Egalitarian Society or Benevolent Dictatorship: The State of Cryptocurrency Governance

Sarah Azouvi, Mary Maller, and Sarah Meiklejohn

University College London
{sarah.azouvi.13,mary.maller.15,s.meiklejohn}@ucl.ac.uk

Abstract. In this paper we initiate a quantitative study of the decentralization of the governance structures of Bitcoin and Ethereum. In particular, we scraped the open-source repositories associated with their respective codebases and improvement proposals to find the number of people contributing to the code itself and to the overall discussion. We then present different metrics to quantify decentralization, both in each of the cryptocurrencies and, for comparison, in two popular open-source programming languages: Clojure and Rust. We find that for both cryptocurrencies and programming languages, there is usually a handful of people that accounts for most of the discussion. We also look into the effect of forks in Bitcoin and Ethereum, and find that there is little intersection between the communities of the original currencies and those of the forks.

1 Introduction

Cryptocurrencies are an alternative to fiat currencies that aim to replace traditional institutions with a digital platform (or *blockchain*) whose rules are enforced largely by consensus, with anyone able to participate (a property typically called *decentralization*) and check that they are being followed (*transparency*). Whereas fiat currencies inherently rely on some degree of trust in central entities such as banks, blockchains thus promise a radical shift away from trusted parties.

In a decentralized system, no one entity can act to censor transactions or prevent individuals from joining the network (as is possible with traditional institutions [17]). Instead, there is a network of peers that is collectively responsible for entering information into the ledger. In theory, each peer in the network has a "vote" proportional to their computational power, which is used to seal transactions into the ledger. Provided the peers controlling the majority of the network's computational power are honest, only valid information enters the ledger, which — if the ledger is made globally visible — can be checked by anyone. It is unclear, however, to what extent this theoretical promise of decentralization and transparency has been achieved in practice. Indeed, previous research has demonstrated that even the enforcement of the rules is not as decentralized and transparent as originally intended [16,12].

Beyond enforcement, we must also consider how the rules governing a cryptocurrency are set in the first place, and who gets to set them. The founders of a cryptocurrency necessarily make numerous decisions regarding both its design and its implementation. For example, they decide on the interval between the generation of blocks, the reward for generating a block, and the size of the blocks. Thus there is a governance structure underlying all blockchains, and these governance structures have a seemingly inherent degree of centralization. Many cryptocurrencies address this by open-sourcing their code and opening their protocols to so-called "improvement proposals," in which anyone can propose changes to the high-level protocol. These improvement proposals serve not only to reduce the degree of centralization in the maintenance of the platform, but also can provide a significant degree of transparency into the decision-making process. As these rules impact the functioning of a cryptocurrency just as much as the actual enforcement of the consensus protocol, it is important to consider not only the decentralization and transparency of the blockchain itself, but also of its underlying governance structure.

Our contributions. In this paper, we study the centralization in the existing governance structures of Bitcoin and Ethereum, which as of this writing are the top two cryptocurrencies by market capitalization [4]. In order to determine whether or not our results should be expected for any open-source software, we also conduct our study on two popular open-source programming languages: Clojure and Rust.

For each platform, we measured two different properties: the number of developers contributing to each file in the codebase, and the number of people contributing to the discussion around the platform by making comments in the relevant part of its GitHub repository. In terms of contributions to the codebase, we found that the distributions amongst the contributors were all different. In terms of the discussion, we found that for all the systems we studied, at any given time at most eighteen contributors accounted for a majority of all comments.

To evaluate the decision-making infrastructure in Bitcoin and Ethereum, we looked into who creates and comments on improvement proposals. Ethereum appears more centralized than Bitcoin in terms of improvement proposals, but is more decentralized in terms of the discussion around its codebase, according to our metrics, . Finally, we compared the communities behind Bitcoin and its fork Bitcoin Cash, and Ethereum and its fork Ethereum Classic, to see whether these forks bring in new people or split the initial community. In both cases, the fork seems to bring in a new community.

2   Related work

Much of the previous research examining decentralization on blockchains focuses specifically on Bitcoin, or on the general governance issues associated with blockchains. In terms of centralization within Bitcoin, Gervais et al. observe that some of the key operations in Bitcoin, in particular the mining process and the

maintenance of the protocol, are not decentralized [16]. Moore and Christin find a high degree of centralization in popular Bitcoin exchanges [22], and observe that popular exchanges are more likely to suffer security breaches. Böhme et al. look at the various centralized intermediaries within the broader Bitcoin ecosystem, such as currency exchanges, wallet providers, mixers, and mining pools [12]. They also evaluate the decisions that the designers make regarding how much money there should be in the system, and de Filippi and Loveluck examine the overall decision-making process of the Bitcoin developers [13]. In particular, they discuss the "block size" debate and the difficulty in deciding whether or not to fork Bitcoin in order to increase the block size.

