ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics
Financial Economics

TRANSCATION FEE MODEL IN THE BITCOIN NETWORK

Anton Braun
492219
Supervisor: Eisert, T.
Second assessor:

Abstract

This research explores a potential model of bitcoin transaction fees. The model attempts to link price movement of transaction fees to congestion in the blockchain infrastructure and is tested in an empirical setting. Insight from the model is used to construct a one-day ahead forecast through an autoregressive distributed lag model. To find the optimal lag length for each independent variable from the initial congestion model, a rolling window approach is applied in combination with a mean squared error as the selection criterion.

Table of contents

1. Introduction 2
    - 1.1. Research objective 2
    - 1.2. Academic and scientific relevance 3
    - 1.3. Research question and hypotheses 3
2. Literature review 4
3. Conceptual framework 6
    - 3.1. Transaction in bitcoin network 6
    - 3.2. Transaction fees in bitcoin network 9
4. Research and methodology 10
    - 4.1. Data 10
    - 4.2. Historical transaction Fee development 11
    - 4.3. Data transformation 14
    - 4.4. Transaction fee model specification and empirical analysis 14
    - 4.5. ADL forecast model specification 17
    - 4.6. ADL forecast model analysis 20
5. Conclusion 22
6. Appendix 24
7. Bibliography 26

1. Introduction

Cryptocurrency is a new financial asset that has developed from an open-source code, not taken serious to a financial asset gaining acceptance by the mainstream financial markets. However, cryptocurrency is still a young and volatile asset, characterized by intense price speculation. As such, the media focus has been put on price development and modern versions of the American Dream from racks to riches. However, a topic that has received less interest by the mainstream is the development of blockchain based currencies into every day online payment methods. As such, the ECB report on a digital euro (2020) proposes to utilize the blockchain technology for issuance and trading purposes. Although, the report highlights that the digital euro will not be a crypto-asset, the statement provides incentive to further investigate fundamentals underlying the blockchain based currencies.

One of those fundamentals in the infrastructure in a blockchain based currency is the presence of transaction fees. Considering the potential start of a new stage in the blockchain development, the entrance into government-controlled currencies, all variables around existing blockchain based currencies will have to be evaluated, to propose a system that combines government controlled and issued currencies with the deregulated peer-to-peer blockchain nature. In the following, an investigation of transaction fees in the blockchain infrastructure is conducted based on the bitcoin, as the largest current blockchain based asset type.

1.1. Research objective

This thesis is aimed at diagnosing transaction fees in cryptocurrencies in the market. The market is proxied by the bitcoin network and to obtain comparable insight, the transaction fees are broken down to a per byte scale. More specifically, this study attempts to construct a model for transaction fees constituting of indicators taken from the bitcoin network and the bitcoin price. Additionally, the second objective is to assess whether the insight of the model can be used to forecast transaction fees in a on day ahead time interval, through an autoregressive distributed lag model. Both objectives attempt to provide insight into the transaction fees of cryptocurrencies in the market, useful for introducing a government affiliated blockchain based currency.

1.2. Academic and social relevance

The research will add to the increasing numbers of papers investigating transaction fees in cryptocurrency networks. This research will extend the analysis of equilibrium transaction fees and game theoretical focus (Huberman, Leshno, and Moallemi, 2017; Easley, O'Hara, and Basu 2019) to utilizing insight in developing a one day ahead forecast model. Thereby, this research adds to the current stock of knowledge on cryptocurrency transaction fees by attempting to provide market insight on cryptocurrency transaction fees for a potential government issued currency based on the blockchain infrastructure. Further, a bridge into practical appliance is provided by the suggested model to forecast transaction fees, helpful to predict and smoothen out fluctuations of transaction fees.

While the last primarily is intended to provide governments with a tool to understand and predict transaction fees in a cryptocurrency network, it holds similar relevance for financial markets and companies involved in cryptocurrencies and potentially users of new type of FIAT currency based on the blockchain. Following current trends of cryptocurrencies and other blockchain based appliances becoming more mainstream, it is beneficial for society to participate in markets that are capable of potentially preventing friction from high and unexpected transaction fees.

1.3. Research question and hypotheses

The market only relatively recently has been introduced to cryptocurrencies through its first representation in form of the bitcoin, a decentralized, peer-to-peer based online payment method (Nakamoto, 2008). Regardless, academic research has been since investigating the bitcoin. Initially, academic research has been focused on discussing whether bitcoin is in fact a real currency (Yermack, 2013; Maurer, Nelms, and Schwartz, 2013) and bitcoins and cryptocurrencies legality concerns and design issues (Grinberg, 2012; King, & Nadal, 2012; Eyal, & Sirer, 2014). Furthermore, methods and techniques have been analyzed to improve bitcoin as a currency (Barber, Boyen, Shi, and Uzun, 2012). With the development of bitcoin to a mainstream financial asset accepted by institutional investors (Reuters, 2021), research extended to components of the bitcoin such as transaction fees, that are investigated in different stages of equilibriums through game theory (Huberman, et al., 2017; Easley et. al., 2019). Their research on equilibrium transaction fees provides incentives to investigate anomalies in transaction fees and gives rise to the following question;

Are there driving factors of transaction fees in the bitcoin network and can fluctuations in fees be forecasted?

To test and provide a possible answer to the question two hypotheses are formulated and further investigated. The first hypothesis constructed is:

(1) Are there factors that drive transaction fees in the bitcoin network?

The second hypothesis is constructed with the recent spike in bitcoin transaction fees in April 2021 in mind. The spike resulted in the largest transaction fee yet present in bitcoin network. The hypothesis follows the bitcoin leading up to the event including the spike in transaction fees in 2017 and states:

(2) Can a forecast model for transaction fees predict an exogenous spike in bitcoin transaction fees?

The accumulated answer to both hypotheses is later discussed and based on the obtained insight a potential answer of the research question formulated. In the remaining, first the literature preceding and related to this thesis will be reviewed in Section 2. Thereafter, in Section 3 the conceptual framework will be presented leading up to the research methodology in Section 4. Lastly, Section 5 concludes this research.

2. Literature review

The research is part of a growing academic interest into digital decentralized currencies based on the blockchain. Pioneered by the introduction and publication of the open-source program for bitcoin based on the blockchain through Nakamoto Satoshi, an unknown individual or group in 2008. Specifically, this research joins recent investigation on transaction fees in the bitcoin network. Huberman et al. (2017) strive to answer the question how the bitcoin system raises revenue and pays for its infrastructure. Specifically, they discuss transaction fees in the presence of diminishing mining rewards. The mining process, essential for the functioning of the bitcoin network and further discussed under 3.1. is partially incentivized through the mining reward. This reward is diminishing under a predetermined rate by the bitcoin protocol. Huberman et al. (2017) express concern on the long-term viability of the mining process and hence the bitcoin infrastructure as a whole, if equilibrium transaction fees are too low. They determine equilibrium fees in a congestion queueing game. They determine congestion in transaction processing as an essential factor to raise revenue to maintain the infrastructure and thereby a key driver of transaction fees. Following the findings of Huberman et al. (2017), congestion and its relationship with factors in the bitcoin network are among other indicators later adopted in this research to potentially constitute and predict changes in transaction fees.

