Ransomware Activity and Blockchain Congestion

Konstantin Sokolov*

Abstract

Theory predicts that a cryptocurrency may fail if blockchain congestion causes users to exit. I examine this prediction using congestion episodes caused by more than 9,000 triggers for ransomware attacks over a four-year period. Congestion leads to exit by some users due to increased transaction costs. Such users often migrate from the blockchain to crypto exchanges. Migration places a constraint on mining rewards, potentially leading to the failure of a cryptocurrency in the long run.

*Correspondence: ksokolov@memphis.edu

I thank Amber Anand, Jonathan Brogaard, Eric Budish, Dobrislav Dobrev, Sean Foley, Joel Hasbrouck, Jiasun Li, Emiliano Pagnotta, Tālis Putniņš, Gideon Saar, Andriy Shkilko, Hao Zhang, and the audiences at the Sydney Market Microstructure Meeting, Toronto FinTech Conference, University of Memphis, and University of Mississippi for their insightful comments. Aldwin Maloto from the Information Security Office at the Rochester Institute of Technology provided detailed guidance on the interpretation of cybersecurity data.

1. Introduction

Block size limit is often perceived as a major blockchain inefficiency. When the number of pending transactions exceeds the maximum block size, blockchain users incur congestion costs. There is, however, no technological limit to block size.[1] Instead, the designer of a blockchain specifies the block size limit. On the one hand, such a limit must be low enough to induce congestion, and thereby engage users in fee competition. These fees provide blockchain miners with compensation for maintaining a reliable blockchain. On the other hand, the block size limit must not impose excessive congestion costs on blockchain users, as some users may choose alternative means of transaction settlement in response. This may lead to the failure of a cryptocurrency if the fees collected from the remaining blockchain users are insufficient to maintain a reliable mining infrastructure. In this paper, I examine how blockchain users respond to exogenous shocks to congestion and explore the common blockchain limitations predicted by the theory.

I rely on the theory of Easley, O'Hara, and Basu (2019) and address implications that could not be explored in their empirical setup. Specifically, Easley, O'Hara, and Basu (2019) show empirically that the number of fee-paying users and transaction fees increase when more users have entered the line for blockchain settlement (also known as the memory pool). Their empirical model assumes that the memory pool size and the number of transactions is exogenous to congestion. I relax this assumption and explore the decisions made by exogenous (ransomware attack-related) and endogenous (ransomware attack-unrelated) blockchain users. This allows testing of the implications of the theory about the exit of endogenous users during the inflow of exogenous ones.

[1] Blockchain protocol can be modified to eliminate congestion. For example, Red Belly Blockchain can process up to 700,000 transactions per second: http://redbellyblockchain.io/Benchmark.html

Consistent with the theory of Easley, O'Hara, and Basu (2019), I find that the endogenous demand for blockchain settlement shrinks during positive shocks to the exogenous demand. Nevertheless, endogenous demand does not typically completely offset exogenous demand shocks. Therefore, the total number of transactions increases during exogenous demand shocks resulting in congestion.

A congested blockchain makes users choose between two options. The first is to compete for execution priority by paying higher transaction fees. Such equilibrium has been proposed by Huberman, Leshno, and Moallemi (2019), Easley, O'Hara, and Basu (2019), and Iyidogan (2019). The second option is to allow more time for transaction confirmation without engaging in competition for execution priority.

The data reveals that the average blockchain user chooses the first option. During congestion episodes, an average fee-paying user has to wait longer for a transaction to be included in the blockchain. The average fee paid per transaction increases by 11.5%, which suggests that users engage in competition for execution priority.

A block miner receives the sum of fees attached to transactions in the block as compensation for mining. The stronger the congestion, the higher the compensation the miner should expect. Transaction fees are, however, not the only source of mining reward. Several coins are minted and distributed to miners when a block is found. The majority of blockchains have caps on the total number of coins minted, and the number of newly minted coins decreases over time. Thus, the congestion-induced reward (CIR) becomes the primary source of mining compensation for miners in the long run.

Huberman, Leshno, and Moallemi (2019) emphasize the role of congestion in creating a reward for mining. Specifically, a mining reward should be high enough to ensure blockchain

reliability. Nevertheless, the CIR alone may be insufficient to ensure such reliability. In other words, users may sooner switch to alternative means of transaction processing than pay fees that are sufficient to provide adequate compensation to a miner for investment in the reliable mining infrastructure. In line with this, Abadi and Brunnermeier (2018) argue that a large mining reward renders the risk of a blockchain failure negligible. Budish (2018) shows theoretically that reliable blockchains should offer a mining reward equal to at least 30% of the highest-value transactions that are possible through the system. Although the data does not allow determining what level of the CIR is sufficient to ensure blockchain reliability, I can estimate the upper bound to the fees per fixed size block that congested users agree to pay. This brings the literature one step closer to answering the question whether the blockchain can be reliable over the long term.

I find that such an upper bound exists. Specifically, the CIR does not typically exceed $2,800 per Bitcoin block, even when the demand for Bitcoin blockchain settlement becomes extremely high.

Finally, I look at the market for the cryptocurrency underlying the congested blockchain. Specifically, I explore how traders on Gemini Exchange react to blockchain congestion. Gemini fits the purpose of this study well because it is very unlikely to be affected by ransom payments mechanically. It is not well suited to processing transactions initiated by ransom payers or payees. The ransomware countdown timer typically allows up to three days for payment, while it takes at least four days to withdraw cryptocurrency from Gemini.[2] This makes Gemini unattractive for ransomware attack victims who are pressured to acquire and pay cryptocurrency immediately. Gemini also does not allow the anonymous withdrawal of

[2] https://medium.com/gemini/instant-ach-deposits-are-here-795c9bdbac1

funds, which makes it unattractive for cybercriminals. As such, it is unlikely that cryptocurrency trading on Gemini is mechanically affected by ransom payments.

Cryptocurrency exchanges, such as Gemini, do not rely on the blockchain for trade settlement within the exchange. A coin circulating on the exchange remains in the centralized exchange wallet until one of the traders decides to withdraw it. Blockchain settlement is required only when a coin is transferred between the exchange and the traders. Within the exchange, coins circulate with sub-microsecond latency between traders' accounts.

Although the exchange architecture allows congestion-free settlement, it is vulnerable to cybersecurity threats. A hacker may compromise the centralized exchange wallet and steal the coins. According to the SEC (2018), wallet thefts cause significant losses to coin investors. For instance, such thefts occurred when hackers compromised the centralized wallets of Mt. Gox (2011-2013) and Coincheck (2018). Biais, Bisiere, Bouvard, Casamatta, and Menkveld (2018) argue that coin thefts constitute a fundamental risk for cryptocurrency investors.

The discovery of software vulnerabilities may increase the chance of theft from the exchange. Although individual wallets are also subject to theft, they are much less attractive to hackers because the resources invested in stealing the wallets and decoding individual passwords may exceed the value of the coins. A cryptocurrency exchange, however, cannot efficiently protect itself from theft by splitting coin reserves into numerous wallets. According to Jain, Felten, and Goldfeder (2018), the cybersecurity risk associated with coin transfer within exchange wallets often outweighs the risk of keeping a large number of coins in a centralized hot wallet.

On the one hand, a cryptocurrency exchange with a centralized wallet may attract a higher demand when ransomware activity leads to blockchain congestion. On the other hand,

traders have the incentive to avoid such an exchange when the coin theft risk increases. The empirical investigation suggests that ransomware activity leads to an increase in exchange coin trading despite the coin theft risk. On average, the exchange volume increases by 10.4% with one standard deviation shock to the number of severe vulnerabilities. Exchange transaction costs measured by the effective spread and implementation shortfall appear to be insensitive to congestion.

The remainder of the paper is as follows. Section 2 discusses the relationship between ransomware attacks and the demand for blockchain settlement. Section 3 describes the data and variables. Section 4 discusses the main results. Section 5 reports robustness tests, and Section 6 concludes.

