The Influence Factors on Ethereum Transaction Fees

Giuseppe Antonio Pierro, Henrique Rocha

▶ **To cite this version:**

The Influence Factors on Ethereum Transaction Fees

Guiseppe Antonio Pierro[*][†], Henrique Rocha[‡]
[*]University of Cagliari, Italy
[†]Inria Lille - Nord Europe, France
[‡]University of Antwerp, Belgium
antonio.pierro@gmail.com, henrique.rocha@gmail.com

Abstract—In Ethereum blockchain, the user needs to set a Gas price to get a transaction processed and approved by Miners. To have the transaction executed, the Gas price has to be greater than or equal to the lowest Ethereum transaction fees. This paper presents a set of data sampled every 15 seconds, from 1 December 2018 to 15 December 2018, coming from different blockchain web APIs. The aim of the paper is to investigate whether and to what extent different variables - such as the number of pending transactions, the value of the USD/Ether pair, average electricity prices around the world, and the number of miners - influence the Ethereum transaction fees. This study is relevant from an economic perspective because more and more companies in different economic fields are adopting Ethereum blockchain. From historical data analysis, we found that only some of these variables do have an influence. For example, the number of pending transactions and the number of miners have a major influence on Ethereum transaction fees when compared to the other variables.

Index Terms—Ethereum, Blockchain, Gas Price, Ethereum Transaction Fees, Granger causality.

I. Introduction

Ethereum [Eth14], [Eth16] is a platform based on a peer-to-peer network that supports an immutable transaction record on a public shared ledger known as blockchain. The blockchain technology promotes interactions among parties without the need of a central trusted authority [Woo18], [LCO+16], [BDLF+16], [JKS16], [O'C17], [BRDD18]. Nowadays, blockchains are not only used for monetary transactions but also for a broader application domain [LCO+16]. For instance, the Ethereum platform can execute Turing-complete programs called Smart Contracts [Eth18], [Eth16], [BRDD18].

Blockchains are made of *blocks*. A block is a container data structure and it is composed of a header and a list of transactions. Every transaction in Ethereum must pay a transaction fee in a special resource called *Gas* [Woo18], [Eth16]. Gas is "fuel" for computational instructions executed in the blockchain. The general idea is to make users pay for the computational costs (*e.g.,* energy, CPU) necessary to execute, create, and approve their transactions. Gas is bought by using Ether, a digital currency. A *Miner* approves a batch of transactions by adding them to the blockchain ledger and gets a reward as well as the transaction fees converted to Ether.

Since Ether is the second most valued cryptocurrency in the world[1] (the most valued is Bitcoin [Nak09]), the transaction fees have a non-trivial cost for the user.

For a Miner, these fees contribute to his/her profit. For example, the average revenue per block[2] (in September 2018) was 3.48 Ether ($785.58 USD[3]). From that amount, 0.48 Ether ($108.35 USD) was just in transaction fees. Every Miner can set his/her own minimum fees (measured in Gas price) and there is no consensus on what value should that be [Woo18].

In the period considered in this paper (from December 1, 2018 to December 15, 2018), the Ethereum transaction fees paid to have the transaction mined in approximately 30 seconds (i.e., two blocks in the Ethereum platform) noticeably changed from 60 GWei to 20 GWei[4]. Figure 1 shows the Gas prices estimated by the Etherchain's API[5] in 8 hours.

It is essential for a user to get a good estimate of the price s/he should set for his/her transaction to be approved in a given time. We assume that users do not want to overpay for Gas. Moreover, if a user does not need that his/her transaction is approved quickly, s/he may wait until it is possible to pay less and thus save money.

The variation of Ethereum transaction fees poses some questions worth considering, such as: what are the factors that influence the Ethereum transaction fees? What is the relationship between the Ethereum transaction fees and other economic indicators? Is there any connection between the Ethereum transaction fees and the number of pending transactions in Ethereum? Is there any connection between the Ethereum transaction fees and the Miners' policy, such as the minimum Gas price and the maximum Gas limit for mining a transaction?

The aim of the paper is to investigate the factors that influence the Ethereum transaction fees and therefore the possible decision making behaviour of blockchain users, miners included. The results are relevant not only from a computer science perspective but also from an economic perspective, because blockchain is a technology adopted by an increasing number of institutions, because it can coordinate strategic activities across different sectors: global agrifood chains, healthcare system, banks, insurance and entertainment companies, etc. Section II presents previous literature on the same research topic. Section III presents the main research question of the

---

[1]https://coinmarketcap.com/coins, 2018-09-28.

