public good being requested in the proposal.

This is done to address a point of contention raised earlier in the paper (Section 2.4) which is the question of whether investment technology should continue to remain an important determinant of the optimal ownership structure of the public good. [41] If a proposal is provisioned by the lowest bidder, does that mean it is being provisioned with the best investment technology because they are able to provision it at the lowest price? But if a bidder charges more could that just mean they may end up doing a better job? This is almost definitely context dependent and will likely vary from proposal to proposal with no real means to make objective determinations
ahead of time such that they may be applied as a uniform policy to the entire system. So, the idea behind asking for a ranked preference list of bidders before the value of the bids are revealed is to try to get the “best” bidder to provision the public good regardless of what they charge. But how do voters know which bidder is the “best”?

This is a valid point of concern. All contributors to a proposal may not be equally informed about the space in which the proposal operates, nor can they be expected to possess this knowledge. If that is true, then how can they be expected to provide an informed ranked preference list of the various bidders for a proposal? As previously mentioned in Section 6.3.1, expert committees (Section 6.8.1.4) will provide assessments of industries and of organizations based on their domain expertise and effectively act as oracles that users can use to guide their voting decisions. In addition to this, as a safeguard, if the results of the voting process deviate from the expected/recommended results from the expert committees by a certain margin, then those decisions will be flagged for audit and shall be reviewed by a specially elected committee or by a third party known as the arbitration committee (Section 6.8.1.5).

6.5.5 Conflict Resolution

At this stage, users and bidders reveal their “hand”. The following pieces of information are revealed: the total amount of money raised, the value of each bid, and the ranked choice preference list of bidders (if there is one). There are a few possibilities once this happens.

1. There are no bids.
   a. A vote is proposed to extend the proposal for a fixed amount of time to give bidders another chance to bid on it. This period of time is predetermined and the same across all proposals. This vote is a special vote of bid extension that can be proposed exactly once. If the vote passes, then the proposal is extended (by that predetermined amount of time). All contributors’ money will continue to remain locked up during this period. If the vote fails, it proceeds as in 1b.
   b. The proposal terminates. All users get their money back.

2. There is one bid.
   a. The proposal raised approximately enough money to satisfy the bid. All contributors to the proposal get to vote as to whether or not the bidder is deemed fit to provision the public good. If the vote passes, then the money is paid out to the bidder in a phased manner as defined by all contributors. There will likely be some involvement from lawyers at this stage to define formal contracts that both parties will abide by. Currently (as of 2023), smart contracts do not hold the same legal precedence as traditional legal contracts, but should they do so, then smart contracts are an acceptable alternative.
   b. The proposal raised less money than what is needed to satisfy the bid. A vote is proposed to attempt to raise the remaining money in a certain period of time. This period of time is, once again, predetermined, and the same across all proposals. This vote is a special vote of second funding that can be proposed exactly once. If the vote passes then the proposal is extended (by that predetermined amount of time) to raise the required funds. If they successfully raise the
funds, the process proceeds as described in 2a. If they fail to raise the funds the process proceeds as in 1b.

c. The proposal raised more money than what is needed to satisfy the bid. The excess funds are given back to users in proportion to how much each user contributed and then the process proceeds as described in 2a.

3. There are multiple bids. Contributors have been asked to provide their ranked preference lists of bidders.
   a. The proposal raised enough money for at least one bid.
      i. The top ranked bidder charges approximately what was raised. The process proceeds as described in 2a.
      ii. The top ranked bidder charges more than what was raised. A vote (the special vote of second funding as in 2b) is proposed to determine if they should attempt to raise more money to reach their top ranked choice. If the vote passes, then the process proceeds as in 2b. If they fail to raise more money to reach their top ranked choice, contributors can vote to proceed with a lower ranked choice (for which they have money), the process then proceeds as in 2a, or they can vote to terminate the process as in 1b.
      iii. The top ranked bidder charges less than what was raised. The process proceeds as described in 2c where the top ranked bid is effectively the only bid.
   b. The proposal did not raise enough money for even a single bid. The process proceeds as described in 2b where the top ranked bid is effectively the only bid. After this process, if they manage to raise enough money for some bid but not their top ranked bid then the process proceeds as in 3a(ii) except they cannot propose a special vote for second funding again since that cannot be repeated. So effectively, they simply vote to proceed with a lower ranked choice, the process then proceeds as in 2a, or they vote to terminate the process as in 1b.
   c. The proposal raised enough money for all bids. The process proceeds as described in 2c where the top ranked bid is effectively the only bid.

This is an iterative process that may take a while to be fully resolved but upon doing so, if a proposal is successfully provisioned then once the final transaction takes place and all money is paid out to the bidder then all users involved with the process get certificates commemorating their participation and contribution (Section 6.10.1.6).

\[\textbf{Property 6}:\text{ Prosocial individuals care more about those around them than proself individuals and thus tend to be far more altruistic and willing to cooperate; but these orientations are not fixed. Prosocial tendencies increase when individuals are allowed to communicate, have a means to build and establish trust, and have universally agreed upon conflict resolution schemes.}\]

The mechanism described in this section will be guaranteed by smart contracts that automatically enforce the rules described above. This is one of the primary conflict resolution schemes in the entire system and blockchains.
allow for these to truly be universally agreed upon and enforced without the need for any manual human intervention.

### 6.6 Assurance Contract

An assurance contract [112] is a game theoretic mechanism and financial technology that helps facilitate the voluntary provision of public goods in the face of the collective action problem. In essence, the entirety of Section 6 thus far has involved designing a system that provides an assurance contract. Section 4.1 on mechanism design was also aimed at designing an assurance contract; and at the end of that section, it was concluded that it is impossible to design a mechanism that can mathematically guarantee four properties that would help create an assurance contract for the specific problem at hand. However, there is some hope given that empirical results did not ever back up these theoretical claims.

Instead of attempting to prove certain strong mathematical properties that end up not holding empirically when tested in the real world, logically reasoning through the set of problems that need to be addressed would likely prove more fruitful. The primary goal of this paper is to create a market for public goods to help the world become more Kaldor-Hicks efficient, and eventually, more Pareto efficient. But why markets? Simply because they are extremely powerful (despite their failures) and are a complementary solution to funding public goods rather than just relying on governments and incredibly inefficient funding mechanisms such as taxation (Section 2.2.1). However, there are several issues that plague free markets preventing them from facilitating the provision of the adequate (i.e., the economically efficient) level of public goods (Section 2.2.3).

A summary of the problems previously discussed follows. The first is the lack of valuation mechanisms that are incentive compatible and elicit truthful valuations as the dominant strategy resulting in a stable Nash equilibrium. The second is public goods having unremunerated externalities that result in deadweight loss and thus, a lack of provision of the efficient quantity of the public good in question. The third is the free rider problem, and large group sizes resulting in social loafing while attempting to collectively provision public goods.

To address the lack of a valuation mechanism there are two avenues. The first is to take advantage of state-of-the-art LLMs and fine tune them to ask a set of questions specific to a proposal and use a user’s responses to those questions as inputs to the Related Goods Method to generate an approximate valuation for the public good in question that the user can then choose to honor should they like to do so. This approach is definitely an

---

**Property 5:** There is no mathematically provable mechanism that is capable of raising adequate funds for public goods in an efficient, direct, and incentive-compatible manner. However, **these results do not hold in practice**, as individuals are not rational (bounded rationality), are often indifferent about lying, and have malleable preferences that change in repeated interactions with stable equilibrium.
The second is to improve on the Stated Preference Method. The primary issues with this method are as follows. First, individuals tend to fear public judgment when providing a valuation for a public good. This can be addressed by guaranteeing pseudonymity (Section 6.4.2). Anonymity is not guaranteed because research previously discussed showed that a reputation mechanism is beneficial for a plethora of reasons and pseudonymity provides that. In addition to which there are regulatory benefits to only guaranteeing pseudonymity and not anonymity.

Second, individuals have no “skin in the game”. When asked to provide a valuation for a public good, individuals may simply spitball the first number that comes to mind or a number influenced in any number of other ways. The only way to address this is to make them “put their money where their mouth is”. The obvious solution to this would be to get people to commit their money (Section 6.5.2) to a proposal when they provide a valuation. However, no one wants to lock up their money for an unknown period of time due to the time value of money (Section 2.2.2). So, a time constraint (Section 6.5.3) can be established.

However, once this is done, everyone would simply wait until the very last minute to make their contributions, again, due to the time value of money, especially in high interest rate markets (as of 2023). The reason this is not optimal, is because, as discussed in Section 4.2.5.1 and 6.5.3, individuals like to see engagement with a specific proposal before getting engaged themselves and the greater the engagement the more engagement from others it elicits. In addition to this, from the perspective of contributors, there is also the concern of individuals contributing money but needing to pull it out in an emergency and such transactions taking days, sometimes even up to weeks, when run through the traditional banking system. The solution to these problems is to use a blockchain (Section 6.5.2).

Third, people have an incentive to lie about their valuation in order to free ride. This third concern with the stated preference method overlaps very closely with the second and third market failures itself and may be addressed as follows. The incentive to free ride emerges from knowledge of the following facts and the ability to deduce information from these facts that would make free riding rational. They include, how much money needs to be raised, how much has been raised thus far, and who has contributed to a proposal. All of this information will be hidden from contributors (Section 6.5.3). Without this information, choosing to free ride would effectively be a random choice or one influenced by factors that cannot be controlled for (i.e., if the person is just personally
inclined to free ride and is indifferent to the consequences). The idea is to eliminate access to all the information that would make free riding an informed decision thereby eliminating a large part of the incentive to lie while simultaneously eliminating any reason for apprehension when contributing money to a proposal.

To reiterate, another way of looking at this is that individuals would not be willing to contribute to a proposal unless they expect others to do so as well. Essentially, individuals do not want to contribute to a proposal for fear that they will be left “exposed” to the risk that others will not contribute fairly. The system architected in this paper should alleviate this concern entirely; and it does. There are predetermined and fixed periods of time during which contributions and withdrawals of contributions can be made at any point with no penalty. During this time, everyone has access to the same information - enough to inform their decision to contribute but nothing that could ever inform the implications of their decision to free ride and not cooperate. So, people might as well make a reasonable and honest contribution within these established time frames in which no one has a leg up on another person or will be able to take advantage of them in an informed manner, if someone did manage to do so, it would have simply been a gamble.

### 6.7 Philanthropy

The system, as described thus far, has involved individuals making contributions to proposals of personal value to them. However, there is nothing stopping people from philanthropically contributing to proposals from which they will not directly benefit (though there are some restrictions as described in Section 6.11.1). As previously mentioned, the fundraising process is a means to determine the value society places on a public good, not just the value that those who will directly benefit from it place on the good. If society values a good at a certain price and there is someone who is willing to provision the good for that price, then the free-market system described in this paper will allow for the good to be provisioned.

**Property 3:** The responsibility to provision public goods often falls on communities who lack the infrastructure and funding to provision public goods themselves. Without infrastructure that helps them collectively organize around issues they care about, the transaction costs involved with provisioning public goods to address those issues is too high and cannot be met without special interest support.

To clarify, the system being proposed in this paper is fundamentally a private contributory system i.e., it does not entirely eliminate the free-rider problem, nor will it completely address the concern of “suppression of goods of great value but only to small communities with a lack of resources”. In fact, this was one of the key properties previously identified i.e., the responsibility to provision public goods often falls on communities that do not have the infrastructure and financial resources to provision these goods themselves. The infrastructure is now being provided; and the finances are also being elicited by allowing for large philanthropic donors to contribute to proposals. This feature is targeted at those philanthropists who have the resources and want to provision these public goods but simply do not know where and how to direct their resources as they often only have vast
amounts of money. These sponsors would love to contribute to a system that guarantees that the money is going to places that truly need it and that it is being spent in the most efficient manner possible [142][143][144]. This is likely why most uber-wealthy individuals tend to have their own charitable organizations. [149][150][151]

### 6.7.1 Charities

A charity is defined as an organization set up to provide help and raise money for those in need. It would seem plausible that there are unavoidable conflicts of interest, regardless of intentions, if the same party is responsible for both raising money and choosing how to spend it, with virtually no supervision. It would also make sense that people are hesitant to give their money to such organizations for the same reason, even some of the world’s richest people are. [142][143][144] Yet if these organizations are truly helping people, they should be able to raise the money they need, and people should be able to freely give to them.

Charities came into existence partly as a result of groups of people wanting to do good simply working together but also because they were a more effective, responsive, and flexible way of provisioning things that governments were expected to provision but simply could not or did not. However, as charities have gotten larger and have sprung up for virtually every imaginable thing, it is clear that they rely on motives that can be difficult to closely align reliably with the common good outside of the relatively small groups in which they were initially founded and designed to be effective within. [138] Charities, especially many today, often instead get captured by status motivations and parochial, even exclusionary interests.

The process of donating to charities is also outdated. Despite technology and software penetrating nearly every aspect of life, it has somehow not revolutionized the most structured institution to help others - charity. The biggest, and most challenging issue is that of charities not being held accountable, nor donors feeling the direct impact of their contributions. With typical donation structures today working very loosely, users throw their money, at worst, into a vague mission statement, or, at best, into a project without explicit terminus, specifications, or budget. This treats charitable donation as mere altruism, turning a blind eye to what actually motivates people to donate.

