


RESEARCH

Open Access



Does a higher hashrate strengthen Bitcoin network security?

Daehan Kim¹, Doojin Ryu^{2*}  and Robert I. Webb^{3,4}

*Correspondence:
doojin.ryu@gmail.com

¹ 22nd Infantry Division,
Republic of Korea Army,
Goseong, Korea

² Department of Economics,
Sungkyunkwan University, Seoul,
Korea

³ McIntire School of Commerce,
University of Virginia,
Charlottesville, VA, USA

⁴ Global Finance Research Center,
Sungkyunkwan University, Seoul,
Korea

Abstract

In the blockchain world, proof-of-work is the dominant protocol mechanism that determines the consensus of the ledger. The hashrate, a measure of the computational power directed toward securing a blockchain through proof-of-work consensus, is a fundamental measure of preventing various attacks. This study tests the causal relationship between the hashrate and the security outcome of the Bitcoin blockchain. We use vector error correction modeling to analyze the endogenous relationships between the hashrate, Bitcoin price, and transaction fee, revealing the need for an additional variable to achieve our aim. Employing a measure summarizing the growth of demand factors in the Bitcoin ecosystem indicates that hashrate fluctuations significantly influence security level changes. This result underscores the importance of the hashrate in ensuring the security of the Bitcoin blockchain.

Highlights

- We investigate the causal relationship between the hashrate and the security outcome of the Bitcoin blockchain.
- We analyze the endogenous relationships in the Bitcoin ecosystem.
- Measuring the growths in Bitcoin demand factors indicates that hashrate movements significantly influence changes in the security level.

Keywords: Bitcoin, Blockchain security, Blockchain sustainability, Financial innovation, Proof-of-work

JEL Classification: G12, G29, L86

Introduction

One of the key features of the Bitcoin network is its proof-of-work consensus mechanism, which requires a significant amount of computational power to solve cryptographic puzzles.¹ The Bitcoin network generates a randomly selected target *hash* value. Bitcoin “miners” compete for the right to add the subsequent block in the blockchain and earn the compensation of guessing possible hash values equal to or less than the target

¹ See Irresberger et al. (2021) for figures related to the numbers of proof-of-work-based and proof-of-stake-based blockchains.

value. The *hashrate* is simply the number of guesses of the hash value per second across all Bitcoin miners. A higher hashrate means that more computational power is being directed toward securing the network, making it more resistant to attacks from malicious actors. Our paper is devoted to empirically examining this security mechanism.

Computer scientists are constantly finding new methods of attacking blockchains. Saad et al. (2020) provide a systematic review of different types of attacks, yet controversy remains concerning the definition of attacks. Our research highlights the hashrate's role in ensuring blockchain security with proof-of-work protocol. Hence, our interest is in attacks associated with peer-to-peer architecture in establishing a consensus on which transactions to validate. Among such attack types of interest, the most famous is the "51% attack," also called a "majority attack." If a single Bitcoin miner or group of miners working in concert could obtain a simple majority of the computational power on a blockchain, the miner(s) could manipulate information on the blockchain. This concern is valid, as instances of 51% attacks have been frequently detected throughout different blockchains, including "Bitcoin Gold" and "Litecoin Cash" (Lee and Kim 2020; Shanaev et al. 2020). Bitcoin is often highlighted for its vulnerability to 51% attacks due to the high block generation time variance (Bissias and Levine 2020; Bazzanella and Gangemi 2023). Budish (2022) discusses two subtypes of 51% attacks, namely, double-spending and *sabotage*, arguing that a double-spending attack is more likely to occur. Many studies of the 51% attack focus on the double spending attack subtype (Aponte-Novoa et al. 2021). Other blockchain attacks include "*eclipse attacks*" (Heilman et al. 2015) and "*selfish mining*" (Eyal and Sirer 2018). Another is the "*Sybil attack*," which is usually implemented to execute other attacks, such as double spending (Zhang and Lee 2019). Some of the listed attacks can increase the uncertainty of transaction confirmations and reduce miners' potential rewards and the incentive to compete in solving hash values. Prevention from attacks is connected to decentralization, which is the spirit of permissionless blockchain. Securing rewards helps small miners survive in the industry, and decentralization disrupts the will of malicious miners.

Even though technological countermeasures to blockchain attacks have been suggested, the most fundamental measure preventing the system from attacks is the hashrate. More precisely, the proof-of-work protocol is designed to *incentivize* and *increase* miners' participation in the process, thereby increasing the total hashrate and making it more expensive to conduct attacks. This approach is especially effective against a 51% attack, where an attacker must control more than 50% of the network's computational power to manipulate the network. As the hashrate of the network increases, the cost of carrying out such an attack also increases, making the network more secure.

This study seeks to uncover the causal relationship between the hashrate and the security outcome within the Bitcoin network. The proof-of-work protocol is designed so that hashrate supports the network security. While prior research, such as Ciaian et al. (2021), has used hashrate to measure Bitcoin security, the strength and significance of the relationship remain uncertain. As frequently modeled in theoretical studies (Easley et al. 2019), a hashrate level may exist that works as a threshold determining the viability of the blockchain system. In that case, if the hashrate level is above the viability threshold, the impact of hashrate fluctuations on the security can become negligible. Various economic approaches have been used to examine blockchains through different

dimensions,² and our empirical investigation will be an interesting contribution to the literature on blockchain economics.

Achieving our objective is difficult because the security level is an unobservable variable. Therefore, we need to use different factors related to the demand side of the Bitcoin blockchain, which may allow us to see users' responses on the demand side. A natural starting point is to use the Bitcoin price and the Bitcoin transaction fee because they are the variables that can increase in response to strengthened security. Nonetheless, our econometric results and analyses in section "[Endogenous relationship within the system](#)" fail to verify or deny the causal relationship between the hashrate and the security level. Thus, we can conclude that the prices of Bitcoin and the Bitcoin transaction fee are not appropriate variables to test the relationship between the hashrate and level of security. Therefore, in section "[Hashrate and the blockchain demand](#)", we compute the first principal component of four factors related to blockchain (or cryptocurrency) users, which is a proxy for blockchain demand growth. Theoretically, four factors are influenced by blockchain security but are not directly related to the hashrate in any other paths. We use the constructed measure to confirm that hashrate fluctuations affect network security levels.