Esaley et al. (2019) conduct a similar equilibrium investigation of transaction fees. However, they examine the bitcoin system, while the mining reward is still present, and has not reached zero or neglectable levels. Therefore, instead of discussing the prevalence of bitcoin in a long run steady state, they discuss whether such a steady state is feasible to reach in the first place. In their research, they determine waiting time as the key driver of transaction fees. Further, they are concerned that waiting times and transaction fees resulting of congestion are too high in equilibrium and thereby discourage participants to utilize bitcoin as a payment system. A different assumption compared to Huberman at al. (2017).

This research does not investigate a steady state bitcoin network and therefore follows the assumptions of Esaley et al. (2019), that increasing congestion in the bitcoin network is a key driver of transaction fees. The following investigation however differs from Esaley et al. (2019), considering that absolute transaction fees are broken down to a per byte scale are investigated compared to the percentage of zero cost transaction fees in Esaley et al. (2019). With the focus of this research being to model current and near future levels of transaction fees, mining rewards are proposed to influence transaction fees, further elaborated, and empirically tested under 4.4.

This is different compared to Huberman et al. (2017) and Esaley et al. (2019), finding that transaction fees are determined in equilibrium regardless and independent of the mining reward.

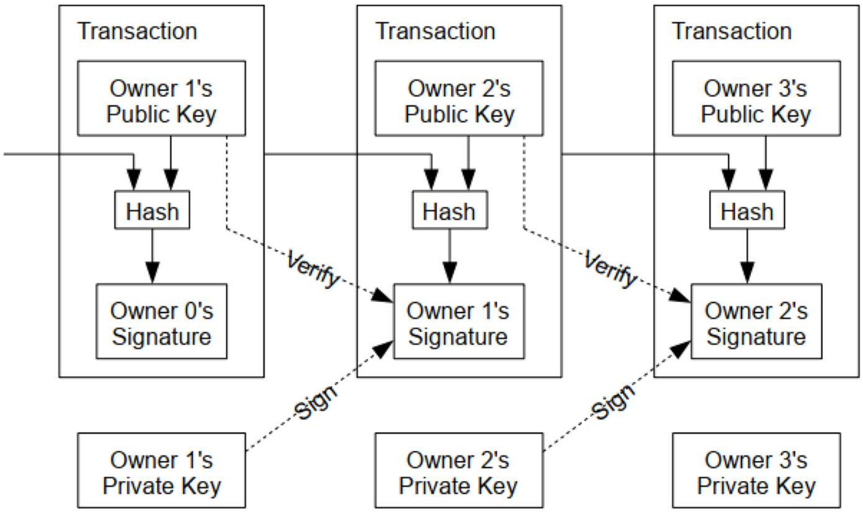
3. Conceptual framework

To follow the later investigation of transaction fees in the bitcoin network, a fundamental understanding of the system underlying bitcoin and cryptocurrencies in general is beneficial. Assuming different levels of expertise among readers on this topic, a summary of the the bitcoin network follows based on Nakamoto (2008) initial publication on the peer-to-peer electronic cash system. Further, a brief introduction of the factors influencing transaction fees follows the introduction of the bitcoin and is leading up the terminology of data utilized in this research.

3.1. Transaction in bitcoin network

Nakamoto motivates the bitcoin with the need for an electronic payment system that is based on cryptographic proof instead of trust. This allows any two willing parties to transact directly with each other, without the need for a trusted third party. To do so an electronic coin, a chain of digital signatures, is proposed. To participate, an individual needs to open an account called a bitcoin wallet, that generates a private key, a randomly generated string of letters. The owner of a wallet transfers coins, by signing a transaction with his or her private key. Specifically, an output referred to as hash of the the previous transaction is signed, which transferred the coins into his or her wallet. The new owner, shows acknowledgment of the transaction by signing with a unique public key, derived from his or her private key (see Figure 1).

Figure 1: Signing transaction in the blockchain with the private key



Note: Each block represents one transaction, with owner 1 being the reciprocate of transaction 1 and owner 0 being the payee in transaction 1 and so for. Each new owner of a bitcoin is verified by the previous owner's public key and signed by the previous owner's private key. Thereby a chain of ownership is established and used to verify the ownership of the bitcoin. A private key is a random string of letters that through a cryptographic algorithm forms the public key in a on way manner. This means retrieving the private key based on the public key is virtually impossible. The private key is created once a wallet, an account for a bitcoin broker has been opened.

A public key can be interpreted as an IBAN in common FIAT currency system and is a mathematical function of the private key. The function (hash function) is a cryptographic algorithm[1] that enables an easy link from the private key to the public key. On the other hand, the function makes it virtually impossible to retrieve the private key, if only the public key is known. The payee in the transaction can verify the signature to prove the chain of ownership but is unable to verify if one of the previous owners did not double spend coins.

Instead of relying on a trusted central authority like a central bank, that checks every transaction for double spending, Nakamoto proposes to record a public history of all transaction. Therefore, next to payee and payer, a further entity takes part in the transaction, the node, who records and saves all transaction and thereby all previous owners of each bitcoin to provide the network consensus.

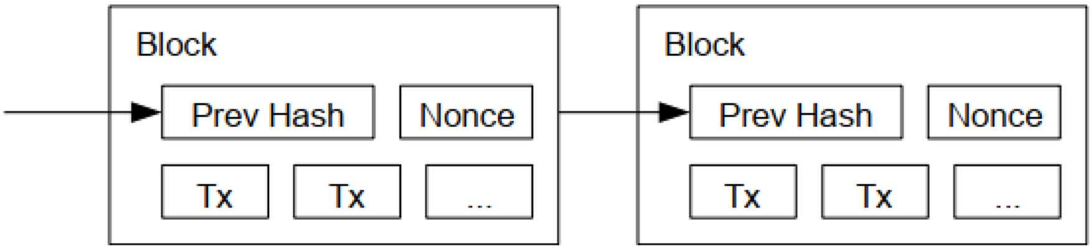
[1] The cryptographic algorithm in bitcoin is called: Secure Hash Algorithm 2, short: SHA-256 (Nakamoto, 2008)

The consensus is established by first, broadcasting new transactions to all nodes (nearby in the network). The nodes store the pending new transaction in the mempool, a virtual waiting line in form of a block. Now, nodes pick up the transactions and compete on finding a solution for the block, referred to as proof-of-work. Nodes that do this are also called miners. The task of mining requires nodes to utilize the cryptographic hash function (SHA-256) on the input of the block of new transactions, the proof-of-work of the previous block (Hash) and nonce, last being random letters and digits. The task is complete when a certain nonce, an alpha-numeric string, is found that gives a 32 characters outcome of the hash function that starts with a certain number of trailing zeros (Badev, & Chen, 2014).