In truth, donations are no less a transaction than paying for groceries, only the value proposition for donors when compared to that of charities has been left entirely asymmetrical up until this point, where a donor gets little more than vague peace of mind and a tax write-off for their contributions, with the charity offering very little in return. What if charities were reimagined as an apparatus for people spending money on good they wanted to see done in the world, but could not necessarily accomplish themselves? This gets closer to what actually inspires donation. In fact, the system proposed in this paper provides a means to achieve exactly that. Anyone, who wishes to do good with the money they have but do not have the time, resources, and/or skills to accomplish that good themselves can use the system to find proposals they deem worthy of their money; and if charities were to now use the system to do the good they wanted to by bidding on proposals that are similar to projects they were already planning to work on, it would solve most of the problems with charities today.
If one wanted to go a step further, they could take a few pages out of the QF (Section 4.1.2) playbook to align themselves more closely with movements toward more democratic forms of charitable giving. The basic idea is for philanthropists and donors with large sums of money to simply provide that money as a fund that matches contributions to projects, scaled quadratically with respect to the number of contributors. The idea is to harness decentralized information to inform the philanthropist’s whims to give away their money. The Open Philanthropy and Effective Altruism movements are based on the idea that donor discretion must be eliminated from philanthropy to the greatest extent possible to make the process efficient. These forms of non-hierarchical democratic mechanisms for charitable giving are increasingly relevant as backlash continues to grow against the top-down dictates of well-intentioned but ultimately “elitist” class of donors with misinformed notions on where, why, and how their money must be distributed. [139][140][141]

### 6.7.2 Tax Benefits

All contributions made on the system will be tax exempt/deductible. However, it is not clear exactly which kind of exemption this system will fall under. Per the IRS, [145] organizations organized and operated exclusively for religious, charitable, scientific, testing for public safety, literary, educational, or other specified purposes and that meet certain other requirements are tax exempt under Internal Revenue Code Section 501(c)(3). Every organization that qualifies for tax-exempt status under Section 501(c)(3) is classified as a private foundation unless it meets one of the exceptions listed in Section 509(a). Private foundations typically have a single major source of funding (usually gifts from one family or corporation rather than funding from many sources) and most have as their primary activity the making of grants to other charitable organizations and to individuals, rather than the direct operation of charitable programs. Organizations that meet specified requirements may qualify for exemption under subsections other than 501(c)(3). These include social welfare organizations, civic leagues, social clubs, labor organizations, and business leagues. The nature of the system described in this paper is fairly complicated, and depending on implementation specifics of the system, it could be interpreted differently in a court of law. Hence, it is left to the foundation to acquire the appropriate certifications for all contributions made through the system to be tax exempt. To reiterate, there is no ambiguity as to whether this is possible, it most definitely is. The only question that remains is under which specific section of the IRS tax code this system’s contributions will fall under.
6.8 Governance

6.8.1 Liquid Democracy

6.8.1.1 Decentralization

Property 9: The ability to collectively govern a public good once provisioned is pivotal to its long term sustainability. Larger groups prefer elected leaders with more power than smaller groups do. Any voting system being used to elect leaders or make decisions must be strategy-proof. The governance system must instill confidence and the feeling that everyone’s vote matters (the mere belief that it does is sufficient).

The governance mechanism constructed must be resistant to attacks and credibly convince a large and untrusting user base that it is a robust and trustable form of governance that empowers users to maintain control. This is especially crucial in this system because unlike traditional blockchain based systems that are able to perform the majority of their operations in a trustless manner on-chain, this system has decisions that need to be made off-chain or decisions that rely on off-chain expertise. Whenever this is the case, addressing trust concerns requires a robust governance system. As previously discussed, (Section 6.3.1) decentralized governance is how this facility can be provided.

Decentralization may be new, but governance is not. As seen in Section 4.4, there is much to learn about governance from history but decentralized governance must not overly rely on the past, for it will suffer the same problems, only exacerbated due to the rapid pace and unique interweaving of public and private elements, whether that be the public nature of for-profit organizations or the open-source nature of systems where customers are the owners in blockchain based systems. The pitfalls of significant reliance on historical forms of governance for decentralized organizations include low participation and concerns about weak oversight and interest-group capture.

The challenges that decentralized governments face are of the same flavor as those faced by societies and organizations for millennia, spanning from the Athenian Ecclesia where citizens collectively made policy decisions, to the Dutch and British East India Companies which distributed risk and capital at scale by adding legal institutions between shareholders and creditors with privatized governance thereby giving rise to the modern corporation. For decentralized governance, such innovation involves, but is not limited to, moving from direct and representative democracy to newer forms of democracy such as liquid democracy to mitigate concerns of low voter participation and asymmetric information that create risks of interest-group capture, building explicit governance institutions beyond rudimentary token-based voting systems to represent all stakeholders equally, and empowering delegates with the tools to build the trust needed to act as representatives while allowing for actions of all users to be audited to help build trust.
6.8.1.2 Historical Governance

The primary problems with representative democracy are that of rational ignorance (Section 4.4.1) and the issue of concentrated benefits and diffused costs (Section 2.2.1). Yet virtually all governing bodies today, including publicly traded corporations, govern via representatives. It is clear that this system works well, and some argue that by asking less of voters (only to vote into power representatives every few years) that representative democracies empower them more. However, anecdotally, no one is ever really pleased with governance models today, especially as it pertains to public governance. Why is that?

It has to do with the fact that the model being used is too strict i.e., either direct democracy where every single person has to vote on every single thing or representative democracy where everyone can only ever vote for representatives every few years who will then vote on the actual issues that matter. Why does it need to be one way or the other? Liquid democracy solves this problem. It is a form of delegative democracy [173], whereby a group of individuals engage in collective decision making through both direct participation and dynamic representation. [173] It utilizes elements of both direct and representative democracy. All voters have the right to directly vote (direct democracy) but also have the option to delegate their votes to someone who will vote on their behalf (representative democracy). [174] The governance mechanism used here will be derived from liquid democracy.

6.8.1.3 Delegating Votes

In this system, any voter may delegate their votes to specially designated voters who have sufficiently high reputation scores (Section 6.10) that it gives them the opportunity to receive delegated votes and vote on others’ behalf. These special voters are called proxies. The delegation of votes may be absolute i.e., an individual divests their vote to a proxy across all issues, policy-specific i.e., an individual divests their vote to a proxy only when the vote concerns a certain type of issue, or time-sensitive i.e., an individual decides to divest their vote for a period of time. [175] Regardless of how votes are delegated, they may be retracted at any time until voting is closed. The appeal of the retractability mechanism is that it alleviates any concerns of proxies misrepresenting the views of their “constituents” and also imposes increased accountability on proxies compared to traditional representatives that have fixed terms. To prevent proxies from misrepresenting the views of their “constituents”, proxies will be required to make their stances either completely publicly known or at least known to those who delegated their vote to the proxy. This will be done in a manner that is binding via some smart contract i.e., once a proxy makes their stance known, they cannot change their stance. There is not much risk to revealing this information (and it thereby leaking the way in which a vote may swing) because no one knows how many votes were delegated to a proxy until after the voting process has concluded. The instantaneous retractability mechanism ensures that no one can be sure of how many votes get delegated to a proxy despite access to historical records (though they could make an educated guess). The retractability mechanism also means that proxies are essentially re-evaluated every time a vote is cast, and users can withdraw their delegations at the first sign of incompetence i.e., the principal agent problem is solved with a flywheel of accountability that rotates much faster than it ever did in representative
democracies. This flexibility also allows users to be involved with proposals as much as their life permits. They could be completely involved one week but delegate their vote to a proxy the following week.

This form of governance does run the risk of devolving into a type of meritocracy with all votes being delegated to those with knowledge on specific subjects or with special experience. Interestingly enough, regardless of whether this form of meritocratic delegation is normatively appealing, a study found that the idea of less informed citizens exclusively delegating their votes to more informed citizens was positively undesirable. [179] In fact, in issues where there did exist a “ground truth”, a subset of supposedly more informed voters within a larger populace would be less adept at identifying the ground truth than if every voter had just voted directly. [179] At the same time, as previously discussed, direct democracy has its own challenges and, in a system where everyone is simply volunteering their time and effort, liquid democracy seems like the appropriate compromise.

6.8.1.4 Expert Committees

While the meritocratic nature of liquid democracy is definitely a reason for concern, it is not as pronounced as it would be otherwise because in this system, votes are being delegated to trusted users (as measured by their reputation score), not domain experts or specially informed ones. As has been previously mentioned in Section 6.3.1 and 6.5.4, there is a separate mechanism, that of expert committees, that helps address this concern. These committees will provide assessments of industries and of organizations based on their domain expertise and effectively act as oracles that users can use to guide their voting decisions. For example, in a given niche, say construction, there could be five different expert committees all of which will independently and simultaneously provide their assessments of the bidders in the space. The final assessment is the median/mean of the individual assessments and deviations between assessments beyond a certain degree are investigated by an arbitration committee (Section 6.8.1.5).

Why not have these expert committees just make these decisions instead? Why do they have to make assessments? Why do people need to inform themselves using these assessments before voting? This goes back to the age-old argument of a government of the people vs one that is able to make informed decisions. Creating representative systems has always been at odds with creating informed ones. Socrates’ distaste for democracy stemmed from the belief that governing required expertise and should not be left to those without it. The famous analogy of the “ship of state” where he argued that just as one would not fire the navigator on a ship and let the unexpert crew navigate, the governance of society should not be left to the supposedly unexpert members of the community. However, the opposite extreme of autocratic rule by self-appointed experts is clearly undesirable and incompatible with free societies as well. As previously discussed, in addition to not being desirable, empirically this was also found to not be ideal. In general, democracies prefer indirect accountability.

The use of expert committees follows this principle of indirect accountability. In governments and society at large, people with the relevant skills and expertise work as full-time employees of the state, not having to seek election themselves, but being subject to appropriate checks and balances from those elected to power. This has two benefits. It helps recruit those who lack the interest to become an elected representative but have the skills to
help the state and it also creates a degree of separation between these employees’ work and the sometimes myopic pressures of the electorate. Corporate governance works similarly. The directors of the board are, loosely, akin to elected delegates of the shareholders at large. They in turn oversee the executives who run the company. The executives do not themselves have to run for election, but at the same time, they are accountable to shareholders indirectly. The board of directors is not expected to be fully expert in the day-to-day decisions that the executives make for the firm, but they are expected to evaluate whether the executives are doing a good job overall.

The same holds for the system here where expert committees can help guide the decisions of voters. It also acts as a safeguard against attacks of many kinds (Section 6.11). If the results of the voting process deviate from the expected/recommended results from the expert committee by a certain margin, then those decisions will be flagged for audit and shall be reviewed by a specially elected committee or by a third party known as the arbitration committee (Section 6.8.1.5). In the same vein, should voters ever feel like they are being misled, then expert committees may be replaced. There should be a mechanism by which if a sufficient quorum is reached, the foundation will be required to replace an expert committee; they must also do this if the committee is ever found to have misled voters.

6.8.1.5 Arbitration Committee

The arbitration committee could be a sub-committee within the DAO, it could be a special committee voted into power by the DAO or could even be a third party. There are benefits and drawbacks to each of these approaches and evaluating those comprehensively are beyond the scope of this paper. This function will initially be performed by the foundation, however, eventually an alternative for succession must be decided upon. The primary role of the arbitration committee is to have the “final” say on any concerns regarding security. For the purposes of this paper, whenever there is a security risk, a proposal will be flagged for audit, and the audit is performed by this committee. There have already been several occasions mentioned thus far where proposals are flagged for audit, and this will continue to happen over the course of the rest of the paper. The arbitration committee’s judgment on an issue is “final”; only because all of their judgments will be valid but in flux, for a fixed period of time, during which users can refute the committee’s decision if they reach a sufficient quorum. If the committee’s decision is refuted, then the foundation will review the details and potentially get a second opinion from (another) third-party that acts as a tie-braker against the current judgment. If the foundation finds any signs of malpractice from the arbitration committee, then they shall be replaced in their entirety.

The majority of cases that the arbitration committee presides over will be those that have to do with deviations between the outcomes of voting processes and those that are predicted/recommended by expert committees. This implicitly also means that the arbitration committee has some degree of oversight over the expert committees and must periodically ensure there are no security breaches on that front. There have been attempts to create generalized arbitration protocols such as Kleros [189] which once again have certain benefits and drawbacks. The possibility of outsourcing the arbitration committee to another decentralized protocol is also worth considering. Regardless, the details here are left to implementation, and to a greater degree, left to the community since during
the initial days, the arbitration committee will be within the foundation and subsequently community members can help decide how such a committee may be formed and overseen.

6.8.2 Token-Based Governance

The majority of decentralized governance systems are token-based and as a result are susceptible to attacks that effectively destroy the ecosystem they were previously governing. Whether that was the Beanstalk flash loan attack [192] or the hostile takeover of Steem and subsequent exodus of Hive [190]. Yet, token-based governance continues to be the dominant form of decentralized governance. This is because it feels credibly neutral, not permissioned or gated in any way; anyone can simply go and buy governance tokens to become a part of the system. In practice, however, token-based governance only appears secure and neutral as a result of one of its imperfections, namely that the supply of governance tokens is often in the hands of a few whales who happen to be genuinely vested in the interests of the systems they help govern. However, these governance systems have not really been stress-tested in mature ecosystems and financial markets, so one should not be lulled into a false sense of security about this method’s resiliency. In addition to which, the often cited solution of off-chain governance as a means to recover from exploits of on-chain governance via means of hard forks is not applicable to a system such as this one where the system is likely going to be built upon an existing L1 or L2 as an “application” and so there is no real notion of a hard fork.

6.8.2.1 Incentives

The primary concern with token-based voting systems is that of the many ways in which it can be attacked (Section 6.8.2.2), however, there are issues with token-based voting systems even in the absence of attacks. [160] These issues generally fall into one of the following buckets. The first is that small groups of wealthy individuals, the so-called “whales”, are better at successfully executing decisions than a large group of small holders. This is because of, as previously discussed, concentrated benefits and diffused costs (Section 2.2.1) as well as rational ignorance (Section 4.4.1). The second is that it incentivizes token-holders to prioritize the interests of themselves at the expense of the community. It is impossible for large communities to all hold the same values, visions, and goals - some will definitely come at the expense of others. This is especially true when token-holders overvalue the goal of making the token value go up (for their personal monetary benefit) even if that involves harmful rent extraction. The third is conflict of interests. This primarily applies to investment funds that hold vast amounts of a particular token, but as they are an investment firm, they also likely hold vast amounts of tokens of competing protocols; and as activist investment firms have done repeatedly, they will often purchase a controlling stake in a company only to downsize it or restructure its business. This is a time-tested strategy to help shape the competitive landscape of an industry in which the firm owns multiple competing businesses; because as long as they come out on top at the end, they are not concerned about the impacts on the firms they invested in. [176][177][178]

6.8.2.2 Attacks

In addition to these issues, there is the most concerning one and that is its primary vulnerability to attacks via vote-buying (and bribing). The fundamental vulnerability with token-based systems is that the tokens in the
protocol bundle two rights into a single asset. It is both a means of representing some kind of economic interest in the protocol’s revenue but also the right to participate in governance. The combination is deliberate. It is a fairly effective means to align power and responsibility. However, it is quite easy to unbundle these two rights from each other, and doing so can have disastrous effects.

