Improving Proof of Stake Economic Security via MEV Redistribution

Tarun Chitra
tarun@gauntlet.network

Kshitij Kulkarni
ksk@eecs.berkeley.edu

August 2022

Abstract

Maximal Extractable Value (MEV) has generally been viewed as a negative, parasitic aspect of economic transactions on blockchains that increases costs for non-strategic users. Recent work has shown that MEV is not *always* bad for social welfare in crypto networks. In this note, we demonstrate that if rational validators in Proof of Stake (PoS) protocols are able to earn a portion of MEV revenue, by a process we call MEV redistribution, they are disincentivized to unstake and lower economic security. We construct a joint staking-lending dynamical system in which a fraction of MEV revenue is used to increase staking returns. We formally show that this MEV redistribution can avoid bad competitive equilibria between staking and lending in which no users stake under benign conditions on the reward inflation schedule of the protocol, and conduct numerical simulations that demonstrate this. This represents another potentially positive externality of MEV, provided that the mechanism for redistribution is well-designed.

1 Introduction

As blockchain applications become increasingly sophisticated, strategic users are able to profit off of non-strategic users by reordering, adding, or censoring tranasctions. This involves strategic users colluding with miners to ensure that particular transaction-level covenants are enforced. An example of such covenants includes front-running transactions, where a strategic user places a transaction in front of a non-strategic user by paying a miner for this queue position. Note that virtually all blockchain consensus protocols only enforce agreement at the block level, allowing transaction processors (*e.g.* miners or validators) to choose which transactions to include and what order they are executed in. This opens the door for strategic users to pay miners via off-chain mechanisms to guarantee transaction inclusion or ordering — any such strategy is called Maximal Extractable Value (MEV).

The total amount of MEV extractable depends on the types of applications that have large

numbers of pending transactions sent to them. However, the most popular form of MEV involves decentralized finance (DeFi) applications, which allow users to trustlessly trade, borrow, and lend digital assets. Most of the time, non-strategic users submit transactions to the public pending transaction pool while strategic users search through possible ways to insert and reorder non-strategic users' transactions to maximize strategic user profit. These types of attacks exist outside of DeFi, however, and also impact non-fungible token mints. In the worst-case scenario, MEV can cause consensus instabilities [6, 19, 18] and be fatal to a decentralized network.

Flashbots Auctions. While MEV has existed since 2019, initially in the form of probabilistic gas auctions [6], it was made systematic by the introduction of Flashbots [17]. Probabilistic gas auctions involves users submitting transactions whose transaction fee was just above (or below) a transaction that the strategic user intended to be executed in front of (or behind of, respectively). These auctions had a negative externality that arose due to strategic users spamming the network in order to guarantee that their transaction was placed in a specific ordering. For instance, if a strategic user wanted to front-run a non-strategic user, they would spam the network to ensure that their transaction was executed immediately before the non-strategic user.

Flashbots introduced an auction mechanism to ensure that transactions only need to be submitted for consensus once, thus reducing spam to full nodes. Flashbots runs a centralized, off-chain auction that allows strategic users to bid in an auction for the placement of *transaction bundles*, which are ordered seqeunces of valid Ethereum transactions. Strategic users bid for placement and upon successfully being auctioned a slot in the next block of transactions pay a fee equal to the intrinsic execution cost of the bundle plus the auction bid. Although other competing auctions exist, Flashbots dominates Ethereum market share with 75-85% of hash power agreeing to Flashbot's auction ordering for transaction bundles [13].

Proof of Stake. While MEV currently exists on Proof-of-Work (PoW) Ethereum, it is expected to exist and increase in magnitude when Ethereum makes the transition to Proof-of-Stake (PoS) [11]. One major change that has been proposed for proof-of-stake Ethereum is the proposer-builder separation (PBS) [2]. This change separates out two tasks that are crucial for MEV extraction. A *builder* creates an ordered list of transactions to be executed in the next block. The builder submits this set of transactions as well as bid to a *proposer*. The proposer does not learn the content of the set of transactions — just the bid and a proof that the builder executed the set of transactions correctly. Once a proposer selects a set of transactions and a bid to accept, the proposer submits a built block (which includes their signatures).

If the builder market is competitive, even if there is a builder who submits a bundle that censors a particular transaction, they will pay extra. Heuristic analysis [2] intimates that regardless of the valuation of the bundle of transactions by either a censoring builder or

an honest builder, there is a 'tax' placed for censoring. Another benefit of PBS is that is allows for so-called MEV redistribution [11]. This involves allowing for the quantity of MEV extracted (which can be measured by the builder) to be distributed partially to all validators and partially to the particular builder for that block. By splitting MEV extracted between the entire network and the individual miner, MEV can be distributed to the entire Ethereum network, thus reducing its net negative externality provided that it is easy to stake Ethereum.

Currently, PoW Ethereum has a single miner do both tasks. Without instruments such as hash power derivatives, it is impossible to securitize a payment that goes to all miners proportional to the hash power they have committed [3]. On the other hand, a PoS stake distribution is publicly known and pro-rata distributions can take place. The benefit of PBS is that the censorship resistant properties of PBS make it expensive for a block builder within a competitive market to try to reduce the network's MEV redistribution payment. That said, while this is more expensive within in-protocol PBS, it is not impossible and in this paper we will assume that the necessary cryptographic and relaying safeguards are taken to ensure that MEV redistribution rebates can be distributed.[1]

Staking, Lending, and Derivatives. While PoS provides the benefit of reduced energy usage for transaction processing and finality, it has notable economic flaws. Firstly, there is an excess concentration of wealth over PoW assets unless inflation schedules are chosen correctly [7]. Secondly, as PoS utilizes locking transaction processors' virtual assets to earn future rewards denominated in virtual assets as opposed to PoW (where one uses locked energy to earn the same rewards), there are a number of rehypothecation risks. One such risk is competition with lending — if a lending protocol can offer a higher yield than the underlying PoS network, then validators can migrate their stake to the lending protocol [3]. Since the security of a PoS network depends on the value of the assets locked within the protocol, such capital flight makes it much easier to execute a double spend attack or reduce liveness. Similar attacks that reduce the net quantity of assets used to secure the network can occur when using staking derivatives [5], which make up over 33.49% of Ethereum 2 stake [12].