Our discussion on relationships among variables within the Bitcoin network indicates a need for another approach to reach this study's goal. Nevertheless, we expect that the discussion in section "[Endogenous relationship within the system](#)" will be helpful to researchers. Due to the complex interactions within the Bitcoin system, Kubal and Kristoufek (2022) use instrumental variables estimation. We also consider the vector error correction model (VECM) for the analyses because it can capture long-run comovements among variables (Sun et al. 2023). Using VECM allows us to incorporate feedback effects among Bitcoin variables both in the short and long run. Our analyses suggest interpretations based on consistent findings from some recent works.

The remainder of this paper is organized as follows. Section "[Data description](#)" describes the sample data for our analyses, and section "[Endogenous relationship within the system](#)" discusses the endogenous relationship among the hashrate, Bitcoin price, and transaction fee within the system based on the VECM model. Section "[Hashrate and the blockchain demand](#)" explains the hashrate and the blockchain demand by analyzing the Bitcoin network's active addresses, transactions, transfers, and wallets.

Data description

Our sample dataset spans from January 1, 2017, to January 5, 2023, covering a recent period for robust analyses. The data on blockchains and cryptocurrency markets are usually recorded in Coordinated Universal Time. Daily hashrate data is originally from Blockchain.com, retrieved from Nasdaq Data Link, and measured in terahashes (trillion hashes) per second. For the Bitcoin price, we use index data in United States dollars made by Coinmarketcap, and the data are obtained through Yahoo! Finance. Since Bitcoin is traded 24 h a day, a rational measure for a daily Bitcoin price is the arithmetic mean of open and closing prices. Transaction fee data is from Blockchair; unlike Blockchain.com, Blockchair provides daily median data as well as daily average data. We use

² See John et al. (2022) for reviews of different topics in blockchain economics.

the daily median transaction fee data to avoid outlier issues. A zero fee is the lower bound for the fee level, whereas there is no upper bound because some non-strategic decision-makers can bid an irrationally high fee level, causing outlier issues. The unit of the median transaction fee value is Satoshi. In the VECM, we use logarithmized data of the three variables.³ We do not call them logarithmic data in section “[Endogenous relationship within the system](#)” for convenience. Data of the three logarithmic variables are cointegrated in order 1.

There are four demand factors (network factors) suggested by Liu and Tsyvinski (2021): the number of active addresses, the number of transactions, the number of transfers (also called payments), and the number of wallets. Data for the first three variables are from Coin Metrics. The data on the number of wallets we download from the Nasdaq Data Link are data on the number of wallets provided by Blockchain.com. The variable we employ in our empirical analysis is the growth rate rather than the raw data. Although various businesses provide different kinds of Bitcoin wallets, we believe that it will not undermine the results of our study.⁴

Endogenous relationship within the system

We use VECM to investigate the endogenous relationship among the three logarithmized variables: hashrate, Bitcoin price, and transaction fee. The Johansen (1991) test result demonstrates a single cointegrating relationship. Using the result, we write the vector error correction system as follows:

$$\begin{aligned}
 \Delta Hashrate_t &= \sum_{i=1}^L (\beta_{Hi}^H \Delta Hashrate_{t-i} + \beta_{Pi}^H \Delta Price_{t-i} + \beta_{Fi}^H \Delta Fee_{t-i}) + \lambda_H Z_t + e_t^H, \\
 \Delta Price_t &= \sum_{i=1}^L (\beta_{Hi}^P \Delta Hashrate_{t-i} + \beta_{Pi}^P \Delta Price_{t-i} + \beta_{Fi}^P \Delta Fee_{t-i}) + \lambda_P Z_t + e_t^P, \\
 \Delta Fee_t &= \sum_{i=1}^L (\beta_{Hi}^F \Delta Hashrate_{t-i} + \beta_{Pi}^F \Delta Price_{t-i} + \beta_{Fi}^F \Delta Fee_{t-i}) + \lambda_F Z_t + e_t^F,
 \end{aligned}
 \tag{1}$$

where $Z_t = Hashrate_{t-1} - 0.7489Price_{t-1} - 1.5583Price_{t-1}$. The superscript of each coefficient denotes the label of the response variable, while the subscript of each coefficient indicates the label of the corresponding variable. λ are speeds of adjustment that can capture how fast the variable can converge to the long-run equilibrium, and e_t are the error terms. We set the lag length in Eq. (1) at $L = 14$. This decision is originally from a theoretical consideration rather than an empirical one because a Bitcoin mining adjustment occurs approximately every two weeks. The adjustment mechanism of Bitcoin will be explained later. Although using a short lag length is parsimonious, it can also be dangerous. When lagged variables are insufficient, e_t^H , e_t^P , and e_t^F can be correlated with their lagged variables, which causes endogeneity in the VECM equations. Therefore, a long lag length is also required in the empirical sense.

We expect that the price and fee’s positive response to the rise in hashrate will be revealed, signaling the Bitcoin security level’s positive response to the rise in hashrate.

³ For convenience, we do not call them logarithmic data in section “[Endogenous relationship within the system](#)”.

⁴ All the raw data are preprocessed using Python, and statistical analyses are done through R.

Indicators associated with on-chain activities, including the hashrate, can affect Bitcoin price dynamics and off-chain activities and sentiments (Kukacka and Kristoufek 2023). The price can reflect a cryptocurrency's valuation, which can depend on the security level of the corresponding blockchain. The transaction fee can also be an essential factor as the hashrate influences it via the level of security. When the security level is strengthened, agents will be more incentivized to remit Bitcoin to another agent's wallet through the Bitcoin network (Kim et al. 2023). When more users try to settle their transactions, congestion occurs, raising the transaction fee (Huberman et al. 2021). Table 1 presents the estimation results.