2. Ransomware attacks and demand for blockchain transactions

Counterparties enjoy anonymity under blockchain technology. A strand of literature has raised concern that such an opaque transaction system has high social cost (Malinova and Park, 2017; Ma, Gans, and Tourky, 2018; Williamson, 2018) and facilitates illegal activities (Brill and Keene, 2014; Böhme, Christin, Edelman, and Moore, 2015; Brown, 2016). For example, Athey, Parashkevov, Sarukkai, and Xia (2016) find that a substantial amount of identifiable Bitcoin blockchain addresses are involved in ransomware money processing, contraband, gambling, and money laundering. Foley, Karlsen, and Putniņš (2019) show that half of all Bitcoin transactions are associated with illegal activities. In fact, Bitcoin proves to be very sensitive to darknet market disruptions.[3]

Hackers take advantage of system vulnerabilities to deploy ransomware on victims' computers. Once ransomware is deployed, it encrypts files with certain extensions such as .txt,

[3] https://www.theguardian.com/technology/2013/oct/03/bitcoin-price-silk-road-ulbricht-value

.pdf, and .jpeg. The purpose of the selective file encryption is to limit a victim's ability to access and modify valuable information while leaving the system bootable. When the reboot happens, ransomware invokes a message specifying the amount of ransom, the hacker's Bitcoin wallet address, and the time remaining for the victim to pay the ransom (see Figure 1). If the victim pays the ransom, the hacker may either release the password for file decryption or continue demanding even higher ransom, thus further increasing the number of transactions if the victim agrees to pay.

[Figure 1]

The life of a severe vulnerability is fleeting. Software vendors rush to come up with updates to patch the holes through which ransomware infiltrates. The short life of exploitable vulnerabilities incentivizes hackers to act opportunistically and attempt large-scale ransomware attacks despite potential frictions due to blockchain congestion. The number of such ransomware attacks may reach 638 million per year.[4] The victims of ransomware attacks often choose to pay the ransom to restore the encrypted files. For instance, Choi, Scott, and LeClair (2016) report that 85% of U.S. police departments affected by ransomware attacks paid the ransom to hackers. As a result, demand for blockchain transactions spikes when hackers discover severe vulnerabilities.

According to Athey et al. (2016), CoinVault and CryptoLocker claimed ransom through the Bitcoin blockchain. Although these pieces of ransomware were active for a relatively short period, their average daily transaction rates ended up among the highest in the sample. Paquet-Clouston, Haslhofer, and Dupont (2019) document that Bitcoin addresses receiving ransom payments experience intense yet short-term spikes of activity. Consistent with the above

[4] https://www.forbes.com/sites/leemathews/2017/02/07/2016-saw-an-insane-rise-in-the-number-of-ransomware-attacks/#3485ad2458dc

results, I find that a discovery of severe vulnerabilities typically ignites demand for Bitcoin blockchain settlement. Figure 2 shows that shocks to the number of newly-discovered severe vulnerabilities are accompanied by a visible spike in Bitcoin blockchain transactions. I use these shocks to test the predictions of theory literature about the implications of blockchain congestion.

[Figure 2]

3. Data and variables

3.1 Blockchain and exchange data

The primary data on blockchain comes from blockchain.info. This source has been used by Biais, Bisière, Bouvard, and Casamatta (2019), Easley, O'Hara, and Basu (2019), and Pagnotta and Buraschi (2018). The data consists of the daily metrics of the Bitcoin blockchain. Hackers typically choose the Bitcoin blockchain as a means of ransom processing due to its anonymity, popularity, and reliability. Although some ransomware may allow victims to pay through other blockchains, the Bitcoin blockchain is the most popular blockchain for ransom payments. According to Symantec, ransomware almost exclusively asks for payment using Bitcoins.[5]

Bitcoin experienced a sharp spike in speculative demand at the end of 2017. To avoid any impact of this spike, I exclude this period from the sample. The resulting sample spans from January 2014 to November 2017. I obtain the limit order book data from Gemini – a US-based cryptocurrency exchange.

One of the most important properties of Gemini is that it is unlikely to be affected by ransom processing mechanically. In contrast to some other active cryptocurrency exchanges,

[5] https://www.symantec.com/connect/blogs/dawn-ransomwear-how-ransomware-could-move-wearable-devices

Gemini does not allow anonymity of counterparties. This makes it unattractive to hackers who need to convert collected ransom into cash. The Gemini design also does not allow quick withdrawal of cryptocurrency, which prompts ransomware attack victims to purchase cryptocurrency elsewhere.

Among all cryptocurrency exchanges, Gemini comes the closest to the conventional electronic trading venues. Gemini implements the Market Data Integrity Policy and distributes the limit order book history through the Chicago Board Options Exchange (CBOE). Moreover, it does not have deposit/withdrawal fees unlike other cryptocurrency exchanges. This makes it even closer to conventional trading venues. The limit order book data covers the interval from January 2016 to November 2017.

3.2 *Variables of interest*

In pursuit of anonymity, hackers typically do not use the same address to collect all ransom payments (Kshetri and Voas, 2017). Moreover, ransom often travels through multiple wallets before the hackers cash it out (Foley, Karlsen, and Putniņš, 2019). As such, regular blockchain addresses are unlikely to be involved in ransom processing. Thus, transactions involving these addresses reflect the endogenous demand for blockchain settlement. The remaining transactions reflect both endogenous and exogenous demand (ransom processing). I split all blockchain transactions into transactions involving regular and ad hoc addresses:

$$nTransAll_t = nTransRegular_t + nTransAdHoc_t,$$

where $nTransRegular_t$ is the daily number of transactions involving the blockchain's one hundred most popular addresses, and $nTransAdHoc_t$ is the number of transactions excluding the regular addresses.

The result of the above decomposition is reported in Table 1, Panel A. On an average day, regular addresses are involved in 7.7 thousand transactions, while 171.5 thousand transactions are processed through ad hoc addresses.

[Table 1]

Blockchain congestion occurs when the number of transactions lined up for confirmation exceeds the maximum block size. When the average number of transactions per block is increasing, there is a higher potential for congestion to occur. I estimate the average daily number of transactions per block as follows:

$$nTransPerBlock_t = \frac{nTransAll_t}{Blocks_t},$$

where $Blocks_t$ is the number of blocks mined on the day t.

The remaining blockchain variables capture the congestion outcome. $ConfTime_t$ measures the average daily processing time of the fee-paying transactions, while $TransFeeBTC_t$ and $FeesPerBlockBTC_t$ show the fees per transaction and the total value of the fees attached per block, respectively.

Limit order book data allows to determine whether blockchain congestion affects the market for underlying cryptocurrency. Specifically, I test whether liquidity providers demand higher compensation when the blockchain is congested. I use two estimates of transaction costs: effective spread and implementation shortfall. Effective spread is defined as follows:

$$EffectiveSpread_t = \frac{2q_t(p_t - mid_t)}{mid_t},$$

where p_t is the execution price, q_t is the trade direction indicator equal to one for buyer-initiated trades and negative one for seller-initiated trades, and mid_t is the average of the best

bid and ask quotes at time t. The intraday effective spread is time-weighted to obtain daily estimates.

The implementation shortfall shows how much a price yields to an order. According to Hendershott, Jones, and Menkveld (2013), the implementation shortfall constitutes the substantial component of execution costs. For every market order, I measure the implementation shortfall as follows:

$$ImpShortfall_t = \frac{2q_t(p_{0,t} - \bar{p_t})}{p_{0,t}},$$

where $p_{0,t}$ is the price of the first quote hit by a market order and $\bar{p_t}$ is the average execution price of a market order. I average the implementation shortfall across all trades to obtain the daily estimates.

Panel B of Table 1 reports limit order book summary statistics. Execution costs are rather low in the sample. The mean effective spread is less than 2bps. On average, the implementation shortfall accounts for 25% of the effective spread.

3.3 Vulnerability data

The data on vulnerabilities comes from the National Vulnerability Database (nvd.nist.gov) – a U.S. government repository of cybersecurity threats. The database is populated with the intention to list all publicly known vulnerabilities and exposures.[6] The National Vulnerability Database (NVD) distributes newly listed vulnerabilities through the public data feed.

The database administration acknowledges that public dissemination of newly discovered vulnerabilities may help hackers exploit them. Despite this, the cybersecurity

[6] http://cve.mitre.org/about/faqs.html#cve_list_contain_all_vulnerabilities

community believes that the benefits of such dissemination outweigh the risks. Numerous cybersecurity-related organizations (including commercial security tool vendors, academia, research institutions, government departments and agencies, and end-users of vulnerability information) support sharing information on vulnerabilities.[7]

In addition to the identification, collection, and dissemination of vulnerabilities, the NVD examines the severity of the newly discovered vulnerabilities. The vulnerabilities are scored by their severity on a scale of 1 to 10 according to the Common Vulnerability Scoring System (CVSS). The score depends on a number of factors such as the vulnerability complexity, the potential to compromise the system, and the attacker's visibility. I discuss further details of the CVSS scoring system in the Appendix.