Consider a wrapper contract with the following rules. Depositing one unit of a governance token XYZ gives back one unit of a wrapped governance token WXYZ that can be redeemed back for the original governance token XYZ at any time, but as long as the governance token is held in its wrapped form, it accrues dividends. These dividends come from revenue generated by auctioning off the governance power of the underlying unwrapped token XYZ that the wrapper contract possesses.

For the majority of holders of a governance token, it is in their best interest to deposit the token into such wrapper contracts that generate revenue for them on a consistent basis. There are large swaths of voters who hold sufficiently small amounts of the token that voting is not worth it due to concentrated benefits and diffused costs (Section 2.2.1) as well as rational ignorance (Section 4.4.1). Furthermore, even if the wrapper contracts were issued by attackers and used to perform an attack on the system then the perceived impact on each individual is very small but they would have been gaining the full benefits of the dividends from the wrapper contract thus far. However, it is only the perceived impact on each individual that is very small but when the impacts of an attack are applied to everyone involved then the net cost incurred by each individual is actually much higher than what they were compensated for.

Suppose that an attacker decides to corrupt a specific DAO vote in their favor and the harm to each individual who is a part of the DAO is $D$, the chance that any individual’s vote affects the outcome is $p$, and the attackers make a bribe of $B$. This example is illustrated as though the attacker is offering voters a bribe instead of tricking them into depositing their tokens into wrapper contracts that effectively do the same thing i.e., this math also holds true in the vote-buying case where $B$ is the dividend issued by the wrapper contract.
<table>
<thead>
<tr>
<th>Decision</th>
<th>Benefit To Individual</th>
<th>Benefit To Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accept Bribe</td>
<td>$B - D \cdot p$</td>
<td>$-999 \cdot D \cdot p$</td>
</tr>
<tr>
<td>Do Not Accept Bribe</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

If $B > D \cdot p$ then an individual will be inclined to accept the bribe, but as long as $B < 1000 \cdot D \cdot p$, accepting the bribe is collectively harmful to the group. So if $p < 1$, and it usually is far below 1 (as mentioned previously, there is only a one in ten million chance that one vote could sway the outcome of a United States Presidential election even in a battleground state [42]), then an attacker could bribe enough users to vote in their favor, compensating each user far less than the collective harm they suffer.

An obvious question that arises is whether users would be willing to accept bribes? Especially in a system such as this one where the goal of the system is to just make the world a better place, it would be hard for someone to so selfishly and blatantly sell out just for some money, right? However, again, the point being missed is that these bribes can be obfuscated in so many different ways. The simplest of which is the wrapper contract described earlier but another example that is in some ways equally simple is that of borrowing from DeFi lending platforms. Though the underlying mechanism differs for flash loans, those attacks are also of a similar vein.
A user could take any cryptocurrency they own, say ETH, and lock it up in a CDP (collateralized debt position) in one of these DeFi lending platforms. Once that has been done, the user can use the CDP to borrow any token of equal value, including a governance token. They can then use the governance token to do whatever they want, and once they are done, return the token back to the CDP, along with some interest, to retrieve their originally deposited ETH. Throughout this process, the borrower has no financial exposure to the governance token so they can use those voting powers to do something that destroys the system, yet it is of no concern to them (or the DeFi lending platform) since they only need to return the token they borrowed (similar to how stocks are borrowed when taking short positions). On the other side of this equation are those that are depositing their governance tokens into a system that gives them some dividends for doing so. Oftentimes, the platforms that offer this will make it seem like they are just safe custodians of customers’ assets, not that they are loaning out these assets when in possession of them. Thus, the unbundling of the two rights have been achieved. The borrowers have governance power without economic interest and the lenders have economic interest with no governance power.

Lastly, in addition to these explicit means of taking advantage of small holders with no real incentive to participate in the voting process, those that deposit their tokens (governance or otherwise) with non-custodial exchanges run the risk of the exchange using the collective governance power of all the tokens they hold to vote in favor of things in their interest. This is clearly a step further than obfuscating loaning out tokens to others. This happened with the hostile takeover of Steem [190], where exchanges used their customers’ tokens to vote for proposals that the community strongly opposed; and this allowed for a takeover of the Steem network. This was resolved via off-chain governance resulting in a mass exodus of the community to a new chain called Hive that was forked from Steem. [191]

These flash-loan or buy-then-vote-then-sell attacks can be slowed down by time-lock mechanisms but ultimately these too can be bypassed [193] by wrapper contracts, and once again, more trivially, a centralized exchange. These kinds of attacks are common but have not plagued every single DAO using token-based governance despite there being a surprising lack of defenses against these kinds of attacks. However, this is primarily because currently, the majority of DAOs are sufficiently small that selling each other out is stopped by social dynamics and close-knit interpersonal relationships, most DAO members are sufficiently technically sophisticated that simple mechanisms such as wrapper contract based attacks do not work well often, those with high concentration of tokens/wealth are actually invested in the well-being of the system, and liquidity in DeFi lending markets is fairly low. However, over time, these factors will naturally only move in the direction of supporting such attacks, not preventing them.

6.8.3 Reputation-Based Governance

There are ways of addressing the drawbacks of token-based governance and the solution deployed for this system will use a combination of these features, all of which are backed by a reputation mechanism (Section 6.10).
6.8.3.1 Limited Governance

The first is the notion of “limited governance”. It involves assigning varying levels of responsibility to stakeholders, depending on what they have at stake. In addition to which, governance could be limited to fixed-parameter choices i.e., based on normal voting processes, changes within certain limits can be made, while anything outside those predefined limits will require a special voting process with time delays to allow the entire community to reconsider decisions they may not agree with and handle them appropriately. Another interesting approach is that of “un-governance” [162] which defines a roadmap for automating away all possible functions that are performed through governance and limiting voting and governance to exclusively those functions that cannot be performed by any automated means whatsoever.

In this system, the reputation mechanism has various tiers, each of which has multiple levels, this is elaborated upon in Section 6.10.1.4. As a user continues to engage with the system, their reputation score rises, and they move through the ranks to higher tiers. The higher their tier, the more responsibility they get. The rationale behind this approach is that those with higher reputation scores have more to lose should they try to exploit the governance system and be caught. The security mechanisms in place all but guarantee that even if someone is not caught immediately, they eventually will be, and the foundation can slash their reputation at any time. The fact that the only way to grow one’s reputation is by committing time and effort to the system and consistently engaging with it should make it very valuable to those who have high scores, and those are the only users in positions to potentially compromise the governance system. In addition to providing this security measure, the tiered reputation system also provides something to work toward for users i.e., it gamifies the system in some way.

The idea of fixed-parameter choices may also be implemented. As previously discussed, there are many scenarios where the outcomes of votes are cross-referenced with expected outcomes as predicted by expert committees (Section 6.8.1.4) and when there are significant deviations, the arbitration committee (Section 6.8.1.5) steps in to resolve them. These are fixed-parameter votes because based on normal voting processes, changes within certain limits can be made (the limits are the accepted deviations from expert committee recommendations), while anything outside those predefined limits will be reviewed by the arbitration committee. The arbitration committee’s decisions can then also be reviewed by the community during a fixed period of time after a judgment is made. In essence, nearly all decisions are either fixed-parameter, or not permanent and have a fairly large window of time during which they may be appealed.

The last idea of “un-governance” is very interesting and worth considering but it is beyond the scope of this paper to define a formal roadmap to achieve this. It is also unclear if explicitly predefining such a roadmap is more useful than just operating under the implicit assumption that the goal is to decentralize (and hence automate away) as much of the system as possible. This is a very new concept that is only just being tested (as is everything discussed in this paper, to a certain extent) so the details on this front are left to implementation and are to be informed by developing literature in the space.
6.8.3.2 Vote-Buying

The solution to eliminate incentives for vote-buying and bribing is to not use a token-based governance system, but if not tokens, then what determines the weight a particular vote has? Reputation. For starters, every voter in this system is verified to be a unique human being i.e., this is a proof of personhood system. Then comes a person’s reputation (Section 6.10). This mechanism is extremely important. The biggest challenge with a reputation system is the balance between money and participation as well as what counts as how much participation. This is addressed in Sections 6.10.1.1 and 6.10.1.2. In addition to this, addressing the risks of collusion are also very important. These mechanisms must be architected such that the systems being used to measure reputation continue to remain non-financial and do not end up inside wrapper contracts or at some centralized entity that can sell reputation, and hence, governance power to the highest bidder.

While the notion of vote-buying in token-based governance systems was broached earlier, the distinction between trustless and trusted vote-buying transactions was not. A transaction may be considered trusted if the buyer and seller can be certain that the other will behave honestly, a trust that usually comes from some kind of social relationship. If they cannot trust each other in this way, then the transaction is trustless. The reason this distinction matters is because reputation systems have a strong built-in defense against trustless vote-buying transactions. The most naive example is that of a user selling access to their account then the buyer simply has to trust that despite the seller having sold their private key to the buyer, the seller will not use it to make votes that are ones that the buyer does not agree with (or even simply change their private key after the sale has concluded). This type of vote-buying clearly requires trust. However, due to the pseudonymous nature of the system and the many restrictions on communication that are outlined in Sections 6.9 and 6.11, it will be virtually impossible for such trust to exist in transactions between users using this system. In fact, if it is possible to restrict all communication to the local communication platform then they may not even be able to set up such a sale let alone execute it. There are more details on the many restrictions on the communication system as well as the rationale for such restrictions in Sections 6.9 and 6.11.

Further, mechanisms to detect shared private keys can also be implemented, for example, if the same private key is used to perform certain actions in quick succession from locations that are very far away from each other, and while location spoofing is possible, employing techniques similar to that Netflix has employed to detect users sharing accounts with those who are not family or those they live with, can help identify any shared private keys. In addition to this imperfect automated means to detect the sale of private keys, the mere fact that either the buyer or the seller could snitch on the other at any time should suffice to game theoretically prevent the trustless sale of private keys (especially if there are incentives to encourage reporting of any such sale). This in combination with the fact that the foundation can slash reputation scores at any time should act as a sufficient deterrent to attempts to buy votes in a reputation-based system.

Despite blockchains themselves being a solution to help facilitate trustless transactions, and a potential solution to the key sharing problem being escrows and smart contracts to transfer ownership of an account securely for
limited periods of time, a blockchain does nothing to prevent the aforementioned risk of snitching, especially if there are incentives to do so. Regardless, as with all things that have the ability to affect something as important as the governance of a system, people will attempt to find loopholes.

The use of reputation mechanisms simply slows this down significantly by preventing the straightforward setup of trustless transactions and threatening any discovery of such vote-buying transactions with a high cost punishment - slashing of one’s reputation and since voting power is correlated with reputation scores, it will be extremely unlikely that an attacker is able to convince any user with a reasonably high reputation score (i.e., one whose vote has the power to exploit the governance system, as described in the prior section on limited governance) that selling their vote is worth it i.e., is worth the risk of losing their reputation.

6.8.3.3 Bribing

In addition to vote-buying, there is also the risk of bribing. In fact, for reputation-based mechanisms, bribing is a safer and more effective means of guaranteeing that a given voter votes in the favor of the attacker. This is because it can theoretically be guaranteed on-chain by means of a contract that escrows the money to be paid out and reads votes once a proposal passes and distributes rewards pro-rata to those who voted in favor. In fact, the worst-case scenario of this form of attack is when a contract is constructed that nominates a proposal to send all funds from a DAO’s treasury (thankfully, there is no such notion of a shared treasury that the DAO controls in this system) to that contract, and then distributes the funds received pro-rata to those who voted yes.

However, similar to vote-buying attacks, there is an element of this system that is a built-in defense against bribing. It eliminates the ability to use a completely trustless method (such as smart contracts) of facilitating these bribes. As mentioned in Section 6.5.3, when votes are being cast in real time, all details of the votes need to be obscured. If everything is run on existing L2 or L1 (like Ethereum), then everything posted to the blockchain may be done in the form of batched updates that are encrypted using a zero-knowledge scheme that allows users to verify that their vote was cast as intended but reveals nothing else. This means that those setting up bribery contracts will have no way of verifying whether votes were cast as intended during the payout process since only each user can personally verify that their vote was cast as intended. This means that there is no trustless mechanism involving the exclusive use of a smart contract to facilitate a bribe.

This once again leads to the exact same scenario as that discussed in the prior section which is that some form of external trust is needed for this kind of attack to be executed successfully. The attacker needs to trust that the bribed voter will in fact vote in their favor and make an advance payment to the voter, or the voter must trust that the attacker will pay them once they provide manual proof that their vote was cast as intended. To make the latter even harder, the ability to verify that one’s vote was cast as intended may be delayed for a certain amount of time (to raise the bar for the level of trust needed for a bribery transaction) since the chances that the system casts a vote incorrectly are very slim (given the system’s code base will be open sourced) so honest users may not even feel the need to verify that their votes were cast correctly all that often, and even if they did, they would likely be okay with a short delay before they can perform this verification.
6.8.3.4 Asymmetric Consequences

The last mechanism used in the governance system will be that of the negative side effects of a bad governance decision asymmetrically impacting those who voted for it more so than those who did not. In token-based systems, when a governance decision goes awry, everyone’s tokens drop to zero but what if only those who contributed to that decision were affected? In this system, only those who contribute to bad decisions, and in fact, repeatedly do so, will be the ones who get punished for it (Section 6.10.1.5) by losing their reputation and eventually being completely banned from ever using the system again; and since their user identity is tied to their government identity, they will truly never be able to execute any kind of attack again if they ever do get a lifetime ban.

