https://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Vol. 69, No. 11, November 2023, pp. 6508–6524
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

StableFees: A Predictable Fee Market for Cryptocurrencies

Soumya Basu,[a] David Easley,[b,*] Maureen O'Hara,[c] Emin Gün Sirer[d]

[a] Department of Computer Science, Cornell University, Ithaca, New York 14850; [b] Departments of Information Science and Economics, Cornell University, Ithaca, New York 14850; [c] College of Business, Cornell University, Ithaca, New York 14850; [d] Ava Labs, New York, New York 10036

*Corresponding author

Contact: sb2352@cornell.edu (SB); dae3@cornell.edu, https://orcid.org/0000-0002-6405-4134 (DE); mo19@cornell.edu, https://orcid.org/0000-0003-2563-7748 (MO); egs@systems.cs.cornell.edu (EGS)

Received: July 11, 2021
Revised: March 13, 2022
Accepted: May 19, 2022
Published Online in Articles in Advance: March 31, 2023

https://doi.org/10.1287/mnsc.2023.4735

Abstract. Blockchain-based cryptocurrencies must solve the problem of assigning priorities to competing transactions. The most widely used mechanism involves each transaction offering a fee to be paid once the transaction is processed, but this discriminatory price mechanism fails to yield stable equilibria with predictable prices. We propose an alternate fee setting mechanism, StableFees, that is based on uniform price auctions. We prove that our proposed protocol is free from manipulation by users and miners as the number of users and miners increases and show empirically that gains from manipulation are small in practice. We show that StableFees reduces the fees paid by users and reduces the variance of fee income to miners. Data from December 2017 show that, if implemented, StableFees could have saved Bitcoin users $272,528,000 USD in transaction fees while reducing the variance of miner's fee income, on average, by a factor of 7.4. We argue that our fee protocol also has important social welfare and environmental benefits.

History: Accepted by Agostino Capponi, Special Section of *Management Science*: Blockchains and Crypto Economics.
Supplemental Material: The data files are available at https://doi.org/10.1287/mnsc.2023.4735.

Keywords: cryptocurrency • blockchain • transaction fees

1. Introduction

Almost all decentralized cryptocurrencies use the same basic mechanism to prioritize transactions. A user who wants their transaction included in the blockchain attaches a fee to the transaction. This fee serves as a bid for block space. The miner who is building the block then chooses which transactions to include in their block and collects the respective transaction fees from the included users. This mechanism gives the miner an obvious incentive to select the highest fee transactions, and it plays a crucial role in rewarding miners for processing transactions.

It is useful to note that, although fees were envisioned in the original Bitcoin protocol, that protocol did not describe exactly how they would work. The fee mechanisms used in various blockchain environments have evolved over time to allow users to signal, and pay for, their desire to have their transactions included in the blockchain before others' transactions are included. The initial mechanism took into account a combination of factors including the age of the coins being spent ("bitcoin days destroyed"), but has by now been almost universally replaced by a discriminatory price auction mechanism of the kind described previously. There is no reason to expect this emergent payment mechanism to result in an efficient, stable, or predictable fee market and that appears to be the case.

The current fee mechanism produces very volatile prices for block space. For example, in Bitcoin, the average daily fee paid in December 2020 ranged from $2.72 to $12.05. In December 2017, during a period of heavy trading activity, the average daily fee ranged from $5.82 to $61.44. This volatility makes it difficult for users to decide what fee to attach to a transaction, and this hampers usability of the crypotcurrency. Users who bid too high have overpaid for their transaction to get confirmed, whereas users who underbid may not be included in the block even if their transaction is more valuable than some transactions in the block. This can cause block space to be inefficiently allocated; low value transactions may be confirmed, whereas high value transactions are still pending.

To provide insight into why this volatility occurs, we note that the current cryptocurrency fee market shares important features with discriminatory price auctions where users act as bidders and miners act as auctioneers.[1] Although there are key differences between the fee market and auctions, prior experience with discriminatory price auctions suggests why the current fee market is unstable. In a typical discriminatory price auction for multiple, identical items, the highest bidder pays his bid and gets the first item, the second highest bidder pays his bid and gets the second item, and so on until either items or bidders are exhausted. These auctions do

not have a dominant strategy equilibrium. Although efficient Bayes-Nash equilibria exist if there are multiple identical items and symmetric bidders, the equlibria require users to model accurately the values of other user's bids, which is a difficult, if not impossible, task in the cryptocurrency environment. In Bitcoin, we see this difficulty reflected in the behavior of users who, instead of revealing the full utility of their transaction in the fee they bid, prefer to bid low at first and only increase their bid if they are waiting too long.[2] The strategic bidding induced by the first-price-like nature of the current mechanism is an underlying cause of fee instability and inefficiently allocated block space in cryptocurrencies.

Adapting insights from uniform price auctions to the cryptocurrency setting could potentially ameliorate the previous drawbacks.[3] In a uniform price auction, with the price determined by the highest losing bid, truthful bidding is a dominant strategy. Therefore, rational users do not need to strategize; instead, they can bid a fee equal to their value of having their transaction included in the block. This would make it possible for miners to more efficiently allocate block space, resulting in a more usable and valuable cryptocurrency.

There are several challenges to modifying existing auction theory for application to the cryptocurrency fee market. First, the fee mechanism can only control what each user will pay for their transaction (given their bid) and the reward earned by the miner given the set of transactions they include in the block. Miners are allowed to place any transactions they want in a block, including fee-paying transactions created on the fly after observing users' bids. Thus, miners, unlike trusted auctioneers, are able to manipulate the fee mechanism. Second, users' payments cannot depend on fees attached to transactions not included in the block as the details of these unused transactions are not externally verifiable. Thus, payments must depend only on accepted bids and so truthful bidding cannot be a dominant strategy. These constraints, inherent to most decentralized cryptocurrencies, prevent us from directly using a standard uniform price auction.[4]

In this paper, we present StableFees—a mechanism inspired by second price auctions. StableFees provides provable nonmanipulation guarantees for both users and miners. We show that as the number of users increases, users' gain from bidding strategically converges to zero. This result demonstrates that in large markets users have a nearly dominant strategy of bidding truthfully. Additionally, we show that miners' gain from manipulating the transactions they include in a block also converges to zero as adoption increases. Taken together, these results demonstrate that truthful bidding by users and nonmanipulation by miners is an ϵ-Bayes Nash equilibrium. We demonstrate these results both theoretically and through simulations on real transaction fee distributions. An empirical analysis of the Ethereum and Bitcoin blockchains suggests that users could have saved $13.2 million and $273 million, respectively, if StableFees was implemented during December 2017. Empirically comparing StableFees with other fee mechanisms using demand curves based on real Bitcoin data, we show that StableFees provides 49%–103% more welfare than comparable schemes. Finally, we argue that our mechanism, by reducing unnecessary mining, is more environmentally sustainable.

1.1. Prior Work

Recent literature has considered alternative protocols for selecting which transactions are placed on the blockchain. The work most closely related to ours is the monopolistic miner protocol proposed in Lavi et al. (2017).[5] They propose a protocol in which the winning miner decides how many transactions to put into the block and charges all of them the lowest fee proposed by any transaction he placed in that block. There are fundamental differences between our approaches stemming from our goals and setup. Lavi et al. (2017) assume a single monopolistic miner and strive to maximize revenue from fees at a cost of lower social welfare. In contrast, our work explicitly targets maximizing social welfare and operates with many miners. In their system, the monopolistic miner is incentivized to leave transactions offering positive fees out of the block even if there is space in the block as including them reduces the uniform price he can charge (such behavior is consistent with the findings of Lehar and Parlour (2020)). This behavior is important for maximizing miner revenue, but we believe that the first criterion for a viable protocol should be to use the blockchain efficiently, as otherwise users are discouraged from participating. Their nonmanipulation result is stronger than the one we obtain from our mechanism as we only obtain declining gain from manipulation as the system grows, but their result comes at a cost of lower social welfare. Finally, in both our protocol and the protocol proposed by Lavi et al. (2017), users' incentive to behave strategically vanishes as the number of users grows.

Buterin (2018) proposes an alternative mechanism (EIP-1559), and Roughgarden (2020) provides an economic analysis of this mechanism. This alternative mechanism is based on miners estimating, and dynamically adjusting, a single fee that is charged uniformly to all transactions within a block, coupled with dynamically varying the block size to accommodate demand. This approach differs from ours in a few key ways. First, it does not aim to maximize social welfare, and instead adopts heuristics to modify two independent variables, fees, and block size. Modifying fees, similar to our work, will maximize transactions cleared subject to any desired block size constraint, determined by any desirable mechanism. However, because block size is a primary determinant of security and centralization, we

believe it is prudent to decouple its management from the fee mechanism. Second, Buterin's approach assumes that the demand curve is relatively stable, so that all transactions meeting the fee threshold can be included in the next block. If the demand curve could be inferred accurately such that all transactions whose utility exceeds the block fee can always be accommodated, then Buterin's proposal would have no incentive issues. However, inferring demand curves is difficult in adversarial, Byzantine environments, which is why auction mechanisms are used. Finally, this approach has not been proven to be resistant to manipulation by users and miners. Indeed, there is an uncoordinated attack where users may artificially manipulate the demand curve to simulate demand spikes, a scenario where Buterin's proposal has comparatively weak guarantees (Tefagh 2021). If it is not resistant to manipulation, then this mechanism will suffer from the same problem as the current discriminatory price mechanism, where users have to solve the fee selection problem.

Roughgarden (2021) provides three desirable properties for a blockchain fee setting mechanism and discusses whether various proposed fee setting mechanisms satisfy these desiderata. No known mechanism satisfies all three of these criteria in all plausible environments, but some satisfy all three in a restricted set of environments (Buterin's EIP-1559 and a new alternative) and others satisfy two of the three (including our StableFees proposal). The criterion offered by Roughgarden that our mechanism can fail is off-chain-agreement proofness. However, our mechanism can fail off-chain-agreement proofness only if it averages over blocks and averaging is not needed if there is a large number of users. We discuss how our mechanism deals with this issue in more detail in Section 3.3.

