*ECAI 2023*
*K. Gal et al. (Eds.)*


*doi:10.3233/FAIA230501*

# Stackelberg Attacks on Auctions and Blockchain Transaction Fee Mechanisms

**Daji Landis[b] and Nikolaj Schwartzbach[a]**

[b]Bocconi University
[a]Aarhus University
ORCiD ID: Daji Landis https://orcid.org/0000-0002-9985-0552,
Nikolaj Schwartzbach https://orcid.org/0000-0002-0610-4455

**Abstract.** We study a multi-unit single-demand auction in a setting where agents can arbitrarily commit to strategies that may depend on the commitments of other agents. Such commitments non-trivially change the equilibria of the auction by inducing a metagame, in which agents commit to strategies. We demonstrate a strategy an attacker may commit to that ensures they receive one such item for free, while forcing the remaining agents to enter a lottery for the remaining items. The attack is detrimental to the auctioneer, who loses most of their revenue. We show that the strategy works as long as the agents have valuations that are somewhat concentrated. The attack is robust to a large fraction of the agents being either oblivious to the attack or having exceptionally high valuations. The attacker may coerce these agents into cooperating by promising them a free item. We show that the conditions for the attack to work hold with high probability when (1) the auction is not too congested, and (2) the valuations are sampled i.i.d. from either a uniform distribution or a Pareto distribution. The attack works for first-price auctions, second-price auctions, and the transaction fee mechanism EIP-1559 used by Ethereum.

## 1 Introduction

Consider $n$ agents participating in an auction with $m$ copies of the same item. Each agent $i$ receives utility $v_i > 0$ by obtaining one of the copies. Assume that all $v_i$ are distinct and ordered $v_1 < v_2 < \cdots < v_n$. Each agent places a bid $b_i \geq 0$ and the $m$ agents with the highest bids receive a copy of the item, at the cost of paying some function of the bids. If there are multiple agents with the same bid, the mechanism chooses uniformly at random among these agents. If $m \geq n$ then all agents receive a copy of the item, in which case the optimal strategy for each agent is to bid $b_i = 0$. Thus, we will assume that $n = (1 + \alpha)\, m$ for some *congestion constant* $\alpha > 0$.

In a first-price auction, an agent pays their own bid which results in untruthful behavior: it is well-known that the best response for an agent $i$ is to slightly outbid agent $n - m$ if their valuation exceeds this bid. That is, agent $i$ will place the following bid.

$$b_i = \begin{cases} v_{n-m} + \varepsilon & \text{if } i > n - m, \\ 0 & \text{if } i \leq n - m. \end{cases} \tag{1}$$

Where $\varepsilon > 0$ is some small constant, representing a negligible amount of money. It is not hard to see that this bidding strategy is indeed an equilibrium (at least up to $\varepsilon$). Of course, this requires that the agents are able to estimate the valuations of other parties. In some applications, this might not be a realistic assumption. Instead, the mechanism can be made truthful by letting each party with a winning bid pay $b_{n-m}$, a second-price[1] auction [40]. In this case, it can be shown that the proposed mechanism is truthful so that each party will bid their valuations [40, 13, 22]. While truthfulness is a desirable property, these auctions may be vulnerable to collusion [35].

**Blockchains.** The auction described is also known as a *transaction fee mechanism* and is used in blockchains to determine which transactions to include in the next block of data to be included in the chain [12]. Here, all pending transactions are public, so it is reasonable to assume agents know the valuations of other parties. Although blockchains canonically store transactions of cryptocurrency between different accounts [10, 11, 4], many blockchains have since generalized this to support arbitrary execution of code, so-called smart contracts [42]. Smart contracts are decentralized programs that run on a virtual machine implemented by the blockchain. A smart contract maintains state, can transfer funds between parties, and responds to queries. A smart contract is guaranteed to be faithful to its implementation by security of the underlying blockchain [26, 25].

**Stackelberg Equilibria.** It is well-known that being able to commit to strategies, in general, changes the equilibria of the game by allowing an agent to commit to acting irrationally in some subgame. The case with one agent being allowed to commit to strategies is known as a Stackelberg equilibrium and was first used in economics by von Stackelberg [41] to model competing firms, where one firm (the leader) has market dominance. This was later generalized to the scenario in which the leader commits to a strategy that depends on the strategies chosen by the other players, in what is known as reverse Stackelberg equilibria [18, 19]. This was further generalized by Hall-Andersen and Schwartzbach [23], who consider a model of 'universal commitments,' wherein all players have smart contracts that can depend on each other sequentially. They show that this constitutes a hierarchy of equilibria that generalizes Stackelberg equilibria and reverse Stackelberg equilibria.

In this paper, we study transaction fee mechanisms involving agents who can universally commit to strategies such as by using smart contracts, either on the blockchain in question or a parallel

[1] Technically, the auction should be called a $(n-m)^{\text{th}}$-price auction, or a Vickrey auction; we stick to second-price for simplicity.

one. This can alter the equilibrium of the game in attacks we call 'Stackelberg attacks.' This suggests the following natural question.

*How do universal commitment to strategies impact the equilibria of transaction fee mechanisms?*

We show that these commitments drastically change the structure of equilibria for various types of auctions, thus showing they are vulnerable to a Stackelberg attack wherein the buyers spontaneously organize to conspire against the auctioneer. In the attack, some agent commits to a strategy that ensures that they receive one of the items for free, while the remaining agents enter into a lottery for the remaining space on the block. The attack benefits all the buyers, but is detrimental to the auctioneer, who stands to lose most of their revenue. Note that while blockchains and smart contracts provide a natural setting in which to study these attacks, in principle the same framework can be used to analyze any setting in which agents can credibly commit to strategies, e.g. through reputation or by staking money. Understanding these attacks may also be important in predicting the behavior of advanced intelligent systems that have access to the internet (and hence access to a blockchain). We stress that the effectiveness of this attack is limited to the effectiveness of the commitment strategy. If the commitments can be undermined, say by the auctioneer, then so too can the attack be undermined.

## 1.1 Our Results

We demonstrate the existence of a Stackelberg attack on the transaction fee mechanism EIP-1559, which is used by Ethereum. This mechanism is a generalization of first-price auctions intended to fix various problems with first-price auctions in the context of transaction fee mechanisms [35]. By corollary, we show an attack on first-price auctions, which are used as transaction fee mechanisms in most other blockchains. The attack allows any agent to ensure they receive the item almost for free, while forcing (most of) the other agents to participate in a lottery for the remaining items. The attack works as long as the valuations are concentrated, in the sense that (most of) the largest values are not too much larger than the middle values. In this case, each agent voluntarily chooses the lottery because doing so will award them the item for free at some cost, while in the auction they would have to pay an amount commensurate with their valuation. If instead the valuations were spread out, the agents with a high valuation would not participate because they would be getting the item for a price much lower than their valuation, but with a degree of uncertainty.

**Theorem 1** (Informal)**.** *Let $v_1 < v_2 < \cdots < v_n$ be the valuations of the agents and $n > m$. Then if for some $k < m$, it holds that,*

$$\frac{v_{n-k+1}}{v_{n-m}} < \frac{n-k}{n-m},$$

*then EIP-1559 (including first-price auctions), as well as second-price auctions, are not Stackelberg resilient.*

This is shown by explicitly demonstrating a strategy that an agent may commit to for which the equilibrium involves most parties entering into a lottery as described. The strategy extends also to second-price auctions.