The reputation mechanism (Section 6.10) punishes a lack of engagement with the system as well as actions that are explicitly ones meant to derail the system in some way. However, that only works for guarding against decisions that are truly extreme. What about smaller-scale decisions that get flagged for audits but do not have any negative impacts (beyond causing some kind of delay as a result of the audit)? If such decisions are not punished, then the reputation mechanism unfairly favors attackers manipulating the system but just not at a scale that is ruinous. Furthermore, an explicit lack of punishment for failure to vote (though this is encompassed within the punishments for lack of engagement) may encourage means to participate just enough to not be affected by such punishments for lack of engagement but still impede progress by not voting and/or not delegating votes to proxies. The solution to this is to implement a futarchy. [171]

A futarchy essentially converts votes to bets. So, voting in favor of a specific bidder is the same as making a bet that the bidder is one that will successfully provision the public good in question; and similarly, voting against is the same as placing a bet that the bidder will not be able to. As will be specified in Section 6.10.2, once the good has been provisioned, all contributors vote on the performance of the bidder and the results of that voting process are used to determine the bidder’s reputation score. This can then be factored back into each user’s own reputation score. If the bidder scored well and a user voted in favor, then their score receives a boost but if the bidder did poorly and a user voted in favor then their score is decreased.

However, there is something cyclic about this process i.e., if a user voted in favor of a bidder and the bidder happened to not provision the good per their specified contract then maybe the user accepts the money they put toward the proposal as a sunk cost and would prefer to not also take a hit to their reputation score so the user says that the bidder did a good job even if they did not. This issue is hard to overcome without making the bidder evaluation process something performed externally by a third party.

However, the benefits of having a community driven reputation system for bidders outweighs the benefits of implementing a futarchy for users. A futarchy is also very close to the line of where the reputation mechanism becomes overbearing and discourages users from using the system because their every action will be penalized. This is definitely a balancing act and there is a chance that the benefits of a futarchy outweigh that of the reputation system currently designed. So, details are left to implementation.
6.8.4 DAOs

In this system, governance will initially be led by the foundation but will eventually transition to the DAO. The goal is to have decentralized governance, but the nature of the system necessitates some centralization early on. This is not to say that complete decentralization from the get-go is impossible, as has been demonstrated by YFI, Compound, Synthetix, Gitcoin etc. [156][157][158][159] but it is hard and is even harder due to the specific requirements of the system at hand. For this reason, complete decentralization from the very start is not proposed here, however, that does not mean that it needs to be eliminated from consideration, the choice is left to implementation. Regardless, eventual decentralization is very much a part of the goal of the system, and the DAO helps achieve that.

6.8.4.1 The DAO

6.8.4.1.1 Responsibilities

This refers to the organization that will assume the responsibilities of the foundation after a certain point in time and is generally responsible for running the system. This role encompasses a very wide range of responsibilities, many of which have already been mentioned and there will be many more that will be mentioned in subsequent sections. The list of responsibilities outlined in this paper are not exhaustive but merely those responsibilities that happened to be of concern given the topics broached in this paper. A good rule of thumb would be to use the kinds of responsibilities outlined in the paper as a guideline for the general tasks that the DAO will be responsible for. This is being kept intentionally vague because the exact set of responsibilities that the DAO will take charge of will change with even the smallest changes to the structure of the system being built so even summarizing the list of responsibilities here would be misleading.

6.8.4.1.2 Features

While those are the responsibilities that the DAO has, what are the privileges it is offered in order to fulfill those responsibilities? Once again, this will be kept abstract by listing eight qualities that the DAO must have in order to be able to do what it needs to. This list is inspired by the Rochdale Principles for the operation of cooperatives. [155][195]

Autopoietic. The DAO must be able to reproduce and maintain itself quickly without a single point of failure. The mythical many-headed hydra is an apt metaphor for DAOs - for every head chopped off, it regrows two. [194] The point of a DAO is that there is no notion of a centralized node which when compromised brings down the whole system. It must be resilient.

Alegal. The DAO must have an organizational structure sufficiently well formalized [196] that its basic characteristics may be mapped to a traditional legally incorporated organization i.e., it must have functional equivalence with the legal structure of traditional organizations. The DAO need not possess such structure at the very start but must at least work toward it.
Hyper scalable. While a DAO may be more inefficient than a corporation at inception, as it scales, its efficiency must scale with it (unlike traditional corporations) to reach large membership sizes. [197] DAOs must defy the theory of the firm and they need to possess a toolkit that allows them to provide for fractal membership growth i.e., each addition to the organization fits in modularly and improves upon the synergy of the whole.

Executable. The DAO must have the ability to run the majority of its operations in a trustable, sustainable, and fully automated manner i.e., via smart contracts. The DAO must also have access to a vertically integrated toolkit (Section 6.9) that allows it to run an organization in a resource efficient manner that would be impossible for traditional corporations.

Permissionless. The DAO must allow for the entry and exit of any individual according to some public criteria. Anyone can join the system as long as they are human (Section 6.4.1). Anyone can join the DAO as long as they meet some predefined criteria (more on this below). The DAO is permissionless but gated to protect against governance exploits.

Aligned. The DAO must have an incentive structure that enables it to encourage groups of individuals to overcome the collective action problem. It must have the ability to apply mechanism design through economic incentives and social norms [198] to achieve consensus while maintaining liveness and safety properties. [199] Community and culture are equally, if not more, powerful tools. [200]

Co-Owned. The DAO exists purely for governing purposes and has no greater claim on ownership of the organization than any other member of the system.

Mnemonic. The DAO only works publicly and there is a permanent record of their activity on the blockchain. This is for both transparency and knowledge sharing and transfer.

6.8.4.1.3 Structure

As previously mentioned, the reputation system will have tiers (Section 6.10.1.4). The higher the tier, the greater the responsibility. To even be eligible to be a part of the DAO, users must reach certain reputation tiers that demonstrate their unwavering commitment to the system. Once a user reaches such a tier, they may nominate themselves to be a part of the DAO. If their nomination receives a certain number of supporting votes, then they are subject to a system-wide vote that chooses whether they get elected to the DAO. There will likely be a minimum quorum of votes (Section 6.8.5) that needs to be reached before any decision can be made.

Once a user is elected to the DAO, there is no fixed term, nor is there a minimum guaranteed time for which the user will be a part of the DAO. At any time, users may submit a vote of no confidence against a member of the DAO. If the proposed vote of no confidence receives a certain number of supporting votes, then the vote of no confidence is opened up as a system-wide vote. Once again, there will likely be a minimum quorum of votes that
needs to be reached before any decision can be made. The decision will be verified by the arbitration committee and then executed. This approach to nomination and removal of DAO members is a bit of a hybrid between representative and liquid democracies. The process of nominating and voting to power members of the DAO is inspired by representative democracies but the lack of official terms of power and the ability to vote out members using votes of no confidence at any time is inspired by liquid democracies.

The DAO also has no limit to its size. As long as someone meets the qualifying criteria and receives a sufficient number of votes from the user base, they may become a part of the DAO. The DAO itself would likely consist of a number of committees each of which have a certain number of members determined based on the total number of required committees and the total number of members in the DAO. These committees could also have specialized functions and candidates could be voted into specific types of committees.

For any decision that needs to be made, a random subset of the members of a committee are selected i.e., sortition is used to decide the specific members that will be voting on a given issue. For any given issue, a random subset of committees is chosen as well. For example, if each committee had 20 members and there were 10 such committees, then for a particular issue, sortition would be used to select three random committees and seven random members from each committee to vote. Note that this model will change in behavior slightly if each of the committees are specialized in their function.

The rationale behind doing this is primarily to have redundancies that protect against any kinds of exploits. The randomness introduced into the process should make it hard to “buy” influence at a reasonable cost and the selection of multiple committees is a means to let them operate independently but see if they arrive at the same/similar conclusions. If their conclusions deviate by a certain margin, then the issue is referred to the arbitration committee (Section 6.8.1.5).

In a system like this, oversight becomes vital. Where does the buck stop? Is it with the DAO or with the arbitration committee? Or is it with a third party? It should be clear by now that it is hard, and unwise to pick a single “node” that presides over all. As previously mentioned, the system must not have a single point of failure. Instead, the model presented here is that of “trust, but verify”. The DAO and the arbitration committees both consist of the most reputable users who were chosen by the community for representation. Yet, the trust can never be complete, and it does not appear magically. Instead, each of these groups of individuals are meant to act as a system of checks and balances for the other. This helps create a delicate equilibrium built on a foundation of credible oversight where the “citizens” always feel empowered to take charge (by being able to revoke the powers of anyone in the DAO or replace the arbitration committee at will).

This is merely a suggestion for the potential structure of the DAO, but the ideas presented here are the fundamental ones needed to build out the structure of a successful DAO that can govern a system such as this one. The “ingredients” are here, “cooking” is left to implementation.
6.8.4.2 Contributors’ “DAO”

**Property 9:** The ability to collectively govern a public good once provisioned is pivotal to its long term sustainability. Larger groups prefer elected leaders with more power than smaller groups do. Any voting system being used to elect leaders or make decisions must be strategy-proof. The governance system must instill confidence and the feeling that everyone’s vote matters (the mere belief that it does is sufficient).

In addition to the DAO that is an overarching body that runs the system, the ability to collectively govern a public good once it has been provisioned is pivotal to its long-term sustainability. When the funding round for a proposal closes, all contributors to that proposal end up forming a DAO themselves, though it is never formally referred to as such in this paper. The reason for that is because these informal DAOs formed for every proposal, are just that, they are informal i.e., there are no formal governing mechanisms or structures that are defined for these groups of contributors. To be clear, anytime the DAO is referenced in this paper, it is specifically referring to the organization defined in the prior section, not these informal DAOs for each proposal.

All the contributors for a given proposal are put into a Discord-esque communication environment that is vertically integrated with all the necessary tools to collectively self-govern. This is elaborated upon in Section 6.9 but allowing communities to organically form without structure being enforced has been incredibly successful so that is the approach followed here. Initially, all contributors will simply be in a single communication environment (Section 6.9) during which they will have the chance to get to know each other, and to vote various times during the bidding process (Section 6.5.4 and 6.5.5). Once this process concludes and a bidder has been chosen to provision the good, the following process ensues.

**Property 1:** The tendency to cooperate and not free ride is much greater in smaller groups than in larger groups; and in smaller groups this goes up when those in the group have close-knit interpersonal relationships and/or care about each other. This can be fostered by encouraging communication but also by establishing means to record past tendency to cooperate i.e., institutions and infrastructure can also enable trust.

**Property 9:** Larger groups prefer elected leaders with more power than smaller groups do. Any voting system being used to elect leaders or make decisions must be strategy-proof. The governance system must instill confidence and the feeling that everyone’s vote matters (the mere belief that it does is sufficient).

**Property 10:** Larger group sizes have a positive effect on contributions (frequency and magnitude), and seeing greater contributions only elicits more, but they have a negative effect on the tendency to not free ride and cooperate. At the same time, groups that are too small lack the bandwidth to effectively self-organize and govern.
Research shows that smaller group sizes have a greater tendency to cooperate, and members tend to free ride less than in larger groups, but at the same time, larger groups have a positive effect on contributions. This is why during the fundraising process, group sizes are not capped and can be as large as they naturally will end up being. However, once the fundraising process has concluded and a bidder has been chosen, group sizes should get smaller. Another reason in favor of smaller groups is the fact that they tend to prefer the lack of an elected leader with a lot of power, which fits in perfectly with the model of liquid democracy that this system uses for governance.

Once a bidder has been selected, all contributors will be allowed to self-organize into a set of functions appropriate for a given proposal. This set of functions may be derived from a generalized list of functions for proposals in specific categories provided by the expert committees. Once users nominate themselves to specific functions, a randomized algorithm splits the larger group into smaller groups, all of which are of the same size, but the size of the smaller groups can vary between proposals (within some margin). The sizes of the groups will have to be experimentally determined such that they are sufficiently large that they have the bandwidth to perform necessary tasks but small enough that free riding is discouraged. The randomization of group assignment and group size is simply another deterrent to any kind of collusion (Section 6.11.2).

Once this happens, each of the smaller groups are allowed to operate independently and matters of concern are assigned to these subgroups automatically by an LLM. If the LLM makes a mistake, then a subgroup can refuse to accept an assigned task, which then gets manually assigned to a subgroup by someone from the foundation. Tasks will generally be assigned to multiple subgroups and the outcomes will be cross-referenced against each other as well as against any expert committee recommendations. Any deviations, shall once again, be flagged for audit by the arbitration committee. Lastly, all subgroups will have a means to reconvene all contributors for votes that require input of all contributors involved with a proposal. The goal of this setup is to strike the right balance between privatization and local self-governance. As discussed in Section 4.2.3, this is very important to the success of these efforts.

### 6.8.5 Voting Mechanisms

*Property 9:* The ability to collectively govern a public good once provisioned is pivotal to its long term sustainability. Larger groups prefer elected leaders with more power than smaller groups do. **Any voting system being used to elect leaders or make decisions must be strategy-proof.** The governance system must instill confidence and the feeling that everyone’s vote matters (the mere belief that it does is sufficient).

As discussed extensively in Section 4.4, the voting system used must be strategy-proof i.e., it must uphold independence of irrelevant alternatives (IIA). The reason for this is because if a voting system allows IIA to be violated then it is not strategy proof. A voting system is strategy proof if no voter can get a better outcome in the election by lying about their preferences. There should be no incentive for anyone to lie about their preferences,
if they do have any incentive to lie, then the chances of building a competitive free market that works is slim, and so, IIA must be upheld by the voting mechanism. There are far too many scenarios where votes are cast for them to be listed here again or even summarized, and while many are only votes involving a binary choice, there are many that are not, so it would be wise to build a general voting mechanism that upholds IIA and can be used in any and all voting scenarios that already exist but could also potentially exist in the future. Defining such a mechanism in detail would be beyond the scope of this paper but a few details of interest are mentioned below.