Figure 1 displays impulse response functions from the hashrate, showing that the hashrate does not significantly impact price but weakly impacts fees. We expect that a hashrate increase will lead to higher security, inducing greater participation by users and ultimately raising the transaction fees. This is the "security path" that we want to identify; however, we obtain an unexpected result. Namely, although the response after 7 or 8 days is positive, the hashrate generally negatively influences the fee. This finding conflicts with our initial prediction that an increase in the hashrate causes a fee increase; however, this result does not preclude the possibility that a higher hashrate strengthens network security and encourages more users. A different path—which may be more direct—from the hashrate to the fee other than the "security path" can exist, and that path may dominate the causal relationship.

Another likely path is when an increase in the hashrate reduces the fee, which would be possible if some mechanism leads to a higher hashrate from a tighter transaction capacity constraint in the Bitcoin network. Findings from recent research mentioned later support the existence of two such mechanisms.

The first mechanism is related to the mining difficulty of the cryptographic puzzle that miners try to solve. Assuming a consistent level of mining difficulty, when miners invest more computational efforts in the system, the total hashrate increases, and the generation speed of a single block will be faster. Similarly, under a constant mining difficulty, when miners invest less effort, the total hashrate will be lower, and the block generation will be slower. To ensure that the interval between two successive block generations stays around $T = 10$ minutes, the protocol within the Bitcoin blockchain automatically controls the mining difficulty, representing Bitcoin's mining difficulty adjustment. The mining difficulty is adjusted every 2,016 blocks, which responds slowly to real-time situations. Furthermore, the adjustment is based on the information from the past 2,016 blocks (Noda et al. 2022).

When the hashrate increases sharply in a short period owing to a positive shock but the adjustment is still a long way off, the time interval between blocks will fall to less than 10 min. This situation results in a larger *capacity* to confirm more transactions in an equivalent time interval. Each Bitcoin user will find it easier to confirm one's transaction at this relaxed capacity constraint, leading to a lower transaction fee bid. Interestingly, the case turns the other way when the adjustment occurs. When the adjustment to limit the capacity comes, a high hashrate may no longer characterize the situation, and the capacity can be heavily limited. This theory may explain why the coefficient estimates of the hashrate with small lags are all negative in the fee equation, whereas the hashrate variables with significant lags are often positive.

Table 1 VECM estimation results

<i>Panel A. hashrate equation</i>					
ECT	-0.0028*** (0.0006)				
Hashrate-1	-0.7714*** (0.0229)	Price-1	0.1870 (0.0994)	Fee-1	-0.0200* (0.0090)
Hashrate-2	-0.6355*** (0.0289)	price -2	-0.0148 (0.1320)	Fee-2	-0.0032 (0.0090)
Hashrate-3	-0.5404*** (0.0324)	Price-3	0.2315 (0.1514)	Fee-3	-0.0188* (0.0092)
Hashrate-4	-0.3802*** (0.0349)	Price-4	-0.1036 (0.1655)	Fee-4	0.0084 (0.0092)
Hashrate-5	-0.3167*** (0.0362)	Price-5	0.1692 (0.1755)	Fee-5	-0.0185* (0.0093)
Hashrate-6	-0.2638*** (0.0369)	Price-6	0.1156 (0.1823)	Fee-6	-0.0106 (0.0093)
Hashrate-7	-0.2075*** (0.0374)	Price-7	-0.1431 (0.1856)	Fee-7	-0.0029 (0.0094)
Hashrate-8	-0.1384*** (0.0374)	Price-8	0.2514 (0.1857)	Fee-8	-0.0031 (0.0093)
Hashrate-9	-0.1068** (0.0371)	Price-9	-0.4063* (0.1826)	Fee-9	-0.0086 (0.0093)
Hashrate-10	-0.0746* (0.0364)	Price-10	0.4321* (0.1761)	Fee-10	0.0025 (0.0092)
Hashrate-11	-0.0732* (0.0352)	Price-11	-0.4323** (0.1662)	Fee-11	-0.0105 (0.0091)
Hashrate-12	-0.0254 (0.0327)	Price-12	0.2260 (0.1523)	Fee-12	-0.0011 (0.0090)
Hashrate-13	-0.0342 (0.0291)	Price-13	-0.0677 (0.1320)	Fee-13	-0.0092 (0.0088)
Hashrate-14	-0.0489* (0.0226)	Price-14	-0.1178 (0.0992)	Fee-14	-0.0130 (0.0087)
<i>Panel B. price equation</i>					
ECT	-0.0003** (0.0001)				
Hashrate-1	0.0057 (0.0050)	Price-1	0.8804*** (0.0216)	Fee-1	-0.0028 (0.0019)
Hashrate-2	0.0087 (0.0063)	Price-2	-0.7617*** (0.0287)	Fee-2	0.0002 (0.0019)
Hashrate-3	0.0063 (0.0071)	Price-3	0.6924*** (0.0329)	Fee-3	-0.0004 (0.0020)
Hashrate-4	0.0099 (0.0076)	Price-4	-0.6233*** (0.0360)	Fee-4	-0.0010 (0.0020)
Hashrate-5	0.0008 (0.0079)	Price-5	0.5828*** (0.0382)	Fee-5	-0.0021 (0.0020)
Hashrate-6	0.0020 (0.0080)	Price-6	-0.4952*** (0.0396)	Fee-6	0.0008 (0.0020)
Hashrate-7	-0.0012 (0.0081)	Price-7	0.3998*** (0.0404)	Fee-7	0.0010 (0.0020)
Hashrate-8	-0.0138 (0.0081)	Price-8	-0.3595*** (0.0404)	Fee-8	-6.1e-05 (0.0020)
Hashrate-9	-0.0123 (0.0081)	Price-9	0.3174*** (0.0397)	Fee-9	-0.0007 (0.0020)
Hashrate-10	-0.0098 (0.0079)	Price-10	-0.2129*** (0.0383)	Fee-10	-0.0004 (0.0020)
Hashrate-11	-0.0031 (0.0077)	Price-11	0.1704*** (0.0362)	Fee-11	0.0006 (0.0020)
Hashrate-12	-0.0014 (0.0071)	Price-12	-0.1245*** (0.0331)	Fee-12	0.0027 (0.0020)

Table 1 (continued)

