THE UNIVERSITY *of* EDINBURGH

BLOCKCHAIN TECHNOLOGY LAB

Do Blockchain Systems Achieve Decentralization?

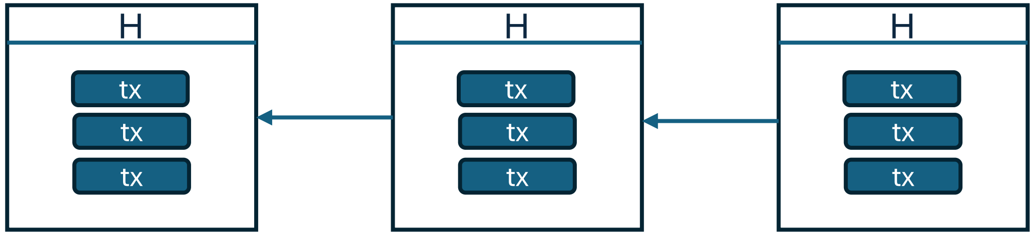
Christina Ovezik

AGT@Blockchains Workshop - WINE 2024

Based on work with Aggelos Kiayias, Dimitris Karakostas, Daniel Woods

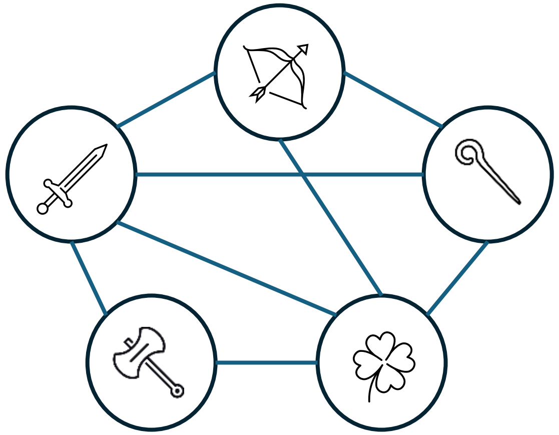
What is a blockchain?

- Digital ledger that satisfies a set of liveness and safety properties, "**without relying on trust**"
- Mechanism design perspective: a protocol with **decentralisation** as an objective



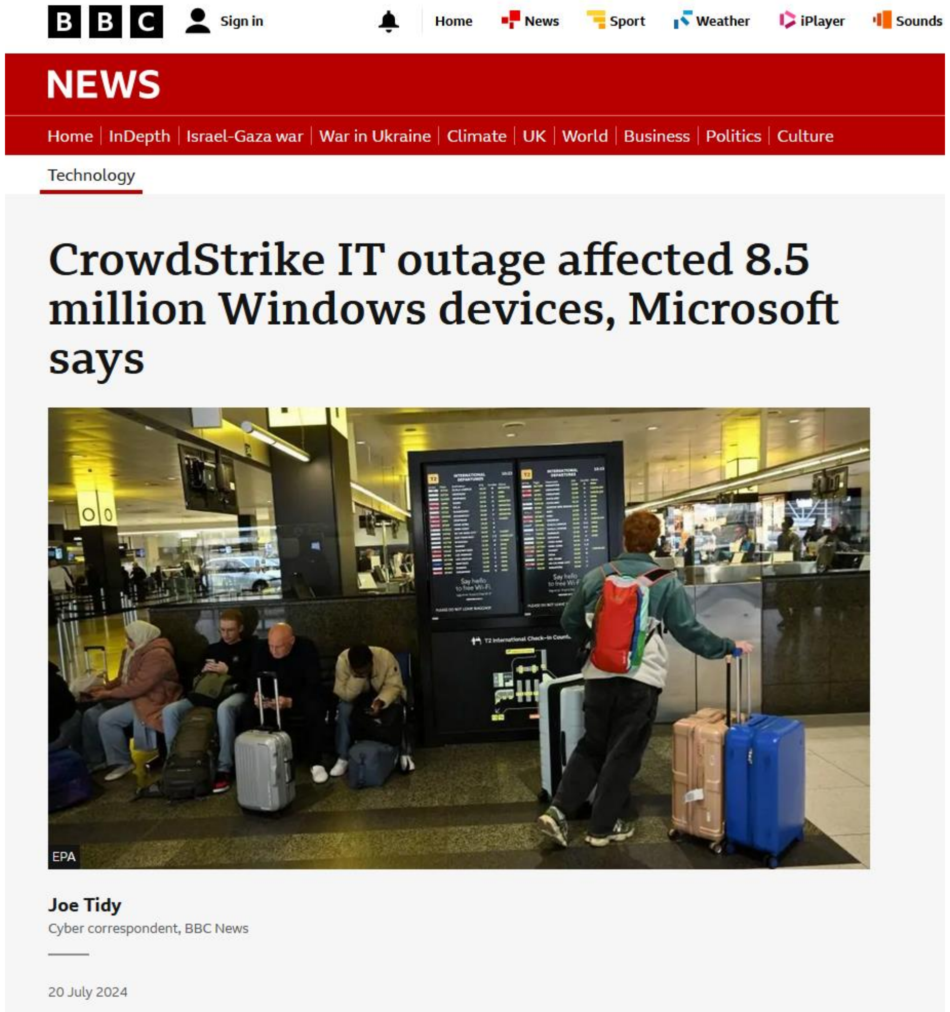
Centralised vs Decentralised systems

- A **single party** has full control over the system (“rules them all”)
- If it misbehaves or crashes, the system is compromised, e.g.:
  - Censorship
  - Inconsistency
  - Data loss / unavailability



- A number of **independent parties** collectively control the system and guarantee its properties
- If only few nodes are malicious / faulty, the system remains operational
  - No single points of failure

Example of centralised system failure


BBC
Sign in
Home
News
Sport
Weather
iPlayer
Sounds
NEWS
Home | InDepth | Israel-Gaza war | War in Ukraine | Climate | UK | World | Business | Politics | Culture
Technology
EPA

CrowdStrike IT outage affected 8.5 million Windows devices, Microsoft says

Joe Tidy

Cyber correspondent, BBC News

20 July 2024

Blockchains are multi-layered systems

- **8 layers of** blockchain systems where **(de)centralisation** can occur

- **Centralisation** in some layer **threatens** desirable system properties, such as **safety, liveness, privacy and price stability**

