Explaining Sustained Blockchain Decentralization with Quasi-Experiments: Resource Flexibility of Consensus Mechanisms

Harang Ju, Madhav Kumar, Ehsan Valavi, Sinan Aral

{harang, madhavk, valavi, sinan}@mit.edu

MIT Sloan School of Management

April 18, 2024

Preliminary. Please do not circulate.

Abstract

Decentralization is a fundamental design element of the Web3 economy. Blockchains and distributed consensus mechanisms are touted as fault-tolerant, attack-resistant, and collusion-proof because they are decentralized. Recent analyses, however, find some blockchains decentralized, others centralized, and trends towards both centralization and decentralization. Despite the importance and variability of decentralization across blockchains, we still know little about what enables or constrains blockchain decentralization. We hypothesize that the resource flexibility of consensus mechanisms is a key enabler of blockchain decentralization. We test this hypothesis using three quasi-experimental shocks—policy-related, infrastructure-related, and technical—to the sustained decentralization of six blockchains. We find strong suggestive evidence that the resource flexibility of consensus mechanisms enables sustained blockchain decentralization and discuss the implications for the design, regulation, and implementation of blockchains.

1 Introduction

> "The coming wave launches immense centralizing and decentralizing riptides *at the same time*... within the next decade, we must anticipate radical flux, new concentrations and dispersals of information, wealth and above all power... this crisis will take the form of a huge, existential bind, a set of brutal choices and trade-offs that represents the most important dilemma of the twenty-first century."
>
> – Mustafa Suleyman, *The Coming Wave*

In their book, *The Coming Wave*, Suleyman and Bhaskar (2023) describe how multiple technology trends will, in the coming decades, push societal institutions between dramatic centralization and unprecedented decentralization. Avoiding these extremes and their potentially dire consequences is, in Suleyman's estimation, humanity's greatest modern challenge. These institutional trends are catalyzed by new forms of organization enabled by technology, like those that depend on blockchains as their organizing foundation. Blockchain technology was conceived as a time-stamping mechanism to prevent back- or forward-dating of digital documents (Haber and Stornetta, 1991). In their modern incarnation, however, blockchain implementations were designed to remove centralized institutions from the process of tracking electronic cash (Nakamoto, 2008) and to execute generalized software programs without intermediaries (Buterin, 2014). Through a model of trust built on distributed consensus, blockchains serve as an alternative to traditional centralized systems that consolidate authority (Ammous, 2018).[1] For example, entrepreneur and investor Marc Andreessen, whose venture capital firm has invested over $7 billion in blockchain and Web3 startups (Dixon, 2022), has argued that "there is a fundamental technological breakthrough in an area of computer science called distributed consensus. It's the ability of people and software on the Internet to form trust relationships in an untrusted environment... To us it looks like... [a] breakthrough technological transformation happening in the tech industry and we take it very

[1] See congressional testimony by Peter Van Valkenburgh of Coin Center explaining Bitcoin's historical context.

seriously."[2] Since 2009, blockchains have witnessed rapid growth, surpassing a trillion-dollar market capitalization in 2021 despite occasional downtrends.[3]

Blockchain theorists claim decentralization as fundamental to the ability of blockchains to operate without trusted intermediaries while maintaining resistance to attacks, faults, and collusion (Buterin, 2017). While many tout the decentralized nature of blockchain technology, it is still unclear whether blockchains are actually becoming more centralized or decentralized over time and, perhaps more importantly: what drives blockchain centralization or decentralization? Indeed, decentralization faces powerful economic headwinds, including economies of scale and coordination costs (Hui and Tucker, 2023). To achieve consensus across distributed nodes of a network, many blockchains require the use of costly resources to verify transactions for proof-based consensus mechanisms, including hardware and electricity for Proof-of-Work (PoW) and native blockchain tokens for Proof-of-Stake (PoS).[4] Using technical virtualization theory (see Section 2), we hypothesized that the flexibility of resources required for consensus mechanisms may be important for sustained blockchain decentralization. We define resource flexibility as the ease with which participants can acquire and deploy the resources needed for consensus mechanisms, which includes not only the initial acquisition but also the ongoing operation and maintenance of such resources. This concept encompasses various factors, such as the cost, availability, and technical expertise required to participate in the consensus process.

In this paper, we investigated the effect of resource flexibility on blockchain decentralization in response to three independent policy-related, infrastructure-related, and technical shocks across six blockchains. To measure these effects, we employed event studies, difference-in-difference (DiD) estimation, and synthetic DiD (SDiD) methods. First, we defined the resource flexibility of distributed consensus mechanisms as the flexibility of resources—whether hard-

[2]See Bloomberg Wealth interview with David Rubenstein. Accessed October 20, 2023.
[3]CoinMarketCap Accessed August 11, 2023.
[4]The Proof-of-Authority consensus mechanism, which relies on a small number of validators selected by a central authority, is an obvious exception to the requirement for costly resources. However, most consensus mechanisms, even the Delegated Proof-of-Stake mechanism which also is relies on a small number of validators, require costly resources for consensus.

ware and electricity or blockchain tokens—used in consensus mechanisms within blockchain systems. Specifically, we examined (1) China's banning of crypto mining on May 15, 2021, (2) Hetzner's shutdown of Solana validators on November 2, 2022, and (3) the Ethereum Merge on September 15, 2022. We focused our analysis on the impact of these shocks on the consensus layers of six different blockchains: four permissionless blockchains (Bitcoin, Ethereum, Solana, and Gnosis) and two permissioned blockchains operated by companies (BNB and Ronin).

We utilized the quasi-experimental variation created by these shocks to explore how different levels of resource flexibility across blockchains, shocks, and time were associated with changes in decentralization. In the first natural experiment, we compared the effect of resource flexibility on decentralization in response to a policy shock across blockchains. In the second, we examined the effect of a technical shock on decentralization and compared the recovery of decentralization across the two shocks. In the third natural experiment, we observed one blockchain as it transitions, in time, from a lower to a higher level of resource flexibility in its consensus mechanism. To identify the impact of resource flexibility on the decentralization of consensus mechanisms, we used variation in the resource flexibility of Bitcoin and Ethereum mining during China's policy change, Hetzner's targeted shutdown of Solana nodes, and the Ethereum Merge as an event study. By comparing the effects of various shocks, we aim to distill fundamental principles of blockchain design that enable and sustain decentralization.

Our analysis provides strong evidence suggesting that the resource flexibility of distributed consensus mechanisms enables sustained blockchain decentralization. We found that China's ban reduced Bitcoin's decentralization in comparison to Ethereum's by around 0.21 bits of Shannon entropy from around 3.7 bits, that Hetzner's shutdown reduced Solana's decentralization by around 0.35 bits,[5] and that the dynamics of the recovery from the initial shocks

[5]To intuit this effect, consider a simple, best-case scenario in which Bitcoin nodes were perfectly equally distributed. Then, an effect of -0.209 bits on Bitcoin would indicate a loss of 1.780 (*i.e.*, $2^{3.722} - 2^{3.722-0.209}$) out of 13.19 (*i.e.*, $2^{3.722}$) nodes.

varied widely across the two shocks. Bitcoin's decentralization is still recovering from the Chinese ban, while Solana's decentralization fully recovered from the Hetzner shutdown in approximately one month. Our analysis also decomposes the entropy into the number of nodes in the network and the distribution of block production across those nodes. We discuss why the difference in recovery rates between Bitcoin and Ethereum in response to the China ban and Solana's rapid recovery from the Hetzner shutdown together support our hypothesis. We also found that the Ethereum Merge, which transitioned Ethereum from a Proof-of-Work (PoW) consensus mechanism (which is less flexible) to a Proof-of-Stake (PoS) mechanism (which is more flexible) increased Ethereum's decentralization by 0.52 bits. Collectively, our findings offer strong suggestive evidence that variation in the level of resource flexibility in consensus mechanisms—from ASICs to GPUs and from PoW to PoS—is associated with sustained levels of blockchain decentralization.

Our work contributes to several areas of research. First, our study is related to the literature on blockchain consensus mechanisms and their implications for firms (Bamakan et al., 2020; Halaburda et al., 2022). Consensus predicates many of the unique features of blockchains, including permissionlessness and trustlessness, which in turn support ownership, native payments, DeFi, and arbitrary smart contracts without central authorities (Catalini and Gans, 2016; Cong and He, 2019; Harvey et al., 2021; John et al., 2022). While previous studies illustrated the benefits of the technology, our paper is among the first to empirically demonstrate whether and how the consensus mechanism supports the promise of blockchains.

Our work also contributes to the literature on the decentralization of blockchains (Gencer et al., 2018; Hoffman et al., 2020; Wu et al., 2020; Cong et al., 2021; Sai et al., 2021; Jia et al., 2022). Many studies have pointed out the centralized nature of blockchain consensus mechanisms (Srinivasan and Lee, 2017), especially in permissionless blockchains (Bakos et al., 2021). Cong et al. (2022) recently causally identified off-chain computation (in Layer 2 "roll-ups") as a driver of decentralization in data providers for oracles, though not for the central mechanism of blockchain consensus. To the best of our knowledge, ours is the first

study to quantify the impact of consensus mechanism design on sustained decentralization.

Our results provide strong evidence in support of our hypothesis that the resource flexibility of consensus mechanisms enables sustained decentralization. Given that a *raison d'être* of blockchains is their independence from centralized intermediaries and their resilience to external shocks, it is essential to understand the conditions under which blockchains are or are not decentralized and shock resistant, especially as more and more economic activity occurs on blockchains, including trading in government bonds (Bank, 2021) and stock exchanges (Meichler, 2023).

2 Theory: Decentralization and Virtualization

The underlying integrity of blockchains depends, in theory and in practice, on their decentralization. The decentralization of blockchain consensus mechanisms theoretically ensures transaction validity without trusted intermediaries by preventing faults, attacks, and collusion in transaction validation (Buterin, 2017). Beyond ledger integrity, the decentralization of blockchains offers key sociotechnical primitives that are not available in traditional systems and form the basis of the so-called Web3 movement (Dixon, 2021).

First, decentralization reduces or entirely eliminates the role of trusted intermediaries and thus enables trustlessness as a foundational feature. Trustlessness not only allows for secure, direct peer-to-peer transactions but also alternative currencies (which can serve as a financial refuge from economic volatility), ownership records in digital ledgers outside of trusted institutions (from digital art to government bonds),[6] and smart contracts that can both codify and enforce arbitrary rules (Halaburda, 2016; Iansiti and Lakhani, 2017; Ammous, 2018; Cong and He, 2019; Halaburda et al., 2022). Second, decentralization enables participation and innovation without permission from gatekeepers, forming a permissionless ecosystem that fosters open access. Blockchain operators can participate in validation without permission (Cong et al., 2021), and any software developer can build applications on blockchains without permission (Halaburda, 2018). Third, when consensus is decentralized, blockchain data is

[6]See Bank (2021).

immutable and censorship-resistant (Nakamoto, 2008; Wahrstätter et al., 2023). Despite the foundational benefits of decentralization, empirical evidence shows significant variation in its application across different blockchains and over time (Gencer et al., 2018; Sai et al., 2021). Given the importance of decentralization to blockchain integrity and the broader potential of decentralized systems, the empirical variations underscore the need to examine the factors driving these variations.

At the theoretical limit, full decentralization is impossible in a permissionless system without an identity management system that gives one vote to each person (Kwon et al., 2019; Bakos et al., 2021). However, even in permissionless systems, Cong and He (2019) and Cong et al. (2022) showed that technological innovations that reduce frictions inherent in participation can increase competition in mining and decentralization in automated "oracles" that enable blockchains to seek and append information from off-chain sources. Among many innovations in computing networks more broadly, virtualization stands out as highly relevant and instrumental in reducing friction in deployment, disaster recovery, and operation on heterogeneous physical computing resources (Chowdhury and Boutaba, 2010). In this section, we illustrate with technical virtualization theory how resource flexibility could bolster blockchain decentralization, thus informing our hypothesis for our empirical study.

