Transaction fee economics in the Ethereum blockchain

Article *in* Economic Inquiry · September 2021

DOI: 10.1111/ecin.13025

2 authors:

Anil Donmez
Simon Fraser University

Alexander K. Karaivanov
Simon Fraser University

Transaction Fee Economics in the Ethereum Blockchain*

Anil Donmez[†] and Alexander Karaivanov[‡]

July 7, 2021

Abstract

We study the economic determinants of transaction fees in the Ethereum blockchain. We estimate an empirical model based on queueing theory and analyze the factors determining the "gas price" (transaction cost per unit of service, "gas"). Using block- and transaction-level data from the Ethereum blockchain, we show that changes in service demand significantly affect the gas price - when there is high block utilization, per-unit fees increase on average, with strong non-linear effect above 90% utilization. The transaction type is another important factor - larger fraction of regular transactions (direct transfers between users) is associated with higher gas price.

JEL codes: G19, G23, C13

Keywords: blockchain, transaction fees, Ethereum, time series, queueing theory

*We thank C. Bidner, L. Petersen and CEA and seminar participants for useful comments and suggestions. Karaivanov acknowledges funding from SSHRC Insight Grant A-435-2018-0111.

†PhD Student, Dept. of Economics, Simon Fraser University, 8888 University Dr., Burnaby BC, V5A 1S6, Canada

‡Corresponding author. Professor, Dept. of Economics, Simon Fraser University, 8888 University Dr., Burnaby, BC, V5A 1S6, Canada. Email: akaraiva@sfu.ca

1 Introduction

We study the economic determinants of transaction fees in the Ethereum blockchain, the second largest blockchain by market capitalization.[1] While Bitcoin, launched in 2009, is the oldest, best-known and most widely researched blockchain platform, Ethereum, started in 2015, has a different design, with broader applicability going beyond digital payments. Similar to other blockchains, Ethereum has an internal digital (crypto) currency with decentralized and scarce supply, called Ether (ETH), that can be used as store of value or transacted between users anywhere in the world. However, unlike other blockchain platforms, Ethereum is much more flexible and programmable, hence many developers have used the platform to create a wide range of decentralized applications ("DApps") and "smart contracts".[2]

A distinctive feature of Ethereum is its internal metering variable called *gas*. Each blockchain transaction has an algorithmically defined *gas requirement* – a pre-specified execution cost expressed in units of gas. The more complex a transaction or a smart contract is, the higher is its gas requirement, which must be paid by the user for the transaction to be recorded and executed on the blockchain. The simplest and most common transaction type, which we call "regular transaction", is a transfer of ETH between two blockchain addresses. Such transaction requires 21,000 gas to be executed, for any transferred ETH amount. Smart contract creations and smart contract calls have substantially higher gas requirements (10 to 20 times higher, see Figure 4).

The key component of transaction costs in Ethereum, and the main focus of our analysis, is the *gas price*. The gas price is a bid price (in ETH) per unit of gas, specified by the user when posting a blockchain transaction (transfer or contract creation/call). Together, the gas requirement and the gas price determine the total transaction cost which equals (gas used)$\times$(gas price). For example, an ETH transfer between two addresses requires exactly 21,000 gas and hence its transaction fee is 21,000$\times$(gas price). This implies that both the transaction complexity/type (its gas requirement)

[1] On December 10, 2020 the market capitalization of Ethereum was $64 bln, second only to Bitcoin with capitalization $341 bln. Source: http://coinmarketcap.com

[2] At the heart of Ethereum's extensive functionality is the "Ethereum Virtual Machine" (EVM), a Turing-complete virtual computation engine, Buterin (2013). Thanks to the EVM design, Ethereum users are not only able to execute simple transactions such as sending digital currency from one address to another, but can also create and make "calls" (e.g., supply data or make automated trade requests) to "smart contracts". A smart contract is computer code that can be used to digitally facilitate, verify, or enforce the negotiation or execution of trades (e.g, see Cong and He, 2019). Examples include "DeFi" (decentralized finance) applications that allow users to borrow, lend and invest digital assets; exchanges for trading digital currencies; Ethereum-based (ERC-20) digital tokens; games; gambling platforms; voting platforms; and supply chain management systems.

and the users' willingness to pay per unit of gas (the gas price) determine the cost of posting a transaction. The transaction type choice, including contract creation or calls, and its interplay with the gas price adds a dimension to our analysis which is mostly absent in the Bitcoin blockchain.

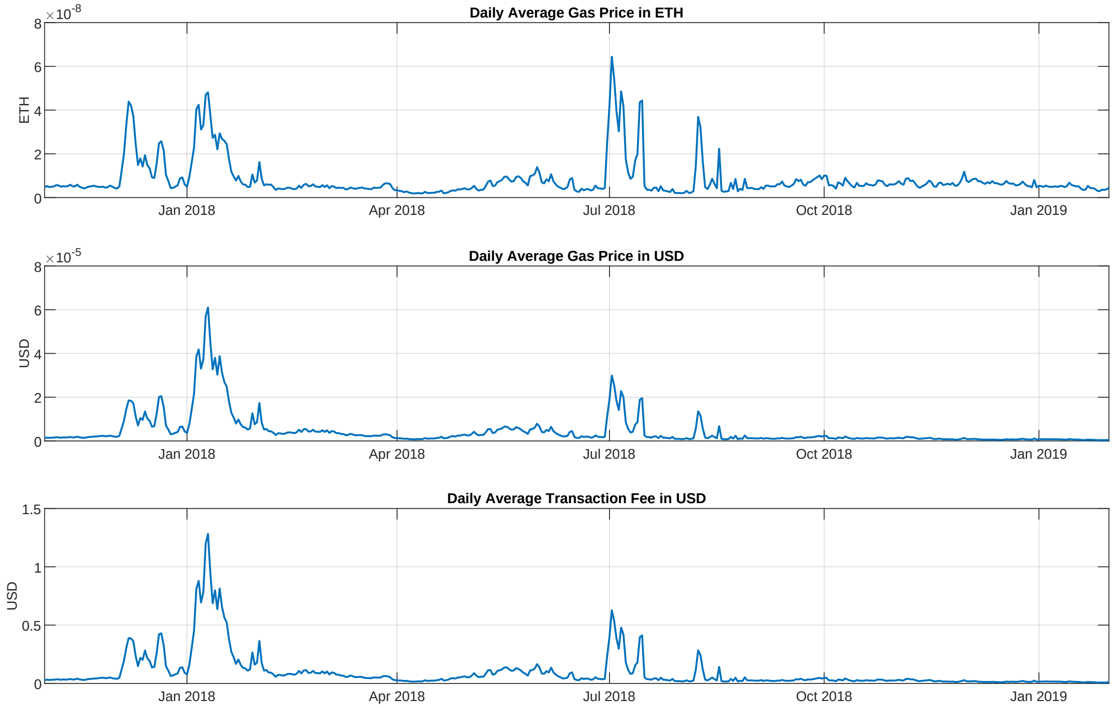
We characterize the economic determinants and dynamics of transaction fees in the Ethereum blockchain by estimating an empirical time-series model based on queueing theory. In contrast to opinions that blockchain activity is highly speculative and volatile,[3] we find that, by and large, gas price levels and dynamics comply with standard economic predictions from queueing theory and supply/demand theory. First, both the marginal Ethereum gas price (the minimum gas price bid within a block) and the median block gas price are higher on average when the blockchain experiences higher utilization or congestion, consistent with the theory. Second, we find that the effect of blockchain utilization on the gas price is strongly non-linear. When the block occupancy is below a threshold (90 percent), the effect of higher utilization on the marginal and median gas price is insignificant. However, beyond that threshold, the effect of blockchain utilization on the gas price is positive and increasing in a convex way. Third, higher marginal and median gas prices are positively and statistically significantly associated with a higher fraction of regular transactions per block and negatively associated with the fraction of contract creations and contract calls.

On Figure 1 we display the daily average value of our main variable of interest – the marginal gas price in ETH or USD, defined as the minimum gas price in each block, over the period of study November 1, 2017 to January 31, 2019. We also show the implied marginal transaction fee for a regular transaction (ETH transfer). The minimum gas price in a block is the bid price per unit of gas at which the last (marginal) transaction was recorded and corresponds to the gas price at which a transaction is just on the margin of being included vs. not being included in the current block (see Section 2.2 for more discussion). We see that the gas price exhibits significant variation over time, allowing us to analyze and identify its determinants.[4] In addition, to manage spikes in the gas price, the Ethereum protocol can increase the service rate by raising the block gas limit to

[3]For example, on April 14, 2021 US Federal Reserve Chairman Jerome Powell said about cryptocurrencies "They're really vehicles for speculation" (CNBC). See Foley et al. (2019), Gandal et al. (2018), Griffin and Shams (2020), Li, Shin and Wang (2020) or Yermack (2014) for formal discussion and analysis.

[4]Although the median value of the marginal gas price is \$0.000018, implying a \$0.04 median fee for a regular transaction (see Table A1), there are times, e.g., in January 2018 or July 2018, when the gas price was substantially higher, implying \$1 or larger fee for a regular transaction. Relatively low transaction fees which are independent of the transfer amount are an important advantage of cryptocurrency platforms compared to conventional payment services providers. For example, Paypal, the largest online payment platform, charges \$5 or more for international transfers. (`https://www.paypal.com/us/webapps/mpp/paypal-fees`).

Figure 1: Gas Price



Notes: The top and middle panels of Figure 1 plot the daily average of the minimum gas price in each block ('marginal gas price') over time, measured in ETH or USD respectively. The bottom panel plots the implied minimum transaction fee in USD for a 'regular transaction' (ETH transfer between two accounts), computed as 21,000×(the daily average USD gas price).

accommodate a larger number of transactions.[5] We control for these algorithmic supply changes in our empirical analysis.

Each Ethereum block has a block gas limit which can be thought of as the block's 'capacity' and which, together with the number of blocks created in a day, determines the gas supply.[6] The gas supply is kept stable by the blockchain algorithm automatically adjusting the cryptographic difficulty to ensure that block creation is spaced out evenly in time. Supply fluctuations are thus small and can be considered exogenous, arising mostly because of the randomness in the time to complete the cryptographic proof-of-work.[7] We do control for the few system-wide gas supply changes implemented via the blockchain code in the sample period (see Appendix B for more details). Given the limited and stable gas supply, the gas prices bid by the users therefore determine

[5]One such increase occurred on December 10, 2017, when the block gas limit was raised by 18%, see Figure 6.

[6]The block capacity (size) in Ethereum is measured in gas (not bytes), unlike Bitcoin which uses a fixed 2MB block size.

[7]The cryptographic problem is very costly to solve but its solution is easy to verify. Hence, the Proof-of-Work mechanism, together with the rest of the blockchain software code, ensures the security, consensus, and stability of the network.

the priority order of transactions included in each block. Too low gas price may cause a transaction's inclusion to be delayed, possibly indefinitely.

When posting any transaction on the Ethereum blockchain, in addition to the gas price, the user must also specify a transaction *gas limit*, that is, the maximum gas the transaction can use up. The transaction would be executed as long as its gas requirement does not exceed the transaction gas limit.[8] In addition, the sum of the gas requirements of all transactions included in a block cannot exceed the block gas limit. The need for a costly gas requirement and gas limit stem from the virtual machine basis of Ethereum and helps avoid problems such as infinite loops, coding errors, sabotaging the network, etc.[9]

The transaction fee (gas price) bid and paid by the users is the key mechanism providing economic incentives to operate the blockchain network and to verify and record transactions (a block creation reward also exists but is being reduced over time).[10] In Ethereum the blockchain transactions are verified and blocks are created by computing pools ("miners"), through a competitive cryptographic problem-solving consensus mechanism called Proof-of-Work. In order to maintain their operations, there must be sufficient financial incentive for the miners, covering their time, equipment and electricity costs.

Related literature

This paper relates to complementary theoretical and empirical research on cryptocurrencies in economics and finance. Huberman et al. (2019) show that Bitcoin's decentralized design protects users from monopoly pricing, derive closed-form expressions for fees and waiting times and compare Bitcoin payments to a traditional payment system. Consistent with our findings for Ethereum, they show that Bitcoin transaction fees increase in non-linear way with congestion, however, we find, in

[8]For example, a user would normally set the gas limit of a regular transaction to 21,000 gas, which would satisfy exactly the gas requirement for such transaction. In contrast, the gas limit for a smart contract creation or call must be set higher, to ensure that the limit covers the gas requirements of the resulting EVM instructions, depending on the contract's complexity.

[9]Ethereum is Turing complete – in theory a program of any complexity can be computed by the Ethereum Virtual Machine (EVM), see Buterin (2013). However, this flexibility can introduce security, stability and resource management problems. For example, Turing-complete systems can be set to run infinite loops. If executing arbitrary code did not face a resource constraint and if a transaction or smart contract causes such a loop, either by mistake or deliberately, this could destabilize or disable the Ethereum network. To prevent such problems, each allowed EVM instruction has a pre-defined cost in units of gas. The execution of any transaction or smart contract is automatically terminated if the gas consumed by running it exceeds the available gas (the gas limit) for the transaction. This 'resource constraint' feature of Ethereum therefore maintains Turing completeness while capping and putting a price on the system resources that any blockchain transaction can consume.

[10]See Catalini and Gans (2019) for further discussion and analysis of the verification and networking costs in blockchains.

addition, that user choice over the transaction type is another key factor.

Easley et al. (2019) develop and test empirically a game-theoretic model of Bitcoin transaction fees and the strategic behavior of miners and users. They find that transaction fees are positively correlated with the average waiting time.[11] Since detailed waiting time data for the Ethereum platform is unavailable, we construct a variable measuring blockchain utilization directly from the blockchain records, which to our knowledge is new in the literature. Möser and Böhme (2015) study 45.7 mln Bitcoin blockchain transaction records and find that changes in the system protocol or the actions of big intermediaries can trigger regime shifts in the transaction fee level. Chiu and Koeppl (2019) examine theoretically the relationship between Bitcoin transaction fees, block size and user characteristics. Their simulation results suggest that users are willing to pay more when their transactions are more urgent and when the block size is smaller.