Designing a system that requires a majority vote (consensus may be far too strict for the liquid democracy approach taken here) but also requires a quorum of voters (this quorum should definitely be a majority quorum but could be quite steep as well, say 75%) is advised. These values need not be statically determined at the very start but instead should change based on the kind of decision being voted upon. For example, nomination of a trusted community member who has already met the reputation requirements to be a part of the DAO can be far more relaxed than that for actually getting elected to the DAO. Similarly, refuting the decision of the arbitration committee or accusing it of malpractice is an issue of far greater importance than alerting the DAO to a meme proposal that needs to be manually taken down. So, building a voting mechanism that has a good degree of flexibility built into it such that it can be optimized in real time based on the kinds of decisions being voted upon as well as the decisions that have the greatest impact (both positively but also in terms of potential attacks and security breaches) would be optimal.

6.9 Communication

The above two properties summarize the majority of the literature on the benefits of community formation that were discussed in detail in prior sections. As far as trust is concerned, reputation mechanisms will be discussed in detail in Section 6.10 and conflict resolution methods are discussed in Section 6.5.5. The only remaining point to be addressed is that of communication. There is no doubt that it is crucial, and in some ways, despite being the least technically complex aspect of this system, it may be the most important determinant of success. The majority of DAOs today use Discord for virtually all administrative tasks including communication. [152] It is hard to nail
precisely why Discord has become the tool that everyone in the community turns to but at least exploring the question should yield useful information on what this system should strive for should it want to be the go-to system for forming decentralized communities that are able to work together to solve some of the toughest problems plaguing society.

For starters, Discord took discrete features traditionally tied to console ecosystems like voice chat (that have historically been wildly successful at building very strong communities between pseudonymous identities) and made them sufficiently adaptable to more formal contexts. Discord set out from the very start with a mission to help build communities, the very same mission of the system being proposed in this paper. A very similar platform in its function - Slack, simply went straight to trying to solve the issues with email, especially those caused by the high volume of messages in modern communication, but failed to ever stop and ask if this form of communication was optimal in the first place and so it ended up becoming exactly what it meant to replace, just a shorter form, more amplified (magnitude of messages) version. It is for this reason that Discord is uniquely well suited to help build communities. It is a tool that encourages it in every tiny detail without being overbearing, a place to stop by, hang out, contribute, and leave per one’s convenience, not something that demands constant engagement.

Discord also allows for bots to be built and deployed with ease for a variety of functions, many of which are surprisingly powerful. DAOs need a large number of tools and technologies to assist with administration and to make onboarding and participation as easy as possible for their members. If the group of contributors working on a given proposal could be considered members of a DAO for that proposal (even though, as previously mentioned, they are never referred to as such), one could foresee the notion of a “Vertical Of One” being very useful to them; it is the idea that there should eventually be a system that is an ensemble of discrete, modular, plug and play technologies. Such as: financial plumbing, dynamic IP ownership, content delivery, walled gardens, community gates, governance tools, and smart contracts operating at different levels of abstraction. Interoperable, yet vertically integrated to produce one (nominally seamless) user experience. This aims at addressing the concern of transaction costs impeding communities without sufficient resources from organizing around issues of importance to them. Discord is the closest thing to this today because it nails the element at the heart of all this - communication, while having bots that can be used to satisfy many of the other tools needed for DAOs to operate successfully.

A part of the goal of the system being described in this paper is to build such a vertically integrated system that possesses the majority of these tools specifically as it applies to solving collective action problems. The model
could then be expanded to be a truly versatile globally applicable suite of tools for decentralized community formation. For example, dynamic IP ownership is a problem not applicable to the system being built in this paper while walled gardens and community gates are particularly relevant (as will be discussed shortly). Everything starts with a communication platform that possesses the majority of the same characteristics as Discord. However, naturally, the question arises, why not use Discord itself? The answer is quite simple, and it has to do with preserving pseudonymity (every user is represented by a pseudonymous ID and their reputation score in the communication system described here) and preventing communication outside of the specific situations in which it is desired, to prevent unhealthy collusion (which is often the primary reason for the unraveling of many DAOs). The consequences of collusion are discussed in detail in Section 6.11.2. In addition to collusion, there is also the risk of too much communication resulting in an awareness of the willingness of others in the group to cooperate which could end up encouraging free riding. However, it is not of major concern since this kind of communication is only possible after money has been raised and is locked in, so, at worst, such free riders could derail the proposal from succeeding purely through lack of participation.

To clarify, this is most definitely an ambitious task, the details provided here are also not nearly enough to architect and engineer such a system, but providing those details is well beyond the scope of this paper. The primary goal of this section is to mention the requirements that need to be satisfied for the communication mechanism for the system as well as to provide the rationale for why these requirements need to be met. A helpful guideline would be to remember that the goal of the communication system is to enable groups of individuals working together to be autopoietic, alegal, hyper scalable, executable, permissionless, aligned, co-owned, and mnemonic. (Section 6.8.4.1.2)

### 6.10 Reputation

**Property 1:** The tendency to cooperate and not free ride is much greater in smaller groups than in larger groups; and in smaller groups this goes up when those in the group have close-knit interpersonal relationships and/or care about each other. This can be fostered by encouraging communication but also by **establishing means to record past tendency to cooperate** i.e., institutions and infrastructure can also enable trust.

This point has been harped on extensively, but a reputation mechanism is integral to this system functioning as intended. Reputation mechanisms help build trust. In a blockchain-based system where everything is permanently recorded and tied to a user’s “identity” (publicly this would be a pseudonym but under the hood it is tied to a user’s government identity i.e., there is no way to erase one’s record - for both users and organizations), the trust in a specific user’s (or organization’s) reputation record will be extremely high. The best way to illustrate the many nuances required when building a reputation mechanism is by example. This is one of the few instances where implementation specifics will be provided in detail, but this is not prescriptive i.e., the details are provided simply for illustrative purposes and are not reflective of an ideal reputation mechanism by any means, but rather encode the properties that make a reputation mechanism ideal.
6.10.1 Users

Every user is given a reputation score. The score is made up of two components. The first component is called “Money” and the second component is called “Activity”. There is a multiplier that gets applied to a person’s reputation score based on “Time”.

6.10.1.1 Money and Activity

The “Money” component. All users receive one point per cent that they contribute and if they are a proxy, receive 0.1 points per cent contributed by users who have delegated their votes. The points accumulated via this component are worth 40% of the total reputation score for a user. The points are issued right after a funding round closes but before any results of the funding round are revealed. The points are issued regardless of whether the proposal succeeds or fails.

The “Activity” component. All users receive 1000 points every time they vote, 100 points for every meaningful (as determined by the bidder) response to a query from a bidder (Section 6.5.4), 10 points per delegated vote, and 10 points per meaningful message sent on the communication platform (as measured by an LLM, with some daily cap to prevent misuse and abuse). The points accumulated via this component are worth 60% of the total reputation score for a user.

The actual point values chosen are arbitrary but the ratios between points awarded for different kinds of activity as well as the specific kinds of activity that are awarded points, are both very important. Implementation should strive to keep these ratios in a similar ballpark and to make sure that all these activities are awarded points. The idea is to provide the highest rewards for personal engagement but if one is making the effort to build a reputation such that they can have votes delegated to them then they must be rewarded for that as well. The 60-40 split between points for activity and money is primarily to not discourage those who do not have massive coffers to commit to every single proposal there is. The point is that even contributing small amounts but engaging with the system actively and using one’s built reputation to become a proxy, will suffice to help build a high reputation and become a trusted member of the community.

6.10.1.2 Local and Global Time

The “Time” component. There are two multipliers applied via this component.

The first is a “local” multiplier that is applied on a weekly basis to all points accrued in a given week. It is modeled by the function $y = 10 \ln(x)$ as shown below:
The y-axis represents the value of the multiplier, and the x-axis represents the number of continuous weeks of engagement with the system. This could be measured by something as simple as using the system for 10 minutes a week or something more complicated involving true measures of engagement via some AI, details are left to implementation. The local multiplier only activates from week five and is capped at week 52 so the highest local multiplier that can ever be active on an account is 39.512%. The local multiplier is applied as a % of points gathered over each week i.e., if a user gains 200 points in week five, the local multiplier for week five which is 16.094% will be applied to the 200 points gathered over the week to give the user a net total of 232.188 points for the week.

The counter recording the number of continuous weeks of activity is reset to zero if a user has been inactive for four weeks (as measured by whatever metric for activity is chosen). For example, if a user has not logged into the system and spent at least 10 minutes on the system at least once a week for four weeks then their counter would be reset to zero. The counter recording the number of continuous weeks of activity is decreased by one for every week of inactivity for the first four weeks of inactivity after which it is reset to zero. As previously mentioned, the local multiplier will only activate at week five i.e., if a user has had five weeks of continuous activity then the local multiplier will activate at the end of week five and be applied to all points gathered in week five but if they are inactive on week six then the counter will go back down to four so they will only receive a week five local multiplier even if they are active on week seven.

The second is a “global” multiplier that is applied on a per proposal basis at the conclusion of a given proposal. It is modeled by the function $y = 2 \ln(x)$ as shown below:
Once again, the y-axis represents the value of the multiplier, and the x-axis represents the number of continuous weeks of engagement with the system. The global multiplier also only activates from week five. The global multiplier is applied as a percentage of all points gathered thus far. The global multiplier has two components. Both components use the exact same function, the only difference being that the positive component adds points to the reputation score while the negative component subtracts points from the score.

The positive component is applied at the end of a proposal a user is involved with since points are anyway only accumulated over the course of involvement with a proposal. This multiplier is applied to all points gathered thus far, i.e., if a user’s reputation score is 341,617 points after a proposal concludes (i.e., the proposal is either a success and proceeds or is not a success and terminates - regardless, for the money contributed as well as votes cast, points would have been awarded on a weekly basis along with local multipliers applied as appropriate) and a user has been active for 64 weeks then the global multiplier is 8.318% so their score will immediately receive a boost to 370,032.702 points.

The negative component however is applied once a week if inactive for more than four weeks. The counter recording the number of continuous weeks of activity is reset to -5 if a user has been inactive for four weeks i.e., for example, if a user has not logged into the system and spent at least 10 minutes on the system at least once a week for four weeks (or whatever activity metric is chosen). It is also applied to all points gathered thus far, i.e., if a user’s reputation score is currently 1,275,896 and they have been inactive for four weeks then their counter recording the number of continuous weeks of activity is set to -5 and at the end of week five, a negative global multiplier of 3.219% is applied to their total score resulting in their score dropping to 1,234,824.91. If they now
become active in week six, then the counter moves to -4 and so on i.e., if a user is inactive for four weeks, it takes a minimum of four weeks to stop being punished for that inactivity and then takes another four weeks to start being rewarded for activity again.

**Property 2:** Reputation mechanisms that have **deterrents (i.e., some punishment mechanism)** with simple heuristics that easily indicate to others the degree of trustworthiness/cooperativeness of those they are working with helps increase the likelihood of cooperation in larger groups. This is especially true when individuals are pseudonymous and may not have previously interacted with each other.

These multipliers were heavily inspired by hotel and airline loyalty rewards programs. The idea behind the local multiplier is to encourage continuous engagement while being involved with a proposal by providing high weekly bonuses that are issued to points accumulated in real time. The analog to this in rewards programs would be point multipliers or bonus points issued for every stay during peak travel times in specific regions. For example, Marriott ran a promotion that awards two elite night credits for every one night in a Marriott property during the summer in the U.S. and may do the same during Cherry Blossom season in Japan. [146] A slightly different incentive mechanism was offered by Hyatt this Spring where such bonuses were only awarded for stays of two nights or more thereby encouraging longer stays at the same property and decreasing costs of checking guests in and out i.e., being required to deep clean rooms. [147] The incentives to be encouraged in this system are that of continuously being engaged with a proposal from when someone first contributes to it all the way through to seeing the proposal succeed. The high weekly bonuses (which could even be limited on a per proposal basis i.e., the weekly engagement criteria would need to be seen on each proposal for the bonus to be awarded for that specific proposal) does exactly this. This is also why this multiplier has a 52-week cap, because despite not being stated explicitly, no proposal must take longer than 52 weeks from beginning through termination, and if it does, it likely means that anything happening post-week 52 should only concern the bidder provisioning what they promised to and not really require much involvement from contributors at all.

The idea behind the global multiplier on the other hand is to provide an incentive for continuous engagement for those with already high reputation scores since the multiplier values are low but are applied to all points accumulated thus far. The idea behind the global multiplier is to mimic something similar to airline and hotel loyalty programs wherein status is computed based on miles flown or nights occupied each year. In fact, these programs are even stricter than the one proposed here because they reset status at the start of every year regardless of commitment in the prior year; however, due to the gradual nature of the negative global multiplier in this
system, it will take a decent amount of time for a few years of committed involvement (scores in the tens of millions) to be completely wiped out. Delta adopted something similar to this with rollover mileage requirements for status during the peak of COVID as their frequent flyers could not fly as much as they normally would have. [148] The multipliers are not a large incentive for those who have just started using the system, because they have nothing to lose (and/or gain), but for those who have been engaged for a while, the stakes are higher. The idea is to punish lack of engagement but not so strictly that it turns people away from using the system altogether, because after all, they are volunteering their time, money, and effort.

Once again, the actual functions chosen for the multipliers were fairly arbitrary but the manner in which they scale as well as the ratio of scaling between the local and global multipliers should remain in the same ballpark in implementation. In addition, the balance between punishment for deterrence, and forgiveness, is an art and not a science. It is possible that the suggested negative multiplier here is too weak, it is also possible that it is too strong. This will have to be determined and adjusted in real time with some experimentation, so keeping the 10 desired properties to be satisfied as the guiding principles will yield the best results. In addition to this form of “punishment”, any kind of collusion, fraud, or stalling will be punished far more severely, and the specifics of those mechanisms are outlined in Section 6.10.1.5.