In terms of other cryptocurrencies, Reyes et al. examine the theft of 3.6 million ether from The DAO in June 2016, and discuss the lessons learned and the potential strengths and weaknesses of decentralized organizations [26]. Gandal and Halaburda analyze how network effects affect competition in the cryptocurrency market [15]. In particular they look at competition between different currencies and competition between cryptocurrency exchanges and observe that there was a "winner takes all" effect in early markets, but not today.

More generally, Atzori gives a critical evaluation of whether blockchains are suitable as political tools [11], and examines to which extent they can mitigate coercion, centralization, and hierarchical structures. Reijers et al. study the question of governance from the perspective of social contract theory and finds that it fails to incorporate aspects of distributive justice [25]. Similarly, Lehdonvirta poses the "blockchain paradox," in which he argues that once you solve the problem of decentralized governance, you no longer need blockchains [20].

Finally, Srinivasan and Lee [30] introduce a metric for measuring the decentralization in cryptocurrencies that they call the Nakamoto coefficient. We use this metric, in addition to several others, in order to compare and contrast the level of decentralization in Bitcoin and Ethereum.

3 Background

3.1 Bitcoin

Bitcoin [23] was created by the pseudonymous Satoshi Nakamoto, who deployed the currency on January 3 2009. In September 2012, some prominent members of the community created the Bitcoin Foundation, a non-profit organization based on the model of the Linux Foundation, but today there is also a significant development effort by Blockstream [3], which is a for-profit company run by the core developers. .We refer the reader to the Bitcoin textbook [24] for a more technical presentation of the Bitcoin protocol.

3.2 Ethereum

Ethereum [6] was created by Vitalik Buterin, and launched on July 30 2015. Its initial development was done by the Ethereum Foundation, a Swiss non-profit,

but today there is also a significant development effort by Parity [8], which is a for-profit company developing one of the main Ethereum clients. Ethereum is designed to support a broader functionality beyond atomic transfers of money from one set of parties to another.

3.3 Improvement proposals

No system is perfect, and cryptocurrency protocols sometimes need updating due to flaws or vulnerabilities. These changes can be fundamental and affect all users. To keep the improvement decision process open and fair, most cryptocurrencies have an *Improvement Proposal* system, where anyone can propose changes to the protocol and discuss existing proposals. If support exists for a proposal, it may be incorporated into the codebase. There is no formal definition of how to agree upon an improvement proposal [1]. The Improvement Proposals process happens mainly on GitHub, but there are many other places for discussion, such as mailing lists, forums and IRC channels.

3.4 Forks

When disagreements occur in cryptocurrency communities, the only way to resolve them might be for the communities to split. Anyone disagreeing with the current core developers can fork the code and create their own currency. This has happened in both Ethereum and Bitcoin. For example, in June 2016 more than 50M USD of ether was stolen due to a code vulnerability in a smart contract [19]. The Ethereum Foundation decided to "roll back time" in order to take the stolen ether back from the hacker. Arguing that this contradicts the fundamental immutability property of blockchains, some members of the community forked Ethereum and Ethereum Classic was born [7]. In Bitcoin, the "block size" debate has been ongoing for years. Arguing that one of the main limitations of the Bitcoin protocol is scalability and that this problem could be solved with larger block sizes, some members of the community forked Bitcoin, resulting in Bitcoin Cash in August 2017 [2].

4 Methodology

4.1 Comparison with programming languages

To determine whether the governance structures of Bitcoin and Ethereum are as decentralized as should be expected, we compare them against those of open-source, general-purpose programming languages. We chose programming languages as, similarly to cryptocurrencies, they tend to have a large amount of participation from their user communities. For an even closer comparison, we sought out programming languages that: (1) have existed for a similar length of time to the cryptocurrency; (2) have a similar number of users (which we measured according to the number of watchers and stars on the GitHub codebase [18]); and (3) are decentralized in the sense that they are maintained by an

online community rather than a private company or government. We could not find programming languages that fully satisfied each of these properties, but we decided that a relatively fair comparison was between Bitcoin and Clojure, and Ethereum and Rust.

Bitcoin and Clojure were both proposed by individuals (or a set of individuals) and were both released in 2009 (Bitcoin in January, and Clojure in May). While Bitcoin has a much larger userbase than Clojure (close to 2000 watchers and 18k stars, as opposed to roughly 700 watchers and 7k stars), we ultimately decided to stick with this comparison rather than use a programming language like Go, which does have a larger userbase, as Go is closely tied to Google.

Ethereum and Rust were both released in 2015 (Ethereum in July, and Rust in May), and are both tied to not-for-profit foundations (Ethereum with the Ethereum Foundation, and Rust with Mozilla). Rust has a larger, but not incomparable, userbase than Ethereum: roughly 1500 vs. 900 watchers, and 24k vs. 8k stars.