Numerous other recent papers analyze the current blockchain protocols, the games they induce, and efficiency or the lack thereof. Easley et al. (2019) and Huberman et al. (2021) analyze transaction fees, the mining game, and waiting times for users in the current Bitcoin protocol. Houy (2014) and Cong et al. (2021) provide analyses of the mining game. Lehar and Parlour (2020) document inefficiencies (nonfilled blocks and excess fees) that can be explained by collusive-price discrimination by miners. Carlsten et al. (2016) describe some consequences of removing the block rewards. Böhme et al. (2015), Harvey (2016), Malinova and Park (2017), Raskin and Yermack (2018), Yermack (2017), and Aune et al. (2017) all examine aspects of the Bitcoin environment. Azevedo et al. (2020) suggest an alternative auction to use to handle transactions bundled off the blockchain and then brought to a block in a batch. Guasoni et al. (2021) develop these (second layer) solutions in more detail and determine when these channels will emerge in cryptocurrencies. Rosenfeld (2011), Eyal and Sirer (2014), Gans and Halaburda (2015), Gandal and Halaburda (2016), Biais et al. (2019), Alsabah and Capponi (2021), and Garratt and van Oordt (2020) consider various design issues of the Bitcoin protocol, although they mostly focus on security rather than the efficiency of the fee mechanism. Chen et al. (2021) provides an introduction to the economic analysis of blockchains. Tsoukalas and Falk (2020), Rosu and Saleh (2020), and John et al. (2021) analyzes games between miners induced by proof of stake protocols and the various mechanisms at play there, although these are orthogonal to the fee market that we focus on.

There are also at least two auction theory papers that are related to the issues we discuss. First, Akbarpour and Li (2018) provide an analysis of mechanisms in which the seller can deviate from the rules of the auction. In this case, the mechanism has to be incentive compatible for the seller. They show that a first price auction is the only credible static auction. Essentially, an auctioneer could announce a different auction, such as a second price auction, but then once bids are received, he can submit a false second highest bid just below the actual highest bid—turning the auction into a first price auction. Credibility also matters for our analysis as our miner can submit own bids; however, our environment differs as there are multiple items for sale, the mechanism can impose some constraints on the miners, and most importantly, miners revenue can depend on the fees generated by a sequence of blocks determined by the protocol. Second, Mezzetti and Tsetlin (2008) show that in identical multiunit auctions, the bidding functions in the highest-losing-bid and lowest-winning-bid auctions converge together as the difference between the number of bidders and the number of units for sale diverges. We demonstrate that if there is a sufficiently large number of users, then truthful bidding by users and nonmanipulation by the miner is an ϵ-Bayes Nash equilibrium of the game induced by StableFees. The user part of our analysis relies on a similar intuition as in Mezzetti and Tsetlin (2008)—As the number of users diverges, the gain from strategic bidding disappears as the likelihood that any bidder sets the price vanishes.

In the remainder of the paper, we first discuss the positive and negative aspects of using a discriminatory price auction for slots on the blockchain, and we describe our goals for an improved mechanism. We then present our model of the cryptocurrency fee market and our mechanism, StableFees, which mitigates issues present in the current fee market. To do this, StableFees incorporates and adapts ideas from uniform price auctions to the trustless, decentralized setting present in today's cryptocurrencies. The key difference between the cryptocurrency fee market and the traditional auction setting is that miners are not trusted auctioneers so they can engage in behavior that manipulates the auction. StableFees accounts for these differences and has provable guarantees about potential manipulation by miners and

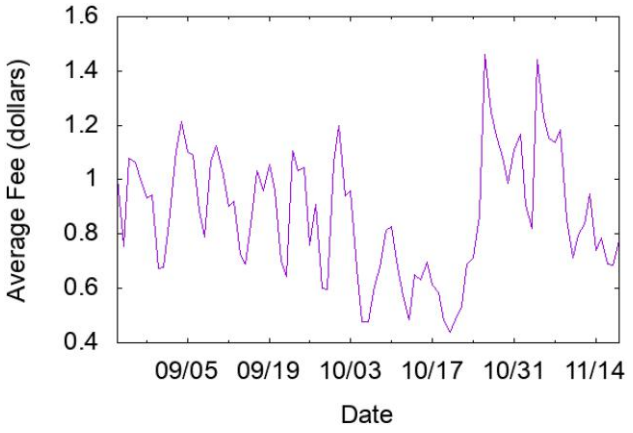
users. Finally, we provide an analysis of the social welfare achieved by StableFees, EIP-1559, and a monopolist miner.

2. Motivation

To motivate our design, we examine the lessons learned from auction design in the sponsored search market and then discuss challenges unique to the fee market.[6] Overture, the first company to use keyword-based advertising, initially sold ads using a discriminatory price auction. In the sponsored search market, advertising slots on the page that appears in response to a search term are sold to advertisers. Slots near the top of the page are preferred to ones down the page, and all of these dominate those on the second page, and so on.[7] Auctions are run frequently to determine whose ad appears where. Overture and its advertisers experienced instability: bids in successive auctions would rise as advertisers priced out in one auction tried to get into the next one; and then they would crash once bids reached levels that discouraged bidding at all. Eventually, discouraged advertisers quit and the auction was clearly producing less revenue than should be possible. Figure 1 illustrates that this erratic behavior of bids is also prevalent in the current Bitcoin auction mechanism.

An important aspect of Google's subsequent success in the sponsored search market was its use of a superior auction form: GSP, Google's generalization of the single-unit second price auction to their multiunit environment. GSP does not have dominant strategies, but it is second-price-like and simple, and it works reasonably well in practice. An earlier generalization of the single item second price auction to multiple items that has dominant strategies is the Vickrey-Clarke-Groves (VCG) procedure, which forms the basis of the auction mechanism used by Facebook.[8]

Figure 1. (Color online) Sawtooth Pattern for Bitcoin Fees over the Period August to November 2015



Source. Generated from BitInfoCharts (2021).
Note. This is similar to the figures and description from Edelman and Ostrovsky (2007) about Overture's first price auction.

2.1. Lessons from Sponsored Search

Auction theory and the experience of the sponsored search market suggest that some generalization of the uniform price auction could improve on the protocols used in the crypto space. Before modifying a uniform price auction to fit these environments, it is useful to first set out our objectives in designing a protocol and then to describe how uniform price auctions work.

We have three objectives. First, the protocol should result in an efficient assignment of slots on each block to users. Therefore, we want to assign all slots to users with the highest values, leaving a user out of a block only if there is no user in the block who has a lower true value than the left-out user. An assignment with this property is called socially optimal. Second, we want the game induced by the protocol to incentivize nonstrategic behavior. Ideally, we would like each user's optimal bid, which is the fee they propose to pay, to be their true value for a slot and we would like the miner building the block to have no profit motive for deviating from the "rules of the auction." Third, we want optimal strategies to be simple and obvious. This last criterion is difficult to quantify, but a protocol that induces a game in which every participant has weakly dominant truth-telling strategies surely satisfies it.

In the standard auction environment, a uniform price auction with the price set by the highest unaccepted bid achieves these goals. A uniform price auction for K identical items to be sold to $N > K$ bidders who each want at most one item works as follows. Bidders are asked to submit bids to the seller or to the algorithm running the auction. The bidders who have submitted the K highest bids each win an item, and they all pay the $(K+1)$st highest bid. If the algorithm, or auctioneer, can commit to this auction form, and if bidder's private valuations for an item are independent, identically distributed (i.i.d.) draws from a fixed distribution, then it is a weakly dominant strategy for each bidder to bid truthfully—submit a bid equal to the value for an item.[9] This auction form has another attractive feature—It guarantees that winning bidders place the highest values on the items, so it results in a socially optimal allocation.[10] The following remark summarizes standard results about multiunit auctions.[11] In the appendix, we provide a direct proof of item 1 as we use this logic elsewhere in our arguments.

Remark 1. Suppose that the auctioneer has K identical items for sale and can commit to an auction form. Suppose also that each bidder, $i = 1, \ldots, N$ with $N > K$, wants at most one item and that bidders' private values V_i are drawn i.i.d. from a distribution on $[0, \overline{V}]$. The auction form chosen by the seller induces a game

between the bidders in which each bidder selects a strategy mapping the bidder's private value, V_i, to a bid $b_i \in [0, \overline{V}]$.

1. If the auctioneer runs a uniform price auction with the K items sold to the K highest bidders at the $(K+1)$st highest bid, then it is a weakly dominant strategy for each bidder to bid truthfully, $b_i = V_i$ for all i, and if each bidder follows this weakly dominant strategy, the assignment induced by the auction is socially optimal.

2. If the number of bidders and the distribution of values is common knowledge, and the auctioneer runs a discriminatory auction—the K items are sold to the K highest bidders and each successful bidder pays his own bid—then there is a Bayes-Nash equilibrium (a list of strategies, one for each bidder, which are mutual best responses) of the game induced by the auction in which the equilibrium assignment is socially optimal.

For the environment described in Remark 1, the discriminatory auction and the uniform price auction both result in socially optimal assignments. In a uniform price auction, each bidder only needs to know his own value and the form of the auction. Bidding truthfully is optimal regardless of who the other bidders are or how they behave. This is not true in the discriminatory auction. Here, the efficiency claim rests on the assumption that play can be described by a Bayes-Nash equilibrium in which each bidder is best responding to each other bidder.

To illustrate the difference in these two auctions, it is useful to examine them in the simplest case in which there is a single item for sale to N bidders with values, V_i, drawn i.i.d. from the uniform distribution on $[0,1]$. In a uniform price auction, it is weakly dominant for each bidder i to simply bid his value V_i. In a discriminatory price auction, there is an equilibrium in which the optimal strategy for a bidder with value V_i is to bid $((N-1)/N)V_i$. This result requires knowledge of the number of bidders, depends on distribution of values being uniform, and is optimal only if all other bidders follow the same strategy. However, it does result in a socially optimal allocation because equilibrium bids are increasing in true values.[12]

2.2. Redesigning the Fee Market

The blockchain environment is a trustless, decentralized system in which there is no mechanism that can force miners to act as if they are a trusted auctioneer. Therefore, any redesigned mechanism has to take into account the incentives of the miners to follow the "rules" of the auction rather than to manipulate it. Furthermore, only the miner knows the transactions and their attached fees in his private mempool. Once the miner writes transactions to the blockchain the details of those transactions are known, but details of the transactions left out are not known—and the protocol cannot credibly call for payments that depend on those left-out transactions.

Most importantly, the miner can also act as a user and include the miner's own transactions in the block, moving money from one of the miner's wallets to another with whatever fee the miner chooses after observing the fees offered by users. All identities on the blockchain (miners, users, etc.) are uniquely identified by a cryptographic key. Thus, it is cheap to create a new identity but hard to assume the identity of another person. This makes it difficult to enforce roles for each participant because it is possible for a miner to also impersonate other, arbitrarily many, identities that act as "users."