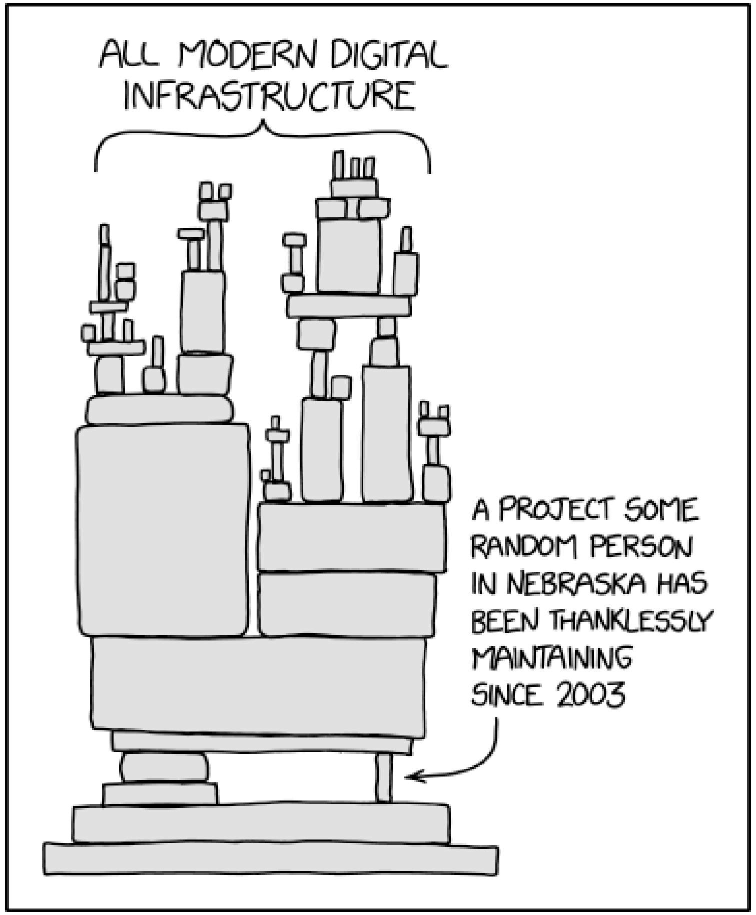

ALL MODERN DIGITAL INFRASTRUCTURE
A PROJECT SOME RANDOM PERSON IN NEBRASKA HAS BEEN THANKLESSLY MAINTAINING SINCE 2003

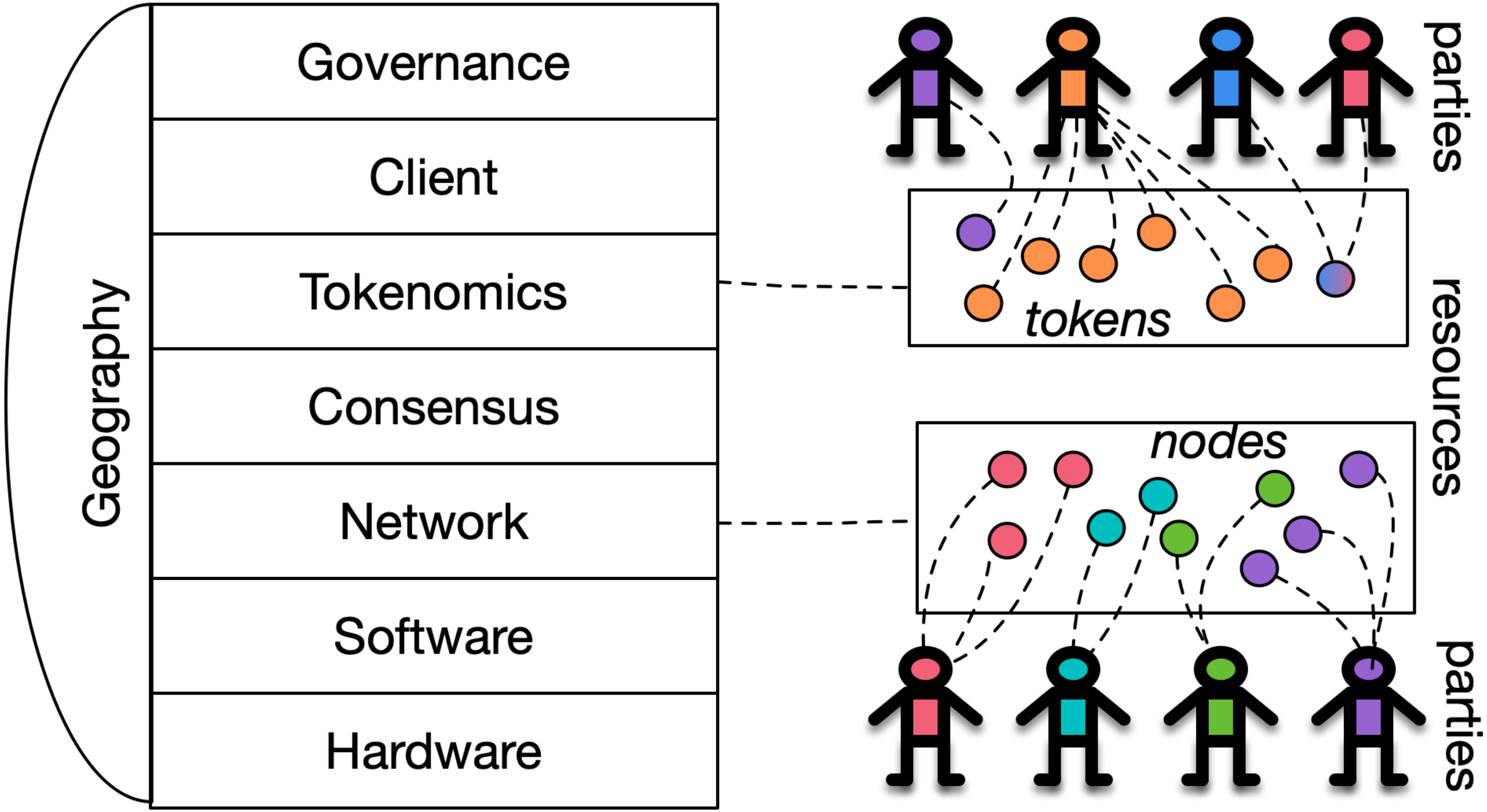
Image credit: https://xkcd.com/2347/

For each layer

Resource

Relevant parties

Properties at risk

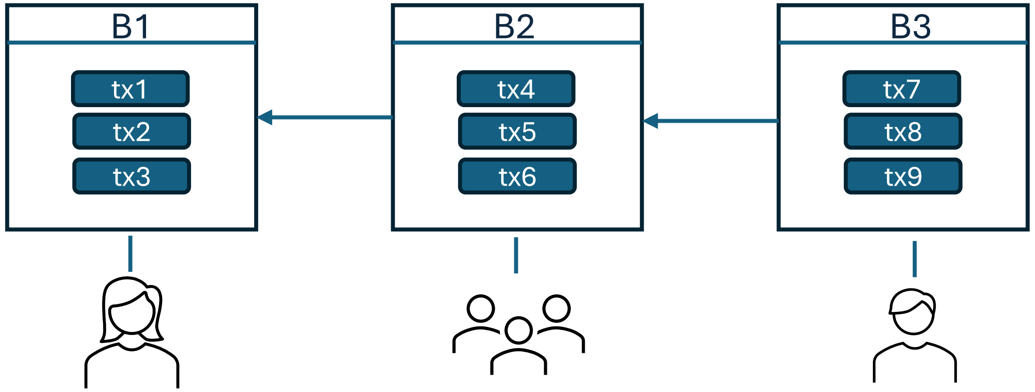

Governance
Client
Tokenomics
Consensus
Network
Software
Hardware
Geography
parties
tokens
resources
nodes
parties

Case study: consensus layer

Measuring the decentralisation of block production

Consensus layer: extending the blockchain

- Block creators:
  - Decide which transactions get included in a block
    - and in what order
  - Receive rewards for each new block
- The more blocks one creates the more influence they have in the system