4.2   Data collection

Name	Repository URL
Bitcoin	`https://github.com/bitcoin/bitcoin`
Bitcoin Cash (ABC)	`https://github.com/Bitcoin-ABC/bitcoin-abc`
Clojure	`https://github.com/clojure/clojure`
Ethereum	`https://github.com/ethereum/`
Parity	`https://github.com/paritytech/parity`
Ethereum JS	`https://github.com/ethereumjs/ethereumjs-lib`
Ethereum Ruby	`https://github.com/cryptape/ruby-ethereum`
Ethereum Classic	`https://github.com/ethereumproject`
Rust	`https://github.com/rust-lang/rust`

Table 1: The open-source repositories for the various cryptocurrencies we consider. For Ethereum and Ethereum Classic, the listed repositories contain the code for the Go, C++, and Python versions of the client. Parity is compatible with both Ethereum and Ethereum Classic.

To quantitatively measure the level of centralization in the maintenance of Bitcoin and Ethereum, we analyzed their codebases, and the extent to which these codebases are produced and maintained in a decentralized fashion. We obtained copies of the open-source repositories for Bitcoin, Bitcoin Cash, Ethereum, Ethereum Classic, Rust and Clojure. A summary of the locations of these repositories is in Table 1.

One notable property of these platforms is that Bitcoin has only one reference client, whereas the others tend to have many. For Ethereum, we collected the

repositories for all the clients as listed in the Ethereum documentation.[1] For Ethereum Classic, we considered the Go, C++ and Python clients, as the ones for JavaScript, Java, and Ruby were not listed. The Parity client supports both Ethereum and Ethereum Classic. For Bitcoin Cash, we picked the most popular one in terms of watchers and stars, which was Bitcoin ABC.

Since contribution to the protocol is also captured through discussions in addition to lines of codes written, we also scraped all the discussion threads for pull requests and issues (both open and closed). The discussions of Improvement Proposals were not included in the Bitcoin and Ethereum repositories themselves, so we also scraped the main pages, pull requests, and issues on the respective GitHub repositories for Bitcoin (BIPS) [1] and Ethereum (EIPS) [5].

4.3 Centrality metrics

Centrality metric	Usage
Interquartile range (IQR)	Measure of spread
Interquartile mean (IQMean)	Mean of the data in the IQR
Kolmogorov-Smirnov test	See if two vectors have same probability distribution
Nakamoto index	Minimum # of contributors making 51% of the data
Satoshi index	Minimum % of contributors making 51% of the data
Sørensen-Dice index	Measure of similarity of two sets

Table 2: Centrality metrics used in this paper.

Table 2 lists some of the centrality metrics used in this paper. In addition to these, we also use the mean and the median. The interquartile range (IQR) represents where the bulk of values lie and is computed as the difference between the 75% and the 25%, and the interquartile mean (IQMean) is the mean of the data in the IQR. The benefit of using the IQMean (as compared to the regular mean) is that, as with the median, it is not affected by outliers.

To confirm the statistical significance of our findings, we use a two-sample Kolmogorov-Smirnov test [28,21], which determines whether or not two vectors of values have the same probability distribution. More specifically, it quantifies the distance between the empirical distribution functions of the two samples. The p-value, used to determine the statistical significance of the test, must be under 0.05 in order to reject the null hypothesis (i.e., in order to show that the two vectors have a different distribution). We used the Bootstrap version of the Kolmogorov-Smirnov test [27], which is designed to work on discrete distributions.

The Nakamoto index, introduced by Srinivasan and Lee [30], represents the minimum number of contributors to a dataset needed to get 51% of the data.

---

[1] http://ethdocs.org/en/latest/ethereum-clients/choosing-a-client.html#why-are-there-multiple-ethereum-clients
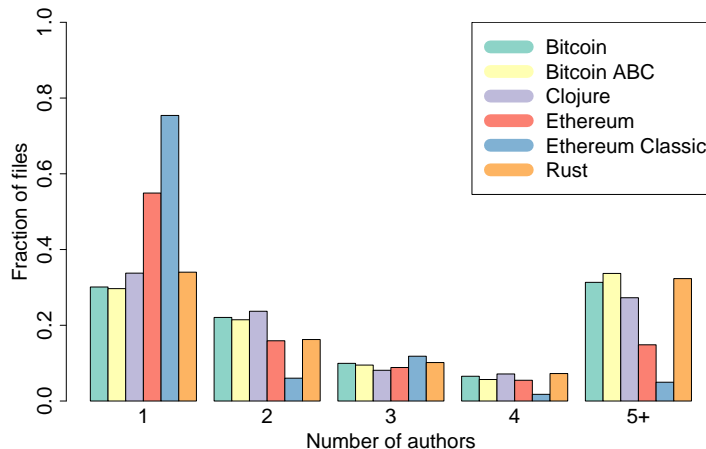
Fig. 1: The coverage of each file in a given repository, as determined by the number of authors that have contributed to that file. Different clients are grouped according to the cryptocurrency they support.