We evaluate the economic efficiency of this new situation and show that, while the attack benefits all users, it is detrimental to the auctioneer. This impact on auctioneer suggests that successful and widespread deployment of the attack would be detrimental to the viability of running the auctions. Therefore, our analysis is grounds for reevaluation of the auctions for transaction fee mechanisms. Formally, we define the *price of defiance* as the ratio between the utility an agent receives by cooperating versus the utility they would receive by deviating (or *defying* the attacker). We give a probabilistic bound on the price of defiance for the attack.

**Theorem 2** (Informal)**.** *Suppose $n$ agents participate in an auction with $m$ identical items, and $n = (1+\alpha)\,m$ for some $\alpha > 0$. If the agents have valuations that are i.i.d. uniform, then with high probability, the price of defiance is at least $1+\alpha$.*

We show that the conditions required to apply Theorem 1 are natural, in the sense that they are satisfied with high probability at certain levels of congestion when the valuations are sampled from two natural distributions.

**Theorem 3** (Informal)**.** *Suppose $n$ agents participate in an auction of $m$ identical items, and $n = (1+\alpha)\,m$ for some $\alpha > 0$. Then the conditions required for Theorem 1 to apply hold with overwhelming probability if either of the following two conditions are satisfied.*

1. *The valuations are sampled i.i.d. from a uniform distribution and,*
$$0 \leq \alpha < 0.53.$$
2. *The valuations are sampled i.i.d. from a Pareto distribution with parameter $p > 1$ and any $0 \leq \alpha < \alpha(p)$ for some function $\alpha(\cdot)$ with,*
$$\lim_{p\to\infty} \alpha(p) \approx 0.69.$$

The problem we study is natural in Web3 systems where agents natively interact using a blockchain. Thus, the agents are capable of deploying smart contracts that commit them to placing certain bids. In particular, the setting of an auction with multiple identical items models the transaction fee mechanisms that are used by blockchains to determine which transactions to include in the next block. Our work demonstrates that these mechanisms, in theory, are vulnerable to these attacks and may be cause for re-evaluation of the use of auctions in transaction fee mechanisms, at least when the networks are not too congested. Our work highlights the difficulty in designing smart contracts and suggests that other smart contracts that have already been deployed on major blockchains may be susceptible to Stackelberg attacks.

## 1.2 Related Work

Stackelberg equilibria are quite well-studied and are important e.g. in control theory [6, 8, 34] and security games [29, 24, 37]. In general, these commitments change the equilibria in highly non-trivial ways [3, 36], and they are known to be hard to compute in general [14, 28, 5], though there are some games for which the Stackelberg equilibrium can be shown to coincide with the subgame perfect equilibrium (SPE) [7]. Variants of Stackelberg equilibria are known for some auction scenarios, e.g. for all-pay auctions with complete information [27], for procurement auctions [17, 16], and for repeated auctions [33]. Reverse Stackelberg equilibria are less studied, however they also find applications in routing [20], in control theory [21, 39], and sparsely in auctions [32] though in a different context than what we consider in this work. In fact, the attack we consider in this work only works for $n \geq 3$ which means it inherently eludes analysis as a (reverse) Stackelberg equilibrium. Little is known of the generalizations of Stackelberg equilibria that we study in this work, aside from the complexity results shown in [23].

In recent years, there has been increased interest in analyzing blockchain transaction fee mechanisms using techniques from classic mechanism design. A line of work, [31, 35], identifies three desiderata of such mechanisms:

1. *user-incentive compatibility (UIC)*. The users are incentivized to bid truthfully;
2. *miner-incentive compatibility (MIC)*. The miners are incentivized the implement the mechanism as prescribed;
3. *off-chain agreement proofness (OCA proofness)*. No coalition of miners and users can increase their joint utility by deviating from the mechanism.

In [35], Roughgarden shows that EIP-1559 satisfies MIC and OCA proofness when the block size is large and shows that it is not UIC, in the sense that users may benefit by bidding strategically. Here, OCA proofness means that the users and the miner cannot benefit by agreeing to off-chain payments and thus captures a specific type of commitment to strategies. Chung and Shi [12] show that no mechanism can simultaneously be UIC and 1-OCA proof. These results are shown in a model where agents cannot universally commit to strategies, and indeed we show that, arguably, neither of these three properties hold in a model where the agent can universally commit to strategies.

## 2 A Stackelberg Attack on Auctions

We will consider a set of $n$ transactions competing for space on a block of size $m$. We assume for simplicity that each transaction is owned by exactly one agent, which we identify with the integers $\{1, 2, \ldots, n\}$. Each agent $i$ has a valuation $v_i > 0$ of their transaction, which is the utility they gain by having their transaction included in the block for free. We assume the agents are rational, risk-neutral, and have a quasilinear utility functions. We will take each $v_i$ to be sampled i.i.d. from some known underlying distribution $D$. It will be convenient to assume that agents know each others' valuations precisely, i.e. we assume the values $v_1, v_2, \ldots, v_n$ are public and known to all the agents. Although this assumption is false in practice, by fixing $D$, the parties can mostly infer the valuations of the other parties, as these values will be highly concentrated around their expectations, if the number of agents is sufficiently large. This approach is used in practice on Ethereum, where several services provide tip estimations based on the current network congestion [15].

We assume each agent is capable of deploying a smart contract capable of bidding on their behalf, and that can condition on the smart contracts deployed by the other agents. To formalize this, we use the model of [23]. First, fix some extensive-form representation of the sealed-bid auction, which could be done as follows: (1) choose an arbitrary order of the $n$ agents, (2) construct the $n$-horizon game with agents in the specified order and where each layer sees the corresponding agent make a bid, which is born out in the ensuing subgame, (3) add information sets to ensure agents are not aware of the bids made by the other agents, (4) add utility vectors corresponding to the type of auction (first-price, second-price, etc.). Then, add 'smart contract moves' to the top of the game tree for each player. These moves are special nodes that are syntactic sugar for the larger 'expanded tree' that results from computing all appropriate cuts in the game tree and reattaching them with a node belonging to that player. Expanding these moves in a bottom-up fashion yields a natural way for contracts to condition on the contracts deployed by other agents and is shown to generalize (reverse) Stackelberg equilibria. For more details, we refer to [23], though we trust that the intuitive understanding of 'contracts that depend on other contracts' suffices for the purposes of this work. An auction that is weakly strategically equivalent (i.e. the equilibrium payoffs are equal) to itself with smart contract moves is said to be *Stackelberg resilient*.

We now give our model of the transaction fee mechanism EIP-1559 used by Ethereum since 2021[2]. It generalizes first-price auctions by including a *base fee* $B \geq 0$ that each agent has to pay, which is burned. The base fee is continuously adjusted by the network to balance the demand to ensure each block is half full (in expectation). A first-price auction with $m$ identical items is retained as a special-case when $B = 0$.

