From Mining to Markets: The Evolution of Bitcoin Transaction Fees

David Easley, Maureen O'Hara, and Soumya Basu*

July 2017

Revised Sept. 2017

We investigate the role that transaction fees play in the Bitcoin blockchain's evolution from a mining-based structure to a market-based ecology. We develop a game-theoretic model to explain the factors leading to the emergence of transactions fees, as well as to explain the strategic behavior of miners and users. Our model also highlights the role played by mining rewards and by volume, and examines how microstructure features such as exogenous structural constraints influence the dynamics and stability of the Bitcoin blockchain. We provide empirical evidence on the predictions of our model and discuss implications for Bitcoin's evolution.

*Easley is in the Departments of Economics and Information Science, Cornell University and UTS Australia; O'Hara is in the S.C. Johnson School of Management, Cornell University and UTS Australia; Basu is in the Computer Science Department, Cornell University and the Initiative for Cryptocurrencies and Contracts (IC3). He thanks the Initiative for Crypto-currencies and Contracts (IC3) and National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1650441for support. We are grateful to Campbell Harvey, Andrew Karolyi, Talis Putnins, Gun Sirer and seminar participants at Cornell University for helpful comments and discussion.

From Mining to Markets: The Evolution of Bitcoin Transactions Fees

In a few decades when the reward gets too small, the transaction fee will become the main compensation for nodes. I'm sure that in 20 years there will either be very large transaction volume or no volume.

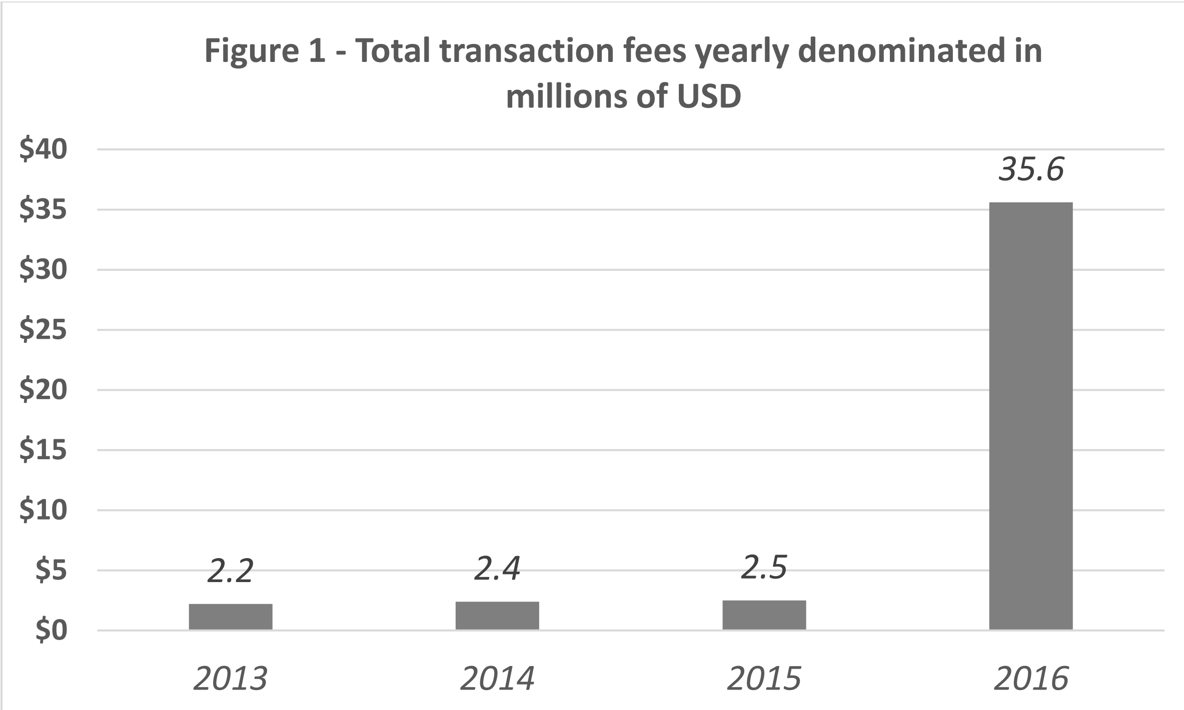
Satoshi Nakamoto[1]

By a variety of metrics, Bitcoin is no longer a financial curiosity. From its "genesis" transaction of 50 bitcoins in January 2009, bitcoins in circulation now number more than 16.3 million. There are estimated to be 35 million Bitcoin wallets held worldwide with 100,000 companies accepting payments in bitcoins, some via the newly issued bitcoin debit card. Daily trading volume in May 2017 at major bitcoin exchanges hit a record $400 million, far exceeding the previous peak of $200 million set in 2014. A recent study (see Hileman and Rauchs [2017]) estimates that 10 million people now hold a material amount of bitcoins as a financial asset. There are estimated to be more than 7,300 global Bitcoin nodes, each containing a complete copy of the Bitcoin blockchain.[2]

Envisioned in 2008 as a decentralized, trustless digital currency and payment system, the Bitcoin blockchain operates on a worldwide basis via a complex set of rules originally proposed in Nakamoto [2008]. Fundamental to the Bitcoin ecology are "miners", who play a crucial role both in creating new bitcoins and in verifying transactions on the Bitcoin blockchain. Mining involves using specialized computer hardware to solve a mathematical problem, with the reward for success being payment in new bitcoins[3]. The amount of such payments, as well as a variety of parameters such as the difficulty of the underlying computational problem and even the total amount of bitcoins that can ever be mined are specified exogenously. The Bitcoin protocol also exogenously specifies a dynamic adjustment process for these payment and difficulty parameters.

As Figure 1 illustrates, an additional form of compensation for miners has recently emerged in the form of transactions fees. These transaction fees are voluntarily appended to bitcoin transactions by buyers and sellers wanting to ensure that their

[1] Cited in https://www.bitcoinmining.com/what-is-the-bitcoin-block-reward/
[2] See https://bitnodes.21.co/
[3] More precisely, the miners use the ASIC chip to run a SHA-256 to find a particular hash function value.



This Figure gives the total amount of transactions fees paid by transactions posted to the Blockchain, denominated in millions of U.S. dollars.

Source: Blockchain.info

transactions are included in the block of transactions the miner attaches to the blockchain. The endogenous development of transactions fees reflects an important step in the evolution of the Bitcoin blockchain from being a mining-based set of rules towards being a market-based system capable of adapting to changing economic conditions. What is less clear is how successful this transition will be.

In this paper, we investigate the evolution of transactions fees in Bitcoin. We build a framework for understanding why such fees developed, and how they influence the dynamics and stability of the Bitcoin blockchain. Examining the Bitcoin ecology is complex as it involves strategic behavior on the part of both miners and users, all packaged within a set of exogenous rules. We develop a game-theoretic model to explain the factors leading to the emergence of transactions fees, as well as to explain the interactions between miners and users. Our model also highlights the roles played by volume and by mining-based revenues known as block rewards. We then provide empirical evidence in support of the predictions of our model.

Our research provides a number of results on the role and behavior of transaction fees in the Bitcoin blockchain. Our model confirms that without transaction fees over time the blockchain would not be viable as miners' profitability eventually requires transaction fees. However, even with transaction fees, there are limits on the size of the blockchain imposed by waiting times confronting users. We show how these waiting times arise in equilibrium, and how they are influenced both by endogenous transactions fees and by exogenous dynamic constraints imposed by the Bitcoin protocol. If the arrival rate of potential transactions is low, transactions without fees attached are written to the blockchaing, but as the arrival rate of potential transactions increases, the equilibrium shifts and only transactions with fees attached are posted to the blockchain. In these equilibria, transaction costs can lead to fragility of the system by inducing user non-participation: the fees directly induce some users to drop out, while increasing wait times cause other fee-paying users to depart as well. Our model suggests that these user participation effects are the reason why transaction fees alone are not a panacea for the dynamic challenges facing the evolving blockchain.

Our empirical work supports this finding. We show that higher transactions fees are being driven by queuing problems facing users, rather than by reductions in block rewards. As users battle to get transactions posted on the blockchain, transaction fees are rising to levels that discourage bitcoin usage, highlighting an important structural issue confronting the blockchain. The recent forking of bitcoin into a new currency named bitcoincash underscores the importance of these issues. Overall, our results delineate the role that transaction fees play in Bitcoin's evolution from a mining-based structure to a market-based ecology.

Our research joins a growing literature examining Bitcoin, digital currencies, and the broader applications of blockchains. A variety of authors (see, for example, Eyal and Sirer [2014]; Gans and Halaburda [2015]; and Halaburda and Gandel [2016]) analyze design issues of the bitcoin protocol, as well as the dynamic interactions between cryptocurrencies. Other recent research analyzes aspects of the bitcoin ecosystem specifically as they relate to finance and the financial markets (see Boehm et al [2015]; Harvey [2016]; Malinova and Park [2016]; Raskin and Yermack [2016]; Aune et al [2017]; Cong and He [2017]; FINRA [2017]).

There is also a literature looking at the game played by miners on the bitcoin blockchain. Huoy [2014], for example, analyzes the mining game but his analysis includes neither users nor transactions fees. Kroll et al [2013] does include transaction fees but argues they have little importance. Huberman et al [2017], in a paper written contemporaneously with ours, analyze a congestion queuing game that includes miners and fees. While similar in orientation to our work, our research differs from theirs in several important dimensions. First, their work is strictly theoretical, while our paper includes both theory and empirical work. Second, their focus is on the operation of the system with respect to infrastructure and usage fees and the determination of feasible revenue-generating mechanisms, whereas our analysis focuses on the microstructure role played by endogenous transaction fees and exogenous mining rewards. Third, our model is concerned with the evolution of these mining rewards and transaction fees in equilibrium, while their analysis looks only at the long-run steady state where mining rewards have disappeared. As mining rewards are currently the predominant source of miners' revenue, our focus is on understanding this evolution from mining to markets.

This paper is organized as follows. The next section gives a brief overview of the Bitcoin blockchain, explaining the various exogenous components of the system. Section 2 then develops game-theoretic models of the games played by the miners and by the users. We find the Nash equilibria in these games and characterize the factors influencing these equilibria, with a particular focus on the important role played by waiting times. We also characterize the equilibrium transaction fee structure, and how it influences user waiting times. Section 3 then provides empirical analyses based on predictions from our model. Section 4 concludes by discussing the current structural challenges confronting Bitcoin and the role played by transaction fees.

I. A Brief Overview of the Bitcoin Blockchain Protocol

As noted in the Introduction, the Bitcoin blockchain set out by Nakamoto [2008] involves a decentralized trustless network in which miners validate transactions and provide the requisite security for the blockchain. Miners are computers using dedicated hardware and software to solve a computational problem. This computation problem essentially involves using a brute force approach to find a specific "hash function" or string of numbers.[4] The first miner to solve the problem is compensated with a fixed number of newly issued bitcoins known as a block reward. The winning miner is then allowed to "post" a block (or collection) of pending bitcoin transactions to the blockchain. The number of transactions put into a block can vary from zero up to a maximum block size currently set at 1MB.[5]

This validation process of posting transactions to the blockchain may also provide another source of revenue to the winning miner via transaction fees embedded in pending bitcoin transactions. In particular, the timing of the validation process involves miners first gathering up pending transactions into a block, then racing to be the first miner to solve the computation problem, and, if successful, finally appending the transactions in the block to the blockchain. Transactions specify the number of bitcoins to be taken from one account and the number to be transferred to another account, and any difference in the two numbers can be kept by the miner as a transaction fee. Note that the decision to append a fee is not specified by the bitcoin protocol but rather is up to the participant in the underlying transaction. We analyze this fee decision in more detail in Section 3.