Consensus – Decentralization Analysis

Resource
=
Blocks

Relevant parties
=
Block
producers

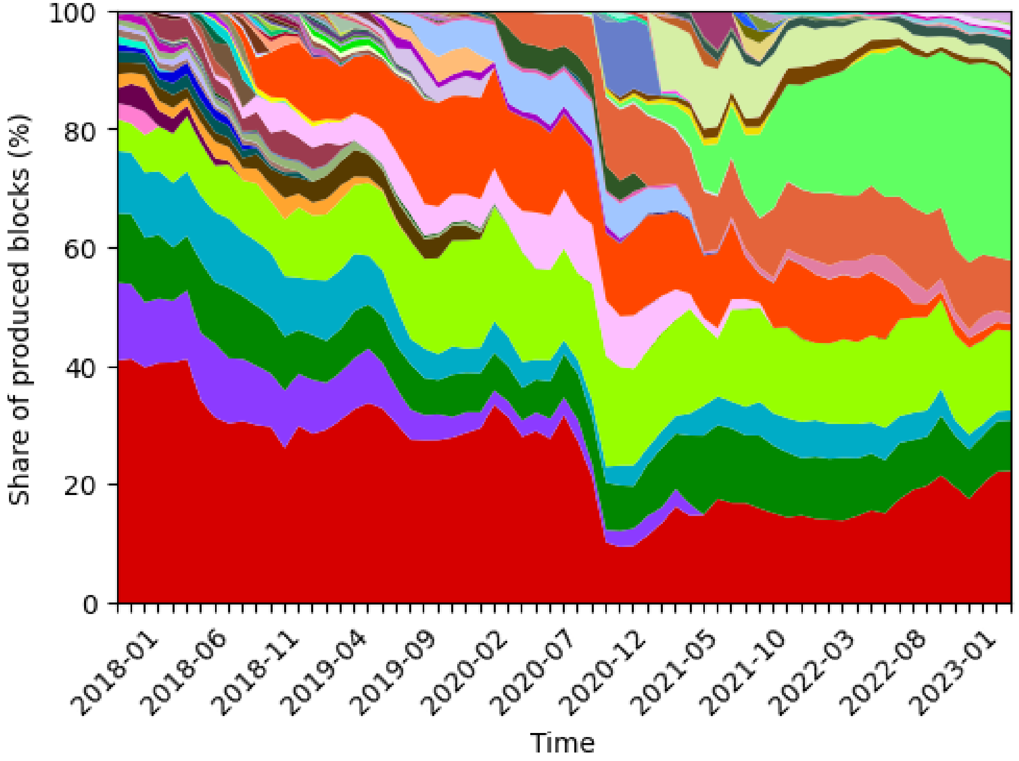
Properties
at risk
=
Safety / Liveness

Mining and the formation of coalitions

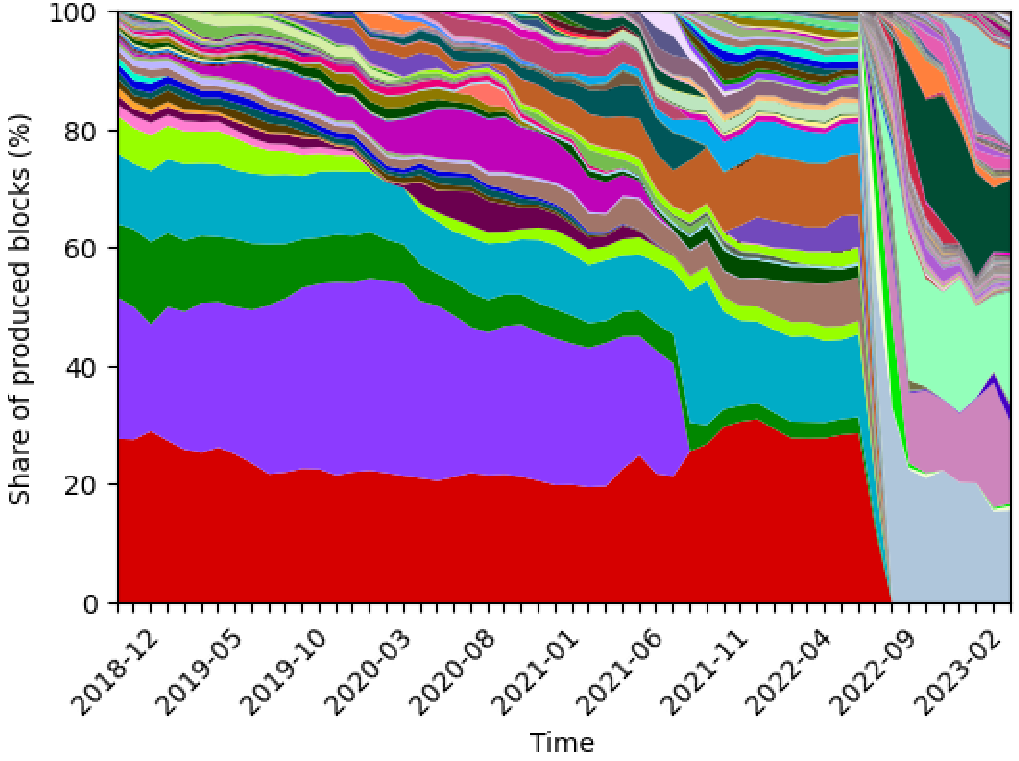
- Economies of scale incentivize the formation of coalitions (**mining / stake pools**) with their leaders consolidating disproportionate power
- We treat each such coalition as a single entity