Our work complements these papers but differs in two important ways. First, we use detailed micro-level data (transaction-level and block-level) directly downloaded from the Ethereum blockchain. Second, unlike most authors who study Bitcoin, we instead analyze Ethereum's internal gas and gas price variables, which have clear economic interpretation as a scarce resource and its endogenous price.[12] We abstract from monetary policy issues related to the blockchain technology on which there exists a large separate literature.[13]

2 Data

2.1 Data sources

We use data obtained directly from the Ethereum blockchain.[14] By construction, the blockchain records a publicly accessible permanent copy of the complete transaction history since Ethereum's

[11] In other related work, Kasahara and Kawahara (2019) model transaction execution as a non-preemptive priority queueing game and show that users' waiting time is determined not only by their own posted fee, but also by the arrival rate of users with higher transaction fees. Li et al. (2018) model transaction confirmation as a single-server queue with batch service and priority mechanism and show that the average waiting time is affected by the share of users from each priority class.

[12] See also Zochowski (2019) for a non-technical review of the network characteristics and transaction fee dynamics in several blockchain platforms including Ethereum.

[13] There is also a finance literature on cryptocurrency prices and returns determination, predictability and volatility, e.g., Ciaian et al. (2015), Kristoufek (2015), Corbet et al. (2018), and Athey et al. (2016) among others.

[14] We used the publicly shared Python scripts by E. Medvedev, github.com/blockchain-etl/ethereum-etl to extract and save the Ethereum blockchain data. The raw data files were then merged, processed and analyzed in Matlab and Stata using code written by the authors and available on request.

launch in 2015. We downloaded both block-level and transaction-level data, including transaction receipts data. The block-level data contain information about each block included in the ETH blockchain. The data contain the block's sequential number, difficulty level[15], the block *gas limit* which determines the maximum capacity of the block, the *gas used* which is the total gas consumed by all transactions included in the block, and a timestamp. The main block variables that we use in the empirical analysis are the *gas limit* and *gas used*, which determine the blockchain network utilization at the time of block creation.

The transaction-level blockchain data include: the block number in which the transaction is recorded, the transaction value in Wei,[16] the sender address, the receiver address, a "nonce" value which indicates the number of prior transactions posted by the sender address, the gas price in Wei set by the sender, the maximum gas amount that the transaction can use (transaction gas limit), and the actual gas quantity used by the transaction (obtained from the transaction receipt data). We merge the block-level and transaction-level data using the block number. The end result is a complete transaction-level blockchain dataset in which block specific information is preserved. For part of the analysis we also combine the blockchain data with additional data on the ETH price in USD or other currencies, obtained from the website min-api.cryptocompare.com at the daily level.

We analyze Ethereum gas prices in the period between November 1^{th}, 2017 and January 31^{th}, 2019. This period does not include significant changes in the core Ethereum blockchain protocol and code. This enables us to focus on the interplay between the blockchain platform and user demand. The first three months of the sample period, which we call the "peak period" and also analyze separately, capture several historic peak events as of the end of 2020 - the highest daily transaction count, the highest ETH price in USD, and highest total market capitalization of ETH.[17]

Our sample period includes 2,688,667 blocks and 309,760,480 transactions in total. The raw data have slightly fluctuating time frequencies, since block creation times differ.[18] To mitigate hour-of-the-day and other high frequency periodical variation in the data, we aggregate the raw transaction data at the daily level. This generates 457 daily observations that we use in the empirical analysis.

[15]Each block has a computational difficulty level corresponding to the algorithmic cryptographic problem solved by miners to verify and confirm the block transactions.

[16]Wei is the smallest denomination in Ethereum and equals 10^{-18} ETH.

[17]The Ethereum blockchain is experiencing another peak period in early 2021.

[18]On average, an Ethereum block is created every 13 to 14 seconds.

2.2 Variables and definitions

Our main variable of interest is the daily *marginal gas price*, defined as the daily average of the minimum observed gas price in each of the blocks recorded in a given day. The minimum gas price in a block is the user bid price per unit of gas at which the last (marginal) transaction was recorded. Hence, the marginal gas price variable captures most closely the price (fee per unit of gas) at which a transaction is just on the margin of being included vs. not being included in the current block. In robustness analysis we also use two alternative gas price variables: the *median gas price* (daily average of the median gas price in each block recorded on that day) and the *lowest 5-th percentile* gas price (daily average of the bottom 5th percentile gas price in each block recorded on that day).

Second, we construct a *blockchain utilization* variable to capture the usage level, or congestion rate, of the platform. Blockchain utilization is defined as the ratio between the sum of the gas requirements of all recorded transactions in a given day and the *gas supply* in the same day. We define the gas supply as the sum of the gas limits of all blocks recorded in a given day. The blockchain utilization variable therefore measures the fraction of total available gas supply that is used up, per unit of time. For example, daily blockchain utilization of 0.8 means that on average 80% of block capacity is used in a given day.

Third, to account for the fact that there are different types of transactions in Ethereum, we define the variable *regular transactions share* as the ratio of the number of regular transactions (simple ETH transfers between two addresses) to the number of all transactions recorded in a given day. This variable captures the possible effect of changes in the composition of posted transactions (for example, more vs. less urgent) on the gas price. In Section 4.4 we show that regular transactions are more likely to have higher gas prices and to be recorded near the top of their blocks; this is consistent with interpreting regular transactions as more urgent on average.

Table A1 in the Appendix displays summary statistics of the variables defined above, and also the ETH price in USD or in terms of a basket of currencies.

2.3 Descriptive findings

Figures 2-6 illustrate the magnitudes and time dynamics of key Ethereum blockchain variables relevant for our empirical analysis, in our sample period. The "peak period" (Nov. 1, 2017 to Jan.

31, 2018) is shaded.

Figure 2 displays the daily average blockchain utilization (block occupancy rate). The Ethereum network operated close to its full capacity at times, particularly in the second half of the peak period and also around July and August 2018. For the remainder of the study period, the blockchain network utilization was well below 1, implying no significant congestion and available room for additional transactions in most blocks.

Figure 3 shows the daily regular transactions share. During the peak period in the beginning of the sample, the share of regular transactions rises steadily from about 0.45 to as high as 0.8. This suggests an increasing number of user transactions transferring ETH from one address to another, as opposed to smart contract creations or calls. In the remainder of the study period the regular transactions share goes back down and stays around 0.5.

Figure 4 displays the daily average gas used (gas requirement) by transaction type. Regular transactions (ETH transfer between two addresses) require and use 21,000 units of gas for execution, regardless of the transferred amount. In contrast, contract calls and contract creations have significantly higher gas requirement, depending on their complexity – in our study period, the daily mean (median) gas requirement for contract calls and creations is 344,000 (324,000) gas.

Figure 5 plots the daily gas supply, defined as the sum of the gas limits of all blocks recorded in a given day. The gas supply variable proxies the total service rate of the Ethereum blockchain, since block space is restricted by the block gas limit. Figure 5 shows that the gas supply is stable, with the exception of the beginning of the peak period (Dec. 2017) and the end of the sample (Jan. 2019). The Dec. 2017 increase was caused by a change in the Ethereum protocol which raised the block gas limit by approximately 18 percent as of December 10, 2017. The decrease in gas supply at the end of the sample was triggered by a system-wide difficulty increase. We control for these events in the empirical analysis (see also Figures 6 and 7 and Appendix B for more details).

3 Model

3.1 Preliminaries

We model the demand for transactions using queueing theory. Transaction execution in the Ethereum blockchain is an example of a priority queueing system. Users who aim to obtain higher priority set higher per-unit transaction fee (gas price). This implies that if there are more users with higher waiting costs then, on average, the marginal gas price would be higher. Similarly, when the blockchain utilization (block occupancy rate) is higher, the cost of waiting would dominate the cost of transaction execution, which results in higher gas prices paid by the users. The predicted effect of the blockchain utilization on the gas price is hence positive.

The supply side of the blockchain consists of so-called "miners", decentralized providers of computing power who service user requests by verifying, executing and recording user-submitted transactions on the blockchain. To maximize their profits, miners sort the submitted transactions in descending gas price order, that is, the transactions with higher gas prices are included first (at the top) of the current block.[19] This means that a transaction's priority and position within its block is determined by the transaction's gas price. Moreover, miners' activity cannot be preempted by external "higher priority" job requests. Consequently, a natural way to model the transaction execution process in Ethereum is as a "non-preemptive priority queueing mechanism" (Shortle et al., 2018).

The Ethereum blockchain algorithm strives to minimize the impact of fluctuations in computer power supply (mining hash power) on transaction fees. Specifically, the Ethereum protocol is designed to keep the service rate (the average time between consecutive blocks) as stable as possible by automatically adjusting the Proof-of-Work cryptographic difficulty, so that higher availability of computational power (e.g., new miners or more computer power coming online) is quickly offset by higher cryptographic difficulty. This automatic adjustment, illustrated on Figure 7, generates a stable service rate (gas supply) in the blockchain platform, as shown on Figure 5. Fluctuations in the gas supply are thus caused by exogenous factors (randomness from small variations in the block creation rate) or updates to the blockchain protocol. As an example of the latter, the observed increase in gas supply at the beginning of the study period occurred because of the implementation, via the blockchain protocol, of a system-wide increase in the block gas limit on December 10, 2017 (see Figure 6). We control for these system-wide changes in the empirical analysis.

[19] 85% of all blocks in our data contain transactions which are perfectly sorted in descending gas price order and the rest of the blocks contain only minor exceptions from descending order. See Appendix B for more details.

3.2 Testable implications

Assume that each block is created at rate μ, called the service rate and that there are K priority transaction classes with corresponding arrival rates λ_i, $i = 1, ..., K$, where $i = 1$ denotes the highest priority class. In our setting, transaction(s) with the highest gas price constitute the highest priority class. The users' priority classes are thus not fixed in Ethereum – we can think of each user being assigned a type depending on the gas price they choose. The users' waiting time is then endogenously determined.

The total transaction arrival rate is $\lambda = \sum_{i=1}^{K} \lambda_i$. To simplify the notation, define the variable, $\rho = \frac{\lambda}{\mu}$, which measures how busy the platform is. To satisfy system stability we must have $\rho \leq 1$.[20] Applying Little's Law[21] yields

$$w_1^q = \frac{\lambda_1}{\mu(\mu - \lambda_1)} \tag{1}$$

where w_1^q is the waiting time of the highest priority class, $i = 1$. For priority classes $i = 2, ..., K$, the average waiting time is therefore

$$w_i^q = \frac{\sum_{j=1}^{i} \lambda_j}{(\mu - \sum_{k=1}^{i-1} \lambda_k)(\mu - \sum_{j=1}^{i} \lambda_j)} \tag{2}$$

There are two takeaways. First, the average waiting time of priority class i increases with the arrival rates of the higher priority classes $k \leq i$. Second, the average waiting time of each priority class $i = 1, ..., K$ decreases in the service rate μ.

The main trade-off in the queuing model is between transaction costs and waiting costs. Users of type i are assumed to have exogenous waiting costs per unit of time, c_i. The total cost incurred by user type i is thus

$$C_i = c_i w_i^q + p_i g_i$$

where p_i is the gas price bid by user i and g_i is the required gas for user i's transaction.

Specifically, if a user sets a low gas price, she would pay a low transaction fee if her transaction is picked by the block's miner. However, the transaction's priority among all other submitted

[20] If $\rho > 1$, i.e., the arrival rate is greater than the service rate, then the system queue will grow infinitely long.

[21] Little's Law, a theorem in Little (1961), asserts that the average number of customers in a queueing system is equal to the rate at which customers arrive and enter the system times the average sojourn time of a customer.

transactions would be low, which would increase the waiting time and consequently the cost of waiting. Alternatively, if a user's unit waiting cost c_i (transaction urgency) is high, then the user would choose to bid a higher gas price p_i to shorten the waiting time which again affects the total cost.

The modeled relationship and trade-off between the gas price, transaction urgency and waiting time generates the following testable hypotheses which we evaluate empirically using the Ethereum blockchain data:

Hypothesis H1. When the blockchain experiences higher utilization or congestion (high λ), the marginal and median gas price are higher, holding all other variables constant.

Hypothesis H2. When there are more urgent transactions (high c_i's), the marginal and median gas price are higher, holding all other variables constant.

The model also implies that when the service rate is higher (high μ), the marginal and median gas price would be lower, holding demand and all other variables constant. As emphasized earlier, the blockchain protocol by design automatically adjusts the cryptographic difficulty to offset changes in computing power supply and maintain a constant service rate and gas supply. Therefore, we do not include the gas supply variable in our main empirical specification and results (we show in a robustness check that including it does not affect our findings).

4 Empirical analysis

4.1 Estimation strategy

We estimate the following empirical specification:

$$p_t = \alpha + f(B_t) + \beta_2 R_t + \beta_3 X_t + \epsilon_t \tag{3}$$

The dependent variable p_t in equation (3) is the natural logarithm of the gas price (daily average) observed in the Ethereum blockchain, measured either in ETH or USD. We report results with the marginal, median and the lowest 5-th percentile gas price, as defined in Section 2.2. The variable B_t is the daily average blockchain utilization (block occupancy) which captures the usage or congestion level in the network. We estimate both a simple linear specification, $f(B_t) = \beta_1 B_t$, as well as a piece-wise linear specification (see Section 4.2) which allows for a threshold effect in blockchain utilization (a quadratic specification is also considered as robustness check). The variable R_t is the regular transactions share, included to control for changes in the composition of submitted transactions (higher vs. lower urgency). In Section 4.4, we show evidence suggesting that regular transactions have higher urgency on average compared to contract calls. Finally, X_t denotes control variables, e.g., the price of ETH in USD in Tables 1 and 2 below.

In the Ethereum blockchain, users make their transaction fee payment in ETH, the internal cryptocurrency of the platform. That is, the ETH gas price determines the transaction's priority among all other waiting transactions. However, it is possible that users are less concerned about the ETH transaction fee they must pay and more concerned about its real-value equivalent in US dollars or the national currency in which they receive income and pay bills. To capture the possible effect of the price (exchange rate) of ETH in terms of conventional currency on the gas price bidding choice of users, we also include the ETH price in USD in the specifications in which the gas price p_t is measured in ETH. An increase in the dollar price of ETH makes the ETH value of a given transaction fee relatively cheaper in USD terms, hence a user may be willing to to bid a higher ETH gas price. The expected correlation between the ETH gas price and the ETH price in USD is negative. In Section 4.3 we also consider a specification using the ETH price in terms of a basket of the three most common local currencies used by Ethereum nodes (USD, CNY and EUR).

Mapping blockchain utilization, B_t and the regular transactions share, R_t in the data to system congestion and the share of urgent transactions, respectively, in the queueing model, hypotheses H1 and H2 imply a positive association of the gas price p_t with both B_t and R_t. The daily gas supply is stable during the study period except for three system-wide events described in Appendix B, see Figure 6.[22]

4.2 Baseline results

Table 1 reports our baseline estimation results using equation (3). The dependent variable is log of the marginal gas price, as defined in Section 2, measured in ETH or USD. To correct for possible heteroskedasticity and autocorrelation in the error terms we use Newey-West standard errors with

[22] We control for these exogenous events affecting gas supply by including separate binary variables that take value of 1 on December 10-12, 2017, on January 3, 2019 and on January 21, 2019, respectively and equal 0 otherwise.

maximum lag 4.