Finding these strings is random and therefore looking for them involves a computational workload. The work required to obtain the required hash is exponential in the number of trailing zeros and the success rate is referred to as hash rate. In adjusting the number of trailing zeros, the system can modify the difficulty[2] and therefore the hash rate of the mining process. The algorithm sets the difficulty to a level such that on average a block is added to the blockchain every ten minutes (Easley et al., 2019). Once a miner completes the proof-of-work and finds the required hash, it broadcasts the block to all other nodes in proximity. Other nodes accept the block if each individual transaction in the block is valid and not already spent. To verify the proof-of-work, only a single hash, the outcome, needs to be executed and checked. To control the computational effort of a block, once the hash is given, is very easy but vice versa is difficult, similar to the process of retrieving the public key from the private key. Now, all transaction that are worked into the block are accepted and other nodes express their acceptance by working on creating the next block in the chain using the hash of the accepted block as the previous hash (see Figure 2).

[2] The difficulty of the proof-of-work is determined by a moving average, to compensate for increasing hardware speed and varying interest in running nodes over time (Nakamoto, 2009). The first block started with the difficulty of 1.

Figure 2: Simplified visualization of the blockchain.



Note: Each block contains a hash function (Prev Hash), that is based on a cryptographic function from the previous block. With the hash function the transaction in the previous block can be verified. Next, it contains all new transaction (Tx) that were received by a node at a certain time. In addition, it contains random strings of numbers (Nonce) that are solved based on random computation in the so-called proof-of-work process, otherwise known as the mining process. The node responsible for the block can only broadcast the block to the blockchain, if a new hash is found that can be used as the Prev Hash for the next block by other nodes to verify all new transaction in the block.

3.2. Transaction fees in Bitcoin network

The computational effort nodes exert to solve the nonce into the required outcome to include a block into the blockchain is financially rewarded. In the following, this reward is referred to as mining reward. The Mining reward is by convention only issued for the first transaction in each block and given to the node that embedded this transaction and hence the creator of the block. Nakamoto (2008) states the steady addition of new coins is analogous to gold miners adding gold to the circulation. However, the pool of available bitcoins for issuance is limited to 21 million. Hence, the block reward is set to decline and therefore halved after every 210.000 blocks that are mined[3]. The current mining reward stands at 6,25 bitcoins per mined block and resulted from the most recent reduction on May 11th, 2020.

The bitcoin and its issuance structure therefore is designed to be deflationary. Notably, a rising bitcoin price could counteract or even offset any mining reward declines and deflationary properties, if the increase in the value of the bitcoin is larger than the decrease in absolute bitcoins rewarded. In addition, the rate at which new bitcoins are mined depends on the difficulty level explained in 3.1., which can reduce the incentives for miners to participate and enable the

[3] The initial mining reward per block of 50 bitcoins was first reduced in November 2012 to 25 bitcoins.

bitcoin infrastructure (Huberman et al., 2017). Therefore, there is need for a second incentive for miners. Here transaction fees enter the bitcoin system. Transaction fees describe an input value for a transaction that exceeds the output value: “the difference is (the) transaction fee that is added to the incentive (…) of the block containing the transaction” (Nakamoto, 2009, p. 4).

The duality of incentives for participation in the blockchain network implies two aspects. Fristly, a possible model of transaction fees can be described, in which transaction fees dependent next to the market demand and supply of transactions on mining revenue and the rate at which blocks are mined. This model is further investigated in the next section. Secondly, the incentive structure of cryptocurrencies can transition entirely to transaction fees and neglect mining revenue. This long run equilibrium is discussed further in Huberman et al. (2017) research. Therefore, a system is feasible that does not add inflation though unexpected levels of currency issuance through the mining process, which in turn is appealing to central banks planning to utilize the blockchain. Not introducing uncontrolled inflation in government issued currencies is in the interest of central banks as it ensures that monetary policy sovereignty is not affected by a government issued currency based on the blockchain.

4. Research Methodology

In order to examine transaction fees in the bitcoin network, in the following, the data collected is introduced, the historical development of transaction fees over the sample period visualized and lastly both hypotheses further investigated.

4.1. Data

To investigate and forecast transaction fees in the bitcoin a sample is constructed that contains daily data from July 25th of 2016 until May 24th of 2021. The data is retrieved from satoshi.info, coinmetrics and blockchain.com. All three data providers, collect the data by running a node in the bitcoin network, that records bitcoin system specific data.

The dependent variable in this research, daily mean transaction fee, is recorded in two ways, once per successful transfer and second per byte, both in USD. On the on hand, fees per successful transfer are considered, given that they represent an intuitive overview, that resembles the classical illustration of transaction fees. On the other hand, a per byte scale is considered, as transaction can differ in their computational effort even if constituting to the same amount of value transferred. The computational effort depends on the size of the transaction in byte, which in return depends on the number of previous owners of the transferred bitcoin(s)[4]. Hence, recent mind bitcoins have less previous owners and require less computational effort to control. A closer comparison between both will proceed in the next section.

Furthermore, the dataset contains the estimated mean daily hash rate. The hashing power is estimated based on the blocks mined in a certain interval and the corresponding difficulty of the bitcoin network[5]. In addition, the dataset contains the mean daily size of the mempool, the daily Bitcoin price in USD reported at 23:00:00 and the daily mean mining reward per byte. The last is estimated by considering the network specified reward per mined block relative to the mean daily transaction per block. The obtained mining reward per transaction is then taken relative to the mean daily byte size of each transaction[6].

4.2. Historical transaction fee development

The sample period starts in July 2016, thereby capturing the turbulent and chaotic time during fall of 2017, when bitcoin reached its first peak at above 19.000 USD per coin. In addition, the sample captures the extreme rise in 2020 and 2021, with bitcoin prices above 60.000 USD,