**Mechanism 4.** (EIP-1559).

1. *Each party $i \in [n]$ submits a transaction of value $v_i > 0$ and makes a deposit of $B + \tau_i$ funds where $\tau_i \geq 0$ is an optional tip.*
2. *A miner finds a block, and selects a $T \subseteq [n]$ with $|T| = m$ that maximizes $\sum_{i \in T} \tau_i$. If there are multiple such $T$'s, it selects $T$ uniformly at random from all suitable sets.*
3. *Each party $i \in T$ has their transactions included in the block and pays their deposit, in total gaining $v_i - B - \tau_i$ money; each party $j \notin T$ is returned their deposit of $B + \tau_j$ currency and gains 0.*
4. *The miner receives $\sum_{i \in T} \tau_i$ currency.*
5. *The network adjusts the base fee $B$ depending on $m$ and $n$.*

In keeping with auction terminology, moving forward we will refer to the miner as the *auctioneer*. As per the introduction, we will let $n = (1 + \alpha)\, m$ for some *congestion constant* $\alpha > 0$. Let $\varepsilon > 0$ be the smallest unit of currency, and assume it is sufficiently small, i.e. $\varepsilon \ll v_i$, to mostly be ignored in calculations. In practice, on Ethereum, as of 2022, we have $\varepsilon \approx \$10^{-12}$.

We now propose a Stackelberg attack on Mechanism 4: essentially, the leading contract agent commits to paying $2\varepsilon$, conditioned on everyone else committing to bidding $\varepsilon$. In this case, the leading contract agent has their transaction included at almost zero cost, while everyone else enters into a lottery. If anyone does not comply, the leading contract agent instead submits the bid they would have submitted without the contracts, or one slightly higher. This forces each other agent to decide between a lottery and a first-price auction. We will show that when the valuations of the transactions are somewhat concentrated, the agents prefer the lottery over the first-price auction, as they would otherwise have to pay a bid commensurate with their valuation, while in the auction they may receive the item for free.

As a warm-up and ongoing example, we look at the case where there are three agents and two slots up for auction, that is $n = 3$ and $m = 2$. This models a case where there are three buyers who wish to purchase two identical items — we may imagine these big buyers to be exchanges that control large quantities of user transactions, such as Coinbase or Binance, which are juggernauts in the industry [2]. Note that in this example we have $\alpha = \frac{1}{2}$. Suppose that agents $1, 2, 3$ have valuations $0 < v_1 < v_2 < v_3$, respectively. In a first price auction, where the valuations of the respective parties are known, the $m$ agents with the highest valuations only need to outbid the agent with $m + 1$ highest valuation, who is unwilling to bid beyond their valuation and receive negative utility. In our example, agents 2 and 3 will bid slightly higher than the valuation of agent 1, yielding the following utilities: $u_1 = 0$, $u_2 = v_2 - v_1 - \varepsilon$, $u_3 = v_3 - v_1 - \varepsilon$.

[2] In practice, the block size of EIP-1559 is variable and we shall let $m$ denote its maximum possible value. In practice, the base fee would be adjusted to ensure that $\mathbb{E}[n] = m/2$, however the case of $n \leq m$ is not interesting (as all transactions will simply be included) so we take $m$ to be the maximum value and assume $n > m$.

We now equip these three agents with contracts. If the agent with the leading contract can make a credible and enforceable threat with the contract, they may force other agents to accept the lottery at the price $\varepsilon$, thereby guaranteeing the leading agent space an item at price of $2\varepsilon$. The viability of such a threat depends on the agents' valuations. Agents will only comply if their expected utility is higher when they cooperate compared to when the threat is executed.

Consider first the case when agent 3 is the leading contract agent. The contract will commit agent 3 to bidding either $2\varepsilon$, if the two other agents commit to playing $\varepsilon$, or to bidding the usual first price bid of $v_1+\varepsilon$ otherwise. If the contract works, agent 3 enjoys utility $v_3-2\varepsilon$, a better result than the first price auction utility of $v_3 - v_1 - \varepsilon$. The desirable outcome is also clear for agent 1: the lottery case yields utility $\frac{1}{2}(v_1 - \varepsilon)$, which is better than the first price auction utility of 0. Therefore, both 1 and 3 will submit to the contract. Agent 2 will cooperate if the first price utility is lower than the lottery utility, that is if $v_2 - v_1 - \varepsilon < \frac{1}{2}(v_2 - \varepsilon)$, which reduces to $v_1 + \frac{1}{2}\varepsilon > \frac{1}{2}v_2$. The attack would not work if the valuations were less concentrated. If agent 2 is the lead contract holder, the attack works if $v_1+\frac{1}{2}\varepsilon > \frac{1}{2}v_3$, a more stringent concentration requirement. If agent 1 has the leading contract, they may threaten to bid $v_2 + \varepsilon$, knowing they will likely not have to pay it. In this scenario, agent 1 has a credible threat if $v_2 + \frac{1}{2}\varepsilon > \frac{1}{2}v_3$, similar to the conditions for agent 3.

The attack generalizes readily to a larger number of agents, although the requirement on the valuations becomes stronger with more agents. In particular, the attack no longer works if even a single agent has a valuation that is significantly higher than the median. However, the leading contract agent may persuade such agents to participate by promising them a free item from the auction, taking some of the spots intended for the lottery. We denote by $C \subseteq [n]$ the *coalition* of agents (with $|C| = k$ for some $k < m$) who are given free items. This significantly loosens the valuation requirement and allows us to show that the attack works even if $k < m$ of the parties have large valuations. The set $C$ may also be used to capture those agents who are oblivious to the attack, thus modeling the (very realistic) scenario where some of the agents are not aware of the attack and cannot respond accordingly. We have not explicitly accounted for this; doing so would give a slightly stronger bound in the following, but would not fundamentally change the analysis. We now describe the attack in more detail.

**Theorem 5.** *Consider $m$ identical items, and let $\varepsilon \ll v_1 < v_2 < \cdots < v_n$ be the valuations of the $n$ buyers, with $n = (1+\alpha)\,m$ for some $\alpha > 0$. If for some $k < m$ it holds that,*

$$\frac{v_{n-k+1} - B}{v_{n-m}} < \frac{n-k}{n-m}, \tag{2}$$

*then EIP-1559 is not Stackelberg attack resilient.*

*Proof.* Assume that each agent has exactly one transaction, and let agent $i$ be the agent associated with the transaction of valuation $v_i$. Suppose the contract agents are ordered $i_1, i_2, \ldots, i_n$, where $i_1$ is the leading contract agent. Now consider the following contract $A^C_u$, parameterized by an integer $u \in [n]$ that represents the index of the contract order and a set $C \subseteq [n]$ with $i_1 \in C$ and $|C| = k$ for some $k \leq m$.