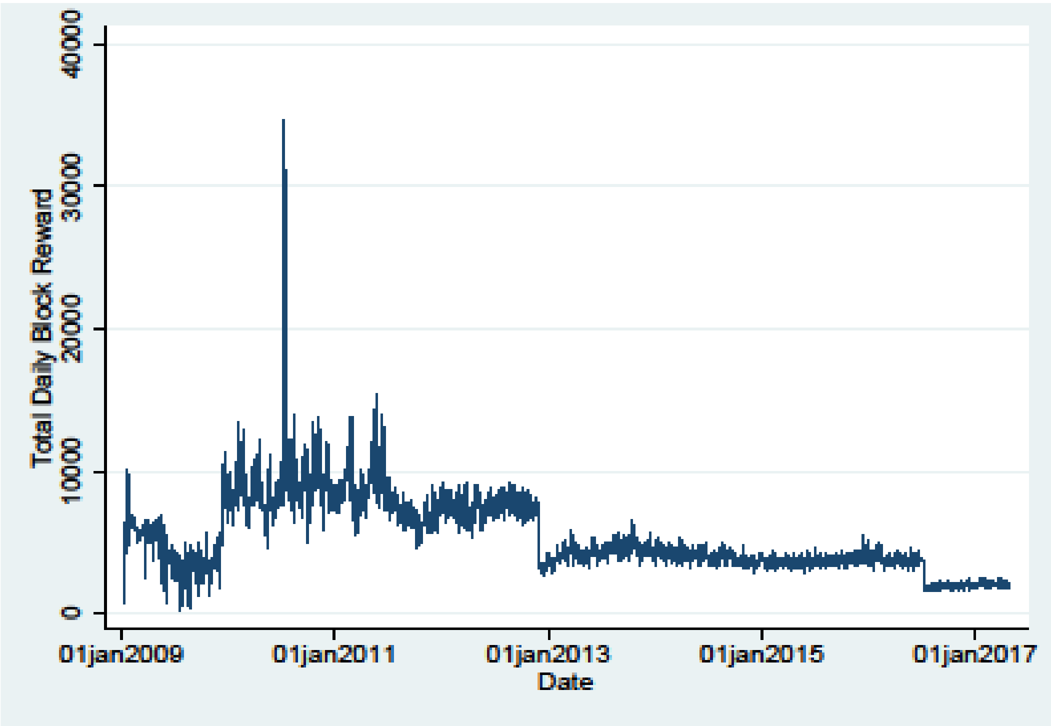
The bitcoin blockchain mechanism is highly structured, with the total number of bitcoins available for issuance limited to 21 million. Because issuance is tied to mining, the block reward is set to decline as the number of bitcoins in existence grows. In particular, the block reward is halved after every 210,000 blocks are mined. Figure 2 shows the daily total block reward earned by miners starting with 2009. The breaks in the figure in November 2012 and July 2016 correspond to the block reward being reduced from its initial level of 50 to 25, and then from 25 to its current level of 12.5. The reward will be reduced 32 more times before eventually reaching zero.[6]

[4] More precisely, passing the transactions augmented with a random number through a cryptographic hashing function and cycling through different random numbers to try to produce a rare result that has many leading zeros.

[5] This block size limit was reduced from 36 MB to 1MB in 2010 to counter risks from spam and denial of service attacks. The optimal block size is currently an issue of controversy in bitcoin circles, and the inability to reach a consensus on the optimal block size led to the "forking" of the blockchain on August 1, 2017 into two blockchains. We discuss the implications of this issue in Section 4. For more technical details see http://www.coindesk.com/what-is-the-bitcoin-block-size-debate-and-why-does-it-matter/

[6] For more discussion of mining rewards see "Bitcoin miners face fight for survival as new supply halves", Reuters, July 8, 2016

Figure 2---Daily total block rewards

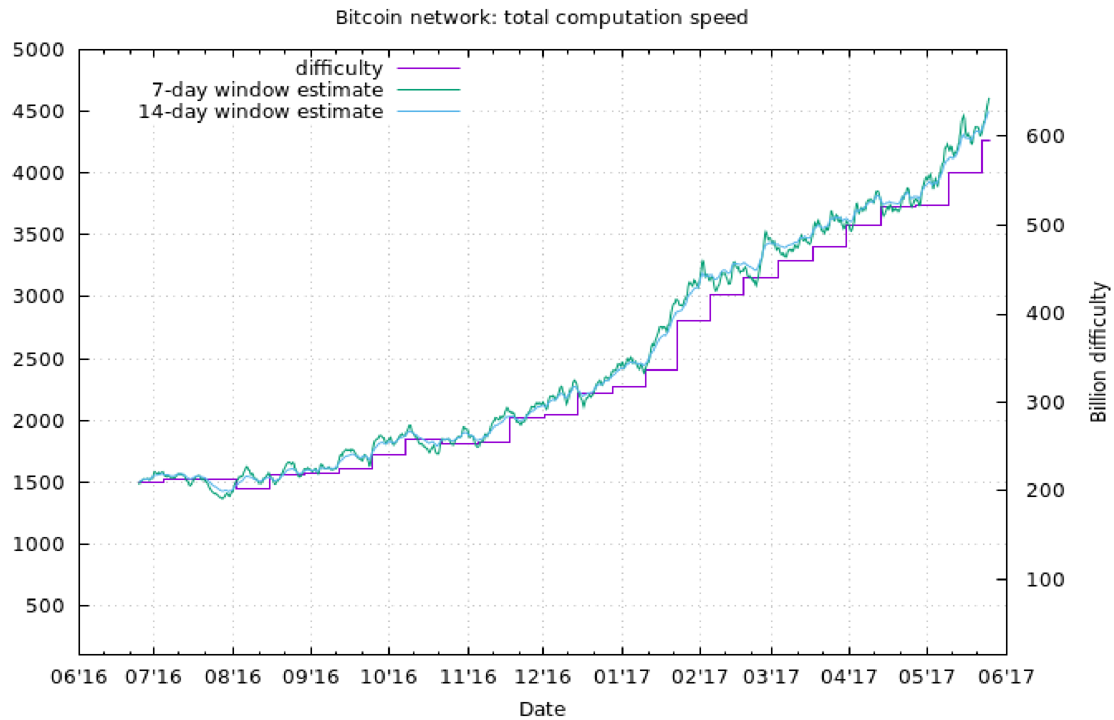


This figure shows the total daily block rewards earned by miners, denominated in bitcoins. The level of the block reward was halved in Novenber 2012 and again in July 2016.The data are extracted from our blockchain node.

The rate at which new bitcoins are issued is also affected by the difficulty of solving the computational problem.[7] In general, the algorithm sets the difficulty level such that on average a block is added to the blockchain every 10 minutes, or approximately 144 blocks a day. If new blocks are added faster than this desired level (perhaps because of an increase in the number of miners or an advance in mining technology), then the difficulty is increased, and conversely it is decreased if the rate is too low. This adjustment is made every 2,016 blocks, or approximately every 14 days. Figure 3 shows the changes in this difficulty level.

[7] Difficulty is a measure of how difficult it is to find a hash below a given target. For more information see https://en.bitcoin.it/wiki/Difficulty

Figure 3---The time path of difficulty levels in the Bitcoin network.



This table shows the adjustment in the difficulty level of solving for the hash function in the bitcoin blockchain. The difficulty level is adjusted to keep the rate at which blocks are posted to the blockchain at approximately 144 blocks per day. Adjustments (if needed) occur after every 2016 blocks.

Source: http://bitcoin.sipa.be/

Transactions to be posted to the blockchain are originally broadcast across the various nodes (designated as either full or partial nodes) in the decentralized Bitcoin network. Full nodes contain an exact copy of the Bitcoin blockchain, and each full node will generally have a holding tank, called a "mempool", in which these transactions are held pending their inclusion in a block. Transactions flow into the mempool from these broadcasts and they leave the mempool when they either are posted to the blockchain or are dropped from the pool (typically after 3 days or so) if their fee is too low to attract a miner.[8] As the volume of transactions in the bitcoin network increases, the flows into these mempools also rise. The flows out of the mempool, however, are circumscribed by the maximum block size and the current difficulty level. The Bitcoin mechanism provides no natural means to equilibrate these potentially disparate flows.

Given this protocol, two problems are apparent. The deterministic decline in the block reward results in miners' revenues (at least as defined in terms of bitcoins) also falling deterministically over time, raising the potential that miners may be unwilling to perform the costly calculations needed to validate the blockchain. The rise in Bitcoin transaction volume, coupled with limits on the number of blocks that can be posted to the blockchain, raise the possibility that some, potentially many, transactions may never be posted to the blockchain, undermining the willingness of users to transact in bitcoin.

With this as a backdrop, the question of interest is whether fees appended to Bitcoin transactions represent a potential market solution to these mechanism shortcomings. In the next section, we develop a game theoretic model to investigate the dynamics of the Bitcoin blockchain and how it is affected by endogenous transaction fees. Our focus is on understanding the evolution from mining to markets in the Bitcoin ecosystem.

II. Model

The main participants in the Bitcoin blockchain are miners and users. Miners are individuals, often operating in pools, who solve computational problems allowing them to put transactions on the blockchain and reap the block reward as well as any fees attached to those transactions. There is free entry into the mining game, so miners play a standard entry-exit game in which any Nash equilibrium will entail zero profit. Users submit transactions to the mempool(s) that they want placed on the blockchain. They play a game in which they can chose to pay a fee to move up in the queue and thus reduce their waiting time, or not pay a fee and experience a longer waiting time. The Nash equilibria of this game can involve none, some, or all of the users paying a fee. Which equilibria occur (there can be multiple

[8] Some nodes have rules over which transactions they will include in their mempool. These restrictions may arise due to limitations on the mempool size or due to setting minimums on the level of transaction costs that transactions must offer. For details, see https://support.21.co/bitcoin/mempools/what-is-the-mempool-size

equilibria) depends on the parameters of the problem and, most importantly, on the waiting times for fee-paying and non-fee-paying users.

A. Miners

Miners are indexed by $m=1,\dots,M$. In our model, M is endogenous. We determine the Nash Equilibrium value of M after solving the mining problem for a fixed M. We assume that all miners are identical in that they have a common fixed cost of mining $F > 0$; incur a depreciation cost of a common amount δ of F for each problem they work on; and have a common variable cost per unit of time spent working on a problem of $e > 0$.[9] These costs are denominated in a common currency which, for simplicity, we denote as dollars. At any moment, the miners are all working on finding a particular hash and having the right to put the next block on the chain. We assume that independently, for each miner working on the problem, success in solving the problem arrives according to a Poisson distribution with arrival rate λ. The miner who first solves the problem gets to put the next block on the chain. Because miners are symmetric and assumed to have identical hashing power, the probability that any individual miner is the first to solve the problem is $1/M$.