Since MEV revenue from on-chain liquidations is often correlated with the yield of a lending protocol [9, 8], one natural question is to ask if *MEV redistribution*, *i.e.* the act of sharing MEV revenue with validators, reduces the impact of these competitive equilibria. Provided there is enough MEV revenue shared with validators, this can over come the competitive effect of derivatives and lending protocols. We formalize this problem and provide a positive solution to within this paper.

This Paper. We pose MEV redistribution as a dynamical system in which lending and staking portfolios of agents in a PoS system are affected by an exogenously chosen parameter that determines how much of the MEV extracted in a block is distributed to staking. In the

[1]Formally, virtually all MEV redistribution mechanisms are not resistant to off-chain agreements if there

absence of such MEV redistribution, it was shown in [3] that exponential reward inflation is necessary to avoid equilibria in which staking portfolios rapidly decrease. Counter-intuitively, we show that feeding back earned MEV revenue to staking returns can incentivize avoiding equilibria in which staking goes to zero *at significantly lower inflation rates*. We also show that the restrictions on the rate of reward inflation that were required without MEV redistribution can be eased.

Explicitly, denote the block reward at height k is $R(k)$ and $\epsilon(k)$ by the quantity of assets burned or held as protocol-owned liquidity[2]. The results of [3] can be summarized as: if $\epsilon(k) = 0$ and $R(k) = o(e^{\lambda k})$, then the network will reach equilibria where the staked supply goes to zero. Note that live networks such as Ethereum are already implementing $\epsilon(k) \neq 0$ via fee-burning mechanisms such as EIP-1559 [14], so expanding these results to $\epsilon(k) \neq 0$ is crucial for designing inflation schedules $R(k)$.
In §3.1, we prove that if $\lim_{k\to\infty} |R(k)/\epsilon(k)|$ is smaller than a constant depending on the risk preferences of validators, then the outcome where the staked supply goes to zero is locally stable for the staking-lending dynamics. This extends the results of [5] to include protocols that try to mitigate lending attacks by creating a treasury with protocol-owned liquidity. However, we then prove that if ones provides stakers with $\alpha\%$ of MEV profits (in addition to inflation rewards) in a pro-rata manner, then one can avoid the economically insecure equilibria where the staked supply is zero.

Conversely, we note that the condition for avoiding these bad equilibria, $R(k)/\epsilon(k) \to c(\alpha)$ is significantly weaker that the exponential inflation requirement of [3]. We demonstrate that adding in MEV redistribution can allow for networks to use *subexponential inflation* schedules while still avoiding the equilibria where the staked supply goes to zero. Furthermore, we validate these results using agent-based simulation in §3.2, where we find that even *deflationary* schedules (*i.e.* similar to that of Bitcoin) avoid bad equilibria provided that there exists some MEV redistribution.

Our main tool is dynamical systems theory, and specifically the analysis of the equilibria of the joint staking-lending dynamics. We expand the works of [3, 5] to include contributions from lending protocol MEV (*e.g.* liquidations), which leads to a feedback system. While our toy model does not represent all MEV, it shows that well-designed redistribution of MEV can have positive impacts on PoS networks. Our model is a derandomized version of [3], as we are able to describe asymptotic system equilibria without needing to analyze the probabilistic fluctuations that occur at finite block height. This simplification is what allows us to utilize dynamical systems to analyze asymptotic stability in staking systems with MEV redistribution.

---

are large enough (in terms of percentage of stake owned) players making agreements (see [2] for more discussion on this point.)

[2]Protocol-owned liquidity is when the decentralized protocol builds up a treasury from fees and/or percentages of block rewards that it uses for incentivizing liquidity or paying participants for development contributions. Formal analyses of protocol-owned liquidity include [10] and [5].

Notation	Denotes
$i \in [n]$	$i \in \{1, \ldots, n\}$
k	Block height, $k \in \mathbb{N}$
$\mathbb{1}$	The all ones vector $[1, \ldots, 1]^\top$
$\sigma_{\max}(A)$	Maximum eigenvalue of a matrix $A \in \mathbf{R}^{n \times n}$
$\mathsf{spec}(A)$	The set of eigenvalues of a matrix $A \in \mathbf{R}^{n \times n}$
$\mathbf{C}^o_+, \mathbf{C}^o_-$	Open right (resp. left) half of the complex plane, $\mathsf{Re}(z) > 0$ (resp. $\mathsf{Re}(z) < 0$)

Figure 1: Table summarizing mathematical notations used within the paper

2 Model

There are n rational agents, and a PoS protocol which has a total supply of its token $S(k) \in \mathbf{R}$ at block height k and a predefined reward schedule, $R(k+1) \in \mathbf{R}$. This implies that the token supply first increases via the linear relationship $\tilde{S}(k+1) = S(k) + R(k+1)$, where $\tilde{S}(k+1)$ is the base token supply at height $k+1$. We further assume that there is a quantity $\epsilon(k) \in \mathbf{R} - \{0\}$ with $|\epsilon(k)| \leq S(k)$ that represents a 'frozen' portion of the token supply that cannot be transferred, staked, or lent. If $\epsilon(k) > 0$, this is equivalent to protocol-owned liqudity (*i.e.* protocol constructs a treasury $\sum_k \epsilon(k)$ from the rewards) whereas if $\epsilon(k) < 0$ this corresponds to burning tokens. The token supply at height $k+1$ is represented as $S(k+1) = \tilde{S}(k+1) + \epsilon(k)$.