This ability to act as a user or many users allows a miner who earns the revenue generated by the block to introduce first-price-like features into a supposedly second-price-like auction at zero cost. To see this in the simplest case, suppose that there is only one transaction per block, a second price auction is announced and all fees that bidders submit can be observed and used by the protocol. The miner can manipulate this auction by including a fictitious transaction paying a fee slightly below the highest offered real fee, so the user with the highest offered fee wins and pays approximately that fee while the miner pays nothing. This makes the single item, "second price" auction with a strategic auctioneer effectively a first price auction. Therefore, bidders should place first price bids and in equilibrium, we should see a first price outcome. Nonetheless, we show that with a large number of bidders and multiple slots on the block, StableFees achieves uniform price-like results: The incentives for both users and miners to manipulate converges to zero as the number of users diverges.

StableFees starts with the intuition from uniform price auctions that the fee to be included in a block should be approximately the minimum bid that was included in that block. However, miners are incentivized to manipulate such an auction to try to maximize their revenue. To discourage this behavior, StableFees spreads out the fee reward from a block across several blocks so that miners are unable to insert transactions into each block for free. This allows StableFees to provide similar guarantees to uniform price auctions in realistic conditions while still being deployable in the cryptocurrency ecosystem.

2.3. Additional Design Issues

The potential for side payments and security issues associated with the choice of block size introduce additional limitations and considerations in any redesign of the fee market. Side payments arise when the issuer of a transaction and the miner of a block arrange a payment outside of the blockchain itself. These payments may be direct, with fiat money or other cryptocurrency tokens changing hands, or indirect, as when a miner

pays out rewards to its pool participants. It is difficult for any auction mechanism to effectively deal with side payments as they are inherently more valuable to one particular miner and that value is not shown in the bid for the transaction. We require our design to be resistant to manipulation due to side payments, but we do not see eliminating side payments entirely as either feasible or desirable.

In cryptocurrencies, the hard cap block size is set according to a variety of factors, including security and the fee level. Blocks that are larger than the hard cap are deemed invalid and are not included on the chain. If blocks are small enough, then their size has no effect on the network security—and this secure size is slowly increasing as the underlying network infrastructure improves. However, beyond a certain safe block size, the larger the block, the longer it takes to transmit and the more likely it is to cause adverse effects on the network security. Protocol designers trade off these properties when choosing the hard cap, and in our view, the choice of a hard cap should be exogenous to the fee market design.[13]

If a miner is unable to fill a block to a sufficient level, then the price of block space should be approximately zero as space on the block is not scarce. As a result, the auction mechanism should charge transactions in underfilled blocks only some nominal transaction fee.

We outline a path toward deployment and other considerations in Appendix A.

3. StableFees

We now present the formal model in which StableFees operates, the StableFees protocol, and performance guarantees provided by StableFees.

3.1. Model

We denote the fixed number of slots in a block by K. These are the slots that can be assigned to transactions by the auction mechanism.[14] We assume that there are $N > K$ users. Users have private values denoted V_i, $i = 1, \ldots, N$ for having their transaction recorded to the current block.[15] There is a common shock, V, to these individual values that has continuous, strictly positive density on $[0, \overline{V}]$. Given V the V_i are i.i.d. and have continuous density $g(\cdot|V)$ on $[0, \overline{V}]$ with $g(V_i|V) \geq \epsilon > 0$ for all $V_i \in [0, \overline{V}]$.[16] This structure allows user values to be unconditionally correlated as there is a common shock to their values. However, the users' payoffs are determined by their private values and not by the common shock. Therefore, no user needs to make inferences about the values of other users. For example, it could be that placing a transaction on the blockchain has a common value component that varies from time to time, perhaps as the price of the cryptocurrency varies, but users also have an idiosyncratic component to their value perhaps reflecting their own liquidity needs. A user who is not included in the current block receives no reward from the current block.[17]

Users attach transaction fees, or bids, to their transactions. We model users as selecting bids after knowing their own value, but without knowing the realization of values or the choice of bids for other users. We assume that distributions of users and values are common knowledge.

The miner selects which transactions to put into the block after seeing the bids attached to those transactions. We do not address how this miner is selected. However, that is done (e.g., by proof-of-work or proof-of-stake) our protocol can be applied to determine fees. We focus on blocks for which the number of users in the pool is greater than the number of slots in the block; for other blocks there is no congestion as all users in the pool can be included in the block.

3.2. Protocol Description

1. Any user who wants a transaction recorded in block b can attach a fee to their transaction. Denote the fee attached by user i by f_i.

2. Users whose transactions are included in block b each pay the minimum fee proposed by any user whose transaction is included in block b. The total paid by the users in block b is the revenue generated by block b.

3. The miner who builds block b is paid the average revenue generated by the $B \geq 1$ most recently mined blocks, including block b, if and only if the miner fills block b. Otherwise, the block is not included in the blockchain.

- A block is defined to be filled if it contains K transactions or if the miner pays a *fill penalty*. The necessary fill level, K, is a parameter which can be chosen to be some fraction, say 80%, of the hard cap. The fill penalty is defined to be the difference between K and the number of transactions in the block times the fee paid by each transaction. A miner can also avoid paying the fill penalty by declaring that there were not enough transactions in the mempool to fill the block, in which case each user is charged the minimum allowable fee for a transaction to be included in the mempool and the block is declared filled.
- A minimum fee required for a transaction to be considered can be included by declaring that transactions are not in the mempool if the proposed fee is below that minimum level.[18]
- The miner has the option to fill the block to capacity with transactions, but only K of them are priced using this auction mechanism. The other transactions are charged no fees for being included in this block. This allows the miner to, for example,

pay pool participants, and is helpful in creating side payment resistance.

In this section, we consider a miner problem in which $B = 1$. In this case, the miner of block b is paid the revenue generated by block b. Averaging over past blocks reduces incentives to manipulate, so in this section, we consider the case that is most demanding for a nonmanipulation result. Given the StableFees protocol, it is obvious that, for any collection of fees submitted by users, if a miner accepts a transaction with fee f then the miner will accept all transactions that offer higher fees. Therefore, for any given fees proposed by users, the miner's strategy needs to specify the accepted minimum fee. The miner can insert a fictitious fee to manipulate the highest nonaccepted fee, so the miner's strategy also needs to specify any fictitious fee.

This protocol induces a Bayesian game with users and the miner as players. Denote user i's strategy by $\sigma_i : [0, \overline{V}] \to [0, \overline{V}]$ a mapping from V_i to a fee, f_i.[19] The payoff to a bidder who proposes a fee f_i is $V_i - f_i$ if the fee is accepted and zero if the fee is not accepted.

The miner's strategy is denoted by $m : [0, \overline{V}]^N \to [0, \overline{V}]^2$, where $m(f_1, \ldots, f_N) = (\overline{f}, f^m)$ specifies the minimum fee the miner accepts, $\overline{f}$, and the fictitious fee the miner enters, f^m, which is 0 if the miner does not enter a fictitious fee. The miner's payoff is the sum of accepted fees proposed by users.

We analyze ϵ-Bayes Nash equilibria of this game. A Bayes Nash equilibrium is a collection of strategies, one for each user and one for the miner, that are best responses to each other. If, given the strategies of the other players, a strategy for a player provides an expected payoff within ϵ of the maximum payoff the strategy is called ϵ-optimal. A collection of strategies that are mutually ϵ-optimal is an ϵ-Bayes Nash equilibrium.

3.3. StableFees Guarantees

If the number of users is small, there are incentives for both users and the miner to manipulate StableFees. As StableFees uses a Kth price auction for K slots, the marginal user sets the price and has an incentive to underbid. Ex ante, any user could be the marginal user, so all users have an incentive to underbid. Miners can manipulate by inserting fictitious transactions into a block, hoping to increase the minimum fee attached to transactions included in the block without losing too many real transactions. We discuss the small numbers case later in this section and address it empirically and with simulations in Section 4. We next consider the large numbers case and show that the incentives for both users and the miner to manipulate decline to zero as the number of users grows relative to the size of a block. Thus, truthful revelation of values by users and nonmanipulation by the miner is a ϵ-Bayes Nash equilibrium.

3.3.1. Truthful User Bidding with a Large Number of Users.

Assuming, as we will demonstrate later for the large numbers case, that miners place the K highest fee transactions on the block and that users whose transactions are placed on the block all pay the Kth highest fee, it is not a dominant strategy for users to bid truthfully. However, the incentive to bid strategically is small if the number of users is large. To see this, suppose that all other users bid truthfully. Let V_{K-1} and V_K be the $(K-1)$st highest bid of others and the Kth highest bid of others. If a user's value is below V_K, there is no possible gain from bidding strategically as the price will be greater than the user's value for any bid. There is a potential profit from strategic bidding only if the user's value is greater than V_K, and this gain is bounded by $V_{K-1} - V_K$. Therefore, for user I, the expected gain to strategic bidding is bounded by $E[V_{K-1} - V_K]$, which converges to zero in the number of users. For example, with draws of user values according to the uniform distribution on $[0,1]$ and N active users, the upper bound on the gain is $1/N$ and with the exponential distribution with parameter λ, it is $1/\lambda(N-K+1)$. That is, with a large number of users, the potential gain to strategic user behavior is small, and it seems plausible that, rather than attempting to follow a manipulation strategy, users will instead follow the simpler, nearly optimal, strategy of truthful bidding.

Proposition 1. *In the previous model, if the StableFees mechanism is used then:*

- *Truthful bidding is not a dominant strategy for users.*
- *Suppose that miners do not manipulate. For any $\epsilon > 0$ there is an N_ϵ such that for any number of users $N \geq N_\epsilon$ truthful bidding is an ϵ-Bayes Nash equilibrium of the induced game between N users.*

Proofs are given in Appendix B.

3.3.2. Nonmanipulation by Miners with a Large Number of Users.

Consider first the case in which miner revenue is not averaged over past blocks; that is, $B = 1$. In this case, the miner of block b is paid the revenue generated by block b. Averaging over past blocks reduces incentives to manipulate within the protocol, so in this section, we consider the case that is most demanding for a nonmanipulation result.