We refer to the normalized version of this index as the Satoshi index, which represents the minimum percentage of all contributors needed to get 51% of the data. Finally, the Sørensen-Dice index [29,14] captures the similarity of two sets. It is defined as $\mathsf{SD}(X, Y) = \frac{2|X \cap Y|}{|X|+|Y|}$, so in particular has a value of 1 for sets that are equal and 0 for sets that are disjoint.

5 Data Analysis

5.1 Contributors to the main codebase

For each repository, we collected all non-hidden files and measured how many distinct authors had contributed to that file throughout its lifetime (in terms of Git commits). The results of this measurement can be seen in Figure 1. While the number of contributors to the Bitcoin and Bitcoin Cash codebases follow a fairly similar pattern, the number of contributors to the Ethereum Classic codebase follows a different distribution to that of the Ethereum codebase, even though Ethereum Classic is a fork of Ethereum. Both Clojure and Rust seem to follow a fairly similar pattern to that of Bitcoin.

For Bitcoin, 30% of all files were written by a single author, and 24% of these files were written by the same author, Wladimir van der Laan. This means that one author wrote 7% of the files. In Ethereum, 55% of all files were written by a single author, and 36% of these files were written by the same author, Tomasz Drwiega. This means that one author wrote 20% of the files.

Table 3 contains the mean, median, IQR, and IQMean for each of the repositories. We see relatively similar metrics for Bitcoin (and Bitcoin Cash) and both

Repository	# Authors per file				# Comments per author			
	Mean	Median	IQR	IQMean	Mean	Median	IQR	IQMean
Bitcoin	5.56	2	5	2.78	27.2	2	4	2.4
Bitcoin Cash	5.48	2	6	2.94	3.8	2	2	1.8
Ethereum	2.58	1	2	1.49	17.0	2	3	1.9
Ethereum Classic	1.69	1	0	1.00	11.1	2	3	1.8
Clojure	4.03	2	4	2.41	1.9	1	1	1.3
Rust	5.17	2	5	2.79	69.8	3	8	3.5

Table 3: Centrality metrics for the number of contributors per file and the number of comments per author in the pull requests and issues.

(a) Satoshi index (b) Nakamoto index

Fig. 2: The evolution of the Satoshi and Nakamoto indexes over time. The values for Ethereum are in blue, for Bitcoin in red, and for Rust in green.

of the programming languages, and see that Ethereum and Ethereum Classic are both lower for all metrics.

We performed a Kolmogorov-Smirnov test to confirm the statistical significance of our findings; the resulting p-values are in Table 7. We see that the results are statistically significant for all the comparisons except for Bitcoin and Bitcoin Cash, indicating that the respective numbers of codebase authors are drawn from different distributions, except for Bitcoin and Bitcoin Cash. This is expected as Bitcoin Cash is a very recent fork of Bitcoin.

We acknowledge, however, that measuring levels of centralization by looking at the codebase is limited in some respects, as it is not practical — or even necessarily important for accountability — to have many people contributing to the same files, and there are likely people looking over and discussing files in ways that are not reflected in Git commits. This is why we look next at the discussion around the code.

5.2 Commenters on the main code base

To get a feeling for the evolution of the distribution of comments over the life of a cryptocurrency, we compute the Nakamoto and Satoshi indexes over time

in Figure 2, using the discussion threads in the pull requests and issues of the GitHub repository. These graphs exclude Clojure, as ultimately we believed the dataset was too small to get any real insights. For example, there were only 72 pull requests, as compared to 11,604 for Bitcoin. Similarly, since Bitcoin Cash and Ethereum Classic are relatively recent forks, and thus have a far smaller level of discussion so far, we also exclude them from this analysis.

In Figure 2, we see that in all the repositories there is a strong tendency towards centralization in the number of commenters, with a handful of people contributing to most of the comments. The Nakamoto indexes for the codebases of Bitcoin, Rust, and Ethereum are consistently relatively low, as every month there are no more than 10 authors contributing to half of the comments for Bitcoin and Rust and 15 for Ethereum. When normalized by the total number of commenters per month, for Bitcoin and Ethereum this is less than a quarter of the commenters each month (as seen in Figure 2a).

	10	20	30	40	50	60	70	80	90	100
Bitcoin	1	1	3	5	8	13	21	41	239	2443
Clojure	2	5	9	15	23	33	45	65	85	104
Ethereum	2	5	8	12	18	29	49	127	467	3139
Rust	1	1	1	2	4	10	21	43	181	3882

Table 4: Minimal number of commenters that contribute to $x\%$ of all the comments.

