Towards a deeper understanding of the Cardano macro-economics

1st Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

3rd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

***Abstract*—In this paper, we conduct an in-depth analysis of the Cardano blockchain, focusing on the Proof of Stake (PoS) protocol, Ouroboros, and its implications for entity wealth and stake delegation dynamics within the ecosystem. Utilizing a heuristic-based address clustering method, previously introduced in the literature, we aggregate Cardano addresses to identify distinct entities within the network. This approach allows us to examine the wealth distribution and staking dynamics at an entity level, offering a unique perspective on the Cardano network's economic behaviors. Our study reveals insightful patterns in the emergence of new entities and their participation in stake delegation, alongside the rewards they accrue. We investigate the distribution of wealth among these entities, employing statistical measures such as the probability density and the Gini index for wealth, delegated stakes, and rewards over the blockchain's history. This analysis provides a comprehensive view of the wealth distribution within the Cardano ecosystem. Moreover, the paper explores how entities of varying wealth levels engage in stake delegation to pools of different sizes and the resultant reward distribution. We also delve into the temporal evolution of the number of pools within the network, examining how pools of various sizes are rewarded and the extent of inequality in reward distribution among them. Through this analysis, the paper sheds light on the complex interplay between wealth distribution, stake delegation, and reward mechanisms in the Cardano ecosystem, offering valuable insights into the economic and staking landscape of one of the leading blockchain networks.**

***Index Terms*—Wealth Distribution, Stake Delegation Dynamics, Entity Recognition, Blockchain Analytics**

I. Introduction

Blockchain technology has culminated in a paradigm shift, paving the way for an ecosystem that supports a decentralized economic model, reducing reliance on central authorities in managing and controlling financial transactions and fund flows [1]. The inception of this technology was marked by Bitcoin, the pioneering blockchain system, which operates on a Proof of Work (PoW) mechanism as described by Nakamoto [2]. In the PoW model, the creation of a new transaction block requires miners to successfully solve a computationally intensive puzzle. However, PoW has raised concerns due to its high computational demands, translating into significant energy consumption and environmental impact. This inefficiency in energy use is often viewed as a resource waste. Additionally, PoW has the potential to lead to centralization [3]. The increasing costs and technical demands of mining make it less accessible to individual users, deviating from the decentralized ethos of blockchain technology. To counter this, smaller miners often form mining pools, combining their resources to improve their chances of mining a block and gaining rewards [4]. However, this off-chain pooling approach, which operates outside the PoW protocol, can lead to the creation of large, dominant mining pools. Such concentration of computational power raises security concerns, including the risk of a 51% attack, where a single entity could potentially control the majority of mining power [5].

Proof of Stake (PoS) [6] presents an alternative protocol, offering a more energy-efficient and environmentally friendly approach. In PoS, the authority to validate new blocks is given to stakeholders, with the probability of being chosen as a validator being proportional to the amount of stake held. This system eliminates the need for intensive computational work, as validators are primarily responsible for verifying transactions in upcoming blocks. For its reduced energy and cost requirements, PoS is seen as a greener and more sustainable option. It also lowers the entry barrier for becoming a validator, often requiring only a specific amount of the cryptocurrency, making it potentially more accessible and promoting decentralization. The distribution of stakes within the network thus becomes a critical factor in determining both the decentralization and the voting power in the network. Recent studies have delved into aspects of stake delegation and reward distribution within PoS blockchains [7]–[9], highlighting the evolving dynamics of these systems. However, a notable concern in PoS, where wealth equates to influence, is the risk of wealth concentration, often referred to as the 'rich get richer' phenomenon [10]. This concentration can lead to an increase in voting power for wealthier participants, potentially contradicting the decentralization ethos that blockchains aim to preserve.

To ensure a more decentralized stake distribution, many PoS variants have been proposed that allow more participants to be rewarded in the staking process by delegating their assets to the validator. This Delegated Proof of Stake (DPoS) should benefit the equality of reward distribution and eventually maintain the system's security [9]. The validators first receive rewards and share them with their delegators [8]. Many protocols proclaim their implementation to be fully decentralized. However, PoS can be implemented in many different ways, and it is not clear whether the specific PoS protocol is decentralized regarding its wealth, stake, and reward distribution.

An important consideration for analysts working with blockchain data is the privacy-preserving features inherent in many public blockchains. These systems allow users to create numerous anonymous addresses, offering a degree of privacy and anonymity. Consequently, one of the primary challenges in accurately analyzing user behavior within a blockchain ecosystem is identifying and clustering addresses that are likely controlled by the same entity. This clustering is crucial because it allows for a more holistic understanding of individual user behaviors and patterns, rather than viewing each address in isolation. Recognizing these clusters is the foundational step in unraveling the complex network of transactions and interactions on the blockchain, providing insights into the economic activities and strategies of entities (users).