The NVD classifies vulnerabilities with a score above 7.0 as highly severe. The discovery of such vulnerabilities is indicative of ransomware attacks. Figure 3 shows that there is an increase in the number of severe vulnerabilities around the day the new ransomware is listed in the Symantec dictionary.[8] The data on vulnerabilities, however, has several advantages over the Symantec dictionary. First, a ransomware attack may not be successful, as it is the severity of the exploited vulnerability that defines the hacker's success. Second, some older ransomware can be recycled to exploit newly discovered vulnerabilities. Third, Symantec does not always make a distinction between ransomware and trojans. Thus, I only use data on severe vulnerabilities in the further analysis.

[Figure 3]

Figure 4 shows the daily number of severe vulnerabilities discovered in the sample period (9,397 in total). Table 2 reports the breakdown of these vulnerabilities by software

[7] http://cve.mitre.org/about/faqs.html#hackers_break
[8] https://www.symantec.com/security-center/a-z

vendors. Although the total number of software vendors exceeds one thousand, the distribution of vulnerabilities is rather skewed. Almost two thirds of all severe vulnerabilities are found in the products of only ten most popular software vendors. Among these top ten, Microsoft, Google, and Adobe are responsible for one third of severe vulnerabilities. This allows ransomware to spread rapidly within the network of common software users (Goyal and Vigier, 2014 and Acemoglu, Malekian, and Ozdaglar, 2016). As such, severe vulnerabilities typically allow hackers to demand ransom from a massive number of victims.

[Figure 4]

[Table 2]

4. Results

4.1 Demand for blockchain settlement

An exogenous shock to demand for blockchain settlement may not lead to congestion if the endogenous users cease their demand substantially. Theory models predict that blockchain users will resort to alternative means of payment when congestion costs increase. This reaction may mitigate the total demand and ease the congestion. I examine how vulnerability releases impact the total and the endogenous demand for blockchain transactions. Specifically, I estimate the following linear model:

$$DepVar_t = \beta_1 Vuln_t + \beta_2 BTC_t + \varepsilon_t,$$

where $DepVar_t$ is one of three variables: the total number of blockchain transactions, transactions involving regular addresses, and transactions involving ad hoc addresses; $Vuln_t$ is the number of severe vulnerabilities; and BTC_t is the price of the Bitcoin cryptocurrency. This variable controls for the non-linear upward trend in demand for the Bitcoin blockchain

settlement. All regressions are estimated with standardized and demeaned variables. Thus, the intercept term is omitted from the model.

The results in Table 3, Column 1, show that the total number of blockchain transactions increases with the number of exploitable vulnerabilities. Specifically, a one standard deviation shock to the number of severe vulnerabilities corresponds to the 5.8% increase in blockchain transactions. As such, the total demand for blockchain settlement increases with an increase in exogenous demand.

[Table 3]

Next, I examine whether there is any evidence of endogenous users forgoing blockchain settlement when the exogenous demand increases. To avoid detection, hackers prefer to collect money to multiple ad hoc addresses instead of using the same address for all ransom processing. Moreover, they often try to hide their activity by moving money between multiple wallets of their own before cashing out. This further increases the demand for blockchain transactions. In contrast, regular blockchain addresses are unlikely to be involved in ransom collection. Therefore, transactions through regular addresses represent the endogenous demand for blockchain settlement.

I decompose all transactions into those involving regular and ad hoc addresses. The results of this decomposition are reported in Table 3, Columns 2 and 3. Vulnerability releases have the opposite effect on the number of transactions involving ad hoc and regular addresses. Consistent with the theory predictions, the endogenous demand from regular addresses declines when the demand from ad hoc addresses increases. Nevertheless, the coefficient on the endogenous demand has lower economic and statistical significance. This suggests that

some endogenous blockchain users choose to continue running transactions through blockchain despite the potential congestion costs.

4.2 Congestion costs

Blockchain congestion occurs when demand for blockchain transactions exceeds blockchain capacity. Specifically, the number of transactions recorded per block has a limit. When the number of transactions exceeds this limit, transaction initiators face a choice between two options.

The first option is to attach a fee to the transaction. This fee will increase the transaction's priority. According to Aune, Krellenstein, O'Hara, and Slama (2017), blockchain users have the incentive to attach a fee, because waiting costs are magnified by the risk of front-running. This fee, however, does not guarantee that the transaction is written into the next block. Transaction initiators can further increase this priority by paying even higher fees. Importantly, the limit on transactions per block does not change with the total value of attached fees. As such, even if all transaction initiators choose to attach very high fees, the congestion will remain unresolved.

The second option is to allow more time for transaction processing without attaching a fee. Although waiting costs incentivize an individual user to attach the fee, the user may decide to forgo the fee to avoid triggering other users to attach even higher fees. In line with this, Pappalardo, Matteo, Caldarelli, and Aste (2017) estimate that the waiting time for 20% of Bitcoin blockchain transactions exceeds thirty days.

The results are consistent with the predictions of the theory models that, on average, users prefer the first option to the second one. Table 4 shows that transaction costs increase by

11.5% when the number of released vulnerabilities increases by one standard deviation. Thus, blockchain users react to congestion by increasing the transaction costs. The increase in transaction costs, however, does not resolve the congestion. An average fee-paying user ends up waiting longer despite the increase in transaction costs.

[Table 4]

4.3 Mining reward

Competitive mining prevents double spending (Abadi and Brunnermeier, 2018; Biais et al., 2019; Budish, 2018 and Pagnotta, 2018) and settlement fails (Chiu and Koeppl, 2018). Although transaction fees increase during congestion, this finding does not imply that transaction fees alone are sufficient to ensure blockchain reliability. It may well be that no attainable level of congestion is sufficient to maintain blockchain reliability if all mining reward comes from fees.[9] Budish (2018) and Huberman, Leshno, and Moallemi (2019) suggest that a blockchain can only be viable when the equilibrium congestion-induced reward (CIR) exceeds the minimum amount required for reliable mining. Although the data does not allow determining what level of CIR is sufficient to ensure blockchain reliability, I can estimate the limit to the fees per block that congested users agree to pay. The limit is reached when a further increase in ransomware activity no longer increases the CIR. I use the following model to find whether such a limit exists:

$$CIR_t = \alpha + \beta_1 Vuln1_t + \beta_2 Vuln2_t + \beta_3 Vuln3_t + \beta_4 Vuln4_t + \beta_5 Vuln5_t + \beta_6 Vuln6_t + \beta_7 PriceBTC_t + \varepsilon_t$$

[9]Cryptocurrency blockchain miners receive compensation in two forms. First, blockchain rewards miners with a number of new coins minted when a block is mined. Second, blockchain users compete for execution priority by paying fees to miners. The number of new coins minted is typically diminishing over time. Therefore, transaction fees will become the primary source of mining reward in the long run.

where CIR_t is the mean daily mining reward per block coming from transaction fees, and $Vuln1_t - Vuln6_t$ are dummy variables equal to 1 if the number of severe vulnerabilities lies within a certain interval. Specifically, $Vuln1_t = 1$ if $0 < N \leq 3$, $Vuln2_t = 1$ if $3 < N \leq 10$, $Vuln3_t = 1$ if $10 < N \leq 20$, $Vuln4_t = 1$ if $20 < N \leq 30$, $Vuln5_t = 1$ if $30 < N \leq 40$, and $Vuln6_t = 1$ if $N > 40$. $PriceBTC_t$ is the Bitcoin price. The intercept corresponds to the days when no severe vulnerabilities are released.

Table 5 contains the model coefficients. Consistent with the findings on transaction costs, the CIR increases with the number of severe vulnerabilities. However, the coefficients β_4, β_5, and β_6 are statistically equal. As such, congestion can increase the mining reward only to a certain limit. I show the economic significance of this limit in Figure 5. On average, congestion does not drive the CIR above $2,800.