Block production dynamics

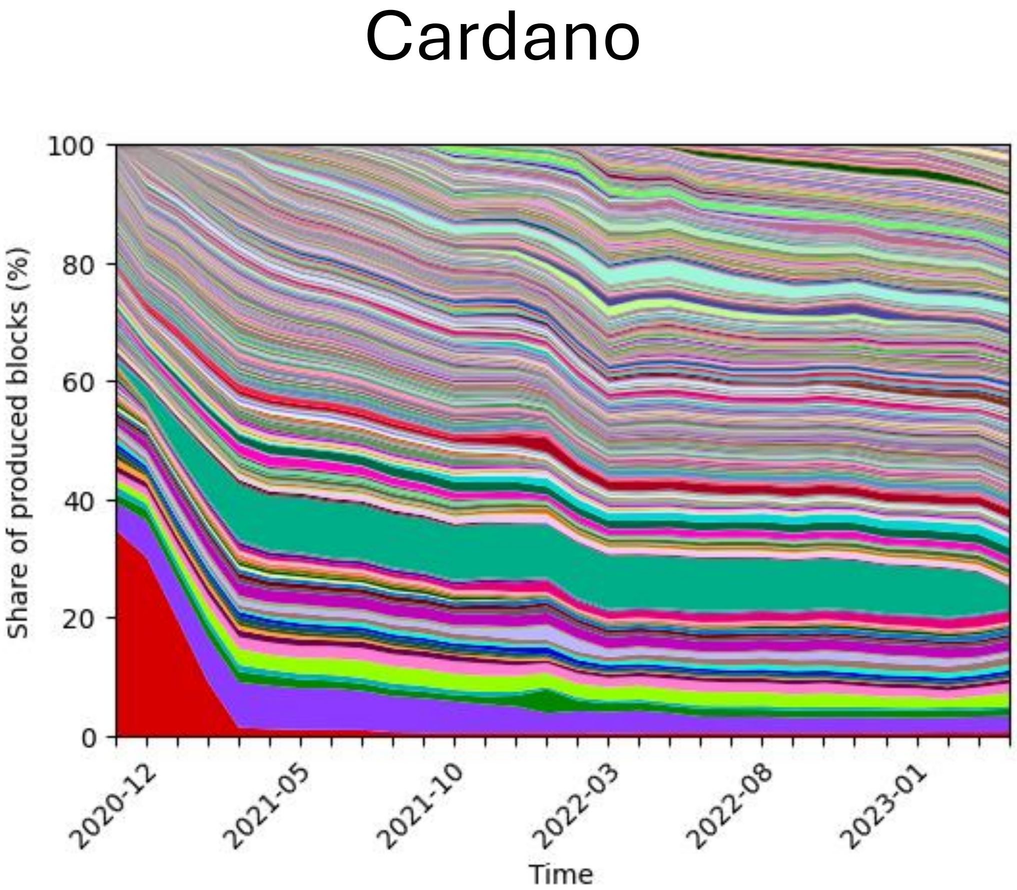
Bitcoin

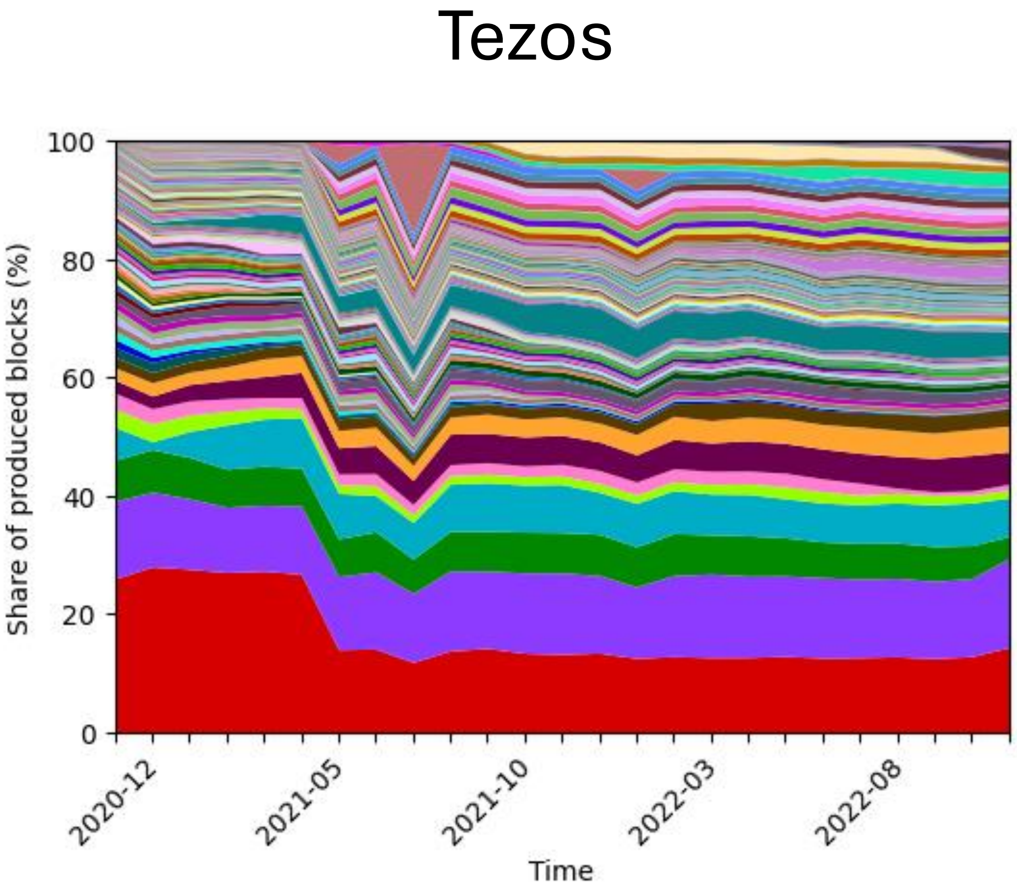


Ethereum



Block production dynamics

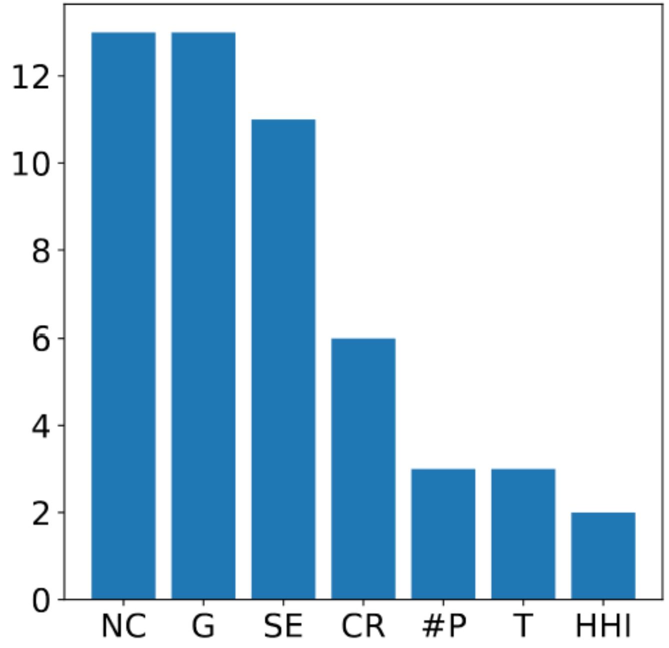




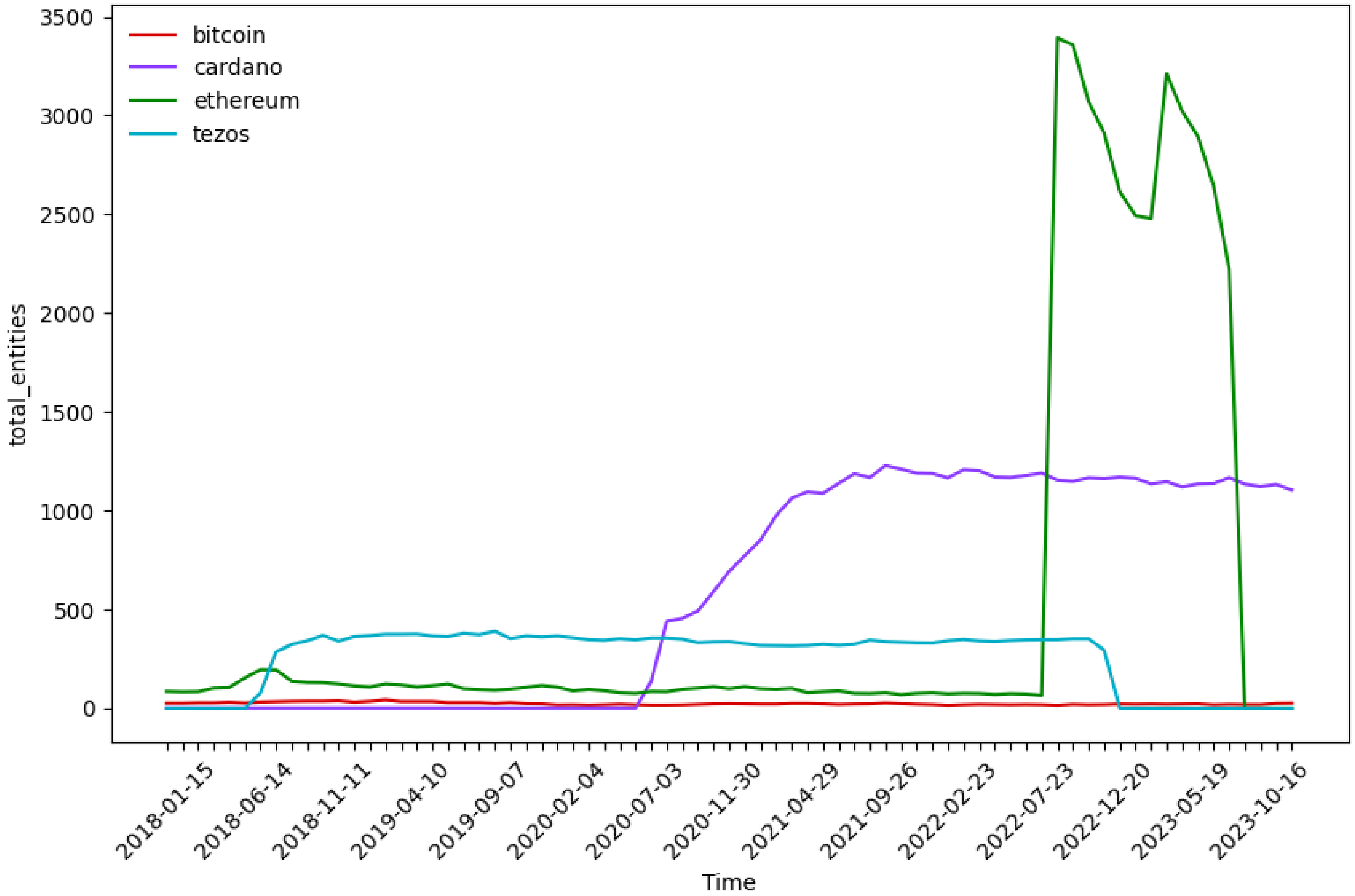
Decentralization metrics

- Assign a value that represents the decentralization of a distribution
- Metrics used in the blockchain decentralization literature:
    - Nakamoto coefficient (NC)
    - Gini coefficient (G)
    - Shannon entropy (SE)
    - Herfindahl–Hirschman index (HHI)
    - Concentration ratios (CR)
    - Number of parties (P)