If a miner accepts a transaction with a bid of b, then the miner clearly accepts all transactions with greater bids. Therefore, a miner who wants to manipulate will insert a fictitious transaction with a fee equal to one of the K highest fees offered.[20] Relabeling the K highest fees from highest to lowest, they are $f_1 \geq f_2 \geq \cdots \geq f_K$. For a miner to not manipulate, we need the revenue generated from K transactions at the Kth highest bid to be greater than the revenue generated from any smaller number of transactions n at the nth highest bid, that is, $Kf_K > (K-1)f_{K-1}$, $Kf_K > (K-2)f_{K-2}$, and so on. This

clearly holds if it holds sequentially, that is, $Kf_K > (K-1)f_{K-1}$, $(K-1)f_{K-1} > (K-2)f_{K-2}$, and so on. This second collection of inequalities can be written as $nf_n > (n-1)f_{n-1}$ for each $n = 2, \ldots, K$. Or $f_n > (n-1)(f_{n-1} - f_n)$ for each $n = 2, \ldots, K$. If users bid truthfully, then for any fixed n, as the number of users diverges, the left-hand side of this inequality converges to $\overline{V}$ almost surely, and the right-hand side converges to zero manipulation almost surely vanishes. Therefore, the miners' gain from manipulation vanishes as the number of users grows.[21]

Proposition 2. *Suppose the StableFees mechanism is used, and users bid truthfully. For any $\epsilon > 0$, there is an N_ϵ such that for any number of users $N \geq N_\epsilon$, nonmanipulation is an ϵ-optimal strategy for the miner.*

These two results together imply the following proposition.

Proposition 3. *Suppose that the StableFees mechanism is used. For any $\epsilon > 0$, there is an N_ϵ such that for any number of users $N \geq N_\epsilon$ truthful bidding by users and nonmanipulation by the miner is an ϵ-Bayes Nash equilibrium of the induced game between N users and miner.*

In a truthful, nonmanipulation, ϵ-Bayes Nash equilibrium bids are increasing in values, and miners fill blocks with the users whose bids, and thus their values, are highest. Therefore, if the number of users is large, there is a simple, ϵ-Bayes Nash equilibrium in which the space on the block is used optimally.

3.3.3. Miner Incentives with a Small Number of Users.

If the number of users is small, then users may not bid truthfully, but regardless of how fees are chosen by users, StableFees provides an nonmanipulation incentive for miners through averaging over B blocks. In the remainder of this section, we discuss the effect of B on this incentive, and in the next section, we address it with simulations.

3.3.4. Fee-Based Ordering.

Recall that the miner's reward is the average revenue generated over the last B blocks, including the block that the miner has just mined.[22] Therefore, the miner receives a fraction of the reward generated by any block that they mine. Thus, when choosing which transactions to include in a block, the miner has an incentive to order the transactions according to the fees they offer and accept the highest fee transactions first. This force alone does not imply that the miner will fill the block; averaging over multiple blocks provides that incentive.

3.3.5. Full Blocks.

StableFees incentivizes miners not to submit fake transactions to fill blocks. To see why, note that the optimal fake-bid manipulation for a miner is to insert a fake bid equal to the minimum fee bid by a transaction that the miner wants to include in the block. Each fake transaction from the miner will require the miner to pay the associated fee. Because each block's reward is computed over the past B blocks, the fees collected from a particular block are spread over the next B blocks, including the current block. Thus, the miner can only expect to get a fraction of the increase in reward back.[23] Exactly how many blocks B to average over is an empirical question that depends on how much mining power the largest miner in a blockchain controls.

The fill penalty in StableFees allows miners to perform a manipulation equivalent to inserting fake bids without filling the block with unnecessary transactions. Miners will prefer to use the fill penalty rather than inserting fake bids as larger blocks are more likely to get forked and excluded from the blockchain due to random chance.[24]

3.3.6. Side Payment Resistance.

One type of manipulation is when users may attempt to pay miners outside of the auction mechanism. StableFees resists these types of side payments by including some block space where miners can place transactions outside of the auction mechanism. A transaction placed in this block space avoids paying fees and thus enables the miner to capture the full value of that transaction. These types of transactions may not necessarily be malicious and have legitimate value, for example, payments to members in their mining pool. However, this space is limited and should only be available at a premium cost higher than the auction clearing price, making it unattractive to users.

As Roughgarden (2021) notes, side payments may become an issue if because of averaging over blocks a miner and users reach a deal outside of the protocol to pay for and assign space on the block. If the number of users is large, we do not need to average over blocks and there is no side-payment issue. In the small numbers case, averaging is helpful in reducing the incentive for miners to manipulate, provided that it does not create too large an incentive for the miner and some users to circumvent the protocol. If this occurs, then how users should bid and how miners should respond becomes more complex—Truthful bidding is not optimal, and the miner should strategize over which users to include in any side deal and how much to charge them. Additionally, miners who attempt to circumvent the mechanism are likely to face social pressure from other miners. If StableFees is used, this behavior is detectable, so social pressure may ameliorate this potential off-chain problem. Whether off-chain agreements are an issue and how much averaging is possible before they matter are empirical questions.

4. Simulations and Stylized Facts

In this section, we evaluate StableFees' properties empirically. First, we use simulations to understand the

miner's incentive to manipulate StableFees. Second, we analyze bids during a period where block space was scarce in Bitcoin and Ethereum to estimate how much users are overpaying relative to StableFees and to estimate the reduction in variance of miners' revenue achieved by StableFees. Finally, we use simulations to understand the performance of StableFees compared with alternative auction mechanisms that were proposed by Lavi et al. (2017) and Buterin (2018).

4.1. Miner Manipulation

We ran a series of simulations to provide insight into a miner's incentive to manipulate StableFees. We show diminishing benefits to miners from manipulation as the number of miners increases or as the number of blocks we average over in determining miner fees increases. Overall, our results show that for reasonable parameters even optimal manipulation by miners has little benefit.

Our simulations take as parameters the number of transactions (N) in the mempool, the maximum number of transactions (K) per block, the number of miners (M), and the number of blocks transaction fees are averaged over (B). We first draw user bids from a Pareto distribution with a median of 2 cents and a mean of 10 cents, which is similar to the actual transaction fee distribution that appeared on the Bitcoin blockchain in July 2018.

The miner chooses which transactions to include to maximize payoff. The miner's optimal manipulation strategy is to accept the j, $1 \le j \le K$, highest bids and fill the rest of the block with fake transactions. Let b_j be the jth highest bid. A full block has K transactions, and thus the total fee generated by this block is Kb_j, of which the manipulating miner receives $(Kb_j)/B$ as a reward. A miner with $1/M$ of the hashpower also, on expectation, will receive an additional $(Kb_j(B-1))/(MB)$ in fees as this miner is expected to mine $1/M$ of the remaining $B-1$ blocks that the transaction fees are averaged over. The cost of the fake transactions that the manipulating miner must insert to perform this manipulation is $(K-j)b_j$. Thus, the benefit to the manipulating miner of only filling the block up to j transactions is $(Kb_j)/B + (Kb_j(B-1))/MB - (K-j)b_j$.

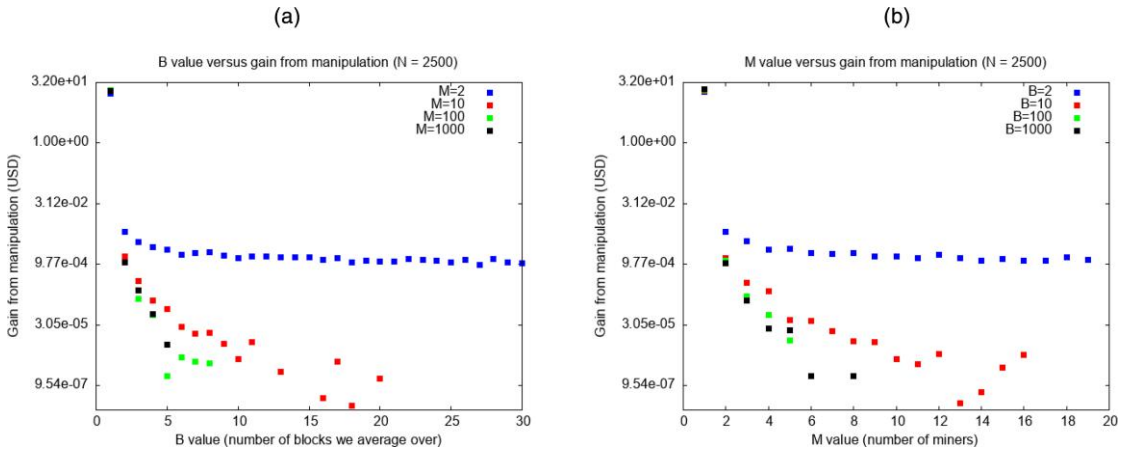
To find the optimal manipulation, we maximize this benefit over j. We define the gain as this maximum value minus the miner revenue if $j = K$, the case where the miner is honest and has not inserted any fake transactions into the block. We average this gain over 1,000 independent trials. In all our simulations, we keep the number of transactions in a block, K, fixed at 2,000.

Figure 2(b) shows the effect of the number of miners on a miner's gain from optimal manipulation for various values of B. As the number of miners increases, the dominant term in the block reward is the reward from transaction fees from the freshly mined block, which decreases as B increases. Thus, the gain from manipulation declines as the number of miners increases, but the decrease is most pronounced with higher values of B.

Figure 2(a) shows that the gain from optimal manipulation declines as the number of blocks averaged over increases and that it is uniformly lower for high numbers of miners. This is because as B increases, the fee rewards obtained from mining a block decreases.