The winning miner receives a block reward of $S > 0$ bitcoins. In addition, the winning miner places the block of B transactions from the mempool onto the blockchain, thereby collecting any fees embedded in those transactions.[10] Let f_i be the fee embedded in transaction i, where all fees are also denominated in bitcoins. We assume that in forming blocks miners choose among transactions according to their fees, taking the ones with the greatest fees first and then proceeding down the list of fees until the block is full.[11] If the last transactions added to the block come from a collection of transactions with a common fee (which may be 0) the miner can choose at random among those transactions. Abusing notation slightly we let B be the block of transactions written to the chain. So the revenue earned in bitcoins by the successful miner is $S+\sum_{i\in B} f_i$. Denoting the exchange rate between the bitcoin

---

[9] As our focus is on transaction fees, our characterization is purposely simplified to capture the basic features of mining. Actual mining involves difference in costs arising from differences in computing technology, cost of electricity, and scale, all of which would be expected to influence specific miner fixed and variable costs. Moreover, mining operations now often feature mining consortia, further complicating entry and exit into the field. These issues are addressed in more detail in Eyal and Sirer [2014].

[10] We treat all miners as having access to identical mempools and refer to this common pool as "the mempool".

[11] For simplicity, we assume that all transactions are the same size, although in reality some transactions require more of the block capacity than others.

value and the dollar as p, the revenue earned by the successful miner in dollars is thus $p(S+\sum_{i\in B} f_i)$. [12]

Miners are assumed to be risk neutral. They play a simple game in which they independently chose to participate or not. We assume that there is free entry into mining, thus in a Nash equilibrium every miner must be making zero expected profit given the choices of all other potential miners. [We ignore integer issues, but with integer constraints the equilibrium condition is that each miner has a non-negative profit and, given the choices of all other miners, would have a non-positive profit from any alternative choice. Thus, the number of miners we find is within one of the equilibrium number expressed as an integer.]

Any miner's expected revenue from attempting to solve the problem is $p[S+\sum_{i\in B} f_i]/M$. There are M miners working on the problem with individual success rates λ , so the expected time until the first success is $1/_{\lambda M}$. Thus, each active miner's expected profit is $\frac{p[S+\sum_{i\in B} f_i]}{M}-\frac{e}{\lambda M}-\delta F$.

In a Nash equilibrium, the expected profit to mining must be 0 as otherwise there would be entry or exit of miners. So in an equilibrium with any active miners (*M>0*) the number of miners must be such that

(1)
$$\frac{p[S+\sum_{i\in B} f_i]}{M}-\frac{e}{\lambda M}-\delta F=0$$

Note, however, that if $p[S+\sum_{i\in B} f_i]-e/_{\lambda}<\delta F$, then no miner can make a non-negative profit even if there are no other miners, and in equilibrium there are no active miners. In this case, the value of M that solves (1) is less than one and we say that the equilibrium number of miners is 0. However, for the blockchain to be secure some number of miners greater than one is necessary. There is debate about the maximum fraction of mining that can be controlled by one miner, or mining consortium, before security becomes an issue.[13] For our purposes, it is sufficient to

[12] We do not take into consideration any delay created by adding transactions to the block. If fees are very low this delay makes it likely that miners put few if any transactions into blocks. As fees grow over time they provide an incentive for miners to fill blocks as they currently do. If, as a first approximation, all miners are working on the same problem then they all face the same delay.

[13] See Eyal and Sirer [2014] for analysis of how many miners are needed to keep the bitcoin blockchain viable.

set the minimum number of miners necessary for the mining and the blockchain to be viable at $\bar{M} > 1$. So, if the number of miners implied by the zero-profit condition is less than $\bar{M}$ we say that mining is not viable, the blockchain fails, and we set the number of miners to 0.

For any given success rate, λ, equation (1) determines the equilibrium number of miners if that number is at least $\bar{M}$. However, the success rate λ is not exogenous. The protocol sets the difficulty of the mining problem so that the arrival rate of new blocks to the chain occurs at approximately some exogenously fixed rate Λ. If there are M miners, each with independent individual success rates of λ, the arrival rate of blocks is λM. So, in equilibrium, the difficulty of the problem, λ, depends on the number of miners

(2) $$\lambda M = \Lambda$$

Of course, this requires that there are active miners. If not, then the success rate is undefined.

Result 1: The equilibrium number of miners and arrival rate of success is the simultaneous solution to equations (1) and (2)

(3)
$$M^* = \begin{bmatrix} \dfrac{p[S + \sum_{i \in B} f_i]}{\delta F + e/\Lambda} \;\; if \geq \bar{M} \\ 0 \;\; otherwise \end{bmatrix}$$
$$\lambda^* = \begin{bmatrix} \Lambda / M^* \;\; if \;\; M^* \geq \bar{M} \\ undefined \;\; otherwise \end{bmatrix}$$

The number of active miners is thus the ratio of the revenue earned by a successful miner to the expected cost of mining. Note that changes in revenue, $p[S + \sum_{i \in B} f_i]$, that leaves $M^* \geq \bar{M}$ affects the equilibrium number of miners, but not the rate at which blocks are added to the chain. Over time the block reward, S, will decline to 0 so, absent fees, mining eventually would fail because miners would be unable to recover their costs. The timing of when this failure will occur is not straightforward as miners care about total revenue and the declining block reward level may be offset (or more) by rising bitcoin prices, at least until the block reward reaches 0. However, as Satoshi predicted, fees have emerged and new blocks continue to be added to the chain. The emergence of fees is the topic of the next section.

Remark 1: (a) With no transactions fees and a constant bitcoin price, a decrease in the block reward S to a level at which mining is still viable reduces the number of miners, and reduces the difficulty level λ, but leaves the rate at which new blocks are added to the blockchain, Λ, unchanged; (b) With a constant block reward level and viable mining, an increase in the bitcoin price or transaction fees increases the number of miners and increases the difficulty level λ, but leaves the rate at which new blocks are add to the blockchain , Λ, unchanged.

B. Transactions and the Mempool

A user who wants to record a transaction moving bitcoins from one account to another account submits a transaction to the mempool. However, that transaction only succeeds in moving bitcoins if it is actually written to the blockchain. We suppose that there are N potential users and the opportunity to engage in bitcoin transactions arises independently for each of them at Poisson rate γ.[14] So transactions flow into the mempool at rate γN . Initially, we consider a world in which no transactions offer a fee to miners, and so miners pick transactions at random from the mempool when they create a block. For notational convenience we also treat a block as consisting of one transaction, B=1.

Transactions flow into the mempool at rate γN and flow out at rate $\lambda^* M^*$. We are interested in the number of transactions in the mempool and in how long transactions wait on average before being recorded to the blockchain. This can be viewed as a queueing problem with random service as when a miner builds a block he selects from the mempool at random instead of taking the transaction in the pool which has been waiting the longest as in a standard first-in, first-out queue.[15] The order in which transactions are removed from the pool does not affect the expected size of the pool, so standard queueing theory results can be applied to determine the size of the pool.

The dynamics of the pool are straightforward. If the arrival rate of transactions is greater than the rate at which transactions are removed from the pool, then the size of the pool grows without bound. Alternatively, if $\gamma N < \lambda^* M^*$ then there is an

[14] Initially we take the number of potential users, N, as fixed and we focus on parameter ranges in which they all chose to participate in bitcoin. In section D we determine N endogenously.

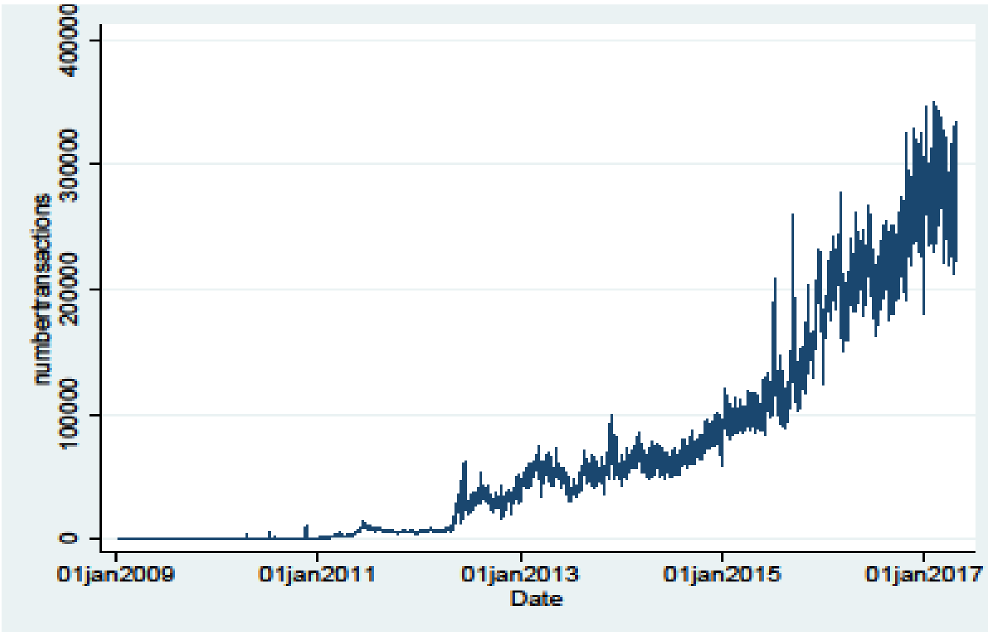
[15] The lack of time priority in the mempool is addressed in Aune, Krellenstein, O'Hara, and Slama [2017].

equilibrium distribution of pool size with mean $N^* = \rho / (1-\rho)$ where $\rho = \gamma N / \lambda^* M^*$. By Little's Law [1961] the long run expected waiting time for a transaction in the mempool to be recorded is $w^* = N^* / \gamma N$.

Result 2: If $\gamma N > \lambda^* M^*$ then the mempool grows at rate $\gamma N - \lambda^* M^*$ and waiting times diverge. Alternatively, if $\gamma N < \lambda^* M^*$ then there is an equilibrium distribution of mempool size with mean $N^* = \rho / (1-\rho)$ and mean waiting time $w^* = [\lambda^* M^* (1-\rho)]^{-1}$ where $\rho = \gamma N / \lambda^* M^*$.