In this paper, we focus on the Cardano PoS protocol, called Ouroboros [11]. We first employ two heuristics tailored specifically for the Cardano blockchain to cluster the addresses owned by the same entities in the network [12]. Afterwards, we explore the following research questions to provide insights about the decentralization state in Cardano:

- How is wealth (stake) distributed among entities?
- How does this wealth distribution relate to the stake delegation patterns to different pools?
- How are staking rewards distributed among delegating entities?

Our paper introduces, for the first time to our knowledge, a pioneering staking analysis of the Cardano network at an entity level.

This paper is structured as follows. We start with a review of the existing literature on address clustering and reward distribution in blockchain networks in Section II. This is followed by an exploration of the key elements of the Cardano framework pertinent to our study in Section III. In Section IV, we detail the methodology of our research, including the specific heuristics used for address clustering, our approach to measuring inequality, and a description of the dataset utilized in our analysis. The findings of our research are then presented in Section V, followed by a discussion of the limitations of our current study and suggestions for future research in Section VI.

II. Literature Review

In this paper, we use a heuristic-based address clustering method to recognize the entities in the network. The notion that multiple input addresses of a transaction may be associated with the same entity was initially suggested by Nakamoto [2]. Building upon this assumption, Reid and Harrigan [13] later introduced the multi-input heuristic as an effective method for identifying identities. In addition, they discussed the change address heuristic, where the change address in a transaction signifies an output address belonging to the sender and receiving the excess input. Androulaki et al. [14] analyzed Bitcoin's privacy provisions by employing multi-input and change address heuristics to identify entities in the network. Meiklejohn et al. [15] assumed that change addresses would only have one input. Zhang et al. [16] discarded change addresses that were subsequently reused as non-change addresses. Liu et al. [17] developed a one-time change address identification algorithm based on multi-conditional recognition. He et al. [18] enhanced the change address heuristic and proposed a new heuristic based on the number of output addresses. Using the heuristics implemented in the BlockSci [19] C++ library as a foundation, Campajola et al. [3] devised their own heuristic-based address clustering algorithms, primarily focusing on detecting change addresses in transactions. By employing logical combinations of five heuristics, including multi-input and change address approaches, they aimed to minimize false positives. Drawing inspiration from these five heuristics, Rafati Niya et al. [20] conducted address clustering for the first time in the Cardano blockchain to identify addresses belonging to the same wallet.

In the scope of reward distribution among different PoS protocols, previous studies have proposed some general PoS consensus protocols and different reward functions on income and wealth distribution from the theoretical point of view [7], [8], [21]. Especially in 2021 Rosu and Saleh [22] showed a counter-intuitive result that the PoS protocol will not cause wealth concentration (rich getting richer) but rather relatively stable shares, by considering that the evolution of investor shares in a general PoS protocol closely parallels the evolution of colour shares in a Pólya's urn. In addition, in terms of the data analytics, Li et al. [8] analyse the fairness (incl, inequality and decentrelised) of reward and stake distribution in the four well-known PoS platforms Tezos, Polkadot, Cardano, and Casper through a data-driven approach. They found a trade-off that the PoS platforms face: the more inclusive (open) the staking process is designed, the unfairer the reward distribution may tend to be. However, all their analyses are based on the address level, without doing any heuristic address clustering to further study from the entity level. In this paper, focusing on Cardano PoS systems, we conduct comprehensive empirical studies about the stake and reward distribution, especially from the entity perspective, which could bring insight into the analysis of the system‘s macro-economics in a more realistic situation.

III. Understanding Cardano Framework

In this section, we provide a background overview of the key concepts and components in Cardano relevant to our study.

A. Cardano Consensus Protocol

The Cardano consensus protocol, Ouroboros, is a PoS protocol that enables secure and scalable distributed networks [11]. The stake is in the form of Cardano native cryptocurrency, which is ADA. Users holding stakes in Cardano can earn rewards by actively or passively participating in the block validation process. Users can participate by either establishing a new stake pool and pledging/delegating stakes to it or by delegating stakes to an existing pool. To delegate their stakes, users must register a stake address and obtain a delegation certificate for that address. A stake pool consists of a stake pool operator, one or multiple stake pool owners, and one or multiple stake delegators.

Time in the Cardano protocol is divided into epochs, each comprising 432,000 slots lasting one second. At the start of each epoch, a secret and random leader election process takes place, assigning stake pools as slot leaders for each slot. The slot leader, chosen randomly by the protocol from the registered pools, is eligible to produce the next block. Stake pools receive rewards for generating blocks and supporting the network, with the rewards distributed proportionally to the amount of stake pledged or delegated to the pool. These rewards are distributed among the pool operator, owners, and delegators. However, the rewards are not immediately usable and need to be withdrawn through transactions making them spendable.