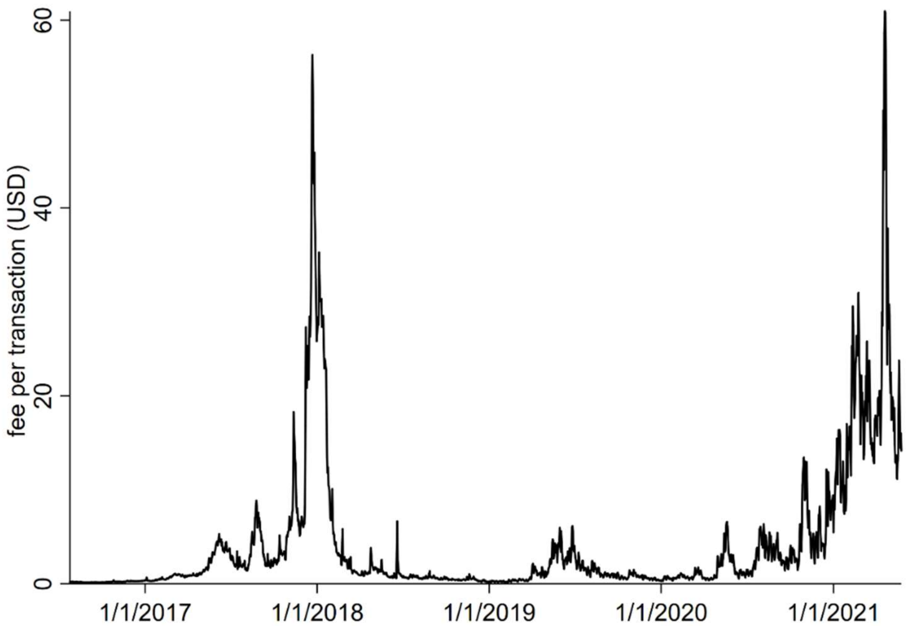
[4]Individual A and B both want to transfer one bitcoin to individual C. Individual A owns a bitcoin that he purchased from 10 previous owner (each owned 0,1 bitcoin), while induvial B received her bitcoin from only one previous owner. If both A and B now transfer their bitcoin, miners have to control the ownership for 10 previous owners for the transaction of owner A, compared to only 1 previous owner under transaction B. Therefore, the computational effort measured in bytes for transaction A is larger.

[5] Ozisik, A. P., Bissias, G., & Levine, B. N. (2017) elaborate on different estimation techniques. They differentiate between network Hash-rate and miner´s hash rate and provide empirical evaluation accuracy. But this goes beyond the scope of this analysis.

[6] The mean daily byte size of transaction is retrieved from the mean daily number of transactions per block and the mean daily byte size of a mined block in the blockchain.

followed by the volatile May of 2021, induced through the ban of cryptocurrency related services in China (Reuters, 2021), that lead to a decline of the bitcoin price. Especially important relating to transaction fees are the Chinas power shortages in April 2021. The shortages were caused by mandatory closings of mines in China, that impacted the energy supply of the large miner network located in China and hence the transaction fees. The first rise and spike of the bitcoin price and the power shortages in China are clearly visible in Graph 1, which visualizes the development of transaction fees. The turbulent bitcoin price during 2020 and 2021 and the power shortage in China are a partial explanation for the largest daily mean transaction fee in the sample, at above 60 Euros in April 2021.

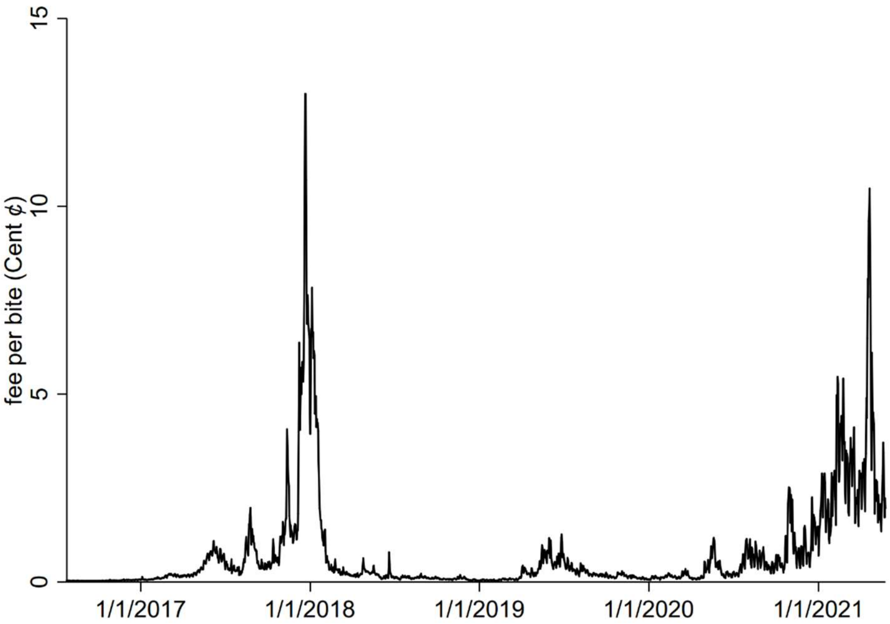
Graph 1: Historical transaction fee per transaction



Note: data is taken from coinmetrics and represents the mean fee in USD per transaction on a daily interval. Transaction represents a bundle of intended actions the bitcoin ledger initiated by a user (bitcoin wallet). Changes to the bitcoin ledger mandated by the bitcoin protocol (not by users) or new issuance issued by controlling entity are not included under transaction, to proxy only the activity of active users. However, in case a transaction is not executed, or does not result in a transfer of native units it is still counted in the daily transaction count, which potentially induces minor measurement errors.

While Graph 1 is intuitive to understand, one issue is that it does not allow to make interpretations on the computational effort exerted and paid for by miners. The representation oversimplifies all transaction as of equal size. However, the size of transaction differs, hence, the computational effort involved in transaction varies. Therefore, Graph 2 visualizes the transaction fee per byte over the sample period.

Graph 2: Historical transaction fee per byte



Note: date taken from coinmetrics and representing the mean transaction fee per byte. Estimated through daily sum of mean transaction fees relative to daily mean byte size of a single transaction. Last, obtained though daily mean byte size of a block on the bitcoin blockchain and the daily mean number of successfully mined transaction per block.

The development of transaction fee per byte is mostly equal to the development of Graph 1. Nevertheless, slight differences persist, with the most obvious one being the transaction fee per byte during the fall of 2017 exceeding the fee per byte paid in April 2021. The reason underlying the difference is in the development of the mean block size in the bitcoin network. An increase in the mean bytes per block is caused by individual transaction that constitute more inputs and outputs resulting from added previous owners.

To provide comparable insights and results among different cryptocurrencies for potential future research of the underlying components of transaction fees, the following research adopts transaction fee per byte as the dependent variable scale.

4.3. Data transformation

Each individual variable is controlled for stationarity. If non-stationarity is detected, the series suffers from persisting shocks, that never die away. Thereby, the current value of the variable would be just an infinite sum of past shocks and a starting value at the beginning of the time interval (Brook, 2008). In other words, the best prediction of non-stationary process is its most recent past value. To prevent this first, the integer of order one is taken from the bitcoin price (BTCP), meaning its current value is deducted by its most recent lagged value. In addition, the natural logarithm is taken to stabilize the residual variance and take advantage of adaptive returns, a process based on Box and Jenkins´ (1976) proposition to utilize logs in constructing forecasting models. The same transformation is applied to the hash rate. Further, the mining reward is expected to contain a trend, therefore the first order integer is applied.

To test whether the transformation has been successful and whether the not transformed variables are stationary, an augmented Dickey Fuller test is applied. The result of the test is represented in Table A in the appendix. The null hypothesis is rejected in favor of a stationary process for both not transformed variables, transaction fee per byte and mempool size and for all three transformed variables.