In the context of blockchain decentralization, we define resource flexibility as the capacity for blockchain resources—such as computational power, storage, or blockchain tokens—to be deployed adaptively in response to changes in operational conditions, thereby reducing the cost of entry and avoiding operational disruption. As virtualization theory illustrates, computing networks, which include blockchains, can achieve more resource flexibility by leveraging technical virtualization. Resource flexibility may explain sustained blockchain decentralization by allowing them to be easily deployed and adaptively managed to shocks to the nodes of the blockchain network. This concept underpins our hypothesis that increased resource flexibility, facilitated by technical virtualization, can contribute to the robust decentralization of blockchain systems, thereby addressing some of the economic challenges

such as economies of scale and coordination costs.

In the realm of technical virtualization, Ameen and Hamo (2013) survey the role of virtualization technologies in improving resource utilization and providing a unified operating platform based on the aggregation of heterogeneous resources. Such technologies transform traditional information technology (IT) infrastructure into a flexible resource pool, thereby enhancing IT agility and scalability (Shirinbab, 2019).[7] Server virtualization also improves resource availability (van Cleeff et al., 2009), and these improvements are particularly crucial for systems like the blockchain, where efficiency, flexibility, and availability are key to promoting and sustaining decentralization. While it has been shown that server virtualization has certain negative impacts on network stability and confidentiality (van Cleeff et al., 2009; Wang and Ng, 2010), these effects are largely mitigated on blockchains, which operate with secure cryptography at orders of magnitude less throughput and higher latency than traditional cloud services and are designed to be resilient to network instability. Technical virtualization also enables rapid deployment and flexibility in contexts beyond cloud computing. In manufacturing, for example, the virtualization and decentralization of supply chains transform manufacturing firms into integrated networks that allow them to flexibly alter and rapidly innovate on products (Brettel et al., 2014).

To explicitly understand the potential role of virtualization in blockchain decentralization, we examine two key contributions of virtualization to decentralization: 1) offering resource flexibility for rapid deployment and disaster recovery, and 2) democratizing participation by reducing barriers to entry. First, virtualization technologies are instrumental in enabling flexible and rapid deployment, as well as re-deployment, of blockchain nodes. This is crucial for disaster recovery and operational resilience, allowing blockchains to adapt and recover swiftly from disruptions. The ability to rapidly deploy nodes offers a time-efficient alternative to hardware-based solutions, which typically require more resources and longer setup times.

[7]Not only does virtualization affect IT flexibility for firms, it also influences how consumers use technologies. For example, cloud document editing services, such as Google Docs or Overleaf, allow users to access and collaborate on any document on any computer that has an Internet browser without needing to download applications or set up and maintain LaTeX environments.

This flexibility not only facilitates greater participation in the network but also enhances its decentralization. Second, virtualization plays a significant role in lowering barriers to entry for participants. By enabling the use of a wide range of computing resources and reducing the overall cost of participation, virtualization democratizes access to the blockchain network. This is especially important in the context of permissionless blockchains, where the objective is to prevent the centralization of control in the face of economies of scale associated with computing infrastructure. These contributions collectively seem to mitigate the risks associated with centralization and may support a more robust, decentralized blockchain ecosystem.

In conclusion, the theory of technical virtualization presents a critical perspective on how resource flexibility may act as a pivotal force in encouraging blockchain decentralization. This perspective highlights the significant role of resource flexibility provided by virtualization technologies, which are essential for the rapid deployment and recovery of blockchain nodes. These capabilities are not only pivotal for enhancing operational resilience but also crucial for fostering wider participation by lowering entry barriers, thus promoting a more decentralized network structure. By affording an array of computing resources and minimizing participation costs, virtualization democratizes access to blockchain networks. This is particularly critical in permissionless blockchains aimed at circumventing the concentration of control due to economies of scale in computing infrastructure. Hence, our investigation delves into how resource flexibility impacts blockchain decentralization across various dimensions, including differences within and between consensus mechanisms, and through significant upgrades such as Ethereum's transition to Proof-of-Stake. Further, virtualization theory can potentially guide future blockchain innovations. For instance, advancements allowing stateless nodes and those reducing data storage demands, as seen in Ethereum's recent upgrades, are expected to lower validator operational costs, thus improving the flexibility of Ethereum validators.[8] In light of these developments, our empirical analysis seeks to examine whether resource

[8]See the Ethereum roadmap at ethereum.org and the Cancun-Deneb upgrade on github.com. Accessed March 26, 2024.

flexibility of consensus mechanisms—the cornerstone of blockchain technology—can explain sustained decentralization in blockchains.

3 Methodology

3.1 Experimental Setting

We studied three events that serve as policy-related, infrastructure-related, and technical shocks to the decentralization of blockchain consensus respectively. Table 1 outlines the three events, the types of shocks, and the resources used in the consensus mechanisms in the events.

Event	Type of Shock	Resource
China bans crypto mining	Policy-related	Hardware (ASIC, GPU)
Hetzner shuts down Solana nodes	Infrastructure-related	Tokens (SOL)
The Ethereum Merge	Technical	Hardware (GPU) to Tokens (ETH)

Table 1: Experimental setting

China's ban on crypto mining. First, we studied the policy shock of China's ban on crypto mining on May 15, 2021, analyzed in Section 4.1 (Shen and Siu, 2021). This official Chinese policy established a blanket ban on all mining activities, including Bitcoin and Ethereum, both of which used PoW validation at the time. Both blockchains were relatively mature when the ban was enacted, with Bitcoin online since January 2009, with a market cap of over $809 billion, and Ethereum online since July 2015, with a market cap of $349 billion.[9]

While the ban applied to both Bitcoin and Ethereum, the key difference between the chains was the hardware used to achieve consensus, which was directly related to their resource flexibility and, in turn, affected their recovery from the shock. We used this quasi-experimental variation in the shock to mining for Bitcoin and Ethereum to identify the impact of virtualization in response to a large policy shock on blockchain decentralization. The technical specification underlying the variation was that Bitcoin used (and still uses) the

[9]See coingecko.com Accessed August 21, 2023.

SHA-256 hash function (Nakamoto, 2008) and Ethereum used Ethash.[10] Bitcoin's SHA-256 function was primarily mined using application-specific integrated circuits (ASICs), while Ethereum's Ethash was designed to be ASIC-resistant and thus was mined primarily using graphical processing units (GPUs).[10] Because ASICs are, by definition, specific to mining, it is much more difficult to replace them during such a shock. On the other hand, GPUs are more general-purpose and can be virtualized and readily accessed on many cloud computing services. Given this variation in the inherent resource flexibility of the technologies supporting the consensus mechanisms underlying Bitcoin and Ethereum, the China ban provides an excellent natural experiment with which to test the impacts of virtualization on blockchain decentralization. For this shock, we analyze data from July 19, 2020, to March 11, 2022. Note that the Bitcoin halving occurred on May 11, 2020, which does not confound our analysis.

While Bitcoin and Ethereum differed in the types of applications used on them, we focused here on the consensus mechanisms used by the blockchains. Bitcoin was primarily used as a store of value and a medium of exchange, and Ethereum had those and additional use cases, such as decentralized finance and fungible and non-fungible tokens. However, the application layer is largely agnostic to the consensus layer; *i.e.*, the PoW consensus mechanisms, which were used at the time by both blockchains, simply manage the creation of blocks (Nakamoto, 2008; Buterin, 2014). The PoW consensus mechanism secures the network by requiring computational work to add new blocks, indirectly influencing economic incentives such as mining rewards and transaction fees that maintain blockchain functionality and integrity. Importantly, this mechanism is agnostic to the type of data stored in the blocks—whether they are transactional records, smart contracts, or token information—and focuses solely on ensuring data integrity and network security. Thus, the economic incentive structure is equivalent for both Bitcoin and Ethereum apart from the different hash functions and hence the differences in required mining hardware.

Given the focus of our analysis on the consensus layer, we tested the robustness of our

[10]See a detailed explanation of Ethash on ethereum.org. Accessed October 13, 2023.

results with key metrics of the consensus layer in Section A of the Appendix. First, the metrics that capture the computational work to add new blocks include the number of transactions, the amount of data stored, and the estimated hashrates. Next, metrics that capture the indirect economic incentives of mining include the token prices of BTC and ETH, transaction fees, and miner rewards. We found that our results were robust to these potential confounders.

Hetzner's shutdown of Solana nodes. Second, we studied Hetzner's shutdown of Solana nodes on November 2, 2022, as described in Section 4.2 (Avan-Nomayo, 2022). Hetzner is a cloud service provider based in Gunzenhausen, Germany. Operators of PoS validators often run their validators on cloud services because of their ease of deployment, high uptime, and low maintenance. Hetzner, a German cloud service provider, had expressly prohibited any cryptocurrency-related activities on its platform as of August 23, 2022.[11] Despite this general ban, the specific targeting of Solana nodes for shutdown, while not applying the same measure to other blockchains like Ethereum. On the day of the shutdown, the delinquency rate among Solana validators spiked to 19.8%, indicating that validators representing nearly a fifth of actively staked SOL could not validate.[12] This event created a unique quasi-experimental scenario that allowed us to explore the effects of significant infrastructure disruptions on the decentralization of blockchain networks. Our study leverages synthetic difference-in-difference and event study methods to aid relevant comparisons and support identification. For this shock, we analyzed data from September 13, 2022, to December 22, 2022, 50 days before and after the shock on November 2, 2022.

This selective enforcement by an infrastructure provider represents an exogenous shock to the Solana network, offering a valuable perspective on infrastructure dependency and the risks posed by centralized cloud services to blockchain decentralization. However, the reasons behind Hetzner's selective enforcement against Solana nodes, as opposed to nodes from

[11] Hetzner's official statement: "Using our products for any application related to mining, even remotely related, is not permitted. This includes Ethereum. It includes proof of stake and proof of work and related applications. It includes trading." See their official statement and X post. Accessed April 14, 2024.
[12] RockawayX Solana Dashboard. Accessed August 22, 2023.

other blockchains, remain unexplained by the company. Hetzner's public communications state a broad prohibition against the use of their services for any crypto-related activities, encompassing both PoW and PoS mechanisms, as well as indirectly related activities such as cryptocurrency trading.[11] The absence of similar actions against other chains during the same period suggests a potentially unique issue with Solana's relatively high resource usage that triggered enforcement. For example, Solana nodes require 256GB of RAM, while Ethereum nodes require just 8 GB of RAM.[13] If indeed Solana was targeted due to its excessive resource usage, this would support the role of resource efficiency of consensus mechanisms in enabling sustained blockchain decentralization.

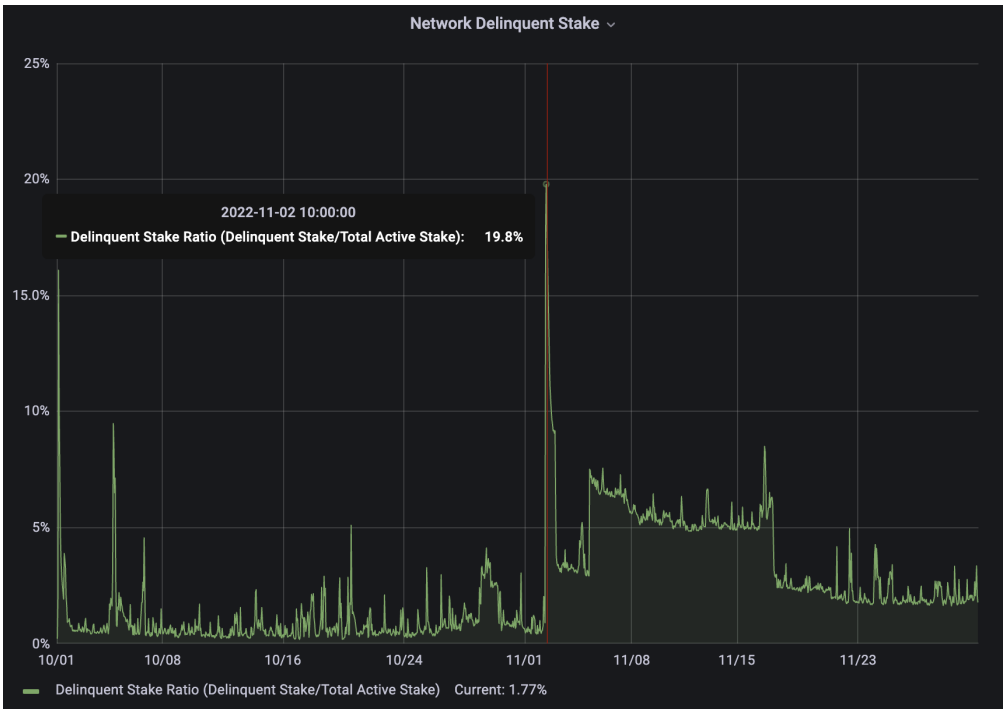

Figure 1: Delinquent stake ratio for Solana during the Hetzner shutdown.