Columns (1) and (3) in Table 1 use a linear specification for blockchain utilization, i.e., $f(B_t) = \beta_1 B_t$ in equation (3). A scatterplot of the marginal gas price and blockchain utilization, Figure 8, however, shows that their relationship is non-linear, with much larger gas prices observed when the blockchain utilization (block occupancy) is close to 1. We formally estimate this non-linear relationship in columns (2) and (4) of Table 1, where we allow for a threshold effect at 90% utilization. Specifically, in addition to $\beta_1 B_t$, we also include in $f(B_t)$ the binary variable D_t equal to 1 if $B_t > 0.9$ and zero otherwise and its interaction with the utilization variable, $TR_t = B_t * D_t$. We selected the 90% threshold using a formal structural break test (Andrews, 1993), which shows that the blockchain utilization series has a break at 0.9. In addition, at 90% utilization there may be insufficient block space to include certain types of transactions.[23] We also perform a robustness check using a quadratic form for $f(B_t)$, see Section 4.3.4 and Table A4.

The results in Table 1, columns (1) and (3) show that the blockchain utilization variable, B_t measuring the network usage or congestion is positively and statistically significantly associated (at the 1% significance level) with the marginal gas price. Columns (2) and (4), which allow for a non-linear utilization effect, further clarify that the utilization impact is negligible below the threshold level of 0.9 (the B_t coefficient is weakly or not statistically significantly different from zero). However, in both the ETH and USD specifications (2) and (4), the coefficient estimate on the threshold interaction term TR_t is positive and statistically significantly different from zero at the 1% level, implying a strongly non-linear relationship between blockchain utilization and the gas price above 90 percent utilization. Quantitatively, the estimate 27.85 in column (4) means that 0.01 increase in blockchain utilization above the threshold level of 0.9 is associated with $100(e^{27.85(0.01)})$ or 32.1% increase in the marginal gas price in USD. Given that the mean USD marginal gas price in our sample is \$0.45(10^{-5}) (see Table A1), a user executing a transaction with complexity 100,000 gas would pay $0.45(10^{-5})(10^5)(32.1\%) = 0.15$ USD (15 cents) more on average following a 0.01 increase in the blockchain utilization when the latter is above 90%.

The regular transactions share, R_t has positive and statistically significant estimates in all four specifications of Table 1. This implies that a larger share of regular transactions is associated on

[23] Contract creations are the most complex transactions, with the largest gas requirement. A direct examination of the data shows that there are contract creation transactions using more than 10% of their block gas limit.

Table 1: Main results

	marginal gas price, ETH		marginal gas price, USD	
	(1)	(2)	(3)	(4)
blockchain utilization, B_t	2.263***	0.090	3.310***	0.885*
	(0.624)	(0.306)	(0.796)	(0.476)
utilization $> 90\%$, TR_t	-	33.03***	-	27.85***
	-	(4.052)	-	(4.868)
regular transactions share, R_t	4.464***	2.539***	8.255***	6.659***
	(1.000)	(0.668)	(0.796)	(0.744)
ETH price in USD, X_t	-0.732***	-0.651***	-	-
	(0.208)	(0.160)	-	-
sample size	457	457	457	457
R-squared	0.293	0.595	0.647	0.760

OLS regressions with daily-level data including a constant. Newey-West standard errors reported in the parentheses. The dependent variable, "marginal gas price" is the natural logarithm of the minimum observed gas price (in 10^{-8} ETH or 10^{-5} USD) over all transactions in a block, averaged over all blocks created on day t. "Blockchain utilization" is the ratio between the total gas requirement for all transactions on day t and the total day-t gas supply. "Utilization $> 90\%$" is the product of the blockchain utilization, B_t and a binary variable which equals 1 if $B_t > 0.9$ and zero otherwise. "Regular transactions share" is the fraction of regular transactions in all day-t transactions. We include dummies for the system-wide gas supply changes on December 10-12, 2017, January 3, and January 21, 2019. "ETH price in USD" is the daily average ETH price in 10^3 USD. *, **, *** denote 10%, 5%, and 1% significance level, respectively.

average with a higher gas price in ETH and USD. For example, the estimate 6.66 in column (4) suggests that a 0.01 increase in the regular transactions share is associated with $100(e^{6.66(0.01)})$ or 6.9 % increase in the marginal gas price in USD. This implies that a user executing a transaction with complexity 100,000 gas would pay $0.45(10^{-5})(10^5)(6.9\%) = 0.03$ USD (3 cents) more, on average, following a 0.01 increase in the regular transactions share, holding all else equal.

We also include the price of ETH in USD in specifications (1) and (2) in Table 1, to measure the responsiveness of blockchain users to the conventional currency cost of a given ETH-denominated transaction fee. We obtain a negative estimate which implies that, when executing a transaction is more expensive in USD terms, the marginal user tends to bid a lower ETH gas price, all else equal. This result is consistent with the notion that users take the real value of transaction costs into consideration in their bidding decisions.

Our main results in Table 1 show strong evidence in support of Hypotheses H1 and H2 from Section 3. Higher blockchain utilization, measured by the average block occupancy rate, is associated with a higher marginal gas price (Hypothesis H1). Higher share of urgent transactions, proxied by the regular transactions share, corresponds to a higher marginal gas price (Hypothesis H2).

So far we interpreted the right-hand side variables in equation (3) as exogenous. However, a potential simultaneity problem may exist – a higher gas price makes all transactions costlier but the associated impact could be larger for more complex transactions such as contract calls and creations.[24] Hence, a higher gas price may raise the share of users choosing regular transactions because of the lower gas requirement.

To address the potential simultaneity between the gas price, p_t and the transaction type choice, we use the first lag of the regular transactions share, R_{t-1} and the daily number of new sender accounts as instruments. Since R_{t-1} was decided at time $t-1$, it does not have a direct causal impact on the dependent variable p_t, while R_t and R_{t-1} are highly correlated (Reed, 2015). Similarly, the number of new sender accounts is highly correlated with the regular transactions share (correlation 0.68) and does not have a direct causal effect on the minimum gas price.

Table 2 reports the IV regression results, using the same specifications (linear or piece-wise linear in blockchain utilization) of equation 3 as in Table 1. The first-stage regresses the regular transactions share R_t on the instruments, blockchain utilization and the price of ETH in USD. The reported coefficient estimates in Table 2 are for the instrumented share. We also report two IV diagnostic statistics – the Cragg-Donald Wald F-statistic for weak instruments and the chi-squared Durbin-Wu-Hausman endogeneity test. The null hypothesis is that all regressors are exogenous which can be rejected if the test statistics are sufficiently large.

In all specifications in Table 2, our main results from Table 1 remain robust and the coefficient estimates change in only minor ways. In addition, the Cragg-Donald Wald F-statistics are much greater than 10, the standard threshold for weak instruments, which suggests that the chosen instruments are not weak (Staiger and Stock, 1997). The DWH endogeneity test value is smaller than the critical level, showing that we fail to reject the null that all regressors are exogenous. Overall, the Table 2 results suggest that our main conclusions are robust to the potential simultaneity

[24]Contracts and regular transactions can be used interchangeably for some purposes, e.g., sending ETH between accounts could be done by both a contract call and a regular transaction.

Table 2: Instrumental variables (IV) results

	marginal gas price, ETH		marginal gas price, USD	
	(1)	(2)	(3)	(4)
blockchain utilization, B_t	2.266***	0.114	3.154***	0.830*
	(0.624)	(0.303)	(0.929)	(0.471)
utilization > 90%, TR_t	-	32.37***	-	26.79***
	-	(4.367)	-	(5.750)
regular transactions share, R_t	5.621***	2.990***	8.830***	7.130***
	(1.494)	(0.981)	(1.419)	(1.218)
ETH price in USD, X_t	-0.980***	-0.743***	-	-
	(0.287)	(0.191)	-	-
sample size	456	456	456	456
F-statistic (C-D weak instrument test)	267.1	239.7	327.7	282.0
χ^2-stat (DWH endogeneity test)	2.455	0.761	0.385	0.408
R-squared	0.285	0.594	0.646	0.759

Notes: Instrumental variables (IV) regressions with daily data including a constant. Newey-West standard errors reported in the parentheses. The first-stage regresses the regular transactions share on its first lag, the total number of new sender accounts and on the blockchain utilization and the ETH price in USD. The dependent variable "marginal gas price" is the natural logarithm of the minimum observed gas price (in 10^{-8} ETH or 10^{-5} USD) over all included transactions in a block, averaged over all blocks created on day t. "Blockchain utilization" is the ratio between the total gas requirement for all transactions on day t and the total day-t gas supply. "Utilization > 90%" is the product of the blockchain utilization, B_t and a binary variable that equals 1 if $B_t > 0.9$ and zero otherwise. "Regular transactions share" is the fraction of regular transactions in all day-t transactions. We include dummies for the system-wide gas supply changes on December 10-12, 2017, January 3, and January 21, 2019. "ETH price in USD" is the daily average ETH price in 10^3 USD. *, **, *** denote 10%, 5%, and 1% significance level, respectively.

problem in the regular transactions share.

4.3 Robustness and sensitivity analysis

4.3.1 Using a basket of currencies

In Tables 1 and 2 we control for the price of ETH in USD (in columns 1 and 2) or use the marginal gas price in USD as the dependent variable (in columns 3 and 4), to incorporate the possibility that the users' choice of gas price may be affected by the real cost of the transaction fee, in terms of the users' local conventional currency. In this section we extend this analysis and check the sensitivity of our results by considering the price of ETH in terms of a basket of currencies (BoC), as opposed to USD only. Specifically, we take the local currencies of the three geographic locations with the largest numbers of active Ethereum nodes (USA, China and the Euro-zone, see Kim et al., 2018) and use the share of active nodes from each location as weights for the respective currencies in the basket, USD, CNY and EUR (see Appendix B for a detailed description).

Table 3: Marginal gas price results – Basket of currencies (BoC)

	marginal gas price, ETH		marginal gas price, BoC	
	(1)	(2)	(3)	(4)
blockchain utilization, B_t	2.263***	0.085	3.322***	0.898*
	(0.624)	(0.306)	(0.795)	(0.475)
utilization > 90%, TR_t	-	33.01***	-	27.83***
	-	(4.043)	-	(4.854)
regular transactions share, R_t	4.421***	2.489***	8.206***	6.611***
	(1.003)	(0.665)	(0.796)	(0.745)
ETH price (BoC), X_t	-0.570***	-0.504***	-	-
	(0.166)	(0.127)	-	-
sample size	457	457	457	457
R-squared	0.291	0.593	0.646	0.759

OLS regressions with daily-level data including a constant. Newey-West standard errors reported in the parentheses. The dependent variable, "marginal gas price" is the natural logarithm of the minimum observed gas price (in 10^{-8} ETH or 10^{-5} BoC, weighted basket of currencies consisting of USD, EUR, and CNY) over all transactions in a block, averaged over all blocks created on day t. "Blockchain utilization" is the ratio between the total gas requirement for all transactions on day t and the total day-t gas supply. "Utilization > 90%" is the product of the blockchain utilization, B_t and a binary variable that equals 1 if $B_t > 0.9$ and zero otherwise. "Regular transactions share" is the fraction of regular transactions in all day-t transactions. We include dummies for the system-wide gas supply changes on December 10-12, 2017, January 3, and January 21, 2019. "ETH price (BoC)" is the daily average ETH price in terms of a basket of currencies consisting of US Dollars, Euro, and Chinese Yuan, weighted by the country/region specific number active nodes (see Appendix B). *, **, *** denote 10%, 5%, and 1% significance level, respectively.

Tables 3 and 4 repeat the analysis from Tables 1 and 2, using the price of ETH in terms of the basket of USD, CNY and EUR currencies (BoC). Our main results are essentially unchanged. The relationship between blockchain utilization and the marginal gas price is positive and strongly non-linear. The regular transactions share remains statistically significantly positively associated with the gas price in all specifications. The estimates on ETH price (BoC) in columns (1) and (2) show that the gas price is negatively associated with the ETH value in terms of the basket of currencies.

Table 4: Instrumental variables (IV) results – Basket of currencies (BoC)

	marginal gas price, ETH		marginal gas price, BoC	
	(1)	(2)	(3)	(4)
blockchain utilization, B_t	2.266***	0.112	3.168***	0.844*
	(0.624)	(0.302)	(0.927)	(0.470)
utilization > 90%, TR_t	-	32.52***	-	26.79***
	-	(4.348)	-	(5.726)
regular transactions share, R_t	5.541***	2.903***	8.772***	7.072***
	(1.493)	(0.968)	(1.417)	(1.212)
ETH price (BoC), X_t	-0.762***	-0.571***	-	-
	(0.228)	(0.151)	-	-
sample size	456	456	456	456
F-statistic (C-D weak instrument test)	267.5	239.8	327.7	282.0
χ^2-stat (DWH endogeneity test)	2.367	0.659	0.378	0.399
R-squared	0.288	0.592	0.645	0.758

Notes: Instrumental variables (IV) regressions with daily data including a constant. Newey-West standard errors reported in the parentheses. The first-stage regresses the regular transactions share on its first lag and the total number of new sender accounts and on the blockchain utilization and the price of ETH. The dependent variable "marginal gas price" is the natural logarithm of the minimum observed gas price (in 10^{-8} ETH or 10^{-5} BoC, weighted basket of currencies consisting of USD, EUR, and CNY) over all included transactions in a block, averaged over all blocks created on day t. "Blockchain utilization" is the ratio between the total gas requirement for all transactions on day t and the total day-t gas supply. "Utilization > 90%" is the product of the blockchain utilization, B_t and a binary variable that equals 1 if $B_t > 0.9$ and zero otherwise. "Regular transactions share" is the fraction of regular transactions in all day-t transactions. We include dummies for the system-wide gas supply changes on December 10-12, 2017, January 3, and January 21, 2019. "ETH price (BoC)" is the daily average ETH price in terms of a basket of currencies consisting of US Dollars, Euro, and Chinese Yuan, weighted by the country/region specific number active nodes (see Appendix B). *, **, *** denote 10%, 5%, and 1% significance level, respectively.