To see whether it was the same people making most of the comments each month or different people every time, we plotted in Figure 3 the number of comments per author every month. For Bitcoin and Rust, we see that there is one commenter that accounts for most of the comments each month (for Bitcoin, Wladimir van der Laan is the top commenter with 13,923 comments in total, followed by Jonas Schnelli with 4,409 comments), and for Ethereum there is a small handful of commenters who stand out from the rest (the top three are Gavin Wood with 3,352 total, Péter Szilágyi with 2,242, and Jeffrey Wilcke with 2,230). Overall for Bitcoin there are only eight people contributing to half of all the comments, which represents 0.3% of all commenters. For Ethereum there are 18 people (or 0.6% of all commenters), and for Rust there are four (or 0.1%). These results are summarized in Table 4.

This centralized trend is confirmed by the values in Table 3, as we see that the mean is much greater than the IQMean or median, which are values that typically ignore outliers. The mean is one order of magnitude higher than the IQMean for Bitcoin, Ethereum, and Rust. This means that the tails of the distribution (i.e., the top 25% of the distribution) differ a lot from the value in the main range. This can also be confirmed by looking at the number of comments for the top commenters, compared to the average number of comments per author. Generally this confirms that a handful of people (less than 10) contribute to most

(a) Bitcoin

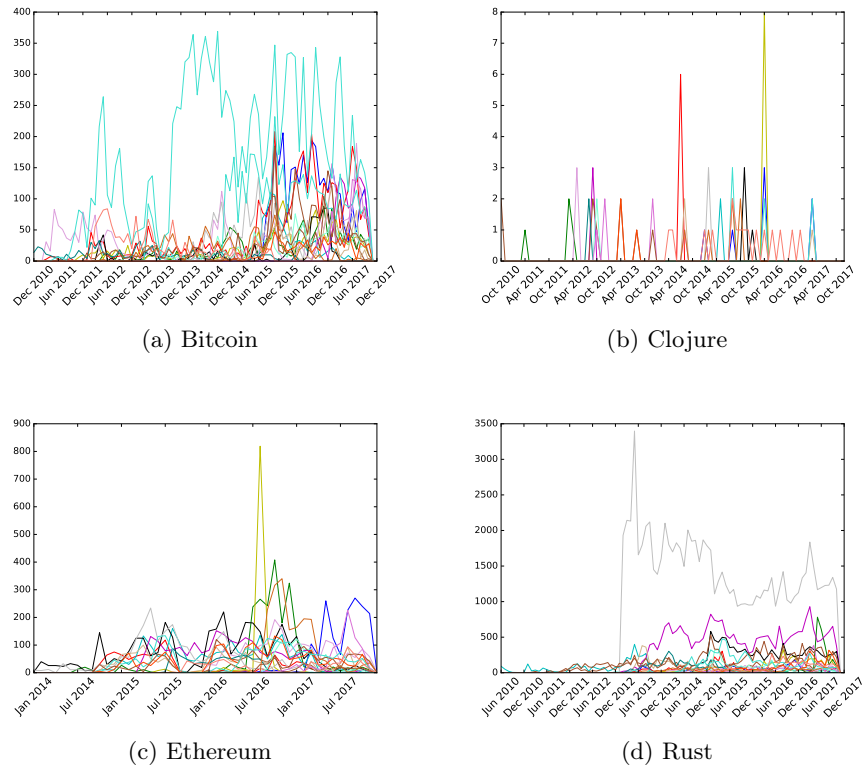(b) Clojure

(c) Ethereum

(d) Rust

Fig. 3: Number of comments per commenters per month.

of the comments. As this is true for all the repositories, we conclude that this is potentially a common (and somewhat natural) feature in open-source systems.

We also computed, in Table 8, a Kolmogorov-Smirnov test on the total number of comments per author. We see that the number of comments per author from Bitcoin, Ethereum, and Rust are drawn from different distributions. In the next two sections, we will focus on Bitcoin and Ethereum, looking more closely at the improvement proposals process in Section 5.3 and comparing the communities behind the main codebases, the improvement proposals, and forks in Section 5.4.

5.3 Improvement Proposals for Bitcoin and Ethereum

In this section, we looked at the improvement proposal (IP) process. Together with pull requests, this is the main road to contributing to the design and development of the currency. For each author we counted how many improvement proposals they made to Ethereum and Bitcoin, and what states these proposals were in (i.e., if they were accepted, rejected, or under review). In Figure 4, we notice that only a handful of people are contributing to Bitcoin improvement
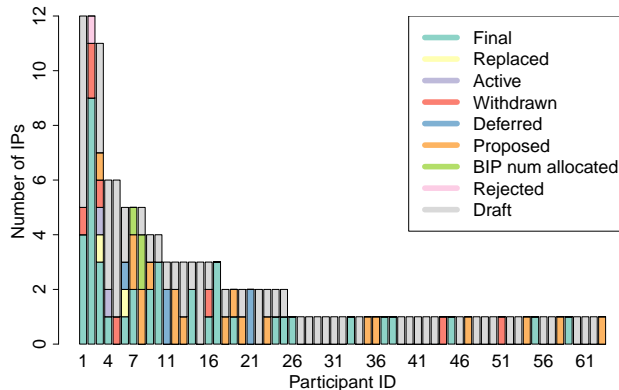
Fig. 4: The authors of BIPs, identified by a unique numeric value, along with the number of proposals they have created and the respective status of those proposals. The top five contributors are Gavin Andresen (with 12 proposals and 9 accepted) and Pieter Wuille (12 proposals, 4 accepted), Luke Dashjr (11 proposals), Eric Lombrozo (6 proposals), and Johnson Lau (6 proposals).