4.4. Transaction fee model specification and empirical analysis

Establishing why there is a need for computational effort in the transaction mechanism of bitcoin in 3.1. and the duality of incentive for nodes to participated in the network under 3.2., I proceed with a potential model describing the transaction fees. The model attempts to proxy a demand and supply estimation through cryptocurrency specific factors.

The model proxies the total demand for transaction in the bitcoin network though all the transaction saved in the mempool in a daily interval. Recall from 3.1., new transactions are stored in the mempool by nodes after being shared publicly. The mempool however proxies a further indication in the bitcoin network. The inflow of new transactions compared to the outflow of transactions. Pending transaction leave the mempool either when posted to the blockchain or dropped from the pool, typically after three days (Easley et al., 2019). If more transactions enter the mempool compared to leaving the mempool, longer waiting times are the result and the network suffers from congestion.

The change in the supply for transaction processing power in the network is included through the hash rate. Recall from 3.1., the hash rate indicates the computational power in the bitcoin network necessary for the mining process, which verifies transactions. An increase in the hash rate for instance can prevent congestion of transaction in the mempool by accelerating the outflow. Moreover, the return of bitcoin is included, to proxy for exogenous variation in the transaction fees from speculation of the bitcoin and other exogenous events represented through bitcoin price changes. Additionally, the mining reward per byte is added to the model. Figure 3 provides a visualization of the described time series OLS regression in a daily interval.

Figure 3: time series OLS regression of transaction fee per byte with daily intervals

$$
\begin{aligned}
&fee_{t,perbyte,cent} \\
&\qquad = \beta_1 mempoolsize_t + \beta_2 \Delta \mathrm{Ln}\,(hashrate_t) + \beta_3 \Delta Ln(BTCP_{t,USD}) \\
&\qquad + \beta_4 \Delta Ln(miningreward_{t,perbyte,USD})
\end{aligned}
$$

t:	*time series indicator, daily interval*
$fee_{t,perbyte,cent}$:	*daily mean transaction fee per byte in the bitcoin network at time t in USD*
$mempoolsize_t$:	*daily mean number of transactions saved in the mempool of nodes in the bitcoin network at time t*
$\Delta Ln\ (hashrate_t)$:	*first order integer of the daily mean natural logarithm of the hash rate at time t*
$\Delta Ln(BTCP_{t,USD})$:	*first order integer of the natural logarithm of the daily mean price of bitcoin at time t in USD, equal to the natural logarithm of daily return of bitcoin at time t in USD*
$\Delta Ln(miningreward_{t,perbyte,USD})$:	*first order integer of the daily mean natural logarithm of the mining reward per byte in USD at time t*

The empirical analysis of the above-described model focuses on endogenous daily observable factors of the bitcoin system and additionally the bitcoin price. Therefore, the mean waiting time for the conformation of transaction is not included. Considering that the data sources of this research only report an average waiting time every three days, which potentially includes measurement or estimation errors and limits the sample size. However, the findings of Easley et al. (2019), who determined waiting time as the key driver for transaction fees in their model, are assumed to be adequately proxied though the mempool size, which is expected to represent most of the movement in average transaction time and congestion in the bitcoin network. This relationship of the mempool size is another finding of Easley et al. (2019). Mempool size can be easily overserved and reported by nodes and therefore, is expected to yield more exact indication compared to a three-day average waiting time factor. Table 1 shows the outcome of the OLS regression described in Figure 3.

Table 1: Outcome of time series OLS regression on transaction fee per byte in daily intervals

Transaction fee per byte in cent		
Independent variable	coefficient	t-statistic
mempoolsize	0,358*** [0,029]	12,28
ΔLn(hashrate)	0,336** [0,158]	2,13
ΔLn(BTCP)	(1,723)** [0,791]	(2,18)
Δminingreward	2,3 [4,749]	0,48
Constant	0,011	0,27
Number of Observation	1.765	
R-squared	0,464	

Note: Scale adjustment; fee per byte in cent; mempoolsize per 10000 to adjust increase by absolute increase by one closer to average value in the sample, Δminingreward per cent. ΔLn(hashrate) and ΔLn(BTCP) no adjustment given percentage change effect of Ln. Number in [ ] *equals the robust standard error, and numbers in regular brackets () are negative.. * Indicate the significance level: *** p<.01, ** p<.05, * p<.1;*

Table 1 adjusts the scale of the variables for visual purpose. Changing the scale of dependent and independent variables does not make a difference to the overall results, since the coefficient estimates will change by an off-setting factor to leave the overall relationship between them unchanged (Gujarati, & Porter, 1999). Hence, there are only relative change to the betas that are relative to the scale adjustment but there are no absolute changes.

Table 1 shows, that the mining reward does not have a significant relationship with transaction fees per byte. Easley et al. (2019) and Huberman et al (2017) provide a similar interpretation in their derivation of equilibrium transaction fees. Fees are not affected by block reward-based levels of mining revenue even if investigated in absolute form. Table 1, however does provide an indication of a significant relationship between the mempool size at one percent. As expected, a larger mempool size has an increasing relationship with the transaction fee per byte. Further, a significant relationship between a one percent change in bitcoin returns and one percent change of the first order integer of hash rate is found at the five percent significant level.

In the sample, an increase of mempool size by 10.000, all else equal, increases transaction fees per byte by 0,358 cent. Furthermore, a one percentage change in the difference of the hash rate indicates a 0,336 increase in fees per byte. This indicates, if governments attempt to keep transaction fees small a constant hash rate and thereby computational power in the blockchain network is beneficial and should be encouraged. Lastly, a one percent change in the bitcoin returns, all else equal, relates to a 1,723 Euro decrease in the transaction fees.

Considering the significant relationship in Table 1, a possible answer for hypothesis one is formulated. The empirical test of the model motivated by economic reason of transaction fees contains factors that show a significant relationship to the level of transaction fees. A similar answer to the first hypothesis, is discussed and empirically tested in Easley et al. (2019). They find factors that as well relate to transaction fees. Therefore, In the next step all significant independent variables are considered in predicting and forecasting changes in the bitcoin transaction fees.

4.5. ADL forecast model specification

The focus of the forecast model is to provide a prediction of the conditional expectation as good as possible. Therefore, an autoregressive distributed lag model (ADL) is considered, which enables inclusion of the dependent variables previously discussed. For simplicity, only the one-step ahead forecast is considered.