4.3.2 Alternative marginal gas price definition

The marginal gas price used in our baseline analysis in Table 1 was defined as the daily average of the *minimum* gas price observed in each block recorded on that day. A possible concern could be that the minimum gas price in a block may be unusually low, e.g., because of miner error or other reasons. This happens very rarely in our data, since executing transactions with very low gas price is not profitable for the miners. Still, to address potential concerns with using the minimum gas price in a block, we instead construct an alternative marginal gas price variable defined as the daily average of the *lowest 5-th percentile* gas price in each block recorded on that day. In this way any extreme or outlier gas price values are avoided, while we still focus on transactions that are at the margin or very close to the margin of being included vs. not included in a block.

Table A1 in the Appendix shows descriptive statistics for the lowest 5-th percentile of the gas price (daily average). As expected, its mean is slightly higher than that of the baseline marginal gas price definition (the block minimum gas price) used in Table 1, but otherwise their distributions are very similar. Table 5 reports estimation results using this alternative marginal gas price definition. We do not observe any notable change in our main results when using the lowest 5-th percentile gas price as the dependent variable. Blockchain utilization has a positive and statistically significant estimate in columns (1) and (3) as before. The relationship between blockchain utilization and the lowest 5-th percentile gas price is strongly non-linear, reflected in the large positive and statistically significant threshold coefficients in columns (2) and (4). Likewise, the share of regular transactions has positive and significant coefficient estimates in all specifications. The negative and statistically significant estimates in columns (1) and (2) shows that the ETH gas price remains negatively associated with the price of ETH in USD, as in Tables 1 and 2. These results show that our main results are not sensitive to outliers in the marginal gas price.

4.3.3 Median gas price

Above we focused on the gas price for transactions on the margin of being included vs. not included in a block. We now look instead at the median transaction in terms of the gas price distribution within a block. We define the *median gas price* as the daily average of the median gas prices in each block recorded on a given day.[25]

Table A1 in the Appendix reports summary statistics for the median block gas price. Its mean value over the studied period is more than twice larger than that of the marginal (block minimum) gas price but its distribution is right-skewed, similar to that of the marginal gas price. Table 6 reports regression results using log of the median gas price as the dependent variable. The results are broadly consistent with the theory hypotheses H1 and H2 and our baseline results for the marginal gas price, except for a few differences explained below. The estimate on blockchain utilization is positive and statistically significant and the non-linear 90% threshold effect is still present and large in magnitude. The regular transactions share is positively and statistically significantly associated

[25]We do not study the mean gas price as it can be heavily influenced by outliers, e.g., abnormally high gas prices. See, for example, cryptonews.com/news/ethereum-transaction-fee-mystery-just-got-more-mysterious-6816.htm.

Table 5: Alternative marginal gas price definition, lowest 5-th percentile

	lowest 5-th percentile, ETH		lowest 5-th percentile, USD	
	(1)	(2)	(3)	(4)
blockchain utilization, B_t	2.239***	0.123	3.324***	0.957*
	(0.619)	(0.305)	(0.795)	(0.485)
utilization > 90%, TR_t	-	32.40***	-	27.18***
	-	(4.071)	-	(4.942)
regular transactions share, R_t	4.587***	2.703***	8.558***	7.000***
	(0.980)	(0.670)	(0.786)	(0.759)
ETH price in USD, X_t	-0.663***	-0.584***	-	-
	(0.204)	(0.160)	-	-
sample size	457	457	457	457
R-squared	0.314	0.598	0.658	0.761

Notes: OLS regressions with daily-level data including a constant. Newey-West standard errors reported in the parentheses. The dependent variable "lowest 5-th percentile" gas price is the natural logarithm of the bottom 5-th percentile of gas prices (in 10^{-8} ETH or 10^{-5} USD) in each block, averaged across all blocks created on day t. "Blockchain utilization" is the ratio between the total gas requirement for all transactions on day t and the total day-t gas supply. "Utilization > 90%" is the product of the blockchain utilization, B_t and a binary variable that equals 1 if $B_t > 0.9$ and zero otherwise. "Regular transactions share" is the fraction of regular transactions in all day-t transactions. We include dummies for the system-wide gas supply changes on December 10-12, 2017, January 3, and January 21, 2019. "ETH price in USD" is the daily average ETH price in 10^3 USD. *, **, *** denote 10%, 5%, and 1% significance level, respectively.

with the median gas price.

One difference from Table 1 when we use the median gas price as the dependent variable is that the price of ETH in USD does not have a statistically significant association with the median gas price. This can be interpreted as the median user being less concerned about the real cost of transaction execution, unlike the marginal user in the left tail of the gas price distribution. Possibly, the median user may be wealthier compared to the marginal user and less sensitive to the transaction fee, or the median user may be more risk-averse or have higher urgency hence bidding a higher gas price to avoid delays. Unfortunately, we do not have data to directly test these possible mechanisms.[26]

[26] The reason for posting a marginal gas price vs. median gas price transaction could differ too - e.g., a non time-sensitive advertising or faucet payment from a website vs. time-sensitive transfer or purchase.

Table 6: Median gas price

	median gas price, ETH		median gas price, USD	
	(1)	(2)	(3)	(4)
blockchain utilization, B_t	0.790*	-0.581**	2.187***	0.572
	(0.490)	(0.268)	(0.701)	(0.523)
utilization > 90%, TR_t	-	24.38***	-	18.81***
	-	(3.819)	-	(5.047)
regular transactions share, R_t	3.954***	2.596***	9.042***	8.329***
	(0.748)	(0.588)	(0.735)	(0.835)
ETH price in USD, X_t	-0.094	-0.037	-	-
	(0.161)	(0.139)	-	-
sample size	457	457	457	457
R-squared	0.371	0.580	0.676	0.726

Notes: OLS regressions with daily-level data including a constant. Newey-West standard errors reported in the parentheses. The dependent variable "median gas price" is the natural logarithm of the median gas price (in 10^{-8} ETH or 10^{-5} USD) in each block, averaged across all blocks created on date t. "Blockchain utilization" is the ratio between the total gas requirement for all transactions on day t and the total day-t gas supply. "Utilization > 90%" is the product of the blockchain utilization, B_t and a binary variable that equals 1 if $B_t > 0.9$ and zero otherwise. "Regular transactions share" is the fraction of regular transactions in all day-t transactions. We include dummies for the system-wide gas supply changes on December 10-12, 2017, January 3, and January 21, 2019. "ETH price in USD" is the daily average ETH price in 10^3 USD. *, **, *** denote 10%, 5%, and 1% significance level, respectively.

4.3.4 Additional robustness checks

We perform four additional robustness checks of our main results.

Quadratic model. First, in Table A4 in the Appendix we consider a quadratic specification for the effect of blockchain utilization on the gas price, the term $f(B_t)$ in equation (3). We demean the blockchain utilization variable to avoid multicollinearity issues and for easier interpretation of the coefficient on the squared term. Consistent with our main results in Tables 1, 5 and 6 we find that blockchain utilization is positively and statistically significantly associated with the gas price, with a strong non-linear effect reflected in the large coefficient on squared utilization. The regular transactions share estimate remains positive and statistically significantly different from zero at the 1% significance level. The price of ETH in USD has a negative estimate which is statistically significant, except for the median gas price, as in Table 6.

Other alternative specifications. In Appendix Table A5 we re-estimate our main specifications from Table 1 by using as dependent variable the marginal gas price in levels instead of logs. Our main results remain robust – there is a significant positive non-linear relationship, with a strong threshold effect above 90% utilization, between the blockchain utilization and the marginal gas price (in ETH or USD). The results about the regular transactions share and the ETH price in USD (in columns 1 and 2) also remain robust.

In Appendix Table A6 we re-estimate equation (3) by additionally including gas supply, defined as the sum of the gas limits of all blocks created in a given day. As in all tables, we control separately for the system-wide blockchain protocol events (jumps) affecting the block gas limit or difficulty via dummy variables. The gas supply estimate not statistically significantly distinguishable from zero which confirms that, once we account for the system-wide events, the blockchain gas supply is stable and any remaining minor fluctuations in it do not affect the gas price on average. Our main results, on the non-linear positive association of blockchain utilization with the gas price, the positive association of the regular transaction share with the gas price, and the negative relationship of the ETH price in USD with the gas price remain robust and very close in magnitude to the baseline estimates in Tables 1 and 2.

Finally, in Appendix Table A7 we re-estimate the ETH specifications with a non-linear utilization effect (column (2) in Tables 1 and 2) by using the first lag of the ETH price in USD which allows for potential delays in the users' transaction posting behavior or bid gas prices in reaction to changes in the ETH/USD exchange rate.[27] Our results remain essentially unchanged.

4.4 Are regular transactions more urgent?

We examine further our results about the regular transactions share and its positive correlation with the observed gas prices. Specifically, we hypothesize that regular transactions are more urgent and associated with higher waiting costs. We construct a variable which measures the average position of regular transactions within a block, defined as follows. Suppose there are n regular transactions in a block containing N total transactions, where $N > n > 0$. Define the average regular transaction position (RTP) as:

$$\text{RTP} = 1 - \frac{\sum\limits_{i \in \text{regular}} pos_i/n - \frac{n+1}{2}}{N - n}$$

[27]For example, many new users joined Ethereum during the peak period when the USD price of ETH went up significantly (see Figures 10 and 12).

where $pos_i \in [1, N]$ denotes the position of regular transaction i in its block relative to the block middle. Remember from Section 3.1 that transactions are ordered in descending gas price order within each block; that is, more urgent transactions, with higher gas prices, are recorded nearer the top of the block. By construction, the RTP measure takes values between 0 and 1. A RTP closer to 1 (larger than 1/2) means that regular transactions are recorded on average nearer the top (beginning) of the block, i.e., their associated gas prices are higher.[28] The opposite is true if RTP is closer to 0 (smaller than 1/2). Figure 9 plots the daily average RTP value. The Figure confirms that regular transactions are indeed located, on average, in the upper (higher gas price) half of the Ethereum blocks.[29] This supports our hypothesis that regular transactions are more urgent on average and likely to be associated with higher waiting costs and justifies using the regular transactions share, R_t as proxy for transaction urgency in the estimation equation (3).

We also checked whether there is any bunching in the gas prices of regular transactions, since blockchain users often use special software ("wallet") to automate transaction execution and set the transaction fee. If the software systematically picks relatively high gas price values, and if many regular transaction are posted by such wallet users, then this could be an alternative explanation for the observed upper-half relative position of regular transactions in their blocks depicted on Figure 9. Direct inspection of the data confirms that there is no noticeable bunching in the gas prices for regular transactions.

4.5 Peak period

In this section we analyze a specific three-month sub-period of our sample, which we call the "peak period", between November 1, 2017 to January 31, 2018. This is a time in which the USD price of ETH increased sharply and the blockchain utilization was very high. Figure 10 shows that the ETH price in USD went up by 400% during this period. In addition, Figure 2 shows a substantial increase in blockchain utilization. This period is also associated with a large and growing number of new Ethereum accounts/addresses (more than 10,000 per day), part of which could be because of new users joining and transacting at least once on the Ethereum platform, see Figure 12. Furthermore,

[28]For example, suppose there are N=6 transactions in a block, n=3 of which are regular transactions. If the positions of the regular transactions are 1, 2 and 4 (i.e., toward the beginning/top of the block), then the RTP measure equals $1-(7/3-2)/(6-3) = 0.89$ which is larger than 0.5. Similarly, if the positions of the regular transactions in the block were instead 3, 5 and 6 (i.e., toward the end/bottom of the block), then RTP equals $1-(14/3-2)/(6-3) = 0.11$.

[29]The mean and median RTP in our data are both equal to 0.57.

the peak period exhibits sharp oscillations in the blockchain conditions: from an individual user's perspective the average per unit transaction cost could vary by more than 200% from one day to the next, see Figure 1. Table A1 in the Appendix reports summary statistics for the main blockchain variables during the peak period. There is a noticeable difference in the gas price levels compared to the full sample – both the marginal and median gas prices set by users are much higher during the peak period. In addition, the share of regular transactions is considerably larger during the peak period compared to the full sample.

Table 7: Peak period (November 2017 - January 2018)

	gas price in ETH					
	marginal		lowest 5-th percentile		median	
	(1)	(2)	(3)	(4)	(5)	(6)
blockchain utilization, B_t	2.787**	0.534	2.921***	0.858	1.519**	0.626
	(1.122)	(0.741)	(1.039)	(0.695)	(0.681)	(0.579)
utilization > 90%, TR_t	-	17.43***	-	16.19***	-	9.554***
	-	(3.097)	-	(2.998)	-	(2.702)
regular transactions share, R_t	3.419**	1.185	3.243**	1.194	1.972*	1.035
	(1.416)	(0.933)	(1.333)	(0.863)	(1.014)	(0.769)
ETH price in USD, X_t	-0.336	0.057	-0.286	0.072	0.346	0.480**
	(0.483)	(0.324)	(0.451)	(0.299)	(0.296)	(0.235)
sample size	92	92	92	92	92	92
R-squared	0.642	0.854	0.684	0.867	0.733	0.819

Notes: OLS regressions with daily-level data including a constant. Newey-West standard errors reported in the parentheses. The dependent variable "marginal gas price" is the natural logarithm of the minimum observed gas price (in 10^{-8} ETH) over all included transactions in a block, averaged across all blocks created on day t. The dependent variable "lowest 5-th percentile" gas price is the natural logarithm of the bottom 5-th percentile of gas prices (in 10^{-8} ETH) in each block, averaged across all blocks created on day t. The dependent variable "median gas price" is the natural logarithm of the median gas price (in 10^{-8} ETH) in each block, averaged across all blocks recorded on day t. "Blockchain utilization" is the ratio between the total gas requirement for all transactions on day t and the total day-t gas supply. "Utilization > 90%" is the product of the blockchain utilization, B_t and a binary variable that equals 1 if $B_t > 0.9$ and zero otherwise. "Regular transactions share" is the fraction of regular transactions in all day-t transactions. We include dummies for the system-wide gas supply changes on December 10-12, 2017, January 3, and January 21, 2019. "ETH price in USD" is the daily average ETH price in 10^3 USD. *, **, *** denote 10%, 5%, and 1% significance level, respectively.

Because the peak period looks different from the full sample in several dimensions, we test the robustness of our main results using the peak period data only. This helps us evaluate whether there are any significant omitted factors determining gas prices in the peak period or whether the observed high gas prices and utilization rate were mainly driven by demand and the other factors we account for, as in the full sample. We run regressions using the specifications in Table 1 for each of the three gas price measures (marginal, lowest 5-th percentile and median). Table 7 and Appendix Table A8 summarize the results. Columns (1) and (2) report results using the daily average of the marginal gas price as the dependent variable while columns (3)-(4) and (5)-(6) report results using the lowest 5-th percentile gas price and the median gas price, respectively.