Metric usage in the literature

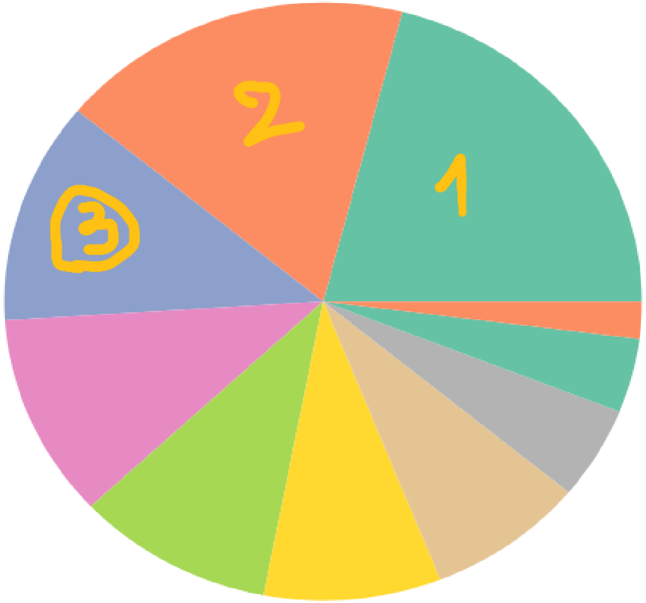


Number of parties

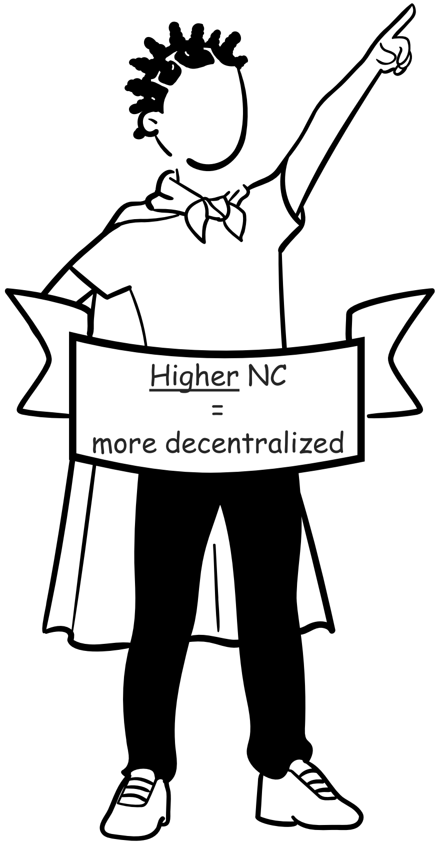


Nakamoto coefficient

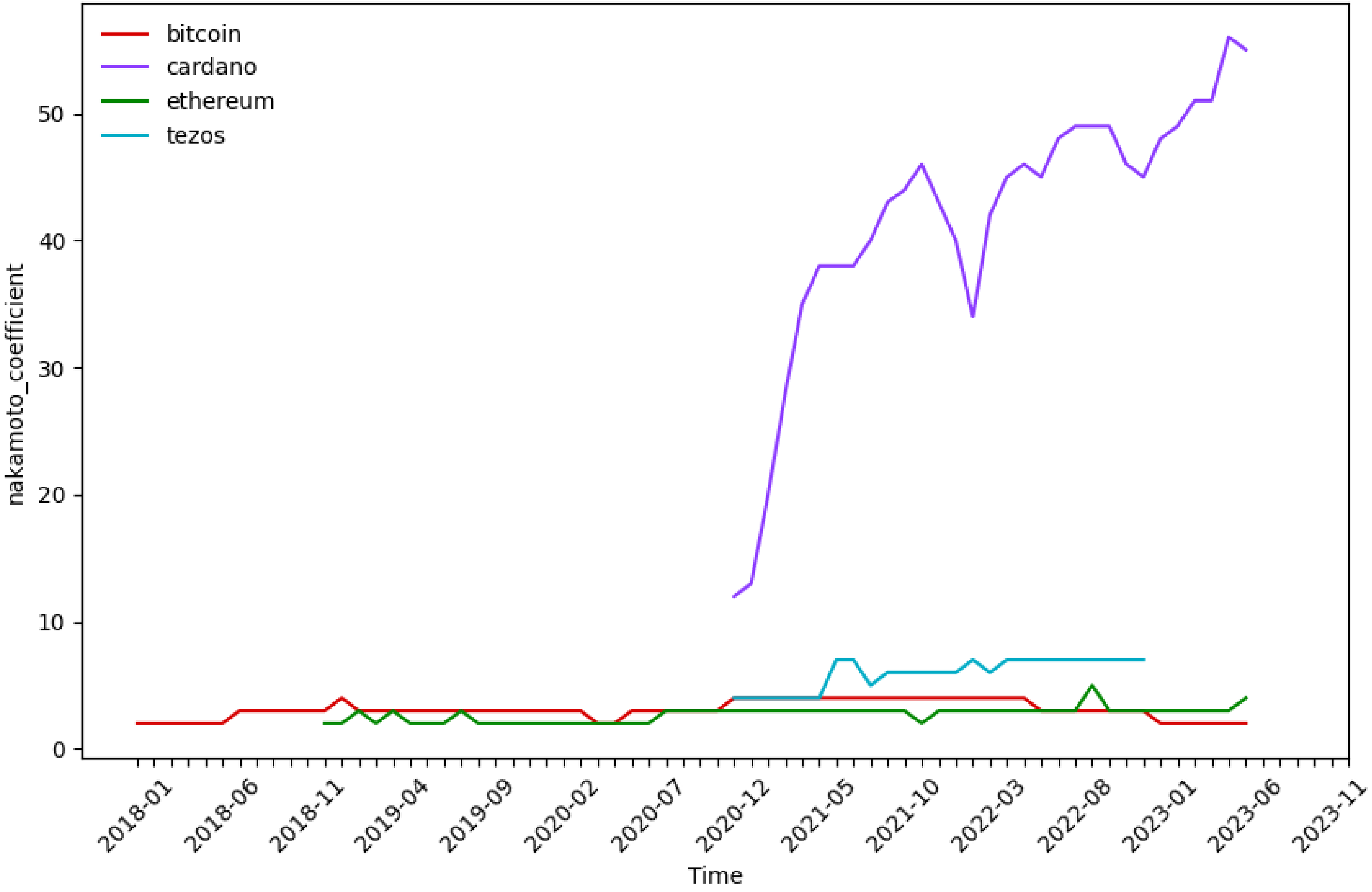
- Represents the **minimum number of entities** that collectively control a **majority of resources** (> 50%)
- ..aka the number of parties that need to collude in order to launch a 51% attack
- The higher the Nakamoto coefficient, the higher the resilience to a majority attack



Nakamoto coefficient

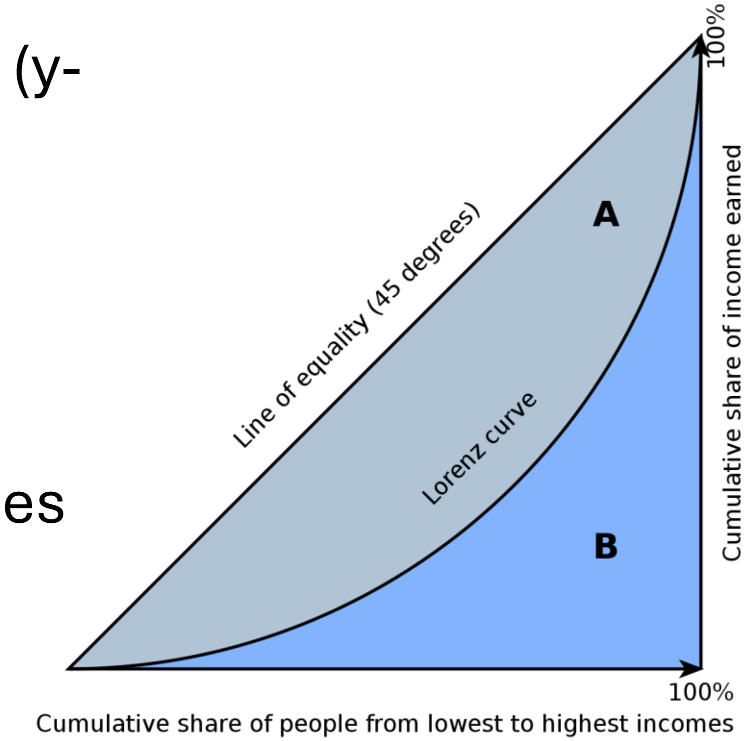


<u>Higher</u> NC
=
more decentralized

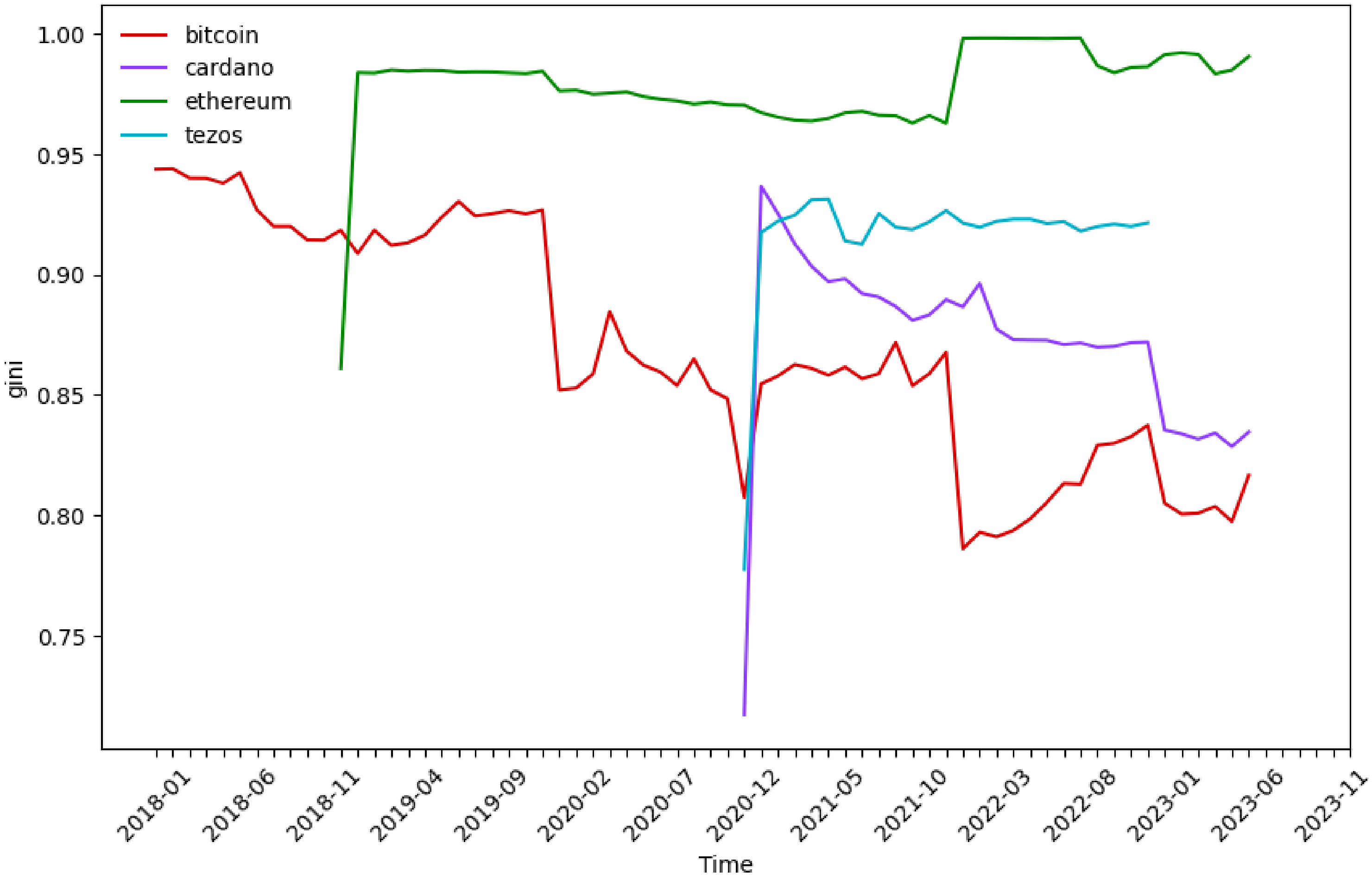


Gini coefficient

- Lorenz curve of resources across entities
  - Points: the cumulative ownership of resources (y-axis) by a percentage of entities (in ascending order) (x-axis)
- Gini coefficient: A / A + B
- Maximum equality ⇒ Gini = 0
  - Every entity holds the same amount of resources
- Maximum inequality ⇒ Gini = 1
  - One entity holds all resources