[2]https://bitinfocharts.com/ethereum/, 2018-09-28.
[3]We are assuming an exchange rate of 1 Ether = $225.742791 USD, based on the values by https://currencio.co/eth/usd/ at 2018-09-28.
[4]1 GWei = 10^{-9} Ether
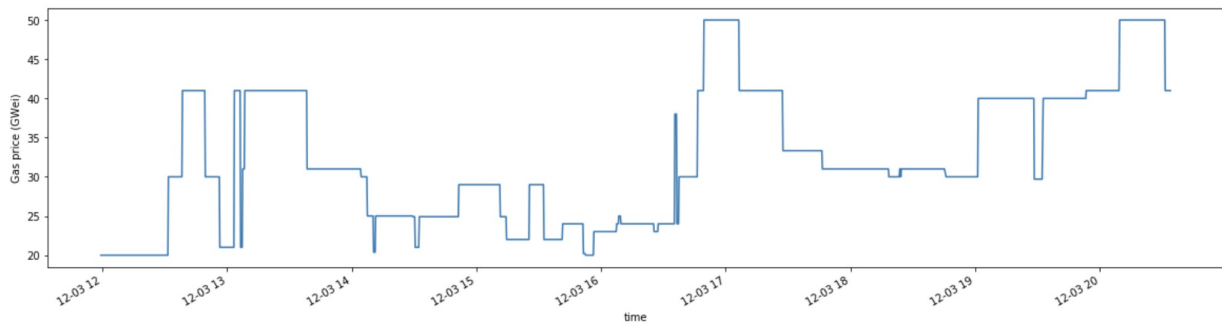[5]https://www.etherchain.org/api/gasPriceOracle

Fig. 1. Ethereum transaction fees variation

paper. Section IV describes some aspects of the Ethereum protocol and Granger causality, needed to understand the relationship among variables that might influence the Ethereum transaction fees. Section V presents the methodology used to gather and analyze the data from Ethereum, to evaluate the variables that affect the Ethereum Gas fees. Section VI presents and discusses the results. Section VII presents the conclusions of the research.

II. RELATED WORK

Several studies examined factors that influence cryptocurrencies prices and fees [Sov18], [GAH18], [Hou14], [MB15], [EOB17]. Sovbetov [Sov18] examines factors that influence the five cryptocurrencies Bitcoin, Ethereum, Dash, Litecoin, and Monero, over 2010-2018 using weekly data. Giudici and Abu-Hashish [GAH18] propose a new model that explains the dynamics of bitcoin prices and models the interconnections among different crypto and classic asset prices. Houy [Hou14] analyses the economics of Bitcoin's transaction fees and shows that a fixed and imposed transaction fee can keep Bitcoin blockchain secure enough when the transaction fee is high enough. Möser and Böhme [MB15] analyse the transaction fees paid within 45.7 million transactions recorded in the public Bitcoin blockchain from the inception of Bitcoin until the end of August 2014. They interpret the heterogeneity and instability of transaction fees as an indication that the protocol's market mechanism fails to set a fair price for transactions. Easley et al. [EOB17] develops a game-theoretic model to explain the factors leading to the emergence of transactions fees, as well as the strategic behaviour of miners and users. He highlights the role played by mining rewards and by transactions volume.

III. RESEARCH QUESTION

Previous studies focus on the factors that influence cryptocurrencies prices and fees in a daily, weekly or monthly time frame, while the present study considers a narrower time frame, in seconds. Moreover, previous literature especially considers the Bitcoin blockchain, while the present study investigates the Ethereum blockchain. Figure 1 shows that the Gas prices estimated by the Etherchain's API changes many times in just 8 hours ranging from 20 to 50 Gwei. Therefore,

we decided to analyse the variables in a 15 seconds time frame. We chose this interval of time, because it is the average time to mine a block in the ethereum network, as shown by the row block_time in Table III. We thought this is the best way to answer our main research question: what are the factors influencing the transaction fee price in the time frame to mine a block?

IV. BACKGROUND

A. Ethereum protocol

The Ethereum protocol [Woo18] defines how the Ethereum network works, how Miners should generally operate, and rules everyone must follow to be a valid part of the network. The protocol is written in general terms such that anyone could implement his/her own version of the protocol into a custom Ethereum client. The most used Ethereum clients are Go-Ethereum and Parity-Ethereum. Go-Ethereum, named Geth, is written in GO, Parity-Ethereum is written in Rust. Figure 2 is a pie chart showing the most used Ethereum clients[6].
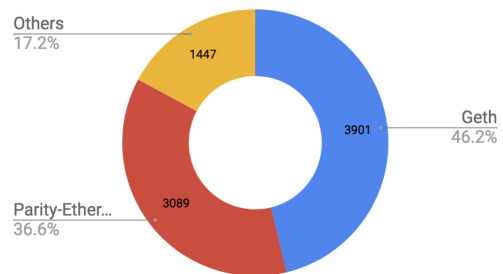
Fig. 2. Ethereum clients

Through the Ethereum clients the Miners can set the conditions that the transactions must satisfy to be accepted and transmitted in the network. For example the Miners can set the minimum Gas price to mine a transaction and the amount of Gas per block to target when mining a new block. Each Ethereum mining client has different default values to mine transactions and blocks. For example, the Geth Ethereum client has 2 GWei as the minimum Gas price and 4,712,388 as the maximum amount of Gas per block. The Miners can change

---

[6]https://www.ethernodes.org/network/1

the settings according to their needs. For example, to gain more money, a Miner can set a higher minimum Gas price or, to prevent malicious denial of service (DoS), a Miner can lower the block Gas limit to 2 million units of Gas.