The results in Tables 7 and A8, using gas prices in ETH and USD respectively, show that the peak-period results remain broadly in line with our full-sample findings, although the peak period estimates are noisier. Blockchain utilization still has a positive and strongly non-linear association with the gas price overall, however the standard errors are larger than in the full sample results. A larger regular transactions share is positively associated with higher gas prices in columns (1), (3) and (5), as in the baseline results, but the estimates are not statistically significantly different from zero in the specifications with a non-linear utilization threshold, columns (2), (4) and (6). The regular transactions share estimates are statistically significantly positive in all columns of Appendix Table A8, when the gas price is measured in USD. The estimate on the ETH price in USD in Table 7 is not significantly different from zero, except in column (6), which could be explained by the possibility that the sharply increasing ETH price in USD in the peak period may have been seen as a (speculative) investment or trading opportunity by many users, despite the high associated transaction costs.

5 Conclusions

Blockchain cryptocurrency platforms like Ethereum are unique examples of financial markets that, because of their decentralized and anonymous nature, remain largely unregulated by official authorities. Furthermore, these digital markets lack external formal contract enforcement, beyond the algorithmic protocol and computer code of the platform itself (see Karaivanov, 2021 and Townsend, 2020 for further discussion). There has also been a lot of debate on the extreme volatility, speculation activity and illicit transactions on these platforms (for example, Foley et al., 2019; Griffin and Shams, 2020; Li et al., 2020). It is therefore a valid empirical question whether blockchain platforms like Ethereum operate as financial markets for payments and related services, subject to standard demand and supply and other economic factors.

We find that the answer is by and large affirmative. We analyze transaction fees (gas prices) in the Ethereum blockchain and find that demand factors, measured by block utilization rates, and the choice of transaction type are the primary economic determinants of observed fees. Blockchain utilization has a positive association with the marginal and median gas price which is strongly non-linear above the 90% block utilization threshold. Additionally, we find that blockchain users endogenously vary the mix of regular transfers vs. smart contract transactions, with a larger fraction of regular transactions observed when the gas price is high.

We abstracted from studying the determinants of the ETH price in USD, instead treating it as a potential factor affecting gas prices. Further research on the joint determinants of blockchain transaction fees and the ETH/USD exchange rate could be beneficial. In addition, more research on the network structure of Ethereum blockchain addresses and on the heterogeneity of recorded transactions (by size, sender address, or recipient address) can provide additional useful insights on the economics of blockchain platforms.

References

[1] Andrews, D. (1993), "Tests for parameter instability and structural change with unknown change point" *Econometrica* 61(4): 821-856.

[2] Antonopoulos, A., and G. Wood (2018), *Mastering Ethereum: building smart contracts and dapps*. O'reilly Media.

[3] Athey, S., I. Parashkevov, V. Sarukkai and J. Xia (2016), "Bitcoin pricing, adoption, and usage: Theory and evidence" Stanford Working Paper No. 3469.

[4] Buterin, V. (2013), "Ethereum white paper" https://ethereum.org/en/whitepaper.

[5] Catalini, C. and J. Gans (2016), "Some simple economics of the blockchain", NBER Working Paper 22952.

[6] Ciaian, P., M. Rajcaniova and A. Kancs (2016), "The economics of Bitcoin price formation" *Applied Economics* 48(19): 1799-1815.

[7] Cong, L. and Z. He (2019), "Blockchain disruption and smart contracts" *Review of Financial Studies* 32(5): 1754-1797.

[8] Corbet, S., B. Lucey, and L. Yarovaya (2018), "Datestamping the Bitcoin and Ethereum bubbles" *Finance Research Letters* 26: 81-88.

[9] Chiu, J., and T. Koeppl (2019), "Blockchain-based settlement for asset trading" *Review of Financial Studies* 32(5): 1716-1753.

[10] Easley, D., M. O'Hara, and S. Basu (2019), "From mining to markets: The evolution of bitcoin transaction fees" *Journal of Financial Economics* 134(1): 91-109.

[11] Foley, S., J. Karlsen, and T. Putnins (2019), "Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies?", *Review of Financial Studies* 32(5): 1798-1853.

[12] Gandal, N., J. Hamrick, T. Moore, and T. Oberman (2018), "Price manipulation in the Bitcoin ecosystem", *Journal of Monetary Economics* 95: 86-96.

[13] Griffin, J. and A. Shams (2020), "Is Bitcoin really untethered?", *Journal of Finance* 75(4): 1913-64.

[14] Huberman, G., J. Leshno, and C. Moallemi (2019), "An economic analysis of the bitcoin payment system." *Columbia Business School Research Paper* 17-92.

[15] Karaivanov, A. (2021), "Blockchains, collateral and financial contracts", SFU working paper dp21-03.

[16] Kasahara, S. and J. Kawahara (2019), "Effect of Bitcoin fee on transaction-confirmation process" *Journal of Industrial & Management Optimization* 15(1): 365-386.

[17] Kim, S., Z. Ma, S. Murali, J. Mason, A. Miller, and M. Bailey (2018), "Measuring Ethereum network peers" in *Proceedings of the Internet Measurement Conference 2018*, pp. 91-104.

[18] Kristoufek, L. (2015), "What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis" *PLOS ONE* 10(4): e0123923.

[19] Li, Q., J. Ma, Y. Chang, F. Ma and H. Yu (2019), "Markov processes in blockchain systems" *Computational Social Networks* 6(1): 1-28.

[20] Li, T., D. Shin and B. Wang (2020), "Cryptocurrency pump-and-dump schemes" *SSRN working paper no. 3267041.*

[21] Möser, M. and R. Böhme (2015), "Trends, tips, tolls: A longitudinal study of Bitcoin transaction fees" *International Conference on Financial Cryptography and Data Security*, Springer, pp. 19-33.

[22] Reed, W. (2015), "On the practice of lagging variables to avoid simultaneity" *Oxford Bulletin of Economics and Statistics* 77(6): 897-905.

[23] Shortle, J., J. Thompson, D. Gross and C. Harris (2018), "Advanced Markovian Queueing Models", in *Fundamentals of Queueing Theory*, vol. 399. John Wiley & Sons, pp. 180-184.

[24] Staiger, D. and J. Stock (1997), "Instrumental variables regression with weak instruments" *Econometrica* 65(3): 557-586.

[25] Takagi, S. (1986), "Pegging to a currency basket" *Finance and Development* 23(3): 41.

[26] Townsend, R. (2020), *Distributed Ledgers*, MIT Press, Cambridge, MA.

[27] Yermack, D. (2014), "Is Bitcoin a real currency? An Economic Appraisal", NBER Working Paper 19747.

[28] Zochowski, M. (2019), "Crypto transaction fee economics primer", *Medium*, Feb. 9, 2019.

Appendix

A. Additional Tables and Figures

Table A1: Summary statistics

Variables	# of obs.	Mean	Std	25^{th} pctile	50^{th} pctile	75^{th} pctile
	Full sample, Nov. 1, 2017 - Jan. 31, 2019					
marginal gas price, 10^{-8} ETH	457	0.82	0.89	0.43	0.55	0.76
marginal gas price, 10^{-5} USD	457	0.45	0.76	0.11	0.18	0.42
marginal gas price, 10^{-5} BoC	457	0.56	0.95	0.14	0.23	0.52
lowest 5 pctile gas price, 10^{-8} ETH	457	0.89	0.96	0.46	0.58	0.80
lowest 5 pctile gas price, 10^{-5} USD	457	0.49	0.84	0.12	0.20	0.47
lowest 5 pctile gas price, 10^{-5} BoC	457	0.62	1.05	0.15	0.25	0.59
median gas price, 10^{-8} ETH	457	1.67	1.39	0.98	1.24	1.58
median gas price, 10^{-5} USD	457	0.94	1.45	0.23	0.44	1.03
median gas price, 10^{-5} BoC	457	1.19	1.81	0.30	0.54	1.32
gas supply, 10^{10} gas/day	457	4.60	0.23	4.52	4.69	4.75
blockchain utilization	457	0.82	0.09	0.76	0.84	0.90
regular transactions share	457	0.52	0.08	0.47	0.51	0.55
ETH price in USD, 10^3 USD	457	0.46	0.29	0.21	0.44	0.67
ETH price (BoC), 10^3 BoC	457	0.58	0.36	0.27	0.55	0.84
	Peak period, Nov. 1, 2017 - Jan 31, 2018					
	# of obs.	Mean	Std	25^{th} pctile	50^{th} pctile	75^{th} pctile
marginal gas price, 10^{-8} ETH	92	1.38	1.20	0.50	0.77	2.09
marginal gas price, 10^{-5} USD	92	1.13	1.32	0.20	0.65	1.45
marginal gas price, 10^{-5} BoC	92	1.42	1.66	0.25	0.80	1.83
lowest 5 pctile gas price, 10^{-8} ETH	92	1.55	1.32	0.56	0.90	2.32
lowest 5 pctile gas price, 10^{-5} USD	92	1.28	1.47	0.22	0.77	1.56
lowest 5 pctile gas price, 10^{-5} BoC	92	1.60	1.84	0.28	0.95	1.98
median gas price, 10^{-8} ETH	92	2.98	1.90	1.54	2.24	3.87
median gas price, 10^{-5} USD	92	2.46	2.48	0.58	1.75	2.87
median gas price, 10^{-5} BoC	92	3.08	3.12	0.72	2.10	3.75
gas supply, 10^{10} gas/day	92	4.36	0.28	4.12	4.34	4.62
blockchain utilization	92	0.82	0.15	0.67	0.87	0.94
regular transactions share	92	0.61	0.09	0.53	0.62	0.70
ETH price in USD, 10^3 USD	92	0.69	0.33	0.41	0.69	1.02
ETH price (BoC), 10^3 BoC	92	0.87	0.41	0.50	0.86	1.25

Notes: summary statistics of the blockchain variables aggregated to the daily level. The full sample (peak period) statistics are reported in the upper (lower) part of the Table respectively. BoC denotes "basket of currencies" – see Section 4.3.1 and Appendix B.

Table A2: Alternative marginal gas price definition, lowest 5-th percentile (BoC)

	lowest 5-th percentile, ETH		lowest 5-th percentile, BoC	
	(1)	(2)	(3)	(4)
blockchain utilization, B_t	2.239***	0.118	3.336***	0.970*
	(0.619)	(0.305)	(0.794)	(0.493)
utilization > 90%, TR_t	-	32.38***	-	27.16***
	-	(4.062)	-	(4.928)
regular transactions share, R_t	4.549***	2.658***	8.503***	6.952***
	(0.982)	(0.667)	(0.787)	(0.759)
ETH price (BoC), X_t	-0.517***	-0.452***	-	-
	(0.163)	(0.127)	-	-
sample size	457	457	457	457
R-squared	0.313	0.597	0.657	0.761

Notes: OLS regressions with daily-level data including a constant. Newey-West standard errors reported in the parentheses. The dependent variable "lowest 5-th percentile" gas price is the natural logarithm of the bottom 5-th percentile of gas prices (in 10^{-8} ETH or 10^{-5} BoC, weighted basket of currencies consisting of USD, EUR, and CNY) in each block, averaged across all blocks created on day t. "Blockchain utilization" is the ratio between the total gas requirement for all transactions on day t and the total day-t gas supply. "Utilization > 90%" is the product of the blockchain utilization, B_t and a binary variable that equals 1 if $B_t > 0.9$ and zero otherwise. "Regular transactions share" is the fraction of regular transactions in all day-t transactions. We include dummies for the system-wide gas supply changes on December 10-12, 2017, January 3, and January 21, 2019. "ETH price (BoC)" is the daily average ETH price in terms of a basket of currencies consisting of US Dollars, Euro, and Chinese Yuan, weighted by the country/region specific number active nodes. *, **, *** denote 10%, 5%, and 1% significance level, respectively.

Table A3: Median gas price – Basket of currencies (BoC)

	median gas price, ETH		median gas price, BoC	
	(1)	(2)	(3)	(4)
blockchain utilization, B_t	0.792*	-0.582**	2.198***	0.585
	(0.480)	(0.268)	(0.699)	(0.519)
utilization > 90%, TR_t	-	24.38***	-	18.79***
	-	(3.817)	-	(5.026)
regular transactions share, R_t	3.955***	2.590***	9.353***	8.281***
	(0.752)	(0.587)	(0.734)	(0.835)
ETH price (BoC), X_t	-0.075	-0.028	-	-
	(0.129)	(0.111)	-	-
sample size	457	457	457	457
R-squared	0.371	0.580	0.676	0.727

Notes: OLS regressions with daily-level data including a constant. Newey-West standard errors reported in the parentheses. The dependent variable "median gas price" is the natural logarithm of the median gas price (in 10^{-8} ETH or 10^{-5} BoC, weighted basket of currencies consisting of USD, EUR, and CNY) in each block, averaged across all blocks created on date t. "Blockchain utilization" is the ratio between the total gas requirement for all transactions on day t and the total day-t gas supply. "Utilization > 90%" is the product of the blockchain utilization, B_t and a binary variable that equals 1 if $B_t > 0.9$ and zero otherwise. "Regular transactions share" is the fraction of regular transactions in all day-t transactions. We include dummies for the system-wide gas supply changes on December 10-12, 2017, January 3, and January 21, 2019. "ETH price (BoC)" is the daily average ETH price in terms of a basket of currencies consisting of US Dollars, Euro, and Chinese Yuan, weighted by the country/region specific number active nodes. *, **, *** denote 10%, 5%, and 1% significance level, respectively.