B. Cardano Address Types

Cardano addresses can be divided into two primary categories [23]. The first category comprises Payment Addresses, which are the only addresses capable of owning and spending transactional funds. Payment addresses can further be classified as either Shelley addresses or Byron addresses. A Shelley address consists of two main components: the payment part, which refers to a payment key, a simple script, or a Plutus smart contract that possesses funds stored in the address, and the delegation part, which may refer to a stake address holding stake rights for the funds stored in the Shelley address. Conversely, Byron addresses are an older address format maintained solely for backward compatibility. The second category consists of Stake Addresses, which possess and control staking rewards and are associated with a stake key.

IV. Methodology and Implementation

This section details our approach to analyzing the Cardano blockchain, from clustering addresses to measure wealth distribution, to the data collection and computational processes employed.

A. Address Clustering Heuristics

In this study, we employ two following heuristics that are specifically designed for the Cardano blockchain [12].

- **Modified Multi-Input Heuristic (Heuristic 1)**: This heuristic assumes that if, within the inputs list of a single transaction, there are Byron payment addresses and Shelley payment addresses with payment parts referring to payment keys (rather than simple scripts or Plutus smart contracts), they are considered to belong to the same entity. This assumption is derived from the behavior observed in widely utilized Cardano wallets such as Daedalus and Yoroi, which restrict transactions to spending from wallets controlled by a single user.
- **Staking Heuristic (Heuristic 2)**: This heuristic assumes that if certain Shelley payment addresses delegate their stake rights to the same stake key, indicated by their delegation part referencing the same stake key, then they are considered to belong to the same entity. The rationale behind this assumption is that the entity in possession of the private key associated with a stake key has complete control over withdrawing the staking rewards stored in the corresponding stake address and transferring them to any desired payment address.

TABLE I: Dataset Characteristics

Blocks Count	Transactions Count	Start Timestamp	End Timestamp
8,299,565	59,508,571	2017-09-23 21:44:51	2023-01-21 17:39:30

B. Gini Index

The Gini index is a common inequality index for income or wealth distribution used to measure financial inequality among a nation's residents [24]. It can theoretically range from 0 (complete equality) to 1 (complete inequality, i.e. one participant has everything, the rest have nothing). The Gini index has been applied several times to blockchain-based systems, e.g. in [3]. In [4], the authors use the Gini index to measure the inequality of mining revenue distribution among miners in several PoW-based systems. Similarly, in the context of PoS protocols, equality refers to the distribution of wealth, stakes, and rewards of the parties in the system. Therefore, we apply the Gini index to the wealth, stakes, and rewards of the parties in the network, i.e.,

$$G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j|}{2n\sum_{i=1}^{n}x_i}, \tag{1}$$

where x_i is either the wealth w_i, stake s_i, or the reward r_i of parties i, and there are n parties in total.

C. Data Collection

To parse and format the Cardano blockchain data, we established a Cardano node integrated with a component known as Cardano DBsync. This DBsync serves as the default indexer for the node, efficiently collecting and storing blockchain data in a PostgreSQL database [25]. After setting the node up, by joining three tables tx, tx_out, and tx_in from the database, the historical transaction data of the blockchain was extracted and used for address clustering. Subsequently, we queried three tables epoch_stake, reward, pool_hash, and stake_address to gather information on staking and rewards, details about the pools, and data on stake addresses delegating to various pools. Our dataset includes all historical data stored on the Cardano blockchain until January 2023. The characteristics of our dataset are summarized in Table I. All our analyses were implemented using Python scripts executed on a Debian 11 machine with 128 AMD EPYC 3.40 GHz CPU cores and 512 GB memory.

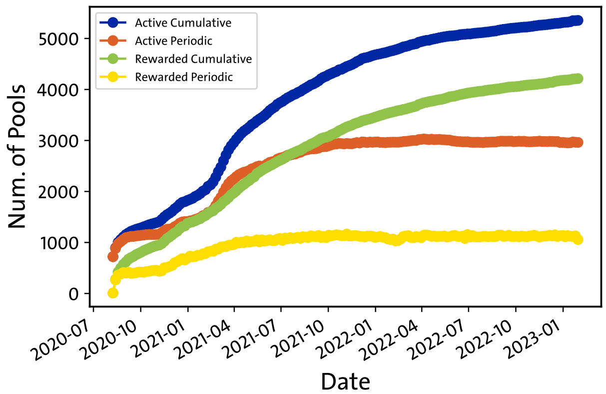


Fig. 1: Periodic and Cumulative number of active and rewarded pools across all epochs

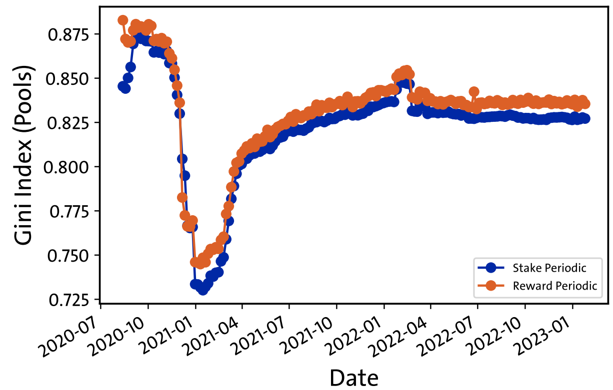


Fig. 2: Gini index of received stakes by active pools and rewards by rewarded pools across all epochs