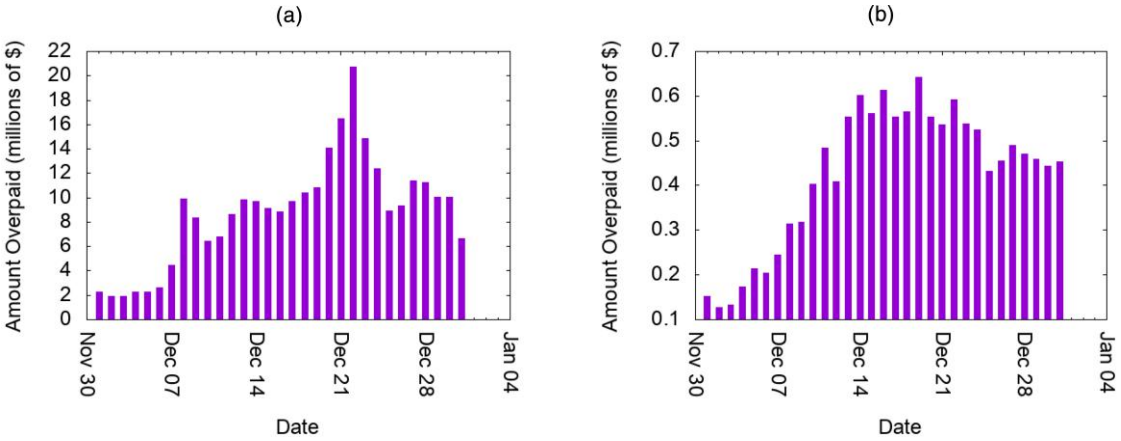
Our simulations show that miners gain little from optimal manipulation for reasonable numbers of miners and blocks averaged over. Miner revenue ranges from \$15 to \$40 over these simulations, and the gain from manipulation never exceeded 3 cents when $B > 2$ and $M > 2$, so the gain from manipulation relative to total

Figure 2. Miner's Gain from Manipulation as We Modulate Different Parameters



Notes. The mempool size, N, is fixed at 2,500 for both experiments. (a) Effect of the number of blocks we average over on the miner's gain from manipulation. The simulation shows that increasing the number of blocks we average over decreases the gain from manipulation with enough miners. (b) Effect of the number of miners on the miner's gain from manipulation. The simulation shows that increasing the number of miners decreases the gain from manipulation as long as we average over enough blocks.

Figure 3. (Color online) How Much (in Millions of USD) That Users Could Have Saved if the Transactions Were Using StableFees Instead of the Current Scheme in December 2017



Notes. We see that the savings in Ethereum are lower due to its higher throughput. (a) Savings in Bitcoin. (b) Savings in Ethereum.

revenue is negligible. Of course, our analysis excludes the miner's incentive to make Bitcoin succeed to maintain the value of their Bitcoin holdings and the ongoing value of the mining operation. Including this dimension would further reduce miner manipulation incentives. We believe that taken together these results suggest that miners are not likely to manipulate StableFees.

4.2. Blockchain Bid Analysis

To provide insight into the benefits of StableFees, we analyzed the bids that appeared on the blockchain during a period of high demand in Bitcoin and Ethereum (December 2017). We can compute the total fees paid by users in the current system, but we do not observe what bids would be if StableFees was to be adopted. The bid distribution would be different, but fortunately we do not need the entire distribution of bids to determine the revenue that would be generated by StableFees. This revenue is determined by the number of transactions in the block and the minimum fee bid among these transactions.

If StableFees was in place, users should bid truthfully, whereas in the current discriminatory price mechanism, users should shade their bids. In either case, if the market is in a Bayes-Nash equilibrium, bids will be increasing in values and the same transactions would be included in each block. This seems a reasonable approximation for StableFees; it is less compelling for the current mechanism, but we will use it to estimate the transactions that will be placed on the blockchain. With a large number of users the amount of shading of bids that users do in an equilibrium should be small.[25] Therefore, we will approximate the uniform price under StableFees by the minimum bid included in the block under the current discriminatory price mechanism.

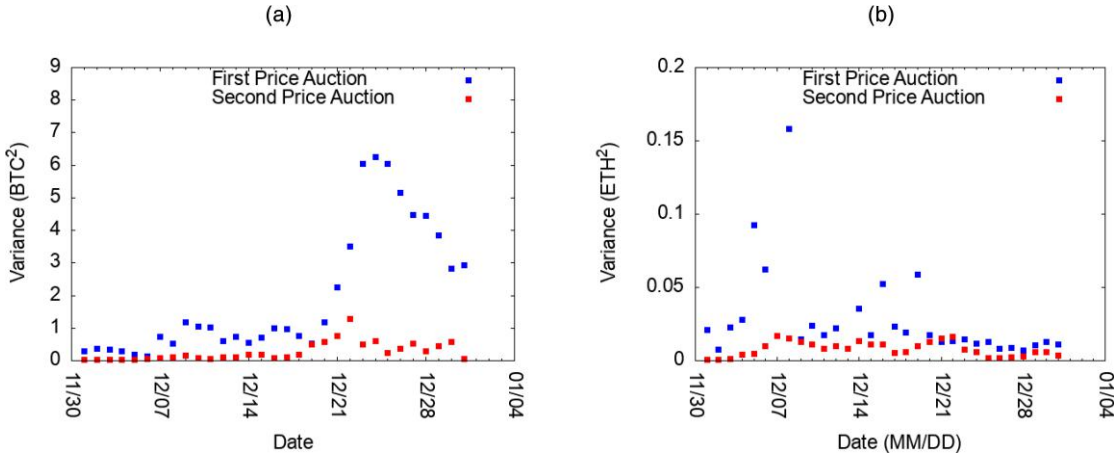
We first examine how much users could save if StableFees was used. To apply our mechanism to a day, we collect the blocks that were mined that day and the transactions in each block. We then calculate the fee per byte for each transaction to normalize the fees paid. Next, we compute the total that would be paid for each block if every transaction on the block pays the smallest fee per byte that appears on the block. We plot the difference between the actual fees paid and the fees that users would pay under StableFees. Figure 3(a) shows that users in this time period could have saved 273 million USD if StableFees was used in Bitcoin. Figure 3(b) shows a similar trend in Ethereum, but the dollar amounts are significantly smaller, which is to be expected as Ethereum has a higher processing capacity than Bitcoin. Even so, users in this time period could have saved 13.2 million USD if StableFees was used in Ethereum.

StableFees has the potential to improve predictability for miners by reducing the variance of miner's rewards. To quantify this, we use the previous calculation to obtain the fees that miners would receive under StableFees. We then compute the variance of the transaction fees over blocks using StableFees and the variance using the current discriminatory price mechanism. Figure 4(a) shows that the variance is lower in Bitcoin when using StableFees, by up to a factor of 20 on some days, with an average of 7.4. Figure 4(b) shows that the same trend holds in Ethereum, with a maximum of 75 times and an average factor of 7.9. In StableFees, payouts are averaged over B blocks, which would further decrease the variance by an additional factor of B^2 relative to Figure 4(a) and (b).

4.3. Social Welfare

The social welfare generated by an auction scheme is the sum of the net user and miner payoffs. In this section,

Figure 4. Variance in Payouts from Transactions Fees to Each Miner in December 2017



Notes. We see that the current mechanism has a significantly higher variance than StableFees, resulting in less stable payouts. (a) Variance in Bitcoin. (b) Variance in Ethereum.

we run a simulation to compare the amount of social welfare generated by StableFees and several alternative auction mechanisms. For this simulation, we assume that there is a fixed number of users all of whom have utility functions that are quasi-linear in money. Each user has a single transaction that they wish to include in the blockchain. The transaction has intrinsic value to the user that is drawn from a random value distribution. The users place a bid according to the optimal strategy for each auction mechanism. The miners then choose a subset of the transactions issued by the users to include in the block.

To calculate the welfare, we start with user payoffs. If a user i's transaction is placed on the block, the user gains the intrinsic value of the transaction, V_i, and pays the fee specified by the mechanism, f_i, yielding payoff $V_i - f_i$. If a user's transaction is not placed on the block, the user's payoff is zero. Thus, the total user welfare is $\sum_{i \in U}(V_i - f_i)$, where U is the set of users whose transactions are included in the block. The miner payoff is the total fees they obtain from the transactions that are placed in the block minus the cost of mining. In the analysis in this section, we keep the total number of miners and thus the cost of mining fixed, so we drop that term in our analysis as it is constant across auction mechanisms. Thus, the net miner payoff is the sum of the fees paid by users, or $\sum_{i \in U} f_i$. Summing the user and miner payoffs, we see that the total welfare generated by the auction mechanism is $\sum_{i \in U} V_i$. Our notion of social welfare is thus equivalent to efficiency; welfare is maximized if and only if the highest value transactions are placed on the block and the block is filled.

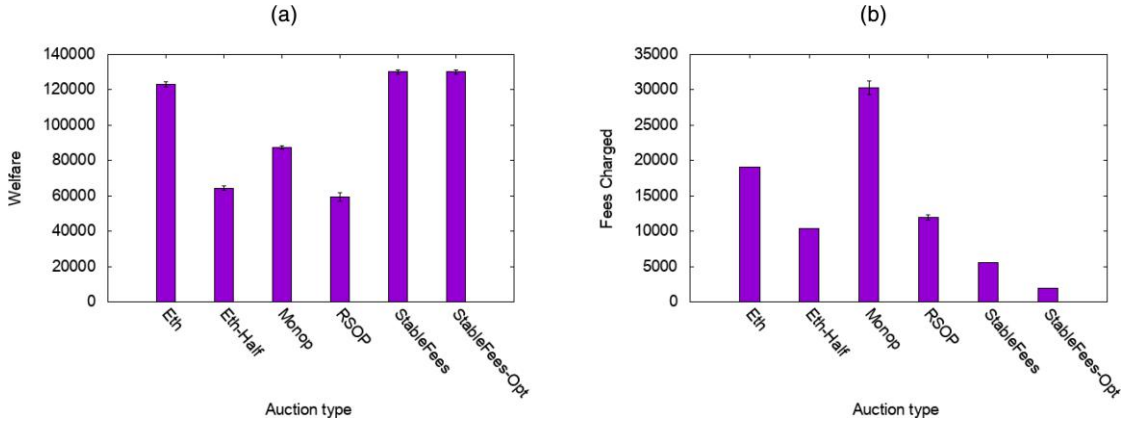
To answer the social welfare question, we ran an experiment where we simulate users and miners, run the auction protocol, and analyze what happened in the trace. In this experiment, we set the number of transactions in a block, K, to be 2,000, with 20,000 users in the system.[26] Once a user's value is chosen, the user submits an optimal bid given the auction mechanism. Then, the miners select the transactions that maximize their fee revenue under that auction mechanism. We repeat the process until 1,000 blocks have been mined. We report results from the last 900 blocks, as the Buterin (2018), or EIP-1559, auction mechanism is sensitive to the initial parameters for the first few blocks.

Although we test four schemes, StableFees and Buterin's mechanism Buterin (2018), which we call Eth, have two interesting parameterizations. Therefore, we consider six auction mechanisms. For each auction mechanism, we outline how it is parameterized, how users choose their bids, and how miners choose which transactions to include in a block.