Every agent $i \in [n]$ has a total wealth of $W_i(k) \in \mathbf{R}^2_+$ and a portfolio $(\pi_i^{\text{stake}}(k), \pi_i^{\text{lend}}(k))^\top \in \mathbf{R}^2_+$, such that $\pi_i^{\text{stake}}(k) + \pi_i^{\text{lend}}(k) = W_i(k)$ (*i.e.* each agent has some fraction of their wealth staked and lent at every time). We denote the total lent supply $\ell(k) = \sum_{i=1}^n \pi_i^{\text{lend}}(k)$ and the total staked supply $T(k) = \sum_{i=1}^n \pi_i^{\text{stake}}(k)$, such that $S(k) = \ell(k) + T(k) + \epsilon(k)$ for all $k \in \mathbb{N}$. Each agent updates their portfolio based on a lending rate γ and a staking rate γ_i^s.

On-chain lending protocols such as Compound and Aave compute interest rates as $\gamma = g(U)$, where g is a deterministic function[3] and U is the *utilization rate*. The utilization is defined as the ratio of the outstanding amount of borrows (*e.g.* loans that are paying interest) to the amount of supplied assets. As in [3], we will make the following assumption on the demand distribution for formal analysis:

Assumption 1. *There exists* $0 < N \leq 1$ *such that the total borrowed quantity of assets is equal to* $NS(k)$.

This assumption can be relaxed and analyzed via the methods of [5, 3], which is left for future work. Given this assumption and a risk-free lending rate $\gamma_0 \in \mathbf{R}_+$, we can compute

[3]The function g is often piecewise linear or quadratic [9, 4] and in this paper we will assume it is linear.

the lending utilization at height k as $U(k) = \frac{NS(k)}{\ell(k)+S(k)}$ and write an update equation for the lending yield, $\gamma(k)$:

$$\gamma(k+1) = (1+\gamma_0)U(k)$$

Due to lending liquidations (*e.g.* liquidating loans that are insolvent), an amount $Q(k)$ of MEV is extracted from the PoS system. We want to understand the impact of distributing a fraction of this MEV to incentivize staking. Therefore, we make the following assumption, which places conditions on the MEV extracted as a function of lending yields:

Assumption 2. *Lending yield is positively correlated with MEV extracted at times in the future. That is, there exists* $c_{\max}$ *such that:*

$$Q(k+1) = \sum_{p=0}^{c_{\max}} \beta_p \gamma(k-p)$$

for coefficients $\beta_p \geq 0$*, and not all* $\beta_p = 0$*.*

For simplicity, we assume going forward that $c_{max} = 0$. This gives the update equation for the return of staking for agent i, $\gamma_i^s(k)$:

$$\begin{aligned}\gamma_i^s(k+1) &= \frac{\pi_i^{\text{stake}}(k)}{T(k)}(R(k+1) + \alpha Q(k+1)) \\ &= \frac{\pi_i^{\text{stake}}(k)}{S(k) - \ell(k) - \epsilon(k)}(R(k+1) + \alpha\beta_0\gamma(k))\end{aligned}$$

This equation shows that the expected return from staking for the ith agent is given by taking the reward $R(k+1)$ and MEV redistribution $\alpha Q(k+1) = \alpha\beta_0\gamma(k)$, and splitting them pro-rata (*e.g.* proportional to the fraction of the total staked amount owned by agent i).

At every block k, agents choose their portfolios $(\pi_i^{\text{stake}}(k), \pi_i^{\text{lend}}(k))^\top$ by solving a Markowitz mean-variance portfolio optimization problem that uses the expected return vectors $\mu_i(k) = (\gamma_i^s(k), \gamma(k))^\top$ and a fixed positive definite covariance matrix Σ_i, $\forall i \in [n]$. Specifically, agents' portfolios solve the following optimization problem:

$$\begin{aligned}(\pi_i^{\text{stake}}(k), \pi_i^{\text{lend}}(k))^\top = \arg\min_{w \in \mathbf{R}^2} \; & w^\top \Sigma_i w - \lambda \mu_i(k)^\top w \\ \text{such that } & w^\top \mathbb{1} = W_i(k)\end{aligned}$$

Optimality conditions for this problem[4] give agent i weights $w_i^\star$:

$$w_i^\star = \lambda \Sigma_i^{-1} \mu_i(k) \tag{1}$$

[4]We compute the optimality conditions without incorporating the constraint $w^\top \mathbb{1} = W_i(k)$, as we find that in practice, the constraint is nearly satisfied, and the analysis is greatly simplified without incorporating

Assumption 3. *The covariance matrices Σ_i are the following for every $i \in [n]$:*

$$\Sigma_i = \begin{bmatrix} \frac{1}{\kappa_i} & 0 \\ 0 & \frac{1}{\eta_i} \end{bmatrix}$$

where $\frac{1}{\kappa_i} \sim \mathsf{Exp}(\tau_{stake})$ and $\frac{1}{\eta_i} \sim \mathsf{Exp}(\tau_{lend})$ for some $\tau_{stake}, \tau_{lend} > 0$.