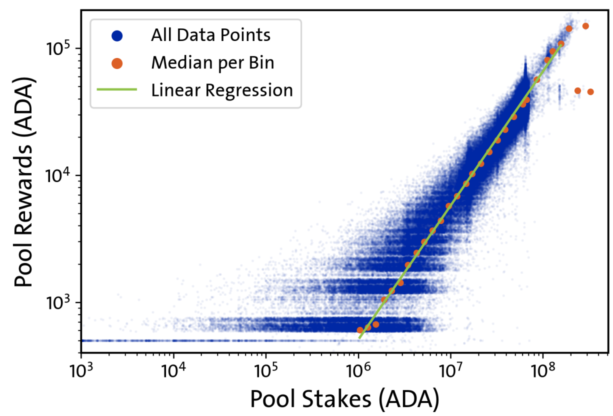


Fig. 3: Received stakes and rewards by pools across all epochs

V. Results

In this section, we present a detailed analysis of the Cardano blockchain, focusing on the dynamics of stake pools and entities, and their interactions. Our findings are supported by a series of figures that shed light on various aspects of the blockchain's evolution. We explore the growth and behavior of stake pools, examining the number of pools over time, their stake and reward distributions, and the correlation between stakes and rewards. The dynamics of entities are also evaluated, where we look at the growth in the number of entities, their wealth distribution, staking behavior, and the rewards they receive. Lastly, we explore the interactions between stake pools and entities, analyzing how the pool's total stakes and the number of pools receiving delegations vary among entities with different wealth levels. This comprehensive examination provides a deeper understanding of the complex ecosystem within the Cardano blockchain.

A. Dynamics of Stake Pools

We examine the evolving landscape of stake pools within the Cardano network. Our analysis starts with tracing the historical growth in the number of stake pools, providing insights into how the network has expanded over time. We also assess the financial aspects of these pools, evaluating the Gini index of the stakes received and rewards earned by the pools across all epochs. Furthermore, our investigation explores the relationship between the stakes delegated to pools and the rewards they receive, revealing crucial patterns and correlations that define the network's staking dynamics.

Fig. 1 depicts the periodic and cumulative number of active pools in the Cardano network that have been active or rewarded in each epoch. We call a pool active in an epoch if it has received delegation stakes from entities during the epoch. This plot shows that by the end of our study period, January 2023, 5,354 distinct registered pools (with distinct pool hash) have received delegation in the network. This number is 4,214 for rewarded pools.

Fig. 2 shows the Gini index of stakes and rewards received by active and rewarded pools in each epoch. This plot reveals a noticeable decrease in the distribution inequality of both delegated stakes and rewards within pools in December 2020 (epoch 234), a trend that might be related to the change of the saturation threshold parameter. Subsequently, although the Gini index begins to rise once more, it does not return to its pre-change level, indicating the lasting significance of the change.

Fig. 3 shows the relation between the amount of stakes delegated to a pool and the staking rewards it receives. Each data point in this figure represents a pool's stakes and rewards in an epoch. Additionally, we applied a logarithmic binning to the pool stakes amount in this figure and then computed the median of the rewards in each bin. This reveals an exponential relation between the two variables. We also employed linear regression to establish a linear relationship on a logarithmic scale for the median values. The resulting line on this log-log plot is characterized by a slope of 1.056 and an intercept of -3.636. This implies that within the analyzed range, we have $log(PoolReward) = 1.056 * log(PoolStake) - 3.636$.

B. Dynamics of Entities

Here we focus on the entities participating in the Cardano ecosystem. We provide a comprehensive view of the cumu-

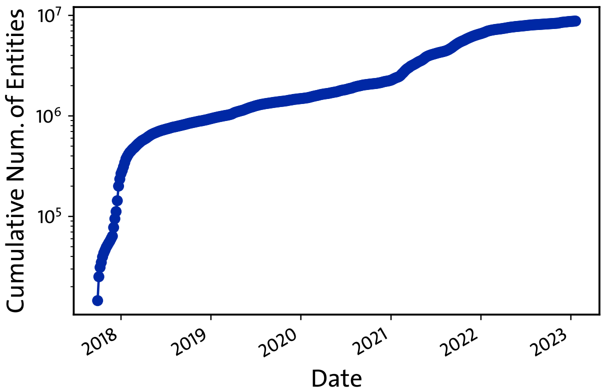


Fig. 4: Cumulative number of entities across all epochs

lative growth of these entities over time. Our analysis further extends to the probability density of wealth, delegated stakes, and received rewards, offering a deep dive into the financial characteristics of these entities. Additionally, we explore the Gini index for wealth, stakes, and rewards, shedding light on the inequality within the network. Moreover, our examination includes an analysis of the number of delegating and rewarded entities, as well as the distribution of rewards among entities with different wealth levels.