B. Life cycle of an Ethereum transaction

Figure 3 presents the life cycle of Ethereum transactions. These are the main stages of the transactions workflow:

1) A user logs into his/her Ethereum account and sends his/her transactions to the Ethereum network, a set of interconnected nodes.
2) Some nodes receive the transactions and each one can pass them to nearby nodes.
3) The nodes that can mine the blocks, i.e. the Miners, select the transactions going to the mempool, according to its settings.
4) A miner picks the transactions up from the mempool, puts them in a block and tries to find the nonce value representing a correct solution to a cryptographic problem.
5) The first miner that finds a solution for its block, broadcasts the solution to all the other nodes.
6) The nodes that received the solution, verify whether it corresponds to the problem of the senders' block. If the solution is correct, the other nodes can confirm that the block can be added to the blockchain.
7) When the majority of the nodes reaches a consensus, the block is added to the blockchain.

C. Gas Price Oracle

The Etherchain Gas Price Oracle[7] is a tool that provides a prediction on the fairest Gas price to pay to get a transaction confirmed within a certain number of blocks. It uses the method developed by EthGasStation[8] to estimate the prices. We decided to use Etherchain instead of EthGasStation mainly because Etherchain provides its oracle data as a REST service, which is easier to acquire automatically. EthGasStation does not provide such interface requiring a manual or text-mining interaction to acquire its prices.

The Etherchain tool provides four recommended Gas prices based on the desired transaction speed and cost: "safe low", "average", "fast", and "fastest".

- "Safe low" is the Gas price intended to be both cheap and successful. It may take a bit longer to get a transaction confirmed with this price, but anyway less than 30 minutes.
- "Average" is the price accepted by the top Miners who account for at least 50% of the blocks. It takes around five minutes to get a transaction confirmed with this price.
- "Fast" is the price accepted by the top Miners which takes approximately one minute to get a transaction approved using this price.
- "Fastest" is the lowest Gas price that is accepted by all top Miners (estimated over the last two days). It takes

[7]https://www.etherchain.org/api/gasPriceOracle
[8]https://ethgasstation.info/

at most 30 seconds to get a transaction confirmed with this price. Paying more than this price, it is unlikely to decrease transaction confirmation time.

The paper will consider the "fastest" prices, as a variable named "oracle_gasprice" from now on. Figure 4 shows the "oracle_gasprice" variation history during the day. The "oracle_gasprice" recommendations are based on the lowest Gas price accepted by the Miners in the last 200 blocks. The goal of the paper is to understand whether other variables can affect the "oracle_gasprice" values, *i.e.,* the alledged fairest Ethereum transaction fees to pay.

D. Granger causality

To understand whether the data series on one variable affects the data series on the other variable, a specific relationship among the series needs to be observed. A time series variable is called causal to another if the ability to predict the second variable is improved by incorporating information about the first one. The notion of causality was first proposed by Wiener [Wie56]. Granger causality is a technique to determine whether one time series is useful in forecasting another. Granger [Gra69], [Gra81] defined causality as follows: A variable Y is causal for another variable X if knowledge of the past history of Y is useful to predict the future state of X in addition to the knowledge of the past history of X itself. So if the prediction of X is improved by including Y as a predictor, then Y is said to be Granger causal for X.

Granger causality between two variables can be unidirectional, bidirectional (or feedback) and neither unidirectional nor bidirectional, *i.e.,* independent or without Granger-causality in any direction. As to what concerns the results of this paper, the Granger causality test is calculated for different lags from 0 to 24 lags. Each lag corresponds to a time interval of 15 seconds. It means that, if the variable x Granger causes the variable y within one lag, it will need 15 seconds before the variable x affects the variable y.

E. Augmented Dickey-Fuller test

Granger causality test can be applied only to statistically stationary time series. A stationary time series is a series whose statistical properties, such as mean, variance, etc., are all constant over time. Most statistical forecasting methods, the Granger causality test included, are based on the assumption that the time series can be made approximately stationary (i.e., "stationarized") via mathematical transformations. If the time series are non-stationary, then the time series model should be applied to temporally differenced data rather than to the original data. Augmented Dickey-Fuller test (ADF) shows whether time series have some upward or downward trend or seasonal effects, i.e. whether mean or variance are not constant over time.