The Ethereum Merge. Third, we studied the Ethereum Merge on September 15, 2022, as described in Section 4.3 (The Ethereum Foundation, 2023). The Paris Upgrade, as it is also known, merged Ethereum's original blockchain with the new PoS consensus mechanism (Beiko, 2022). At that point, each new block was no longer mined by PoW miners but was

[13]See hardware requirements for Solana and Ethereum. Accessed April 14, 2024.

from then on validated by PoS validators. While in PoW, miners used servers with graphical processing units (GPUs) to mine new blocks through energy-intensive computations, anyone after the PoS upgrade could start validating on any server using 32 ETH. This upgrade is perhaps the biggest technical shock to Ethereum's consensus mechanism to date. Since we could not assume or validate parallel trends with any other blockchain during this period, we analyzed this shock as an event study to identify the effect of a major software upgrade from PoW to PoS on blockchain decentralization. For this shock, we analyze data from January 1, 2022, to September 15, 2022.

3.2 Data

We used node-level data from six different blockchains to study the drivers of blockchain decentralization. A blockchain is a linked list of blocks, each containing a list of transactions. A node is a server that operates as a validator in PoS consensus or a miner in PoW consensus. In general, around every N seconds, depending on the blockchain specifications, a node produces a block that is appended to the blockchain.

The data we used consists of the daily number of blocks produced by each node. Here, we focused on six blockchains—Bitcoin, Ethereum, Solana, BNB, Ronin, and Gnosis—that facilitate causal identification. Bitcoin, Ethereum, Solana, and Gnosis are all *permissionless* blockchains in the sense that anyone can operate a node and join the network to produce blocks on the blockchains. However, the three blockchains vary in their designs, which allowed us to tease apart potential drivers of decentralization. Bitcoin is still PoW and uses specialized hardware (i.e., application-specific integrated circuits *ASICs*) and energy-intensive computations to achieve consensus. Ethereum relied on PoW until September 15, 2022, when it was upgraded to PoS, which uses staked native tokens (*i.e.*, ETH) to achieve consensus. Solana also uses PoS along with Proof-of-History (PoH), which scales transaction throughput while maintaining transaction sequences.

In contrast to the other chains, Ronin and BNB are *permissioned* blockchains that have

both centrally designated and elected validators.[14] Importantly, permissioned blockchains can be affected by policy-related or infrastructure-related shocks. Hence, we used the permissioned PoS blockchains as control units for Solana in the analysis of Hetzner's shutdown (an infrastructure shock) and not for the technical shock of the Ethereum Merge, where permissioned blockchains may be less relevant.

We obtained Bitcoin and Ethereum hashrates from `blockchain.com` and `etherscan.io`, respectively. The hashrate is a measure of the collective computing power in a Proof-of-Work blockchain network; specifically, it is the total number of hash functions that all nodes in a network compute per second. While it cannot be directly measured, it is estimated from the number of blocks currently being mined and the current block difficulty. The computational power of an individual node can be affected by gradual hardware degradation, hardware upgrades, changes in efficiency due to electricity or cooling, or, most importantly for our context, changes in the amount of resources used by the node, which can be hash power or blockchain tokens.[15]

We aggregated daily counts of miners per day for Bitcoin, Ethereum, BNB, Gnosis, and Solana on Dune Analytics (dune.com). We obtained the same data for Ronin from Moku (moku.gg). For Bitcoin, we divided the attribution of the block proportionally among the recipients of the block reward to account for mining pools that distribute rewards in the coinbase transaction; for all others, we attributed each block to a single node.[16] The first block of Bitcoin was recorded on January 3, 2009, of Ethereum on July 30, 2015, of Solana on March 16, 2020, of BNB on August 29, 2020, of Gnosis on October 8, 2018, and of Ronin on January

[14] Ronin used Proof of Authority (PoA), which has a limited number of validators chosen by a central entity (often a company), but transitioned to Delegated Proof of Stake (DPoS), which has a limited number of elected and centrally designated validators, in April 2023.

[15] Most crypto miners are profit-maximizing and thus operate at maximum efficiency given the high costs of ASICs and GPUs, even in geographic regions with cheap electricity. See compassmining.io.

[16] The coinbase transaction is the first transaction in a block on a blockchain network such as Bitcoin, not to be confused with the eponymous cryptocurrency exchange Coinbase. This transaction is unique because it creates new coins and awards them to the miner who successfully added the block to the blockchain. It also typically includes the distribution of transaction fees collected from other transactions in the block, which are likewise awarded to the miner. In the context of mining pools, the coinbase transaction distributes these rewards proportionally among all pool participants, based on their contributed computational power.

25, 2021. Bitcoin produces 1 block approximately every 10 minutes, Ethereum around every 13 seconds, Solana every 800ms, BNB every 5 seconds, Gnosis every 3 seconds, and Ronin every 3 seconds.[17] We analyzed data collected from the start of each blockchain until July 31, 2023. The data include 2,359,386 blocks and 84,017 nodes for Bitcoin, 17,930,827 blocks and 17,680 nodes for Ethereum, 14,074,4978 blocks and 3,363 nodes for Solana, 31,372,303 blocks and 60 nodes for BNB, 30,315,947 blocks and 1168 nodes for Gnosis, and 22,973,856 blocks and 39 nodes for Ronin.

3.3 Variable Construction

Decentralization Variables: We used multiple measures of decentralization for the robustness and interpretability of our estimated effects: Shannon entropy, the number of nodes, the Gini coefficient, the Nakamoto coefficient, and the Herfindahl–Hirschman index (HHI).

We define the measures as follows. The number of nodes, which we refer to as Nodes_{it}, is simply the number of nodes N_{it} of blockchain i that produced at least one block on day t.

Shannon entropy, which we refer to as Entropy_{it}, is a measure of the unpredictability or randomness in the distribution of blocks among nodes. Formally, it is defined as

$$\text{Entropy}_{it} = -\sum_{x \in X_{it}} p(x) \log p(x),$$

where X_{it} is the set of the number of blocks produced by each node in blockchain i on day t (Lin et al., 2021; Jia et al., 2022). It has units of bits (as in *gigabit*) and is bounded by $[0, \log_2(N)]$. A higher entropy means more decentralization. Importantly, Shannon entropy captures and can be decomposed into the number of nodes *and* the distribution of block production over those nodes, measured by the Gini coefficient, the Nakamoto coefficient, and HHI. Thus, it is a holistic measure of decentralization. A convenient intuition for entropy is that any system with an entropy of H (*e.g.*, an entropy of 8) is equivalent in measure to a

[17] See specifications for each blockchain at bitinfocharts.com, etherscan.io, docs.solana.com, and ronin-chain.com, respectively. Accessed September 5, 2023.

system with $\log_2(H)$ equally contributing entities (*e.g.*, 3 nodes with the same number of blocks validated in a day).

The Gini coefficient Gini_{it} is a measure of inequality within a distribution and quantifies the disparity in block production across nodes in blockchain i on day t. Formally, it is defined as

$$\text{Gini}_{it} = \frac{\sum_{k=1}^{N_{it}} \sum_{l=1}^{N_{it}} |x_k - x_l|}{2N_{it}^2 \overline{x}_{it}},$$

where x_k and x_l are values from X_{it} and $\overline{x}_{it}$ is the average of all x_k in X_{it}. It is unitless and is bounded by $[0, 1]$. A higher Gini coefficient means less decentralization.

The Nakamoto coefficient Nakamoto_{it} measures the minimum number of nodes required to collectively control over 51% of the block production on a given day. Formally, it is defined as

$$\text{Nakamoto}_{it} = \min\{n \in [1, ..., N_{it}] : \sum_{k=1}^{n} p_k \geq 0.51\},$$

where p_k is the fraction of blocks produced by the k^{th} node on day t (Srinivasan and Lee, 2017). The Nakamoto coefficient is a key measure for assessing the resilience of the blockchain against potential collusion or majority attacks. It has units of nodes and is bounded by $[1, N]$. A higher Nakamoto coefficient means more decentralization.

Finally, HHI_{it} indicates the concentration of block production among nodes in blockchain i on day t. Formally, it is defined as

$$\text{HHI}_{it} = \sum_{k \in K_{it}} p_k^2.$$

HHI_{it} is calculated as the sum of the squares of individual market shares of all participants. It is unitless and is bounded by $[1/N, 1]$. A higher HHI means less decentralization.[18]

[18]Note: We do not take into account potential Sybil attacks in our analysis, as no scalable anti-Sybil systems exist for blockchains. See Vitalik Buterin's blogs on hard problems in cryptocurrency and proof of personhood. Hence, our measures of decentralization are technically the best-case approximations of decentralization. However, substantial Sybil attacks are highly unlikely, especially for the large blockchains that we study here since blockchains, by design, require resources to participate in consensus. For Proof-of-Stake blockchains, as an example, Sybil costs are scaled by the market capitalization of the tokens, which

Treated Blockchain Variable: The treated blockchain variable Chain_i is a binary indicator variable for blockchain i. Chain_i equals one if the observation is from the blockchain that receives the treatment or zero if the observation is from the compared blockchain that does not. For example, in our analysis of China's mining ban, $\text{Chain}_{\text{Bitcoin}} = 1$, and $\text{Chain}_{\text{Ethereum}} = 0$. For simplicity, the tables display $\text{Chain}_{\text{Bitcoin}}$ as Bitcoin for China's mining ban and $\text{Chain}_{\text{Solana}}$ as Solana for Hetzner's shutdown.

Post-Treatment Variable: The post-treatment variable Post_t is a binary indicator variable for date t. It is zero if the observation is before the treatment date and one if the observation is on or after the treatment date.

Treatment Variable: The variable Treatment_{it} is the combination of both the post-shock and blockchain variables. It is formulated as the cross-term ($\text{Post}_t \times \text{Chain}_i$) and is shown as Treatment_{it} for convenience and clarity. This variable is used to estimate the average treatment effects on the treated (ATT) δ. For example, in our analysis of China's mining ban, the coefficient δ of the treatment variable represents the difference in Bitcoin's and Ethereum's entropy after the ban.

Exposure Variable: China's mining ban disproportionately impacted Bitcoin's network compared to Ethereum's, as evidenced by the drop in hashrates starting at the time of the ban (Figure 2). To control for the difference in exposure to the shock, we included the variable Exposure_{it} as a covariate in our analysis of China's mining ban. It is formulated as the cross-term ($\text{Post}_t \times \text{Exposure}_i$), where Exposure_i is the fractional decrease in the hashrate of blockchain i from its peak hashrate at around the time of the ban to the minimum hashrate in the months following the ban.[19]

Non-random exposure to exogenous shocks can bias estimates, and methods have been devised to avoid such omitted variable bias (Borusyak and Hull, 2020). These methods involve estimating the expected exposure by simulating the process by which a unit receives the shock

are in the trillions to tens of trillions of dollars.

[19] Because the exposure lasts for months after the peak, we also perform multi-period difference-in-difference estimation in Section B of the Appendix (Callaway and Sant'Anna, 2021).