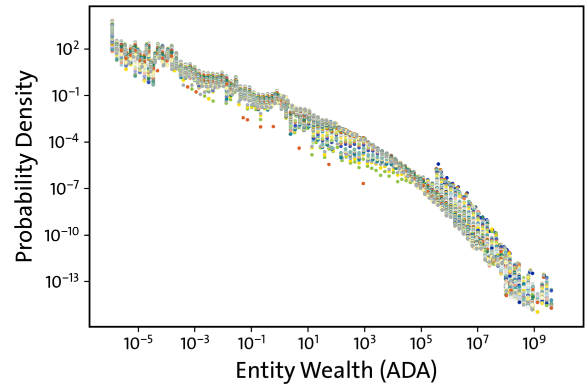
Fig. 4 displays the cumulative count of entities appearing as transaction outputs in each epoch.

Fig. 5 presents the probability density of the entities' wealth, delegated stakes, and received rewards. Each color in the plot represents the probability density across a different epoch. The wealth, stakes, and rewards of entities are calculated at the beginning of each epoch.

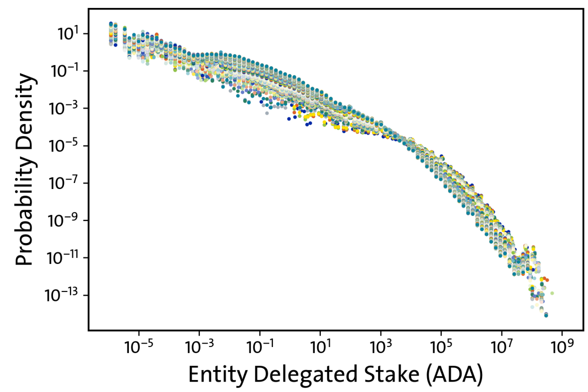
In Fig. 6, the Gini index of the entities' wealth, delegated stakes, and received rewards in each epoch is illustrated. It reveals that after the launch of the Shelley mainnet upgrade in July 2020, when Cardano enabled the stake delegation mechanism, there has been a steady increase in the inequality of stake and reward distribution among entities, as indicated by the rising Gini index. By September 2021, this index had largely stabilized, remaining above 0.97 until January 2023.

Fig. 7 shows the number of entities that have delegated stakes to the stake pools as well as the entities that have received rewards for their contribution to the staking mechanism in each epoch. Except for small drops in six epochs in 2021 and 2022, these numbers have typically been on the rise.

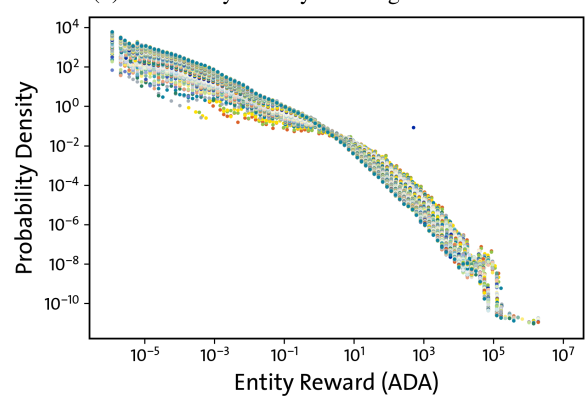
Fig. 8 illustrates the relation between the wealth of an entity and the staking rewards it receives. Each data point in this figure represents an entity's wealth and the total rewards it receives in an epoch. This plot aggregates rewards for each entity across all its stake addresses per epoch, acknowledging that entities may own multiple addresses. Additionally, we applied a logarithmic binning to the entity wealth in this figure and then computed the median of the rewards in each bin. It appears that the two variables are related exponentially. We also applied linear regression to model the relationship between the logarithms of the median values. On the log-log scale, the regression line is defined by a slope of 1.0019 and an intercept of -3.241. This implies that for the range examined, we have $log(EntityReward) = 1.0019 * log(EntityWealth) - 3.241$.



(a) Probability density of wealth



(b) Probability density of delegated stakes



(c) Probability density of received rewards

Fig. 5: Probability density of wealth, delegated stake, and received reward for entities across all epochs

C. Interaction of Stake Pools and Entities

We explore the relationship between the stake pools and the entities that interact with them. This part of the analysis includes examining the ratio of active and rewarded pools to the number of entities, which provides insights into the level of engagement and reward distribution in the network.