Hashrate-13	−0.0054 (0.0063)	Price-13	0.0826** (0.0287)	Fee-13	−0.0035 (0.0019)
Hashrate-14	−0.0007 (0.0049)	Price-14	−0.0357 (0.0216)	Fee-14	0.0024 (0.0019)
<i>Panel C. fee equation</i>					
ECT	0.0038** (0.0015)				
Hashrate-1	−0.0169 (0.0579)	Price-1	0.6550** (0.2515)	Fee-1	−0.0083 (0.0227)
Hashrate-2	−0.0527 (0.0732)	price -2	−0.0050 (0.3343)	Fee-2	−0.2297*** (0.0227)
Hashrate-3	−0.1779* (0.0821)	Price-3	−0.3575 (0.3832)	Fee-3	−0.1065*** (0.0233)
Hashrate-4	−0.2731** (0.0883)	Price-4	0.5176 (0.4189)	Fee-4	−0.1286*** (0.0233)
Hashrate-5	−0.1675 (0.0916)	Price-5	0.0335 (0.4442)	Fee-5	−0.1211*** (0.0235)
Hashrate-6	−0.1238 (0.0935)	Price-6	−0.0483 (0.4614)	Fee-6	−0.0501* (0.0237)
Hashrate-7	0.1419 (0.0946)	Price-7	0.1968 (0.4698)	Fee-7	0.1954*** (0.0237)
Hashrate-8	0.0934 (0.0946)	Price-8	−0.4939 (0.4702)	Fee-8	−0.0225 (0.0236)
Hashrate-9	−0.0195 (0.0938)	Price-9	1.0076* (0.4622)	Fee-9	−0.0481* (0.0235)
Hashrate-10	0.0670 (0.0922)	Price-10	−1.2092** (0.4458)	Fee-10	−0.0330 (0.0233)
Hashrate-11	0.0557 (0.0890)	Price-11	1.3248** (0.4208)	Fee-11	−0.0791*** (0.0230)
Hashrate-12	−0.0808 (0.0828)	Price-12	−0.7185 (0.3855)	Fee-12	−0.0601** (0.0228)
Hashrate-13	−0.0499 (0.0738)	Price-13	0.5532 (0.3343)	Fee-13	−0.0287 (0.0223)
Hashrate-14	0.0689 (0.0572)	Price-14	0.2486 (0.2511)	Fee-14	0.1767*** (0.0220)

We analyze the endogenous relationships within the Bitcoin ecosystem through a vector error correction model for the hashrate, price, and fee. ECT denotes an error correction term. Inside the parentheses are standard errors. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively

The “slow and backward-looking adjustment” theory can help understand the “higher (lower) hashrate, lower (higher) fee” puzzle, but there is another hypothesis that can explain the puzzle even in the long run. As a criticism against the traditional models where the degree of competition in the Bitcoin mining industry does not influence the supply of transaction capacity, Lehar and Parlour (2022) show that a miner with high hashing occupancy can exercise its market power, strategically controlling the transaction validations. They find that miners intentionally leave blank spaces in generated blocks and do not always prioritize transactions with higher fees attached. This theory is consistent with Shao and Rajapaksa (2023), who also find that miners leave room in the blocks to increase their revenues. If less (more) hashing power is associated with higher (lower) mining concentration, the theory of “strategic miners” can account for the phenomenon of a higher hashrate reducing the fee miners earn. Many shocks in the mining industry are local events. Lehar and Parlour (2022) give an example of the 2021 coal mine disaster in Xinjiang, China. A coal mine flooded,

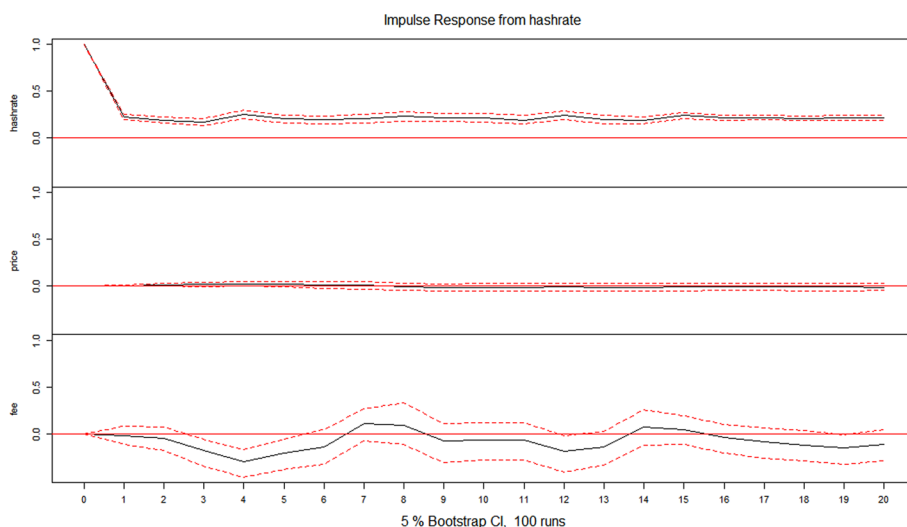


Fig. 1 Impulse response from total hashrate in Bitcoin. Notes: The three panels of this figure represent the functions of impulse response from the hashrate on the hashrate, price, and fee, respectively

and the Xinjiang region suffered a power outage, leading to the shutdown of regional Bitcoin miners. This event gave higher hashing occupancies to the miners operating in the other regions, while dropping total hashing power for several consecutive days (Makarov and Schoar 2022). The miners’ strategic behaviors can be another explanation for why a higher hashrate accompanies an increase in the fee.

Notably, a hashrate shock has no impact on the price of Bitcoin. Pagnotta (2022) argues that increasing the hashrate raises the blockchain security level, increasing the Bitcoin price. A simplistic view of this unexpected result would be that the price level depends on factors other than the blockchain’s security level. Price may be heavily driven by factors related to users’ network activity (Liu and Tsyvinski 2021) or attention (Goczek and Skliarov 2019; Koch and Dimpfl 2023).

Another interpretation, based on economic principles, is that a price rise attributed to strengthened security on the blockchain will ironically undermine its security. As Budish (2022) emphasizes, the decision of a miner to be dishonest is a matter of cost and benefit. Higher security can raise the price, but the attackers will find attacks more lucrative as Bitcoin increases in value. The higher chance of attacks will negatively influence security, offsetting the previous rise in the security level. This situation explains why the security change owing to the hashrate change has no significant impact on the price. If this interpretation is accurate, using the Bitcoin price to analyze the security level is limited.