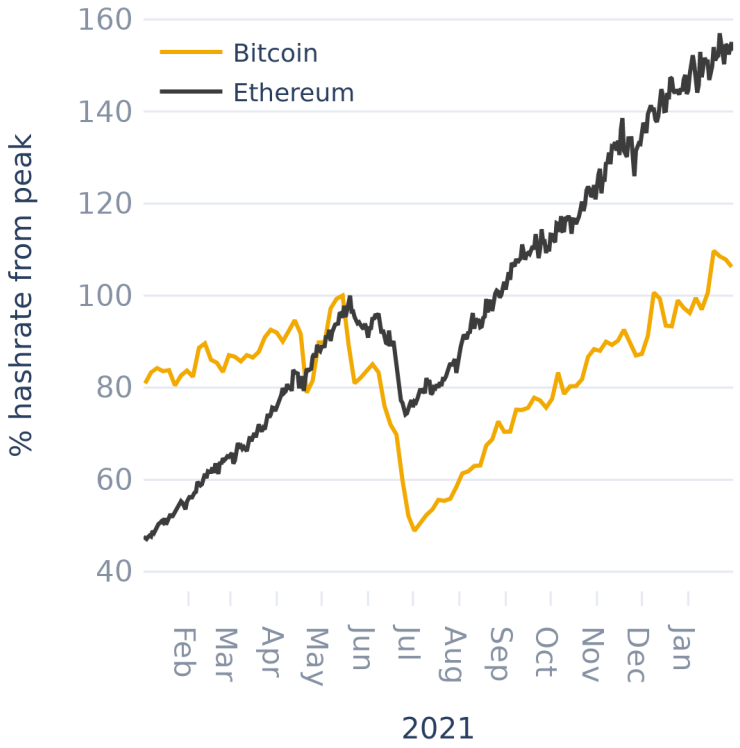

Figure 2: Percentage hashrates for Bitcoin and Ethereum from the local peak around China's ban on crypto mining.

and constructing a recentered instrumental variable based on that estimate. Fortunately, we can directly measure the exposure for both Bitcoin and Ethereum by estimating the drawdown of the total daily hashrate from the peak around May 15$^{\text{th}}$ (Figure 2).[20] We found that $\text{Exposure}_{\text{Bitcoin}} = 0.511$ and $\text{Exposure}_{\text{Ethereum}} = 0.258$, and we used these estimates as control and treatment exposure covariates. We used the date of the peak hashrate for Bitcoin on May 15, 2021, as the date for the shock. As a robustness check, we validated our exposure measures using country-level geographic monthly hashrates for Bitcoin, serving as a proxy for both chains, in Section C of the Appendix.

Time Variable: The time variable Day_t is a running variable used in our event study. It is the number of days after the treatment date and is set to zero at the treatment date.

[20]This direct measurement is a clear feature of our identification strategy as simulated exposure methods are considerably noisier than directly measured exposure proxies.

3.4 Model Specifications

To identify causal effects in our analysis of the three shocks, we employed three empirical methods: difference-in-difference estimation (DiD), synthetic difference-in-difference (SDiD), and an event study.

Difference-in-Difference (DiD) Estimation We used DiD for the policy shock of China's mining ban and the infrastructure shock of Hetzner's shutdown of Solana nodes. The DiD model was specified as follows:

$$\mathrm{Y}_{it} = \delta \mathrm{Treatment}_{it} + \beta_1 \mathrm{Chain}_i + \beta_2 \mathrm{Post}_t + \epsilon_{it}, \tag{1}$$

where Y_{it} is a measure of decentralization (*i.e.*, $\mathrm{Entropy}_{it}$, Nodes_{it}, Gini_{it}, $\mathrm{Nakamoto}_{it}$, or HHI_{it}) of blockchain i on day t and ϵ_{it} is the error term. Including the exposure variable as a covariate yields:

$$\mathrm{Y}_{it} = \delta \mathrm{Treatment}_{it} + \beta_1 \mathrm{Chain}_i + \beta_2 \mathrm{Post}_t + \beta_3 \mathrm{Exposure}_{it} + \epsilon_{it}. \tag{2}$$

For China's ban, the variable Chain_i is one for Bitcoin and zero for Ethereum. The variable $\mathrm{Treatment}_{it}$ is the cross-term (Bitcoin $\times$ Post_t) and indicates one for Bitcoin after the ban or zero for Bitcoin before the ban and for Ethereum before and after the ban. For Hetzner's shutdown of Solana nodes, the variable Chain_i is one for Solana or zero for the synthetic control, which is explained below. The variable $\mathrm{Treatment}_{it}$ indicates one for Solana after the shutdown or zero for Solana before the shutdown and for the synthetic control before and Post the ban, or simply (Solana $\times$ Post_t). The coefficient δ of the treatment variable $\mathrm{Treatment}_{it}$ represents the difference in the treated chain and control chain following the shock.[21]

In addition to estimating the treatment effects before and after the shocks, we aimed to

[21] For China's ban on mining, we perform multi-period difference-in-difference analysis in Section B of the Appendix to account for the months-long exposure to the shock (Callaway and Sant'Anna, 2021).

estimate the dynamics of decentralization after the shocks by identifying the leading and lagging treatment effects. In our context, this analysis reveals if the blockchain recovered from the shock and, if so, how quickly. Thus, we modified Equation 1 by summing over the time-shifted treatment effects as specified by

$$\mathrm{Y}_{it} = \sum_{\lambda=-T}^{T} \delta_{\lambda} \mathrm{Treatment}_{it\lambda} + \beta_1 \mathrm{Chain}_i + \beta_2 \mathrm{Post}_t + \epsilon_{it}, \tag{3}$$

where λ is the lead and lag effects for $\lambda < 0$ and $\lambda > 0$, respectively and T is the maximum lead and lag.

Synthetic Difference-in-Difference (SDiD) Estimation For the infrastructure shock of Hetzner's shutdown of Solana nodes, Solana exhibited no comparable control due to its unique entropy measures, which were significantly higher than those of other blockchains. Thus, we employed a synthetic difference-in-difference (SDiD) estimation approach using a synthetic control (Arkhangelsky et al., 2021). This control is constructed as a weighted sum of the entropy values from four other Proof-of-Stake blockchains: Ethereum, Gnosis, BNB, and Ronin. The purpose of the synthetic control is to create an optimally estimated counterfactual to Solana's decentralization scenario in the absence of Hetzner's shutdown.

The SDiD model is specified as follows:

$$\mathrm{Y}_{it} = \tau_{it} \mathrm{Treatment}_{it} + \beta_1 \mathrm{Chain}_i + \epsilon_{it}, \tag{4}$$

where Y_{it} represents the decentralization of blockchain i at time t (see Section 3.3), τ_{it} is the treatment effect of Hetzner's shutdown on the unit i at time t, $\mathrm{Treatment}_{it}$ is a binary indicator which equals 1 for Solana post-shutdown and 0 otherwise, Chain_i is the blockchain fixed effects, and ϵ_{it} is the error term. To estimate the average treatment effect τ where the treatment was applied, we calculate τ as the average of τ_{it} over the observations with $\mathrm{Treatment}_{it} = 1$. All treated units begin treatment simultaneously. We computed the

synthetic DiD estimators using the R package `synthdid` (Arkhangelsky et al., 2021).

Event Study We used an event study to estimate the effect of the software shock of the Ethereum Merge and to test the robustness of the difference-in-difference estimation for Hetzner's shutdown. For the Ethereum Merge, we could not find parallel trends between the decentralization of Ethereum and other blockchains around the event (as this was an Ethereum-only shock). The Event Study model was specified as follows:

$$\text{Entropy}_t = \delta \text{Post}_t + \beta_1 \text{Day}_t + \beta_2 (\text{Post}_t \times \text{Day}_t) + \epsilon_t, \tag{5}$$

where the superscripts, subscripts, and variables are the same as in Equation 1 and explained in Section 3.3. We do not include the subscript i in Equation 5 because it is estimated for only a single blockchain at a time.

We used an *event study* and not *regression discontinuity in time* to implement this causal estimation strategy because the effect was not localized around a specific time frame. In our case, the running variable, time, only moves forward and does not allow for a local comparison around the event to isolate a causal effect. The effect of the Ethereum Merge was not localized at a specific time frame immediately around the event but was persistent after the event. This made an event study more suitable than regression discontinuity in time for our analysis, as we are capturing both the immediate impact of the Ethereum Merge and the longer-term trends associated with the Ethereum Merge (Hausman and Rapson, 2017).

4 Results

In this section, we present the empirical results quantifying the causal impact of resource flexibility on the decentralization of blockchain consensus in response to three shocks—policy, infrastructure, and technical. For each shock, we first present model-free evidence highlighting the pre- and post-shock values of decentralization. Then, we present DiD, SDiD, and event study estimates of the impact of the shock on decentralization. Taken together, the dynamics

of the shocks and the comparisons between the three shocks provide strong suggestive evidence that virtualization plays a role in influencing blockchain decentralization.

4.1 China Bans Crypto Mining

First, we studied the effect of virtualization on the decentralization of the consensus of Bitcoin versus Ethereum in response to China's ban on crypto mining on May 15, 2021, as explained in Section 3.1. Figure 3 shows the decentralization of Bitcoin and Ethereum as measured by entropy in bits (as used to measure the quantity of digital information). Since our empirical analyses are based on DiD estimates, they rely on the parallel trends assumption. First, we show visual proof of the parallel trends assumption before the shock in Figure 3. While the entropy measures seem to display some seasonality (perhaps due to varying electricity costs for nodes across geographies), the decentralization metrics are similar in absolute terms as well as parallel in trends. We note the possibility of anticipation by miners in the few weeks leading up to the shock.

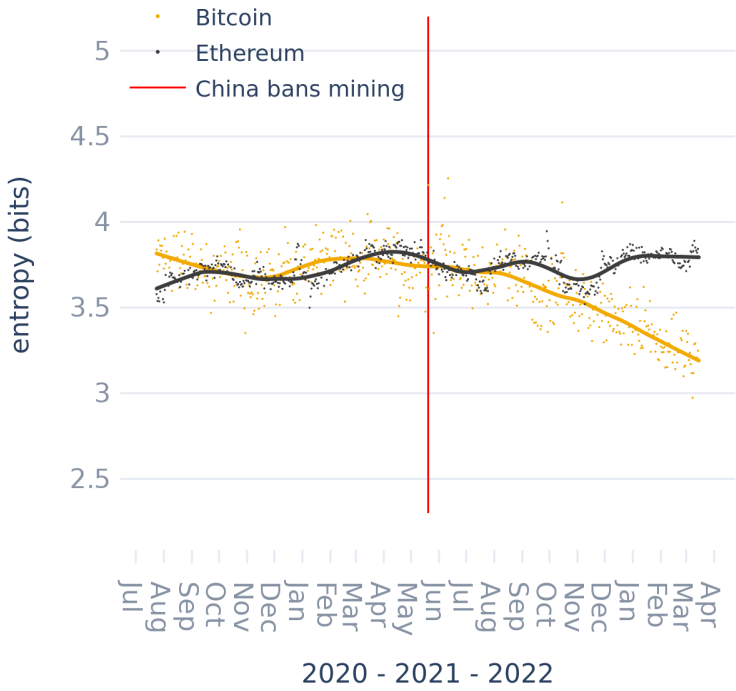

Figure 3: Decentralization (in bits) of Bitcoin and Ethereum around China's ban on crypto mining.

Post China's ban on crypto mining, we observe a significant decline in entropy for Bitcoin

over many months (Figure 3). In contrast, we see little to no decline in entropy for Ethereum. To quantify the effect, we analyzed the results of the DiD specifications in Equation 1 to estimate the impact of resource flexibility on decentralization in response to the shock, by using the variation in hardware requirements as a proxy for resource flexibility, as explained in Section 3.1. We denote Bitcoin = $\text{Chain}_{\text{Bitcoin}}$ for simplicity. In our analysis, we found that Bitcoin's entropy was reduced by 0.209 bits in comparison with Ethereum, from 3.722 bits before the ban (Table 2). To further explore this effect, we present case studies in Section D of the Appendix where we can observe individual nodes and mining pools known to be based in China that experience declines in block production.

Dependent	*Entropy*		*Nodes*		*Gini*		*Nakamoto*		*HHI*	
variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	-0.209***	-0.199***	5.050**	5.162**	0.047***	0.047***	-0.414**	-0.387**	0.012*	0.011*
	(0.061)	(0.059)	(1.595)	(1.580)	(0.011)	(0.011)	(0.138)	(0.134)	(0.005)	(0.005)
Exposure		-0.042***		-0.444**		0.003		-0.106***		0.003***
		(0.012)		(0.154)		(0.002)		(0.025)		(0.001)
Bitcoin	0.020	0.020	-33.980***	-33.980***	-0.327***	-0.327***	1.577***	1.577***	-0.037***	-0.037***
	(0.026)	(0.026)	(1.311)	(1.312)	(0.010)	(0.010)	(0.062)	(0.062)	(0.003)	(0.003)
Post	0.020	0.031	-6.893***	-6.778***	-0.033***	-0.034***	-0.060	-0.032	0.001	0.000
	(0.027)	(0.027)	(1.439)	(1.438)	(0.008)	(0.008)	(0.084)	(0.084)	(0.004)	(0.004)
Intercept	3.722***	3.722***	53.740***	53.740***	0.801***	0.801***	2.990***	2.990***	0.131***	0.131***
	(0.020)	(0.020)	(1.282)	(1.282)	(0.007)	(0.007)	(0.010)	(0.010)	(0.002)	(0.002)
Observations	1202	1202	1202	1202	1202	1202	1202	1202	1202	1202

Notes: *p<0.05; **p<0.01; ***p<0.001. Standard errors are clustered by blockchain and month.