The interpretation of this assumption is that agents independently assess the risk of staking and lending via their variables and the exponential distribution captures population-level differences in risk tolerance. With this assumption, we explicitly solve the Markowitz problem $\forall i \in [n]$ and write the staking and lending portfolios:

$$\pi_i^{\text{stake}}(k) = \lambda \kappa_i \gamma_i^s(k)$$
$$\pi_i^{\text{lend}}(k) = \lambda \eta_i \gamma(k)$$

This allows us to write out the update equations for the mean staking return and lending yield as a feedback system in only the state variables $\gamma_i^s(k), \gamma(k)$ for $i \in [n]$. We use the notation $\gamma^s(k)$ to refer to $[\gamma_1^s(k), \ldots, \gamma_n^s(k)]^\top$, the vector of all the staking returns and $\gamma_{-i}(k)$ to refer to the staking returns of all agents but agent i. Given this setup, we can write:

$$\begin{aligned} \gamma_i^s(k+1) &= f_i(\gamma_i^s(k), \gamma_{-i}^s(k), \gamma(k), \alpha) \qquad (2) \\ &= \frac{\lambda \kappa_i \gamma_i^s(k)}{\sum_{j=1}^n \lambda \kappa_j \gamma_j^s(k) - \epsilon(k)} (R(k+1) + \alpha \beta_0 \gamma(k)) \\ \gamma(k+1) &= \hat{f}(\gamma^s(k), \gamma(k)) \qquad (3) \\ &= (1+\gamma_0) \frac{NS(k)}{\sum_{j=1}^n \lambda \eta_j \gamma(k) + S(k)} \end{aligned}$$

Here, the system's state is $(\gamma_1^s(k), \ldots, \gamma_n^s(k), \gamma(k)) \in \mathbf{R}^{n+1}$, and α, the fraction of lending MEV to be distributed for lending is an exogenously chosen parameter. We also write the total supply as $S(k) = \sum_{j=1}^n \lambda \kappa_i \gamma_i^s(k) + n\lambda \eta_i \gamma(k)$.

3 Results

Our results include theoretical results about the equilibria of (2)-(3) that MEV redistribution is able to avoid as well as simulation results. These simulation results provide numerical evidence that our theoretical results hold under various realistic parametrizations.

it. We note that [3] also solves the unconstrained problem and demonstrates that the equilibria of the unconstrained and constrained problems are close given some constraints on the initial stake distribution $\pi^{\text{stake}}(0)$.

3.1 Formal Results

We seek to understand the behavior of the dynamics (2)-(3) as a function of the parameter α. We seek to show that for large enough α, MEV redistribution causes the system to avoid the tendency to unstake. Stated informally, we find given sufficient conditions on the sequences $R(k)$ and $\epsilon(k)$, such that there exists an $\alpha \in (0,1)$ under the dynamics (2)-(3) where the equilibrium point at which all agents unstake completely, i.e. $\pi_i^{\text{stake}}(k) = 0$, and correspondingly $\gamma_i^s(k) = 0$ for all i, is locally *unstable*. In other words, this point is locally avoided by the joint staking-lending dynamics.

We first characterize the set of undesirable equilibria:

Lemma 1. *There exists a $\gamma^\star \in \mathbf{R}$ such that $(\gamma_1^{s,\star}, \ldots, \gamma_n^{s,\star}, \gamma^\star) = (0, \ldots, 0, \gamma^\star)$ is an equilibrium point of the dynamical system (2)-(3).*

Proof. We compute this explicitly. From the definition of an equilibrium point of a discrete-time system [15], any equilibrium point satisfies the $n+1$ fixed point equations:

$$\gamma_i^{s,\star} = \frac{\lambda\kappa_i\gamma_i^{s,\star}}{\sum_{i=1}^n \lambda\kappa_i\gamma_i^{s,\star} - \epsilon(k)}(R(k+1) + \alpha\beta_0\gamma^\star)$$
$$\gamma^\star = (1+\gamma_0)\frac{NS^\star}{\sum_{j=1}^n \lambda\eta_j\gamma^\star + S^\star}$$

where $S^\star = \sum_{i=1}^n \lambda\kappa_i\gamma^\star + \lambda\eta_i\gamma_i^{s,\star}$. It is clear that when $\epsilon(k) \neq 0$ for all k, $\gamma_i^{s,\star} = 0$ satisfy the first n equations because

$$0 = \frac{\lambda\kappa_i 0}{-\epsilon(k)}(R(k+1) + \alpha\beta_0\gamma^\star)$$

To show the existence of $\gamma^\star$ that satisfies the last equation, we see that when $\gamma_i^{s,\star} = 0$ for all i, $S^\star = \sum_{i=1}^n \lambda\kappa_i\gamma^\star + \lambda\eta_i\gamma_i^{s,\star} = \sum_{i=1}^n \lambda\eta_i\gamma^\star$, which reduces the last equation to:

$$\gamma^\star = (1+\gamma_0)\frac{N\sum_{i=1}^n \lambda\eta_i\gamma^\star}{2\sum_{j=1}^n \lambda\eta_j\gamma^\star}$$

Cancelling $\sum_{i=1}^n \lambda\eta_i\gamma^\star$ gives:

$$\gamma^\star = \frac{(1+\gamma_0)N}{2}$$

□

This lemma establishes that the system (2)-(3) has an equilibrium point when all agents' staking returns go to zero. This also implies that the staked portfolios, $\pi_i^{\text{stake}}(k)$ are identically zero for all i when the agents are at this equilibrium, as the Markowitz portfolio update uses the staking returns, which are zero, to compute portfolios as $\pi_i^{\text{stake}}(k) = \lambda\kappa_i\gamma_i^s(k)$.