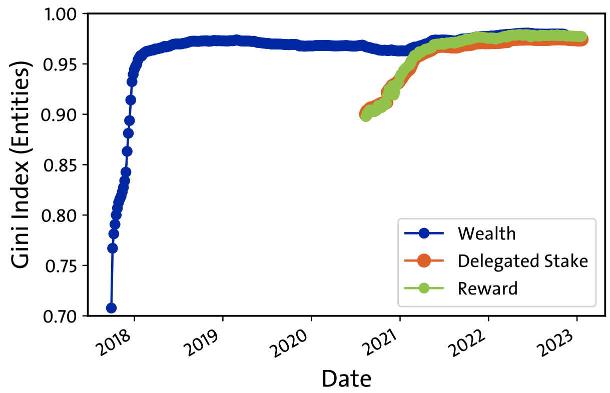


Fig. 6: Gini index of wealth, delegated stake, and received reward for entities across all epochs

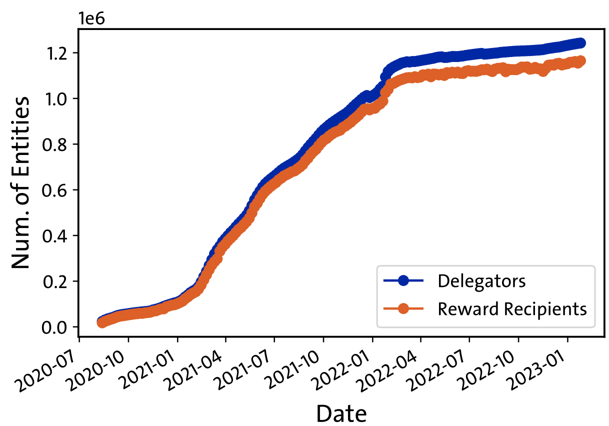


Fig. 7: Number of delegator and rewarded entities across all epochs

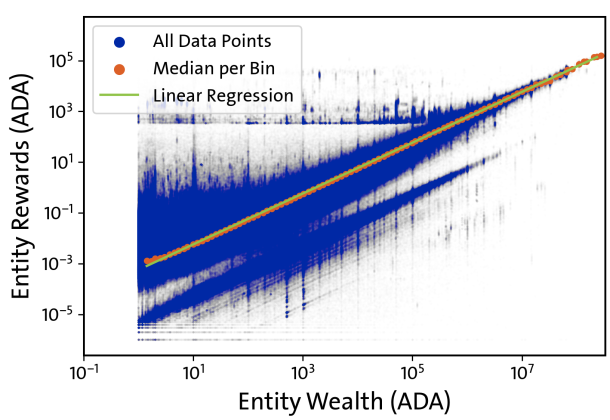


Fig. 8: Rewards received by entities with varying wealth levels across all epochs

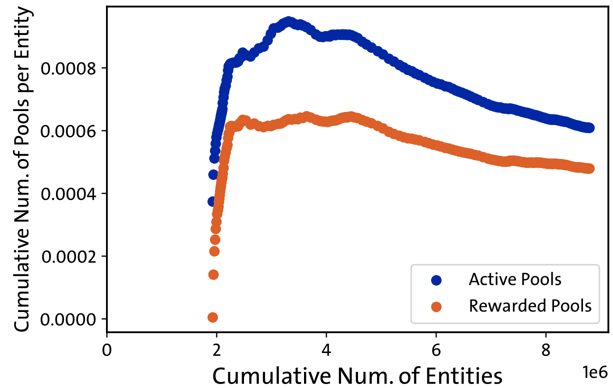


Fig. 9: Number of active and rewarded pools divided by the number of entities across all epochs

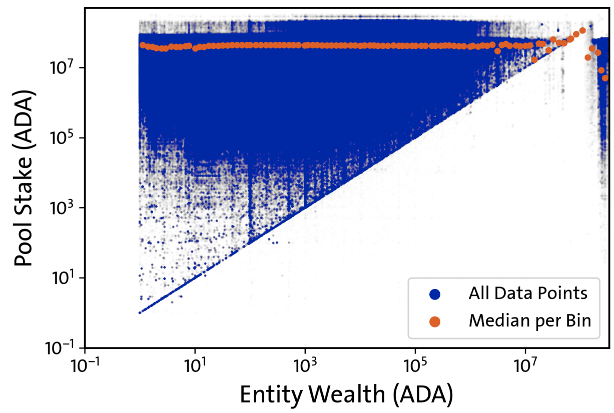


Fig. 10: Total stakes of pools received delegation from entities with varying wealth levels across all epochs

We also look at the total stakes of pools that have received delegations from entities of varying wealth levels, as well as the number of pools that each wealth group delegates to. This analysis helps in understanding how the behavior of entities with different financial standings influences delegation patterns in the Cardano ecosystem.

Figure 9 separates active from rewarded pools, and illustrates the ratio of cumulative pool count to cumulative entity count in each epoch, whereas the cumulative number of entities is subject to change. It shows a peak in May 2021, when the total number of entities in the network reached 3,480,397. In this figure, the x-axis represents a cumulative number of entities that have received funds in the blockchain according to the transaction history.