6.10.1.3 Example Scores

Consider the following moderately enthusiastic user who has just joined the system. They find one proposal attractive and contribute $150 to that proposal. They are consistently engaged with the proposal over the course of 12 weeks and vote 10 times during this period. Their reputation score may be computed as follows:

<table>
<thead>
<tr>
<th>Reason</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Welcome Bonus (Optional Incentive)</td>
<td>10000</td>
</tr>
<tr>
<td>“Money” Component</td>
<td></td>
</tr>
<tr>
<td>- Initial Funding Round Closes, Awarded Points for $150 Contribution = 15000</td>
<td></td>
</tr>
<tr>
<td>- Local Multiplier Bonus (Week 6) Is Applied to Weekly Earnings</td>
<td></td>
</tr>
<tr>
<td>“Activity” Component</td>
<td></td>
</tr>
<tr>
<td>- 50 Meaningful Messages Per Week from Weeks 6 through 11 = 500/Week</td>
<td></td>
</tr>
<tr>
<td>- Local Multiplier Bonus (Weeks 6-11) Is Applied to Weekly Earnings</td>
<td></td>
</tr>
</tbody>
</table>

\[
\begin{align*}
\text{Welcome Bonus (Optional Incentive)} &= 10000 \\
\text{“Money” Component} &= 15000 \times 0.40 \times 1.17198 \\
\text{“Activity” Component} &= ((500 \times 1.17198) + (500 \times 1.19459) + (500 \times 1.20794) + (500 \times 1.21972) + (500 \times 1.23026) + (500 \times 1.23979)) \times 0.60 \\
&= 2179.176
\end{align*}
\]
“Activity” Component
- Voting Period Closes, Awarded Points for 10 Votes Cast = 10000
- 5 Meaningful Responses to Bidder In Week 8 = 500
- Local Multiplier Bonus (Week 8) Is Applied to Weekly Earnings

\[10500 \times 0.60 \times 1.20794 = 7610.022\]

Proposal Is a Success/Failure
- Global Multiplier Bonus (Week 12) Is Applied to All Points Gathered Thus Far

\[(10000 + 7031.88 + 2179.176 + 7610.022) \times 0.0497 = 1333.00758\]

TOTAL SCORE ~28,154

It is clear that the majority of points are awarded from monetary contributions and voting, with voting outweighing the points received from monetary contributions. These categories then receive significant boosts from consistent engagement via the local and global multipliers. As intended, a user who has just started using the system receives far higher point boosts from the local multiplier than from the global multiplier. The points received from engaging with the community via meaningful messages is significantly lower than other categories, and this is intentional. This is to both discourage any attempts to misuse the system but also as a preemptive measure against potential misuse until means to distinguish “meaningful” messages from spam are determined to be accurate. Furthermore, the system does not require users to dedicate multiple hours each week to collaborating on solving collective action problems, a mere few minutes each day will likely suffice, and so the potential to earn points from such engagement is both capped on a daily basis but also not as high as other forms of more meaningful engagement.

Now consider a far more experienced user who already has a score of 15,000,000 and has crossed required thresholds to become a proxy. They are, obviously, fairly actively involved and it would be reasonable to assume they have 500 votes delegated to them.

```
<table>
<thead>
<tr>
<th>Reason</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Money” Component</td>
<td></td>
</tr>
<tr>
<td>- Initial Funding Round Closes, Awarded Points for $250 Contribution = 25000</td>
<td>75000 \times 0.40 \times 1.17198 = 35159.4</td>
</tr>
<tr>
<td>- Awarded Points for $5000 Delegated Contributions = 50000</td>
<td></td>
</tr>
</tbody>
</table>
```
<table>
<thead>
<tr>
<th>Component</th>
<th>Calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Local Multiplier Bonus (Week 6)</strong></td>
<td>Applied to Weekly Earnings</td>
</tr>
<tr>
<td><strong>Activity Component</strong></td>
<td><strong>Weekly Earnings</strong></td>
</tr>
<tr>
<td>- 500 Meaningful Messages Per Week from Weeks 6 through 11 = 5000/Week</td>
<td>$$((5000 \times 1.17198) + (5000 \times 1.19459) + (5000 \times 1.20794) + (5000 \times 1.21972) + (5000 \times 1.23026) + (5000 \times 1.23979)) \times 0.60$$ (Activity Has 60% Weightage) = 21791.76</td>
</tr>
<tr>
<td>- Local Multiplier Bonus (Weeks 6-11) Is Applied to Weekly Earnings</td>
<td></td>
</tr>
<tr>
<td><strong>Activity Component</strong></td>
<td><strong>Weekly Earnings</strong></td>
</tr>
<tr>
<td>- Voting Period Closes, Awarded Points for 10 Votes Cast = 10000</td>
<td>62000 \times 0.60 (Activity Has 60% Weightage) \times 1.20794 (Local Multiplier) = 44935.368</td>
</tr>
<tr>
<td>- Awarded Points for 500 Delegated Voters (10 Votes/Voter) = 50000</td>
<td></td>
</tr>
<tr>
<td>- 20 Meaningful Responses to Bidder in Week 8 = 2000</td>
<td></td>
</tr>
<tr>
<td>- Local Multiplier Bonus (Week 8) Is Applied to Weekly Earnings</td>
<td></td>
</tr>
<tr>
<td><strong>Money Component</strong></td>
<td><strong>Weekly Earnings</strong></td>
</tr>
<tr>
<td>- Second Funding Round Closes, Awarded Points for $200 Contribution = 20000</td>
<td>30000 \times 0.40 (Money Has 40% Weightage) \times 1.23026 (Local Multiplier) = 14763.12</td>
</tr>
<tr>
<td>- Awarded Points for $1000 Delegated Contributions = 10000</td>
<td></td>
</tr>
<tr>
<td>- Local Multiplier Bonus (Week 10) Is Applied to Weekly Earnings</td>
<td></td>
</tr>
<tr>
<td><strong>Activity Component</strong></td>
<td><strong>Weekly Earnings</strong></td>
</tr>
<tr>
<td>- Voting Period Closes, Awarded Points for 5 Votes Cast = 5000</td>
<td>30000 \times 0.60 (Activity Has 60% Weightage) \times 1.23979 (Local Multiplier) = 22316.22</td>
</tr>
<tr>
<td>- Awarded Points for 500 Delegated Voters (5 Votes/Voter) = 25000</td>
<td></td>
</tr>
<tr>
<td>- Local Multiplier Bonus (Week 11) Is Applied to Weekly Earnings</td>
<td></td>
</tr>
<tr>
<td>Proposal Is a Success/Failure</td>
<td><strong>Global Multiplier Bonus (Week 12) Is Applied to All Points Gathered Thus Far</strong></td>
</tr>
<tr>
<td>- Global Multiplier Bonus (Week 12) Is Applied to All Points Gathered Thus Far</td>
<td>$$(15000000 + 35159.4 + 21791.76 + 44935.368 + 14763.12 + 22316.22) \times 0.0497$$ (Global Multiplier) = 752,406,604</td>
</tr>
<tr>
<td><strong>TOTAL SCORE</strong></td>
<td>~15,891,371.9</td>
</tr>
</tbody>
</table>
Once again, it is clear that the majority of points are awarded from monetary contributions and voting, with voting outweighing the points received from monetary contributions. These categories then receive significant boosts from consistent engagement via the local and global multipliers. As intended, the points received as a result of others delegating their votes to the user (in their capacity as a proxy) forms the largest chunk of points earned. The user also earns a significant boost in points as a result of their global multiplier being applied to all points earned thus far, again, as intended. The points received from engaging with the community via meaningful messages continues to be significantly lower than other categories despite a 10x increase in engagement compared to the prior example, and this is also intentional for the same reasons as earlier.

These examples are not comprehensive in the least. There are several aspects of the scoring mechanism that were described but not used when computing these sample scores. This was purely for illustrative purposes. Part of why the reputation scoring system is elucidated in such detail is to highlight the importance of making these scoring systems sufficiently explicit that people are willing to work toward improving their scores and it feels aspirational but at the same time, must not crowd out their willingness to cooperate [53][54], as mentioned in Section 4.2.3. The specifics chosen here are mere guesses at balancing this appropriately, but this will have to be determined experimentally and adjusted as needed in real time. So, final details are left to implementation.

6.10.1.4 Simple Heuristics

Property 2: Reputation mechanisms that have deterrents (i.e., some punishment mechanism) with simple heuristics that easily indicate to others the degree of trustworthiness/cooperativeness of those they are working with helps increase the likelihood of cooperation in larger groups. This is especially true when individuals are pseudonymous and may not have previously interacted with each other.

This scoring system may feel somewhat convoluted, and that is because it must be, in order to encompass all the nuances of the way the system works. However, providing reputation scores as arbitrary numbers may not be the most helpful way to indicate to others the degree of trustworthiness of a given user. A simple heuristic is necessary. One of the simplest ways of providing this heuristic is to create a tiered leveling system where users are bucketed into certain levels/tiers based on their reputation score. In fact, the numerical reputation score could even be completely hidden from users. Based on the sample scores for users demonstrated above, one could develop the following tiered system. Once such a system is established, each tier and level must be given a name that is consistent with the theme and “look and feel” of the system, so details are left to implementation.
<table>
<thead>
<tr>
<th>Points To Reach Each Tier</th>
<th>Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier 1</td>
<td></td>
</tr>
<tr>
<td>- Level 1.a - 10,000</td>
<td></td>
</tr>
<tr>
<td>- Level 1.b - 25,000&lt;sup&gt;1&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>- Level 1.c - 50,000&lt;sup&gt;2&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Tier 2</td>
<td></td>
</tr>
<tr>
<td>- Level 2.a - 70,000 (+20,000)&lt;sup&gt;3&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>- Level 2.b - 175,000</td>
<td></td>
</tr>
<tr>
<td>- Level 2.c - 350,000</td>
<td></td>
</tr>
<tr>
<td>Tier 3</td>
<td></td>
</tr>
<tr>
<td>- Level 3.a - 400,000 (+50,000)</td>
<td></td>
</tr>
<tr>
<td>- Level 3.b - 1,000,000</td>
<td></td>
</tr>
<tr>
<td>- Level 3.c - 2,000,000</td>
<td></td>
</tr>
<tr>
<td>Tier 4</td>
<td></td>
</tr>
<tr>
<td>- Level 4.a - 2,500,000 (+500,000)</td>
<td></td>
</tr>
<tr>
<td>- Level 4.b - 6,250,000</td>
<td></td>
</tr>
<tr>
<td>- Level 4.c - 12,500,000</td>
<td></td>
</tr>
<tr>
<td>Tier 5</td>
<td></td>
</tr>
<tr>
<td>- Level 5.a - 15,000,000 (+2,500,000)</td>
<td></td>
</tr>
<tr>
<td>- Level 5.b - 37,500,000</td>
<td></td>
</tr>
<tr>
<td>- Level 5.c - 75,000,000</td>
<td></td>
</tr>
<tr>
<td>Tier 6</td>
<td></td>
</tr>
<tr>
<td>- Level 6.a - 80,000,000 (+5,000,000)</td>
<td></td>
</tr>
<tr>
<td>- Level 6.b - 200,000,000</td>
<td></td>
</tr>
<tr>
<td>- Level 6.c - 400,000,000</td>
<td></td>
</tr>
<tr>
<td>Tier 7</td>
<td></td>
</tr>
<tr>
<td>- Level 7.a - 450,000,000 (+50,000,000)</td>
<td></td>
</tr>
<tr>
<td>- Level 7.b - 1,125,000,000</td>
<td></td>
</tr>
<tr>
<td>- Level 7.c - 2,250,000,000</td>
<td></td>
</tr>
<tr>
<td>Tier 8</td>
<td></td>
</tr>
<tr>
<td>- Level 8.a - 2,500,000,000 (+250,000,000)</td>
<td></td>
</tr>
<tr>
<td>- Level 8.b - 6,250,000,000</td>
<td></td>
</tr>
<tr>
<td>- Level 8.c - 12,500,000,000</td>
<td></td>
</tr>
<tr>
<td>Tier 9 - 50,000,000,000,000</td>
<td></td>
</tr>
</tbody>
</table>

<sup>1</sup> The sub-b level is 2.5x the sub-a level threshold.

<sup>2</sup> The sub-c level is 2x the sub-b level threshold.

<sup>3</sup> The sub-a level of the following tier is higher than the prior tier’s sub-c level and this value is in parenthesis next to each sub-a level.
This tiered leveling system also aids with the need to create limited governance (tiers that allow users to become proxies, to nominate themselves to be a part of the DAO etc.) as discussed in Section 6.8.3.1. This is definitely not the only way to create an easily understandable heuristic for the reputation mechanism, so details are left to implementation, this was purely to illustrate what this could look like in practice. It is important to remember that whatever system is deployed, it must also function as a means to determine roles and scope of impact in limited governance schemes.

6.10.1.5 Fraud, Collusion and Stalling

Property 2: Reputation mechanisms that have deterrents (i.e., some punishment mechanism) with simple heuristics that easily indicate to others the degree of trustworthiness/cooperativeness of those they are working with helps increase the likelihood of cooperation in larger groups. This is especially true when individuals are pseudonymous and may not have previously interacted with each other.

Property 7: Reputation systems encourage cooperation and altruism. These systems must allow for the punishment of defectors; but have to strike a balance between doing so in a manner sufficiently strict that it deters defectors yet has some room for forgiveness. These systems cost money to run but people are willing to pay for them (altruistic punishment).