Next, we characterize the stability properties of this equilibrium point via its (Jacobian) linearization. The linearization of a nonlinear system provides valuable information about the behavior of the system around an equilibrium point. In particular, we show that the eigenvalues of the linearization of (2)-(3) at the undesirable equilibrium at $(0, \ldots, 0, \gamma^\star)$ are an increasing function of α. We will later use this fact to establish the existence of an α such that the equilibrium point becomes unstable.

Lemma 2. *Consider the dynamics:*

$$\begin{bmatrix} \gamma_1^s(k+1) \\ \cdots \\ \gamma_n^s(k+1) \\ \gamma(k+1) \end{bmatrix} = df^\alpha|_{(0,\ldots,0,\gamma^\star)} \begin{bmatrix} \gamma_1^s(k) \\ \cdots \\ \gamma_n^s(k) \\ \gamma(k) \end{bmatrix}$$

where $df^\alpha|_{(0,\ldots,0,\gamma^\star)} \in \mathbf{R}^{(n+1)\times(n+1)}$ *is the Jacobian linearization*[5] *of* (2)-(3) *at the point* $(0, \ldots, 0, \gamma^\star)$ *as a function of* α. *Then,* $|\sigma_{max}(df^\alpha|_{(0,\ldots,0,\gamma^\star)})|$ *is an increasing function of* α.

Proof. The main step of the proof is to show that $df^\alpha|_{(0,\ldots,0,\gamma^\star)}$ is a lower triangular matrix, and in particular, this means that its eigenvalues are just its diagonal entries [16]. We first compute the linearization of the dynamics, $df^\alpha \in \mathbf{R}^{(n+1)\times(n+1)}$. This matrix has the following diagonal entries for $i \in [n]$:

$$df^\alpha_{i,i} = \frac{\lambda\kappa_i(R(k+1) + \alpha\beta_0\gamma(k))}{\sum_{j=1}^n \lambda\kappa_j\gamma_j^s(k) - \epsilon(k)} - \frac{\kappa_i^2\lambda^2\gamma_i^s(k)(R(k+1) + \alpha\beta_0\gamma(k))}{(\sum_{j=1}^n \lambda\kappa_i\gamma_i^s(k) - \epsilon(k))^2}$$

and the following remaining diagonal entry:

$$df^\alpha_{n+1,n+1} = \frac{N(1+\gamma_0)\sum_{i=1}^n \lambda\eta_i}{\ell(k) + S(k)} - \frac{N(1+\gamma_0)2\sum_{i=1}^n \lambda\eta_i S(k)}{(\ell(k) + S(k))^2}$$

Next, we compute the off-diagonal entries $df^\alpha_{i,j}$ for $j > i$ (these are the entries in the strictly upper triangular part of df^α):

$$df^\alpha_{i,j} = \frac{-\kappa_i^2\lambda^2\gamma_i^s(k)(R(k+1) + \alpha\beta_0\gamma(k))}{(\sum_{j=1}^n \lambda\kappa_j\gamma_j^s(k) - \epsilon(k))^2}$$

[5]For a dynamical system $x_{t+1} = f(x_t)$ with $x \in \mathbf{R}^n$, $f : \mathbf{R}^n \to \mathbf{R}^n$ the Jacobian linearization $df|_{x_0}$ is the matrix $\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \ldots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \ldots & \ldots & \ldots \\ \frac{\partial f_n}{\partial x_1}(x_0) & \ldots & \frac{\partial f_n}{\partial x_n}(x_0) \end{bmatrix}$

Evaluating these three kinds of entries at the equilibrium point $(0, \ldots, 0, \gamma^\star)$, we have:

$$\begin{aligned} df^\alpha_{i,i}|_{(0,\ldots,0,\gamma^\star)} &= \frac{\lambda\kappa_i(R(k+1)+\alpha\beta_0\gamma^\star)}{-\epsilon(k)} \\ df^\alpha_{n+1,n+1}|_{(0,\ldots,0,\gamma^\star)} &= \frac{N(1+\gamma_0)\sum_{i=1}^n \eta_i}{2\ell(k)} - \frac{N(1+\gamma_0)2\sum_{i=1}^n \lambda\eta_i}{4\ell(k)} \\ &= 0 \\ df^\alpha_{i,j}|_{(0,\ldots,0,\gamma^\star)} &= 0 \text{ for } j > i \end{aligned}$$

This establishes that $df^\alpha|_{(0,\ldots,0,\gamma^\star)}$ is a lower triangular matrix, and thus its eigenvalues are simply its diagonal entries, precisely given by the terms $df^\alpha_{i,i}|_{(0,\ldots,0,\gamma^\star)}$ and $df^\alpha_{n+1,n+1}|_{(0,\ldots,0,\gamma^\star)}$. Noting that the maximum eigenvalue, therefore, is :

$$\sigma_{\max}(df^\alpha|_{(0,\ldots,0,\gamma^\star)}) = \lim_{k\to\infty} \lambda \max_i \kappa_i \frac{(R(k+1)+\alpha\beta_0\gamma^\star)}{-\epsilon(k)} \tag{4}$$

we see immediately that $|\sigma_{\max}(df^\alpha|_{(0,\ldots,0,\gamma^\star)})|$ is an increasing function of α, completing the proof. □

We now recall the following two theorems from [15] regarding the linearization $A = df|_{x^\star}$ of a nonlinear system:

$$\dot{x} = f(x) \tag{5}$$

for $x \in \mathbf{R}^n$, $x(0) = x_0$, at an equilibrium point $x^\star = 0$ (that is $f(0) = 0$), and where f is continuously differentiable.