Figure 2 shows that a positive price shock causes positive impacts on both the hashrate and the fee. It is clear why a higher price leads to a higher hashrate; the price surge means an improvement in profitability. On the contrary, it is unclear why the price shock raises the fee (even significantly). Kim et al. (2023) advance a model where a single user denoted by i decides to request its transaction be added to the Bitcoin network according to the following equation: $V > a_i E[w_i] + p\gamma_i$. Here, $E[w_i]$ is the user i ’s expected waiting time in the mempool, and a_i is a parameter denoting i ’s subjective cost of a unit waiting time. γ_i is the fee that i attaches to the potential

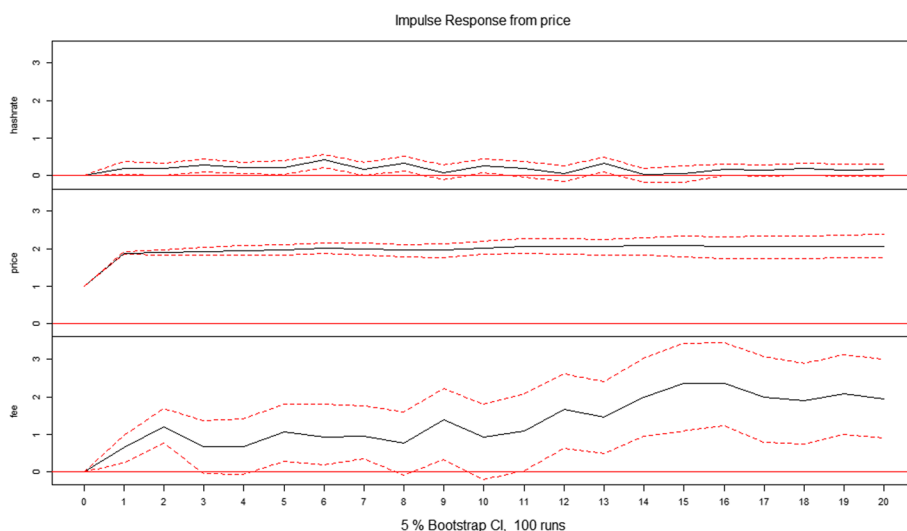


Fig. 2 Impulse response from bitcoin price. Notes: The three panels of this figure represent the functions of impulse response from the price on the hashrate, price, and fee, respectively

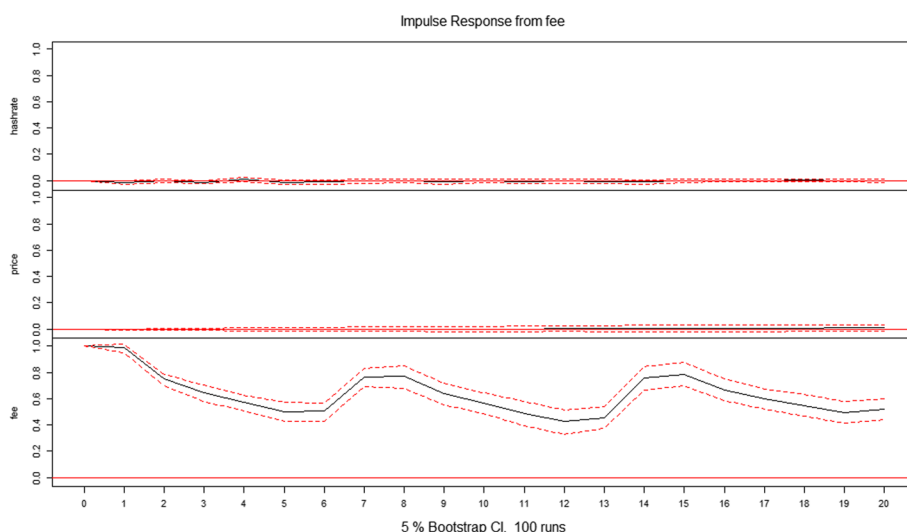


Fig. 3 Impulse response from transaction fee in the Bitcoin network. Notes: The three panels of this figure represent the functions of impulse response from the transaction fee on the hashrate, price, and fee, respectively

transaction, and p is the price of a unit of Bitcoin. Thus, the total cost from confirmation delay, $a_i E[w_i]$, plus the market value of the fee bid, $p\gamma_i$, should not exceed the expected transactional benefit, V . Accordingly, the price surge should raise the hurdle of user participation, lowering the fee by mitigating user competition.

The unexpected result of a positive price shock raising the fee may be due to a confounding factor—Bitcoin’s popularity might have raised both investors’ buying and users’ transaction demands. Another highly likely interpretation is that V in the model above may be a function of p . When the Bitcoin price increases, the market value of the transacted Bitcoins can be higher. Then, users will find the confirmation of a single transaction more valuable; thus, the value of V can rise.