This result should be interpreted with caution. First, the costs of alternative means of transaction processing may decrease over time, thereby lowering the expected CIR limit per block. Second, fully endogenous demand may be less constrained in the selection of transaction processing means than demand consisting of both endogenous and exogenous components. This may further decrease the expected CIR limit.

[Table 5]

[Figure 5]

4.4 Cryptocurrency market

Cryptocurrency exchanges do not typically process internal transactions through the blockchain. For instance, Gemini is registered as a fiduciary and keeps its own record of changes in coin ownership. The coins circulating within the exchange remain in the centralized

exchange wallet until the traders decide to withdraw them. This practice allows coin traders to avoid blockchain congestion costs and trade with sub-microsecond latency. Centralized exchange wallets are, however, vulnerable to cybersecurity breaches. There have been multiple instances when centralized exchange wallets have been hacked and coins stolen. Although individual wallets are also subject to theft, they are much less attractive to hackers because the resources invested in stealing the wallets and decoding the individual passwords may exceed the value of coins. A cryptocurrency exchange, however, can not eliminate cybersecurity risk by splitting funds between numerous wallets, because these wallets will remain under supervision of the same entity. Moreover, frequent transfers of cryptocurrency between the exchange wallets can cause even higher cybersecurity risks than keeping substantial funds in the centralized hot wallet (Jain, Felten, and Goldfeder, 2018).

When an increase in ransomware activity causes blockchain congestion, coin traders face a choice between two options. The first option is to transfer their coins to the centralized exchange wallet and enjoy congestion-free settlement despite the coin theft risk. The second option is to avoid such exchanges and continue trading with individual wallets despite the congestion costs. If more coin traders prefer the first option to the second one, then the demand for exchange settlement should increase with ransomware activity.

The results in Table 6 reveal that the demand for exchange cryptocurrency trading increases with congestion despite the coin theft risk. Both the volume and number of transactions grow by 10.4% and 9%, respectively, with one standard deviation shock to the number of severe vulnerabilities. Moreover, it appears that market makers do not typically demand higher compensation for liquidity provision when more vulnerabilities are discovered.

Shocks to severe vulnerabilities do not have a visible impact on the effective spread and implementation shortfall.

[Table 6]

5. Robustness

5.1 Regression in first differences

In this section, I address a concern about potential seasonality and persistence in the sample variables. Hodrick and Prescott (1997) and Novy-Marx (2014) show that regressions on cyclical and persistent data may produce spurious results. I address this issue by running the main tests in first differences. The first difference regression specification is as follows:

$$\Delta DepVar_t = \alpha + \beta_1 \Delta Vuln_t + \beta_2 \Delta BTC_t + \varepsilon_t,$$

where $\Delta DepVar_t$ is the daily log difference in the number of transactions involving all blockchain addresses; the number of transactions involving the addresses other than one hundred most popular network addresses; the number of transactions involving one hundred most popular addresses; and the volume and number of trades on the Gemini exchange. The independent variable is adjusted to capture the most significant shocks to the number of severe vulnerabilities. Specifically, $\Delta Vuln$ is the dummy variable equal to one if the daily difference in the number of newly discovered severe vulnerabilities exceeds five. ΔBTC_t is the log return on Bitcoin cryptocurrency.

The coefficients in Table 7 are consistent with the main sample specification. As such, the results are unlikely to be driven by a spurious correlation in levels. The economic significance of the coefficients, however, should be interpreted with caution due to the natural persistence of blockchain congestion. First, the congestion may persist for some time after the

demand shock caused by ransomware activity is exhausted. Second, it may occasionally take considerable time to address the vulnerability allowing for ransomware infiltration. Greenwood, Hanson, and Stein (2010) and Cochrane (2018) show that differencing such variables leads to loss of information due to reduction in data variation. Thus, although the results of first difference regressions are important to rule out the possibility of spurious correlation in levels, they are unsuitable to replace the main analysis.

[Table 7]

5.2 Ransomware-unrelated cybersecurity risks

The results on blockchain congestion will still hold even if there exists an alternative explanation of how vulnerabilities may create shocks to demand for blockchain settlement. I address this alternative because it may carry a different economic interpretation. Specifically, coin theft risk may incentivize users to sell their coin holdings, thus increasing the number of blockchain transactions.

I find this explanation rather unlikely. First, there appears to be no relationship between the discovery of severe vulnerabilities and Bitcoin price volatility (the results are available on request). Second, the results in Table 6 indicate that the bid-ask spread and implementation shortfall do not increase with the number of newly discovered severe vulnerabilities. This suggests that vulnerabilities do not systematically tighten up constraints on the cryptocurrency market maker inventory.

To further address this concern, I run the main analysis with the sample excluding the days around the discovery of Bitcoin-related vulnerabilities and major coin thefts. I obtain the release dates of Bitcoin-related vulnerabilities by screening vulnerability descriptions in the

NVD database for the word "Bitcoin." The coin theft dates come from magoo.github.io/Blockchain-Graveyard – a website tracing the major coin theft events. I remove three days around the dates of the coin thefts and Bitcoin-related vulnerability releases from the sample. In total, 91 non-overlapping days are excluded from the sample. According to Panel A of Table 8, the main results are not affected by the elimination of the days around the coin thefts and Bitcoin-related vulnerability releases.

[Table 8]

5.3 Severe vs. weak vulnerabilities

Next, I test whether vulnerability severity matters. Weak vulnerabilities typically have low exploitability and should not cause significant shocks to ransomware activity. The NVD classifies vulnerabilities with a score below 4.0 as low-severity vulnerabilities. I use these weak vulnerabilities to examine whether my results are indeed driven by ransomware activity rather than by commonalities between the patterns of blockchain congestion and the NVD vulnerability releases.

Naturally, cybersecurity experts spot weak and severe vulnerabilities at the same time. As a result, the discovery of severe and weak vulnerabilities is correlated ($\rho = 0.46$). Despite this correlation, the results in Panel B of Table 8 indicate that weak vulnerabilities do not cause shocks to blockchain or cryptocurrency markets. Specifically, weak vulnerabilities do not have significant explanatory power on top of severe vulnerabilities. Thus, the results are consistent with the main finding that ransomware activity causes shocks to the demand for blockchain settlement.

5.4 Placebo test

Finally, I estimate the likelihood of false results. To do so, I generate multiple placebo samples of randomly drawn severe vulnerabilities and use these samples in the main regression setup. Specifically, I use the following procedure: First, the series of severe vulnerabilities are randomly permuted. Then, the main regression runs with the placebo sample of vulnerabilities as an independent variable. Next, the value 1 is assigned if the placebo coefficient sign is consistent with the sign of the coefficient in the original regression and 0 otherwise. Finally, the statistical significance of the placebo coefficient is captured. The above procedure runs 1,000 times.

Table 9 reports the percentage of placebo coefficients consistent with the original ones. Among the 1,000 runs, there are no instances when all placebo coefficients are consistent with the original ones at the 10% significance level (Panel A). A separate estimation with 100,000 runs shows only two such instances. As such, the joint probability that placebo coefficients are consistent with the results is rather low. The instances when coefficients are individually significant and consistent with the original ones are also rare (Panel B).

[Table 9]

6. Conclusion

A crucial problem preventing blockchain technology from replacing conventional means of electronic payment is the inherent capacity limit. Conventional electronic payment networks can increase capacity and reduce the risk of congestion without compromising reliability. In contrast, blockchains may become unreliable in the absence of congestion. As a result, congestion costs will remain a unique feature differentiating established blockchains from other means of electronic transaction settlement.

I study how blockchain users respond to cybersecurity-related shocks to demand for blockchain settlement. Consistent with the theory predictions, some endogenous users forgo blockchain settlement when exogenous demand arrives. The remaining users engage in competition for execution priority and raise the level of transaction fees. An increase in transaction fees also implies an increase in mining reward. Mining reward, however, does not increase with congestion infinitely. I document the existence of the upper bound on the congestion-induced reward for mining. Although the empirical investigation relies on a cryptocurrency blockchain data, the findings are relevant to any blockchain settlement system that requires congestion to support the mining infrastructure. Finally, I find that demand for cryptocurrency exchange congestion-free settlement increases during blockchain congestion shocks despite the risk of theft from the centralized exchange wallet.

References

Abadi, J. and Brunnermeier, M., 2018, Blockchain Economics, Working paper.