V. METHODOLOGY

The selection of variables taken into account in this study is based on different works related to factors that influence cryptocurrencies prices, such as Bitcoin, presented in
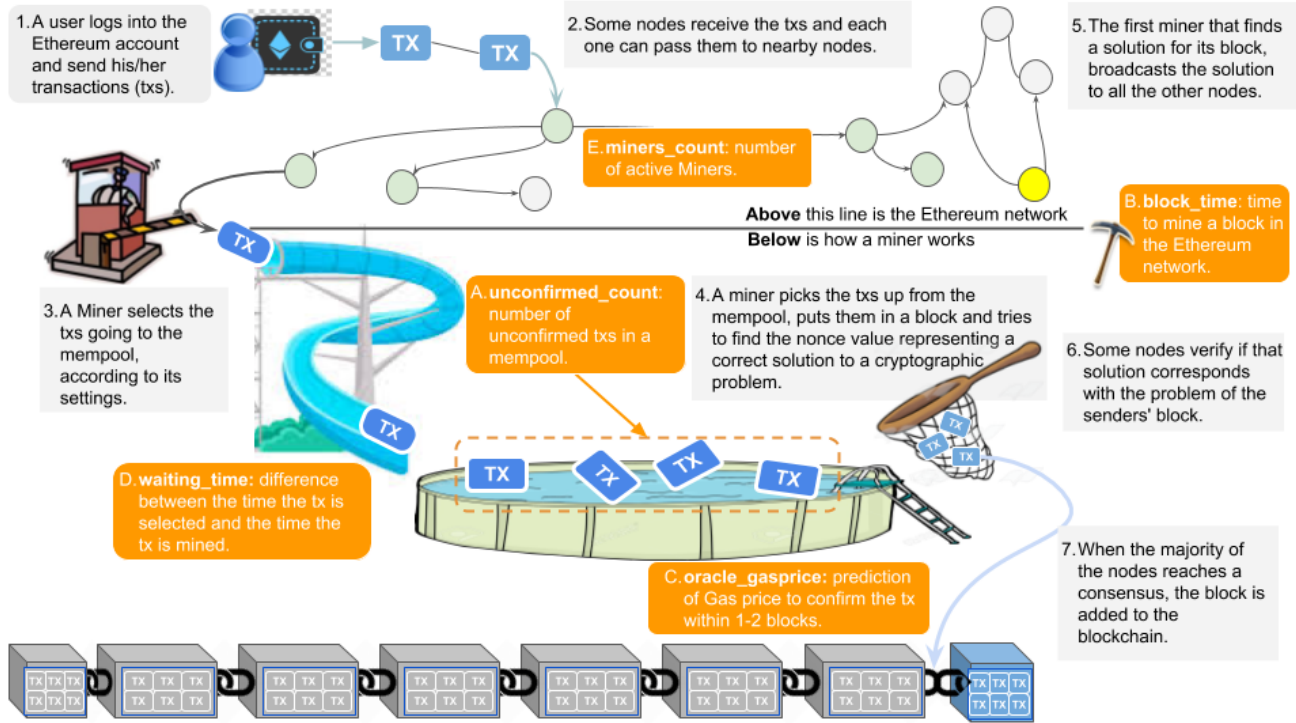
Fig. 3. Life cycle of an Ethereum transaction (tx). Orange boxes represent the variables possibly influencing the Ethereum txs fee. Grey numbered boxes represent the stages of the txs workflow.

[Sov18], [GAH18], [Hou14], [MB15], [EOB17] and listed in Table I. Data were analysed along the following variables: the Gas price oracle (oracle_gasprice), the number of unconfirmed transactions (unconfirmed_count), the block time (block_time), the number of active Miners (miners_count), the current hashrate of the network, a unit measured in hashes per second or H/s (hashrate), the current difficulty of the network (difficulty), the value of the USD/Ether ratio (eth_usd), the value of the BTC/Ether ratio (eth_btc).

- The variable oracle_gasprice, measured in Wei, is the Gas paid to have the transaction confirmed within 1 to 2 blocks time (around 15/30 seconds). In our case it is the Ethereum transaction fees.
- The variable hashrate it is the speed at which a miner solves the Ethereum code. In December 2018, the hash rate of the network was approximately 300 billion H/s or 300 GH/s.
- The variable unconfirmed_count refers to the number of unconfirmed transactions in a particular memory pool. The number of waiting transactions of a particular memory pool differ from the total number of waiting transactions in the Ethereum network. Moreover, the memory pool of each node might differ from the memory pool of other nodes: while there is a consensus on the mined transactions, there is no enforced consensus on what is stored in the memory pool of each node. However, for the aims of this paper, it is reasonable to assume that the waiting transactions trend in the memory pool is

representative of the general waiting transactions trend in the global network.
- The variable difficulty indicates how difficult it is to find the hash of a new block. The difficulty is adjusted periodically as a function of how much hashing power has been deployed by the network of miners.
- The variable block_time refers to the time spent to mine a block in the Ethereum network.

The data are sampled each 15 seconds and were sourced from the different RESTful services listed in Table II.

A. Dataset

For each variable of the dataset there are 92,160 observations, collected from December 1, 2018 to December 15, 2018 (Table II). The dataset is publicly available at Github.[9]

Table III shows statistics on our dataset. For each variable we measured the mean, the standard deviation (SD), minimum (min), the 25th, 50th, and 75th percentiles and maximum (max).