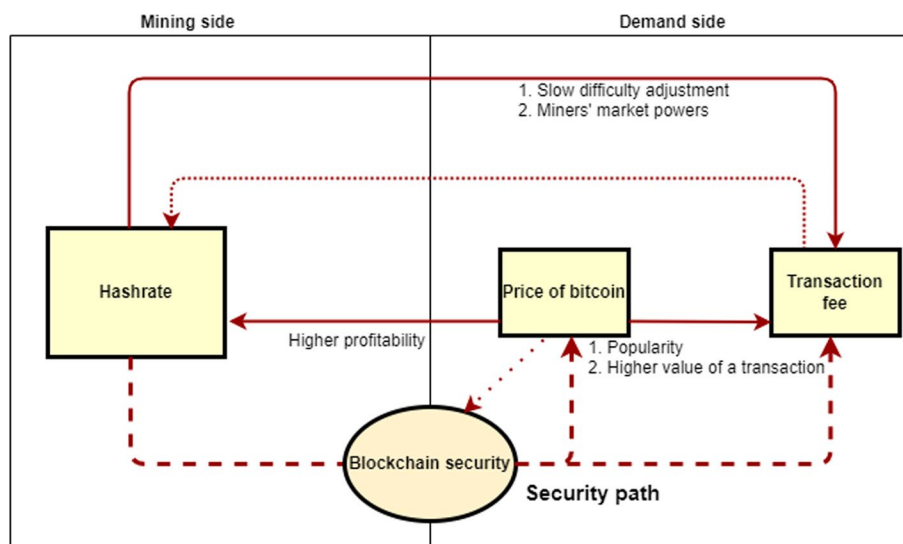


Fig. 4 Analysis of the vector error correction model results. Notes: Solid lines are directions of impacts revealed by VECM. Dashed lines represent the “security path” by which hashrate increases the security level, raising the price and fee. The densely dotted line represents an effect of fee on hashrate, which can be supported by theory but are absent in the VECM results. Lastly, a sparsely dotted line starting from the bitcoin price and pointing to security indicates a hypothetical chance that an increase in price harms security

The last impulse analysis is the impulse response from the network transaction fee shown in Fig. 3. The VECM estimation result shows that a positive fee shock does not significantly impact the hashrate and the price. Since the block reward occupies most of the total mining reward, the miners may not be actively responsive to the fluctuations in the fee.

The VECM analysis and interpretation implemented in this section are organized visually in Fig. 4. Solid lines in the figure represent directions of impact found through VECM. For each solid line, we attach a reason for the corresponding impact. Dashed lines represent the “security path” through which the hashrate boosts the security level and eventually raises the price and fee. The densely dotted line represents the effect the fee would have on the hashrate, as suggested by theory but was absent in the VECM results. Lastly, a sparsely dotted line starting from the Bitcoin price and pointing to security indicates the possibility that an increase in price harms the security. Complex interactions exist between the three variables used in the model. We conclude that the price and fee are unsuitable variables for identifying the hashrate’s impact on network security. The panels suggest that the price is determined independently of the hashrate and fee. The influence of the hashrate on the fee is “contaminated” by different paths; thus, we need a better variable dependent on the network security but not related to the hashrate through a path other than the “security path.”

The analysis in this section involves the Bitcoin hashrate, price, and transaction fee. Since the cryptocurrency mining technology is evolving toward producing more hashes with the same amount of electricity (Gundaboina et al. 2022), the total hashrate is affected by the electricity cost of hashing, implying that unit mining cost should be included in the analysis. Despite the need, electricity costs differ by region; thus, we find it challenging to construct a suitable measure for the unit mining cost. We exclude

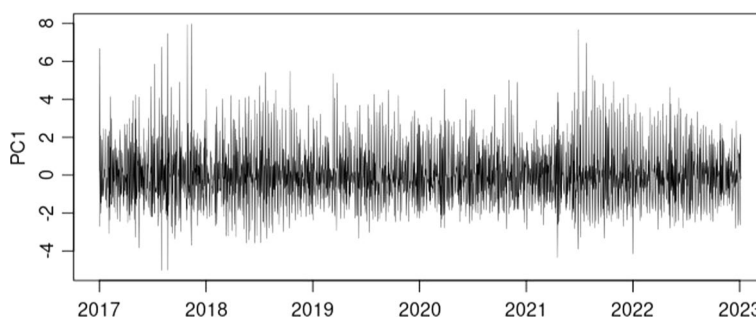


Fig. 5 The first principal component of the growth of four Bitcoin blockchain demand factors. Notes: This is the first principal component computed from PCA with growths of four Bitcoin blockchain demand factors. We call this measure “demand growth proxy.” This is a stationary $I(0)$ process

the cost associated with electricity from the analysis. Furthermore, we do not consider implemented soft forks, such as SegWit, possibly related to electricity consumption (Schinckus et al. 2022).

Hashrate and the blockchain demand

Principal component analysis (PCA) can reduce the number of dimensions by transforming the original variables. Following Liu and Tsyvinski (2021), we compute the first principal component of the growth of four demand factors. The four factors are growth rates of the number of “active addresses,” “transactions,” “transfers,” and “wallets” in the Bitcoin network. Liu and Tsyvinski’s measure covers various activities of blockchain users, making it a better alternative to Google searches (Nasir et al. 2019; Ibikunle et al. 2020) or trading volume in cryptocurrency exchange (Marmora 2022). We use linear interpolation when missing values occur in the wallet count data. There is also a limitation in interpreting the results of PCA. We identify the first component as a proxy for the demand growth in blockchain users. This proxy is an $I(0)$ process, as plotted in Fig. 5. As most trades of major cryptocurrencies, such as Bitcoin, are currently being executed on centralized exchanges (Aspris et al. 2021), some distance exists between the computed demand growth proxy and speculative demand for cryptocurrency. The demand growth proxy mainly reflects the activities done by payments or remittances.⁵

We build a simple linear regression model as follows:

$$Demand_{Growth_t} = \mu + \theta \Delta Hashrate_t, \tag{2}$$

where θ is the coefficient on hashrate growth and μ is an intercept parameter. We use hashrate growth data rather than the hashrate per se because the computed demand growth variable is an $I(0)$ process. Testing the coefficient in Eq. (2) with the traditional ordinary least squares (OLS) method is subject to inconsistency problems arising from serial correlation. We apply the Newey–West (1987) estimator to test the coefficients’ significance with OLS estimates. It is one of the HAC (heteroskedasticity- and autocorrelation-consistent) estimators used to estimate covariance matrix when data violate

⁵ Some of the remittances may be related to users engaged in illicit activities who are more interested in anonymity rather than security; however, as shown by Song et al. (2023), Bitcoin is not actively used as a financial instrument for illicit activities.

traditional OLS assumptions, including heteroskedasticity and autocorrelation. Using the Newey–West estimator, we get the estimate of 4.8630 and the t-statistic of 12.4627, which shows a high level of statistical significance.