	Mean	Median	IQR	IQMean
BIPS	11.41	2.0	6.5	2.95
EIPS	9.16	2.0	5.0	2.56

Table 5: Centrality metrics for the number of comments per author.

proposals (BIPS). In Figure 5, there is mostly just one person, Vitalik Buterin, that is contributing to Ethereum improvement proposals (EIPS).

There are usually many people contributing to the discussion for every proposal, so we measured the level of centrality in terms of the number of comments in pull requests for each user in the BIPS and EIPS repositories. The results are in Table 5. The trend here is similar to the one observed in the previous section: the datasets contain many outliers, corresponding to the top 25% of commenters who comment significantly more than the rest.

5.4 Diversity of communities

In this section, we look at whether or not the same people contribute to the discussion in the main codebase and in the improvement proposals, and whether or not there is any similarity between the community behind a cryptocurrency and its fork; i.e., any resemblance between Bitcoin and Bitcoin Cash and between Ethereum and Ethereum Classic. Because Ethereum Classic does not have a separate implementation for every client, we focus in this section only on the Go client for each platform, as it is the most popular.

To do this, we first computed the Sørensen-Dice index on the set of the 30 top commenters, which account for roughly 75% of all the comments in the
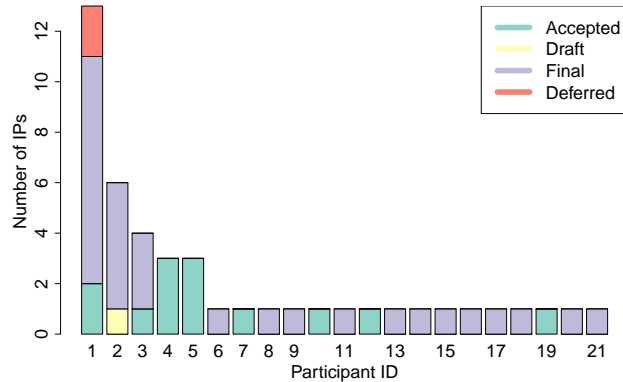
Fig. 5: The authors of EIPs, identified by a unique numeric value, along with the number of proposals they have created and, if in the main set of EIPs, the status of those proposals. The top three contributors are (1) Vitalik Buterin, with 13 proposals, 11 of which were accepted or finalized; (2) Alex Beregszaszi, with 6 proposals; and (3) Nick Johnson, with 4 proposals.

Fig. 6: Improvement Proposals for Bitcoin and Ethereum.

relevant repositories (see Table 4 in the Appendix). As we see in Table 6, the set of main commenters in the main Bitcoin repository and in the BIPS repository overlap, with a Sørensen-Dice index of 0.5. This means that out of the 30 main commenters of Bitcoin and BIPS, 15 are in both communities. This is much more than for Ethereum compared to EIPs, with 7 commenters in both sets.

To include all commenters, we use a weighted version of the whole set of commenters. To do so, we weighted commenters by their number of comments and then computed the Sørensen-Dice index on these augmented sets. The results are in Table 6. Taking the weight (and all the commenters) into account, the similarity between Ethereum and EIPS is still meaningful, with an index of 0.108. The value for Bitcoin vs BIPS, however, drops to 0.069. Therefore, although half of the main commenters for Bitcoin also comment on BIPS, they do not write as many comments in the BIPS repository.

The overlap in the communities of Bitcoin and Bitcoin Cash, and Ethereum and Ethereum Classic, is small. The Sørensen-Dice index was 0.033 for both. Hence only one of the top commenters of the main repository is also a commenter in the forked one. This low value shows that the forked currency is really the formation of a new community rather than a separation of the initial one.

6 Discussion

According to our metrics on the number of contributors per file in the code-bases, we found that Bitcoin, Rust, and Clojure were all more decentralized

| | Sørensen-Dice index | |
Repositories	Top 30	All (weighted)
Bitcoin / BIPS	0.50	0.0686
Ethereum / EIPS	0.23	0.1077
Bitcoin / Bitcoin Cash	0.03	0.0050
Ethereum / Ethereum Classic	0.03	0.0030

Table 6: Sørensen-Dice index for the top 30 commenters, and weighted Sørensen-Dice index for all the commenters.

than Ethereum, even given the fact that Ethereum has many more reference clients. The distributions of the number of authors for all the codebases was different except for Bitcoin and Bitcoin Cash, which is not surprising given that Bitcoin Cash is a recent fork of Bitcoin and thus their codebases are very similar. Interestingly, while Ethereum Classic is a fork of Ethereum, the number of authors on these two codebases is still quite different. However, this fork happened a longer ago and there have been numerous changes to the Ethereum Classic codebase since the fork occurred. Our data implies that one cannot necessarily assume a natural pattern for the number of authors on an open source code base.