Table 2: China bans crypto mining

It is important to note that Bitcoin and Ethereum experienced different exposures to the mining ban, which could have led to omitted variable bias (Borusyak and Hull, 2020). We therefore controlled for exposure by including the covariate Exposure_{it}, described by Equation 2. Then, we observed that the magnitude of the effect size for Treatment decreased slightly to -0.199 while the Exposure covariate saw a small but significant effect of -0.042 (Table 2, Column 2). The slight decrease in the effect size after controlling for exposure underscores the robustness of our findings, suggesting that the observed decline in Bitcoin's decentralization is not solely attributable to differences in exposure to the mining ban but also likely influenced by the underlying virtualization characteristics of the blockchain.

We also estimated the effects for the other metrics (Table 2). Surprisingly, we observed that the number of nodes increased for Bitcoin after the shock by around 5 nodes.[22] Despite the increase in the number of nodes, all other metrics, which measure the *distribution* of block production among nodes, indicate that Bitcoin became more centralized after the shock. The Gini coefficient increased by 0.047, the Nakamoto coefficient decreased by around 0.414, and HHI increased by 0.012, all of which indicate the post-shock centralization of Bitcoin.[23] Thus, the overall findings underscore the significant negative impact of the mining ban on the decentralization of Bitcoin.

Moreover, we were surprised to see the extent to which Bitcoin's decentralization decreased over time after the shock. Indeed, we observe visually that Bitcoin did not fully recover from the shock in the months following the shock. To quantify the anticipatory and post-treatment effects of the shock on Bitcoin's decentralization, we performed DiD analysis on the same dataset as previously but with Equation 3, which includes the lead and lag terms, $\sum_{\lambda=-T}^{T} \delta_\lambda \text{Treatment}_{it\lambda}$. In Figure 4, we show that the treatment lag lasts for over 180 days after the shock, and the coefficients are displayed in Section E of the Appendix. We attribute this lagged effect to the rolling enforcement of the mining ban and the consensus participation behavior across nodes, which we demonstrate with multi-period difference-in-difference estimation in Section B and node-level decomposition in Section D of the Appendix. As we will see in the next shock, this long-lasting effect in Bitcoin sharply contrasts with the fast recovery in the decentralization of Solana, further suggesting the potential role of virtualization.

4.2 Hetzner Shuts Down Solana Validators

Second, we studied the effect of Hetzner shutting down Solana validators on November 2, 2022, on the decentralization of Solana's consensus mechanism, as explained in Section 3.1.

[22] This positive effect on the number of nodes is possibly due to the dissolution of mining pools in China due to the ban. See Section D for case studies of China's mining pools.

[23] Note that Gini and HHI are bounded between 0 and 1. The units are in bits for Entropy and nodes for Nodes and Nakamoto.

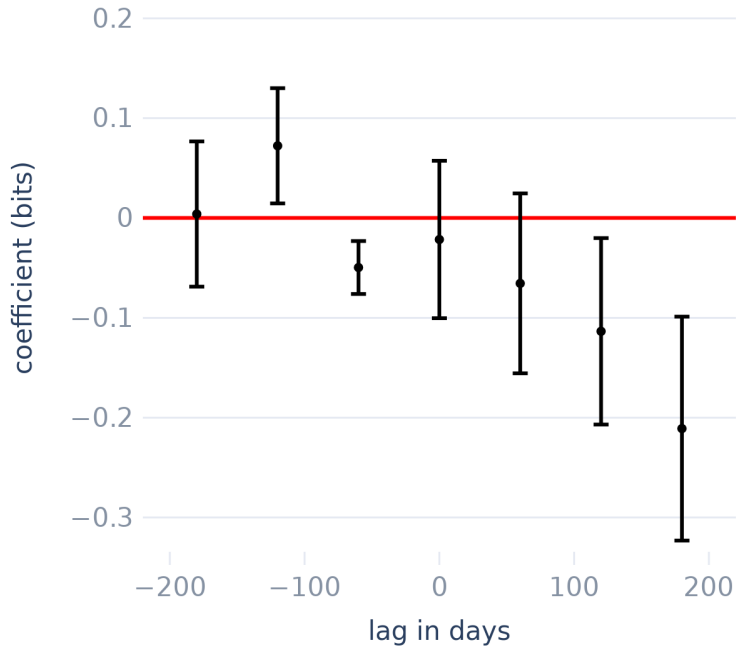

Figure 4: Coefficients for the lead and lag effects Treatment$_\lambda$ of China's ban on crypto mining for λ lags in days. Error bars indicate 95% confidence intervals.

We compared Solana to a synthetic control constructed from a weighted average of the other blockchains which can all experience infrastructure shocks. We note that this difference-in-difference estimation does not estimate the causal effect of virtualization, as does our analysis of China's mining ban, but the effect of the shock itself on decentralization. As such, we can compare the dynamics of the recovery in blockchain decentralization between this and the previous shock. Figure 5 shows the entropy values of Solana and the synthetic control. Before estimating DiD effects, we visually demonstrate the parallel trends assumption between Solana and the synthetic control by showing the daily entropy of the blockchains over the period before the shock.

On the day of and following Hetzner's shutdown of Solana validators, we observed an immediate, visible drop in the entropy of Solana. By comparison, we saw no shock to the entropy of the synthetic control. To quantify this effect, we estimated the synthetic difference-in-difference model in Equation 4 with Solana and a synthetic control of four other Proof-of-Stake blockchains. We denote Chain$_{\text{Solana}}$ as Solana for clarity. We found that

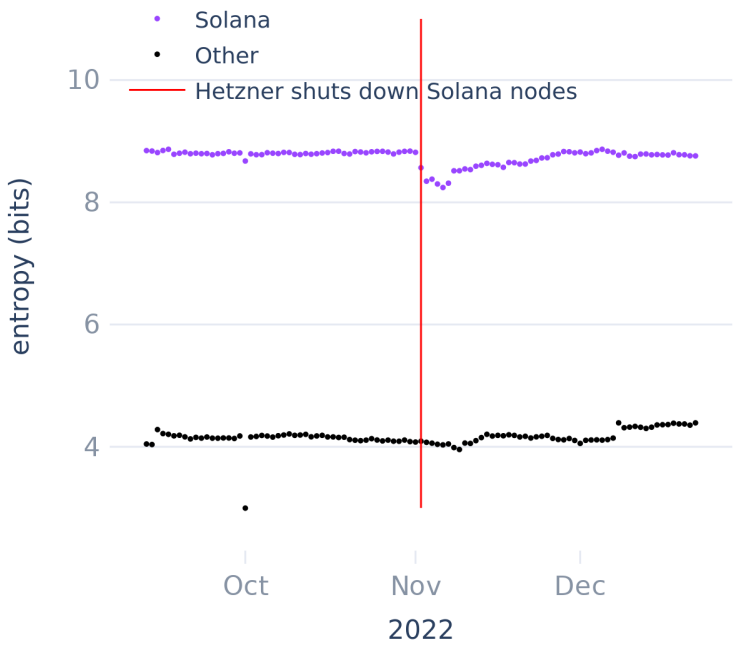

Figure 5: Decentralization (in bits) of Solana and a synthetic control around Hetzner's shutdown of Solana validators.

the shutdown reduced 0.345 bits of entropy of Solana ten days before and after the shock, in comparison with the synthetic control (Column 1, Panel A in Table 3). Then having decomposed the entropy measure into the nodes and measures of the distribution of block production across nodes, we found that around 316 nodes went offline in this period (Column 1, Panel B in Table 3). This decline was accompanied by an increase in the Gini coefficient of 0.042 and a decrease in the Nakamoto coefficient of -6.105 nodes but no significant change in HHI (Column 1, Panels C, D, and E in Table 3). These findings demonstrate a substantial negative impact of the infrastructure disruption on the decentralization of Solana.

To compare Hetzner's shutdown to China's mining ban, we examined the dynamics of the impact of Hetzner's shutdown on Solana's decentralization. We estimated lagged effects using Equation 3.[24] In our analyses, we found that while the impact of China's ban was present six months after the shock, the impact of Hetzner's shutdown was short-lived. In Figure 5, we visually observe a decrease and a subsequent recovery of Solana's entropy within

[24]Note that we used blockchain fixed effects instead of a synthetic control in SDiD since we wanted to estimate multiple treatment effects, unlike in Table 3.

Dependent variable	±10 days (1)	±20 days (2)	±30 days (3)	±40 days (4)	±50 days (5)
Panel A: Entropy					
Treatment	-0.345	-0.317	-0.230	-0.178	-0.218
	(0.090)	(0.225)	(0.178)	(0.197)	(0.200)
Panel B: Nodes					
Treatment	-316.8	-274.0	-237.7	-189.0	-115.1
	(3.8)	(11.5)	(2.3)	(35.3)	(97.1)
Panel C: Gini					
Treatment	0.042	0.025	0.009	-0.014	-0.074
	(0.006)	(0.007)	(0.001)	(0.058)	(0.184)
Panel D: Nakamoto					
Treatment	-6.105	-5.646	-4.269	-2.875	-0.919
	(0.064)	(0.379)	(0.450)	(0.125)	(0.662)
Panel E: HHI					
Treatment	-0.001	0.002	0.002	0.001	-0.004
	(0.005)	(0.009)	(0.008)	(0.000)	(0.011)
Observations	42	82	122	162	202

Notes: Standard errors are derived from placebo tests. P-values are not provided in this analysis.

Table 3: Hetzner shuts down Solana nodes.

around 40 days of the shock. We also quantified the recovery empirically in two ways. First, we estimated the DiD specification in Equation 1 at ±10 to ±50 days before and after the shock in Table 3. We see that the effect at ±50 days is only 63% of the effect at ±10 days. Second, we estimated the leading and lagging treatment effects using the specification in Equation 3. In Figure 6, we see that the shock reduced entropy by 0.359 bits immediately after the shock, but entropy increased in the 40 days following the shock.[25] The coefficients for Entropy, Nodes, Gini, Nakamoto, and HHI are displayed in Section F of the Appendix.

To further test robustness, we estimated Equation 5 of Solana's entropy as an event study which revealed results consistent with our difference-in-difference estimation. Our results, shown in Table 4, demonstrate that the shutdown had a negative effect of 0.344 bits before the

[25]Note that the post-shock positive lagged effects in Figure 6B are measuring the increase in entropy following the initial large dip in response the shock; Table 3 shows the average post-shock effect.