Theorem 1. *[[15] Theorem 5.41: Stability from Linearization] If*

$$\lim_{|x|\to 0} \frac{|f(x) - Ax|}{|x|} = 0 \tag{6}$$

then the equilibrium point $x^\star$ is a locally asymptotically stable equilibrium point of the nonlinear system (5) *if it is an asymptotically stable equilibrium point* [6] *of $\dot{z} = Az$, $z \in \mathbf{R}^n$, that is,* $\mathsf{spec}(A) \in \mathbf{C}^o_-$.

Theorem 2. *[[15] Theorem 5.42: Instability from Linearization] If the linearization $A = df|_{x^\star}$ has at least one eigenvalue in $\mathbf{C}^o_+$, then the equilibrium* 0 *of the nonlinear system* (5) *is unstable.*

Jointly, these theorems give conditions on when the stability or instability of an equilibrium point of a *nonlinear* system may be deduced from its linearization at that point. The first result says that provided a suitable growth condition on the remainder of the linearization, $f(x) - Ax$ (which implies that the remainder is not $o(|x|)$), the local asymptotic stability of

[6] For a definition of asymptotic stability see [15].

the equilibrium point of the nonlinear system can be deduced by the asymptotic stability of its linearization. The second theorem says that if the linearization of a continuous time nonlinear system at an equilibrium point has an eigenvalue in the open right half plane, then the equilibrium point of the nonlinear system is unstable. In both theorems, one can replace the condition on the eigenvalues of the linearization for continuous time systems by an equivalent condition for discrete time systems. In Theorem 1, one replaces the condition that $\mathsf{spec}(A) \in \mathbf{C}_-^o$ by $\mathsf{spec}(A) \in D_1(0) = \{z \in \mathbf{C} : \|z\| < 1\}$. In Theorem 2, one replaces the existence of at least one eigenvalue of the linearization being in the $\mathbf{C}_+^o$ with at least one eigenvalue being outside the unit disk $D_1(0)$. Motivated by this, we show in the following that provided suitable conditions on $R(k)$ and $\epsilon(k)$, there exists an α that makes $|\sigma_{\max}(df^\alpha|_{(0,\ldots,0,\gamma^\star)})| > 1$, thereby rendering the nonlinear system (2)-(3) unstable at $(0,\ldots,0,\gamma^\star)$.

The following proposition characterizes the behavior of the dynamical system (2)-(3) around the equilibrium $(0,\ldots,0,\gamma^\star)$ in the regime when $\alpha = 0$, that is, when no MEV revenue is redistributed to incentivize staking in the PoS system. We show that if the staking rewards are not inflating fast enough, i.e. $R(k+1)$ is not growing faster than $\epsilon(k)$, the undesirable equilibrium may be locally stable (and therefore be reached) for the staking-lending dynamics. Let $\hat{\kappa} = \max_{i\in[n]} \kappa_i$. Then:

Proposition 1. *When*

$$\left|\lambda\hat{\kappa} \lim_{k\to\infty} \frac{R(k+1)}{-\epsilon(k)}\right| < 1$$

and $\alpha = 0$, the equilibrium point $(0,\ldots,0,\gamma^\star)$ of (2)-(3) is locally asymptotically stable.

Proof. The proof is directly by substituting $\alpha = 0$ in the formula for the maximum eigenvalue of the linearization (4) from Lemma 2 and invoking Theorem 1. □

This proposition says that if $R(k)$ is not growing fast enough relative to the burned amount or protocol owned liquidity, then it may be possible to reach the undesirable equilibrium at $(0,\ldots,0,\gamma^\star)$. It is always the case that $\frac{1+\gamma_0}{2} < 1$ (γ_0 is a percentage risk-free lending rate), so in order for $|\sigma_{\max}(df^\alpha|_{(0,\ldots,0,\gamma^\star)})|$ to be made larger than one without any MEV redistribution ($\alpha = 0$), we must have:

$$\lambda\hat{\kappa} \lim_{k\to\infty} \frac{R(k+1)}{\epsilon(k)} > 1 \tag{7}$$

For example, if the protocol is burning its tokens at a rate $\epsilon(k) = e^{r_0 k}$, then for the undesirable equilibrium to be rendered unstable, the inflation schedule must be $R(k) = e^{rk}$ for $r > r_0 + \log\frac{1}{\lambda\hat{\kappa}}$.

The following theorem demonstrates that to render the undesirable equilibrium $(0,\ldots,0,\gamma^\star)$ unstable using MEV redistribution ($\alpha > 0$), it is sufficient to use reward schedules $R(k)$ that inflate at a slower rate than the required rate in the $\alpha = 0$ case above.

Theorem 3. *Suppose that the PoS protocol burns the tokens at a rate $\epsilon(k) = e^{r_0 k}$ and has a reward schedule $R(k) = e^{rk}$. Then, for $\alpha > 0$, the equilibrium $(0, \ldots, 0, \gamma^\star)$ of (2)-(3) is unstable if:*

$$r > 2r_0 + 2\log\frac{1}{\lambda\hat{\kappa}} - \alpha\beta\gamma^\star \tag{8}$$

Proof. The equilibrium $(0, \ldots, 0, \gamma^\star)$ is unstable for $\alpha > 0$ if:

$$\frac{\lambda\hat{\kappa}(e^{rk} + \alpha\beta_0\gamma^\star)}{e^{r_0 k}} > 1 \tag{9}$$