There was a greater number of participants in the Bitcoin improvement proposals than in those of Ethereum. Although the distribution of the number of comments on the Bitcoin codebase was different from the one on the BIPS, the distribution from Ethereum was similar to EIPS (Table 8). The intersection between the main commenters on Bitcoin's main codebase and the commenters on the BIPS was greater than the intersection between the commenters on Ethereum's main code base and the commenters on the EIPS. However, when considering the weighted intersection, we found the opposite applied. Generally there are very few people that account for most of the comments for Bitcoin, where for Ethereum this number is higher.

Finally, both the Bitcoin and Ethereum communities seem relatively unaffected by the hard forks. The number of people commenting was not significantly different before and after the forks, and there was little intersection between the people participating in the original codebases and the forked codebases. This implies that the forks did not split the communities, and that a large proportion of the community decided to stay with the original codebases. However in our discussion we only considered Bitcoin ABC, the most popular client for Bitcoin Cash, which could limit our results. We leave for future work the study of all the Bitcoin Cash clients. Our data implies that there could feasibly be a natural pattern in the number of comments per author in cryptocurrencies.

7   Conclusions

Measuring levels of centralization by looking at the codebase or by looking at specific sources is inherently limited. While our measurements captured the number of people writing code changes and commenting on the GitHub files, they do

not capture the number of people voting on whether or not changes should be accepted. We also did not capture conversations appearing in other places such as on Reddit, the main forums, or the mailing lists. We considered only two main cryptocurrencies, but there are a multitude of other ones, and it would be interesting to see whether similar patterns appear in these other cryptocurrencies, or indeed in other open-source projects in general.

We are aware of two projects that aim to tackle the centralization in governance structures of cryptocurrencies: Tezos [10], a decentralized system that incorporates governance into the consensus protocol, and Steemit [9], a decentralized social media platform in which users are incentivized to post and curate content by receiving a reward in the native cryptocurrency. However, we are not aware of any studies that analyze these solutions. Mostly we hope this work will encourage other work that proposes metrics for centrality, or other empirical studies on the governance structures of decentralized platforms.

Acknowledgments

References

1. BIPS GitHub. `github.com/bitcoin/bips`.
2. Bitcoin Cash. `https://www.bitcoincash.org/`.
3. Blockstream. `https://blockstream.com/`.
4. Cryptocurrency market capitalizations. `https://coinmarketcap.com/`. Accessed: 2017-11-03.
5. EIPs GitHub. `github.com/ethereum/EIPs`.
6. Ethereum. `https://ethereum.org/`.
7. Ethereum Classic. `https://ethereumclassic.github.io/`.
8. Parity. `https://parity.io/`.
9. Steemit. `steemit.com`.
10. Tezos. `tezos.com`.
11. M. Atzori. Blockchain technology and decentralized governance: Is the state still necessary?, 2015.
12. R. Böhme, N. Christin, B. Edelman, and T. Moore. Bitcoin: Economics, technology, and governance. *The Journal of Economic Perspectives*, 29(2):213–238, 2015.
13. P. De Filippi and B. Loveluck. The invisible politics of bitcoin: governance crisis of a decentralised infrastructure. *Internet Policy Review*, 5, 2016.
14. L. R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.
15. N. Gandal and H. Halaburda. Can we predict the winner in a market with network effects? Competition in cryptocurrency market. *Games*, 7(3):1–21, 2016.
16. A. Gervais, G. Karame, S. Capkun, and V. Capkun. Is Bitcoin a decentralized currency? In *IEEE Security and Privacy*, 2014.

17. A. Greenberg. Visa, Mastercard move to choke WikiLeaks, Dec. 2010. `http://www.forbes.com/sites/andygreenberg/2010/12/07/visa-mastercard-move-to-choke-wikileaks/`.

18. Y. Hu, J. Zhang, X. Bai, S. Yu, and Z. Yang. Influence analysis of GitHub repositories. *SpringerPlus*, 5(1):1268, 2016.

19. C. D. Jesus. The DAO heist undone: 97% of eth holders vote for the hard fork, July 2016. `https://futurism.com/the-dao-heist-undone-97-of-eth-holders-vote-for-the-hard-fork/`.

20. V. Lehdonvirta. The blockchain paradox: Why distributed ledger technologies may do little to transform the economy, 2016. `www.oii.ox.ac.uk/the-blockchain-paradox-why-distributed-ledger-technologies-may-do-little-to-transform-the-economy/`.