To estimate coefficients of the ADL model, a rolling window analysis is applied. The method ensures, that if parameters change at some point during the sample, the rolling estimates capture the instability, and the lag length chosen considers potential instability. A rolling window size of 1.386 days is taken, based on the change in the mining reward per block in the sample (after 1386 days), a potentially endogenously caused break in the bitcoin fee structure[7]. The objective is to forecast transaction fees for the current block reward of 6,25 bitcoins per block. Hence, the decision of which lag length to use is made after the break. In addition, the larger rolling window size is expected to result in smoother estimations compared to a shorter window size and allows to capitalize on the large set of data available in the sample. For each variable subsequently investigated, lags up to two weeks are considered in determining the best performing forecast model. The two-week consideration period is applied, following that in the sample on average every two weeks (13,6 days) the endogenous difficulty level for mining a block in the blockchain system is adjusted. To compare the predictive performance of rival models (different lag lengths), the mean square error (MSE) is considered as a measure of accuracy. For each set of variables and lag lengths, the rolling window analysis conducts a pseudo-out of sample forecast in a subsample one-step ahead forecasts. In total 833 of these pseudo-out of sample forecasts are constructed for each rolling window. The forecasts are compared to the actual transaction fees and the MSE is retrieved.

The base model and default option for the performance comparison is an auto regressive model of up to seven lags, determined by the lowest MSE in the rolling window analysis. Table 2 shows the MSE for all lag lengths of each independent variable. Each model stage includes next to the investigated independent variable the independent variable(s) and lag(s) of the previous model stage. In total, Table 2 includes all the MSE´s of the progressively added independent variables and their distributed lags and shows the final chosen ADL model at the last model stage.

[7] On the 11th of May 2020, the mining reward per block was again halved from 12,5 to 6,25 bitcoin per block.

Table 2; Mean squared errors for different stages of the ADL model

	MSE						
Model stage	lag1	lag(1-2)	lag(1-3)	lag(1-4)	lag(1-5)	lag(1-6)	lag(1-7)
AR	11,534	11,955	12,076	12,208	12,006	12,047	12,155
ADL1*	11,024	10,538	10,243	10,255	10,281	10,298	10,237
ADL2**	9,831	9,775	9,780	9,787	9,793	9,797	9,768
ADL3***	9,732	9,731	9,737	9,712	**9,701**	9,726	9,745
	MSE						
Model stage	lag(1-8)	lag(1-9)	lag(1-10)	lag(1-11)	lag(1-12)	lag(1-13)	lag(1-14)
AR	**11,295**	11,313	11,456	11,488	11,501	11,577	12,046
ADL1*	10,253	10,261	10,279	10,311	10,104	**10,023**	10,035
ADL2**	**9,728**	9,734	9,740	9,749	9,752	9,751	9,764
ADL3***	9,747	9,744	9,763	9,769	9,769	9,783	9,790

Note: all values are in 1000; e^-03. Lag1 indicates the the variable at t, while Lag (1-2) indicates inclusion of variables at time; t-1 and t-2 to predict transaction fee per byte at time; t. Bolt numbers **111** represent the minimum MSE for the specified model stage. Each model stage, takes on the previously determined number of lags that obtained the minimum MSE; in following order:

*: 8 lags of transcation fee per byte

**: 8 lags of transcation fee per byte, 13 lags of mempool size

***: 8 lags of transcation fee per byte, 13 lags of mempool size , 8 lags of Δln(hash rate)

The selected ADL model, includes an autoregressive term of eight lags from the dependent variable, transcation fee per byte. The distributed lag term constitutes of thirteen lags from the mempool size, eight lags of the hash rate and five lags of the bitcoin returns. The final model has an MSE of $9{,}701^{e\text{-}0{,}3}$ for the one day ahead forcast. Figure 4, shows the final suggested ADL model regression term to predict the one day ahead forecast of transcation fees per byte in cent at time t. The one day ahead forecast is at time time t, hence, it is predicted soley based variable lags from time t-1 or further in the past. Otherwise, the model would use infomration from the period attempted to be predicted.

Figure 4: ADL model; one day ahead forecast of transcation fee per byte in USD

$$fee_{t,per\ byte,cent} = \sum_{l=1}^{l=8} \beta_l^1 * L._{t-l}\, fee_{perbyte_1} + \sum_{l=1}^{l=13} \beta_l^2 * L._{t-l}\, mempoolsize + \sum_{l=1}^{l=8} \beta_l^4 * L._{t-l}\, \Delta\ln\,(hashrate) + \sum_{l=1}^{l=5} \beta_l^5 * L._{t-l}\, \Delta\ln\,(BTCP)$$

$fee_{t,perbyte,cent}$:	transaction fee per byte at time t in cent
β_l^n:	OLS beta coefficient at, n = variable (1-5) and l = lag of variable
t:	time series indicator, daily interval
l:	timeindicating variable that takes on the value per variable defined in operator $\sum$.
$L._{t-l}\, fee_{perbyte}$:	the lag of transcation fee per byte, the time of the lag = $t - l$
$L._{t-l}\, mempoolsize$:	the lag of once differenced mempool size, the time of the lag = $t - l$
$L._{t-l}\, \Delta\ln(hashrate)$:	the lag of once differenced natural logarithm of the hash rate, the time of the lag = $t - l$
$L._{t-l}\, \Delta\ln\,(BTCP)$:	the lag of once differenced natural logarithm of the Bitcoin price, the time of the lag = $t - l$

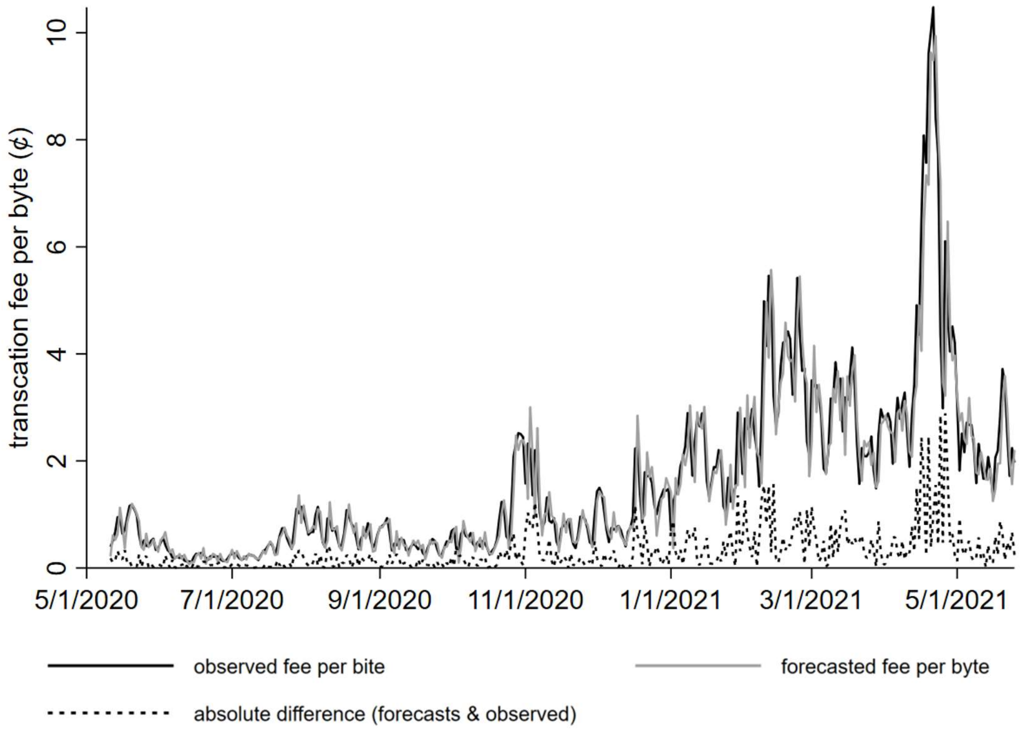
4.6. ADL forecast model analysis

Next, the chosen ADL model is analyized to provide a possible suggestive answer for the second hypothesis, whether a forcast model could be used to predicte bitcoin fees in a one day ahead interval. In addition, the question is raised whether such insight could have predicted the spike in bitcoin transction fees in April of 2021. Therfore, the performance of the ADL forecast model is further investigated in two ways.