In addition to having a negative multiplier applied for inactivity, there are several situations in which users, or more precisely, sufficiently large groups of users, may cause havoc on and strain the system’s resources. These activities need to be punished in a manner far more severe than something as tame as a negative multiplier that chips away at a user’s score. The penalties must essentially be unreasonably high, something that is so high that would otherwise be absurd but is fine because no honest individual would ever be hit by such a penalty and so they should not be bothered by it. Furthermore, there must be some kind of permanent record that flags a user as a past offender for a certain period of time (even if their account were to be reinstated after a probationary period and/or a massive hit to their reputation score was the only penalty they received), likely similar to collections and delinquencies on credit reports. These various mechanisms will, at least initially, be run entirely by the foundation, and can later be modified by the DAO.

At various places over the course of the paper so far, the points at which, proposals may be flagged for audit, users are able to vote, and situations in which users can report things such as potentially suspicious activity etc. have been referenced. However, all of these public moderation mechanisms can, and will, be exploited; many of these scenarios are described in Section 6.11. If someone wants to stall a proposal for any number of reasons, they can do so. For all such violations, a first offense must render a substantial hit to a user’s reputation score, say 50%. The second offense will garner a 100% hit to their reputation i.e., their reputation score will now be zero and a temporary suspension (at least a year) will be levied on their account, however, there will still be no public record.
of such offenses. This is not a static rule, it is simply an arbitrary choice of number of forgiven offenses. As previously mentioned, the reputation mechanism must balance forgiveness with punishment appropriately.

Once again, arbitrarily, all subsequent offenses will render a permanent notation on the user’s reputation record that describes the degree of their “crimes” in detail. All of these “crimes” are generally considered to be a part of the same “class” i.e., they are all equally bad. Hence, all offenses beyond the third one will receive longer and longer suspensions alongside permanent notations on their reputation records, until eventually being banned for life. A given user should not be “convicted” of more than 10 fraud, collusion, and stalling offenses before being banned for life. To reiterate, these numbers are arbitrary but should likely remain in the same ballpark (i.e., not 10 convictions before a lifetime ban, but 20 could be reasonable). Additionally, the possibility of forgiven convictions appearing on permanent records may also be considered since GTFT (Section 4.2.2) only works when people are aware that it is being implemented and are able to see evidence of it.

In essence, the idea is that any kind of user activity which can be considered a security breach i.e., users acting disingenuously and against the system, must be punished severely while allowing for some room for forgiveness and reconsideration. The details are largely left to implementation because again, this is a balancing act, an art, not a science, but this form of “punishment” must be comprehensive, no stone should be left unturned as far as the list of potential “crimes” that fall under this form of “punishment” are concerned. A comprehensive list is not provided because there are simply so many such situations that have been described throughout the course of this paper, and there are likely more that were not covered, or rather unique situations that will arise as a result of specific implementation details - and these must all be handled in the same manner.

6.10.1.6 Certification

As previously mentioned, once a proposal is successfully provisioned or terminated unsuccessfully, users receive commemorative tokens for their contribution and participation. Prior to this, every user has been issued a number of tokens (implementing ERC-20, ERC-777 or ERC-1155) equal to the amount they donated in USD (Section 6.5.2). Once a proposal concludes, these tokens have no utility. Users can exchange these tokens for NFT’s that act as certificates commemorating their participation and contribution that are displayed on their profile. Users will receive a single NFT (implementing ERC-721 or ERC-1155) in exchange for their tokens.

Users can choose if they want the NFT to indicate any identification details should they want this as a “perfect” proof of contribution and participation, but they can also choose to not have any identification on it if they intend on selling it (some offset “revenue” may be possible for contributors via such sales). It is important to note
that the NFT is proof of both contribution and participation, both elements are weighted equally and details in the NFT will reflect this. The NFTs may be created by collaborating with famous artists (in the case of special proposals, say those with lasting global impact) but could end up being far simpler too and really just serve as proof of contribution and participation.

This process of issuing the NFTs in addition to a user’s contribution and participation already being factored into their reputation score is an attempt to gamify the process in some way. It also plays into the “hype” associated with NFTs/crypto/web3, and the recent rise in public virtue signaling on social media, but it is also not unrealistic to envision a future where cryptographically verifiable digital credentials are highly valued (in all facets of life). Regardless, this is largely only a novelty feature, but could enhance the user experience in a meaningful way that contributes to greater coordination and participation.

### 6.10.1.7 Special Users

The reputation mechanism as described thus far, has only specified expected behavior for the average user. It has not broached how these mechanisms will change for users with some kind of special status, be that those who are proxies, or part of the DAO, or part of the arbitration committee etc. This will not be addressed in detail, but some important considerations are as follows. First, when it comes to the basic components of the reputation score, namely “Money” and “Activity” as well as the local and global multipliers from “Time”, there must be some kind of mapping between time spent on the system and the points that can be earned from these components. That can be used as a guiding metric to adjust for the points that users in these special positions receive as they will likely no longer be able to contribute the same amount of time to proposals on a daily basis due to these additional responsibilities. Second, as far as punishments are concerned, the stakes are upped. If found guilty of a “crime” then for their first offense, these users are stripped of their position and their reputation score is completely wiped. This is also permanently notated on their reputation record. They are also permanently banned from ever holding such special positions again. Any subsequent offenses will result in a lifetime ban. This is far harsher than the punishment scheme for regular users but so is their betrayal. Lastly, in addition to the NFTs being awarded to everyone that was involved, there will also be a special set of NFTs issued to those who were pivotal to the success of a given proposal. This includes those with particularly high engagement and participation, the user who nominated the proposal in the first place, those who alerted the foundation to possible security breaches, members of the DAO or arbitration committees that resolved issues etc. The exact set of “awards” are up to implementation depending on the facilities provided for users to be engaged in such pivotal roles. These NFTs will also be issues in a manner similar to that described in the prior section.

### 6.10.2 Bidders

The reputation scores for bidders can be far simpler than that for users. There are a few metrics on which bidders must be evaluated. These include responsiveness when communicating with users, pricing and quality of work compared to competitors, ability to deliver work on par with prior work samples, ability to deliver work per specified contract, and ability to deliver work on time. Bidders may be evaluated on these metrics by contributors
to a proposal once the bidder has completed their work. All contributors will be able to rate bidders on a scale of 1 to 10 and the scores that bidders receive will be aggregated across all contributors. Bidders will have the opportunity to contest the scores they receive should they not agree with them. Such contestations will be evaluated by the arbitration committee, and if the bidders have valid evidence to refute their scores, those users who deviated from a reasonable score for the bidder will be penalized; and remaining contributors will be asked to vote to rescoring the bidder.

The work of a bidder on a proposal does not cease once the good has been provisioned. If there are maintenance concerns or anything else for which contributors need to contact the bidder in the future, then upon failure to respond promptly and address concerns, contributors can vote to re-evaluate the bidder. If the vote passes then the bidder will be re-evaluated by the same process as described earlier, and once again, a bidder can contest the re-evaluation. It may even make sense to have some kind of permanent notation on a bidder’s reputation score should such situations arise due to the importance of self-governing communities being able to guarantee the sustainability of public goods they helped provision.

Bidders will also be permanently banned from the system should they not deliver on agreed upon contracts, though if that does happen, they will likely face far more severe consequences in court (since legal agreements for agreed upon deliverables will be drawn up prior to commencing any work) so receiving a lifetime ban from the system should be the least of their concerns. Despite the juxtaposition of the complexities in reputation scoring systems for users and bidders, rest assured that this “simple” publicly sourced reputation system for bidders shall suffice. In fact, given the importance of this information in informing contributors when they are voting on which bidder a proposal gets assigned to, this information should be as simple, straightforward, and succinct as possible to not confuse contributors at all.

6.11 Attacks

The class of attacks discussed here are those that may be defined as “in-protocol” i.e., they cannot be solved through cryptography, but rather, only by means of careful mechanism design. The basic framework that can be used to address most attacks can be captured as follows:

\[
\text{Profit To Attacker} = \text{Payoff From Attack} - \text{Cost Of Acquiring Voting Power} - \text{Cost Of Executing Attack}
\]

Not all attacks will require acquiring voting power nor will all have any costs outside of just doing that, but the general principles involve decreasing the payoff from attacks, increasing the costs of acquiring voting power, and increasing the costs of executing an attack. It should be clear by now that there has been a conscious effort to design every part of the system to do exactly this. The payoffs from attacks are limited by the fact that most governance decisions only have limited latitude, beyond which, they need to be reviewed by an arbitration committee; and the decisions of the committee can be repealed for a certain period of time once they have been
made. The cost of acquiring voting power has been made as high as possible by replacing token-based governance with reputation-based governance where reputation is incredibly hard to acquire (i.e., there is a large time component to it). Lastly, the cost of executing attacks has also been made very high by requiring every user to be a verified unique human being and allowing the entire community to act as a policing force to help flag any and all suspicious behavior. Despite these design decisions, and in some cases, as a result of them, there are a few remaining challenges worth addressing in greater detail.

6.11.1 Fraud

When considering means by which practical implementations of the system could be fraudulently exploited, the most prevalent problem would be that of attempting to exploit the system for tax benefits and/or money laundering. Since there is a tax benefit to contributing money to these proposals, one could foresee a company/individual that wants to avoid paying taxes (or just wants to launder money, or both) that uses the system for these purposes. They could achieve this by creating dubious and obscure proposals that appear legitimate but the projects themselves are sufficiently unknown that they do not attract significant public attention, then they would create numerous accounts through which they contribute money to the proposal. Since they were the primary contributors, they presumably trust that they will hold a majority of the voting power. They then proceed to place bids (from a shell company) approximately equivalent to the money they contributed, and when votes are cast on bidders, since they hold a majority, they can make sure their (shell) company is selected as the provisioner for that proposal. The company/individual has now effectively moved a certain amount of money through the system and that money is now considered tax-deductible in the U.S. (and eventually other countries where donations to nonprofits would be considered tax-deductible) but the company that made the bid is just a shell company incorporated in Monaco, The Cayman Islands etc. (some tax haven) and so when the attackers (through the shell company) receive the money via the system they will pay no taxes on it. They have effectively laundered taxable money to a tax haven in a “legal” manner. They will also likely have “stolen” some amount of money from innocent contributors to the proposal. This can only happen once, after which all the accounts involved in this fraud effort would be banned for life, however, this is still a fairly large problem because regulators will sue once this is discovered as will innocent contributors who were cheated out of their money.

There are a few crucial elements that attackers would need working in their favor for this to be successful. It would be hard for the foundation to identify the aforementioned “dubious and obscure proposals” and eliminate them from the system. However, there are a few possible solutions to attempt to, at the very least, restrict the chances of such proposals having much success. One solution to this would be to place some kind of threshold as to required local participation for a given proposal. For example, users can only nominate proposals in their own country (or maybe even state), and a given proposal must have a certain percentage of contributors be local to the region (however that is defined). This threshold could even be so high as to guarantee that locals have majority voting power. However, all users who are able to vote and donate to a specific proposal should not be limited to specific geographic regions for the many reasons discussed in Section 6.7 that elucidate the philanthropic benefits
of globalizing the system. In addition to this, users (with sufficiently high reputation scores) could be allowed to flag proposals they believe are not genuine. These proposals will be halted and reviewed by the arbitration committee. To prevent misuse of this system, repeated false positive reports will result in penalties on a user’s reputation score, and the fact that only users with sufficiently high reputation scores may make these reports should deter most abuse.

However, these are merely deterrents and could be overcome as long as attackers are able to create numerous “fake” accounts that bypass any rules (including geographic specificity). Since KYC/AML information is required to sign up, the system would need to have fairly lax identity requirements or be operational in countries where identity is hard if not impossible to verify accurately, for attackers to be able to create such accounts. In fact, even in the U.S., more than 20 million people have no ID at all [133] and nearly 60% have no federal ID [134]. With Real ID (federal ID) becoming a requirement for U.S. domestic air travel by May 2025, one would hope that the number of U.S. adults with federal ID will go up, however this date has been pushed back from October 2021 to May 2023 and now out to May 2025. [135] Regardless, the point is that federal ID is the only form of ID that can easily be electronically verified, and it is a form of ID that is not held by the majority of the population in a first world country like the U.S. This can be contrasted with India which lacked any uniform ID scheme in 2009 at which point it created the Unique Identification Authority of India (UIDAI). UIDAI created the “Aadhaar” identification system that is now used by the majority of the country and can be electronically verified both online and offline. [136] So, the policies in place to prevent Sybil attacks will need to vary from region to region and will primarily rely on secure identification methods that can confirm that individuals attempting to use the system are in fact unique human beings. Despite this proof of humanness requirement, the system is still susceptible to attacks from stolen identities but the assumption that such attacks are fairly low in number is rational given that it would be very hard and expensive for attackers to garner sufficiently high numbers of stolen identities in order to execute such an attack.

The prior steps should prevent the vast majority of these forms of attacks from proliferating, but assuming that it is unable to, the final steps to performing this attack successfully involve using shell companies (or real ones) incorporated in tax havens. A good first step to stopping this would be to have some form of verification for private companies before they are able to place bids on proposals. This would be a manual review/verification process with some overhead cost funded by contributors that aims to prevent shell companies from placing bids. The reason this was not proposed in Section 6.5.4 is because doing so, especially when the system is first launched could prove to be a burden to private companies slowing their onboarding process and deterring them from using the system. This kind of manual review process would likely involve submitting proofs of incorporation, tax fillings, records of past work etc. which is something a company may not want to do if they are being asked to join the system as a “favor”. However, even if this were done, it would not be a comprehensive verification mechanism, because one could have the right documents and even be running a legitimate company except it is just one that is incorporated in a tax haven. There is nothing illegal about running companies incorporated in tax havens, so details are left to implementation as to how this should be handled, but due to the nature of the system and its mission, it would be advised to prevent companies registered in tax havens from joining the system and/or
only allow companies registered in a given country to place bids for proposals in that country - to keep things local and support (legitimate) small businesses.