100%
Line of equality (45 degrees)
A
Lorenz curve
B
Cumulative share of income earned
100%
Cumulative share of people from lowest to highest incomes

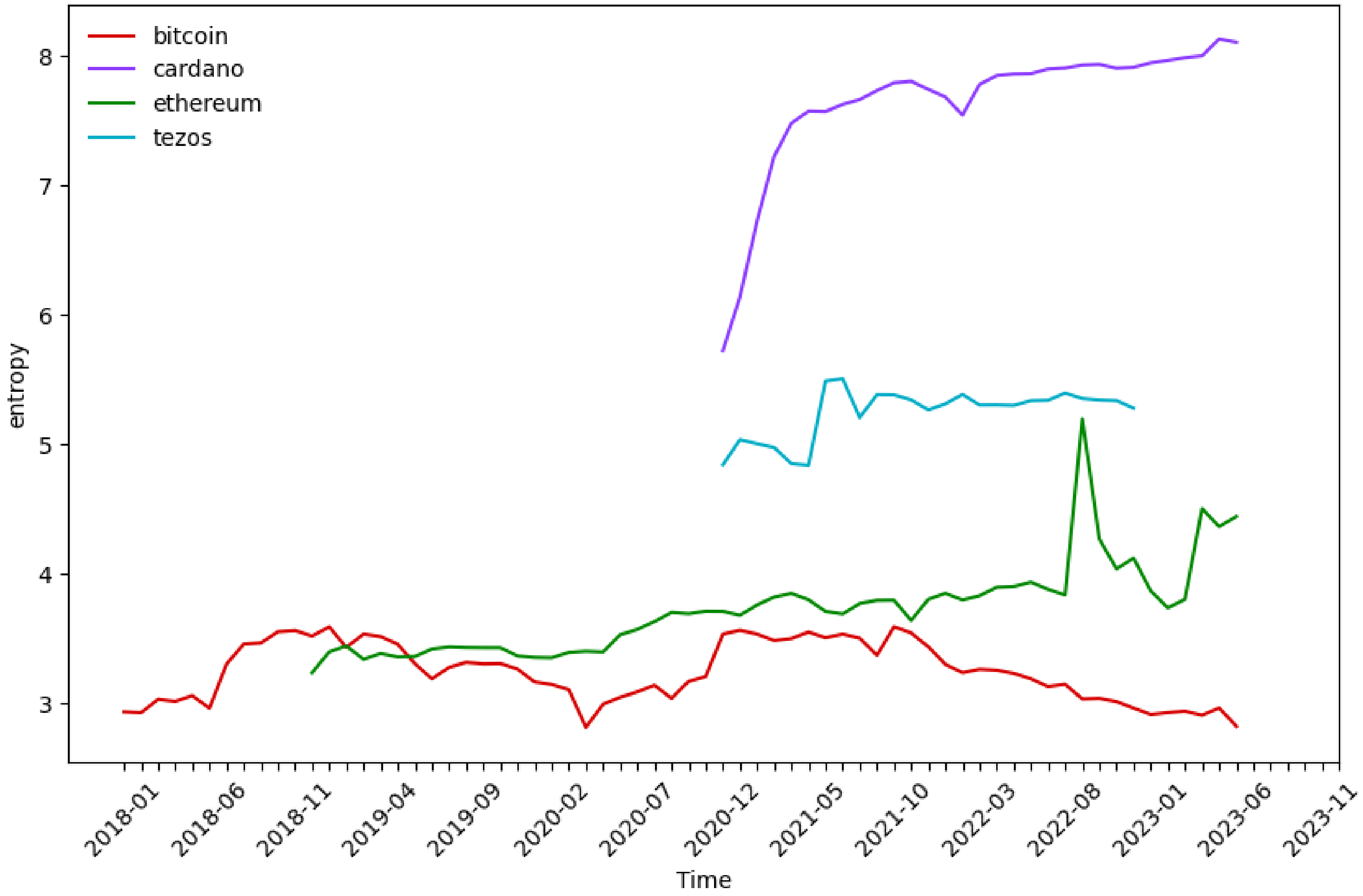
Gini coefficient



Shannon entropy

- The level (bits) of "information" in a distribution
    - $H(X) = - \Sigma (p(x_i) * \log_2 p(x_i))$ (X: random variable, p(x): probability of event X=x)
- In our case:
    - $\text{entropy}(S) = - \Sigma (f(S_i) * \log_2 f(S_i))$
    - $f(S_i)$: the relative resources of entity S_i compared to all resources (as a percentage)
- Resources centralized around a few entities ⇒ Lower entropy
- Resources distributed among many entities ⇒ Higher entropy
    - Max entropy: when resources are evenly distributed among all entities

Shannon entropy



Herfindahl–Hirschman index (HHI)

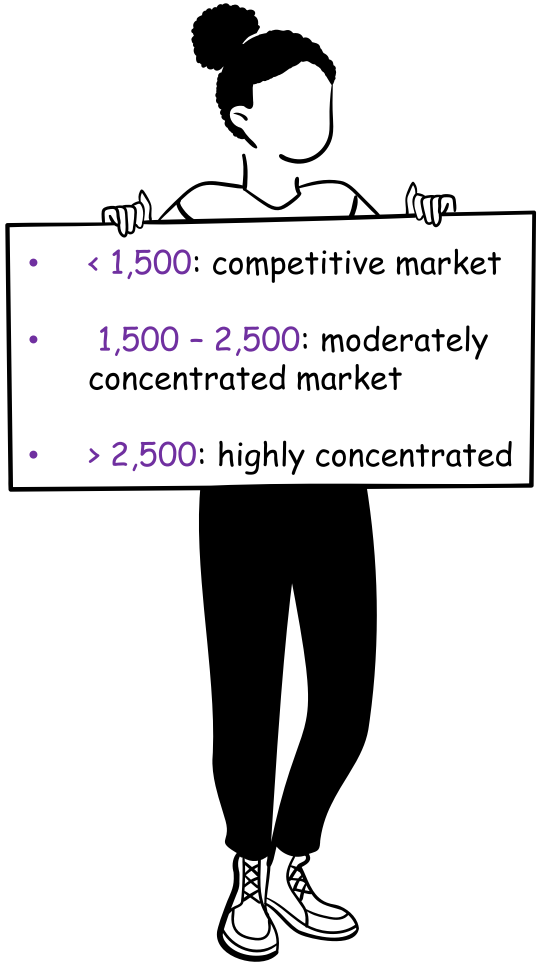
- Market concentration metric
- Can be calculated as follows:

$$HHI = \sum_{i=1}^{n} s_i^2$$

Where s_i is the market share of firm i (as a whole number, e.g. 20 for 20%)