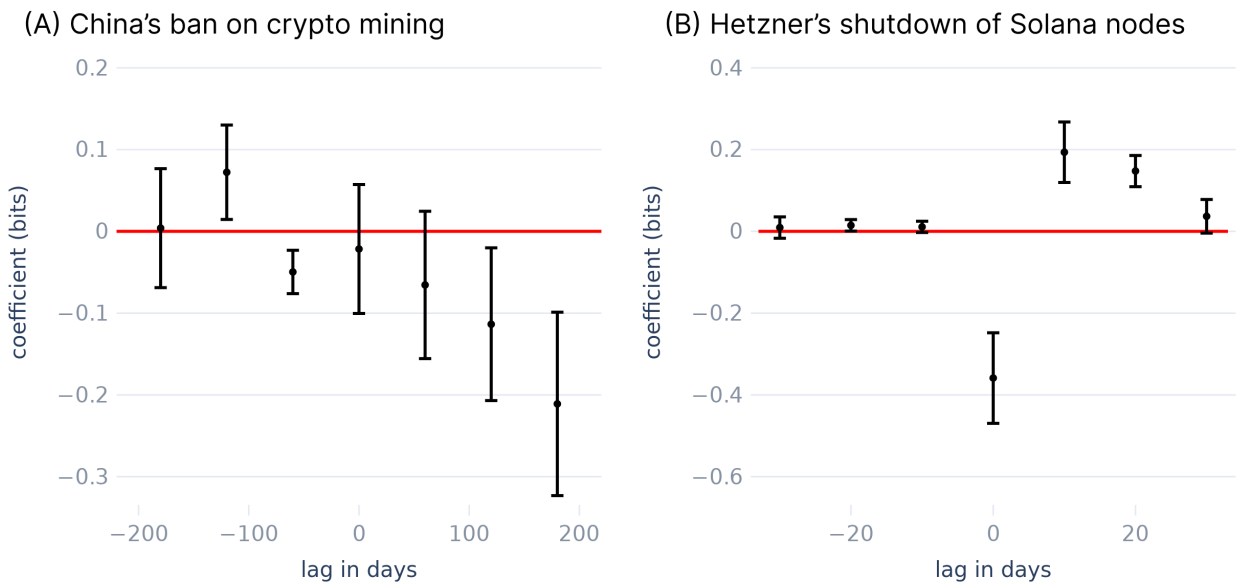

Figure 6: Coefficients for the lead and lag effects Treatment_λ on the entropy of (A) China's mining ban and (B) Hetzner's shutdown of Solana validators for λ lags in days. Error bars indicate 95% confidence intervals.

shock. All other metrics also show centralizing effects after the shutdown, with the number of nodes decreasing by around 338, the Gini coefficient increasing by 0.028, the Nakamoto coefficient decreasing by -7.919, and HHI increasing by 0.001. Then, the slope of the entropy after the shock is small but positive and significant at 0.008^{***} bits per day 60 days after the shock. By dividing the shock of 0.344 bits by the positive slope post-shock of 0.008 bits per day, we calculated that Solana took 43 days to fully recover from the shock. These results are again consistent with our DiD estimation.

The difference in the dynamics of China's ban and Hetzner's shutdown may be attributed to two key factors: one in the nature of the shock and one in the design of the blockchains. First, China's ban was a policy ban while Hetzner's shutdown was an infrastructure shutdown. Policy changes take time to comply with and enforce, especially for distributed networks such as Bitcoin.[26] In contrast, Hetzner themselves locked the servers that were validating for Solana.[27] So the impact of an infrastructure shock may be more immediate, as we have seen

[26]We show more evidence of this rolling ban of Bitcoin miners in Sections B, C, and D of the Appendix.
[27]See tweets of Hetzner's notice by @foldfinance and @solblaze_org. Accessed August 23, 2023.

Dependent variable	*Entropy*	*Nodes*	*Gini*	*Nakamoto*	*HHI*
Post	-0.344***	-338.330***	0.028***	-7.919***	0.001***
	(0.033)	(32.535)	(0.005)	(0.608)	(0.000)
Days	0.000	3.608***	0.000***	0.005	0.000**
	(0.000)	(0.095)	(0.000)	(0.010)	(0.000)
(Days × Post)	0.008***	-1.789	-0.001***	0.243***	-0.000***
	(0.001)	(0.958)	(0.000)	(0.022)	(0.000)
Intercept	8.812***	2065.837***	0.721***	55.878***	0.006***
	(0.006)	(2.595)	(0.001)	(0.237)	(0.000)
Observations	121	121	121	121	121

Notes: *p<0.05; **p<0.01; ***p<0.001. Standard errors are clustered by blockchain and month.

Table 4: Event study on Hetzner shutting down Solana nodes

in the data.

Second, Bitcoin and Solana differ in their consensus mechanisms. Bitcoin uses specialized hardware (i.e., application-specific integrated circuits *ASICs*) and energy-intensive computation to achieve PoW consensus. ASICs are, by definition, specific to Bitcoin mining and are not readily available in other cloud facilities. By comparison, Solana uses staked native tokens, SOL, to achieve PoS consensus. In Solana, one can relatively easily create new validators by transferring their tokens to another cloud server or even a home server.[27] Our results thus show that the flexibility of resources used in the consensus mechanism—*i.e.*, tokens *versus* hardware—appears to have reduced frictions in recovering from shocks, suggesting a role in influencing sustained decentralization.

4.3 The Ethereum Merge

In the two previous natural experiments, we examined the effects of virtualization on decentralization in response to shocks across blockchains that differed in their level of virtualization. In this section, we investigate the effect of a technical upgrade—the Ethereum Merge on September 15—on the decentralization of a single blockchain over time, as explained in Section 3.1. The Merge was a software upgrade that changed Ethereum's consensus mechanism from PoW to PoS. Thus, the Merge removed the need for physical hardware,

thereby making Ethereum more virtualized by definition, and allowed anyone with 32 ETH and even consumer-grade computers to participate in Ethereum's consensus. Figure 7 shows the entropy of Ethereum's consensus before and after the shock. Before estimating the effects of the event study, we can visually observe the stepwise increase in both the means and the variance of decentralization specifically at the time of the Merge.

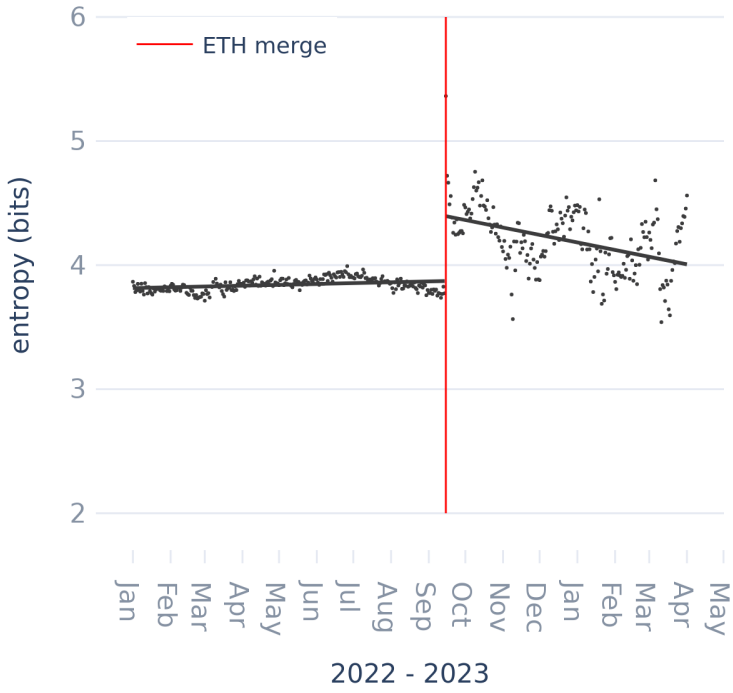

Figure 7: Daily decentralization (in bits) of Ethereum's consensus around the Ethereum Merge.

To quantify this impact on Ethereum's decentralization, we estimated the specification for the Event Study in Equation 5. We found that the Merge increased Ethereum's decentralization by 0.523 bits of entropy from 3.872 bits (Table 5). This increase in entropy can be attributed to the addition of around 405 new nodes and an increase in the Gini coefficient of 0.134, which indicates a rise in the inequality of the distribution of block production. Interestingly, the analyses show no significant change in the Nakamoto coefficient or the HHI. This implies that the number of nodes that control 51% of the network remained similar despite the increase in the broader base of the network, consisting mainly of smaller new entrants. Despite the insignificant change in Nakamoto and HHI, the overall results add to

the evidence that resource flexibility plays a role in influencing decentralization, since PoS, by not relying on hardware for validation unlike PoW, offers a form of virtualization that allows for broader participation. The transition to Proof of Stake (PoS) facilitates an increase in validator numbers due to its lower barriers to entry, which diversifies participation and enhances network entropy.

Dependent variable	*Entropy*	*Nodes*	*Gini*	*Nakamoto*	*HHI*
Post	0.523***	405.953***	0.134***	0.233	-0.002
	(0.103)	(12.077)	(0.004)	(0.153)	(0.008)
Days	0.000	0.023***	0.000***	0.000	-0.000**
	(0.000)	(0.004)	(0.000)	(0.000)	(0.000)
(Days × Post)	-0.002***	-0.348***	-0.000**	-0.002	0.000*
	(0.001)	(0.081)	(0.000)	(0.001)	(0.000)
Intercept	3.872***	52.104***	0.774***	3.028***	0.122***
	(0.030)	(0.617)	(0.002)	(0.026)	(0.003)
Observations	456				

Notes: *p<0.05; **p<0.01; ***p<0.001. Standard errors are clustered by blockchain and month.

Table 5: The Ethereum Merge

To understand the small but statistically significant negative slope (-0.002^{***}) in Ethereum's entropy after the Merge, we first examined whether the daily number of validators that produced blocks decreased after the Merge. In Figure 8, we see that the number of nodes after the Merge did not decrease, though there were some large downswings (the second of which was perhaps due to the Shanghai Upgrade which allowed validators to withdraw their staked ETH). In fact, we see a positive slope of 2.27 validators per day continuing through the Merge. To further test whether there was any relationship between entropy and the number of validators after the Merge, we computed Pearson's correlation coefficient between entropy and validator count and found no correlation (i.e., Pearson's R $= -0.05$, $p = 0.44$).

Next, we investigated whether any particular group of validators had a role in the negative slope of entropy after the Merge, especially given the positive slope in the number of validators over time. To do so, we simulated "knockout" experiments in which we removed particular

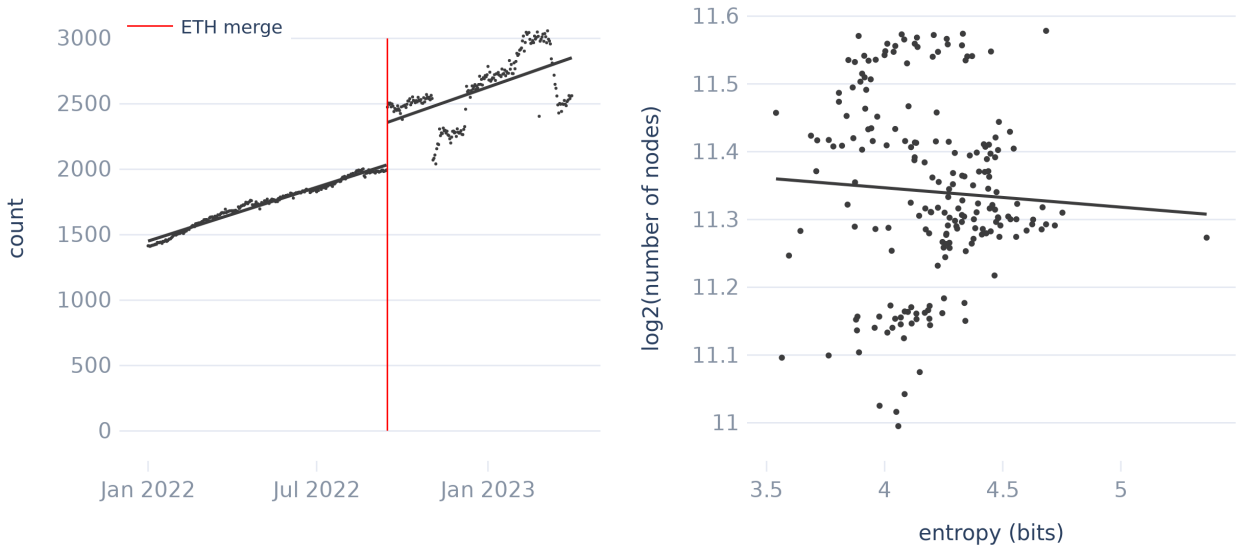

Figure 8: (Left) Daily number of nodes that have mined at least one block. (Right) No significant correlation between the $\log_2$ daily number of nodes and daily entropy (Pearson's R $= -0.05$, $p = 0.44$).

groups of validators from the analysis after the Merge and recomputed the entropy. The groups included centralized exchanges (CEX), Staking Pools such as Lido, MEV (maximal extractable value) builders, and "Others" which include individuals or are unlabeled.[28] We see in Figure 9 that removing MEV builders from the data had a substantial positive effect on both the slope and the absolute value of entropy. This result suggests that MEV builders are a *centralizing* force on Ethereum's consensus. Removing CEX and Staking Pools validators has negligible effects while removing "Others" has a negative effect, as one may expect from removing individuals from consensus.