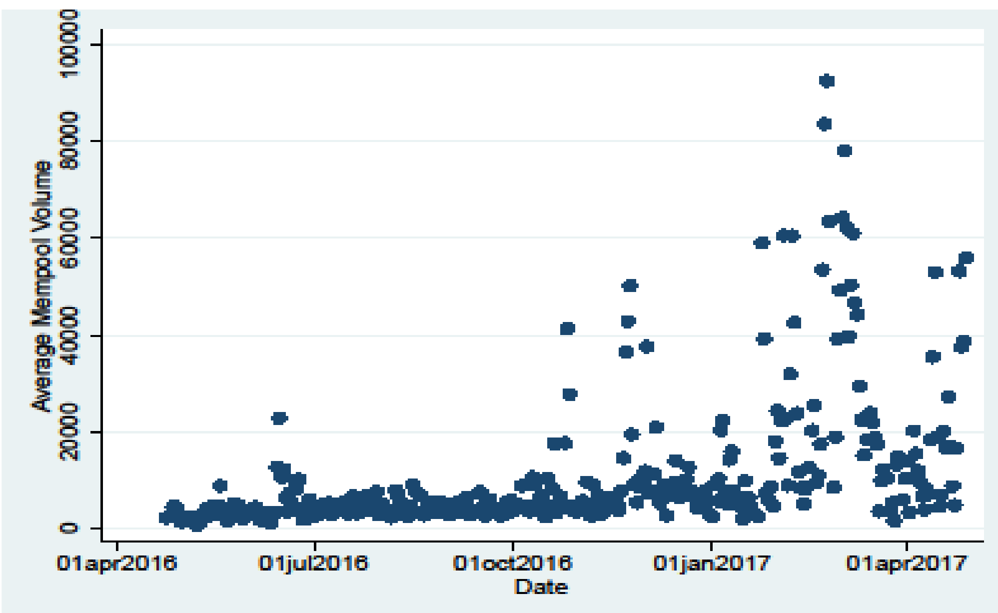
This analysis is done with fixed rates of transactions flowing into and out of the mempool and it focuses on the equilibrium pool size and waiting time. Actually, as Figure 4 shows, over time the rate at which transactions have been flowing into the mempool has increased and the rate at which they are flowing out is not increasing commensurately. As a result, ρ is growing, and as it approaches 1 the expected size of the mempool grows and average waiting times diverge. Figure 5 illustrates the growth in the size of this pool. This is a problem for users, as a transaction does not result in bitcoins moving from one account to another until it is written to the blockchain. Or, put another way, potential transactions not eventually written to the blockchain simply cease to exist.

Figure 4---The number of Bitcoin transactions per day 2009 – 2017



These data are Bitcoin transactions drawn from our blockchain node and reflect the daily number of on-chain transactions.

Figure 5---The size of the mempool daily July 2015- July 2017



This figure shows the size of the mempool on a daily level. The data are estimates provided by blockchain. info

C. Waiting Times and Transaction Fees

Once waiting times become significant, one might expect some users to try to get their transactions recorded ahead of others. One way that a user could do this is to implicitly attach a fee to their transaction by moving more bitcoins out of one account than they move into another account; the difference can be kept by the miner who writes the transaction to the blockchain.[16] This difference can be interpreted as the fee that is offered to the miner who records the transaction. As noted earlier, as long as mining is viable, an increase in fees does not affect the rate at which transactions are written to the blockchain. Fees increase the number of miners, but the protocol increases the difficulty of the problem to keep constant the rate at which successes occur and blocks can be created.

We view the users as playing a game in which each user decides whether to offer a fee taking as given the decisions of all other users and the rate at which transactions are written to the blockchain. To determine the Nash equilibrium fraction of users offering fees, we assume that any transaction written to the blockchain generates a benefit V to the user but that this benefit is reduced by the amount of any fee paid and by the waiting time. We let user expected payoffs be $V - pf - aw$ for a user who pays a fee in bitcoins of f, with dollar value pf, and whose transaction has an expected delay of w periods before it is recorded to the blockchain. A user who does not pay a fee still gets the benefit of V if his transaction is recorded to the blockchain, but this benefit is reduced by his longer waiting time. We initially consider only a single fee level $f > 0$ and we focus on parameter ranges in which the net benefit of being a user is positive.

The fee game can, in principle, have Nash equilibria in which none, some, or all users pay the fee. We consider each of these possibilities in turn. If all users offer the same fee, then nothing changes other than the equilibrium number of miners. Alternatively, if some users offer higher fees than other users offer, then their transactions are selected ahead of those who offer lower, or zero, fees.

If no one offers to pay a fee and the expected waiting time is w, then no user deviates to paying fee f in order to be first in the queue (rather than being randomly selected) if the reduction in expected waiting time is too small. A user who is selected first (as only he offers to pay the fee) has an expected waiting time of

[16] Actual transactions can be more complex in that they specify multiple parties, i.e. x btc paid to seller A, y btc returned to buyer B, and the remainder retained by the miner).

$(\lambda^* M^*)^{-1}$, the expected waiting time for the first miner to solve the problem. So there will be Nash equilibrium in which no fees are paid if

(4) $$V - pf - a[\lambda^* M^*]^{-1} < V - aw^*$$

At the equilibrium with mempool size N^* this is

(5) $$pf/a > \frac{\rho}{(1-\rho)\lambda^* M^*}$$

If the mempool is not too large ($N^* = \rho/1-\rho$), and thus mean waiting times are short, no one offers to pay a fee. But once the mempool is large enough, and thus mean waiting times are large, there is no Nash equilibrium in which no one pays a fee.

Alternatively, if the mempool is large and waiting times are sufficiently long, then in a Nash equilibrium everyone pays the fee. That is, if all other users offer to pay the fee, no one wants to deviate to not paying a fee. Calculation shows that this occurs if

$$pf/a < \frac{\rho}{\lambda^* M^* (1-\rho)^2} \quad .$$

To derive equilibria in which some, but not all, users pay the fee we need to compute the waiting times for fee and no-fee transactions. Users who offer a fee have their transactions recorded before any transactions without fees. So the waiting time for fee-paying users is not affected by the number of non-fee-paying users. Suppose that some fraction α of users offer to pay the fee. [We ignore integer constraints on this fraction as the number of users is large and the constraints do not add any insights to the analysis.] The expected waiting time for these users is thus the waiting time in Result 2 with the arrival rate of transactions reduced to $\alpha\gamma N$

(6) $$w_f(\alpha) = [\lambda^* M^* (1 - \rho_f(\alpha)]^{-1}$$

where $\rho_f(\alpha) = \alpha N\gamma / \lambda^* M^*$. Note that, $\lambda^* M^*$, the equilibrium rate at which transactions are written to the blockchain, is not affected by fees.

The non-fee-paying users face a different situation as their expected waiting time depends on both the number of fee-paying and non-fee-paying users. To derive this waiting time, note that fees do not affect the long run mean size of the mempool; it remains at $N^* = \rho / 1-\rho$.

Remark 2: Suppose that fees and waiting times are such that all mining is viable and that all N potential users continue to participate. Then fees do not affect the equilibrium size of the mempool or the equilibrium rate at which transactions are written to the blockchain.

Let $\rho_n(\alpha) = (1-\alpha)\gamma N / \lambda^* M^*$ and note that for any $\alpha \in [0,1]$, $\rho = \rho_f(\alpha) + \rho_n(\alpha)$. Letting the equilibrium numbers of fee-paying and non-fee paying users be $N_f(\alpha)$ and $N_n(\alpha)$, respectively, we have

$$N^* = N_f(\alpha) + N_n(\alpha) = \frac{\rho_f(\alpha) + \rho_n(\alpha)}{1-(\rho_f(\alpha) + \rho_n(\alpha))}$$

As $N_f(\alpha) = \dfrac{\rho_f(\alpha)}{1-\rho_f(\alpha)}$ we have

$$N_n(\alpha) = \frac{\rho_n(\alpha)}{(1-\rho_f(\alpha))(1-\rho)}$$

Then by Little's Law the expected waiting time for non-fee-paying users is

$$w_n(\alpha) = [\lambda^* M^* (1-\rho_f(\alpha)(1-\rho)]^{-1} \quad . \tag{7}$$

Calculation shows that for any $\alpha \in (0,1)$, $w_n(\alpha) > w^* > w_f(\alpha)$. So fee-paying users do not wait as long as non-fee-paying users. We next use these waiting time results to describe Nash equilibria with a mix of fee-paying and non-fee-paying users.

There is a Nash equilibrium in which an interior fraction of users pay the fee if and only if

$$V - pf - aw_f(\alpha) = V - aw_n(\alpha)$$

for some $\alpha \in (0,1)$. This is possible if and only if

$$\frac{\rho}{(1-\rho)^2 \lambda^* M^*} > pf / a > \frac{\rho}{(1-\rho)\lambda^* M^*}$$

and in this case the Nash equilibrium fraction of users paying the fee is
$\alpha^* = \rho^{-1}[1 - \frac{a\rho}{pf(1-\rho)\lambda^* M^*}]$.

These conclusions are summarized in the Result below.[17]

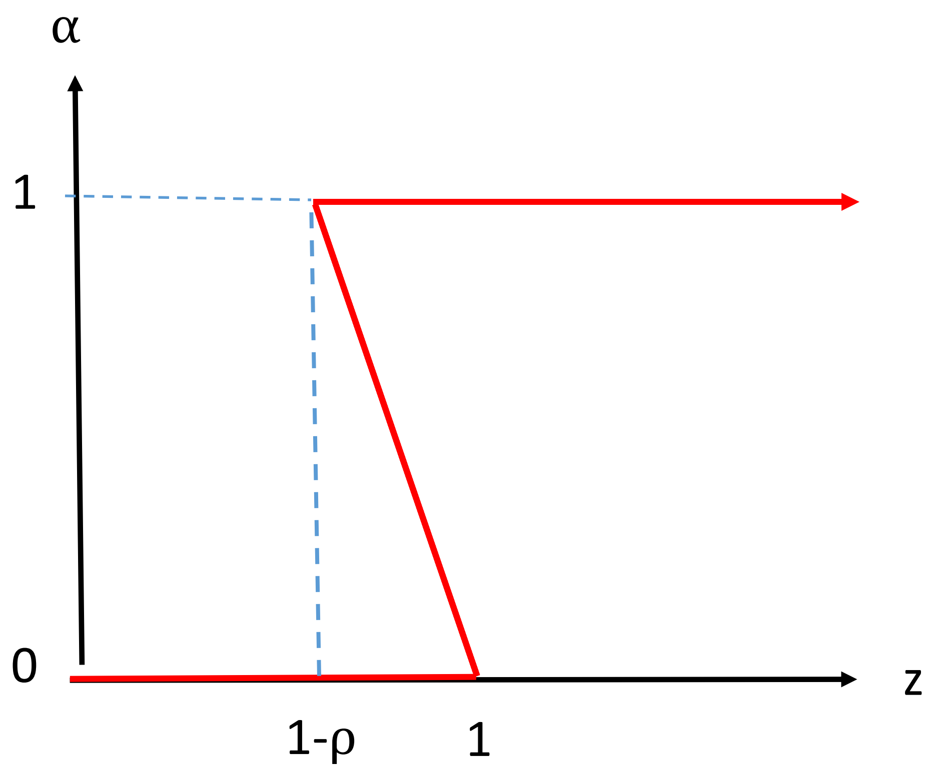
Result 3: Suppose parameters are such that mining is viable and all N potential users continue to participate. For any parameters in this range, the fee-paying game has at least one Nash equilibrium. Let $z = \frac{a\rho}{pf(1-\rho)\lambda^* M^*}$. Nash equilibria in the fee-paying game are described by:

1. If $z < 1-\rho$, the unique equilibrium is one in which no user pays a fee, ($\alpha = 0$).
2. If $z > 1$, the unique equilibrium is one in which every user pays a fee, ($\alpha = 1$).
3. If $1-\rho \le z \le 1$, there are three equilibria: $\alpha = 0$, $\alpha = 1$ and one in which an interior fraction of the users pay a fee.