The data were analyzed to determine the Granger causality between a specific variable, *i.e.,* oracle_gasprice (the Gas price to have the transaction mined in 2 blocks max), and all the other variables of tab. III. A test was previously conducted to ensure that the data were stationary in terms of the variables used. In the case of non-stationary series, the first difference of the series was performed just once to make the series

---

[9]https://github.com/apierr/gas-price

Variable name	Description
oracle_gasprice	Gas paid to have the transaction confirmed within 1 to 2 blocks time
unconfirmed_count	Number of unconfirmed transactions in a particular memory pool
block_time	Time spent to mine a block in the Ethereum network
miners_count	Number of active Miners
hashrate (Hash/s)	Speed at which a miner solves the Ethereum code
difficulty	Number expressing 'how difficult' it is to find a new block
eth_btc	Value of the BTC/Ether ratio
eth_usd	Value of the USD/Ether ratio

TABLE II
RESTFUL SERVICES LIST

Variable name	Service Name	URI
oracle_gasprice	Etherchain's API	https://www.etherchain.org/api/gasPriceOracle
unconfirmed_count	BlockCypher's Ethereum API	https://api.blockcypher.com/v1/eth/main
block_time	Ethpool's API	https://api.ethpool.org/networkStats
miners_count		https://api.ethpool.org/poolStats
hashrate (H/s)		https://api.ethpool.org/poolStats
difficulty		https://api.ethpool.org/networkStats
eth_btc	Etherscan's API	https://api.etherscan.io/api?module=stats&action=ethprice
eth_usd		https://api.etherscan.io/api?module=stats&action=ethprice

TABLE III
STATISTICAL DESCRIPTION OF SAMPLE DATA

	Mean	SD	min	25%	50%	75%	max
oracle_gasprice	33.59	8.83	20	26	31	41	60
unconfirmed_count	91,768	6,312	70,772	88,429	92,915	95,582	111,252
block_time	14.46	0.91	11.70	13.8	14.4	15.1	17.5
miners_count	785	58.64	699	747	770	803	1,031
hashrate (TH/s)	173	4.15	167	169	175	177	181
difficulty ($\div 10^{15}$)	2.37	0.0718	2.22	2.32	2.37	2.44	2.55
eth_btc	0.0268	0.0008	0.0246	0.0263	0.0267	0.0277	0.0283
eth_usd	98.41	8.64	82.56	91.66	96.64	107.33	113.78

stationary. A pair Granger causality test was then performed for all the series.

VI. RESULTS AND DISCUSSION

When we analyzed the data, some series were stationary (ex. data of oracle_gasprice variable), whilst other series revealed to be non-stationary (ex. data of eth_usd variable). Figure 4 is a graphical representation of the stationarity vs. non stationarity of raw data concerning a sample of variables, *i.e.,* hashrate, difficulty, eth_usd, eth_btc, unconfirmed_count, miners_count, block_time, and oracle_gasprice.

In Figure 4, the first five (hashrate, difficulty, eth_usd, eth_btc, and unconfirmed_count) are non-stationary series, whilst the remaining represents stationary series (miners_count, block_time, and oracle_gasprice).

A. ADF test

To confirm whether the series were stationary or not, we perform a Augmented Dickey-Fuller (ADF) test. ADF test results show that the data series on variables hashrate, difficulty, eth_usd, eth_btc, and unconfirmed_count are all non-stationary. However they are all stationary after the first differentiation, therefore it is possible to state that they are integrated at the first order. The ADF test results are summarized in Table IV.