Dividing by $\lambda\hat{\kappa}$, writing $\alpha\beta_0\gamma^\star = e^{\log\alpha\beta_0\gamma^\star}$, taking log on both sides, and writing $\log(e^{rk} + e^{\log\alpha\beta_0\gamma^\star}) = \log(1 + e^{\log\alpha\beta_0\gamma^\star} - rk) + rk - \log e^{r_0 k}$ we have:

$$\log(1 + e^{\log\alpha\beta_0\gamma^\star} - rk) + rk - \log e^{r_0 k} > \log\frac{1}{\lambda\hat{\kappa}}$$

Using the bound $\frac{1}{2}x > \log 1 + x$, we have:

$$\begin{aligned} 0.5e^{\log\alpha\beta_0\gamma^\star} + 0.5rk &> \log e^{r_0 k} + \log\frac{1}{\lambda\hat{\kappa}} \\ \implies r &> 2r_0 + 2\log\frac{1}{\lambda\hat{\kappa}} - \alpha\beta\gamma^\star \end{aligned}$$

□

The above theorem demonstrates that MEV redistribution reduces the rate of reward inflation needed to make the equilibrium point $(0, \ldots, 0, \gamma^\star)$ unstable. We note that this bound is not tight, and can be made sharper, which we leave as future work. For a fixed reward schedule $R(k)$, this theorem also provides conditions on the existence of an $0 < \alpha < 1$ such that the condition in (8) is met. Note that if β_0 and $\gamma^\star$ are large enough, then it may be possible to make the reward schedule have *subexponential* growth, i.e. $r < 1$.

3.2 Simulation Results

We simulate the staking dynamics of equations (2) and (3) by choosing an initial wealth distribution $W_i(0)$ and sampling from the distribution $\hat{\pi}_i(k) = \frac{\pi_i^{\text{stake}}(k)}{\sum_{i=1}^n \pi_i^{\text{stake}}(k)}$. This is an agent-based simulation, as we model each user $i \in [n]$ independently and not simply the distribution of the aggregate stake $S(k) - \ell(k) - \epsilon(k)$. We perform the simple Markov update $i(k) \sim \hat{\pi}(k), \pi_i^{\text{stake}}(k+1) \leftarrow \pi_i^{\text{stake}}(k) + R(k+1)$ to update the stake distrbution. Afterwards, we update the Markowitz weights using the constrained version of (1). In this simulation, we set $\epsilon(k) = c > 0$ for all k so that we can make a direct comparison to the empirical results of [3].

Using $J = 10,000$ simulated runs for each sampled value of α (each of $T = 100,000$

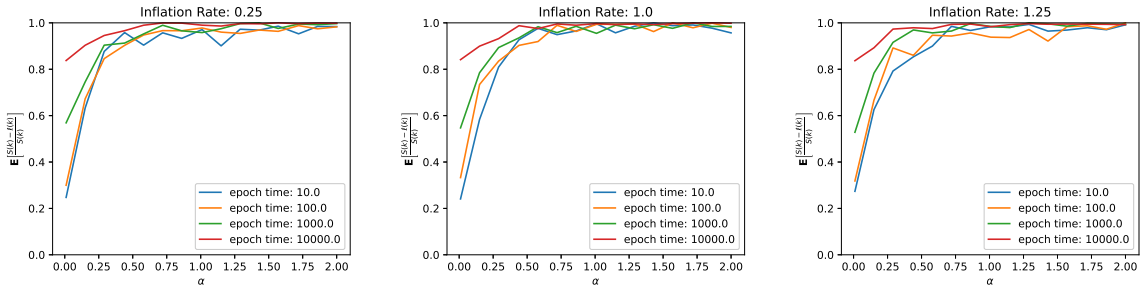

Figure 2: Percentage of supply staked, $\mathbf{E}\left[\frac{(S(k)-\ell(k)}{S(k)}\right]$ for different inflation rates, $r \in \{0.25, 1.0, 1.25\}$ (*i.e.* $R(k) = r^k$). We see that regardless of the inflation rate, whether deflationary ($\lambda = 0.25$), constant ($\lambda = 1.0$), or inflationary ($\lambda = 1.25$), increasing α leads to the staked percentage going to 1. The values where $\alpha > 1$, while mathematically feasible, are unlikely to be feasible in practice since $\alpha = 1$ means that 100% of MEV revenue goes to stakers. One can view these points as adding an extra subsidy to the block reward; we include them strictly for illustration purposes only.

timesteps), we receive trajectories of portfolios $(\pi_{ij}^{\text{stake}}(k), \pi_{ij}^{\text{lend}}(k))$ for $i \in [n], j \in [J]$ and compute the average of the percentage of staked supply:

$$\mathbf{E}\left[\frac{S(k)-\ell(k)}{S(k)}\middle|\alpha\right] \approx \frac{1}{T}\sum_{k=1}^{T}\sum_{j=1}^{J}\sum_{i=1}^{n}\frac{\pi_{ij}^{\text{stake}}(k) - \pi_{ij}^{\text{lend}}(k)}{\pi_{ij}^{\text{stake}}(k) + \pi_{ij}^{\text{lend}}(k)}$$

We plot this in Figure 2, where the y-axis is $\mathbf{E}\left[\frac{S(k)-\ell(k)}{S(k)}\middle|\alpha\right]$ and the x-axis varies α. One can clearly see an even stronger result than those of §3.1 — even fully deflationary, uniformly bounded inflation schedules with $R(k) \to 0$ as $k \to \infty$ still have high staking rates when $\alpha > 0.25$. This suggests that even mild MEV revenue sharing can avoid economically unsafe equilibria for PoS systems.