21. S. Mehta. *Statistics Topics*. Createspace Independent Pub, 2014.

22. T. Moore and N. Christin. Beware the middleman: Empirical analysis of Bitcoin-exchange risk. In *Proceedings of Financial Cryptography and Data Security*, pages 25–33, 2013.

23. S. Nakamoto. Bitcoin: A peer-to-peer electronic cash system, 2008.

24. A. Narayanan, J. Bonneau, E. Felten, A. Miller, and S. Goldfeder. *Bitcoin and cryptocurrency technologies*. Princeton University Press, 2016.

25. W. Reijers, F. O'Brolcháin, and P. Haynes. Governance in Blockchain technologies and social contract theories. *Ledger*, 1(0):134–151, 2016.

26. C. L. Reyes, N. G. Packin, and B. P. Edwards. Distributed governance, 2016.

27. J. S. Sekhon. Multivariate and propensity score matching software with automated balance optimization: The Matching package for R. *Journal of Statistical Software*, 42(7):1–52, 2011.

28. R. Simard, P. LÉcuyer, et al. Computing the two-sided Kolmogorov-Smirnov distribution. *Journal of Statistical Software*, 39(11):1–18, 2011.

29. T. Sø rensen. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on danish commons. biologiske skrifter/kongelige danske videnskabernes selskab 5: 1-34. 5:1–34, 01 1948.

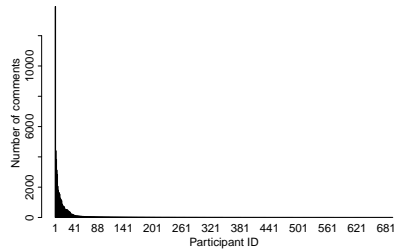30. B. Srinivasan and L. Lee. Quantifying decentralization, 2017. `https://news.earn.com/quantifying-decentralization-e39db233c28e`.

A   Statistical Tables and Figures

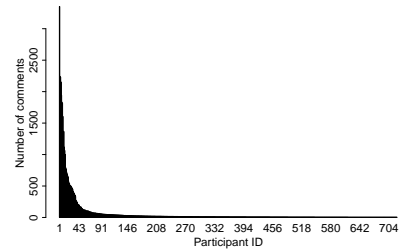|	Bitcoin Cash	Clojure	Ethereum	Ethereum Classic	Rust
Bitcoin	0.4749	0.001	$< 10^{-16}$	$< 10^{-16}$	0.001
Bitcoin Cash		0.003	$< 10^{-16}$	$< 10^{-16}$	0.002
Clojure			$< 10^{-16}$	$< 10^{-16}$	0.028
Ethereum				$< 10^{-16}$	$< 10^{-16}$
Ethereum Classic					$< 10^{-16}$

Table 7: p-values for the Kolmogorov-Smirnov test on the number of authors per file.

	Bitcoin ABC	BIPS	Clojure	Ethereum	Ethereum Classic	EIPS	Rust
Bitcoin	0.045	0.04	0.113	0.029	0.583	0.414	$< 10^{-16}$
Bitcoin ABC		0.008	0.142	0.027	0.12	0.041	$< 10^{-16}$
BIPS			0.015	0.434	0.958	0.285	0.712
Clojure				0.033	0.043	0.07	0.021
Ethereum					0.857	0.536	$< 10^{-16}$
Ethereum Classic						0.854	0.873
EIPS							0.044

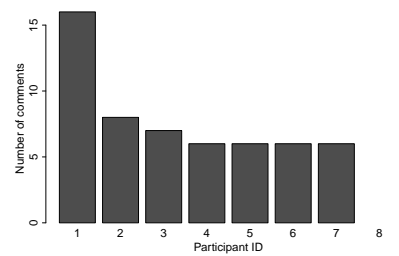Table 8: p-values for the number of comments per author
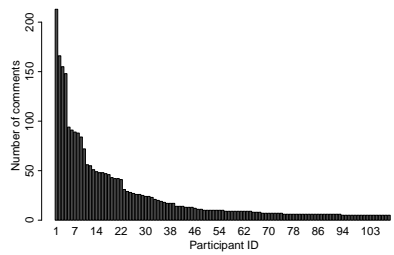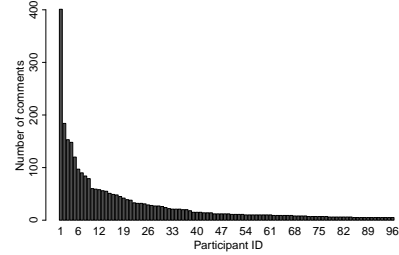

(a) Bitcoin

(b) Ethereum

(c) Rust

(d) Clojure

(e) EIPs

(f) BIPS

Fig. 7: The number of commenters for each repository, ranked from most to fewest comments, ignoring commenters with less than 5 comments.