Acemoglu, D., Malekian, A. and Ozdaglar, A., 2016, Network Security and Contagion, Journal of Economic Theory 166, 536-585.

Aune, R., Krellenstein, A., O'Hara, M. and Slama, O., 2017, Footprints on a Blockchain: Trading and Information Leakage in Distributed Ledgers, Journal of Trading 12, 5-13.

Athey, S., Parashkevov, I., Sarukkai, V. and Xia, J., 2016, Bitcoin Pricing, Adoption, and Usage: Theory and Evidence, Working paper.

Biais, B., Bisiere, C., Bouvard, M. and Casamatta, C., 2019, The Blockchain Folk Theorem, Review of Financial Studies 32, 1662-1715.

Biais, B., Bisiere, C., Bouvard, M., Casamatta, C. and Menkveld, A., 2018, Equilibrium Bitcoin Pricing, Working paper.

Böhme, R., Christin, N., Edelman, B. and Moore, T., 2015, Bitcoin: Economics, Technology, and Governance, Journal of Economic Perspectives 29, 213-38.

Brill, A. and Keene, L., 2014, Cryptocurrencies: The Next Generation of Terrorist Financing? Defence Against Terrorism Review 6, 7-30.

Brown, S., 2016, Cryptocurrency and Criminality: The Bitcoin Opportunity, The Police Journal: Theory, Practice and Principles 89, 327-339.

Budish, E., 2018, The Economic Limits of Bitcoin and the Blockchain, Working paper.

Chiu, J. and Koeppl, T., 2018, Blockchain-Based Settlement for Asset Trading, Working paper.

Choi, K., Scott, T. and LeClair, D., 2016, Ransomware Against Police: Diagnosis of Risk Factors via Application of Cyber-routine Activities Theory, International Journal of Forensic Science and Pathology 4, 253-258.

Cochrane, J., 2018, A brief parable of over-differencing, Working paper.

Easley, D., O'Hara, M. and Basu, S., 2019, From Mining to Markets: The Evolution of Bitcoin Transaction Fees, Journal of Financial Economics, forthcoming.

Foley, S., Karlsen, J. and Putniņš, T., 2019, Sex, Drugs, and Bitcoin: How Much Illegal Activity is Financed Through Cryptocurrencies? Review of Financial Studies 32, 1798-1853.

Goyal, S. and Vigier, A., 2014, Attack, Defence, and Contagion in Networks, Review of Economic Studies 81, 1518-1542.

Greenwood, R., Hanson, S. and Stein, J., 2010, A gap-filling theory of corporate debt maturity choice, Journal of Finance 65, 993-1028.

Hendershott, T., Jones, C. and Menkveld, A., 2013, Implementation Shortfall with Transitory Price Effects, Working paper.

Huberman, G., Leshno, J. and Moallemi, C., 2019, An Economic Analysis of the Bitcoin Payment System, Working paper.

Iyidogan, E., 2019, An Equilibrium Model of Blockchain-Based Cryptocurrencies, Working paper.

Jain, S., Felten, E. and Goldfeder, S., 2018, Determining an Optimal Threshold on the Online Reserves of a Bitcoin Exchange, Journal of Cybersecurity 4, 1-12.

Kshetri, N. and Voas, J., 2017, Do Crypto-Currencies Fuel Ransomware? IT Professional 5, 11-15.

Ma, J., Gans, J. and Tourky, R., 2018, Market Structure in Bitcoin Mining, Working paper.

Malinova, K. and Park, A., 2017, Market Design with Blockchain Technology, Working paper.

Pagnotta, E., 2018, Bitcoin as Decentralized Money: Prices, Mining Rewards, and Network Security, Working Paper.

Pagnotta, E. and Buraschi, A., 2018, An Equilibrium Valuation of Bitcoin and Decentralized Network Assets, Working paper.

Pappalardo, G., Di Matteo, T., Caldarelli, G. and Aste, T., 2018, Blockchain Inefficiency in the Bitcoin Peers Network, EPJ Data Science, 7, 1-13.

Paquet-Clouston, M., Haslhofer, B. and Dupont, B., 2019, Ransomware payments in the bitcoin ecosystem, Journal of Cybersecurity 5, 1-11.

SEC, 2018, Chairman's Testimony on Virtual Currencies: The Roles of the SEC and CFTC, February 6, 2018. https://www.sec.gov/news/testimony/testimony-virtual-currencies-oversight-role-us-securities-and-exchange-commission

Sockin, M. and Xiong, W., 2018, A Model of Cryptocurrencies, Working paper.

Williamson, S., 2018, Is Bitcoin a Waste of Resources? Federal Reserve Bank of St. Louis Review 100, 107-15.

Table 1: Bitcoin blockchain and cryptocurrency exchange summary statistics

The table reports descriptive statistics for the Bitcoin blockchain (Panel A) and Gemini cryptocurrency exchange (Panel B). The blockchain sample spans January 2014 to November 2017. The exchange sample spans January 2016 to November 2017. Both samples are limited to business days due to little activity occurring on weekends and holidays. The variables are defined as follows: $nTransAll$ is the number of transactions involving all blockchain addresses; $nTransAdHoc$ is the number of transactions involving the addresses other than one hundred most popular network addresses; $nTransRegular$ is the number of transactions involving one hundred most popular addresses; $TransFeeBTC$ is the average daily fee attached to blockchain transactions; $ConfTime$ is the average time it takes to process a fee-paying transaction; $nTransPerBlock$ is the average daily number of transactions per block; $FeesPerBlockBTC$ is the average value of fees attached per block; $VolumeBTC$ and $nTrades$ is the daily volume and the number of trades on Gemini exchange; $ESpread$ is the effective spread; $ImplShortfall$ is the implementation shortfall defined as twice the signed difference between the corresponding best quote when the market order arrived to the exchange and the realized execution price of the market order.

	Mean	Median	Std
Panel A: Blockchain Statistics			
$nTransAll$	179,281	187,202	89,692
$nTransAdHoc$	171,545	180,232	88,249
$nTransRegular$	7,736	6,150	7,614
$TransFeeBTC$ x 1000	0.370	0.215	0.357
$ConfTime$, minutes	9.567	8.466	3.211
$nTransPerBlock$	1,174	1,147	617
$FeesPerBlockBTC$	0.571	0.224	0.775
#Days	987		
Panel B: Exchange Statistics			
$VolumeBTC$	4,461	2,123	5,863
$nTrades$	3,266	1,148	4,344
$ESpread$, bp.	1.747	1.177	19.478
$ImplShortfall$, bp.	0.438	0.291	0.478
#Days	482		

Table 2: Vulnerability summary statistics

The table reports summary statistics for the sample of severe vulnerabilities (severity score above 7.0) distributed in the data feed of the National Vulnerability Database (NVD). The NVD assigns a severity score on a 1 to 10 scale according to several exploitability factors. The vulnerability scoring system adopted by the NVD classifies vulnerabilities scored above 7.0 as highly severe. The sample spans January 2014 to November 2017. The sample is limited to business days due to the small number of vulnerabilities released on weekends and holidays. Panel A contains the daily mean, median, and standard deviation of the newly released severe vulnerabilities. Panel B reports the breakdown of vulnerabilities by software vendor.

Panel A: Descriptive Statistics				
	Mean	Median	Std	
$Vuln$	9.52	5	15.01	
#Days	987			
Panel B: Breakdown of Vulnerabilities by Software Vendor				
	#Vuln	Vuln Share (%)	Mean Score	Median Score
Microsoft	1,252	13.32%	8.70	9.3
Google	928	9.88%	8.74	9.3
Adobe	836	8.90%	9.64	10
IBM	617	6.57%	8.09	7.5
Apple	526	5.60%	8.85	9.3
Oracle	423	4.50%	8.30	7.6
Cisco	357	3.80%	8.09	7.8
Linux	261	2.78%	8.00	7.6
Huawei	132	1.41%	8.18	7.8
HP	112	1.19%	8.37	7.6
Other	3,953	42.07%	8.14	7.5
#Vendors	1,229			

Table 3: Ransomware attacks and demand for blockchain settlement

The table examines the sensitivity of blockchain transactions to the number of newly released severe vulnerabilities. In the pursuit of anonymity, hackers typically do not use the same address to collect all ransom payments. As such, regular blockchain addresses are unlikely to be involved in ransom processing. Thus, transactions involving these addresses ($nTransRegular$) reflect endogenous demand for blockchain settlement. The remaining transactions reflect both endogenous and exogenous demand ($nTransAdHoc$). I use the following model to formally test the relationship between ransomware activity and the demand for blockchain settlement:

$$DepVar_t = \beta_1 Vuln_t + \beta_2 BTC_t + \varepsilon_t,$$

where $DepVar_t$ is one of the three variables: $nTransAll$ (Column 1) is the number of transactions involving all blockchain addresses; $nTransAdHoc$ (Column 2) is the number of transactions involving the addresses other than one hundred most popular network addresses; and $nTransRegular$ (Column 3) is the number of transactions involving one hundred most popular addresses. $Vuln$ is the number of newly discovered severe vulnerabilities, and BTC_t is the price of Bitcoin cryptocurrency. All regressions are estimated with standardized and demeaned variables. The standard errors (in parentheses) are adjusted for heteroskedasticity and autocorrelation. Asterisks ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels.