Table A4: Non-linear effect of blockchain utilization – quadratic model

	gas price					
	marginal		lowest 5-th percentile		median	
	ETH	USD	ETH	USD	ETH	USD
blockchain utilization (demeaned)	3.770***	5.273***	3.711***	5.255***	1.758***	3.633***
	(0.772)	(0.855)	(0.768)	(0.856)	(0.651)	(0.802)
blockchain utilization squared	17.86***	23.48***	17.45***	23.09***	11.47**	17.30***
	(4.627)	(4.888)	(4.581)	(4.861)	(3.969)	(4.497)
regular transactions share, R_t	3.828***	7.334***	3.966***	7.652***	3.546***	8.723***
	(0.932)	(0.747)	(0.916)	(0.749)	(0.725)	(0.764)
ETH price in USD, X_t	-0.732***	-	-0.687***	-	-0.110	-
	(0.208)	-	(0.200)	-	(0.160)	-
sample size	457	457	457	457	457	457
R-squared	0.389	0.714	0.404	0.721	0.427	0.712

OLS regressions with daily-level data including a constant. Newey-West standard errors reported in the parentheses. The dependent variable "marginal gas price" is the natural logarithm of the minimum observed gas price (in 10^{-8} ETH or 10^{-5} USD) over all included transactions in a block, averaged across all blocks created on day t. The dependent variable "lowest 5-th percentile" gas price is the natural logarithm of the bottom 5-th percentile of gas prices (in 10^{-8} ETH or 10^{-5} USD) in each block, averaged across all blocks created on day t. The dependent variable "median gas price" is the natural logarithm of the median gas price (in 10^{-8} ETH or 10^{-5} USD) in each block, averaged across all blocks recorded on day t. "Blockchain utilization (demeaned)" is the ratio between the total gas requirement for all transactions on day t and the total day-t gas supply, centered around its mean to avoid multicollinearity. "Regular transactions share" is the fraction of regular transactions in all day-t transactions. We include dummies for the system-wide gas supply changes on December 10-12, 2017, January 3, and January 21, 2019. "ETH price in USD" is the daily average ETH price in 10^3 USD. *, **, *** denote 10%, 5%, and 1% significance level, respectively.

Table A5: Gas price in levels instead of log

	marginal gas price, ETH (1)	(2)	marginal gas price, USD (3)	(4)
blockchain utilization, B_t	3.303***	0.376*	2.073***	0.241
	(0.950)	(0.216)	(0.526)	(0.164)
utilization > 90%, TR_t	-	51.03***	-	37.62***
	-	(7.838)	-	(6.859)
regular transactions share, R_t	4.550***	1.691*	5.266***	3.453***
	(1.507)	(0.997)	(1.356)	(0.676)
ETH price in USD, X_t	-0.605**	-0.486***	-	-
	(0.291)	(0.179)	-	-
sample size	457	457	457	457
R-squared	0.266	0.630	0.469	0.723

OLS regressions with daily-level data including a constant. Newey-West standard errors reported in the parentheses. The dependent variable "marginal gas price" is the minimum observed gas price (in 10^{-8} ETH or 10^{-5} USD) over all included transactions in a block, averaged across all blocks created on day t. "Blockchain utilization" is the ratio between the total gas requirement for all transactions on day t and the total day-t gas supply. "Utilization > 90%" is the product of blockchain utilization, B_t and a binary variable D_t that equals 1 if $B_t > 0.9$ and zero otherwise. D_t is also included in columns (2) and (4) separately. "Regular transactions share" is the fraction of regular transactions in all day-t transactions. We include dummies for the system-wide gas supply changes on December 10-12, 2017, January 3, and January 21, 2019. "ETH price in USD" is the daily average ETH price in 10^3 USD. *, **, *** denote 10%, 5%, and 1% significance level, respectively.

Table A6: Robustness – including gas supply

dependent variable: marginal gas price	OLS regression		IV regression	
	ETH	USD	ETH	USD
blockchain utilization, B_t	0.077	0.768	0.100	0.705
	(0.363)	(0.584)	(0.356)	(0.565)
utilization > 90%, TR_t	33.09***	28.342***	32.58***	27.305***
	(4.219)	(5.106)	(4.510)	(5.934)
regular transactions share, R_t	2.542***	6.671***	2.997**	7.146***
	(0.659)	(0.739)	(0.973)	(1.203)
gas supply, G_t	0.012	0.104	0.018	0.111
	(0.133)	(0.228)	(0.130)	(0.216)
ETH price in USD, X_t	-0.652***	-	-0.744***	-
	(0.159)	-	(0.190)	-
sample size	457	457	456	456
R-squared	0.595	0.760	0.594	0.759

OLS and IV regressions with daily-level data including a constant. Newey-West standard errors reported in the parentheses. The IV first stage regresses the regular transactions share on its first lag and the total number of new sender accounts together with other exogenous variables. The dependent variable "marginal gas price" is the natural logarithm of the minimum observed gas price (in 10^{-8} ETH or 10^{-5} USD) over all included transactions in a block, averaged across all blocks created on day t. "Lowest 5-th percentile" gas price is the natural logarithm of the bottom 5-th percentile of gas prices in each block, averaged across all blocks created on day t. "Median Gas Price" is the natural logarithm of the median gas price in each block, averaged across all blocks recorded on day t. "Blockchain utilization" is the ratio between the total gas requirement for all transactions on day t and the total day-t gas supply. "Utilization > 90%" is the product of the blockchain utilization, B_t and a binary variable that equals 1 if $B_t > 0.9$ and zero otherwise. D_t is also included in all columns separately. "Regular transactions share" is the fraction of regular transactions in all day-t transactions. "Gas supply" is the sum of the gas limits of all blocks created in day t. We include dummies for the system-wide gas supply changes on December 10-12, 2017, January 3, and January 21, 2019. "ETH price in USD" is the daily average ETH price in 10^3 USD. *, **, *** denote 10%, 5%, and 1% significance level, respectively.

Table A7: Robustness – using lagged ETH price in USD

dep. variable: marginal gas price, ETH	OLS regression	IV regression
blockchain utilization, B_t	0.086	0.114
	(0.307)	(0.303)
utilization > 90%, TR_t	32.89***	32.37***
	(4.041)	(4.367)
regular transactions share, R_t	2.488***	2.913***
	(0.667)	(0.981)
lagged ETH price in USD	-0.630***	-0.716***
	(0.157)	(0.188)
sample size	457	456
R-squared	0.593	0.592

OLS and IV regressions with daily-level data including a constant. Newey-West standard errors reported in the parentheses. The IV first stage regresses the regular transactions share on its first lag and the total number of new sender accounts together with other exogenous variables. The dependent variable "marginal gas price" is the natural logarithm of the minimum observed gas price (in 10^{-8} ETH) over all included transactions in a block, averaged across all blocks created on day t. "Blockchain utilization" is the ratio between the total gas requirement for all transactions on day t and the total day-t gas supply. "Utilization > 90%" is the product of the blockchain utilization, B_t and a binary variable that equals 1 if $B_t > 0.9$ and zero otherwise. D_t is also included in all columns separately. "Regular transactions share" is the fraction of regular transactions in all day-t transactions. We include dummies for the system-wide gas supply changes on December 10-12, 2017, January 3, and January 21, 2019. "lagged ETH price in USD" is the first lag of the daily average ETH price in 10^3 USD. *, **, *** denote 10%, 5%, and 1% significance level, respectively.

Table A8: Peak period (November 2017 - January 2018) – USD

	gas price in USD					
	marginal		lowest 5-th percentile		median	
	(1)	(2)	(3)	(4)	(5)	(6)
blockchain utilization, B_t	4.020***	2.802***	4.188***	3.143***	3.228***	3.372***
	(0.848)	(0.818)	(0.791)	(0.795)	(0.607)	(0.767)
utilization > 90%, TR_t	-	17.837***	-	16.623***	-	10.768**
	-	(3.789)	-	(3.746)	-	(4.708)
regular transactions share, R_t	5.335***	4.188***	5.253***	4.226***	5.173***	4.895***
	(1.547)	(0.951)	(1.472)	(0.934)	(1.316)	(1.102)
sample size	92	92	92	92	92	92
R-squared	0.833	0.898	0.852	0.904	0.832	0.852

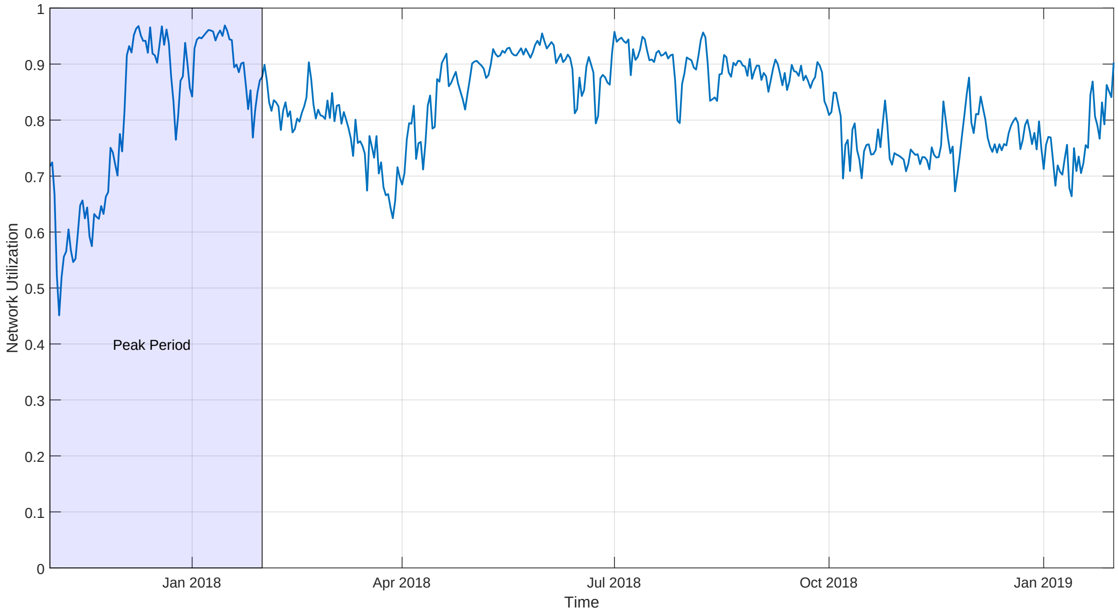
OLS regressions with daily-level data including a constant. Newey-West standard errors reported in the parentheses. The dependent variable "marginal gas price" is the natural logarithm of the minimum observed gas price (in 10^{-5} USD) over all included transactions in a block, averaged across all blocks created on day t. "Lowest 5-th percentile" gas price is the natural logarithm of the bottom 5-th percentile of gas prices in each block, averaged across all blocks created on day t. "Median Gas Price" is the natural logarithm of the median gas price in each block, averaged across all blocks recorded on day t. "Blockchain utilization" is the ratio between the total gas requirement for all transactions on day t and the total day-t gas supply. "Utilization > 90%" is the product of the blockchain utilization, B_t and a binary variable that equals 1 if $B_t > 0.9$ and zero otherwise. "Regular transactions share" is the fraction of regular transactions in all day-t transactions. We include dummies for the system-wide gas supply changes on December 10-12, 2017, January 3, and January 21, 2019. *, **, *** denote 10%, 5%, and 1% significance level, respectively.

Table A9: Peak period – Basket of currencies (BoC)

	gas price in ETH					
	marginal		lowest 5-th percentile		median	
	(1)	(2)	(3)	(4)	(5)	(6)
blockchain utilization, B_t	2.759**	0.501	2.895***	0.830	1.515**	0.636
	(1.127)	(0.741)	(1.043)	(0.695)	(0.676)	(0.584)
utilization > 90%, TR_t	-	17.38***	-	16.15***	-	9.610***
	-	(3.117)	-	(3.019)	-	(2.759)
regular transactions share, R_t	3.346**	1.124	3.177**	1.139	1.959*	1.038
	(1.439)	(0.928)	(1.356)	(0.860)	(1.032)	(0.771)
ETH price (BoC), X_t	-0.236	0.069	-0.199	0.078	0.279	0.381**
	(0.389)	(0.258)	(0.362)	(0.238)	(0.232)	(0.186)
sample size	92	92	92	92	92	92
R-squared	0.641	0.854	0.683	0.867	0.734	0.820

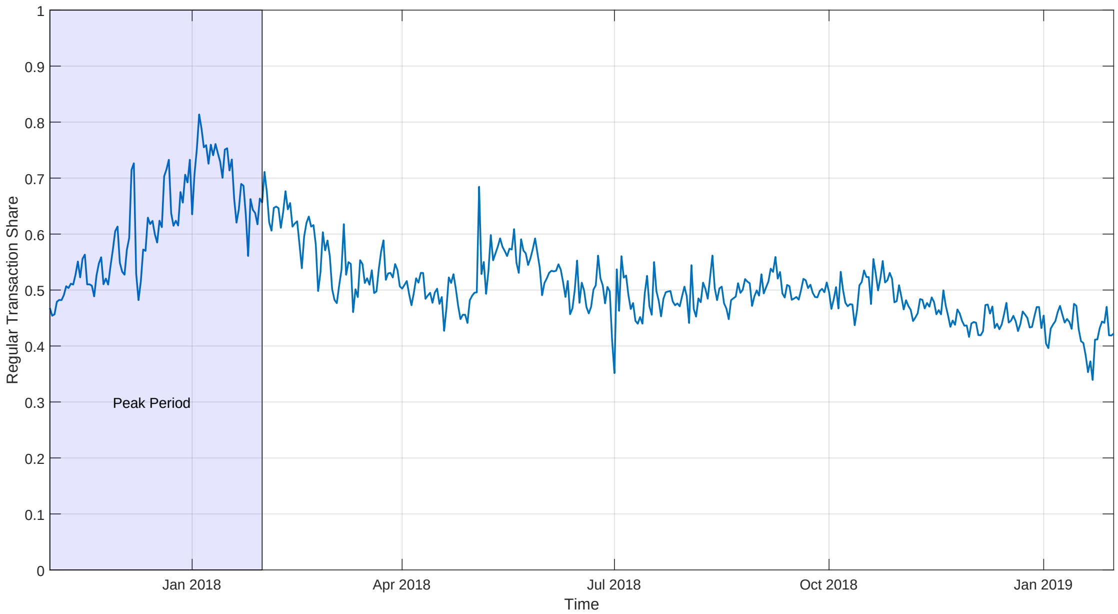
Notes: OLS regressions with daily-level data including a constant. Newey-West standard errors reported in the parentheses. The dependent variable "marginal gas price" is the natural logarithm of the minimum observed gas price (in 10^{-8} ETH) over all included transactions in a block, averaged across all blocks created on day t. The dependent variable "lowest 5-th percentile" gas price is the natural logarithm of the bottom 5-th percentile of gas prices (in 10^{-8} ETH) in each block, averaged across all blocks created on day t. The dependent variable "median gas price" is the natural logarithm of the median gas price (in 10^{-8} ETH) in each block, averaged across all blocks recorded on day t. "Blockchain utilization" is the ratio between the total gas requirement for all transactions on day t and the total day-t gas supply. "Utilization > 90%" is the product of the blockchain utilization, B_t and a binary variable that equals 1 if $B_t > 0.9$ and zero otherwise. "Regular transactions share" is the fraction of regular transactions in all day-t transactions. We include dummies for the system-wide gas supply changes on December 10-12, 2017, January 3, and January 21, 2019. "ETH price (BoC)" is the daily average ETH price in terms of a basket of currencies consisting of US Dollars, Euro, and Chinese Yuan, weighted by the country/region specific number active nodes. *, **, *** denote 10%, 5%, and 1% significance level, respectively.