**Contract 6.** ($A^C_u$).

1. *If $u = n$, play $\varepsilon$.*
2. *If $u < n$, play $v_{n-m} + \varepsilon$ if $v_{i_u} > v_{n-m} + \varepsilon$ and 0 otherwise in every subgame where any agent $i_v$ with $v > u$ does not play the contract $A^C_v$; otherwise play $2\varepsilon$ if $u \in C$, and $\varepsilon$ if $u \notin C$.*

Now suppose the leading contract agent deploys the contract $A^C_1$ with $|C| = k < m$ and $i_1 \in C$. If they are successful, their transaction will be added with certainty for a cost of $2\varepsilon$, thus gaining $v_{i_1} - 2\varepsilon$. Consider the strategy of agent $j$ when every other agent submits, that is, plays Contract 6. If $j \in C$, then clearly for small $\varepsilon$, agent $j$ will comply with the threat. If instead $j \notin C$, they will play Contract 6 to obtain a value of $v_j - \varepsilon$ with probability $\frac{m-k}{n-k}$. If they do not play Contract 6, all agents revert to a first-price auction, in accordance with their contracts. Then agent $j$ can either bid too little to win or bid at least $v_{n-m} + \varepsilon$ to have their transaction included. If $j \leq n - m$, this exceeds their valuation, and will thus prefer Contract 6, as its expected payoff is $\frac{(m-k)(v_i-B-\varepsilon)}{n-k} > 0$. If instead $j > n - m$, they can choose not to comply with the threat to gain $v_j - v_{n-m} - B - 2\varepsilon$ utility. It follows that such an agent will comply with the threat if $v_j - v_{n-m} - B - \varepsilon > \frac{(m-k)(v_j-B-\varepsilon)}{n-k}$, which, when ignoring $\varepsilon$s, solves to $\frac{v_j-B}{v_{n-m}} < \frac{n-k}{n-m}$. But this is guaranteed to hold by Eq. (2), since $v_j \leq v_{n-k+1}$ for any $j$. Thus, complying with the threat is an equilibrium, implying EIP-1559 is not Stackelberg resilient. □

Note that by letting $B = 0$ we obtain a regular first-price auction, and hence Theorem 5 implies that the transaction fee mechanisms of Ethereum, Bitcoin, and most other blockchains are not Stackelberg resilient, regardless of whether there is a base fee or not. We observe that the attack works also for second-price auctions.

**Theorem 7.** *Consider a second-price auction with $m$ identical items, and $n$ buyers, in keeping with Theorem 5. If Eq.* (2) *holds, then the auction is not Stackelberg attack resilient.*

*Proof (Sketch).* Consider the same attack, Contract 6. As we have seen, in the first price setting, bidders who have perfect information must only bid just enough to outbid the $(n-m)^{\text{th}}$ highest player, with a bid of $v_{n-m} + \varepsilon$. In the first price auction, these agents will be charged the amount they bid. In the second price auction, they can either bid their valuation or stick with $v_{n-m} + \varepsilon$. In any case, if they are included, the agent will pay $v_{n-m}$, a slight discount on the $v_{n-m} + \varepsilon$ cost in the first price setting. Thus Contract 6 can be used, and the scenario in which the attack works will look the same. If the attack does not work, agents revert to the equilibrium as it would be without contracts. This equilibrium would have the slightly different, second price cost. Note that, in the proof of Theorem 5, we drop the epsilons that constitute the difference between the first and second price auctions. So by the proof of Theorem 5, second price auctions are also not Stackelberg resilient. □

**Risk Aversion.** It is natural to wonder if the attack will still work if the agents are risk averse. To model risk aversion, agents have some concave utility function $u = U(\cdot)$. If, for example, an agent gets a slot for free at valuation $v_i$, their utility would be defined to be $u = U(v_i)$. For $U(\cdot)$ to be concave, we must have $U((1-p)x + py) \geq (1-p)U(x) + pU(y)$ where $(x, U(x))$ and $(y, U(y))$ are two points on the utility function and $p \in [0, 1]$. Graphically, this implies that any point on the line between $(x, U(x))$ and $(y, U(y))$ is on or below the utility curve. This straight line below the curve traces out the utility of a coin toss with probability $p$ between $U(x)$ and $U(y)$. This models risk aversion because the utility of any outcome based on a coin toss between two outcomes will be on or below the curve, which in turn represents the utility of outcomes that are certain. If we make the assumption that $x = U(x) = 0$ and set $y = v_i$, we have $U(pv_i) \geq p\,U(v_i)$. Note that, in the proof of Eq. (2), we required the

condition, here simplified, that $v_i - v_{n-m-k+1} > pv_i$. If we instead had some concave utility function, this would be $U(v_i - v_{n-m}) > pU(v_i)$. Given that $U(pv_i) \geq pU(v_i)$, the condition found in Eq. (2) is necessary, but not necessarily sufficient, for the contact attack to still be viable. Finding the exact condition requires $U(\cdot)$ to be known.

## 3 Everyone Benefits Except for the Auctioneer

In the following, we will assume that $k = 1$ and that $\varepsilon = 0$. As $k$ increases, more agents with high valuations get free entry when $\varepsilon = 0$. Thus, their relatively high valuations are counted into social welfare. As long as this elite group is relatively small, this will have little impact on the chances of the lottery players, meaning allowing a relatively small $k > 1$ would only increase social welfare.

We define the *price of defiance*, a ratio of sets of equilibrium that is related to the price of anarchy [30]. Let $S$ be the set of all strategy profiles in the game and take two sets $C \subseteq S$, some set of strategies, and $E \subseteq S$, the set of all equilibria of the game. We take the set $C$ to be the set of equilibria after a successful contract attack has been deployed. Define,

$$PoD = \frac{\max_{s \in C} \text{Welf}(s)}{\min_{s \in E} \text{Welf}(s)}. \tag{3}$$

This is the ratio between the best of a subset of possible outcomes and the worst equilibrium. It differs from the price of anarchy in that we compare some subset of strategies, here those that become equilibria due to the introduction of a contract attack, rather the optimal solution, to the game's usual equilibria. We have $PoD \leq PoA$.

Our set $C$ is the set of equilibrium arising from agents having and complying with Contract 4.1. There are up to $n$ equilibria in the set, one for each choice of agent with leading contract. To analyze the price of defiance we will need concentration bounds on the valuations of the parties. Order the players with valuations $v_1 < v_2 < \ldots < v_n$, then $v_i \sim \text{Beta}(i, n+1-i)$. Say a function $f$ is *negligible* if $f(x) = o(x^c)$ for every constant $c \in \mathbb{R}$, i.e. if it grows slower than the inverse of any polynomial. We will make use of the following concentration bound on order statistics from the uniform distribution.

**Lemma 8** (Skorski, [38])**.** *Let $X \sim$ Beta$(\alpha, \beta)$ for $\alpha, \beta > 0$, and define,*

$$v^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+2)}, \qquad c_0 = \frac{|\beta - \alpha|}{(\alpha+\beta)(\alpha+\beta+2)}.$$

*Then for any $\varepsilon > 0$, it holds that,*

$$\Pr\left[|X - \mathbb{E}[X]| > \varepsilon\right] \leq 2\exp\left(-\frac{\varepsilon^2}{2v^2 + 2\varepsilon \max\{v, c_0\}}\right).$$

**Lemma 9.** *Let $X_1, X_2, \ldots, X_n \sim U[0,1]$, and let $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ be the $n$ order statistics. Then,*

$$\left| X_{(i)} - \frac{i}{n+1} \right| = \tilde{O}(1/n), \qquad \textit{for every } i = 1 \ldots n,$$

*except with negligible probability in $n$.*

*Proof.* We make use of Lemma 8 to bound the error term and must therefore first find the relevant values of $v$ and $c_0$. It is a fact that such order statistics have the distribution Beta$(i, n+1-i)$, in particular $\alpha = i$ and $\beta = n+1-i$. Thus, for all values of $i$, we must have $\alpha + \beta = n+1$. It is easy to see that we find the largest value $v^2$ from Lemma 8 when $\alpha = \beta = \frac{n+1}{2}$. This case yields

$$v^2 \leq \frac{\frac{n+1}{2}\frac{n+1}{2}}{\left(\frac{n+1}{2} + \frac{n+1}{2}\right)^2\left(\frac{n+1}{2} + \frac{n+1}{2} + 2\right)} = \frac{1}{4(n+3)}.$$