The parameter z summarizes the costs and benefits to the user of paying the fee (pf/a), the no-fee mean size of the mempool ($\rho/1-\rho$) and the expected waiting time for a miner to arrive. Note that given the rate at which transactions flow out of the mempool (which is bounded by the maximum block size specified by the Bitcoin protocol) the factor that drives changes in the size of the mempool and thus waiting times is the arrival rate of transactions. If that rate increases over time then z increases. For small and large values of z the game has a unique equilibrium in which either no user pays a fee or all users pay a fee. For intermediate values of z there are three equilibria: no-fee, all-fee and an interior fraction paying the fee.

[17] We consider cases in which mining may not be viable or some users chose not to participate in section D.

Figure 6---Nash equilibrium and Fee-paying Users



This figure relates the fraction (α) of users paying a fee as function of the parameter $z = \frac{a\rho}{pf(1-\rho)\lambda^* M^*}$, where z summarizes the costs and benefits to the user of paying the fee (pf/a), the no-fee mean size of the mempool ($\rho/1-\rho$), and the expected waiting time for a miner to arrive.

Figure 6 summaries these results, with the red curve describing Nash equilibrium fractions of users paying a fee as a function of the parameter z. This graph also has interesting implications for the dynamics of fees. Suppose that over time indexed by t=1, 2, … the arrival rate of transactions varies exogenously and thus the parameter z_t varies exogenously. If initially the arrival rate is low, then z_t is low and the only equilibrium is one with a zero fraction of transactions paying a fee, i.e. $\alpha_t = 0$. As z_t increases over time, the equilibrium in the graph, (z_t, α_t), moves along the lower branch of the figure and if the dynamics are locally smooth the system stays on this branch until $z_t = 1$. At this point, any further increase in the arrival rate of transactions causes the equilibrium to jump to the upper branch. Note, however, that if the dynamics are again locally smooth along this upper branch then if z_t falls the equilibrium fraction of users paying a fee stays at 1 until z_t falls all the way to $1-\rho$. That is, once an equilibrium with fees being paid is established it is locally stable. Of course, this model has many simplifying assumptions and we would expect real dynamics to be more complex, but assuming that our critical assumptions capture the main points of the bitcoin system our implications for dynamics should be approximately correct.

D. Extensions

The model used in previous sections is simplified so that we could illustrate our points about Bitcoin without distractions that added little to the analysis. Here we consider a few extensions and show how they modify the details of our analysis, but do not change the qualitative points we make.

1. Multiple Fees and Heterogeneous Users

The analysis in section C focused on a single fee and users who chose either to pay this fixed fee or to offer no fee. This analysis can be generalized to multiple fee levels. Suppose that users consider fee levels (in bitcoins) of $f_1 > f_2 > \cdots > f_K = 0$ and let the fraction of users who chose fee f_k be α_k. Then the $\alpha_k N$ users who pay the highest fee are selected first, those who pay the second highest fee are selected next once the highest fee-paying users are removed, and so on. Their waiting times can be defined recursively as in section C. If users are identical, then having positive fractions paying multiple fee levels requires indifference and thus fractions at each level that equalize the costs and benefits of each fee level.

More realistically, users may be heterogeneous with respect to their utility cost of waiting time. In this case, equilibria in which users are separated by fee levels are possible. This possibility can be illustrated even with only two fee levels, *f* and 0, as in section C. Suppose that there are two types of users with waiting costs $a_1 > a_2$ and that fraction α_1 of the users have the high waiting cost. Impatient users will pay the fee, f, and patient ones will pay no fee if

(8) $$pf/a_2 \geq w_n - w_f \geq pf/a_1$$

where w_n, w_f are the waiting times for non-fee-paying and fee-paying users respectively. Calculation shows that

(9) $$w_n - w_f = \frac{N^*}{\lambda^* M^* - \alpha_1 N\gamma}$$

The right-hand side of this equation is the equilibrium size of the mempool divided by the difference in rates between the flow out of the mempool and the flow of impatient users into the mempool. So inequality (8) relates the difference in waiting costs to the difference between the flow out of the pool and the flow of impatient users into the pool.

2. Exit by Users

In section C we focused on equilibria in which all potential users chose to submit transactions to the mempool whenever opportunities arise. However, if waiting times are too long, or if some transactions never get recorded on the blockchain, those users may chose not to submit transactions to the pool; presumably, they would instead conduct their transactions in an alternative currency. Remark 1 implies that the dollar value of the fee necessary to make mining viable is (normalized to the fee versus net cost per transaction added to the blockchain)

(10) $$pf \geq p\bar{f} = \begin{bmatrix} \delta F + e/\Lambda\bar{M} - pS/\bar{M} \;\; if \;\; \delta F + e/\Lambda\bar{M} - pS/\bar{M} \geq 0 \\ 0 \;\; otherwise \end{bmatrix}$$

At this fee per transaction and the waiting time if all users participate,

$N^* = \rho/(1-\rho)$, the net benefit of transactions may be so low that not all users will chose to participate. To see this, suppose that users can obtain net benefit $\bar{v}$ from conducting their transaction without using bitcoin. Then the waiting time for a user paying fee *f* must be no greater than $V - \bar{v} - pf/a$ as otherwise the user would not submit the transaction to the mempool. As we have an exogeneous lower bound on *f*, this implies an upperbound on waiting times of $\bar{w} = V - \bar{v} - p\bar{f}/a$. This upperbound on waiting times, in turn, implies an upperbound on the number of users who would be willing to participate in using bitcoins for transactions

(11) $$N \leq \frac{\lambda^* M^* - \bar{w}^{-1}}{\gamma}$$

So the equilibrium flow of transactions into the mempool (γN) is bounded by the flow of transactions being written to the blockchain ($\lambda^* M^*$) minus the inverse of the maximum waiting time that users are willing to bear.

In summary, our model provides a variety of insights into the driving forces behind the dynamics of the Bitcoin blockchain. Most importantly, it provides a framework for understanding the emergence of fees as the rate of arrival of transactions and thus waiting times have increased. It also suggests how fees play an important role in the viability of mining and why fees, in turn, limit the potential growth of bitcoin usage. Finally, it demonstrates why a diversity of fees emerges as heterogeneous users use fees to battle for shorter waiting times. In the next section, we provide empirical evidence on these issues.

3. Empirical Analysis

Our model suggests a variety of empirical relationships affecting the operation and stability of the Bitcoin blockchain. In this section, we examine predictions of our model with respect to the role played by transactions fees. Specifically, we provide empirical evidence on the predicted relationship between waiting times and zero-fee equilibrium, and we investigate the predicted relationship between model variables and transaction fee levels.

Before turning to the data, it is useful to raise several important caveats. One is that volume is very hard to measure accurately in Bitcoin because you can only count how many bitcoins go from one account to another.[18] A single person often has many accounts for privacy purposes, so transfers between such accounts will add to total reported volume. As Bitcoin has grown, it is likely that the clientele using Bitcoin has also changed and that can influence volume statistics. For example, an increasing fraction of users with single wallets would reduce the double counting relative to earlier years, while an increased focus on privacy could have the opposite effect by increasing the number of multiple wallet users. Having noted this, it is standard to use reported volumes and we do so in our empirical work.

A second caveat is that due to the newness of the Bitcoin system, the quality and quantity of some data series are limited. In particular, while data drawn directly from the Blockchain exists from the genesis transaction, the Blockchain does not record many variables of interest to our study. Data series such as mempool

[18] This concern relates to calculating on-block volume. There is also off-block volume that does not affect the block chain, but this volume is essentially impossible to calculate.

volume and average waiting times can be found from other sources, but often feature non-existent or spotty data in earlier time periods and more complete data in recent periods. This results in series having missing data. In our empirical work, we adjust for this when appropriate, and restrict our sample periods to only complete data intervals when data quality concerns are too high. The non-availability of data from the earlier periods of bitcoin limits our ability to address some of the more interesting dynamic issues revealed by our model.

Finally, a third issue has to do with censored data series. Data recorded to the blockchain only comes from transactions actually placed in blocks by miners. Transactions in the mempool that do not make it onto the blockchain are not captured in any currently available data set. For many empirical questions, this may not be a problem, but for others, such as analyses of confirmation waiting time or transaction fees, this censoring is an issue. Mempool waiting time data, for example, only captures those "successful" transactions, and omits the ones that miners opt not to select, imparting a downward bias to the waiting time data series. Conversely, transaction fee data may be upwardly biased because transactions that attach too low a fee to attract a miner are excluded. More complete data series may emerge over time, but for now these issues preclude testing some of the predictions of our model.

A. Data and Summary Statistics

The data for our analysis come from several sources. The Bitcoin blockchain stores all transaction information from when the system began, and a copy of this information is stored on every node so that it can validate new transactions as they come up. We started a bitcoin node and allowed it to download and validate all of the transactions from when Bitcoin started up until the end of April 2017. Each transaction in the bitcoin node had the following information: time when the transaction settled, the difficulty of the block, the value of the coins transferred, the block reward, and the mining fees paid. We obtained most of the information in this paper directly from the blockchain or inferred from data on the blockchain. To answer transient queries, we use an outside data source: blockchain.info, which runs many bitcoin nodes and reports on various metrics about the network as a whole. We use blockchain.info for the inflow rate of transactions, the size of the transaction mempool and the median confirmation time for a transaction. Finally, we looked at a popular exchange, CoinDesk, to get the bitcoin to USD conversion rate. Since different exchanges have different exchange rates, we simply used the average rate across all exchanges for this paper.

Table 1 provides summary statistics for bitcoin mining and transactions over the period 2011-2016. The data show that Bitcoin transaction volume has grown over this period, but growth has not been steady. Volume in 2016 is approximately the same as in 2012, it fell in 2013 and 2014, and then resumed its growth. Mining fee revenues (the coinbase rewards) decreased consistently over this period, while total transaction fee revenues at first increased (particularly if one looks back to 2010 and 2011), then fell, and are increasing again. Transaction fees in 2016 were still a small fraction of total miner revenue with block rewards averaging nearly 50 times as large as fees. That ratio is, of course, declining as the block reward declines to zero and fees grow, and more recent data suggest fees now comprise more than 10% of total miner revenues.