First, the predicted fees are compared to the observed fees and visulaized in Graph 3, over the forecasted sample period. In addition, Graph 3 indicates the absolute difference between forecast and observed fees. Secondly, the beta coefficients of the ADL model (see Figure 4) are estimated to forecast transcation fees for the 19.04.2021 in Table B, Appendix.

Graph 3 visualizes, that the combination of autoregressive terms and blockchain specific distributed lag indicators provide a prediction of the rise in transaction fee in April of 2021. The inital increase is adequatly forecasted. However, an increase in the absolute deviation is observed that persisted for several days even after the inital spike as passed. A similar increase in absolute deviation is observed in other smaller spikes in trancation fee per bytes previous to April 2021.

Graph 3: ADL Forcast performance



Note: forecasted values retireved from rolling window prediction for selected (final) ADL model. Absolute difference equals the absolute differnece between observed and forecasts transcation fee per byte in cent.

The deviations during shocks can be partially explained by examining the coefficient estimations of Figure 4, in Table B Appendix. The estimation conducted for the 19th of April is choosen as an example given that from the 18th of April to the 19th of April transcation fee (per transcation) increased by 14,2 Euros to 58,58 Euros, the largest increase in a single day for the transcation fees in the bitcoin network. Figure 4 indicates, that the forecast predomiantly emerges of the autoregressive term. Specifically 88,92 percent of the forecast value for April the 19th equals to the forecast of the autoregressive term. Hence, the recognition of exogenous shocks is lagging behind. Graph B in the Appendix visualizes this lagging of the forecast by focusing on the 2021 April forecast values.

Follwing the two investigations of the ADL model, a mixed answer for the second hypothesis is proposed. On the one hand, the transcation fees per byte can be forecasted. On the other hand, the forecast is predominantly based on autoregressive properties of transcation fees and the prediction of exogenous shocks is therefore only limited. The resulting deviations can have large accumulated implications, considering the daily scale of transcations in a government controlled and issued currency system.

5. Conclusion

The paper shows, that a model of factors influencing transaction fees can be determined. The model investigated under the first hypothesis, follows the findings of Huberman et al. (2017) and Easley et al. (2019), that transcation fees are determined and emerge in the absence of the mining reward. Furthermore, internal congestion proxied through the hash rate and the mempool drives transcation fees in the absence of exogenouse shocks modeled through the bitcoin price movements. A government issued currency distributed through the blockchain technology should, to minimize transcation fees and thereby market friction, aim to prevent congestion on the blockchain.

The market is not oblivious to the issue of congestion on the blockchain and the resulting limitation to scalability. A possible solution discussed is to increase in the absolute bitcoin block size (Easley et al. 2019), or the introduction of a lighenting network that facilitates an instantaneous transfer of coins (Miraz, & Donald, 2019). The compatibility of such methods in a government affiliated currency based on the blockchain and its effect on congestion is required to be further researched.

Moreover, this paper proposes a forecast model for transaction fees in a one-day ahead interval. While the applied ADL model, provides a general indication of the transcation fee development, its performance is limited considering the prediction of exogenouse shocks. This is due to a strong reliaance on the autoregressive term in the forecast. Further research, could attempt to construct either another linear prediction model including moving average terms (e.g. ARIMAX) or a non-linear lag selection process.

6. Appendix

Table A: Augmented Dickey Fuller test

Dickey fuller test with trend		
	test statistic	Dickey-Fuller 1% critcal value
fee per byte	(5,38)	(3,96)
mempool size	(8,79)	(3,96)
Δmining reward	(58,32)	(3,96)
ΔLn(hash rate)	(69,30)	(3,96)
ΔLn(BTCP)	(44,09)	(3,96)
Dickey Fuller with 14 lags		
	test statistic	Dickey-Fuller 1% critcal value
fee per byte	(5,23)	(3,42)
mempool size	(4,87)	(3,42)
Δmining reward	(11,23)	(3,42)
ΔLn(hash rate)	(15,31)	(3,42)
ΔLn(BTCP)	(10,16)	(3,42)

Note: Dickey fuller test to test for unit root process H0: not-stationary process with a unit root. If test statistic < critical value, here 1%, H0 is rejected in favor of stationary process. The augmented Dickey Fuller test can control for trend (adjust critical values) and for lags in variable that is to be tested. 14 lags are included, considering that later in the forecast a test period of 14 days is considered, see Section 7. Numbers in regular brackets () are negative values.

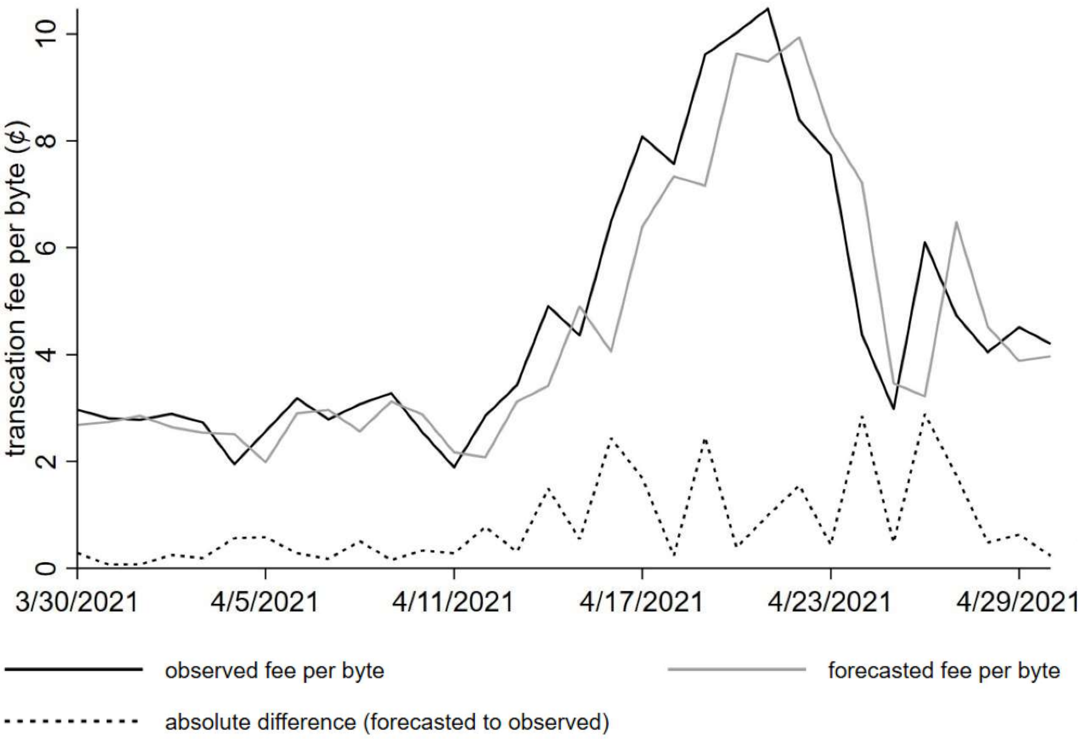
Table 3: Linear regression of ADL coefficients