The value of $c_0$ is largest when the numerator is largest, which is clearly when $|\beta - \alpha| = n-1$. Note that this is a specifically different case from when $v^2$ is largest. When we go on to find the error bounds on specific $v_i$'s, we will refine the bound at this step. Thus, we have the following bounding value,

$$c_0 \leq \frac{n-1}{(n+1)(n+3)}.$$

It is easy to see that $c = \max\{v, c_0\} = c_0$. Thus, we can write down the bound for any $i$,

$$\begin{aligned}
\Pr\left[\left|X_{(i)} - \mathbb{E}[X_{(i)}]\right| > \delta\right] &< 2\exp\left(-\frac{\delta^2}{2v^2 + 2c\delta}\right) \\
&\leq 2\exp\left(-\frac{\delta^2}{2\frac{1}{4(n+3)} + \frac{2\delta(n-1)}{(n+1)(n+3)}}\right) \\
&= 2\exp\left(-\frac{\delta^2\, 2(n+3)(n+1)}{(n+1) + 4\delta(n-1)}\right) \\
&< 2\exp\left(-\frac{\delta^2\, 2n^2}{(n+1) + 4\delta n}\right) \\
&\approx 2\exp\left(-\frac{\delta^2\, 2n}{1 + 4\delta}\right) \\
&= 2\exp(-\Omega(\delta n)).
\end{aligned}$$

If we take $\delta = \frac{\log^2 n}{n+1} = \tilde{O}(1/n)$, we obtain the bound,

$$\Pr\left[\left|X_{(i)} - \mathbb{E}[X_{(i)}]\right| > \delta\right] < 2\exp(-\omega(\log n)), \tag{4}$$

which is negligible in $n$. We conclude with a union bound on all $n$ valuations. □

**Theorem 10.** *For uniformly distributed valuations, the price of defiance is at least $1 + \alpha - o(1)$, except with probability negligible in $n$.*

*Proof.* It is easy to see that the maximal choice $s \in C$ occurs when the agent with the highest valuation has the contract. There is only one choice for equilibrium $s \in C$. Thus we have,

$$\begin{aligned}
PoD = \frac{\max_{s \in C} \text{Welf}(s)}{\min_{s \in E} \text{Welf}(s)} &= \frac{\left(\sum_{j=1}^{n-1} \frac{m-1}{n-1}(v_j - \varepsilon)\right) + v_n - 2\varepsilon}{\left(\sum_{i=n-m+1}^{n} v_i - v_{n-m} - \varepsilon\right)} \\
&\approx \frac{\frac{m-1}{n-1}\left(\sum_{j=1}^{n-1} v_j\right) + v_n}{\left(\sum_{i=n-m+1}^{n} v_i\right) - m v_{n-m}}.
\end{aligned} \tag{5}$$

If the contract attack works, that is if the valuations are in keeping with the in condition from Theorem 5, we have $PoD > 1$. This can be seen mathematically by substituting the condition into the denominator of Eq. (5) above. Intuitively, given that the threat is just the usual first price auction, the other agents will acquiesce only if their utility would be higher in the lottery. Thus, total lottery welfare, the numerator, must be higher than the auction, the denominator, leading to a $PoD > 1$ in the general case. Each $v_i$ is the $i^{th}$ order statistic of a uniformly distributed random variable, that is $v_i = X_{(i)}$ where $X_i$

is sampled i.i.d. from the uniform distribution on $[0,1]$. By linearity of expectation, we have that,

$$\mathbb{E}\left[\sum_{i=n-m+1}^{n} v_i\right] = \sum_{i=n-m+1}^{n} \frac{i}{n+1} = \frac{1}{n+1}\left(\sum_{i=0}^{n} i - \sum_{k=0}^{n-m} k\right) = \frac{n}{2} - \frac{(n-m)(n-m+1)}{2(n+1)},$$

and that,

$$\mathbb{E}\left[\sum_{j=1}^{n-1} v_j\right] = \sum_{j=1}^{n-1} \frac{j}{n+1} = \frac{(n-1)n}{2(n+1)}.$$

We proceed to lower bound $PoD$ using Lemma 9 to yield,

$$PoD \geq \frac{\frac{m-1}{n-1}\left(\frac{(n-1)n}{2(n+1)} - (n-1)\delta\right) + \frac{n}{n+1} - \delta}{\frac{n}{2} - \frac{(n-m)(n-m+1)}{2(n+1)} + m\delta - m\left(\frac{n-m}{n+1} + \delta\right)} = \frac{n(m+1) - 2m(n+1)\delta}{m(m+1) + 4m(n+1)\delta}$$

We now condition on the errors of the valuations being bounded by $\delta = \frac{(m+1)\log^2 n}{2m(n+1)}$, which we know to happen except with negligible probability by Lemma 9. Then we obtain the following bound,

$$PoD \geq \frac{n - \log^2(n)}{m + \log^2(n)} = 1 + \alpha - o(1),$$

as desired. □

Arguably, this suggests that lotteries should be used as transaction mechanisms when the valuations are believed to be of similar size. In the $n = 3$, $m = 2$ case, we have

$$PoD = \frac{\frac{v_1}{2} + \frac{v_2}{2} + v_3 - 3\varepsilon}{v_2 + v_3 - 2v_1}$$

If the condition for the contract to work from the example in Section 2 holds, that is, if $v_1 + \frac{1}{2}\varepsilon > \frac{1}{2}v_2$, the ratio becomes

$$PoD > \frac{\frac{v_1}{2} + \frac{v_2}{2} + v_3 - 3\varepsilon}{v_3 + \varepsilon},$$

which is clearly larger than one.

It is important to note that while the attack benefits all the agents with transactions, it is detrimental to the auctioneer, who loses essentially all of their revenue. Auctioneers that find themselves subject to such an attack might respond by not allowing the smart contracts to be deployed, but this could be remedied if agents use a different blockchain to deploy the attack. Auctioneers may also find themselves obligated to include the contract moves due to a staking scheme [9]. Continuing with our $n = 3$, $m = 2$ example, we readily see the auctioneer will earn $2(v_1 + \varepsilon)$ in the auction case. If the contract attack is successfully executed, the auctioneer income will be $3\varepsilon$, $2\varepsilon$ from the leading contract holder, regardless of which agent this is, and $\varepsilon$ from the winner of the lottery. Thus, almost all the revenue is lost; the auctioneer will miss out on $2v_1 - \varepsilon$ income. If there were a base fee and all agents had a valuation larger than said base fee, i.e. $v_1 > B$, the first price revenue would be $2(v_1 + \varepsilon - B)$. The lottery revenue will continue to be $3\varepsilon$ and the income lost to the attack will be $2(v_1 - B) - \varepsilon$.

## 4 The Attack Works for Natural Distributions

In this section, we show that the conditions required for the attack are satisfied with high probability under reasonable assumptions. We will assume that $B = 0$. The results obtained are qualitatively similar when the valuations are much larger than the base fee.