Table 1 Summary Statistics for Bitcoin 2011-2016

Variable	Total	2016	2015	2014	2013	2012	2011
Est Transaction Volume (BTC)	3184607965	895508635	491390095	219888649	387003097	898729132	292088357
Monthly Min Transaction Volume	-	36838116	21347546	11972307	16123300	9241879	3420186
Monthly Max Transaction Volume	-	218196073	102435247	25374095	50063614	335046349	72694347
Mining Fees	11089683	1051363	1366225	1472451	1601099	2619012	2979533
Transaction Fees	60318	22338	8200	4626	15274	6797	3083
Median Bitcoin Price (USD)	-	582.19	248.07	501.63	112.23	6.78	3.27

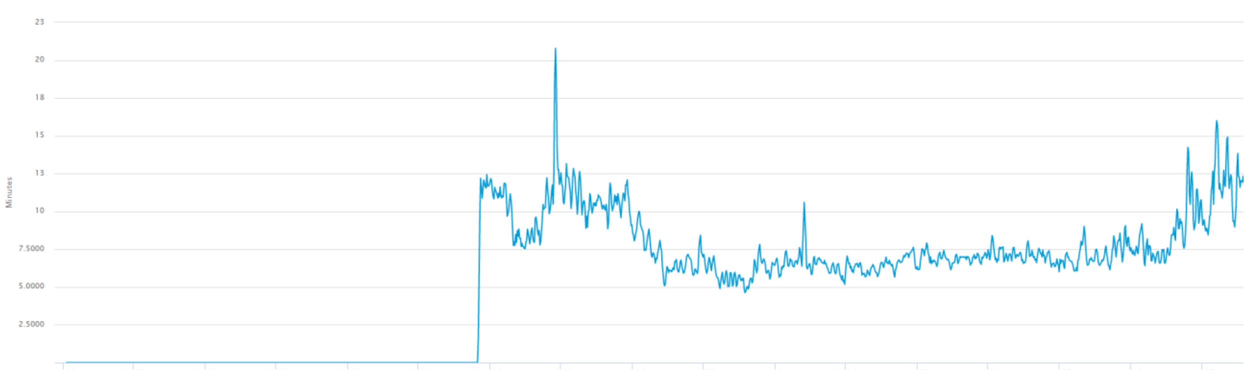
Source: Bitcoin blockchain node data

B. Transactions Fees and Equilibria

Our model shows that the arrival rate of transactions to the mempool exogenously drives growth in the mempool size, waiting times, and fees. We model this arrival rate as exogenous, but growing, up to the point at which users are discouraged from using Bitcoin by the long waiting time or fee necessary to get their transaction posted to the blockchain. Figure 6 depicted the model's prediction that as the waiting time for transactions to be confirmed increases, the equilibria should switch from one with zero fees to one in which every transaction pays a fee, with perhaps an intermediate stage in which some but not all transactions pay a fee. We test this prediction by examining how confirmation waiting times relate to the fraction of transactions paying fees.

We first look at how queuing times, defined by median confirmation times for transaction paying fees to be added to the blockchain, behave over time. Figure 7 Panel A gives this information for the available blockchain history, while Panel B gives data for the past year. As is apparent from Panel A, data on queuing times is not available before late 2011. As the blockchain has developed, median confirmation times have varied, particularly in the early years, but have generally been increasing since late 2012. Our model predicts that over time this should result in a transition from few, if any, users paying fees in late 2012 to most, if not all, users paying fees recently.

Figure 7. Median confirmation times for transactions paying fees.



A. Median confirmation times for transactions paying transaction fees 2009 - 2017

Figure 7 (cont.). Median confirmation times for transactions paying fees.

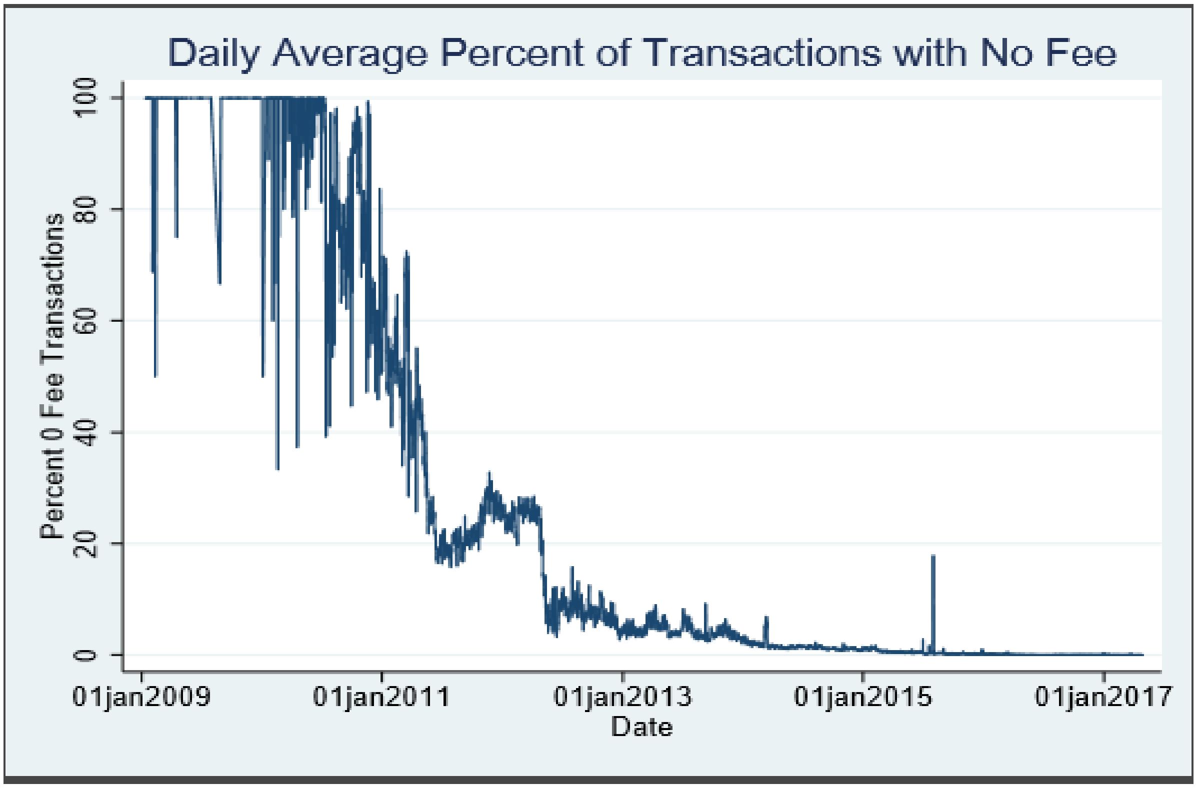
B. Median confirmation times for transactions paying transaction fees 2017

This figure shows the daily median confirmation time for transactions paying fees that are posted to the blockchain. Data on median confirmation times is not available prior to 2011. Panel A gives median confirmation times for the entire blockchain history; Panel B gives median confirmation times for 2017. Source: Blockhain.info

Figure 8 shows that the fraction of fee-paying transactions posted to the blockchain follows the pattern suggested by our theory. Prior to approximately the end of 2011, very few transactions paid fees, consistent with our no-fee equilibrium. This changed over 2011 and 2012, where transactions with attached fees moved from being a relatively small fraction to predominating trading. Over this interval, an equilibrium with both fee-paying and non-fee-paying transactions prevailed. By late 2014, this changed again, with an only-fee-paying equilibrium emerging.[19] Recall that our model showed that once such a fee-paying equilibria emerged, it was generally stable. Thus, even though bitcoin volumes (and waiting times) have fluctuated, the equilibrium in which everyone pays a fee remained. As predicted by our model, the good old days of zero-fee transactions are now essentially obsolete.

[19] One point to note is that prior to 2011 most transactions did not offer a fee and there was little financial incentive for miners to actually fill a block with transactions (a successful miner earned the block reward even if he did not put any transactions on the blockchain). As result, there are empty blocks during this period. The percent of zero-fee transactions for these blocks is undefined so we treat these as missing observations. Other blocks had transactions in them, but these transactions did not pay fees so these other early blocks have 100% zero-fee transactions.

Figure 8---The percentage of transactions written to the blockchain without attached fees



This Figure gives the percentage of transactions posted to the blockchain without attached fees. Source: Data from our bitcoin node

Direct evidence on the effects of waiting times on the incidence of zero transactions fees can be obtained by regression analysis. While it is tempting to simply regress the percentage of zero fee transactions on median confirmation times, our model shows that median confirmation time is endogenous, leading to the natural concern that such a simple OLS regression is mis-specified.[20] To address this concern, we ran two-stage least squares regressions using instrumental variables to capture the influence of exogenous model variables on the endogenous median confirmation times.

The ideal instrument would be highly correlated with the endogenous variable but uncorrelated with the error term. In our setting, instrument choice is also complicated both by the limited number of variables available and by the restricted

[20] The hypothesis of exogeneity of median waiting time in the regressions reported in Table 1 Panel A is rejected by the Durbin test (score) chi2(1) = 40.0948 (p = 0.0000) and the Wu-Hausman test F(1,1305)= 41.2029 (p = 0.0000).

sample periods of some data series. Median confirmation time data are available from Dec. 2, 2011 – April 28, 2017 so we initially select instruments that allow for estimation over this period. Using this period we have 1,309 daily observations. This a subperiod of our entire data set so we report summary statistics for it (and for the other subperiod used later) in the Appendix.

We ran the following 2SLS regressions:

$$(MWT)t = \alpha + \gamma_1(BTC\ price_t) + \gamma_2(block\ reward_t) + \gamma_3(numtrans_t) + \varepsilon$$

$$(\%zero\text{-}fee\ transactions)_t = \alpha + \beta_1\ (EMWT)t + \beta_2(BTC\ price_t) + \beta_3(block\ reward_t) + \varepsilon$$

where MWT is the median waiting time variable and EMWT is its estimated value from the first stage regression, and the instruments bitcoin price (BTC price), block reward (number of newly issued bitcoins awarded to the winning miner), and daily number of transactions (numtrans) capture exogenous effects on median confirmation times. We chose daily number of transactions as an instrument because more transactions should result in larger mempools and ultimately longer confirmation times, but there is no obvious direct effect of the number of transactions on the percent of transactions not paying a fee. Bitcoin price and block reward were included to allow for exogenous effects on bitcoin trading. Table 1 Panel A gives these estimation results.

The first stage results show that higher bitcoin price and greater numbers of transactions increase median waiting times, results that are statistically significant and in the expected direction. Most importantly, a one standard deviation increase in the number of transactions increases the median waiting time by 1.41 minutes, which is substantial relative to the sample mean of 9.42 minutes for median waiting time. Block reward also has a positive and significant effect. Specifications tests show that the instruments chosen have high correlation with MWT, and the F-test clearly rejects that the variables add no explanatory power.

The second stage results show that estimated median waiting times have a negative and significant effect on the percentage of zero-fee transactions. Thus, as median waiting times increase, the percentage of transactions posted to the blockchain with no transaction fee decreases. A one standard deviation increase in median waiting time increases the percentage of zero-fee transactions by 1.94 percent; a large effect relative to the sample mean of 3.13 percent. This confirms the prediction of our model that as median waiting times increase, the equilibrium shifts away from zero-fee transactions.

TABLE 1---2SLS Regression of Percent Zero Fee Transactions on Mean Waiting Time

Panel A: Sample period Dec.2, 2011 – April 28, 2017

First Stage Regression: F(3,1306)=143.34, Adj R-squared = 0.246

MWT	Coef.	Std Error	t	p>|t|
BTC price	.003653	.00040	9.11	0.000
block reward	.213656	.01111	19.22	0.000
numtrans	.000015	1.4e-06	10.99	0.000
constant	.366134	.44679	0.82	0.413

Second Stage Regression: Wald Chi2(3)=1379.75, R-squared = 0.446

% zero-fee	Coef.	Std Error	z	p>|z|
EMWT	-.60886	.14978	-4.06	0.000
BTC price	.00494	.00104	4.74	0.000
block reward	.55262	.02992	18.47q	0.000
constant	-7.353	.76382	-9.63	0.000

Panel B: Sample period April 24, 2016 – April 28, 2017

First Stage Regression: F(3,362)=115.68, Adj R-squared = 0.485

MWT	Coef.	Std Error	t	p>|t|
BTC price	.005038	.00082	6.17	0.00
block reward	-.015075	.031852	-0.47	0.64
Avg mempool size	.000122	.000012	9.82	0.00
constant	5.89376	.909274	6.48	0.00

Second Stage Regression: Wald Chi2(3)=17.05, R-squared = 0.0022

% zero-fee	Coef.	Std Error	z	p>|z|
EMWT	-.002422	.000919	-2.64	0.008
BTC price	8.36e-06	.000011	0.78	0.433
block reward	-.000853	.000287	-2.97	0.003
constant	.053370	.008548	6.24	0.000

This table gives the results of 2-stage least squares regressions of percentage zero fee transactions on mean waiting time. In panel A, the number of transactions is used as an instrument for mean waiting times. In panel B, the average mempool size is used as an instrument for mean waiting times.