StableFees and *StableFees-Opt* are both instantiations of the StableFees auction mechanism. We set B to be 10 in both instantiations. In StableFees, we pessimistically assume a large miner with 20% of the total hashpower, whereas, for StableFees-Opt, we assume that each miner has a very small hashpower and consequently will not win any further blocks outside of the one they mined. StableFees is a pessimistic parameterization of the StableFees algorithm, whereas StableFees-Opt describes the best-case performance of StableFees. In both instantiations, the miner attempts to manipulate the auction to maximize the total fee revenue as in Section 4.1. Users also set their individual bid to be their private value, $b_i = V_i$, as this is approximately optimal with a large number of users.

RSOP (Lavi et al. 2017) is a randomized sampling auction mechanism. To initialize RSOP, a parameter α is set similarly to our B value. However, to prevent our choice of α from skewing the results, we assume that miners do not misbehave and simply include the 2,000

Figure 5. (Color online) Welfare and User Fees Paid Under a Constant Demand Curve



Notes. (a) Welfare. (b) User fees paid.

highest pending bids in the block. This will overestimate the social welfare produced by this auction mechanism. However, in RSOP, not all transactions that are included in the block are actually confirmed. In RSOP, users are incentivized to bid their true value; we set each user's bid to be their private value as well, so $b_i = V_i$.

In *Monop* (Lavi et al. 2017), miners are paid the lowest bid in the block times the number of transactions in the block. As a result, miners will choose the number of transactions $j \leq K$ such that jb_j is maximized, where b_j is the jth highest bid placed by a user. Under the large number of users assumption, users bid truthfully in this auction mechanism, so user i's bid will be V_i.

The auction mechanism proposed in Buterin (2018) uses the number of transactions in each block to estimate the demand for block space and consequently does not have a fixed block size. Blocks have a minFee parameter, which is the price that each transaction included in the block is charged. Miners include all transactions that offer to pay more than the minFee. As the user has very little influence on the fee that they will be charged, we again assume that the users bid truthfully in this auction. The minFee is set and adjusted according to the target block size. Blocks larger than the target block size decrease the minFee, and blocks smaller than the target block size increase the minFee. The hard capacity of a block is twice the target block size.

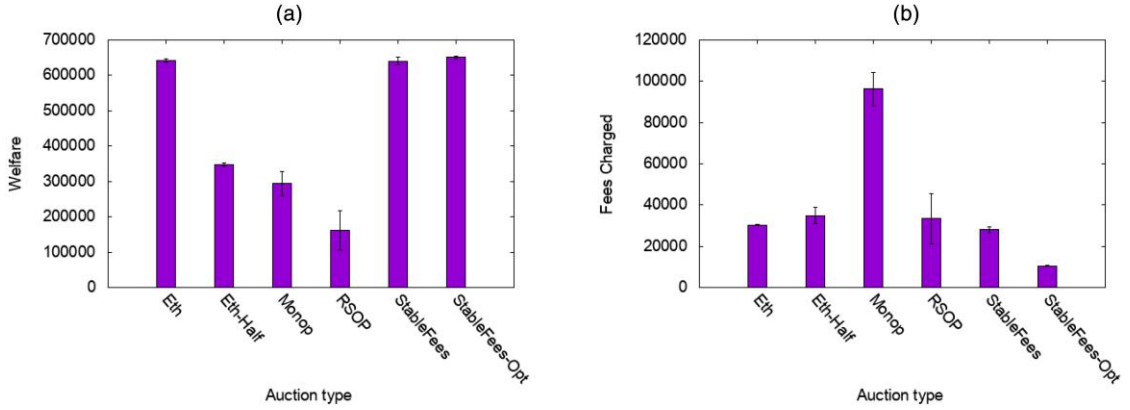
All other auction mechanisms have a hard capacity of 2,000 transactions. To guarantee that blocks produced by the auction mechanism of Buterin (2018) will obey that hard capacity, the target block size is set to 1,000. As the typical block will be half-full in this parameterization, we call this auction *Eth-Half*. Increasing the target size above 1,000 allows this auction mechanism to process more transactions on average at the cost of having some blocks larger than 2,000 transactions, which harms the security of the system. The largest such target size is 2,000, where the typical block is 2,000 transactions, which we call *Eth*. The true social welfare produced by the auction mechanism of Buterin (2018) in practice will lie somewhere between *Eth* and *Eth-Half* depending on how the mechanism designer wants to trade off a loss of security for greater social welfare.

In the simulation used to generate Figure 5(b), the demand curve is constant and the same as the demand curve we used to study miner manipulation—User values are drawn from a Pareto distribution with a median of 2 cents and a mean of 10 cents. Among the most realistic parameterizations (StableFees, Eth-Half, RSOP, Monop), we see that StableFees produces 49% more welfare than the second-best parameterization, Monop. Among all parameterizations, we see that both variations of StableFees and Eth produce the greatest amount of social welfare. However, StableFees charges much lower user fees. Most importantly, 33.6% of blocks in the Eth auction contained more than 2,000 transactions, which weakens the security of the system.

In the simulation used to generate Figure 6, we take the demand curve from Figure 5(b) and multiply the bids by a fluctuating factor from 2 to 10 to randomly increase and decrease demand throughout the experiment. Again, StableFees performs better than all other realistic parameterizations by producing 85% more welfare than Eth-Half, which was the second-best realistic parameterization. Among all parameterizations, we see that both variations of StableFees and Eth produce the greatest amount of social welfare. However, the fees charged to the user are now comparable between the three auction mechanisms due to Eth undercharging users for block space. This comes at a cost of decreased security, as 64.9% of blocks contained more than 2,000 transactions in the Eth auction.

In the simulation used to generate Figure 7, bids are either drawn from the same Pareto distribution as previously discussed or multiplied by a factor of 10. We

Figure 6. (Color online) Welfare and User Fees Paid Under a Fluctuating Demand Curve



Notes. (a) Welfare. (b) User fees paid.

fluctuate between these two extremes throughout the experiment to simulate demand spikes. StableFees is again the best performing realistic parameterization, with 103% more welfare produced than Eth-Half. Similarly, both variations of StableFees and Eth produced the largest social welfare at the cost of 46.1% of blocks containing more than 2,000 transactions. The Monopolistic miner and RSOP produce the lowest welfare while charging the highest fees. This is likely due to the higher chance of manipulation with demand spikes and the larger fee spreads.

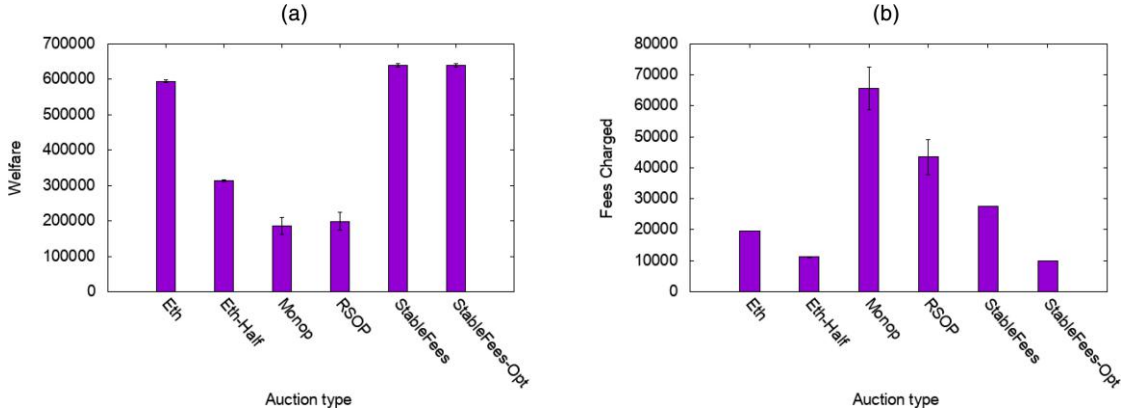
Overall, we see that StableFees consistently produces the largest social welfare among these mechanisms. Moreover, StableFees does comparatively better when the demand fluctuates more, with the relatively best results for StableFees coming when auctions were tested using large demand spikes. Additionally, StableFees, StableFees-Opt, and Eth are the three mechanisms that produce the greatest welfare, although Eth does it by reducing security. Finally, one interesting point to note is that StableFees is very close to StableFees-Opt, with the largest discrepancy being in Figure 6, where StableFees-Opt produces 1.6% more welfare than StableFees. Therefore, even if a cryptocurrency has a few large miners, StableFees still performs reasonably well.

4.4. Welfare with an Endogenous Number of Miners

In the previous analysis, we held the number of miners fixed, but in the long run, it should adjust in response to changes in the fees they earn. In an equilibrium with free entry, expected profit to mining must be zero, and this condition, along with expected revenue and costs per miner, determines the equilibrium number of miners. For our purposes here, the details of that determination are not important (see Easley et al. (2019) for details). What matters is that, holding cost per miner fixed, an increase in revenue per miner will increase the equilibrium number of miners and similarly a decrease will reduce the equilibrium number of miners.

Some number of miners, $\overline{M}$, is needed for secure posting of transactions to the blockchain. Let $BR + F \geq 0$

Figure 7. (Color online) Welfare and User Fees Paid Under a Demand Spike



Notes. (a) Welfare. (b) User fees paid.

be the minimum per block revenue necessary to have $\overline{M}$ miners in equilibrium where BR is the exogenous block reward, and F is total fees collected by the miner who writes a block. If the equilibrium number of miners is less than $\overline{M}$, then the blockchian fails and the social welfare it generates is zero. Clearly, any fee payments in excess of F are socially wasteful. They are paid by users, but they are then wasted in competition by the excessive number of miners necessary to drive expected profit to zero. Therefore, social welfare is

$$\sum_{i \in Block} (V_i - f_i) + \min\left\{BR + F, BR + \sum_{i \in Block} f_i\right\}$$

if the equilibrium number of miners is at least $\overline{M}$ and zero otherwise.

Stablefees reduces fees relative to current levels and relative to the fees proposed by alternative mechanisms. This will increase social welfare if the resulting equilibrium number of miners is at least $\overline{M}$; that is, if the rewards induced by StableFees are at least $BR + F$. It could create an issue with security if revenue is reduced below $BR + F$. We suspect that currently this is not an issue as most (currently approximately 90%) of Bitcoin miner revenue comes from the block reward, BR, rather than from fees.[27] As Bitcoin block rewards decline over time, having miner rewards remain sufficient for security could become an issue—and then a minimum fee (as is allowed in StableFees) may be needed. A force in the other direction (arguing for fewer miners than now) is that mining uses a large amount of electricity in what is clearly a socially wasteful competition. See Benetton et al. (2021) for more discussion of these environmental issues. One advantage of StableFees is that miner revenue is generally lower and so our protocol is environmentally friendly.