We continue with our illustration of the $n = 3$, $m = 2$ case, now assuming that the players have valuations that are uniformly distributed on $[0,1]$. As before, we have three valuations $v_1 < v_2 < v_3$ and we can now make use of the distribution. The ordered valuations in are order statistics, that is $v_i = X_{(i)}$ where all $X_i$ are sampled i.i.d. from the uniform distribution on $[0,1]$. Using the well known fact that order statistics on the uniform distribution follow specific beta distributions, we get the following distributions and their expectations: $v_1 \sim \text{Beta}(1,3)$ yielding, $\mathbb{E}[v_1] = \frac{1}{4}$; $v_2 \sim \text{Beta}(2,2)$, yielding $\mathbb{E}[v_2] = \frac{1}{2}$; and $v_3 \sim \text{Beta}(3,1)$, with $\mathbb{E}[v_3] = \frac{3}{4}$. Note that the variance for all the distributions is $\text{Var}[v_i] \leq 1/20$ and we will not take it into account moving forward. In the first price auction, we can see that if agents 2 and 3 bid just enough to outbid agent 1, i.e. $\frac{1}{4}+\varepsilon$, they will secure their slots as cheaply as possible. So in the first price auction the players will have the expected utilities $\mathbb{E}[u_1] = 0$, $\mathbb{E}[u_2] = \frac{1}{4} - \varepsilon$, and $\mathbb{E}[u_3] = \frac{1}{2} - \varepsilon$.

If agent 1 has the leading contract, they can threaten to outbid agent 2 with a bid of $\frac{1}{2} + \varepsilon$. If the threat were to be carried out, agent 2 would lose their slot and receive utility 0 and agent 3, secure in the top spot, but given they outbid agent 2, will receive $\frac{1}{4} - \varepsilon$. If agents 2 and 3 comply with the threat, i.e. bid $\varepsilon$ and enter a lottery, they will have expected utilities $\frac{1}{4} - \frac{\varepsilon}{2}$ and $\frac{3}{8} - \frac{\varepsilon}{2}$, respectively. These utilities are more desirable than ignoring the threat, and the attack can be executed. Agent 1 will enjoy expected utility $\frac{1}{4} - 2\varepsilon$. Note that agents 1 and 2 have higher utility than they would have had in the first price auction, but agent 3 is hurt by the attack.

If agent 2 has the leading contract, their best threat is outbidding agent 1 with a bid of $\frac{1}{4} + \varepsilon$. This is no threat at all as it simple coincides with their first price strategy. If instead agent 3 has the leading contract, we once again have a viable attack. Since agent 3 already outbids the others, their contract endowed strategy is more a proposition for mutual benefit than a greedy attack. If the other two parties enter into a lottery at price $\varepsilon$ and agent 3 bids $2\varepsilon$, we have expected utilities $\mathbb{E}[u_1] = \frac{1}{8} - \frac{\varepsilon}{2}$,$\mathbb{E}[u_2] = \frac{1}{4} - \frac{\varepsilon}{2}$, and $\mathbb{E}[u_3] = \frac{3}{4} - 2\varepsilon$. It can be easily seen that everyone benefits in this situation and the attack will work. It is an easy calculation to find that $PoD \approx 3/2$. Regardless of which agent has the leading contract, if the attack works, the total tip paid to the auctioneer will be $3\varepsilon$. In the first price auction, the expected auctioneer payout is $\frac{1}{2} + 2\varepsilon$. The difference constitutes an almost complete loss of revenue.

**Lemma 11** (Xu, Mei, Miao, [43])**.** *Let $X_1, X_2, \ldots, X_n \sim U(0,1)$ be i.i.d. Let $i < j$ and define $R_{ij} = \frac{X_{(j)}}{X_{(i)}}$ and let $f(\cdot)$ be its density function with support $[1,\infty)$. Then for every $r \geq 1$,*

$$f(r) = \frac{n!(r-1)^{j-i-1}}{(i-1)!(j-i-1)!(n-j)!r^j} \int_0^1 (1-u)^{j-1} u^{n-j}\, \mathrm{d}u.$$

**Theorem 12.** *Suppose $n$ buyers participate in an auction of $m$ identical items where $n = (1+\alpha)m > m+1$. If the valuations of the items are sampled uniformly at random and $0 \leq \alpha < 0.53$, then first-price auctions are not Stackelberg resilient, except with probability negligible in $m$.*

*Proof.* We will show that Eq. (2) holds except with probability $\mathsf{negl}(n)$. Suppose w.log. that the valuations are sampled uniformly

from $[0,1]$ and let $v_1 < v_2 < \cdots < v_n$ be their valuations. The value $v_i$ equals the $i^{\text{th}}$ order statistic, the distribution of which is well-known for uniform values. We are interested in the ratio $R = v_{n-k+1}/v_{n-m}$, so let $f(\cdot)$ be its density function. Let $k = m\delta$ for some $0 < \delta < 1$. By Lemma 11, noting that we have $j = (1+\alpha-\delta)m+1, i = \alpha m$, we get that,

$$\begin{aligned} f(r) &= \frac{n!\,(r-1)^{m-1}}{(\alpha m-1)!((1-\delta)m)!\,r^n}\int_0^1 (1-u)^{(1+\alpha-\delta)\,m-1}u^{\delta m-1}\,\mathrm{d}u \\ &= \frac{((1+\alpha-\delta)\,m)!}{((1-\delta)\,m)!\,(\alpha m-2)!}\frac{(r-1)^{(1-\delta)\,m}}{r^n}. \end{aligned}$$

We denote by $H(p) = -p\lg p-(1-p)\lg(1-p)$, the binary entropy function, defined on $[0,1]$. Note that $H(p) \le 1$ for every $p \in [0,1]$. A useful upper bound is given by the following.

$$H(x) \le 2\sqrt{x\,(1-x)}. \tag{6}$$

The binary entropy function is useful because it allows us to upper bound the binomial coefficient as follows.

$$\binom{n}{k} \le 2^{nH(k/n)} \tag{7}$$

We bound the probability that Eq. (2) does not hold as follows.

$$\begin{aligned} &\Pr\left[R' \ge \frac{n-k}{n-m}\right] = \int_{\frac{1+\alpha-\delta}{\alpha}}^{\infty} f_R(r)\,\mathrm{d}r \\ &= \frac{((1+\alpha-\delta)\,m)!}{((1-\delta)\,m)!\,(\alpha m-2)!}\int_{\frac{1+\alpha-\delta}{\alpha}}^{\infty}\frac{(r-1)^{(1-\delta)\,m}}{r^n}\,\mathrm{d}r \\ &\le \frac{\alpha}{\alpha+\delta}\binom{(1+\alpha-\delta)\,m}{\alpha m}\left(\frac{1+\alpha-\delta}{\alpha}\right)^{1-(\alpha+\delta)\,m} \end{aligned}$$

We now apply Eq. (7) and collect the terms in the exponent.

$$\le \frac{\alpha}{\alpha+\delta}\exp\left(H\left(\frac{\alpha}{1+\alpha-\delta}\right)(1+\alpha-\delta)\,m + \log\left(\frac{1+\alpha-\delta}{\alpha}\right)(1-(\alpha+\delta)m)\right)$$