B. Granger causality test

The pair-wise Granger causality test was thereafter performed for oracle_gasprice variable versus the variables listed in Table V.
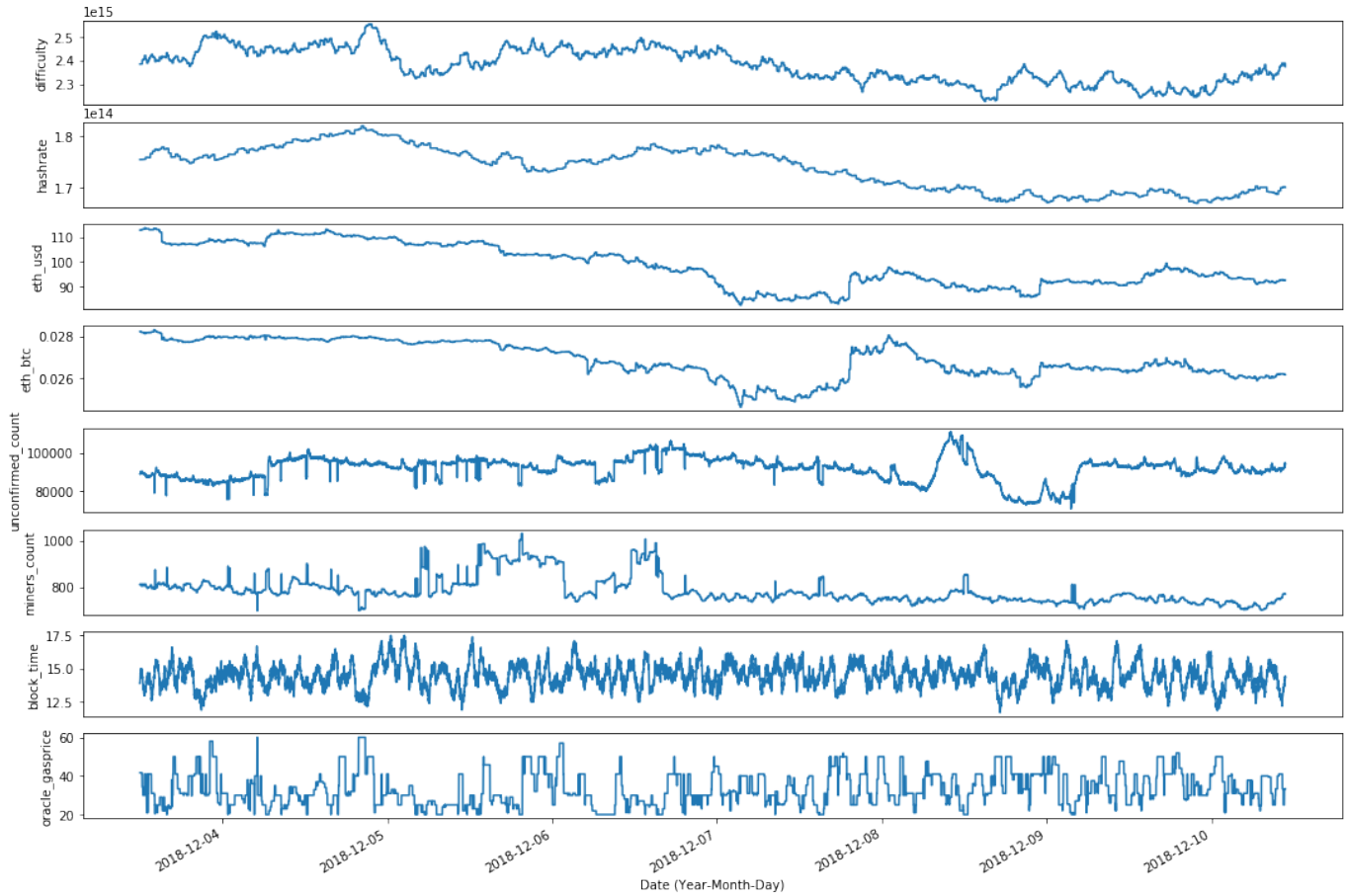
Fig. 4. Time series datasets

Variables	ADF Statistic	p-value	Conclusion
oracle_gasprice	-10.00	**0.000***	stationarity
unconfirmed_count	-2.958	0.058	non-stationarity
block_time	-10.92	**0.000***	stationarity
miners_count	-4.427	**0.001***	stationarity
hashrate	-0.672	0.853	non-stationarity
difficulty	-0.105	0.948	non-stationarity
eth_btc	-1.928	0.318	non-stationarity
eth_usd	-1.686	0.437	non-stationarity

*$p < 0.05$ means that the null hypothesis is rejected, indicating that the data are stationary

C. Pair-wise comparison: oracle_gasprice vs. hashrate, difficulty, block_time, eth_btc, and eth_usd

The results in Table V show that the data series on the variable hashrate does *not* Granger cause the data series on the variable oracle_gasprice, because the p-value (0.1576) is not significant. They also show that the data series on the variable

oracle_gasprice does *not* Granger cause the data series on the variable difficulty, because the p-value (1.000) is also not significant. To sum up, there is no Granger causality between the series, running from oracle_gasprice to difficulty and the other way, because difficulty does not affect oracle_gasprice and the converse is also true. Hence, the Granger causality is independent or non-directional between the two series.

Similar results were obtained for the hashrate variable. Indeed, when the hashrate increases, as a result, also the difficulty of validating newly added blocks increases; when the hashrate decreases, also the difficulty decreases, to speed up the time needed to mine a block. Based on the total hashrate, the difficulty is adjusted by the nodes and the time to mine a block is kept constant (15 seconds). This is the reason why these variables do not influence the transactions fees.

Finally, we obtained similar results for the relationship between oracle_gasprice and the following variables: eth_btc, eth_usd, and block_time.

D. Pair-wise comparison: oracle_gasprice vs. unconf_count

The results in Table V also show that the data series on the variable oracle_gasprice *does* Granger cause the data series on the variable unconfirmed_count, because the p-value (0.0215) is significant. They also show that the data series