Another way to cope with serial correlation issues is to use an autoregressive distributed lag (ARDL) model with a sufficient lag size. The ARDL model combines autoregressive (AR) and distributed lag (DL) components. The AR component captures serial correlation, and the DL component captures the lagged effects of an explanatory variable. Unlike simpler models, this model enables us to explore dynamic effects since there can be some delays in the Bitcoin network. Even if the hashrate determines the network security, the security level may not quickly respond to the hashrate fluctuations. Furthermore, significance of a coefficient on a lagged regressor cannot be interpreted as a sign of reverse causality. For the stated reasons, we use an ARDL model as follows:

$$Demand_{Growth_t} = \delta + \sum_{j=1}^p \phi_j Demand_{Growth_{t-j}} + \sum_{j=0}^q \theta_j \Delta Hashrate_{t-j}, \tag{3}$$

where ϕ_j represent the coefficients on lagged demand growth variables. θ_j are the coefficients on hashrate growth and lagged hashrate growth, and δ is an intercept parameter. The lag length is selected to minimize the Bayesian information criterion. We obtain a lag length of $p = 9$ for the lagged demand growth and a length of $q = 8$ for the lagged hashrate growth rate. This set of lag orders minimizes the Akaike information criterion as well. To estimate Eq. (3), we use the *ARDL* package used in Natsiopoulou and Tzeremes (2022). As in Table 2, the coefficients on the hashrate growth with no lag, lag order one, and lag order two are significant even under a 0.01 significance level. If the

Table 2 Results of the ARDL model regressing demand growth proxy on hashrate growth

Regressor	Coefficient	Regressor	Coefficient
<i>Intercept</i>	−0.0223 (−1.053)	$\Delta Hashrate_t$	4.7553*** (22.404)
$Demand_{Growth_{t-1}}$	−0.5277*** (−24.691)	$\Delta Hashrate_{t-1}$	1.4801*** (5.205)
$Demand_{Growth_{t-2}}$	−0.4688*** (−20.149)	$\Delta Hashrate_{t-2}$	1.6270*** (5.150)
$Demand_{Growth_{t-3}}$	−0.3282*** (−13.559)	$\Delta Hashrate_{t-3}$	0.8061* (2.454)
$Demand_{Growth_{t-4}}$	−0.3154*** (−12.738)	$\Delta Hashrate_{t-4}$	1.0893** (3.294)
$Demand_{Growth_{t-5}}$	−0.3636*** (−14.862)	$\Delta Hashrate_{t-5}$	0.9605** (2.955)
$Demand_{Growth_{t-6}}$	−0.2006*** (−8.082)	$\Delta Hashrate_{t-6}$	0.3605 (1.158)
$Demand_{Growth_{t-7}}$	0.4088*** (17.016)	$\Delta Hashrate_{t-7}$	−2.4213** (−8.624)
$Demand_{Growth_{t-8}}$	0.2365*** (10.087)	$\Delta Hashrate_{t-8}$	−0.7431** (−3.139)
$Demand_{Growth_{t-9}}$	0.0582** (3.047)		

To take lags into consideration, we use an ARDL model regressing the demand growth proxy on the hashrate growth. The coefficients on $\Delta Hashrate_t, \Delta Hashrate_{t-1}, \Delta Hashrate_{t-2}$ are estimated to be positive and tested to be significant. Inside the parentheses are t-statistics. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively

only significant coefficient was the one on the regressor with no lag, we cannot discern between hashrate growth's impact on demand growth and demand growth's impact on hashrate growth; however, coefficients with lagged regressors are notably tested to be significant. The results of the two models show that the system hashrate positively affects blockchain users' demand. The results indicate that the hashrate movement matters greatly for security. One may counterargue that the users' positive responses toward a higher hashrate growth are not the outcome of security improvement; however, there is no other way in which the hashrate can affect users' demand for blockchain uses. The "hashrate" is designed to sustain the security of a blockchain. It is reasonable to believe that rational users know and respond to the blockchain's security situation.

Conclusion

This study aims to empirically identify the hashrate's impact on a blockchain's security level. We find complex relationships between variables in the Bitcoin ecosystem through the VECM for the Bitcoin hashrate, transaction fee, and price. Instead of using the price and fee as Bitcoin demand growth measures, we use a measure that summarizes the growth of some blockchain demand factors. This measure depends on the hashrate change, showing a strong linear relationship. Furthermore, blockchain users recognize the change in the network's security level caused by the change in the hashrate.

We alert designers of blockchain systems of the importance of the hashrate in ensuring the security of the blockchain with proof-of-work consensus protocol, contributing to the sustainability of financial innovations. Additionally, ongoing research on endogenous relationships within the Bitcoin ecosystem will highly benefit from our discussion in section "[Endogenous relationship within the system](#)."

Our empirical analysis only focuses on the Bitcoin system; however, many permissionless blockchains exist with proof-of-work consensus protocols, such as Litecoin and Bitcoin Cash. This limited focus is the principal limitation of our study. The four factors mentioned above regarding cryptocurrency users are highly related to Bitcoin (Liu and Tsyvinski 2021); thus, our analysis in section "[Hashrate and the blockchain demand](#)" might be unsuitable for some other blockchains. Since Ethereum is another popular blockchain, the four factors might be highly related to the Ethereum; however, Ethereum is a blockchain currently adopting a proof-of-stake protocol.

Abbreviations

AR	Autoregressive
ARDL	Autoregressive distributed lag
DL	Distributed lag
HAC	Heteroscedasticity- and autocorrelation-consistent
OLS	Ordinary least squares
PCA	Principal component analysis
VECM	Vector error correction model

Acknowledgements

We appreciate helpful comments from Zhi Zhuo and J. Leon Zhao (Editors-in-Chief), Gang Kou (Managing Editor-in-Chief), and Alexander Webb (University of Wollongong).

Author contributions

DK: proposal and original idea. DK, DR, RW: conceptualization; DK, DR: methodology; DR: validation; DR: resources; DK, RW: literature review; DR, RW: economic and business implication; DK, DR: writing—original draft preparation; DR, RW: writing—review and editing; RW: discussion; DR: project administration. All authors have read and agreed to the published version of the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2022S1A5A2A01044485).

Availability of data and materials

All the raw data used in this study can be found at Coinmarketcap, Blockchain.com, Nasdaq Data Link, and Blockchair. Any data underlying this article can be shared on reasonable request to the authors.

Declarations**Competing interests**

We declare no competing interests.