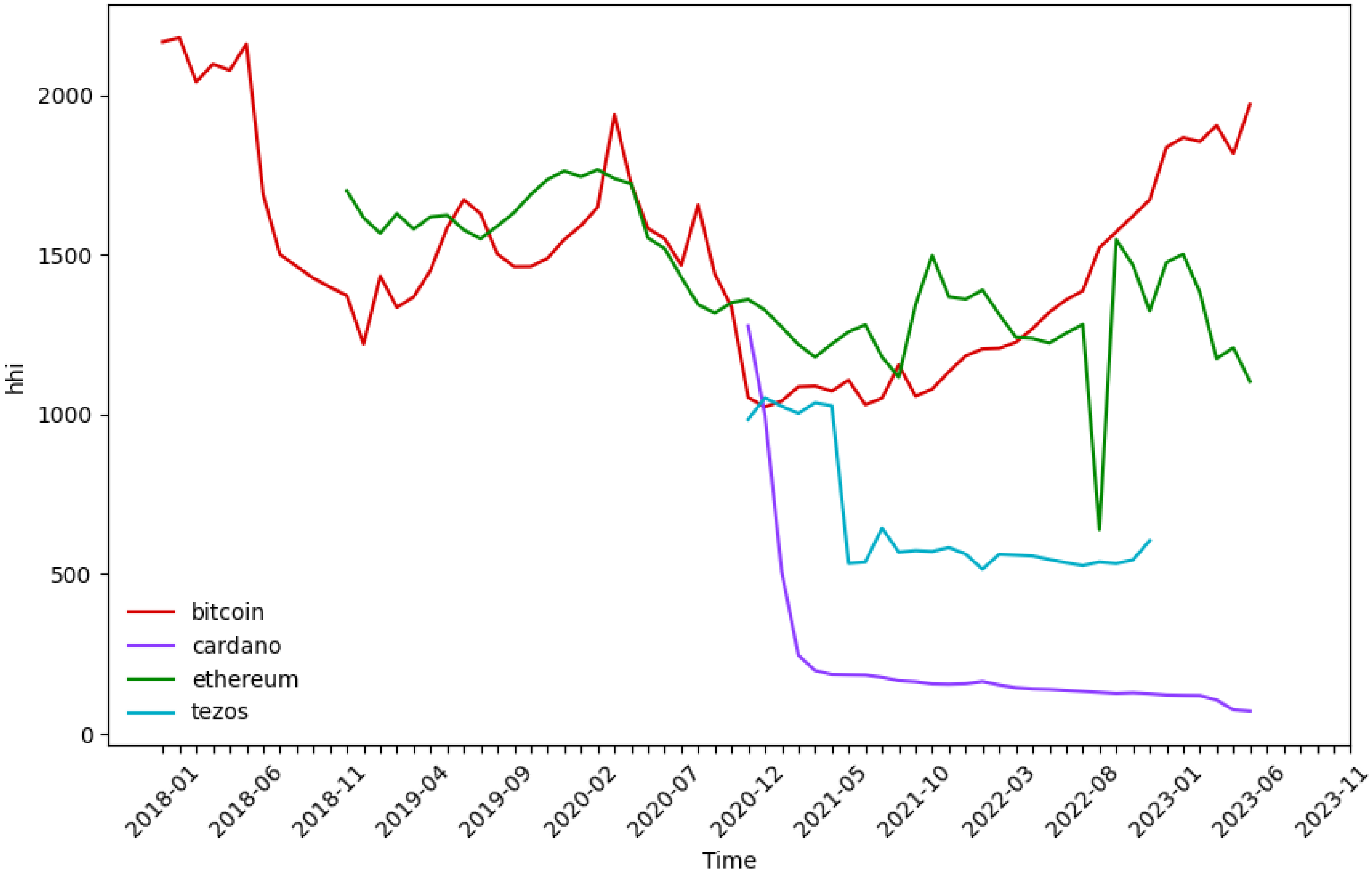
- U.S. Department of Justice guidelines:
  - HHI < 1,500: competitive market
  - 1,500 ≤ HHI ≤ 2,500: moderately concentrated market
  - HHI > 2,500: highly concentrated market
- Values chosen for traditional markets, may need different thresholds for blockchains

Herfindahl–Hirschman index (HHI)

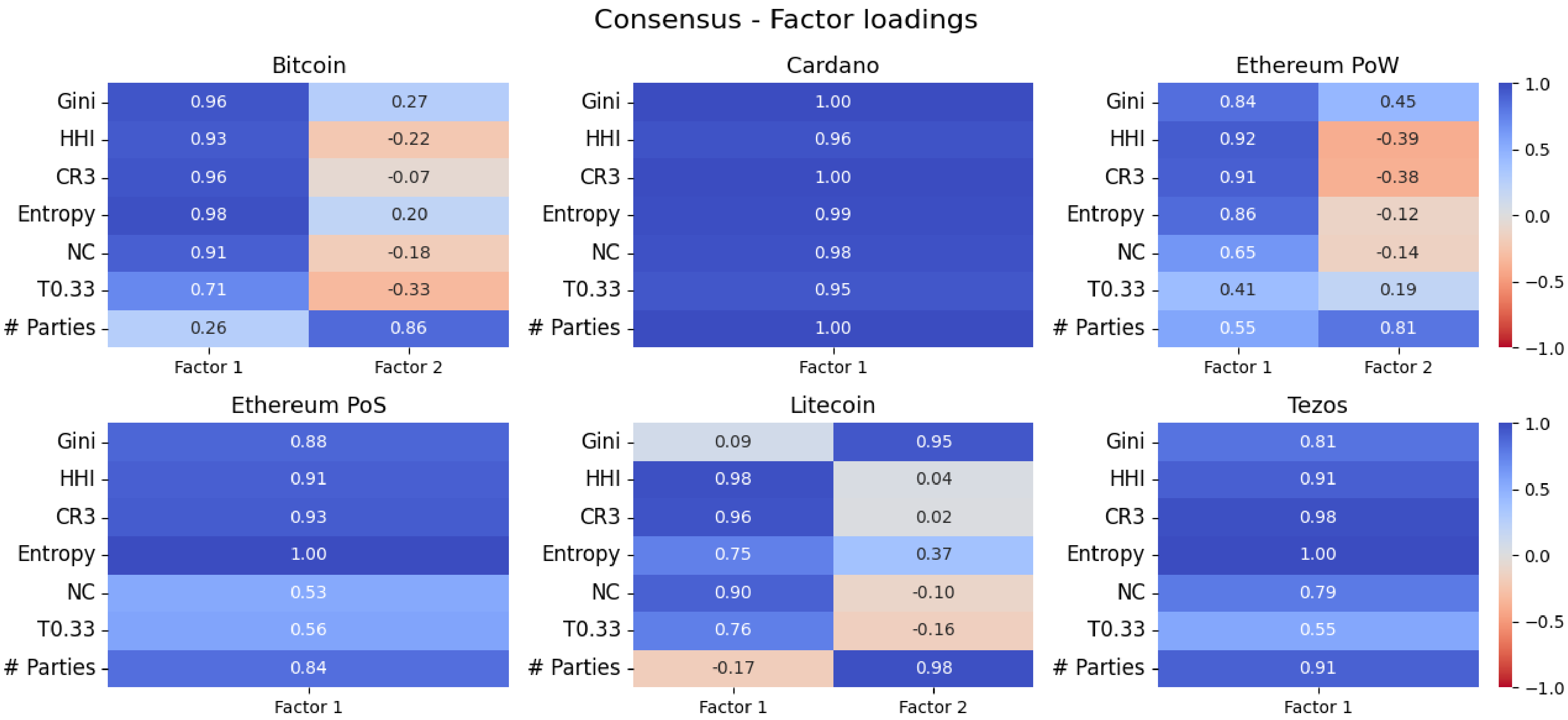


- < 1,500: competitive market
- 1,500 – 2,500: moderately concentrated market
- > 2,500: highly concentrated

U.S. Department of Justice



Do all metrics capture the same signal?



Consensus - Factor loadings

Bitcoin

	Factor 1	Factor 2
Gini	0.96	0.27
HHI	0.93	-0.22
CR3	0.96	-0.07
Entropy	0.98	0.20
NC	0.91	-0.18
T0.33	0.71	-0.33
# Parties	0.26	0.86

Cardano

	Factor 1
Gini	1.00
HHI	0.96
CR3	1.00
Entropy	0.99
NC	0.98
T0.33	0.95
# Parties	1.00

Ethereum PoW

	Factor 1	Factor 2
Gini	0.84	0.45
HHI	0.92	-0.39
CR3	0.91	-0.38
Entropy	0.86	-0.12
NC	0.65	-0.14
T0.33	0.41	0.19
# Parties	0.55	0.81

Ethereum PoS

	Factor 1
Gini	0.88
HHI	0.91
CR3	0.93
Entropy	1.00
NC	0.53
T0.33	0.56
# Parties	0.84