TABLE V
GRANGER CAUSALITY TEST RESULTS

Null Hypothesis:	F-statistic	Prob.	Decision
unconfirmed_count does not Granger cause oracle_gasprice	2.6274	0.0723	Accepted
oracle_gasprice does not Granger cause unconfirmed_count	5.2830	**0.0215***	**Reject**
eth_usd does not Granger cause oracle_gasprice	0.6094	0.5437	Accepted
oracle_gasprice does not Granger cause eth_usd	0.1961	0.6579	Accepted
eth_btc does not Granger cause oracle_gasprice	0.5923	0.4415	Accepted
oracle_gasprice does not Granger cause eth_btc	1.5087	0.2193	Accepted
miners_count does not Granger cause oracle_gasprice	1.0331	**0.0309***	**Reject**
oracle_gasprice does not Granger cause miners_count	0.0035	0.9527	Accepted
difficulty does not Granger cause oracle_gasprice	1.2373	0.2499	Accepted
oracle_gasprice does not Granger cause difficulty	-32186	1.0000	Accepted
block_time does not Granger cause oracle_gasprice	1.8749	0.0323	Accepted
oracle_gasprice does not Granger cause block_time	0.3624	0.5472	Accepted
hashrate does not Granger cause oracle_gasprice	1.9968	0.1576	Accepted
oracle_gasprice does not Granger cause hashrate	-7.2403	1.0000	Accepted

*$p < 0.05$ means that the null hypothesis is rejected, indicating that the effect of the lagged values (value coming from an earlier point in time) of the other variable is statistically significant.

on the variable unconfirmed_count does not Granger cause the data series on the variable oracle_gasprice, because the p-value (0.0723) is not significant. To sum up, there is Granger causality between the series, running from oracle_gasprice to unconfirmed_count, but not the other way, because unconfirmed_count does not affect oracle_gasprice. Hence, the Granger causality is unidirectional between the two series.

E. Pearson correlation: unconf_count vs. oracle_gasprice

To better understand the unidirectional relationship of Granger causality between oracle_gasprice and uncon-firmed_count variables, a Pearson correlation test was performed. The Pearson correlation coefficient is equal to -0.6. Therefore the Pearson correlation test result suggests that there is an inverse relationship between oracle_gasprice and unconfirmed_count. A plausible explanation for this result is that, when the oracle suggests a high price, users are not encouraged to submit the transactions. It is indeed reasonable to assume that people that "can" wait, will do so if the Oracle price is too high.

F. Pair-wise comparison: oracle_gasprice vs. miners_count

Furthermore, the results in Table V show that the data series on the variable miners_count *does* Granger cause the data series on the variable oracle_gasprice, because the p-value (0.0309) is significant. They also show that the data series on the variable oracle_gasprice does *not* Granger cause the data series on the variable unconfirmed_count, because the p-value (0.9527) is not significant. To sum up, there is Granger causality between the series, running from miners_count to oracle_gasprice and not the other way, because oracle_gasprice does not affect miners_count and the converse is not true.

Hence, the Granger causality is unidirectional between the two series.

G. Pearson correlation: miner_count vs. oracle_gasprice

To better understand the unidirectional relationship of Granger causality between miner_count and oracle_gasprice variables, a Pearson correlation test was performed. The Pearson correlation coefficient is equal to -0.41. Therefore the Pearson correlation test result suggests that there is an inverse relationship between oracle_gasprice and miner_count: the more the number of miners, the lower the price predicted by the Oracle. A plausible explanation might come from the fact that an increase of the number of miners entails an increase of the mining competition to get the transactions. As a consequence of the supply/demand balance, the increase in competition entails a decrease of the prices.

VII. SUMMARY AND CONCLUSIONS

There is much work [Sov18], [GAH18], [Hou14], [MB15], [EOB17] investigating the factors influencing cryptocurrencies prices and fees, focusing on the most common cryptocurrencies, such as Bitcoin. Previous studies analyzed the factors on weekly or daily data. On the other hand, this paper analyzes instead Ethereum and adopts a finer time frame, to investigate the factors that influence the Ethereum transaction fees in the average time to mine one block (i.e., approximately 15 seconds).

The main objective of the paper is to analyze the Granger causality relationship between the data series on the variable oracle_gasprice and other variables, such as unconfirmed_count, miners_count and eth_usd. A ADF test and a pair-wise Granger causality test were performed to establish whether there is

a Granger causality between the data series on the variable oracle_gasprice and the data series on other variables.

The results of the ADF test showed that the data series on eth_usd and unconfirmed_count did not present any stationarity. After taking the first difference of the series, the results of the ADF test showed a stationarity.

In light of the results of the pair-wise Granger causality test (see Table V), a non-directional causality relationship was observed between the data series on the oracle_gasprice variable and the data series on the block_time variable and between the data series on the oracle_gasprice variable and the data series on the hashrate variable. This means that the past history of both the oracle_gasprice and the block_time variables cannot help in respectively predicting their future values.

In a similar vein, it is possible to conclude that ethtbc, eth_usd, hashrate and difficulty variables cannot be used to forecast the values of the oracle_gasprice variable, and also the converse is true (see Table V).