In the regressions reported in Panel A of Table 1, we chose the daily number of transactions as an instrument because more transactions should result in larger mempools and ultimately longer confirmation times. A more direct instrument is average daily mempool size which should affect median waiting times and which does not have an obvious direct effect on the percent of zero-fee transactions. However, that variable is only available on a daily basis for the period April 24, 2016 – April 28, 2017. Using this more limited sample period with 365 observations, we ran the regressions:

$$(MWT)t = \alpha + \gamma_1(BTC\ price_t) + \gamma_2(block\ reward_t) + \gamma_3(average\ mempool\ size_t) + \varepsilon$$

$$(\%zero\text{-}fee\ transactions)t = \alpha + \beta_1\ (EMWT)t + \beta_2(BTC\ price_t) + \beta_3(block\ reward_t) + \varepsilon$$

where MWT is the median waiting time variable and EMWT is its estimated value from the first stage regression, and the instruments BTC price, block reward, and average mempool size capture the exogenous effects on median confirmation times.

Table 1 Panel B provides these results. The first stage regression shows that both bitcoin price and average mempool size have positive and significant effects on median confirmation time, but now block reward does not have a significant effect. A one standard deviation increase in the average mempool size increases median waiting time by 1.74 minutes. The R^2 is higher in this specification than with number of transactions, specifications tests again show that the instruments chosen have high correlation with MWT, and the F-test clearly rejects that the variables add no explanatory power.

Turning to our main regression, we again find that estimated median waiting time has a negative and significant effect on the percentage of zero fee transactions. A one standard deviation increase in median waiting time increases the percent of zero-fee transactions by 0.009. This number, while small in absolute value, equals 43% of the mean percent of zero-fee transactions in this sample of 365 observations. These results are consistent with the model's demonstration that the presence of transaction fees in equilibrium is influenced by user queueing effects.

As a robustness check on our analysis, we also ran specifications including an additional instrumental variable. In particular, in our first set of regressions, we added the lagged value of the number of transactions to our regressions:

$$(MWT)t = \alpha + \gamma_1(BTC\ price_t) + \gamma_2(block\ reward_t) + \gamma_3(numtrans_t) + \gamma_4(numtrans_{t-1}) + \varepsilon$$

$$(\%zero\text{-}fee\ transactions)_t = \alpha + \beta_1\ (EMWT)t + \beta_2(BTC\ price_t) + \beta_3(block\ reward_t) + \varepsilon$$

In the second set of regressions, we added lagged average mempool size to the regressions:

$$(MWT)t = \alpha + \gamma_1(BTC\ price_t) + \gamma_2(block\ reward_t) + \gamma_3(average\ mempool\ size_t) + \gamma_4(average\ mempool\ size_{t-1}) + \varepsilon$$

$$(\%zero\text{-}fee\ transactions)t = \alpha + \beta_1\ (EMWT)t + \beta_2(BTC\ price_t) + \beta_3(block\ reward_t) + \varepsilon$$

The estimation results in both specifications are essentially unchanged from our previous specifications and so for sake of brevity we do not report the results. The advantage of having additional instrumental variables, however, is that it allows us to test for over-identifying restrictions in the model using a Sargan test. Here the results clearly show that we cannot reject the null hypothesis that the over-identifying restrictions are valid. Using the daily and lagged number of transactions as instruments, the Sargan test statistic is Chi(1) = 1.039, p-value = 0.308. Using the daily and lagged average mempool size as instruments, the test statistic is Chi(1) = .2355, p-value = .6275. These results confirm that our instruments are both relevant and valid.

C. Queueing, Mining, and Transaction Fees

Our model shows that transaction fees can potentially solve two problems: they can incentivize miners to participate by offsetting declining mining revenues; and they can solve a queueing problem for users. Consequently, both mempool waiting times and the block reward level could influence fee levels.

Figure 9 Panel A shows that transaction fees per byte were initially 0 in 2011, showed substantial volatility in the next two years and in more recent times are increasing again. This recent behavior, illustrated in Panel B, shows transaction fees per byte over the past year. This past year also saw a decrease in the block reward (see Figure 2) as well as increases in the median daily confirmation time (see Figure 7). Thus, an interesting question is how do these variables influence average transaction fee levels?

The block reward is an exogenous variable, but the median confirmation times are endogenous. From our previous analysis, we address this endogeneity issue with simultaneous equations estimation. As before, Median confirmation time data is available from Dec. 2, 2011 – April 28, 2017 (1,309 daily observations) so we initially select instruments for this variable that allow for estimation over this period. We ran the following 2SLS regressions:

$$(MWT)t = \alpha + \gamma_1(BTC\ price_t) + \gamma_2(block\ award_t) + \gamma_3(numtrans_t) + \varepsilon$$

$$(txfeemean)_t = \alpha + \beta_1\ (EMWT)t + \beta_2(BTC\ price_t) + \beta_3(block\ award_t) + \varepsilon$$

where *txfeemean* is average daily transaction fee per byte, MWT is the median waiting time variable and EMWT is its estimated value from the first stage regression, and block reward is the reward earned by the winning miner. The instruments BTC price, block reward, and daily number of transactions capture the exogenous effects on median confirmation times. The results are given in Table 2 – Panel A.[21]

As an alternative specification, we used the instrument average mempool size which is only available on a daily basis for the period April 24, 2016 – April 28, 2017. Using this more limited sample period, we ran the two-stage least squares regression:

$$(MWT)t = \alpha + \gamma_1(BTC\ price_t) + \gamma_2(block\ reward_t) + \gamma_3(average\ mempool\ size_t) + \varepsilon$$

$$(txfeemean)t = \alpha + \beta_1\ (MWT)t + \beta_2(BTC\ price_t) + \beta_3(block\ reward_t\ ) + \varepsilon$$

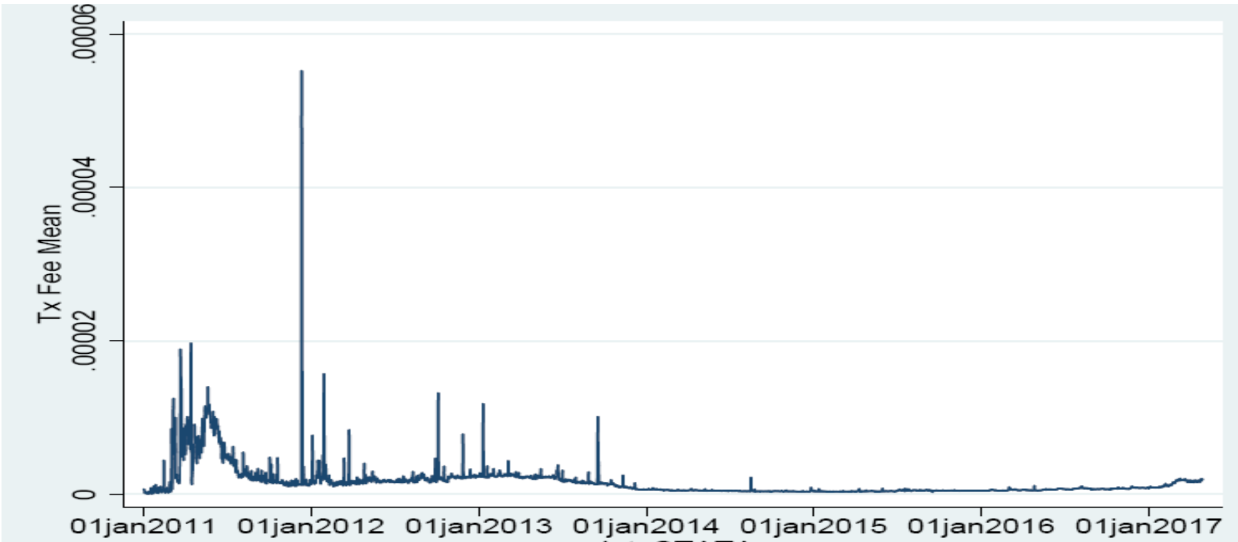
where the variables are as previously defined. The instruments BTC price, block reward, and average mempool size capture the exogenous effects on median confirmation times. Table 2 – Panel B provides these results.

We also ran specifications including an additional instrumental variable. In our first set of regressions, we added the lagged value of the number of transactions to our regressions, while in the second set of regressions we added the lagged

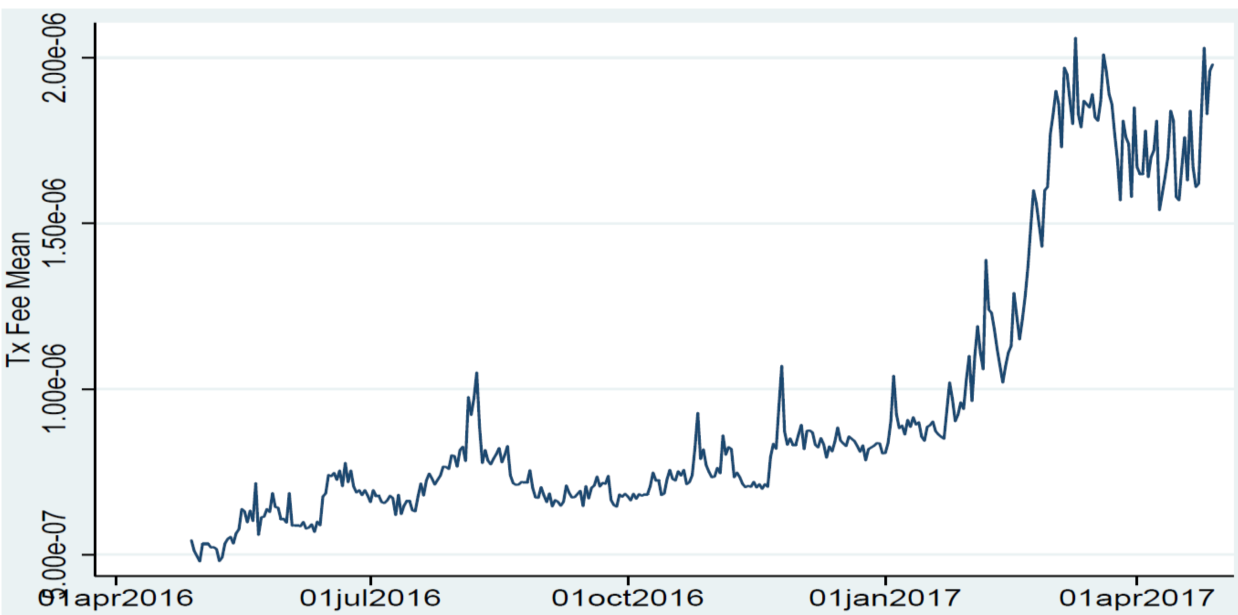
[21] The first stage of this regression is, of course, the same as the first stage reported in Table 1. It is included here only for ease of comparison.