Given that realistic PoS systems only allow stakers to deposit or withdraw stake on epochs (*i.e.* minimum unit of time that stake must be locked into the network in order to earn rewards), we also simulated the behavior under different epoch times, E. This corresponds to only allowing rebalances (*e.g.* updates via equation (1)) at block heights $h \equiv 0 \bmod E$, which is equivalent to the multi-period Markowitz problems [1]. Figure 2 illustrates that for epoch times $E \in \{10, 100, 1000, 10,000\}$, the staked supply percentage has the same behavior for large α. Note that larger epoch times naturally reduce rebalancing out of staking, but are capital inefficient for validators. This implies that MEV redistribution provides positive impact to staking network security, regardless of how epoch times are chosen.

4 Conclusion and Future Work

In this paper, we present the first formal analysis of MEV redistribution, a proposed mechanism in Ethereum. We demonstrate that MEV redistribution, if feasible in practice, can reduce the likelihood of economically insecure equilibria for PoS systems. These equilibria correspond to scenarios where most users with assets that could be staked do not stake them, making attacks (such as double-spending) cheaper to execute as the cost of attack depends on the amount staked in the network. We demonstrate both formally and via simulation that as the redistribution increases, the likelihood of reaching such equilibria go down. Our results extend the framework of [5] and demonstrate that bad equilibria can be avoided with *sub-exponential* and even deflationary block reward schedules if MEV revenue can be shared.

Future work to improve this model includes, but is not limited to: analyzing the stochastic fluctuations of expected reward at finite block height (akin to [4]), using different utilization formulas, analyzing the choice of burn or protocol-owned liquidity models more thoroughly, and proving that deflationary (not simply subexponential) reward schedules work in preventing undesirable economic equilibria of the joint staking-lending dynamics. The last point effectively would formally prove that the simulation results for the top panel of Figure 2 match the formal results of §3.1.

References

[1] Stephen Boyd, Enzo Busseti, Steve Diamond, Ronald N Kahn, Kwangmoo Koh, Peter Nystrup, Jan Speth, et al. 2017. Multi-period trading via convex optimization. *Foundations and Trends® in Optimization* 3, 1 (2017), 1–76.

[2] Vitalik Buterin. 2022. State of research: Increasing censorship resistance of transactions under proposer/builder separation (PBS). `https://notes.ethereum.org/@vbuterin/pbs_censorship_resistance`.

[3] Tarun Chitra. 2021. Competitive Equilibria Between Staking and On-chain Lending. `https://cryptoeconomicsystems.pubpub.org/pub/chitra-staking-lending-equilibria`. *Cryptoeconomic Systems* 0, 1 (April 2021).

[4] Tarun Chitra, Guillermo Angeris, Alex Evans, and Hsien-Tang Kao. 2021. A Note on Borrowing Constant Function Market Maker Shares. (2021).

[5] Tarun Chitra and Alex Evans. 2020. Why stake when you can borrow? *arXiv preprint arXiv:2006.11156* (2020).

[6] Philip Daian, Steven Goldfeder, Tyler Kell, Yunqi Li, Xueyuan Zhao, Iddo Bentov, Lorenz Breidenbach, and Ari Juels. 2019. Flash boys 2.0: Frontrunning, transaction reordering, and consensus instability in decentralized exchanges. *arXiv preprint*

arXiv:1904.05234 (2019).

[7] Giulia Fanti, Leonid Kogan, Sewoong Oh, Kathleen Ruan, Pramod Viswanath, and Gerui Wang. 2019. Compounding of wealth in proof-of-stake cryptocurrencies. In *International conference on financial cryptography and data security*. Springer, 42–61.

[8] Watson Fu, Tarun Chitra, Rei Chiang, and John Morrow. 2021. Aave Market Risk Assessment. (2021). `https://gauntlet.network/reports/aave`.

[9] Hsien-Tang Kao, Tarun Chitra, Rei Chiang, and John Morrow. 2020. An analysis of the market risk to participants in the compound protocol. In *Third International Symposium on Foundations and Applications of Blockchains*.

[10] Kshitij Kulkarni, Theo Diamandis, Tarun Chitra, et al. 2022. Towards a Theory of Maximal Extractable Value I: Constant Function Market Makers. *arXiv preprint arXiv:2207.11835* (2022).

[11] Alex Obadia and Taarush Vemulapalli. 2022. MeV in ETH2 - an early exploration. `https://hackmd.io/@flashbots/mev-in-eth2`.

[12] ratedw3b. 2022. Eth2 Liquid Staking Dashboard. `https://dune.com/ratedw3b/Eth2-Liquid-Staking`

[13] BitMEX Research. 2022. Flashbots. `https://blog.bitmex.com/flashbots/`.

[14] Tim Roughgarden. 2021. Transaction Fee Mechanism Design. *SIGecom Exch.* 19, 1 (2021), 52–55. `https://doi.org/10.1145/3476436.3476445`

[15] Shankar Sastry. 2013. *Nonlinear systems: analysis, stability, and control*. Vol. 10. Springer Science & Business Media.

[16] Gilbert Strang. 1993. *Introduction to linear algebra*. Vol. 3. Wellesley-Cambridge Press Wellesley, MA.

[17] Flashbots Team. 2021. Flashbots Explore v0. `https://explore.flashbots.net`.

[18] Aviv Yaish, Gilad Stern, and Aviv Zohar. 2022. Uncle Maker:(Time) Stamping Out The Competition in Ethereum. *Cryptology ePrint Archive* (2022).

[19] Aviv Yaish, Saar Tochner, and Aviv Zohar. 2022. Blockchain Stretching & Squeezing: Manipulating Time for Your Best Interest. In *Proceedings of the 23rd ACM Conference on Economics and Computation*. 65–88.