5. Conclusion

Cryptocurrencies cannot go mainstream if constructing a transaction imposes a cognitive load or requires strategic bidding behavior. We show that the fee mechanism currently used in a variety of coins encourages users to use strategic bidding strategies. We then present StableFees, an alternative that obviates this need and offers a more stable, predictable fee market.

Both a discriminatory price auction and StableFees work well in equilibrium if there is a large number of users relative to the capacity of blocks. However, in StableFees, the transaction fee offered by a user only affects what a successful user pays if the user has the, potentially unique, Kth highest bid. Otherwise, the fee only affects whether the user is in the block or not in it. Therefore, the gain to strategic bidding is small if there are many users. In a discriminatory price auction, every user pays their bid if their transaction is in the block. Here, strategic bidding is inescapable, although the equilibrium gain from it does converge to zero as the number of users grows. Also, we see in our simulations that miner revenue will have lower variance under StableFees. However, what happens to the actual payout with real users and miners is unclear, at least in part because of the nonrobustness of the discriminatory price procedure.

If StableFees is adopted, then with a large number of users and miners, there are simple nonmanipulative strategies that are nearly optimal and taken together these strategies form a ϵ-Bayes Nash equilibrium of the game induced by StableFees. In our model, if participants follow these strategies, then StableFees provides a social welfare maximizing allocation. It is in this sense an optimal mechanism for the large numbers environment. We view this as a step in the direction of more broadly optimal mechanisms that could take into account the impact of the fee setting mechanism on aspects of the environment that we hold constant, particularly security issues and the number of users.

Finally, our analysis applies to proof-of-work protocols such as those used in Bitcoin, Ethereum, and many others. Alternative protocols are being considered and used in a variety of different digital currencies. Most notably, Ethereum is considering a switch to proof-of-stake. Regardless of the protocol, cryptocurrencies will need to prioritize transactions somehow, and StableFees can be applied to solve this problem. Most cryptocurrencies currently charge fees to use the network and induce a discriminatory auction. Thus, they face the same problems described previously. One potential difference between proof-of-work and proof-of-stake protocols is the differing incentives they introduce for miners to preserve the value of the cryptocurrency. The link between the value of being a miner and preservation of the value of the cryptocurrency is stronger with a proof-of-stake system and that may introduce beneficial incentives for miners to behave nonstrategically in the proof-of-stake system.

Acknowledgments

The authors thank Phil Daian, Amani Moin, Andrew Morgan, David Yermack, and participants at the Georgetown Virtual Seminar Series on FinTech and the Crypto and Blockchain Economics Research Conference (CBER) for helpful comments; our discussant at CBER, Mariana Khapko, for valuable comments and suggestions; Dhruba Basu for assistance with data and graphs; and the referees and editors for many helpful suggestions and comments.

Appendix A. Parameterizing StableFees

Deploying StableFees requires us to set multiple parameters: the fill level, the minimum fee, and the number of blocks we average the rewards over (B). We briefly discuss some issues around how to set these parameters for Bitcoin as a case study. The fill level should be set to something very high, such as 80% or 90%, as miner transactions should not take up

a large amount of block space. Our work assumes that all transactions are equally sized, so the fill level refers to the resource being used to determine whether a block is full, and bids are per unit resource. In Bitcoin today, this is the block weight rather than the block size Lombrozo et al. (2017). Similarly, Ethereum uses gas rather than block size. To set the minimum fee, one can use the default minimum relay fee or any other similarly small fee. As Bitcoin often operates at full capacity, the minimum fee does not matter much. For cryptocurrencies that are not operating at their full capacity, the minimum fee should be set to the amount of resources consumed by a transaction. Finally, to determine B, we note that for Bitcoin, the largest mining pool, has about 21% of the hashpower (Gencer et al. 2018), implying that $M = 5$ for the largest miner. Thus, from Figure 2(b), we see that setting $B = 10$ allows the miner very little gain from manipulation and setting B higher does not make StableFees significantly more robust.

Appendix B. Proofs

Proof of Remark 1. We provide a simple, direct proof of this claim as we use the logic elsewhere in the paper. Consider bidder i with value V_i. We need to show that bidding more than V_i or less than V_i cannot increase the profit of bidder i.

Consider a bid $b_i > V_i$. Bidder i's bid only affects whether I wins or loses the auction; it does not affect the price I pays conditional on winning. Therefore, this high bid only changes the payoff to i if bidder i would not win with a bid of V_i and would win with a bid of b_i. That is, only if $b_i > V^K > V_i$, where V^K is the Kth lowest bid of the other bidders. In this case, i wins with a bid of b_i but pays $V^K > V_i$ as V^K is now the $K+1$st highest bid. Therefore, high bidding reduces i's payoff.

Alternatively, suppose that i bids $b_i < V_i$. This only affects i's payoff if i would have won with a bid of V_i and does not win with a bid of b_i. That is, only if $V_i > V^K > b_i$. In this case i would have won with a bid of V_i and paid $V^K < V_i$ and does not win with a bid of b_i. Therefore, a low bid also reduces i's payoff.

Proof of Proposition 1. Consider user i with value V_i and suppose that all other users bid truthfully. Let V_{K-1} and V_K be the $K-1$st highest bid of others and the Kth highest bid of others. If $V_i > V_{K-1}$ and $V_{K-1} > V_K$, then a bid by i of b_i such that $V_{K-1} > b_i > V_K$ gives i a slot on the block at price b_i, whereas a truthful bid gives i a slot on the block at price $V_{K-1} > b_i$. Therefore, truthful bidding is not a dominant strategy.

To show that truthful bidding is an ϵ-Bayes Nash equilibrium, we need to show that if all other bidders submit truthful bids then truthful bidding is ϵ-optimal for bidder i. If bidder i's value is below V_K, there is no possible gain from bidding strategically as the price will be greater than the bidder's value for any bid. If $V_i \geq V_K$, then bidder i would have a positive payoff if i obtains a slot on the block. The price of that slot will be V_{K-1} if $b_i \geq V_{K-1}$, as V_{K-1} will be the lowest successful bid, and it will be b_i if $V_{K-1} > b_i > V_K$, as in this case b_i will be the lowest successful bid. Therefore, the gain that user i can earn from strategic bidding (a nontruthful bid) is bounded by $(V_{K-1} - V_K)$, and this maximal gain can be earned only if $V_{K-1} \geq V_i \geq V_K$. Alternatively, in the event $V_i \notin [V_K, V_{K-1}]$ user i's maximal gain is zero. Thus, user i's expected gain from strategic bidding is bounded by $E[(V_{K-1} - V_K)|V_i] = E[(V_{K-1} - V_K)]$. This value is decreasing and converges to zero in the number of users. Therefore, for any $\epsilon > 0$, there is a N_ϵ such that truthful bidding is ϵ-optimal.

Proof of Proposition 2. To show that nonmanipulation is an ϵ-optimal strategy for the miner, we need to show that the miner's expected gain from manipulation is decreasing and converges to zero in the number of users, N.

Relabeling the K highest fees from highest to lowest, they are $f_1 \geq f_2 \geq \cdots \geq f_K$. The gain to a miner from optimal manipulation is the maximum difference between the revenue generated from K transactions at the Kth highest bid and the revenue generated from any smaller number of transactions n at the nth highest bid, that is, $(K-1)f_{K-1} - Kf_K$, $(K-2)f_{K-2} - Kf_K$, and so on. For nonmanipulation to be optimal, we want each of these terms to be negative. This clearly holds if $Kf_K > (K-1)f_{K-1}$, $(K-1)f_{K-1} > (K-2)f_{K-2}$, and so on. This collection of inequalities can be written as $nf_n > (n-1)f_{n-1}$ for each $n = 2, \ldots, K$. Or $f_n > (n-1)(f_{n-1} - f_n)$ for each $n = 2, \ldots, K$. If users bid truthfully, then for any fixed n, as the number of users (N) diverges, the left-hand side of this inequality converges to $\overline{V}$ almost surely, and the right-hand side converges to zero almost surely. Therefore, for any $\epsilon > 0$, there is an N_ϵ such that the miners' gain from manipulation is almost surely less than *epsilon* for all $N \geq N_\epsilon$.

Proof of Proposition 3. Follows immediately from Propositions 1 and 2.

Endnotes

[1] We consider, and the fee market uses, a discriminatory price auction in which the pricing rule is that each winning bidder pays his bid. For the case of a single unit of the good, these are known as first price auctions.

[2] The difficulty of determining what to bid has led to the development of bid prediction firms whose models suggest where the bid might be in upcoming blocks. Such predictions, however, can serve as focal points and lead to bidders to try to outbid the predicted bid.

[3] We use the term "uniform price" to refer to an auction of K identical goods in which the uniform price is the $(K+1)$st highest bid. If $K = 1$, this auction is a second price auction.

[4] More precisely, a $(K+1)$st price auction for K slots on the block.

[5] Yao (2018) provides proofs of conjectures from Lavi et al. (2017) about the general incentive compatibility and profitability of their monopolistic miner protocol.

[6] See Easley and Kleinberg (2010, chapter 15), for a discussion of the sponsored search market and references to the literature on sponsored search.

[7] In the sponsored search setting, slots do not have equal value to advertisers, whereas in the blockchain setting, slots on the block do have equal to users. Because of this difference, the immediate generalization of the single-unit second-price auction to the two settings differs. It's the VCG procedure in the sponsored search setting and a simplification to a uniform price auction in the blockchain setting.

[8] See Varian and Harris (2014) for a discussion of the development of GSP and Google and the use of VCG by Facebook.

[9] In our description of standard results, we restrict our attention to the independent values case. Later we consider an environment with

a large number of bidders where we have results for correlated, but private, values.

[10] This auction is equivalent to the VCG procedure. It is simpler to explain than VCG, but its optimality does depend on the items being identical. VCG does not require that restriction.

[11] See Weber (1983) and Milgrom (1985).

[12] Both auctions also yield the same expected revenue for the seller (in equilibrium). For the example in the text, a simple calculation shows this, but it is true much more generally according to Myerson's revenue equivalence theorem (Myerson 1981).