MEV builders are software applications that aggregate and order transactions in such a way as to maximally extract value from the transactions. For example, when a user places a transaction for a large purchase of token A, an MEV bot will perform a "sandwich attack" in which the bot purchases token A before the large trade raises the price of token A, and sells token A right after the trade increases the price of token A. These bots and others pay fees to MEV builders to order transactions in a block, and MEV builders in turn pay validators to produce that block on the blockchain. As a result, an MEV builder first receives the reward for

[28] See ethereum.org for a detailed explanation of MEV.

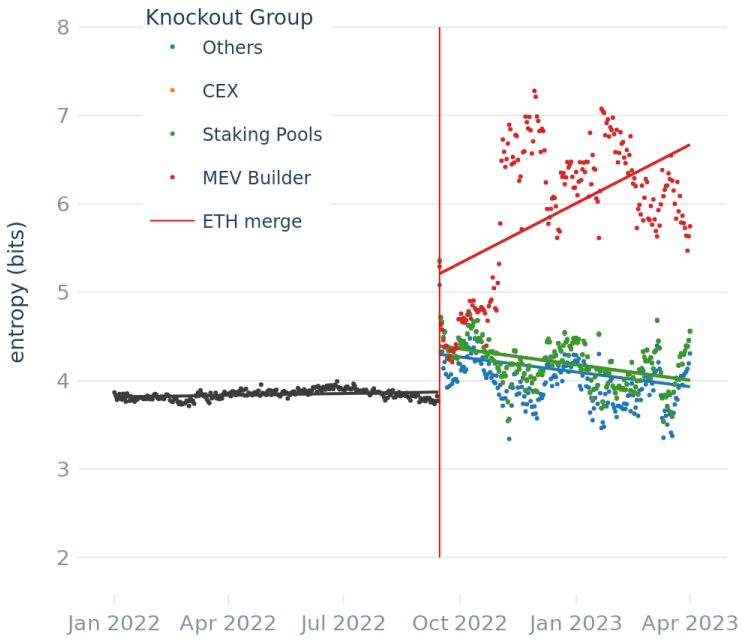

Figure 9: Knockout simulations of Ethereum's entropy after the Merge.

building a block and then sends the block reward to the validator who "sold" its block space to the MEV builder. Because MEV builders are trusted entities, MEV builders are a substantial centralizing force in Ethereum's consensus. Thus, the results suggest that the composability of protocols through proposer-builder separation (PBS) or otherwise—while introducing opportunities for open protocols, market competition, and interoperability—opens up the risk of centralization in each component. This unexpected discovery, which emerged from our analysis, does not directly implicate the virtualization hypothesis but is an interesting element of the centralization and decentralization of blockchains that warrants further investigation. While such analysis is beyond the scope of this paper, we encourage such work in future research.

5 Discussion and Conclusions

We used empirical data from major blockchain networks to investigate the impact of significant shocks to blockchain consensus decentralization. Specifically, we analyzed three policy-related, infrastructure-related, and technical shocks in the blockchain industry, and our findings

strongly suggest that resource flexibility plays a significant role in the decentralization dynamics of blockchains and enables sustained blockchain consensus decentralization.

The results indicate a statistically and economically significant influence of resource flexibility on the sustained decentralization of blockchain consensus. The difference is apparent when we juxtapose the time-lagged effects of the natural experiments in Figure 6. The China crypto mining ban led to a notable decline in Bitcoin's decentralization, while the Hetzner shutdown of Solana nodes illustrated the greater resilience of more resource-flexible consensus mechanisms. Our results suggest that the nature of the shocks, combined with the inherent design of the blockchains, played a crucial role in determining their decentralization trajectories. In particular, Bitcoin, with its reliance on specialized hardware for PoW consensus, faced challenges in adapting to policy shocks, whereas Solana's PoS mechanism allowed for a more agile response to infrastructure disruptions. While we have robust evidence on the immediate effects of these shocks, further analysis may elucidate the long-term implications and user behaviors in response to these events.

5.1 Implications for Blockchains

Our study provides empirical insights into the factors affecting blockchain decentralization, which may inform various stakeholders in the blockchain ecosystem. First, the data suggests that design elements emphasizing resource flexibility could be associated with a blockchain's ability to maintain decentralization in the face of external shocks. However, this flexibility may come with trade-offs in other aspects such as security or efficiency. Second, operational decisions appear to have a measurable impact on a blockchain's decentralization. These decisions could either enhance or compromise a blockchain's structural integrity, contingent on its specific design. Lastly, policy decisions targeting blockchains have the potential for broader, systemic effects. Therefore, if the goal is to maintain a balance between innovation and safety, policymakers might consider the unintended consequences of their regulations on the decentralization and overall stability of multiple blockchains.

Blockchain Design Our study provides empirical evidence that suggests design choices in blockchain architecture have implications for how these systems respond to external shocks. For instance, our data, as shown in Figure 3 and Table 2, indicates that Bitcoin's decentralization was more adversely affected by China's crypto mining ban compared to Ethereum's. This observed difference may be linked to their respective architectural designs; Bitcoin's reliance on specialized hardware could make it more vulnerable to policy-induced shocks. Conversely, Solana's quicker recovery after the Hetzner shutdown could be indicative of the benefits of a more virtualized architecture.

Furthermore, alternative design approaches like stateless nodes and blobs could potentially increase participation in blockchain consensus, as suggested by our findings and supported by existing literature (Rustgi, 2023).[29] Traditional full nodes often require substantial storage and computing power, typically necessitating centralized servers. Stateless nodes, by contrast, could lower these requirements, possibly enabling operation on less powerful devices like smartphones. If the goal is to enhance resilience against external shocks, designs that lower barriers to participation while maintaining the integrity of the system could be considered. However, it is important to note that these design choices come with their own sets of trade-offs that warrant further investigation. For example, our analysis of the Merge shows that despite the increase in network entropy and participation, there remains a concentration of power among the largest validators, highlighting the ongoing challenges in achieving further decentralization. This emphasizes the need for further refinement of consensus mechanisms for further decentralization.

Blockchain Operations The empirical evidence from the three types of shocks—policy, infrastructure, and technical—provides insights into the operational dynamics of blockchain consensus. In this context, we define blockchain operators as individuals or entities involved in the management of validators, either directly or indirectly, through activities like monitoring, coordination, or infrastructure development (Dunbar, 2023). Our findings indicate

[29] See ethereum.org's explanation on statelessness and EIP-4844 for details on blobs. Accessed April 15, 2024.

that reliance on a limited number of infrastructure providers could expose blockchains to vulnerabilities, as evidenced by the prolonged impact on Bitcoin's decentralization following China's ban. This suggests that if resilience against regulatory changes is a priority, operators might consider diversifying their infrastructure across multiple jurisdictions.

Similarly, the recovery dynamics observed in Bitcoin and Solana after the respective shocks point to the role of the consensus mechanism in determining a blockchain's resilience to external factors. Depending on the consensus mechanism in use, different operational strategies may be more or less effective in mitigating the impact of such shocks. For instance, operators of PoW blockchains might need to consider factors like hardware supply, while those managing PoS blockchains could focus on the geographic distribution of nodes. Monitoring efforts by organizations like the Solana Foundation indicate that ongoing scrutiny of decentralization metrics can provide valuable data for operational decision-making.[30] However, it is important to note that these strategies come with their own sets of trade-offs and limitations that require further study.

Blockchain Policy Our findings indicate that policy changes can exert a significant influence on the level of decentralization in blockchain networks, as evidenced by the significant centralization of Bitcoin in response to China's policy change (Figure 3, Table 2). This suggests that policymakers should be cognizant of the potential ripple effects that their decisions could have on the blockchain ecosystem, particularly in areas like decentralization and network security. While clear regulatory guidance could facilitate a more stable environment for blockchain-related services, it is important to note that such policies also come with trade-offs that could affect the broader financial ecosystem. As an increasing number of financial assets, including securities and government bonds, are being recorded on blockchains, understanding the implications of policy decisions on both the integrity and the decentralization of these platforms becomes increasingly important. However, it should be noted that the adoption of blockchain for financial assets is still a subject of ongoing debate and scrutiny.

[30] See the Solana Foundation's March 2023 Report and nakaflow.io. Accessed September 6, 2023.

5.2 Limitations and Future Research

While our study offers insights into the dynamics of blockchain design, operations, and policy, there are some limitations to consider. Our biggest limitation is that our analyses are of natural experiments in which we cannot vary just the resource flexibility of different blockchains. To better understand this limitation, we considered alternative hypotheses to ours. First, the type of event varied across shocks, which include policy-related, infrastructure-related, and technical shocks. For example, the lasting effect of China's mining ban compared to the short-lived effect of Hetzner's shutdown could be because China's mining ban was a policy shock and Hetzner's shutdown was an infrastructure shock. Intuitively, policy changes take days to weeks to implement and months to enforce, whereas infrastructure shocks can occur within days. However, we observed that Ethereum experienced the same shock as Bitcoin, but that Ethereum, whose consensus mechanism was inherently more virtualized and thus more flexible, saw little to no negative effect on its decentralization. Moreover, we perform multi-period difference-in-difference estimation and node-level decomposition in Sections B and D of the Appendix, respectively, to test for robustness of the results at finer resolutions of both time and units, further providing strong evidence for our virtualization hypothesis.

Second, public perception and market sentiment may impact a blockchain's ability to recover from shocks. Though BTC and ETH prices are highly correlated (see Section G), positive market sentiment may lead to quicker recovery as participants are more likely to reinvest, stake, or engage in the network, thereby bolstering its decentralization and resilience. To test this "market sentiment" hypothesis, we estimated Equation 5 on the price of BTC ± 60 days around China's ban, SOL ± 60 days around Hetzner's shutdown, and ETH ± 60 days around the Ethereum Merge. We found that the shocks had either no effect or even negative effects on the cryptocurrencies' prices, indicating that market sentiment did not facilitate a rapid recovery. The price of BTC hardly dropped around the ban by \$1,288.26 from \$49,350 ($p = 0.378$), and the price of SOL and ETH dropped by \$7.25 from \$30.41

($p \ll 0.001$) and by \$350.02 from \$1686.78 ($p \ll 0.001$), respectively. The coefficients for the estimation are displayed in Section H of the Appendix. The zero to negative price effects suggest that, if anything, Bitcoin should have recovered faster than Ethereum or Solana, again suggesting virtualization is the more likely explanation for the decentralization effects we estimated.

Beyond the primary limitation surrounding our hypothesis, our study has some other limitations to consider. First, our research focused on specific aspects of blockchain architecture, such as consensus mechanisms. A comprehensive understanding would require delving into other facets like smart contract design, interoperability, and scalability (Sai et al., 2021; Cong et al., 2022). Second, the quasi-experimental nature of our study might have overlooked potential behavioral patterns of forward-looking blockchain participants or the potential framing effects of initial blockchain adoption phases. Third, our observations do not span the entirety of the vast space of blockchains and their ecosystems, and it is essential to recognize that blockchain ecosystems evolve rapidly. Future research should aim to assess the long-term implications of design choices on blockchain adoption and decentralization. It would be particularly valuable to determine if the observed benefits of certain consensus mechanisms, such as PoS or PoH, remain consistent over extended periods. Fourth, our study mainly centered on well-established blockchains, potentially overlooking the dynamics of emerging or less popular blockchain platforms. We anticipate that future research will address these gaps, further refining our understanding of the virtualization hypothesis, optimal blockchain design, and the effects of blockchain policies in various global contexts.

References

Ameen, Radhwan Y., Asmaa Y. Hamo. 2013. Survey of server virtualization. URL http://arxiv.org/abs/1304.3557.

Ammous, Saifedean. 2018. *The Bitcoin Standard: The Decentralized Alternative to Central Banking*. John Wiley & Sons.

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, Stefan Wager. 2021. Synthetic difference-in-differences. *American Economic Review* **111**(12) 4088–4118.

Avan-Nomayo, Osato. 2022. 1,000 solana validators go offline as hetzner blocks server access. *TheBlock* URL https://www.theblock.co/post/182283/1000-solana-validators-go-offline-as-hetzner-blocks-server-access. Accessed: August 21, 2022.

Bakos, Yannis, Hanna Halaburda, Christoph Mueller-Bloch. 2021. When permissioned blockchains deliver more decentralization than permissionless. *Commun. ACM* **64**(2) 20–22. doi:10.1145/3442371. URL https://doi.org/10.1145/3442371.