Figure 2: Blockchain utilization



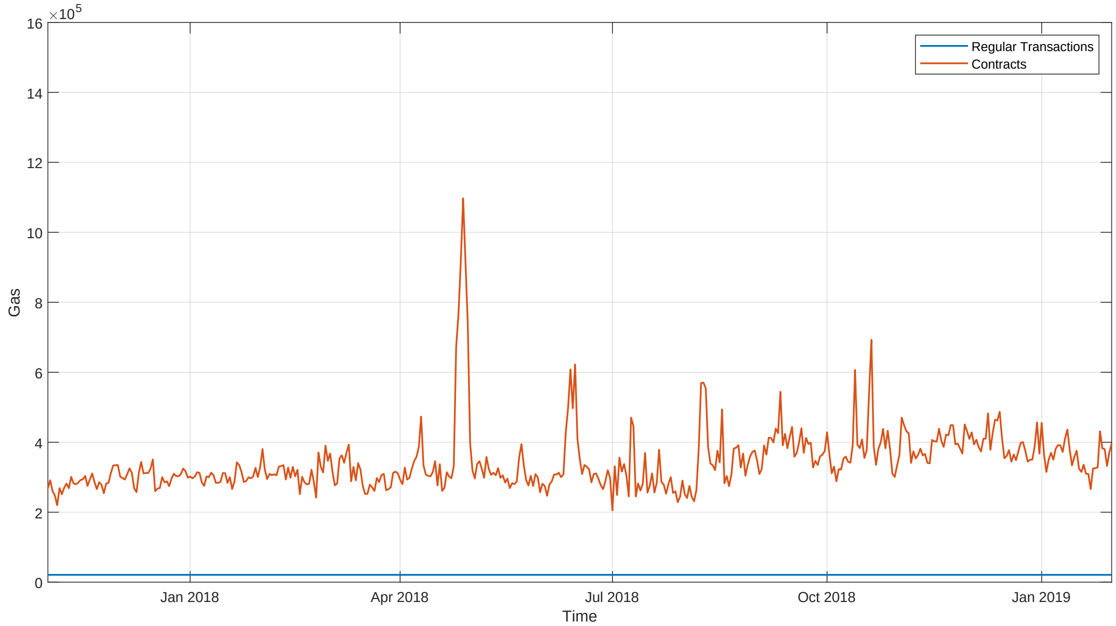
Notes: Blockchain utilization is the ratio between the total gas requirement for all recorded transactions in a given day and the total gas supply in the same day. The shaded region highlights the peak period between Nov 1, 2017 and Jan 31, 2018.

Figure 3: Regular transactions share



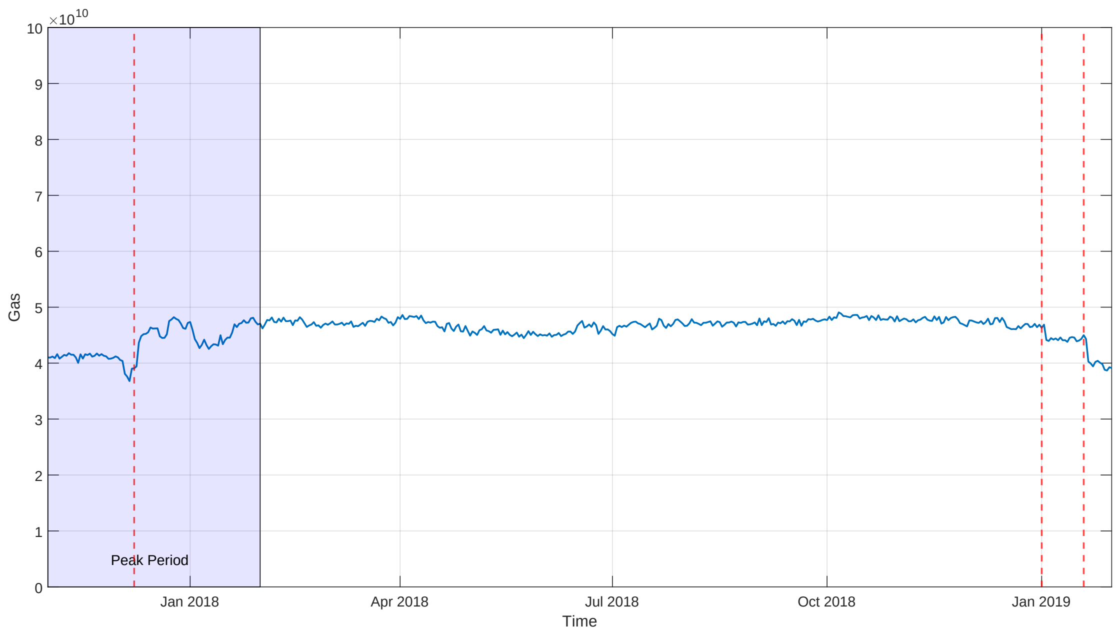
Notes: The regular transactions share is defined as the ratio between the number of regular transactions (direct ETH transfer between two addresses) and the number of all transactions recorded in a given day. The shaded region denotes the peak period between Nov 1, 2017 and Jan 31, 2018.

Figure 4: Required gas by transaction type



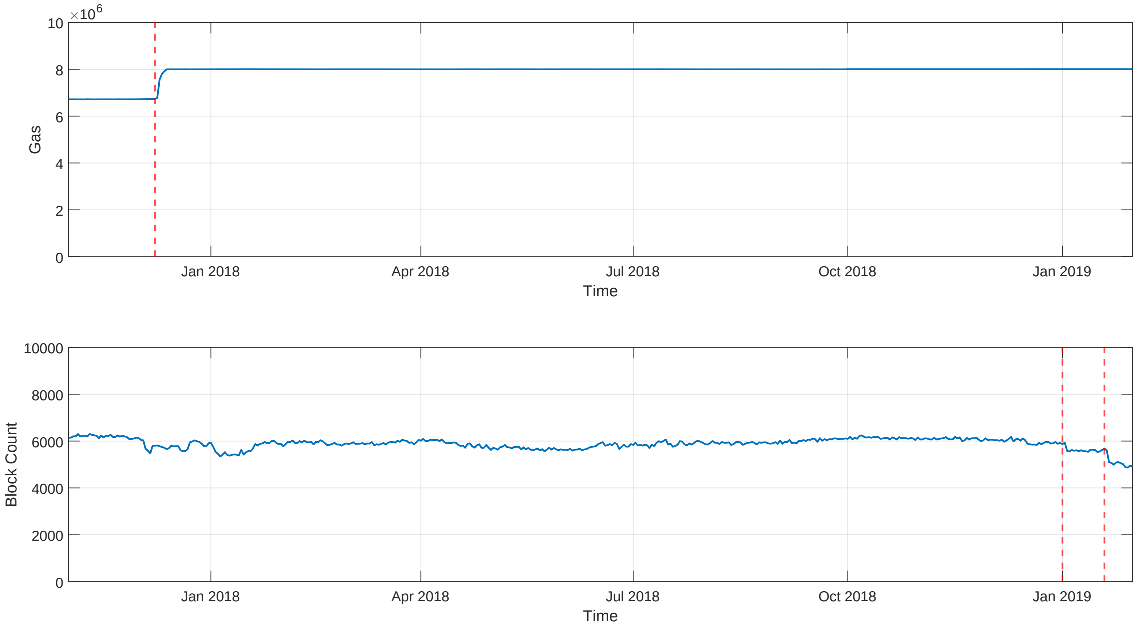
Notes: Daily average of the required gas amounts (gas requirements) by transaction type. The blue (red) line denotes regular transactions (contract calls and contract creations).

Figure 5: Gas supply



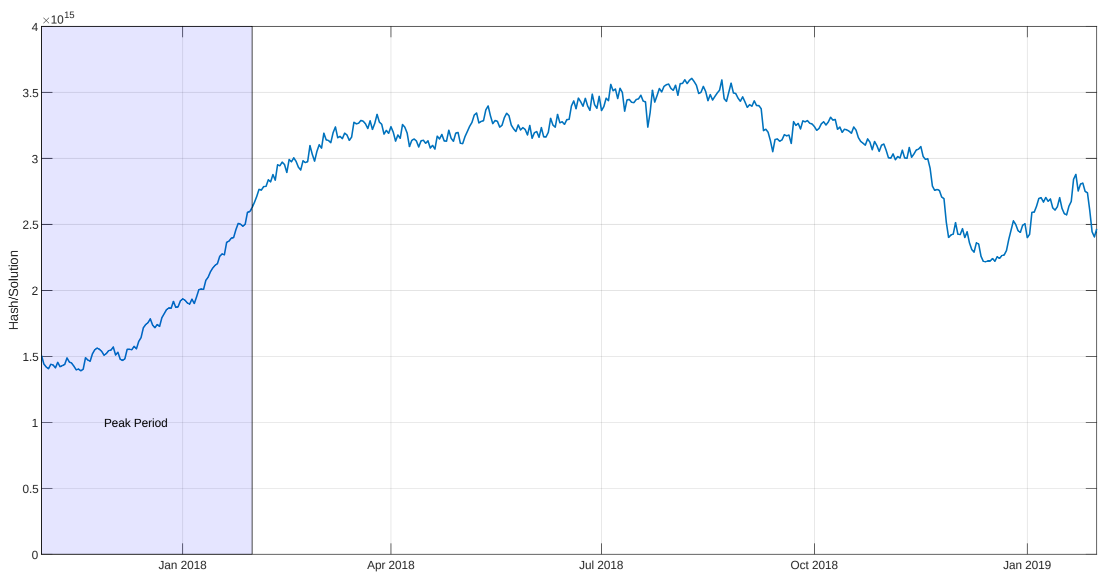
Notes: Gas supply is the sum of the gas limits of all blocks recorded on the blockchain in a given day. The shaded region denotes the peak period between Nov 1, 2017 and Jan 31, 2018. The vertical dashed lines denote exogenous system-wide changes in the gas supply, see also Fig. 6.

Figure 6: Average block gas limit and daily block count



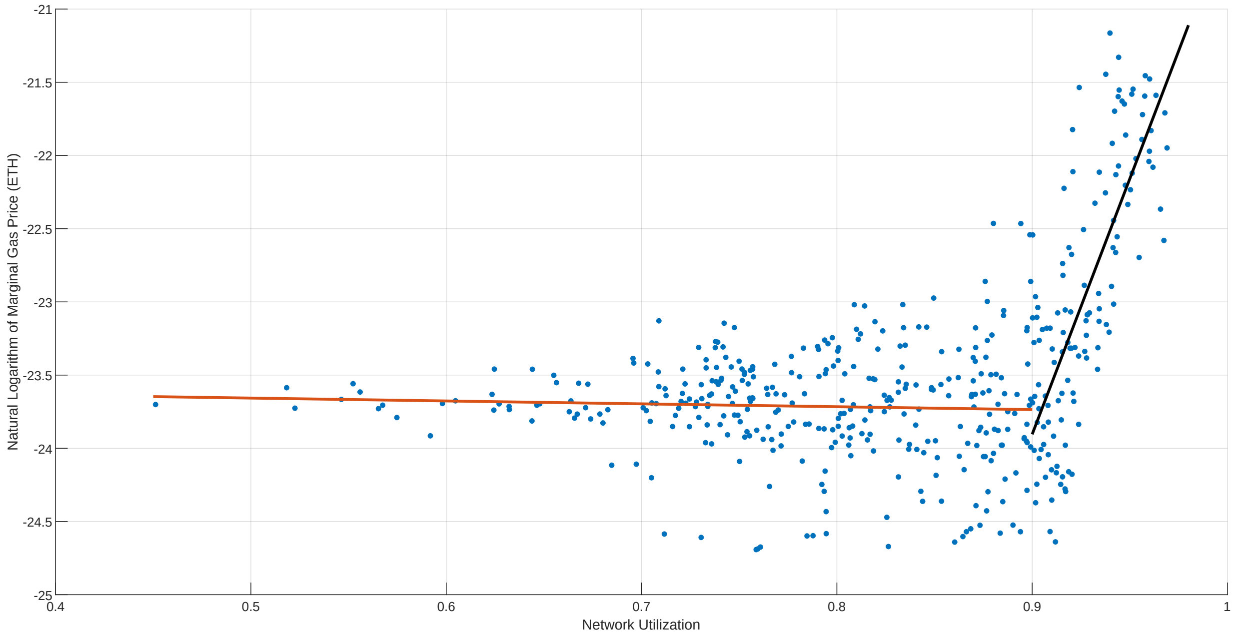
Notes: The average block gas limit (upper panel) is the daily average of the gas limits of all blocks recorded on a given day. The vertical dashed line highlights Dec. 10, 2017, when the Ethereum network initiated a system-wide increase of the block gas limit of approximately 18% that took three days to complete. The daily block count (lower panel) is the total number of blocks recorded on the blockchain in a given day. The vertical dashed lines denote Jan 3 and 21, 2019, the two dates on which the Ethereum protocol increased the cryptographic difficulty resulting in a decline in the number of blocks created per day.

Figure 7: Cryptographic difficulty



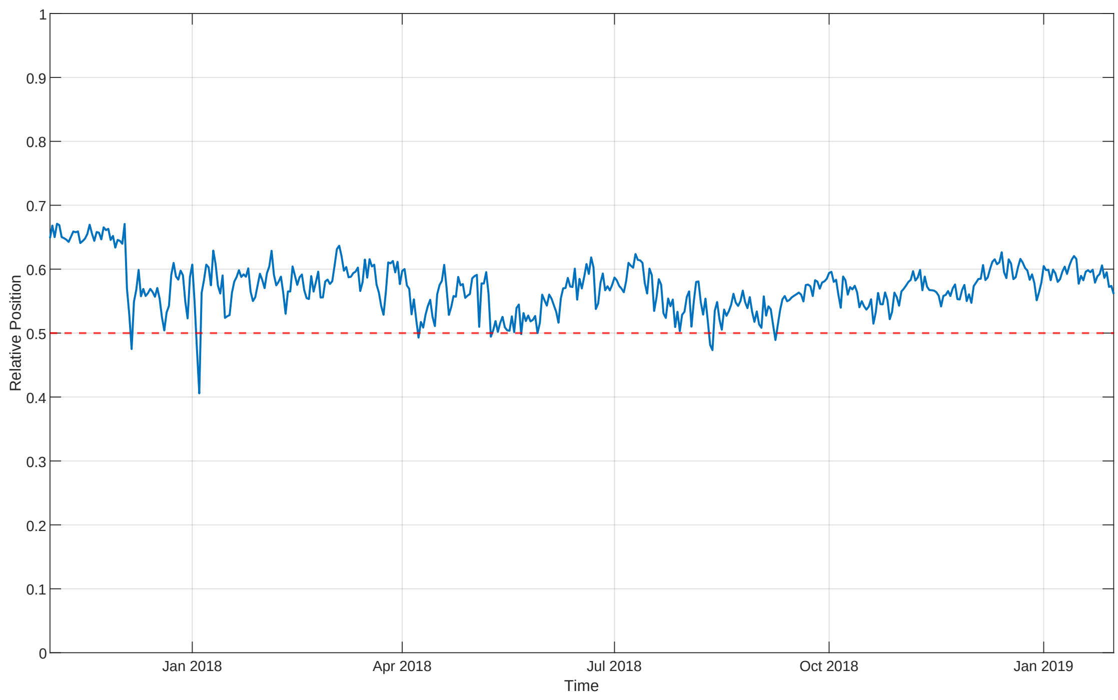
Notes: The figure plots the daily average cryptographic difficulty measured as the number of hash function evaluations per block. The peak period is shaded.