	(1)	(2)	(3)
$Vuln$	.117***	.124***	-.052**
	(.028)	(.028)	(.024)
BTC	.226***	.224***	.063***
	(.019)	(.020)	(.015)
R^2	.194	.193	.015

Table 4: Ransomware attacks and blockchain congestion

The table shows whether congestion costs are increasing with ransomware activity. I report coefficients of the following model:

$$DepVar_t = \beta_1 Vuln_t + \beta_2 BTC_t + \varepsilon_t,$$

where $DepVar_t$ is one of the three variables: $TransFeeBTC$ (Column 1) is the average daily fee attached to the blockchain transactions; $ConfTime$ (Column 2) is the average time it takes to process a fee-paying transaction; $nTransPerBlock$ (Column 3) is the average daily number of transactions per block. $Vuln$ is the number of newly discovered severe vulnerabilities and BTC_t is the price of Bitcoin cryptocurrency. All regressions are estimated with standardized and demeaned variables. The standard errors (in parentheses) are adjusted for heteroskedasticity and autocorrelation. Asterisks ***, **, and * denote statistical significance at the 1%, 5% and 10% levels.

	(1)	(2)	(3)
$Vuln$	.119***	.081***	.116***
	(.034)	(.030)	(.027)
BTC	.289***	.192***	.221***
	(.055)	(.033)	(.020)
R^2	.306	.135	.185

Table 5: The limit to congestion-induced reward

This table estimates the limit to the fees per block that congested users agree to pay. The limit is reached when further increase in ransomware activity no longer increases the mining reward per block coming from transaction fees. I use the following model to find whether such a limit exists:

$$CIR_t = \alpha + \beta_1 Vuln1_t + \beta_2 Vuln2_t + \beta_3 Vuln3_t + \beta_4 Vuln4_t + \beta_5 Vuln5_t + \beta_6 Vuln6_t + \beta_7 PriceBTC_t + \varepsilon_t$$

where CIR_t is the mean daily mining reward per block coming from transaction fees. $Vuln1_t - Vuln6_t$ are dummy variables equal to 1 if the number of severe vulnerabilities lies within a certain interval. Specifically, $Vuln1_t = 1$ if $0 < N \leq 3$, $Vuln2_t = 1$ if $3 < N \leq 10$, $Vuln3_t = 1$ if $10 < N \leq 20$, $Vuln4_t = 1$ if $20 < N \leq 30$, $Vuln5_t = 1$ if $30 < N \leq 40$, and $Vuln6_t = 1$ if $N > 40$. BTC_t is the price of Bitcoin cryptocurrency. The intercept corresponds to days when no severe vulnerabilities are released. All non-dummy variables are standardized. The standard errors (in parentheses) are adjusted for heteroskedasticity and autocorrelation. Asterisks ***, **, and * denote statistical significance at the 1%, 5% and 10% levels.

$Intercept$	-.140***
	(.047)
$Vuln1$	.032
	(.049)
$Vuln2$	.165**
	(.072)
$Vuln3$	.201**
	(.097)
$Vuln4$	.412**
	(.169)
$Vuln5$	.415**
	(.176)
$Vuln6$	.391***
	(.145)
BTC	.648***
	(.107)
$H: Vuln1 = Vuln2 = Vuln3 = Vuln4$	Rejected***
F-value	3.77
$H: Vuln4 = Vuln5 = Vuln6$	Not Rejected
F-value	.01
R^2	.466

Table 6: Blockchain congestion and cryptocurrency exchange liquidity

The table shows how blockchain congestion affects cryptocurrency exchange trading. Specifically, I estimate the following model:

$$DepVar_t = \beta_1 Vuln_t + \beta_2 BTC_t + \varepsilon_t,$$

where $DepVar_t$ is one of the four variables: $VolumeBTC$ (Column 1) and $nTrades$ (Column 2) is the daily volume and the number of trades on the Gemini exchange; $ESpread$ (Column 3) is the effective spread; $ImplShortfall$ (Column 4) is the implementation shortfall defined as twice the signed difference between the corresponding best quote when the market order arrived to the exchange and the realized execution price of the market order. $Vuln$ is the number of newly discovered severe vulnerabilities, and BTC_t is the price of Bitcoin cryptocurrency. All regressions are estimated with standardized and demeaned variables. The standard errors (in parentheses) are adjusted for heteroskedasticity and autocorrelation. Asterisks ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels.

	(1)	(2)	(3)	(4)
$Vuln$	.079**	.068**	.009	-.028
	(.036)	(.027)	(.038)	(.031)
BTC	.294***	.420***	-.048	-.029
	(.045)	(.066)	(.035)	(.044)
R^2	.329	.655	.020	.004

Table 7: First differences

The table addresses a concern about potential seasonality and persistence in the sample variables. The literature shows that regressions on cyclical and persistent data may produce spurious results. I address this issue by running the main regression in first differences. Specifically, I estimate the following model:

$$\Delta DepVar_t = \alpha + \beta_1 \Delta Vuln_t + \beta_2 \Delta BTC_t + \varepsilon_t,$$

where $\Delta DepVar_t$ is the daily log difference in one of the five variables: $\Delta nTransAll$ (Column 1) is the number of transactions involving all blockchain addresses; $\Delta nTransAdHoc$ (Column 2) is the number of transactions involving the addresses other than one hundred most popular network addresses; $\Delta nTransRegular$ (Column 3) is the number of transactions involving one hundred most popular addresses; $\Delta VolumeBTC$ (Column 4) and $\Delta nTrades$ (Column 5) is the volume and the number of trades on the Gemini exchange. $\Delta Vuln$ is the dummy variable equal to one if the daily difference in the number of newly discovered severe vulnerabilities exceeds five, and ΔBTC_t is the log return on Bitcoin cryptocurrency. The standard errors (in parentheses) are adjusted for heteroskedasticity and autocorrelation. Asterisks ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels.

	(1)	(2)	(3)	(4)	(5)
$\Delta Vuln$	.033***	.037***	-.065*	.231**	.184***
	(.010)	(.010)	(.038)	(.074)	(.058)
ΔBTC	.211*	.151	.102	.343	-1.284*
	(.099)	(.101)	(.434)	(.934)	(.691)
$Intercept$	.030***	.019***	.052	-.037*	.022
	(.003)	(.003)	(.017)	(.019)	(.018)
R^2	.018	.019	.002	.012	.014

Table 8: Robustness

Panel A examines whether the main results are sensitive to ransomware-unrelated cybersecurity risks. Coin theft risk may incentivize users to sell their coin holdings, thus increasing the number of blockchain transactions. To address this concern, I run the main analysis with the sample excluding three days around the discovery of Bitcoin-related vulnerabilities and major coin thefts. The specification is as follows:

$$DepVar_t = \beta_1 Vuln_t + \beta_2 PriceBTC_t + \varepsilon_t,$$

where $DepVar_t$ is one of the following variables: $nTransAll$ (Column 1); $nTransAdHoc$ (Column 2); $nTransRegular$ (Column 3); $VolumeBTC$ (Column 4); $nTrades$ (Column 5). $Vuln$ is the number of newly discovered severe vulnerabilities, and BTC_t is the price of Bitcoin cryptocurrency.