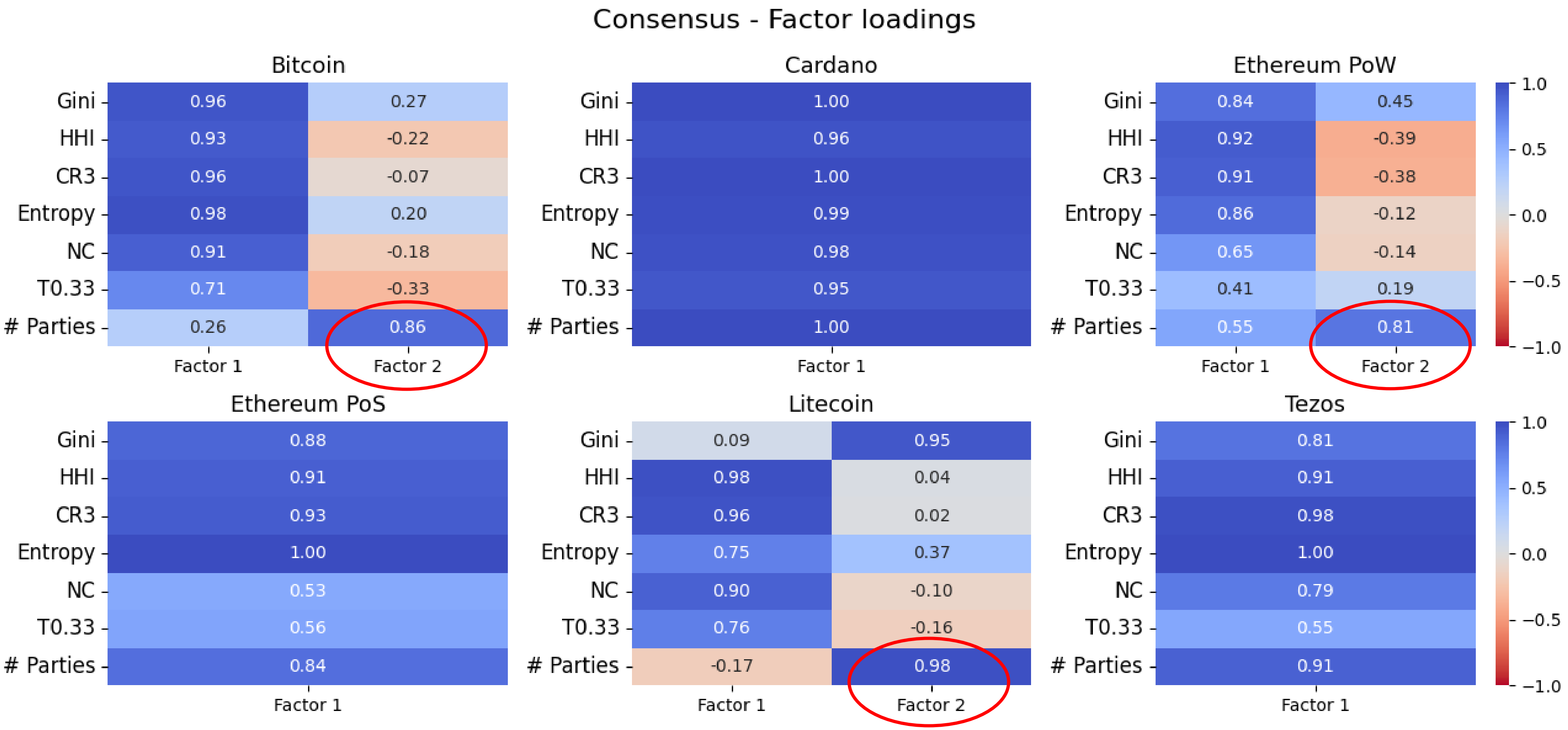
Litecoin

	Factor 1	Factor 2
Gini	0.09	0.95
HHI	0.98	0.04
CR3	0.96	0.02
Entropy	0.75	0.37
NC	0.90	-0.10
T0.33	0.76	-0.16
# Parties	-0.17	0.98

Tezos

	Factor 1
Gini	0.81
HHI	0.91
CR3	0.98
Entropy	1.00
NC	0.79
T0.33	0.55
# Parties	0.91

Do all metrics capture the same signal?



Consensus - Factor loadings

Bitcoin

	Factor 1	Factor 2
Gini	0.96	0.27
HHI	0.93	-0.22
CR3	0.96	-0.07
Entropy	0.98	0.20
NC	0.91	-0.18
T0.33	0.71	-0.33
# Parties	0.26	0.86

Cardano

	Factor 1
Gini	1.00
HHI	0.96
CR3	1.00
Entropy	0.99
NC	0.98
T0.33	0.95
# Parties	1.00

Ethereum PoW

	Factor 1	Factor 2
Gini	0.84	0.45
HHI	0.92	-0.39
CR3	0.91	-0.38
Entropy	0.86	-0.12
NC	0.65	-0.14
T0.33	0.41	0.19
# Parties	0.55	0.81

Ethereum PoS

	Factor 1
Gini	0.88
HHI	0.91
CR3	0.93
Entropy	1.00
NC	0.53
T0.33	0.56
# Parties	0.84

Litecoin

	Factor 1	Factor 2
Gini	0.09	0.95
HHI	0.98	0.04
CR3	0.96	0.02
Entropy	0.75	0.37
NC	0.90	-0.10
T0.33	0.76	-0.16
# Parties	-0.17	0.98

Tezos

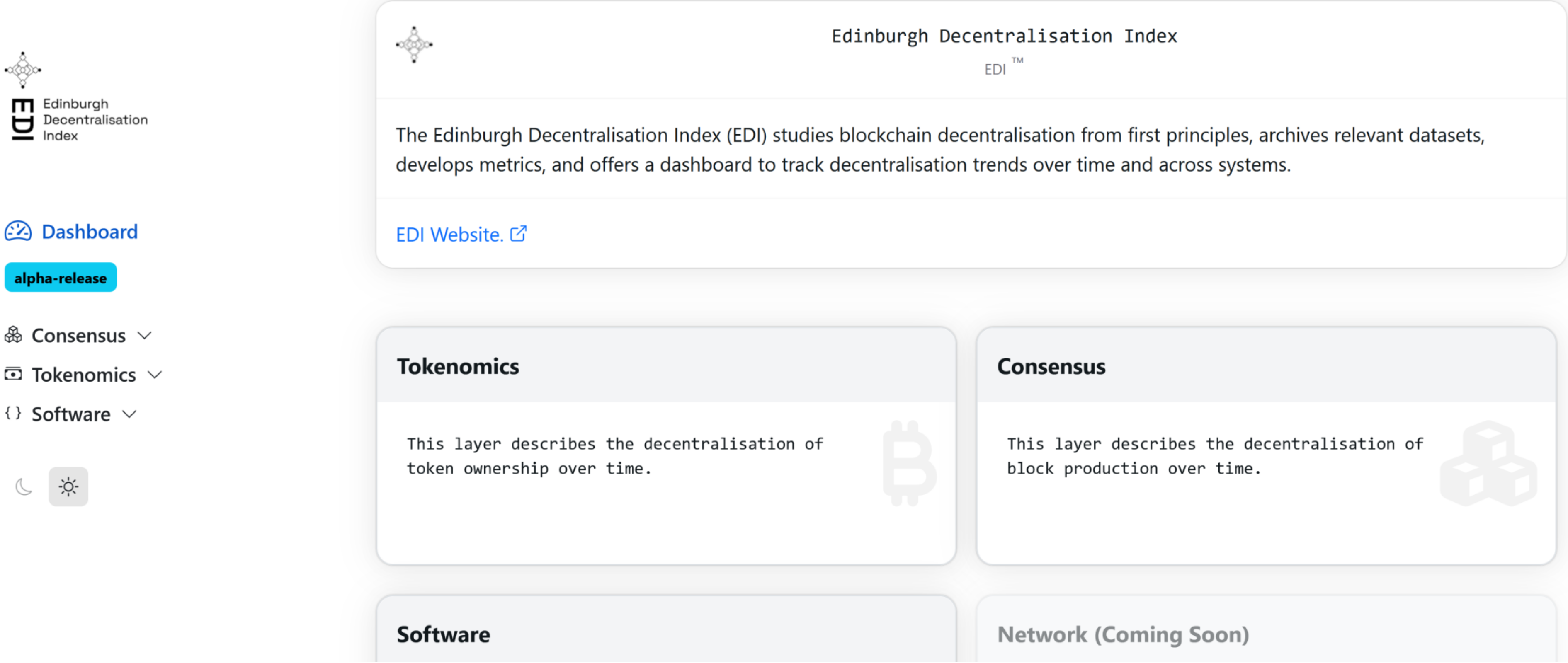
	Factor 1
Gini	0.81
HHI	0.91
CR3	0.98
Entropy	1.00
NC	0.79
T0.33	0.55
# Parties	0.91

Open questions & future work

- How do different design choices or components impact decentralization?
  - Proof of Work vs Proof of Stake
  - Simple vs sophisticated reward schemes
  - Proposer-builder separation
- Can we merge all relevant metrics and layers into a single index that provides a holistic representation of a blockchain system's decentralization?

Edinburgh Decentralization Index (EDI) Public dashboard

- Live public dashboard for multiple layers and systems: http://blockchainlab.inf.ed.ac.uk/edi-dashboard/



Edinburgh Decentralisation Index

EDI ™

The Edinburgh Decentralisation Index (EDI) studies blockchain decentralisation from first principles, archives relevant datasets, develops metrics, and offers a dashboard to track decentralisation trends over time and across systems.

EDI Website.

Tokenomics

This layer describes the decentralisation of token ownership over time.

Consensus

This layer describes the decentralisation of block production over time.

Software

Network (Coming Soon)

Thank you!

Questions?

christina.ovezik@ed.ac.uk