Furthermore, the reputation mechanism for bidders (Section 6.10.2) should help further inform users of the credibility of a bidder during various voting phases (Sections 6.5.3 and 6.5.4). Initially, bidders will of course, not have a reputation, so reliance on proof of prior work will be paramount. In addition to this, voters can rely on the assessment received from the expert committees. As previously mentioned, as an added security measure, should voters’ choices deviate from that suggested by expert committees by a certain margin, then those proposals will be flagged for audit by the arbitration committee to make sure there is not anything untoward going on.

While it may seem like that final rule should guarantee that this loophole is closed, it is not. There is still the possibility of tax arbitrage. As previously mentioned, this mechanism could be executed by an individual as well. Since income tax rates are higher than corporate tax rates in most countries including the U.S., if this were executed by an individual in the U.S. and the company used to place bids was also incorporated in the U.S., then that individual would be arbitraging the difference between individual and corporate tax rates. The vast majority of exploits related to fraud will stem from tax exploit attempts and it is beyond the scope of this paper to explore every possible loophole in the tax code that could be exploited via this system. As part of the due diligence one would perform with a lawyer prior to launching a system such as this one, anything related to tax benefits and resulting exploits must be very carefully studied and accounted for.

Outside of the specific fraud scenario outlined here, there will likely be others depending on implementation specifics. Some general guidelines that will help deter fraud and assist with fraud prevention are to audit proposals with relatively low numbers of contributing users and general activity (as measured by the reputation mechanism) that raise large amounts of funding, as well as audits of those individuals who have consistent involvement with (only) such kinds of proposals. Whenever fraud is detected the penalties that are levied must be so large as to deter anyone considering it (Section 6.10.1.5).

Property 7: Reputation systems encourage cooperation and altruism. These systems must allow for the punishment of defectors; but have to strike a balance between doing so in a manner sufficiently strict that it deters defectors yet has some room for forgiveness. These systems cost money to run but people are willing to pay for them (altruistic punishment).

In addition to these static rules for detecting fraud, one could build in various mechanisms that allow for users to act as a public policing/moderation force. There is evidence that people will actively do this without any monetary incentive to do so and will even go so far as to be willing to pay for it themselves even if they get no direct benefit from the system. So, users will have the ability to flag/halt proposals for audit by the arbitration committee when they suspect any of the above scenarios happening; and as previously mentioned, to prevent misuse of this system,
repeated false positive reports will result in penalties on a user’s reputation score, and the fact that only users with sufficiently high reputation scores may make these reports should prevent significant abuse of this system.

6.11.2 Collusion

In the versions of game theory previously discussed, it was always assumed that the actors were making individual choices i.e., all decisions were made independently and that meant that groups of agents could not work together. In these games, there are mathematical proofs [203] that show that there is at least one stable Nash equilibrium (i.e., situations in which neither party can profit by only changing their action). However, there is also a version of game theory that allows for groups of agents to work together called cooperative game theory. This is a large class of games [204] that do not have any such stable Nash equilibrium i.e., do not have stable situations that coalitions cannot profitably deviate from.

These are formally described as games of \( N \) agents where any subset of more than half of them can capture a fixed reward and split it amongst themselves. This should sound familiar, even outside this paper, there are many situations in life that are a part of this set of inherently unstable games. [204] In essence, if there is a situation with a fixed pool of resources and a means to distribute those resources amongst community members, then a majority of those members can conspire to seize control of the resources; no matter what the circumstances are, there will always be a scenario where doing so will be profitable for that group of members. This is quite abstract, and to make it more concrete, exploring the implications of collusion specific to this system are useful.

The first and foremost issue is that of colluding agents working together to convince other users contributing to a given proposal that a certain bidder is better than another because the colluding group has a vested interest in the specific bidder for whom they are advocating. On the one hand, it is unlikely that even if this did happen that it would end badly, because if it did, contributors would have legal recourse as anyone does today when they hire someone to do something, but they do not deliver. The individual/company that does not deliver would also be permanently banned from ever bidding on the system again (Section 6.10.1.5). On the other hand, it does mean that money will not be used in the most efficient way, and it is an unexpected intrusion to a carefully architected process; it could disturb equilibrium that would have otherwise been reached. One means of assisting with this, are the expert committees (Section 6.8.1.4) whose information can help guide voters but also act as a security measure by having arbitration committees manually investigate voting outcomes that differ (by a certain margin) from that predicted by the expert committees - which is what would happen if, for some reason, a colluding group of users managed to convince the rest of the contributors that a specific bidder was better than the rest even if they were not.

However, instead of trying to solve a symptom of collusion, addressing why it is even possible is more important. If perfect harmonization of interests were possible then purely private contributory mechanisms (even the most naive and primitive ones) would lead to optimal outcomes. The premise of this paper is that collective action problems prevent this from happening so designing a system that overcomes that lack of harmonization is
required. However, this assumption that the perfect harmonization of private interests is not possible is only an assumption (it could be true globally but that is unlikely), and so, if such perfect harmonization were possible, then the system falls prey to manipulation. This is a bit of a paradox because to overcome the assumed lack of harmonization, community formation is fostered and communication is encouraged naturally at every step of the process (Section 6.9), but in doing this, there is an assumption about the lack of strong community outside the system, for there were, then collusion would become impossible to stop.

In fact, due to the explicit measures taken to encourage communication and community formation due to its many benefits supported by literature from varying fields (Section 5), collusion is a very subtle and pernicious problem to root out. Once again, small highly funded proposals, users with repeated involvement exclusively with such proposals, as well as repeated involvement in proposals with the same groups of users should be audited. There are various other metrics that can be developed through experimental observation that can help detect collusion. For example, how often do people in a given geographic region contribute to proposals unrelated to that region, and if they do so, what is their general degree of commitment to governance proceedings? Such data can be acquired and set as thresholds for expected behavior, which when deviated from, flags accounts and proposals for audit. As more data is collected one could derive a model of the expected behavior of a fair and truthful user and flag anything deviating from that model. The extent to which user behavior is modeled and flagged for review is left up to implementation and will largely be dictated as a function of how much one cares about the implications of collusion.

While trying to prevent collusion in the face of encouraging communication and community formation may be a tricky problem to solve, there are also several steps that can be taken in the interim that create barriers to collusion. The first being to prevent people from sharing any external contact information with each other in any way. All communication that happens on the system will be automatically reviewed by a mechanism that simply censors contact information. This has two advantages. The first being that it means that just because people worked together once, does not mean they can collude in the future. As wallet addresses are pseudonymous and randomly generated for every proposal a user is a part of (barring those who go through the effort of breaking that pseudonymity by cross-referencing reputation records as described in Section 6.4.2), it should work fairly well to prevent identity leakage (in combination with the aforementioned feature of censoring contact information). The second is that this policy ensures that users continue to use the system, its communication platform, and vertically integrated set of tools for everything related to the proposal and do not jump to third party solutions such as Discord.

The truth is that users will likely find a way around this, the simplest would be to encode contact information using some encryption scheme and then make it known to others that they are using a specific encryption scheme to send this restricted information. So, the censoring mechanism will have to be robust and will need an AI component derived from some LLM that is able to “understand” conversations and make real-time assessments on the degree of information leakage and censor things as needed. However, again, even the best LLMs have their limitations and eventually, there will be some information leakage that could result in sharing of contact
information for external communication that could lead to collusion on future proposals. The only solace is that since larger groups will be split into smaller groups when working on a specific proposal as described in Section 6.8.4.2, even those who do intend on colluding with each other will be randomly split up amongst the various subgroups making it hard to work together to social engineer the rest of the contributors to pick a specific bidder (or any other activity that impedes the progress of proposals), and even if they managed to do so, it is not the “end of the world” (if it was, it would be handled in courts of law), it just means that the system is not operating as efficiently as possible.

6.11.3 Stalling
Over the course of the paper, and especially Section 6.11, there have been multiple occasions where it has been mentioned that proposals may be flagged for audit by users to be reviewed by an arbitration committee. This happens by one of two mechanisms. The first of which, involves setting a bunch of rules that describe the expected behavior of users (this includes cross referencing voting outcomes with those predicted by expert committees), and any deviation from this expected behavior by a certain margin, is flagged and reviewed manually by the arbitration committee. Once a proposal is flagged, it is halted from further progress until the arbitration committee determines that it is safe to proceed. Initially, these rules will be fairly rudimentary but as the system continues to operate, it will be easy to create stronger and more robust models that describe the “normal” behavior of users and proposals, and to flag any deviations while also not flagging certain kinds of behavior that are completely randomly distributed (i.e., cannot be modeled nicely by any probabilistic distribution).

However, despite how robust any such models can be, there will always be situations that seem to fit expected behavior norms but are actually problematic. The use of LLMs and advanced AI in general to tease out some of this behavior is possible (as has been mentioned repeatedly throughout the paper). Yet again, these models have their own limitations. These techniques have sufficient loopholes that allowing for users to report any deviations from expected behavior may be a worthwhile provision. Once again, this has been mentioned on multiple occasions throughout the paper but allowing for users to report various kinds of activity as suspicious and flag proposals for audit (the same audits that automated systems were previously flagging proposals for) is highly recommended. It comes with an obvious downside which is that the system can be abused to stall proposals for any number of reasons.

The only means to reduce the chances it is abused are to only allow users with a certain reputation score (using the limited governance model) to make these reports or to weigh users with higher reputation scores reporting incidents much more heavily than just the average user. Furthermore, repeated false positive reports by users will be penalized with heavy hits to reputation scores and beyond a certain number, will lead to temporary and then eventually permanent bans from the system (Section 6.10.1.5). This should act as a sufficient barrier to abuse, and the tradeoff of potential abuse of this facility is worth it against the benefits of the reports that come in through this mechanism that will save the entire system from larger security breaches that could effectively destroy the system altogether.
Conclusion

This paper presented specifications for a standardized framework for the provision of public goods through free markets via infrastructure (blockchains) that is decentralized; so as to establish trust via complete transparency, aid in establishing self-governance and endogenous community formation, and to eliminate as many perverse incentives as possible that contribute to the collective action problem. The issues surrounding the provision of public goods have been extensively studied in several domains. A lot of this research was addressed directly here but there is a substantial portion of research that was also not. The scope of this problem is far too large to have been covered comprehensively in this paper. There are several topics that are unfairly simplified and treated superficially. Even amongst those discussed in greater detail, there is significant room for improvement, especially as it pertains to implementation details.

A lot of these ideas need to be further fleshed out after extensive large-scale experimentation. During that process it is possible that certain domains of applications seem particularly well suited to this structure. For example, it could function as “honesty as a service” exclusively for charities, it could also exclusively act as a means for organization for local governments (while funding continues to be provided via taxation). This paper laid out the specifications for a system that would work as intended if every single aspect of the process behaved as expected. However, it is extremely unlikely that is the case, as it is unlikely that building a system of this scope as a means to validate the many hypotheses presented would be time well spent. Instead, the idea is to encourage renewed interest around exploring possibilities for public goods to be funded and provisioned with the means of blockchain technology. Many of these ideas can be narrowly validated and iterated upon with speed previously not possible in this domain.

The provision of public goods is a problem that has plagued society from its very origins; however, this may be yet another window of time during which, as a society, there is the chance to, at least, take another step toward provisioning them more efficiently.
Bibliography


[21] Rittenberg and Tregarthen. Principles of Microeconomics, Chapter 6, Section 4, p. 2


[34] "Martin J. Osborne", economics.utoronto.ca.


[63] "This Modern World". Salon.com Comics.


[97] https://www.landregistry-titledeeds.co.uk/frequently-asked-questions/information/public-rights-of-way.asp#:~:text=A%20Byway%20Open%20to%20All,%2C%20cycling%20or%20horse%20riding


[100] https://www.weforum.org/agenda/2020/12/trust-is-the-glue-of-a-healthy-society-heres-how-to-bring-it-back/


[133] https://www.aclu.org/fact-sheet/oppose-voter-id-legislation-fact-sheet


[135] https://www.dhs.gov/news/2022/12/05/dhs-announces-extension-real-id-full-enforcement-deadline

[136] https://insights.yale.edu/insights/what-happens-when-billion-identities-are-digitized


[142] https://twitter.com/mcuban/status/1587547362188316675?s=46&t=QtKZqR--ZHxLY3LmJU2Xw

[143] https://twitter.com/elonmusk/status/1454809318356750337?s=20&t=1It-wch8GXal5IKkhZ4BVA

[144] https://twitter.com/lexfridman/status/1516100448285036546?s=20&t=ZIske2cHbeczqYWFwBkWiQ


[146] https://thepointsguy.com/deals/marriott-promo/


[149] https://www.gatesfoundation.org
[150] https://www.bezosfamilyfoundation.org

[151] https://www.muskfoundation.org

[152] https://www.thejaymo.net/2020/12/23/discord-daos-dweb/

[153] https://www.newyorker.com/culture/cultural-comment/slack-is-the-right-tool-for-the-wrong-way-to-work

[154] https://www.thejaymo.net/2020/10/28/verticals-of-one/

[155] https://www.ourmachine.net/writing/eight-qualities-daos/

[156] https://gov.yearn.finance


[158] https://blog.synthetix.io/transition-to-decentralised-governance/

[159] https://www.gitcoin.co/blog/introducing-gtc-gitcoins-governance-token


[162] https://docs.reflexer.finance/ungovernance/governance-minimization-guide


[164] https://app.proofofhumanity.id

[165] https://www.brightid.org

[166] https://github.com/idena-network

[167] https://fundoss.org

[168] https://bounties.gitcoin.co/grants

[169] https://downtownstimulus.com

[170] https://clr.fund/#/

[171] https://blog.ethereum.org/2014/08/21/introduction-futarchy

[172] https://medium.com/daostack/reputation-vs-tokens-6d7642c7a538


[180] https://twitter.com/0xSisyphus/status/1545113732585558018

[181] https://twitter.com/hasufl/status/1528740714145890304

[182] https://twitter.com/verbine/status/1315959108684722176


[185] https://graymirror.substack.com/p/optimal-autonomous-organizations


[189] https://kleros.io

[190] https://decrypt.co/38050/steem-steemit-tron-justin-sun-cryptocurrency-war