Figure 8: Network utilization and the marginal gas price



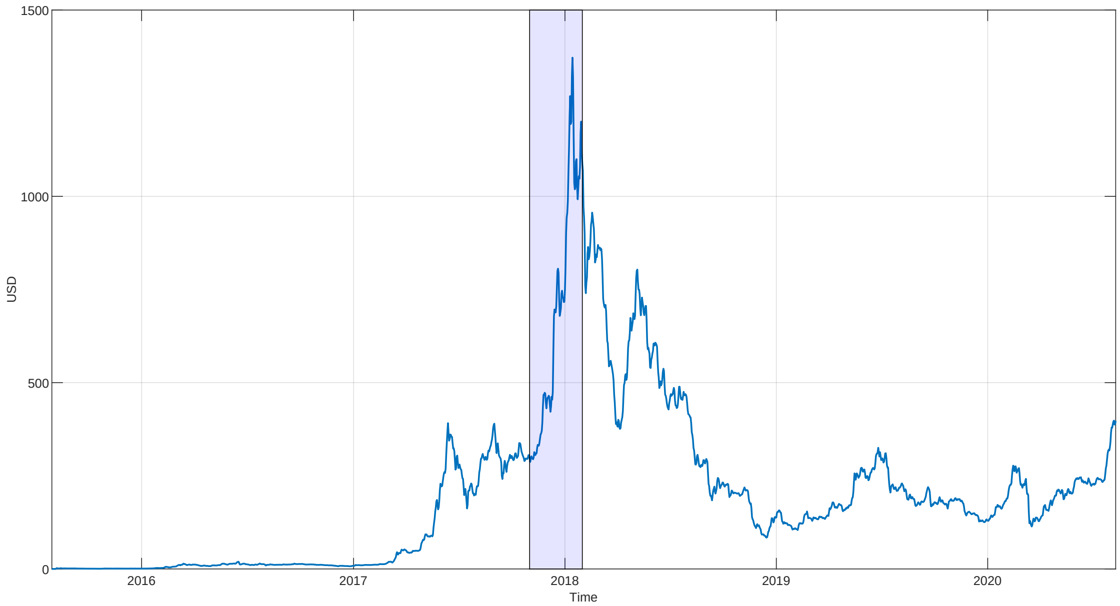
Notes: Scatterplot of blockchain utilization, B_t and log of the marginal gas price, p_t, with fitted lines.

Figure 9: Position of regular transactions within blocks



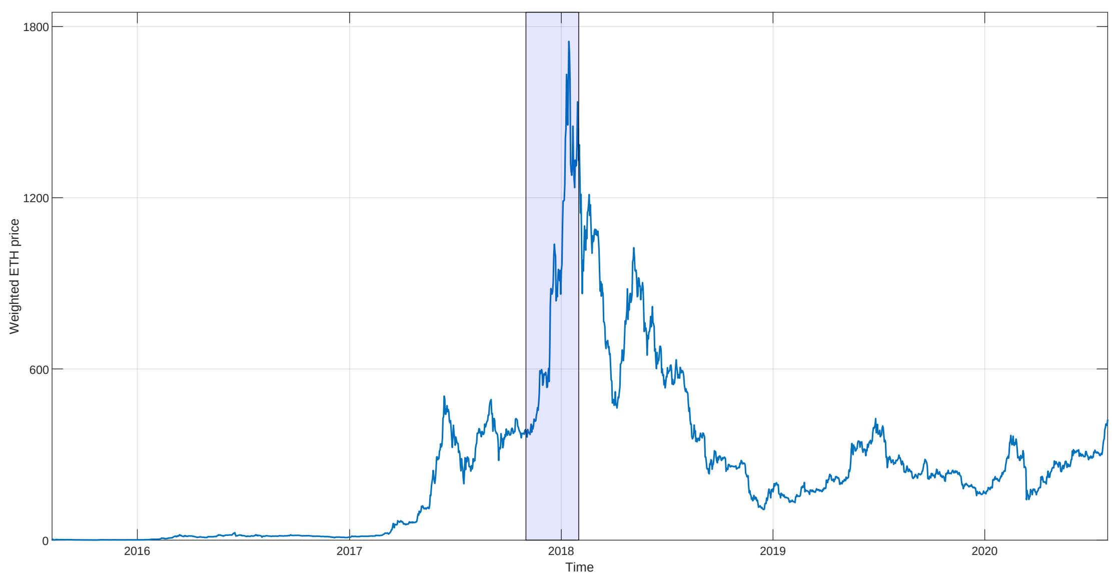
Notes: Regular transaction position (RTP) within blocks, daily average. RTP close to 1 (0) indicates that regular transactions are located close to the top (bottom) of the block and their associated gas prices are higher (lower). The dashed line indicates the mid-block position, RTP = 1/2.

Figure 10: Average ETH price in USD



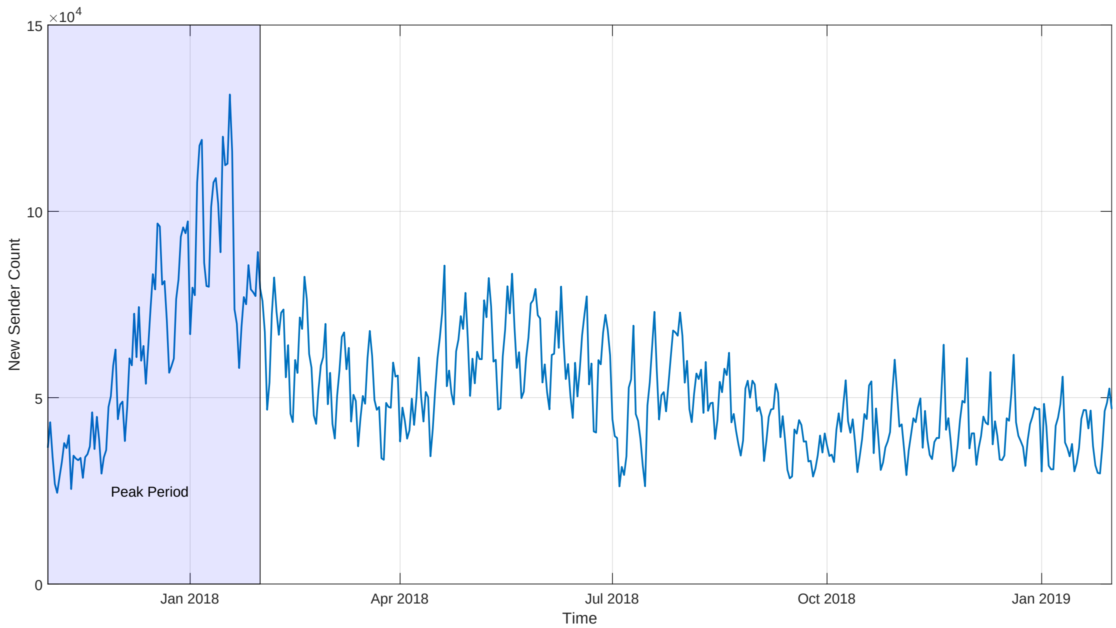
Notes: The Figure plots the daily average ETH price in USD. The shaded region highlights the peak period between Nov 1, 2017 and Jan 31, 2018.

Figure 11: Average ETH price in terms of USD, EUR and CNY basket of currencies (BoC)



Notes: The Figure plots the daily average weighted exchange rate between ETH and a basket of currencies (BoC) consisting of US dollars (USD), Euro (EUR), and Chinese Yuan (CNY) (see Appendix B). The shaded region highlights the peak period between Nov 1, 2017 and Jan 31, 2018.

Figure 12: New sender accounts



Notes: The Figure plots the daily number of new sender accounts, defined as the daily number of transactions with nonce value equal to 0. The nonce field in the transaction-level data records the past transaction count of a sender address. The shaded region highlights the peak period between Nov. 1, 2017 and Jan. 31, 2018.

B. Additional details

Gas supply events

In December 2017, the gas price increased significantly, associated with very high activity in the Ethereum platform. To compensate, the Ethereum protocol began increasing the block gas limit on December 10, 2017 (marked with a vertical dashed line) to allow higher number of transactions to be included in each block, see Figure 6. The upper panel of Figure 6 shows the daily average block gas limit. During a 3-day period, the block gas limit was increased by 18% boosting the network service rate. A second event affecting gas supply occurred in January 2019 when the Ethereum developers engaged in a planned increase in the cryptographic difficulty with the intention to switch the consensus algorithm from Proof-of-Work, PoW to Proof-of-Stake, PoS.[30] On January 3 and 21, 2019, the Ethereum blockchain protocol increased the cryptographic difficulty twice. The lower panel of Figure 6 displays the corresponding daily block count before and after the difficulty increases. Together, these difficulty increases lowered the daily gas supply by about 17%. We control for these system-wide gas supply events in the regression specifications by using dummy variables which equal 1 at the event dates and zero otherwise.

Transactions order within blocks

By design, each block of the Ethereum blockchain has a capacity limit measured in gas, called

[30] In a Proof-of-Stake system, there is no costly competition among miners and instead the block creator is chosen by an algorithm based on the user's "stake", or total ETH balance. In order to change the consensus mechanism from PoW to PoS, the blockchain developers algorithmically increase the cryptographic difficulty which, in theory, would make mining less profitable and provide incentives for the introduction of a Proof-of-Stake consensus mechanism. A switch to PoS has not yet occurred in Ethereum as of December 2020, though it remains planned.

the block *gas limit*. The sum of the gas requirements of all included transactions in a block cannot exceed the block gas limit. This implies that miners maximize their profits by ordering the submitted transactions by their gas price.[31] We checked all blocks in our data to verify whether the transactions are always sorted in descending gas price order. We confirmed that this is the case in 85% of the blocks in our data, with the rest of the blocks featuring only minor exceptions from descending gas price order. If a user sets a low gas price, her transaction request may not be written to the next block. Transaction requests that fail to be executed join the 'pending transactions' pool and wait to be recorded in later blocks. This means that Ethereum users face a trade-off between waiting costs and transaction fees and are competing to obtain higher priority among other waiting transactions in terms of the *gas price* and not in terms of the total transaction fee.[32]

Construction of basket of currencies and weighted ETH price

In Section 4 we argued that users may be concerned about the size of transaction fees in terms of their local currency, instead of or in addition to the fee's ETH value. To capture the possible effect of the ETH price in terms of conventional currencies on the gas price bidding choice of users, we create a basket of currencies consisting of the US dollar, the Euro, and the Chinese Yuan and define a weighted exchange rate between ETH and this basket. The selection of currencies and their associated weights was determined by the number of active Ethereum nodes (node refers to a computer running special software and being active part of the Ethereum network) during our study period.[33] We then compute the weighted exchange rate between the ETH cryptocurrency and the defined basket of currencies using the geometric average method (Takagi, 1986),

$$\text{ETH price}_t = \left[\frac{E_t^{\text{USD,ETH}}}{E_{base}^{\text{USD,ETH}}}\right]^{\text{W}^{\text{USD}}} \text{x} \left[\frac{E_t^{\text{EUR,ETH}}}{E_{base}^{\text{EUR,ETH}}}\right]^{\text{W}^{\text{EUR}}} \text{x} \left[\frac{E_t^{\text{CNY,ETH}}}{E_{base}^{\text{CNY,ETH}}}\right]^{\text{W}^{\text{CNY}}}$$

Above, ETH price_t denotes the weighted exchange rate between ETH and the basket of currencies at time t, $\text{E}_t^{\text{USD,ETH}}$, $\text{E}_t^{\text{EUR,ETH}}$, and $\text{E}_t^{\text{CNY,ETH}}$ are the prices of ETH in US Dollars, Euro, and Chinese Yuan at time t, respectively; $\text{E}_{base}^{\text{USD,ETH}}$, $\text{E}_{base}^{\text{EUR,ETH}}$, and $\text{E}_{base}^{\text{CNY,ETH}}$ are the prices of ETH in US Dollars, Euro, and Chinese Yuan in the base period, respectively; and the weights W^{USD}, W^{EUR}, and W^{CNY} equal the shares of active nodes running in the US, Euro-zone countries, and China, respectively. We chose August 7, 2015 as the base date, i.e., we set each ratio $\frac{E_t^{\text{c,ETH}}}{E_{base}^{\text{c,ETH}}}$ for $c = USD, EUR, CNY$ equal to 1 at $t = base$.[34]

Acknowledgments We thank C. Bidner, L. Petersen and CEA and seminar participants for useful comments and suggestions.

[31]The miners need to solve the Proof-of-Work cryptographic problem as quickly as possible to be selected as the successful creator of the next block and collect the associated transaction fees and block reward. The miners use algorithms which sort and select the transactions to include in the current block in descending gas price order.

[32]This process can also be thought as follows: the successful miner sells space in the block that they mined. Users are buyers who submit bids to purchase space in the block. Assuming that there are only N units of space in the current block and that each user needs one unit on average, then only the users submitting the N highest gas price bids will have their transactions recorded in the block.

[33]We use data from Kim et al. (2018) about the number of active nodes on the Ethereum network in 2018. We only consider locations with share of active nodes greater than 10%. This filtering retains the USA (43.2% share), China (12.9%) and the Euro-zone countries (11.7%). Scaling the shares to 100% yields weights of 63.7%, 19%, and 17.3% for the US, China, and the Euro-zone, respectively.

[34]August 7, 2015 is the first date on which ETH was valued positively against conventional currencies according to min-api.cryptocompare.com.