We use the fact that $H(p) \le 2\sqrt{p(1-p)}$ as per Eq. (6) to obtain,

$$\le \frac{\alpha}{\alpha+\delta}\exp\left(\log\left(\frac{1+\alpha\delta}{\alpha}\right) + m\left[2\sqrt{\alpha(1-\delta)}-(\alpha+\delta)\log\left(\frac{1+\alpha\delta}{\alpha}\right)\right]\right).$$

We note that the exponent is negative for sufficiently large $m$, and hence the probability is negligible if,

$$1+\alpha-\delta-(\alpha+\delta)\log\left(\frac{1+\alpha\delta}{\alpha}\right) < 0.$$

Which solves to $0 < \alpha < 0.529914$ for $\delta = 0.69$. □

**Lemma 13** (Adler, [1])**.** *Let $X_1, X_2, \ldots X_n$ be i.i.d. Pareto distributed with parameter $p > 0$. Let $i < j$ and define $R_{ij} = \frac{X_{(j)}}{X_{(i)}}$ and let $f(\cdot)$ be its density function with support $[1,\infty)$. Then for every $r \ge 1$,*

$$f(r) = \frac{p\,(n-i)!}{(j-i-1)!(n-j)!}\left(1-\frac{1}{r^p}\right)^{j-i-1}\frac{1}{r^{p(n-j+1)+1}}.$$

**Theorem 14.** *Suppose $n$ buyers participate in an auction of $m$ identical items where $n = (1+\alpha)\,m$ for some $\alpha > 0$. If the valuations of the items are sampled according to a Pareto distribution with parameter $p > 1$ and $0 \le \alpha \le \alpha(p) < 0.69$, then first-price auctions are not Stackelberg resilient, except with probability negligible in $m$.*

*Proof.* Suppose for the sake of the argument that $n$ is even, and let $C$ be the $m\delta$ players with the largest valuations for some constant $0 < \delta < 1$. Let $R = (v_{n-k+1}/v_{n-m})$ and let $f(\cdot)$ be its density function. By Lemma 13, it is given by,

$$\begin{aligned} f(r) &= \frac{pm!}{(m-k-2)!(k-1)!}\left(1-\frac{1}{r^p}\right)^{m-k-2}\frac{1}{r^{pk+1}} \\ &= pk(m-k-1)(m-k)\binom{m}{k}\left(1-\frac{1}{r^p}\right)^{m-k-2}\frac{1}{r^{pk+1}}. \end{aligned}$$

We bound the probability that Eq. (2) does not hold as follows.

$$\begin{aligned} &\Pr\left[R \ge \frac{n-k}{n-m}\right] \\ &= pk(m-k-1)(m-k)\binom{m}{k}\int_{\frac{1+\alpha-\delta}{\alpha}}^{\infty}\frac{\left(1-\frac{1}{r^p}\right)^{m-k-2}}{r^{pk+1}}\,\mathrm{d}r \\ &\le pk(m-k-1)(m-k)\binom{m}{k}\int_{\frac{1+\alpha-\delta}{\alpha}}^{\infty}\frac{1}{r^{pk+1}}\,\mathrm{d}r \\ &= m((1-\delta)m-1)(1-\delta)\binom{m}{\delta m}\left(\frac{1+\alpha-\delta}{\alpha}\right)^{-p\delta m} \end{aligned}$$

We now bound the binomial coefficient using Eq. (7) and collect the terms in the exponent.

$$\le m((1-\delta)m-1)(1-\delta)\exp\left(m\left[H(\delta)-\delta p\log\left(\frac{1+\alpha-\delta}{\alpha}\right)\right]\right)$$

We note that the exponent is negative, and hence the function negligible, if the following inequality is satisfied.

$$\delta p\log\left(\frac{1+\alpha-\delta}{\alpha}\right) > H(\delta).$$

By Eq. (6), it suffices then to establish the following bound.

$$\delta p\log\left(\frac{1+\alpha-\delta}{\alpha}\right) > 2\sqrt{\delta(1-\delta)}.$$

We now let $\delta = \frac{5}{p^2+4}$, and note that this inequality is satisfied for any $p > 1$ whenever the following inequality holds.

$$0 < \alpha < \frac{p^2-1}{(4+p^2)\left(\exp\left(\frac{2\sqrt{\frac{p^2-1}{p^2}}}{\sqrt{5}}\right)-1\right)}$$

Denote the rhs by $\alpha(p)$. Note that $\alpha(p) > 0$ for any $p > 1$ and evaluates to $\frac{1}{2}(\coth(1/\sqrt{5})-1) \approx 0.69$ in the limit as $p \to \infty$. □

The Pareto distribution, which follows the 80/20 rule, is the more natural distribution in this context. It is widely used in economics and it makes intuitive sense that transactions, and therefore valuations, would tend to be small, with a tail of rarer, but large transactions. The uniform distribution can be seen as a lower bound, because the real distribution would have fewer high valuation transactions, and this can only favor the attack.

## Acknowledgments

NS was funded by VILLUM FONDEN under the Villum Kann Rasmussen Annual Award in Science and Technology under grant agreement no 17911.

## References

[1] André Adler, 'Limit theorems for arrays of ratios of order statistics', *Bull. Inst. Math. Acad. Sin.(NS)*, **33**(4), 327, (2005).

[2] Carol Alexander, Daniel F Heck, and Andreas Kaeck, 'The role of binance in bitcoin volatility transmission', *Applied Mathematical Finance*, **29**(1), 1–32, (2022).

[3] Rabah Amir and Isabel Grilo, 'Stackelberg versus cournot equilibrium', *Games and Economic Behavior*, **26**(1), 1–21, (1999).

[4] Christian Badertscher, Juan Garay, Ueli Maurer, Daniel Tschudi, and Vassilis Zikas, 'But why does it work? a rational protocol design treatment of bitcoin', in *Advances in Cryptology – EUROCRYPT 2018*, eds., Jesper Buus Nielsen and Vincent Rijmen, pp. 34–65, Cham, (2018). Springer International Publishing.

[5] Yu Bai, Chi Jin, Huan Wang, and Caiming Xiong, 'Sample-efficient learning of stackelberg equilibria in general-sum games', *Advances in Neural Information Processing Systems*, **34**, 25799–25811, (2021).

[6] Tamer Basar and Hasan Selbuz, 'Closed-loop stackelberg strategies with applications in the optimal control of multilevel systems', *IEEE Transactions on Automatic Control*, **24**(2), 166–179, (1979).

[7] Kaushik Basu, 'Stackelberg equilibrium in oligopoly: an explanation based on managerial incentives', *Economics Letters*, **49**(4), 459–464, (1995).

[8] Michael Bloem, Tansu Alpcan, and Tamer Basar, 'A stackelberg game for power control and channel allocation in cognitive radio networks', in *1st International ICST Workshop on Game theory for Communication networks*, (2010).

[9] Vitalik Buterin. Slasher: A punitive proof-of-stake algorithm, 2014.

[10] David Chaum, 'Verification by anonymous monitors', in *Advances in Cryptology: A Report on CRYPTO 81, CRYPTO 81, IEEE Workshop on Communications Security, Santa Barbara, California, USA, August 24-26, 1981*, ed., Allen Gersho, pp. 138–139. U. C. Santa Barbara, Dept. of Elec. and Computer Eng., ECE Report No 82-04, (1981).