The scatter plot in Fig. 10 depicts the relation between the wealth of delegating entities and the total stakes of the pools to which these entities have delegated. Each data point in this figure represents an entity's wealth and the total stakes of the pool to which the entity has delegated in an epoch. Additionally, we applied a logarithmic binning to the entities' wealth in this figure and then computed the median of the pool stakes in each bin. The figure shows that entities with

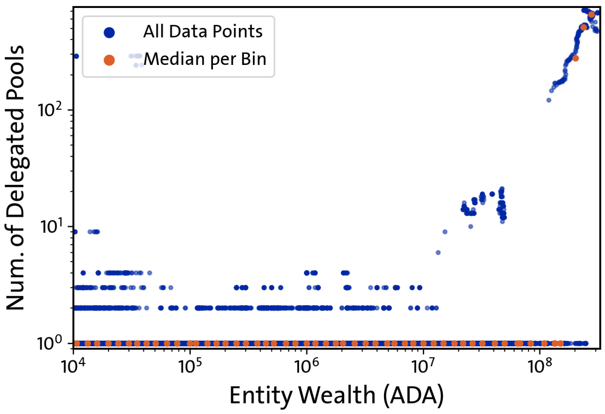


Fig. 11: Number of pools received delegation from entities with varying wealth levels across all epochs

different amounts of wealth are more likely to put their stakes in pools that already have a larger amount of stakes. This trend indicates a common preference for choosing pools with higher stakes, regardless of the entity's own wealth level. Here, we can observe that the entities wealthier than 10^8 ADA have typically delegated to the pools with fewer stakes. This may be linked to the stake pools' saturation threshold, a global parameter that defines the maximum capacity for stake pools to accept delegations while still yielding rewards. The saturation threshold for a stake pool in the Cardano network is calculated using the formula S_t/K, where S_t represents the total supply of ADA in circulation, and K is a parameter used to control the desired number of stake pools, which was set to 500 starting from epoch 234. During our analysis period, the total ADA supply in circulation was approximately 32 billion. It means that the saturation threshold is $6.4 * 10^7$, which is consistent with our observation of the delegators' staking behavior reflected in Fig. 10.

Fig. 11 shows the number of the pools to which the entities with varying wealth have delegated. Each data point in this figure represents the wealth of an entity and the number of pools it has delegated some stakes in an epoch. We again applied a logarithmic binning to the entities' wealth in this figure and then computed the median of the pool stakes count in each bin. The findings indicate a tendency among entities in the Cardano network to delegate their stakes predominantly to a limited number of pools, often choosing to delegate to only one pool. This pattern suggests a preference for concentrated delegation strategies within the network. However, the wealthiest entities have chosen multiple pools for staking. This observation that wealthier entities are diversifying their stakes across multiple pools can be partially attributed to the saturation threshold of the stake pools in the Cardano network. By spreading their holdings across various pools, these entities likely aim to maximize rewards while avoiding the diminishing returns that occur when a single pool reaches its saturation point.

VI. Limitations and Future Work

In this study, we utilized heuristic-based methods to cluster addresses and identify entities. While effective, this approach has limitations, notably the potential for false positives. Identifying and rectifying these inaccuracies is crucial for improving the reliability of our results. Future research should concentrate on this aspect, employing advanced techniques to enhance entity recognition. Specifically, applying community detection methods and analyzing mixed transactions from various users will be instrumental in reducing these false positives. These strategies will not only address current limitations but also significantly advance the accuracy and efficacy of entity recognition in blockchain analysis.

Conclusion

This paper presents an in-depth analysis of wealth distribution and staking dynamics within the Cardano blockchain. Our initial approach involved identifying entities by clustering Cardano addresses using a heuristic-based method. This set the foundation for investigating several key research questions. We explored the patterns of new entities joining the network, their contributions to stake delegation, and the rewards they earned over time. A significant part of our analysis focused on the distribution of funds among these entities. This included calculating the probability density of wealth, delegated stakes, and rewards, along with assessing the Gini index for these metrics across the blockchain's history. Moreover, our study explored the delegation behaviors of entities with varying wealth levels, examining how they allocate stakes to pools of different sizes (with different amounts of stakes delegated to them) and the consequent rewards they receive. We also studied the temporal evolution of the number of pools in the network, the reward distribution across pools of various sizes, and the extent of inequality in reward distribution among these pools. Collectively, these analyses offer valuable insights into the economic and staking landscape of the Cardano blockchain.

References

[1] D. Garcia, C. J. Tessone, P. Mavrodiev, and N. Perony, "The digital traces of bubbles: feedback cycles between socio-economic signals in the bitcoin economy," *Journal of the Royal Society Interface*, vol. 11, no. 99, p. 20140623, 2014.

[2] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Decentralized business review*, 2008.

[3] C. Campajola, R. Cristodaro, F. M. De Collibus, T. Yan, N. Vallarano, and C. J. Tessone, "The evolution of centralisation on cryptocurrency platforms," *arXiv preprint arXiv:2206.05081*, 2022.