Interestingly, a unidirectional causality was observed from the data series on the oracle_gasprice variable to the data series on the unconfirmed_count variable (see Table V). This result shows that the past history of the oracle_gasprice variable is useful to forecast the number of waiting transactions, even though the converse is not true. The results of the Pearson correlation test showed that they are inversely correlated: when the oracle price increases, the number of waiting transactions in the Ethereum network decreases. It stands to reason that when the oracle suggests a high price to pay, the users wait to submit a transaction, thus decreasing the overall number of pending transactions in their memory pools. As to what concerns the results of this paper, the oracle_gasprice variable Granger causes the unconfirmed_count variable when the number of lags is greater than 6. This means that the oracle_gasprice variable does not immediately affect the unconfirmed_count variable. This result is compatible with the fact that the user cannot be immediately aware of the variation of the oracle_gasprice. Moreover we need to consider the time taken by the user to submit the smart contract to the blockchain and the time required by the transactions to propagate in the network. When choosing a time frame greater than 15 seconds, it is not possible to appreciate the time needed for the system to equilibrate the variation of a variable.

Finally, a unidirectional causality was also found from the data series on the miners_count variable to the data series on the oracle_gasprice variable (see Table V). The result of the Pearson correlation test showed that the number of Miners and the oracle_gasprice variables are inversely correlated: when the number of Miners increases, the oracle price decreases, as per the supply/demand balance.

Overall, the results of the paper are useful to improve the predictions on the Ethereum transaction fees at a given time, because they shed a light on how different variables might interact and influence the Oracle Gas Price. Knowing that Oracle predictions are biased by some variables might be useful for blockchain users, to reach a more mindful and efficient use of the platform. Such results are relevant not only from a computer science perspective but also from an economic perspective, because they show the financial mechanisms of blockchain which is adopted by both public and private institutions. Further research is anyway needed to build a model that precisely provides the users with an estimate of the best Gas Price to pay to have the transaction executed in a given time lapse, taking into account the variables that influence the overall Ethereum transaction fees.

References

[BDLF+16] Karthikeyan Bhargavan, Antoine Delignat-Lavaud, Cédric Fournet, Anitha Gollamudi, Georges Gonthier, Nadim Kobeissi, Natalia Kulatova, Aseem Rastogi, Thomas Sibut-Pinote, Nikhil Swamy, and Santiago Zanella-Béguelin. Formal verification of smart contracts: Short paper. In *2016 ACM Workshop on Programming Languages and Analysis for Security*, PLAS '16, pages 91–96, New York, NY, USA, 2016. ACM.

[BRDD18] S. Bragagnolo, H. Rocha, M. Denker, and S. Ducasse. Smartinspect: solidity smart contract inspector. In *2018 International Workshop on Blockchain Oriented Software Engineering (IWBOSE)*, pages 9–18, mar 2018. Electronic ISBN: 978-1-5386-5986-1.

[EOB17] David Easley, Maureen O'Hara, and Soumya Basu. From mining to markets: the evolution of bitcoin transaction fees. *SSRN Electronic Journal*, pages 1–55, jan 2017.

[Eth14] Ethereum Foundation. Ethereum's white paper., 2014.

[Eth16] Ethereum Community. Ethereum homestead documentation, 2016.

[Eth18] Ethereum Foundation. Solidity documentation release 0.4.24, 2018.

[GAH18] Paolo Giudici and Iman Abu-Hashish. What determines bitcoin exchange prices? a network var approach. *Finance Research Letters*, 2018.

[Gra69] C W J Granger. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*, 37(3):424–438, jul 1969.

[Gra81] C W J Granger. Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, 16(1):121–130, may 1981.

[Hou14] Nicolas Houy. The economics of bitcoin transaction fees. Technical report, feb 2014. Working paper GATE.

[JKS16] Ari Juels, Ahmed Kosba, and Elaine Shi. The ring of gyges: Investigating the future of criminal smart contracts. In *2016 ACM SIGSAC Conference on Computer and Communications Security*, CCS '16, pages 283–295, New York, NY, USA, 2016. ACM.

[LCO+16] Loi Luu, Duc-Hiep Chu, Hrishi Olickel, Prateek Saxena, and Aquinas Hobor. Making smart contracts smarter. In *CCS'2016 (ACM Conference on Computer and Communications Security)*, 2016.

[MB15] Malte Möser and Rainer Böhme. Trends, tips, tolls: A longitudinal study of bitcoin transaction fees. In *Financial Cryptography Workshops*, pages 19–33, jan 2015.

[Nak09] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system., 2009.

[O'C17] Russell O'Connor. Simplicity: A new language for blockchains. In *2017 Workshop on Programming Languages and Analysis for Security*, PLAS '17, pages 107–120, New York, NY, USA, 2017. ACM.

[Sov18] Yhlas Sovbetov. Factors influencing cryptocurrency prices: Evidence from bitcoin, ethereum, dash, litcoin, and monero. *Journal of Economics and Financial Analysis*, 2:1–27, feb 2018.

[Wie56] N. Wiener. The theory of prediction. In *Modern mathematics for engineers, Series I*. Beckenham, E. F., 1956.

[Woo18] Gavin Wood. Ethereum: A secure decentralised generalised transaction ledger. *Ethereum Yellow Paper. Byzantium Version e94ebda*, pages 1–39, jun 2018.