[13] The block size for a cryptocurrency has implications for the security and performance for a cryptocurrency (Croman et al. 2016, Gencer et al. 2018).

[14] For ease of exposition, our model uses the terminology used in popular proof of work cryptocurrencies, such as Bitcoin or Ethereum. Most cryptocurrencies operate on batches of transactions analogous to blocks, and transaction priority is decided by user's bids.

[15] We take waiting time into account indirectly, through users differing values for block space, and we do make use of a sequence of blocks, but we analyze the users' fee setting game only one block at a time.

[16] We keep the distribution of individual values fixed as the number of users increases. We can allow the dispersion of values to increase as N increases as long as the expected difference between the $(K-1)$st and Kth order statistics converges to zero as N diverges.

[17] For simplicity of exposition, we treat all transactions as taking the same amount of space on the blockchain. In practice, the fees we discuss are normalized in some way. For example, in Bitcoin, this would be the fee per byte. Dependent transactions, such as child pays for parent, can be handled by charging the average fee for both transactions.

[18] Including a minimum bid is standard in sponsored search auctions. For example, the minimum bid that Google currently uses is one cent per click.

[19] We do not consider mixed strategies as they are not necessary for our existence results.

[20] Inserting a fee between two existing fees is clearly dominated by making the fictitious fee equal to the higher of two nearby fees. Inserting multiple fictitious fees above the Kth highest fee removes some number of transactions from the block and sets the price at the lowest fee remaining in block and so is equivalent to a single fictitious fee strategy.

[21] Inserting a fictitious transaction with fee greater than the Kth highest real fee is equivalent to choosing to not fill the K slots on the block. Therefore, this argument also covers the potential manipulation of restricting the supply of slots.

[22] The fill penalty is considered part of the reward for a block.

[23] In a proof of work system his fraction is the fraction of mining power that this miner has.

[24] Forking occurs when two blocks are mined that cannot both be included in the blockchain (e.g., when they have the same height). Any block that is mined has a chance of being forked, but larger blocks have a higher chance of being forked. The probability of a fork occurring depends on the particular chain's properties.

[25] For example, if $K = 1$ and there are N users in a discriminatory price auction, there is a Bayes Nash equilibrium in which a bidder with value V_i bids $(1 - 1/N)V_i$.

[26] To have a fair comparison for each protocol, we assume a fixed number of transactions per block. Protocols that fluctuate the block size slightly are parameterized to target the same block size to ensure a fair comparison. The block size has implications for the security of the protocol as larger blocks take longer to propagate through the network and should be fairly independent of the fee mechanism used.

[27] As of January 25, 2021, the 30-day averages were \$3.62 for fees and \$35.85 for total revenue according to Blockchain.com.

References

Akbarpour M, Li S (2018) Credible mechanisms. Tardos E, ed. *Proc. ACM Conf. on Econom. and Comput.* (ACM, New York), 371–371.

Alsabah H, Capponi A (2021) Pitfalls of bitcoin's proof-of-work: R&D arms race and mining centralization. Preprint, submitted June 17, http://dx.doi.org/10.2139/ssrn.3273982.

Aune RT, Krellenstein A, O'Hara M, Slama O (2017) Footprints on a blockchain: Trading and information leakage in distributed ledgers. *J. Trading* 12(3):5–13.

Azevedo EM, Pennock DM, Waggoner B, Weyl EG (2020) Channel auctions. *Management Sci.* 66(5):2071–2082.

Benetton M, Compiani G, Morse A (2021) When cryptomining comes to town: High energy-use spillovers to the local economy. Preprint, submitted May 15, https://dx.doi.org/10.2139/ssrn.3779720.

Biais B, Bisière C, Bouvard M, Casamatta C (2019) The blockchain folk theorem. *Rev. Financial Stud.* 32(5):1662–1715.

BitInfoCharts (2021) Bitcoin avg. transaction fee historical chart. Accessed May 11, 2021, https://bitinfocharts.com/comparison/bitcoin-transactionfees.html.

Böhme R, Christin N, Edelman B, Moore T (2015) Bitcoin: Economics, technology, and governance. *J. Econom. Perspective* 29(2):213–238.

Buterin V (2018) Blockchain resource pricing. Accessed February 10, 2019, https://ethresear.ch/uploads/default/original/2X/1/197884012ada193318b67c4b777441e4a1830f49.pdf.

Carlsten M, Kalodner HS, Weinberg M, Narayanan A (2016) On the instability of bitcoin without the block reward. *Proc. ACM SIGSAC Conf. on Comput. and Comm. Security* (ACM, New York), 154–167.

Chen L, Cong LW, Xiao Y (2021) A brief introduction to blockchain economics. *Information for Efficient Decision Making: Big Data, Blockchain and Relevance* (World Scientific Publishing Company, Singapore), 1–40.

Cong LW, He Z, Li J (2021) Decentralized mining in centralized pools. *Rev. Financial Stud.* 34(3):1191–1235.

Croman K, Decker C, Eyal I, Gencer AE, Juels A, Kosba A, Miller A, et al. (2016) On scaling decentralized blockchains. *Proc. Internat. Conf. on Financial Cryptography and Data Security* (Springer, Berlin), 106–125.

Easley D, Kleinberg J (2010) *Networks, Crowds, and Markets: Reasoning About a Highly Connected World* (Cambridge University Press, New York).

Easley D, O'Hara M, Basu S (2019) From mining to markets: The evolution of bitcoin transaction fees. *J. Financial Econom.* 134(1):91–109.

Edelman B, Ostrovsky M (2007) Strategic bidder behavior in sponsored search auctions. *Decis. Support Syst.* 43(1):192–198.

Eyal I, Sirer EG (2014) Majority is not enough: Bitcoin mining is vulnerable. *Proc. 18th Internat. Conf. of Financial Cryptography and Data Security* (Springer, Berlin), 436–454.

Gandal N, Halaburda H (2016) Can we predict the winner in a market with network Effects? Competition in cryptocurrency market. *Games* 7(3):1–21.

Gans JS, Halaburda H (2015) Some economics of private digital currency. *Economic Analysis of the Digital Economy* (University of Chicago Press, Chicago), 257–276.

Garratt R, van Oordt MRC (2020) Why fixed costs matter for proof-of-work based cryptocurrencies. Preprint, submitted May 13, https://dx.doi.org/10.2139/ssrn.3572400.

Gencer AE, Basu S, Eyal I, van Renesse R, Sirer EG (2018) Decentralization in bitcoin and Ethereum networks. Meiklejohn S, Sako K, eds. *Financial Cryptography and Data Security*, Lecture Notes in Computer Science, vol. 10957 (Springer, Berlin).

Guasoni P, Huberman G, Shikelman C (2021) Lightning network economics: Channels. Preprint, submitted May 6, https://dx.doi.org/10.2139/ssrn.3840374.

Harvey C (2016) Cryptofinance. Preprint, submitted January 16, https://dx.doi.org/10.2139/ssrn.2438299.

Houy N (2014) The Bitcoin mining game. Preprint, submitted March 13, https://dx.doi.org/10.2139/ssrn.2407834.

Huberman G, Leshno JD, Moallemi CC (2021) Monopoly without a monopolist: An economic analysis of the bitcoin payment system. *Rev. Econom. Stud.* 88(6):3011–3040.

John K, Rivera T, Saleh F (2021) Economic implications of scaling blockchains: Why the consensus protocol matters. Preprint, submitted February 25, https://dx.doi.org/10.2139/ssrn.3750467.

Lavi R, Sattath O, Zohar A (2017) Redesigning Bitcoin's fee market. Preprint, submitted September 26, https://arxiv.org/abs/1709.08881.

Lehar A, Parlour C (2020) Miner collusion and the BitCoin protocol. Working paper, University of Calgary, Calgary, Alberta, Canada.

Lombrozo E, Lau J, Wuille P (2017) Segregated witness. Retrieved June 2017, https://github.com/bitcoin/bips/blob/master/bip-0141.mediawiki.

Malinova K, Park A (2017) Market design with blockchain technology. Preprint, submitted July 27, https://dx.doi.org/10.2139/ssrn.2785626.

Mezzetti C, Tsetlin I (2008) On the lowest-winning-bid and highest-losing-bid auctions. *J. Math. Econom.* 44:1040–1048.

Milgrom PR (1985) The economics of competitive bidding: A selective survey. Hurwicz L, Schmeidler D, Sonnenschein H, eds. *Social Goals and Social Organization: Essays in Memory of Elisha Pazner* (Cambridge University Press, Cambridge, UK), 261–292.

Myerson RB (1981) Optimal auction design. *Math. Oper. Res.* 6(1): 58–73.

Raskin M, Yermack D (2018) Digital currencies, decentralized ledgers and the future of central banking. *Research Handbook on Central Banking*, chapter 22 (Edward Elgar Publishing, Cheltenham, UK). https://www.elgaronline.com/view/edcoll/9781784719210/9781784719210.00028.xml.

Rosenfeld M (2011) Analysis of bitcoin pooled mining reward systems. Preprint, submitted December 21, https://arxiv.org/abs/1112.4980.

Rosu I, Saleh F (2020) Evolution of shares in a proof-of-stake cryptocurrency. *Management Sci.* 67(2):661–672.

Roughgarden T (2020) Transaction fee mechanism design for the ethereum blockchain: An economic analysis of EIP-1559. Working paper, Columbia University, New York.

Roughgarden T (2021) Transaction fee mechanism design. Working paper, Columbia University, New York.

Tefagh M (2021) Path-dependence of EIP-1559 and the simulation of the resulting permanent loss. Accessed May 11, 2021, https://ethresear.ch/t/path-dependence-of-eip-1559-and-the-simulation-of-the-resulting-permanent-loss/8964.

Tsoukalas G, Falk BH (2020) Token-weighted crowdsourcing, *Management Sci.* 66(9):3843–3859.

Varian HR, Harris C (2014) The VCG auction in theory and practice. *Amer. Econom. Rev.* 104(5):442–445.

Weber RJ (1983) Multi-object auctions. Engelbrecht-Wiggans R, Shubik M, Stark RM, eds. *Auctions, Bidding, and Contracting: Uses and Theory* (New York University Press, New York), 165–191.

Yao AC-C (2018) An incentive analysis of some bitcoin fee designs. Preprint, submitted November 11, https://arxiv.org/abs/1811.02351.

Yermack D (2017) Corporate governance and blockchains. *Rev. Finance* 21(1):7–31.