[11] David Chaum, 'Blind signatures for untraceable payments', in *Advances in Cryptology: Proceedings of CRYPTO '82, Santa Barbara, California, USA, August 23-25, 1982*, eds., David Chaum, Ronald L. Rivest, and Alan T. Sherman, pp. 199–203. Plenum Press, New York, (1982).

[12] Hao Chung and Elaine Shi, *Foundations of Transaction Fee Mechanism Design*, 3856–3899.

[13] Edward H. Clarke, 'Multipart pricing of public goods', *Public Choice*, **11**(1), 17–33, (1971).

[14] Vincent Conitzer and Tuomas Sandholm, 'Computing the optimal strategy to commit to', in *Proceedings of the 7th ACM Conference on Electronic Commerce*, EC '06, p. 82–90, New York, NY, USA, (2006). Association for Computing Machinery.

[15] Anil Donmez and Alexander Karaivanov, 'Transaction fee economics in the ethereum blockchain', *Economic Inquiry*, **60**(1), 265–292, (2022).

[16] Dinesh Garg and Y Narahari, 'Mechanism design for single leader stackelberg problems and application to procurement auction design', *IEEE Transactions on Automation Science and Engineering*, **5**(3), 377–393, (2008).

[17] Dinesh Garg and Yadati Narahari, 'Design of incentive compatible mechanisms for stackelberg problems', in *Internet and Network Economics: First International Workshop, WINE 2005, Hong Kong, China, December 15-17, 2005. Proceedings 1*, pp. 718–727. Springer, (2005).

[18] Noortje Groot, Bart De Schutter, and Hans Hellendoorn, 'Reverse stackelberg games, part i: Basic framework', in *2012 IEEE International Conference on Control Applications*, pp. 421–426, (2012).

[19] Noortje Groot, Bart De Schutter, and Hans Hellendoorn, 'Reverse stackelberg games, part ii: Results and open issues', in *2012 IEEE International Conference on Control Applications*, pp. 427–432. IEEE, (2012).

[20] Noortje Groot, Bart De Schutter, and Hans Hellendoorn, 'Toward system-optimal routing in traffic networks: A reverse stackelberg game approach', *IEEE Transactions on Intelligent Transportation Systems*, **16**(1), 29–40, (2014).

[21] Noortje Groot, Georges Zaccour, and Bart De Schutter, 'Hierarchical game theory for system-optimal control: Applications of reverse stackelberg games in regulating marketing channels and traffic routing', *IEEE Control Systems Magazine*, **37**(2), 129–152, (2017).

[22] Theodore Groves, 'Incentives in teams', *Econometrica*, **41**(4), 617–631, (1973).

[23] Mathias Hall-Andersen and Nikolaj I. Schwartzbach, 'Game theory on the blockchain: A model for games with smart contracts', in *Algorithmic Game Theory*, eds., Ioannis Caragiannis and Kristoffer Arnsfelt Hansen, pp. 156–170, Cham, (2021). Springer International Publishing.

[24] Debarun Kar, Thanh H Nguyen, Fei Fang, Matthew Brown, Arunesh Sinha, Milind Tambe, and Albert Xin Jiang, 'Trends and applications in stackelberg security games', *Handbook of dynamic game theory*, 1–47, (2017).

[25] Thomas Kerber, Aggelos Kiayias, and Markulf Kohlweiss, 'Kachina – foundations of private smart contracts', *2021 IEEE 34th Computer Security Foundations Symposium (CSF)*, 1–16, (2021).

[26] Aggelos Kiayias, Alexander Russell, Bernardo David, and Roman Oliynykov, 'Ouroboros: A provably secure proof-of-stake blockchain protocol', in *Advances in Cryptology – CRYPTO 2017*, eds., Jonathan Katz and Hovav Shacham, pp. 357–388, Cham, (2017). Springer International Publishing.

[27] Kai A Konrad and Wolfgang Leininger, 'The generalized stackelberg equilibrium of the all-pay auction with complete information', *Review of Economic Design*, **11**, 165–174, (2007).

[28] Dmytro Korzhyk, Vincent Conitzer, and Ronald Parr, 'Complexity of computing optimal stackelberg strategies in security resource allocation games', in *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 24, pp. 805–810, (2010).

[29] Dmytro Korzhyk, Zhengyu Yin, Christopher Kiekintveld, Vincent Conitzer, and Milind Tambe, 'Stackelberg vs. nash in security games: An extended investigation of interchangeability, equivalence, and uniqueness', *Journal of Artificial Intelligence Research*, **41**, 297–327, (2011).

[30] Elias Koutsoupias and Christos Papadimitriou, 'Worst-case equilibria', *Comput. Sci. Rev.*, **3**(2), 65–69, (may 2009).

[31] Ron Lavi, Or Sattath, and Aviv Zohar, 'Redesigning bitcoin's fee market', *ACM Trans. Econ. Comput.*, **10**(1), (may 2022).

[32] Thomas Nedelec, Jules Baudet, Vianney Perchet, and Noureddine El Karoui, 'Adversarial learning for revenue-maximizing auctions', *arXiv preprint arXiv:1909.06806*, (2019).

[33] Thomas Nedelec, Clement Calauzenes, Vianney Perchet, and Noureddine El Karoui, 'Robust stackelberg buyers in repeated auctions', in *International Conference on Artificial Intelligence and Statistics*, pp. 1342–1351. PMLR, (2020).

[34] Tim Roughgarden, 'Stackelberg scheduling strategies', in *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*, STOC '01, p. 104–113, New York, NY, USA, (2001). Association for Computing Machinery.

[35] Tim Roughgarden, 'Transaction fee mechanism design', in *Proceedings of the 22nd ACM Conference on Economics and Computation*, EC '21, p. 792, New York, NY, USA, (2021). Association for Computing Machinery.

[36] Hanif D Sherali, Allen L Soyster, and Frederic H Murphy, 'Stackelberg-nash-cournot equilibria: characterizations and computations', *Operations Research*, **31**(2), 253–276, (1983).

[37] Arunesh Sinha, Fei Fang, Bo An, Christopher Kiekintveld, and Milind Tambe, 'Stackelberg security games: Looking beyond a decade of success'. IJCAI, (2018).

[38] Maciej Skorski, 'Bernstein-type bounds for beta distribution', *Modern Stochastics: Theory and Applications*, **10**(2), 211–228, (2023).

[39] Mohammad Amin Tajeddini, Hamed Kebriaei, and Luigi Glielmo, 'Decentralized hierarchical planning of pevs based on mean-field reverse stackelberg game', *IEEE Transactions on Automation Science and Engineering*, **17**(4), 2014–2024, (2020).

[40] William Vickrey, 'Counterspeculation, auctions, and competitive sealed tenders', *The Journal of Finance*, **16**(1), 8–37, (1961).

[41] Heinrich von Stackelberg, *Marktform und Gleichgewicht*, Verlag von Julius Springer, 1934.

[42] Gavin Wood, 'Ethereum: A secure decentralised generalised transaction ledger', *Ethereum project yellow paper*, **151**, 1–32, (2014).

[43] Shoufang Xu, Changlin Mei, and Yu Miao, 'Limit theorems for ratios of order statistics from uniform distributions', *Journal of Inequalities and Applications*, **2019**(1), 303, (Nov 2019).