Figure 9

Panel A. Average Daily Transaction Fee 2011-2017



Panel B. Average Daily Transaction Fee April 2016- April 2017



This figure shows the average daily transaction fee for transactions added to the bitcoin blockchain. Source: bitcoin blockchain node data

TABLE 2---2SLS Regression of Mean Daily Transaction Fee on Mean Waiting Time

Panel A

First Stage Regression: F(3,1306)=143.34, Adj R-squared = 0.246

MWT	Coef.	Std Error	t	p>|t|
BTC price	.003653	.00040	9.11	0.000
block reward	.213656	.01111	19.22	0.000
numtrans	.000015	1.4e-06	10.99	0.000
constant	.366134	.44679	0.82	0.413

Second Stage Regression: Wald Chi2(3)=28.21, R-squared = 0.013

Txfeemean	Coef.	Std Error	z	p>|z|
EMWT	-5.08e-08	5.76e-08	-1.01	0.314
BTC price	1.87e-10	4.00e-10	0.47	0.640
block reward	4.32e-08	1.15e-08	3.76	0.000
constant	4.53e-07	2.94e-07	1.54	0.123

Panel B

First Stage Regression: F(3,362)=115.68, Adj R-squared = 0.485

MWT	Coef.	Std Error	t	p>|t|
BTC price	.005038	.00082	6.17	0.000
block reward	-.015075	.031852	-0.47	0.636
Avg mempool size	.000122	.000012	9.82	0.000
constant	5.89376	.909274	6.48	0.000

Second Stage Regression: Wald Chi2(3)=11.35, R-squared = 0.7708

Txfeemean	Coef.	Std Error	z	p>|z|
EMWT	2.31e-08	7.39e-09	3.12	0.002
BTC price	1.36e-09	8.58e-11	15.82	0.000
block reward	2.89e-09	2.31e-09	1.25	0.211
constant	-4.04e-07	6.88e-08	-5.88	0.000

This table gives the results of 2-stage least squares regressions of mean daily transaction fees on mean waiting time. In panel A, the number of transactions is used as an instrument for mean waiting times. In panel B, the average mempool size is used as an instrument for mean waiting times.

mempool size. Using the daily and lagged number of transactions as instruments, the Sargan test statistic is Chi(1) = .3350, p-value = 0.563. Using the daily and average mempool size as instruments, the test statistic Chi (1)=.23pp, p-value = .6275. The results clearly show that you cannot reject the hypothesis that the over-identifying restrictions are valid. As before, the results of these regressions are essentially the same and so are not reported.

Turning to the regression results in Table 2, the results of the first set of regressions in Panel A show that only the block reward has a significant effect, but its coefficient is now positive, and not negative as might have been expected. Estimated median confirmation times do not have a significant effect on the daily average transaction fee. Given that median confirmation times did have a significant and negative effects on the fraction of fee- paying transactions, these results may suggest that over the entire sample period the mean changes in fees are not well captured by either waiting time or block reward effects. Certainly, the volatility of fees in the 2012-3013 period suggest that a variety of factors may be at play.

Turning to the second set of regressions in Panel B, we reach a different conclusion. Here we do find estimated median confirmation times play the role suggested by our model: the coefficient is positive and significant, consistent with higher daily transaction fees being associated with longer waiting times. We find that a one standard deviation increase in median waiting time increases the mean transaction fee by $7.5e^{-08}$ which is 8% of the sample average mean transaction fee. Moreover, block reward is not significant, suggesting that the behavior of transaction fees over this more recent time-frame is not driven by changes in the block reward.   Overall, these results suggest that transaction fee levels are significantly affected by queuing problems confronting users.

4. Conclusions and Policy Implications

Nakamoto conjectured that over time the bitcoin blockchain would have very large volumes or none at all.[22] We have shown in this paper that his reasoning was mostly correct; in the absence of transaction fees, over time the blockchain is not viable. But even with transactions fees, there is an upper bound on the size of the blockchain imposed by the waiting time confronting users in the mempool. As this

[22] Cited in https://www.bitcoinmining.com/what-is-the-bitcoin-block-reward/

waiting time becomes large, users exit the blockchain in much the way that miners exit the blockchain when their revenues no longer generate profits. Thus, the equilibrium in the bitcoin blockchain is a complex balancing of user and miner participation. We have shown that transaction fees play a crucial role in affecting both of these clienteles, and thus in influencing the stability of the blockchain.

As the bitcoin ecology migrates to a more market-based system, a variety of interesting issues become apparent. One such issue is the role played by microstructure features such as exogenous structural constraints. While constraints limiting the growth of new bitcoin issuance are in line with the system's original design, the constraints on block size are a relatively recent addition intended to decrease the system's vulnerability to attack. A perhaps unintended consequence is that this constraint exacerbates the imbalances between mempool inflow and outflow, potentially leading to instability in the blockchain. The recent dramatic increase in bitcoin transaction fees are viewed by some as evidence of just such effects.[23]

Such instability recently led to debate in the bitcoin community over whether to change the bitcoin block size.[24] Proposals to increase the block size from 1 MB to as much as 20 MB were floated, but little consensus emerged, in large part because different clienteles have different needs of, uses for, and even philosophies regarding bitcoin and the blockchain.[25] Interestingly, Caffyn (2017) notes "[some] maintain that limiting block size in the short-term will create a self-regulating market for transaction fees" and that "will increase miners' incentive to process transactions which will benefit the health of the system." Here our results suggest otherwise. Increasing transaction fees will increase the number of miners, but this in turn will trigger increases in the difficulty level to control the creation rate of new blocks, thereby raising the costs to miners. In a competitive market, this will

---

[23] Fees have increased steadily since 2016, but they reached historic levels in September 2017. For example, fees on Sept.6, 2016 averaged $.165 but were $4.506 on September 6, 2017.

[24] These debates involved a variety of technical issues in addition to block size. One solution, SegWit2, appeared to have sufficient backing, but ultimately was not able to prevail. For discussion see https://www.coindesk.com/explainer-what-is-segwit2x-and-what-does-it-mean-for-bitcoin/

[25] The issue is complicated both by technological issues and by governance issues. As Caffyn [2017] notes, Chinese miners, who account for more than half the computing (or hashing) power of the network, argued against larger sizes due to concerns about the bandwidth. Other parties, such as exchanges and wallet companies, have different views on the optimal size. Still others call for dynamic increases in the block size over time tied to the growth of volume. Because governance issues are settled by computing power "votes", disagreement can lead to potential "forks" in the block chain

not lead to an overall increase in compensation. Moreover, as we have shown here, transaction fees also reflect queueing problems facing users. As median confirmation times rise, users may choose to forego transacting in Bitcoin, perhaps opting instead for one of many other crypto-currencies.

These structure issues highlight the increasingly important role that microstructure will play in the evolution of bitcoin from mining to markets. Designing markets to operate efficiently is always challenging, and it is becoming particularly so in Bitcoin as divergent clienteles emerge. The recent "forking" of the blockchain caused by the introduction of Bitcoin Cash (BCC) reflects exactly such issues, as some users opt to change the rules to make the blockchain better meet their needs.[26] Whether the bitcoin blockchain remains a single entity or fragments into a collection of bitcoin-linked blockchains may well depend upon how these market microstructure issues are resolved. Certainly, these are important issues for future research.

[26] Bitcoin cash began on August 1, 2017 with the production of a 1.9MB block that was not valid on the bitcoin network. For more discussion, see https://www.bitcoincash.org/

Appendix. Data Used in 2SLS Regressions

Panel A: Sample Period Dec 2, 2011-April 28, 2017

Variables	mean	Standard deviation	Min	Max
% zero-fee	3.132257	6.117838	0	28.98364
MWT	9.422252	3.184706	4.766667	47.73333
numtrans	134959.6	91874.28	4738	350368
BTC price	410.3659	318.9638	2.79	1329.19
Block reward	25.68755	10.98438	12.5	50
Txfeemean	1.09e-06	1.76e-06	2.89e-07	.0000552

Panel B: Sample Period April 24, 2016-April 28, 2017

Variables	mean	Standard deviation	Min	Max
% zero-fee	.0206269	.0244841	0	.3642179
MWT	10.90779	3.769638	6	29.25
Avg mempool	10825.19	13950.13	693.25	92429.7
BTC price	775.5763	235.0773	436.73	1329.19
Block reward	15.06849	5.057577	12.5	25
Txfeemean	9.42e-07	4.07e-07	4.80e-07	2.06e-06

References

Aune, R., A. Krellenstein, M. O'Hara, and O. Slama, (2017), "Footprints on a Blockchain: Trading and Information Leakage in Distributed Ledgers," Journal of Trading, *12(2),* Summer 2017.

Bohme, R., N. Christin, B. Edelman, and T. Moore, (2015), "Bitcoin: Economics, technology, and governance," The Journal of Economic Perspectives 29, 213-238.

Caffyn, G., (2017) "What is the Bitcoin Block Size Debate and Why Does it Matter?" available at http://www.coindesk.com/what-is-the-bitcoin-block-size-debate-and-why-does-it-matter/

Cong, L. W. and H., Zhiguo, 2017, "Blockchain Disruption and Smart Contracts" available at SSRN: https://ssrn.com/abstract=2985764

Houy, N., (2014), "The bitcoin mining game," available at http://ssrn.com/abstract=2407834.

Ittay Eyal and E.G. Sirer, (2014) "Majority is not enough: Bitcoin mining is vulnerable", in International Conference on Financial Cryptography and Data Security pp. 436-454,Springer.

FINRA, 2017, "Distributed Ledger Technology: Implications of Blockchain for the Securities Industry," http://finra.org/sites/defaultfiles/FINRA_Blockchain_ Report.pdf

Gans, J. S. and H. Halaburda (2015), "Some economics of private digital currency," in Economic Analysis of the Digital Economy (University of Chicago Press.

Gandel, N. and H. Halaburda, 2016, "Can we predict the winner in a market with network effects? Competition in cryptocurrency market", Games 7(3).

Harvey, C., (2016), "Cryptofinance", SSRN http://ssrn.com/abstract=2438299

Hileman, G. and M. Rausch, (2017), "Global Cryptocurrency Benchmarking Study," Cambridge Center for Alternative Finance, Cambridge University.

Huberman, G., Leshno, J. D. and C. Moallemi (2017), “Monopoly with a Monopolist: An Economic Analysis of the Bitcoin Payment System”, Working Paper, Columbia Business School.

Kroll, J. A., Davey, I., and E. Felton (2013), “The economics of bitcoin mining, or bitcoin in the presence of adversaries,” in Proceedings of WEIS, Vol. 2013.

Little, J.D.C. (1961), “A Proof for the Queuing Formula: L=λW”, Operations Research 9,3, 383-387.

Malinova, K. and A. Park (2016), “Market Design for Trading Blockchain Technology, SSRN.

Nakomoto, S., (2008), “A Peer-to-Peer Electronic Cash System”, White Paper, https://bitcoin.org/bitcoin.

Raskin and Yermack (2016), “Digital Currencies, Decentralized Ledgers and the Future of Central Banking, Discussion Paper, National Bureau of Economic Research.