Panel B examines whether the main results are sensitive to the discovery of weak vulnerabilities. Weak vulnerabilities typically have low exploitability and should not cause significant shocks to ransomware activity. The NVD classifies vulnerabilities with the score below 4.0 as low-severity vulnerabilities. I use these weak vulnerabilities to examine whether my results are indeed driven by ransomware activity rather than by the commonalities between the patterns of blockchain congestion and the NVD vulnerability releases. The specification is as follows:

$$DepVar_t = \beta_1 Vuln_t + \beta_2 VulnWeak_t + \beta_3 PriceBTC_t + \varepsilon_t,$$

where $VulnWeak$ is the number of newly discovered vulnerabilities with the severity score below 4.0. All the remaining variables are as previously defined. All regressions are estimated with standardized and demeaned variables. The standard errors (in parentheses) are adjusted for heteroskedasticity and autocorrelation. Asterisks ***, **, and * denote statistical significance at the 1%, 5% and 10% levels.

	(1)	(2)	(3)	(4)	(5)
Panel A: Ransomware Unrelated Cybersecurity Risks					
$Vuln$	.116***	.122***	-.054*	.087**	.077***
	(.030)	(.030)	(.028)	(.038)	(.027)
BTC	.218***	.216***	.067***	.289***	.413***
	(.019)	(.019)	(.016)	(.046)	(.036)
R^2	.191	.190	.017	.332	.659
Panel B: Severe vs Weak Vulnerabilities					
$Vuln$	.103***	.105***	-.068**	.086*	.071**
	(.032)	(.032)	(.034)	(.049)	(.034)
$VulnWeak$	.026	.024	.028	.001	.010
	(.036)	(.037)	(.026)	(.045)	(.032)
BTC	.217***	.215***	.066***	.289***	.412***
	(.019)	(.019)	(.016)	(.046)	(.037)
R^2	.194	.190	.018	.332	.659

Table 9: Placebo test

The table estimates the likelihood of false results. I generate multiple placebo samples of the randomly drawn severe vulnerabilities and use these samples in the main regression setup. The specific procedure is as follows: first, the series of severe vulnerabilities are randomly permuted. Then, the main regression runs with the placebo sample of vulnerabilities as an independent variable. Next, the value 1 is assigned if the placebo coefficient sign is consistent with the sign of the coefficient in the original regression and 0 otherwise. Finally, the statistical significance of the placebo coefficient is captured. The above procedure runs 1,000 times. Panel A reports the percentage of instances when all placebo coefficients are consistent with the original ones. Panel B reports the percentage of instances when placebo coefficients are individually consistent with the original ones. The results in Panel B are split between the columns as follows: $nTransAll$ – Column 1; $nTransAdHoc$ – Column 2; $nTransRegular$ – Column 3; $TransFeeBTC$ – Column 4; $ConfTime$ – Column 5; $nTransPerBlock$ – Column 6; $VolumeBTC$ – Column 7; $ESpread$ – Column 8.

Panel A: Joint Probability of Consistent Significant Results, %								
10% *significance level*	.000							
5% *significance level*	.000							
1% *significance level*	.000							
Panel B: Marginal Probability of Consistent Significant Results, %								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
10% *significance level*	.059	.060	.096	.034	.035	.056	.021	.031
5% *significance level*	.031	.033	.052	.015	.015	.033	.005	.012
1% *significance level*	.006	.006	.017	.003	.000	.008	.001	.003