[4] S.-N. Li, Z. Yang, and C. J. Tessone, "Proof-of-work cryptocurrency mining: a statistical approach to fairness," in *2020 IEEE/CIC international conference on communications in China (ICCC workshops)*, pp. 156–161, IEEE, 2020.

[5] G. O. Karame, E. Androulaki, and S. Capkun, "Double-spending fast payments in bitcoin," in *Proceedings of the 2012 ACM conference on Computer and communications security*, pp. 906–917, 2012.

[6] C. T. Nguyen, D. T. Hoang, D. N. Nguyen, D. Niyato, H. T. Nguyen, and E. Dutkiewicz, "Proof-of-stake consensus mechanisms for future blockchain networks: fundamentals, applications and opportunities," *IEEE access*, vol. 7, pp. 85727–85745, 2019.

[7] N. Dimitri, "Monetary dynamics with proof of stake," *Frontiers in Blockchain*, vol. 4, p. 443966, 2021.

[8] S.-N. Li, F. Spychiger, and C. J. Tessone, "Reward distribution in proof-of-stake protocols: A trade-off between inclusion and fairness," *IEEE Access*, vol. 11, pp. 134136–134145, 2023.

[9] S. M. S. Saad and R. Z. R. M. Radzi, "Comparative review of the blockchain consensus algorithm between proof of stake (pos) and delegated proof of stake (dpos)," *International Journal of Innovative Computing*, vol. 10, no. 2, 2020.

[10] G. Fanti, L. Kogan, S. Oh, K. Ruan, P. Viswanath, and G. Wang, "Compounding of wealth in proof-of-stake cryptocurrencies," in *Financial Cryptography and Data Security: 23rd International Conference, FC 2019, Frigate Bay, St. Kitts and Nevis, February 18–22, 2019, Revised Selected Papers 23*, pp. 42–61, Springer, 2019.

[11] A. Kiayias, A. Russell, B. David, and R. Oliynykov, "Ouroboros: A provably secure proof-of-stake blockchain protocol," in *Annual international cryptology conference*, pp. 357–388, Springer, 2017.

[12] M. Chegenizadeh, S. Rafati Niya, and C. J. Tessone, "Heuristic-based address clustering in cardano blockchain," in *Proceedings of the 1st ChainScience Conference in Boston, MA, USA*, 2023.

[13] F. Reid and M. Harrigan, *An analysis of anonymity in the bitcoin system*. Springer, 2013.

[14] E. Androulaki, G. O. Karame, M. Roeschlin, T. Scherer, and S. Capkun, "Evaluating user privacy in bitcoin," in *Financial Cryptography and Data Security: 17th International Conference, FC 2013, Okinawa, Japan, April 1-5, 2013, Revised Selected Papers 17*, pp. 34–51, Springer, 2013.

[15] S. Meiklejohn, M. Pomarole, G. Jordan, K. Levchenko, D. McCoy, G. M. Voelker, and S. Savage, "A fistful of bitcoins: characterizing payments among men with no names," in *Proceedings of the 2013 conference on Internet measurement conference*, pp. 127–140, 2013.

[16] Y. Zhang, J. Wang, and J. Luo, "Heuristic-based address clustering in bitcoin," *IEEE Access*, vol. 8, pp. 210582–210591, 2020.

[17] F. Liu, Z. Li, K. Jia, P. Xiang, A. Zhou, J. Qi, and Z. Li, "Bitcoin address clustering based on change address improvement," *IEEE Transactions on Computational Social Systems*, 2023.

[18] X. He, K. He, S. Lin, J. Yang, and H. Mao, "Bitcoin address clustering method based on multiple heuristic conditions," *IET Blockchain*, vol. 2, no. 2, pp. 44–56, 2022.

[19] H. Kalodner, M. Möser, K. Lee, S. Goldfeder, M. Plattner, A. Chator, and A. Narayanan, "Blocksci: Design and applications of a blockchain analysis platform," in *29th USENIX Security Symposium*, 2020.

[20] S. R. Niya, I. Mesić, G. Anagnostou, G. Brunini, and C. J. Tessone, "A first analytics approach to cardano," in *2023 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, pp. 1–5, IEEE, 2023.

[21] Y. Wang, G. Yang, A. Bracciali, H.-f. Leung, H. Tian, L. Ke, and X. Yu, "Incentive compatible and anti-compounding of wealth in proof-of-stake," *Information Sciences*, vol. 530, pp. 85–94, 2020.

[22] I. Roşu and F. Saleh, "Evolution of shares in a proof-of-stake cryptocurrency," *Management Science*, vol. 67, no. 2, pp. 661–672, 2021.

[23] "Cip 19 - cardano addresses." https://cips.cardano.org/cips/cip19.

[24] J. Morgan, "The anatomy of income distribution," *The review of economics and statistics*, vol. 44, pp. 270–283, 1962.

[25] IOHK, "Schema documentation for cardano-db-sync." https://github.com/input-output-hk/cardano-db-sync/blob/master/doc/schema.md.