Received: 22 February 2023 Accepted: 27 December 2023

Published online: 05 June 2024

References

- Aponte-Novoa FA, Orozco ALS, Villanueva-Polanco R, Wightman P (2021) The 51% attack on blockchains: a mining behavior study. *IEEE Access* 9:140549–140564
- Aspris A, Foley S, Svec J, Wang L (2021) Decentralized exchanges: the “wild west” of cryptocurrency trading. *Int Rev Financ Anal* 77:101845
- Bazzanella D, Gangemi A (2023) Bitcoin: a new proof-of-work system with reduced variance. *Financ Innov* 9(1):1–14
- Bissias G, Levine BN (2020) Bobtail: improved blockchain security with low-variance mining. In: NDSS.
- Budish EB (2022) The economic limits of Bitcoin and anonymous, decentralized trust on the blockchain. University of Chicago, Becker Friedman Institute for Economics Working Paper (83)
- Ciaian P, Kancs DA, Rajcaniova M (2021) The economic dependency of Bitcoin security. *Appl Econ* 53(49):5738–5755
- Easley D, O'Hara M, Basu S (2019) From mining to markets: the evolution of Bitcoin transaction fees. *J Financ Econ* 134(1):91–109
- Eyal I, Sirer EG (2018) Majority is not enough: Bitcoin mining is vulnerable. *Commun ACM* 61(7):95–102
- Goczek Ł, Skliarov I (2019) What drives the Bitcoin price? A factor augmented error correction mechanism investigation. *Appl Econ* 51(59):6393–6410
- Gundaboina L, Badotra S, Bhatia TK, Sharma K, Mehmood G, Fayaz M, Khan IU (2022) Mining cryptocurrency-based security using renewable energy as source. *Secur Commun Netw* 2022:1–13
- Heilman E, Kendler A, Zohar A, Goldberg S (2015) Eclipse attacks on Bitcoin's peer-to-peer network. In: Proceedings of the 24th USENIX security symposium, pp 129–144
- Huberman G, Leshno JD, Moallemi C (2021) Monopoly without a monopolist: an economic analysis of the Bitcoin payment system. *Rev Econ Stud* 88(6):3011–3040
- Ibikunle G, McGroarty F, Rzayev K (2020) More heat than light: Investor attention and Bitcoin price discovery. *Int Rev Financ Anal* 69:101459
- Irresberger F, John K, Mueller P, Saleh F (2021) The public blockchain ecosystem: An empirical analysis. NYU Stern School of Business, New York
- Johansen S (1991) Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* 1991:1551–1580
- John K, O'Hara M, Saleh F (2022) Bitcoin and beyond. *Annu Rev Financ Econ* 14:95–115
- Kim D, Ryu D, Webb RI (2023) Determination of equilibrium transaction fees in the Bitcoin network: a rank-order contest. *Int Rev Financ Anal* 2023:102487
- Koch S, Dimpfl T (2023) Attention and retail investor herding in cryptocurrency markets. *Financ Res Lett* 51:103474
- Kubal J, Kristoufek L (2022) Exploring the relationship between Bitcoin price and network's hashrate within endogenous system. *Int Rev Financ Anal* 84:102375
- Kukacka J, Kristoufek L (2023) Fundamental and speculative components of the cryptocurrency pricing dynamics. *Financial Innovation* 9(1):61
- Lee S, Kim S (2020) Proof-of-stake at stake: predatory, destructive attack on PoS cryptocurrencies. In: Proceedings of the 3rd workshop on cryptocurrencies and blockchains for distributed systems, pp 7–11
- Lehar, A., & Parlour, C. A. (2022). Miner collusion and the Bitcoin protocol. SSRN 3559894.
- Liu Y, Tsyvinski A (2021) Risks and returns of cryptocurrency. *Rev Financ Stud* 34(6):2689–2727
- Makarov I, Schoar A (2022) Blockchain analysis of the Bitcoin market. SSRN 3942181
- Marmora P (2022) Does monetary policy fuel Bitcoin demand? Event-study evidence from emerging markets. *J Int Financ Markets Inst Money* 77:101489
- Nasir MA, Huynh TLD, Nguyen SP, Duong D (2019) Forecasting cryptocurrency returns and volume using search engines. *Financ Innov* 5(1):1–13
- Natsiopoulou K, Tzeremes NG (2022) ARDL bounds test for cointegration: Replicating the Pesaran et al. (2001) results for the UK earnings equation using R. *J Appl Econom* 37(5):1079–1090
- Newey WK, West KD (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55(3):703–708
- Noda S, Okumura K, Hashimoto Y (2022) An economic analysis of difficulty adjustment algorithms in proof-of-work blockchain systems. SSRN 3410460
- Pagnotta ES (2022) Decentralizing money: Bitcoin prices and blockchain security. *Rev Financ Stud* 35(2):866–907

- Saad M, Spaulding J, Njilla L, Kamhoua C, Shetty S, Nyang D, Mohaisen D (2020) Exploring the attack surface of blockchain: a comprehensive survey. *IEEE Commun Surv Tutor* 22(3):1977–2008
- Schinckus C, Nguyen CP, Chong FHL (2022) Cryptocurrencies' hashrate and electricity consumption: evidence from mining activities. *Stud Econ Finance*
- Shanaev S, Shuraeva A, Vasenin M, Kuznetsov M (2020) Cryptocurrency value and 51% attacks: evidence from event studies. *J Altern Invest* 22(3):65–77
- Shao E, Rajapaksa D (2023) Miner competition and transaction fees. SSRN 3867552
- Song Y, Chen B, Wang X-Y (2023) Cryptocurrency technology revolution: are Bitcoin prices and terrorist attacks related?. *Financ Innov* 9:29. <https://doi.org/10.1186/s40854-022-00445-3>
- Sun Y, Amanda C, Centana BC (2023) The effect of hashrate, transaction volume, social media and macroeconomics on Bitcoin before and during the COVID-19 pandemic. *Asian J Account Res* 8(3):293–306
- Zhang S, Lee JH (2019) Double-spending with a Sybil attack in the Bitcoin decentralized network. *IEEE Trans Ind Inf* 15(10):5715–5722

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.