Figure 1: The WannaCry decryptor message

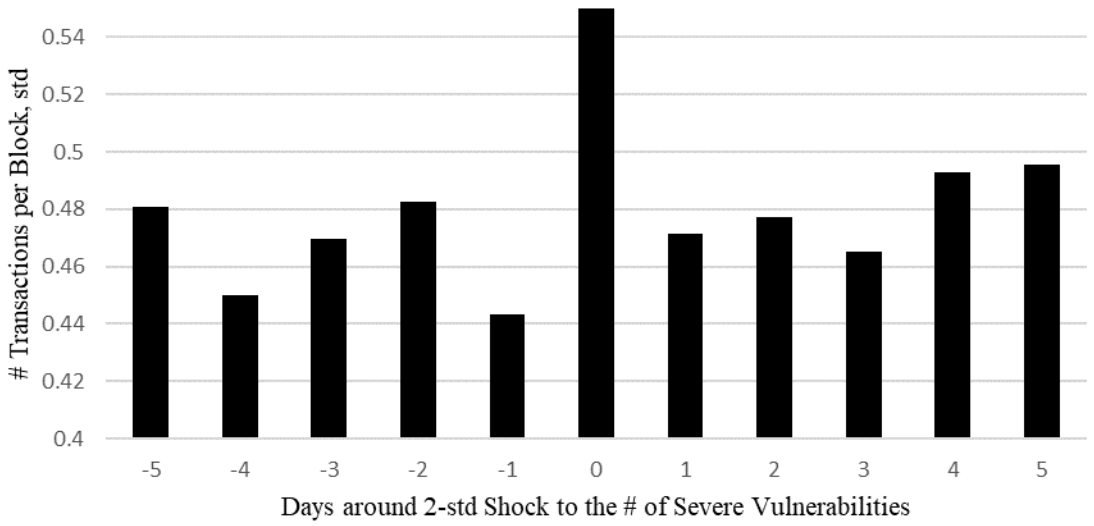

Figure 2: Severe vulnerabilities and Bitcoin blockchain transactions

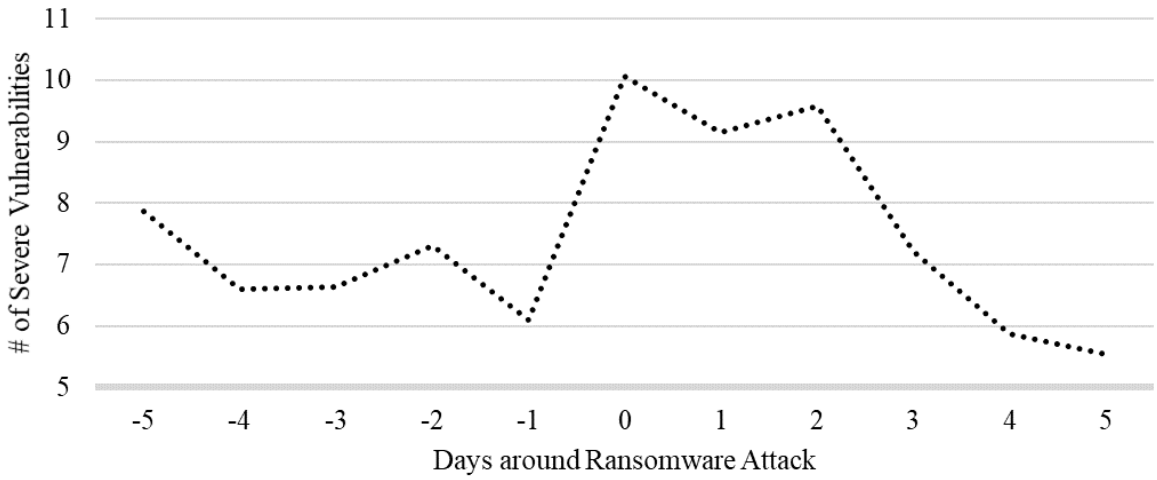

Figure 3: Severe vulnerabilities and ransomware releases

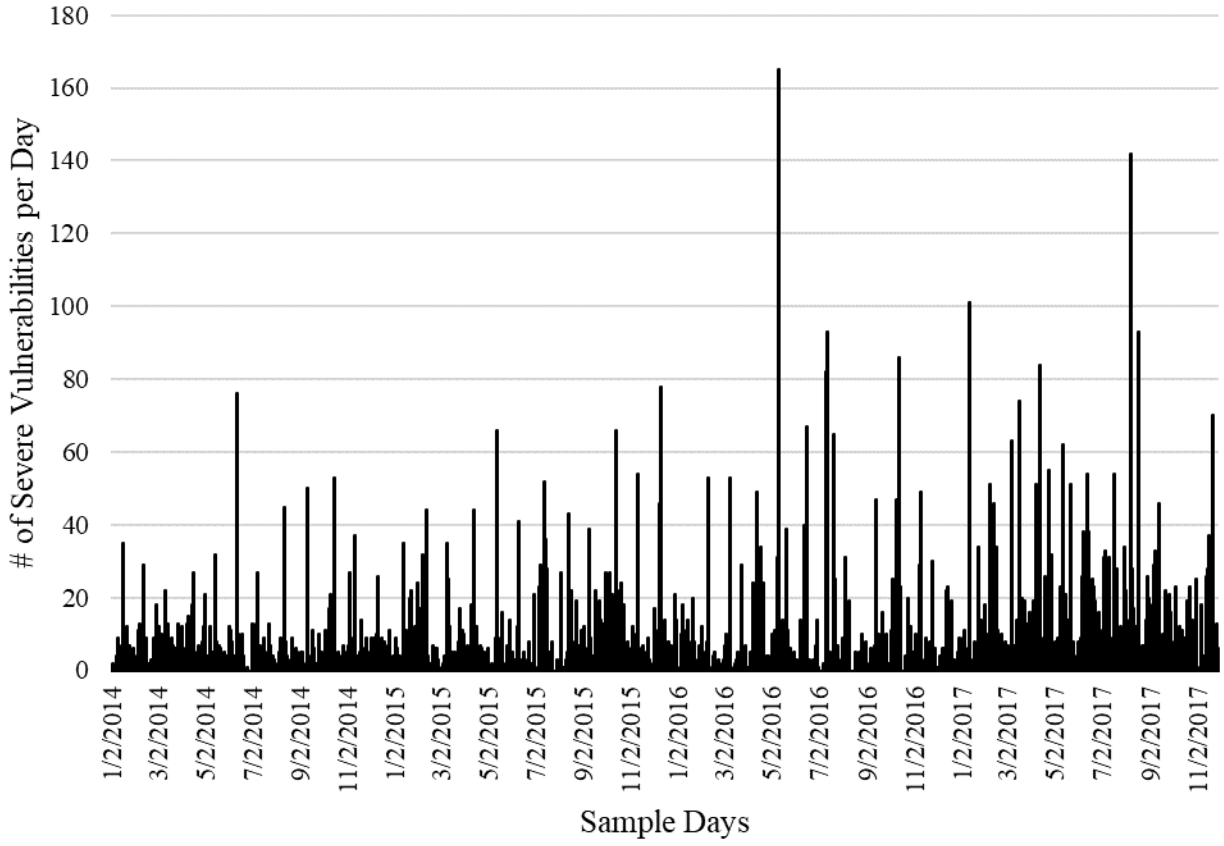

Figure 4: The daily number of severe vulnerabilities discovered during the sample period

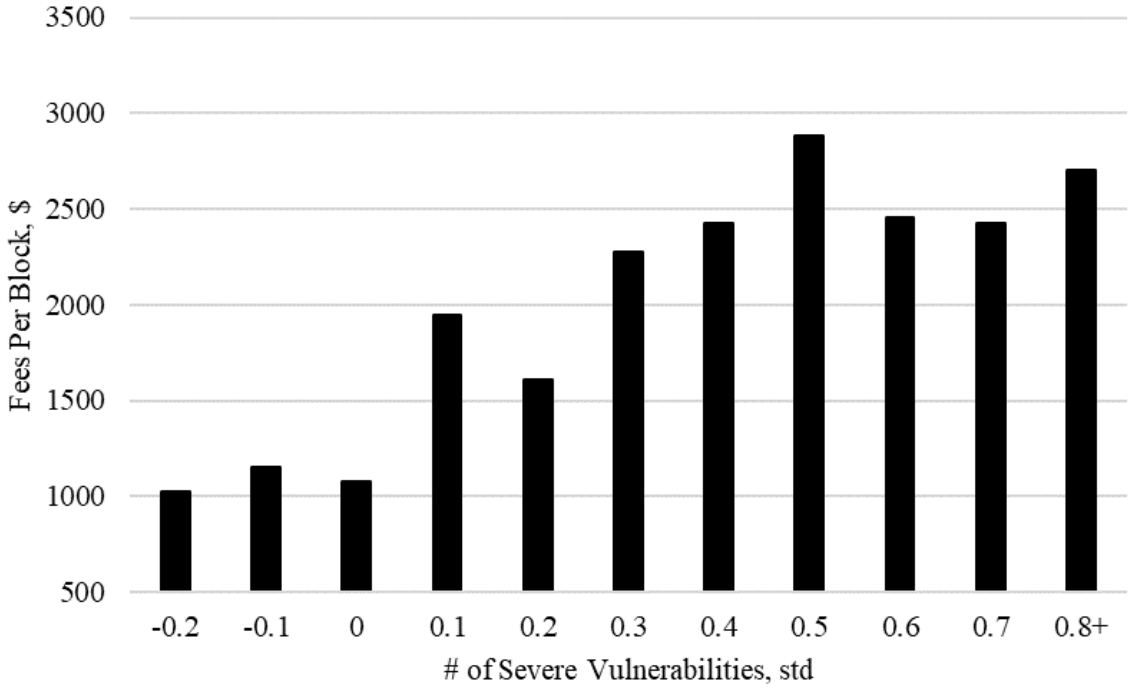

Figure 5: The distribution of the average value of fees attached per block

Appendix

NVD data feeds report a severity score of 1 to 10 for every vulnerability. This score is a function of the following exploitability factors:

$$Score = f(AccComp, Auth, AccVect, ConfImp, IntegImp, AvailImp),$$

where the inputs to the scoring function are:

- $AccComp$ is the vulnerability access complexity. High access complexity implies that a hacker needs to undertake many steps to exploit such vulnerability. Higher complexity lowers the overall score.
- $Auth$ is the number of times a hacker must pass authentication to a target to exploit a vulnerability. The higher this number, the lower the overall score.
- $AccVect$ is the means of access to the target. Some vulnerabilities require a hacker to have physical access to a target computer, while the others can be exploited remotely. Remotely exploitable vulnerabilities receive a higher score ceteris paribus.
- $ConfImp$ is the extent to which the confidentiality of a target system is compromised. An increase in this metric increases the overall score.
- $IntegImp$ is the extent to which the integrity of a system is compromised. This metric captures whether a hacker can modify information on a target computer by exploiting the vulnerability. An increase in this metric increases the overall score.
- $AvailImp$ is the impact the attack may have on the availability of the target system. Specifically, it reflects the potential of a hacker attack to disrupt the target system. A higher availability impact increases the overall score.

Table A.1 reports examples of vulnerabilities with different scores. Vulnerabilities with a high severity score typically allow a hacker to take control of the target system and install malicious programs such as ransomware. For instance, the vulnerability CVE-2017-8543 has a severity score of 10.0, and can be exploited to deploy a ransomware similar to WannaCry[10]. Vulnerabilities scored around 5.0 do not usually allow ransomware attacks. However, a chain of such vulnerabilities may occasionally compromise the system. For example, while CVE-2017-3764 reveals only usernames, another vulnerability may reveal passwords. That being said, the probability of this is rather low. Finally, vulnerabilities with a low severity score have little exploitability. For example, a hacker must rely on help from a network insider to exploit CVE-2017-3318.

[10] https://security.berkeley.edu/news/windows-search-remote-code-execution-vulnerability-cve-2017-8543

Table A.1: Typical vulnerabilities and severity scores

ID	Description	Score
CVE-2017-8543	Microsoft Windows XP SP3, Windows XP x64 XP2, Windows Server 2003 SP2, Windows Vista, Windows 7 SP1, Windows Server 2008 SP2 and R2 SP1, Windows 8, Windows 8.1 and Windows RT 8.1, Windows Server 2012 and R2, Windows 10 Gold, 1511, 1607 and 1703, and Windows Server 2016 allow an attacker to take control of the affected system when Windows Search fails to handle objects in memory, aka "Windows Search Remote Code Execution Vulnerability".	10.0
CVE-2017-3764	A vulnerability was identified in Lenovo XClarity Administrator (LXCA) before 1.4.0 where LXCA user account names may be exposed to unauthenticated users with access to the LXCA web user interface. No password information of the user accounts is exposed.	5.0
CVE-2017-3318	This difficult to exploit vulnerability allows a high privileged attacker with a logon to the infrastructure where MySQL Server executes to compromise MySQL Server. Successful attacks require human interaction from a person other than the attacker. Successful attacks of this vulnerability can result in unauthorized access to critical data or complete access to all MySQL Server accessible data.	1.0