dependent variable: fee per byte USD			
	Coefficient	Standard Error	t-statistic
fee per byte USD			
L1	0,996***	[0,030]	33,37
L2	(0,158)***	[0,043]	(3,71)
L3	0,099**	[0,043]	2,30
L4	(0,129)***	[0,043]	(2,99)
L5	(0,004)	[0,043]	(0,10)
L6	0,049	[0,043]	1,13
L7	0,294***	[0,042]	6,93
L8	(0,274)***	[0,029]	(9,37)
Mempoolsize			
L1	1,36E-07***	[1,13E-08]	12,05
L2	(1,66E-07)***	[1,55E-08]	(10,65)
L3	8,25E-08***	[1,64E-08]	5,02
L4	7,25E-09	[1,66E-08]	0,44
L5	(1,40E-08)	[1,66E-08]	(0,85)
L6	3,17E-08*	[1,65E-08]	1,92
L7	(4,15E-08)**	[1,65E-08]	(2,51)
L8	5,04E-09	[1,62E-08]	0,31
L9	(2,71E-09)	[1,60E-08]	(0,17)

L10	6,68E-09	[1,48E-08]	0,45
L11	(1,78E-08)	[1,46E-08]	(1,22)
L12	5,69E-09	[1,40E-08]	0,41
L13	4,25E-08***	[9,75E-09]	4,36
ΔLn(Hashrate)			
L1	3,59E-03***	[9,77E-04]	3,67
L2	(1,68E-03)	[1,21E-03]	(1,39)
L3	7,02E-04	[1,34E-03]	0,52
L4	1,66E-03	[1,40E-03]	1,18
L5	1,69E-03	[1,40E-03]	1,21
L6	2,79E-03**	[1,34E-03]	2,09
L7	4,30E-03***	[1,21E-03]	3,55
L8	2,74E-03***	[9,73E-04]	2,82
ΔLn(BTCP)			
L1	(2,42E-04)	[2,08E-03]	(0,12)
L2	(1,81E-03)	[2,10E-03]	(0,86)
L3	5,49E-04	[2,10E-03]	0,26
L4	4,09E-03*	[2,10E-03]	1,95
L5	5,84E-03***	[2,09E-03]	2,79
Constant	(3,42E-04)***	[1,26E-04]	(2,71)
Number of observations		1.387	
R-squared		0,954	

Note: estimation of beta coefficient of selected ADL model to forecast 19.04.2021. Sample period considered is 1366 days with a daily interval. *Number in [] equal the robust standard error, while number in regular brackets () are negative. * Indicate the significance level: *** p<.01, ** p<.05, * p<.1;*

Graph B: ADL performance for April 2021



Note: forecasted values retireved from rolling window prediction for selected (final) ADL model. Absolute difference equals the absolute differnece between observed and forecasts transcation fee per byte in cent.

7. Bibliography

Badev, A. I., & Chen, M. (2014). Bitcoin: Technical background and data analysis. Retrieved from https://cointhinktank.com/upload/bitcoin%20technical%20background%20and%20data%20analysis.pdf

Barber, S., Boyen, X., Shi, E., & Uzun, E. (2012, February). Bitter to better—how to make bitcoin a better currency. *In International Conference on Financial Cryptography and Data Security* (pp. 399-414). Springer, Berlin, Heidelberg.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: forecasting and control (Fifth Edidtion). *John Wiley & Sons.*

Brooks, C. (2019). Introductory Econometrics for Finance (Fourth Edition). *Cambridge University Press*. DOI: 10.1017/9781108524872

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

European Central Bank (2020) Report on a digital euro. *ECB Research & Publication.* Retrieved from https://www.ecb.europa.eu/pub/pdf/other/Report_on_a_digital_euro~4d7268b458.en.pdf#page=4

Eyal, I., & Sirer, E. G. (2014, March). Majority is not enough: Bitcoin mining is vulnerable. *In International conference on financial cryptography and data security* (pp. 436-454). Springer, Berlin, Heidelberg.

Grinberg, R. (2012). Bitcoin: An innovative alternative digital currency. *Hastings Sci.& Tech*. *LJ*, 4, 159.

Gujarati, D. N., & Porter, D. C. (1999). Essentials of econometrics (Vol. 2). *Singapore: Irwin/McGraw-Hill.*

Huberman, G., Leshno, J., & Moallemi, C. C. (2017). Monopoly without a monopolist: An economic analysis of the bitcoin payment system. *Bank of Finland Research Discussion Paper*, (27).

Irrera, A., & Lewis, B. (2021, July) More institutional investors expect to buy digital assets, study finds. *Thomson Reuters*. Retrieved from https://www.reuters.com/business/most-institutional-investors-expect-buy-digital-assets-study-finds-2021-07-20/

King, S., & Nadal, S. (2012). Ppcoin: Peer-to-peer crypto-currency with proof-of-stake. *self-published paper*, August, 19(1)

Maurer, B., Nelms, T. C., & Swartz, L. (2013). "When perhaps the real problem is money itself!": The practical materiality of bitcoin. *Social Semiotics*, 23(2), 261-277.

Miraz, M. H., & Donald, D. C. (2019, April). LApps: technological, legal and market potentials of blockchain lightning network applications. In *Proceedings of the 2019 3rd International Conference on Information System and Data Mining* (pp. 185-189).

Ozisik, A. P., Bissias, G., & Levine, B. N. (2017). Estimation of miner hash rates and consensus on blockchains. *arXiv preprint arXiv:1707.00082*.

Shen, S., & Siu, T. (2021, April) China bans financial, payment institutions from cryptocurrency business. *Thomson Reuters*. Retrieved from https://www.reuters.com/technology/chinese-financial-payment-bodies-barred-cryptocurrency-business-2021-05-18/

Yermack, D. (2015). Is Bitcoin a real currency? An economic appraisal. *In Handbook of digital currency* (pp. 31-43). Academic Press