Bamakan, Seyed Mojtaba Hosseini, Amirhossein Motavali, Alireza Babaei Bondarti. 2020. A survey of blockchain consensus algorithms performance evaluation criteria. *Expert Systems with Applications* **154** 113385. doi:https://doi.org/10.1016/j.eswa.2020.113385. URL https://www.sciencedirect.com/science/article/pii/S0957417420302098.

Bank, European Investment. 2021. Eib issues its first ever digital bond on a public blockchain. URL https://www.eib.org/en/press/all/2021-141-european-investment-bank-eib-issues-its-first-ever-digital-bond-on-a-public-blockchain. Accessed: August 28, 2023.

Beiko, Tim. 2022. Paris upgrade specification. https://github.com/ethereum/execution-specs/blob/master/network-upgrades/mainnet-upgrades/paris.md. GitHub repository.

Borusyak, Kirill, Peter Hull. 2020. Non-random exposure to exogenous shocks: Theory and applications. Working Paper 27845, National Bureau of Economic Research. doi:10.3386/w27845. URL http://www.nber.org/papers/w27845.

Brettel, Malte, Niklas Friederichsen, Michael Keller, Marius Rosenberg. 2014. How virtualization, decentralization and network building change the manufacturing landscape: An industry 4.0 perspective. *World Academy of Science, Engineering and Technology, International Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering* **8** 37–44. URL https://api.semanticscholar.org/CorpusID:2964334.

Buterin, Vitalik. 2014. Ethereum: A next-generation smart contract and decentralized application platform. URL https://ethereum.org/669c9e2e2027310b6b3cdce6e1c52962/Ethereum_Whitepaper_-_Buterin_2014.pdf. Accessed on October 13, 2023.

Buterin, Vitalik. 2017. The meaning of decentralization. Medium. URL https://medium.com/@VitalikButerin/the-meaning-of-decentralization-a0c92b76a274. Accessed October 13, 2023.

Callaway, Brantly, Pedro H.C. Sant'Anna. 2021. Difference-in-differences with multiple time periods. *Journal of Econometrics* **225**(2) 200–230. doi:https://doi.org/10.1016/j.jeconom.2020.12.001. URL https://www.sciencedirect.com/science/article/pii/S0304407620303948. Themed Issue: Treatment Effect 1.

Catalini, Christian, J. Gans. 2016. Some simple economics of the blockchain. *IEEE Transactions on Network and Service Management* doi:10.1145/3359552. URL https://dx.doi.org/10.1145/3359552.

Chowdhury, N.M. Mosharaf Kabir, Raouf Boutaba. 2010. A survey of network virtualization. *Computer Networks* **54**(5) 862–876. doi:https://doi.org/10.1016/j.comnet.2009.10.017. URL https://www.sciencedirect.com/science/article/pii/S1389128609003387.

Cong, Lin, Xiang Hui, Catherine E. Tucker, Luofeng Zhou. 2022. Scaling smart contracts via layer-2 technologies: Some empirical evidence. *SSRN* doi:10.2139/ssrn.4312355.

Cong, Lin William, Zhiguo He. 2019. Blockchain disruption and smart contracts. *The Review of Financial Studies* **32**(5) 1754–1797. doi:10.1093/rfs/hhz007. URL https://doi.org/10.1093/rfs/hhz007.

Cong, Lin William, Zhiguo He, Jiasun Li. 2021. Decentralized mining in centralized pools. *The Review of Financial Studies* **34**(3) 1191–1235. doi:10.1093/rfs/hhaa040. URL https://doi.org/10.1093/rfs/hhaa040.

Dixon, Chris. 2021. Why web3 matters. URL https://a16zcrypto.com/posts/article/why-web3-matters/. Accessed October 23, 2023.

Dixon, Chris. 2022. Crypto fund 4. URL https://a16zcrypto.com/posts/announcement/crypto-fund-four/. Accessed on October 5, 2023.

Dunbar, Stephanie. 2023. Evaluating validator decentralization: Geographic and infrastructure distribution in proof-of-stake networks. Messari. URL https://messari.io/report/evaluating-validator-decentralization-geographic-and-infrastructure-distribution-in-proof-of-stake-networks. Accessed: September 6, 2023.

Gencer, Adem Efe, Soumya Basu, Ittay Eyal, Robbert Van Renesse, Emin Gün Sirer. 2018. Decentralization in bitcoin and ethereum networks. *Financial Cryptography and Data Security*, vol. 10957. 439–457. doi:10.1007/978-3-662-58387-6_24.

Haber, Stuart, W. Scott Stornetta. 1991. How to time-stamp a digital document. *J. Cryptology* **3** 99–111. doi:10.1007/BF00196791.

Halaburda, Hanna. 2016. Digital currencies: Beyond bitcoin. *SSRN* doi:10.2139/ssrn.2865004. URL https://ssrn.com/abstract=2865004.

Halaburda, Hanna. 2018. Blockchain revolution without the blockchain? *Commun. ACM* **61**(7) 27–29. doi:10.1145/3225619. URL https://doi.org/10.1145/3225619.

Halaburda, Hanna, Guillaume Haeringer, Joshua Gans, Neil Gandal. 2022. The microeconomics of cryptocurrencies. *Journal of Economic Literature* **60**(3) 971–1013.

Harvey, Campbell R., Ashwin Ramachandran, Joey Santoro. 2021. *DeFi and the future of finance*. Wiley.

Hausman, Catherine, David S Rapson. 2017. Regression discontinuity in time: Considerations for empirical applications. Working Paper 23602, National Bureau of Economic Research. doi:10.3386/w23602. URL http://www.nber.org/papers/w23602.

Hoffman, Michał R., Luis-Daniel Ibáñez, Elena Simperl. 2020. Toward a formal scholarly understanding of blockchain-mediated decentralization: A systematic review and a framework. *Frontiers in Blockchain* **3**. doi:10.3389/fbloc.2020.00035. URL https://www.frontiersin.org/articles/10.3389/fbloc.2020.00035.

Hui, Xiang, Catherine E. Tucker. 2023. Decentralization, blockchain, artificial intelligence (ai): Challenges and opportunities. https://ssrn.com/abstract=4601788. doi:10.2139/ssrn.4601788.

Iansiti, Marco, Karim R. Lakhani. 2017. The truth about blockchain. *Harvard Business Review* URL https://hbr.org/2017/01/the-truth-about-blockchain. Accessed October 17, 2023.

Jia, Yongpu, Changqiao Xu, Zhonghui Wu, Zichen Feng, Yaxin Chen, Shujie Yang. 2022. Measuring decentralization in emerging public blockchains. *2022 International Wireless Communications and Mobile Computing (IWCMC)*. 137–141. doi:10.1109/IWCMC55113.2022.9825341.

John, Kose, Maureen O'Hara, Fahad Saleh. 2022. Bitcoin and beyond. *Annual Review of Financial Economics* **14**(1) 95–115. doi:10.1146/annurev-financial-111620-011240. URL https://www.annualreviews.org/doi/10.1146/annurev-financial-111620-011240.

Kwon, Yujin, Jian Liu, Minjeong Kim, Dawn Song, Yongdae Kim. 2019. Impossibility of full decentralization in permissionless blockchains. *Proceedings of the 1st ACM Conference on Advances in Financial Technologies (AFT '19)*. 110–123. doi:10.1145/3318041.3355463.

Lin, Q., C. Li, X. Zhao, X. Chen. 2021. Measuring decentralization in bitcoin and ethereum using multiple metrics and granularities. *2021 IEEE 37th International Conference on Data Engineering Workshops (ICDEW)* 80–87doi:10.1109/ICDEW53142.2021.00022. URL https://doi.ieeecomputersociety.org/10.1109/ICDEW53142.2021.00022.

Meichler, Mattis. 2023. London stock exchange group plans blockchain-powered digital markets business. URL https://decrypt.co/154852/london-stock-exchange-group-plans-blockchain-powered-digital-markets-business. Last accessed: September 5, 2023.

Nakamoto, Satoshi. 2008. Bitcoin: A peer-to-peer electronic cash system. URL https://bitcoin.org/bitcoin.pdf.

Rustgi, Nivesh. 2023. Future ethereum upgrades could allow full nodes to run on mobile phones: Vitalik buterin. Decrypt. URL https://decrypt.co/154990/future-ethereum-upgrades-could-allow-full-nodes-to-run-on-mobile-phones-vitalik-buterin. Accessed: September 6, 2023.

Sai, Ashish Rajendra, Jim Buckley, Brian Fitzgerald, Andrew Le Gear. 2021. Taxonomy of centralization in public blockchain systems: A systematic literature review. *Information Processing & Management* **58**(4) 102584. doi:https://doi.org/10.1016/j.ipm.2021.102584. URL https://www.sciencedirect.com/science/article/pii/S0306457321000844.

Shen, Samuel, Twinnie Siu. 2021. China's top regulators ban crypto trading and mining, sending bitcoin tumbling. *Reuters* URL https://www.reuters.com/technology/chinese-financial-payment-bodies-barred-cryptocurrency-business-2021-05-18/. Accessed: October 30, 2023.

Shirinbab, S. 2019. Performance implications of virtualization. Ph.D. thesis, Blekinge Tekniska Högskola. URL https://urn.kb.se/resolve?urn=urn:nbn:se:bth-17217.

Srinivasan, Balaji S., Leland Lee. 2017. Quantifying decentralization. URL https://news.earn.com/quantifying-decentralization-e39db233c28e. Accessed: 10-08-2023.

Suleyman, Mustafa, Michael Bhaskar. 2023. *The coming wave: AI, power and the twenty-first century's greatest dilemma*. The Bodley Head.

The Ethereum Foundation. 2023. The merge. https://ethereum.org/en/roadmap/merge/. Page last updated: August 17, 2023.

van Cleeff, André, Wolter Pieters, Roel J. Wieringa. 2009. Security implications of virtualization: A literature study. *2009 International Conference on Computational Science and Engineering*, vol. 3. 353–358. doi:10.1109/CSE.2009.267.

Wahrstätter, Anton, Jens Ernstberger, Aviv Yaish, Liyi Zhou, Kaihua Qin, Taro Tsuchiya, Sebastian Steinhorst, Davor Svetinovic, Nicolas Christin, Mikolaj Barczentewicz, Arthur Gervais. 2023. Blockchain censorship. URL http://arxiv.org/abs/2305.18545.

Wang, Guohui, T. S. Eugene Ng. 2010. The impact of virtualization on network performance of amazon ec2 data center. *2010 Proceedings IEEE INFOCOM*. 1–9. doi:10.1109/INFCOM.2010.5461931.

Wu, Keke, Bo Peng, Hua Xie, Shaobin Zhan. 2020. A coefficient of variation method to measure the extents of decentralization for bitcoin and ethereum networks. *International Journal of Network Security* **22**(2) 191–200. doi:10.6633/IJNS.20200322(2).02.

Appendix

A Robustness Check with Consensus Layer Covariates

To test the robustness of our main difference-in-difference results of China's mining ban in Section 4.1, we included consensus layer covariates in our analysis. As explained in more detail in Section 3.1, our analyses focus on the consensus layers of Bitcoin and Ethereum. While the two blockchains differ in the types of applications they support, the consensus layer is agnostic to the application layer. The consensus layer simply manages the secure production of blocks (Nakamoto, 2008; Buterin, 2014) and indirectly the economic incentives of block production.

In our robustness tests, we added daily metrics related to block production and economic incentives as covariates. For block production, metrics include the number of transactions, the amount of data stored in megabytes (MB), and the estimated hashrates in megahashes per second (MH/s).[31] For economic incentives, metrics include the token price, average transaction fee, and block reward, all in US dollars ($).[32] The Bitcoin metrics were obtained from blockchain.com, and the Ethereum metrics were obtained from etherscan.io.[33] We removed the Exposure variable since it is directly captured by and derived from the hashrates.

In our robustness check, we found that our overall results are robust to these covariates. The treatment effects on entropy, the Nakamoto coefficient, and HHI are all qualitatively similar, with slightly reduced effects on entropy and HHI and a slightly increased effect on the Nakamoto coefficient. Interestingly, the treatment effects disappear for the Nodes and Gini metrics. This result suggests two nuanced effects and highlights the importance of using multiple decentralization metrics to understand decentralization. First, the shock had an indistinguishable effect on the number of nodes in both Bitcoin and Ethereum. Second, the

[31] See Section 3.2 for an explanation of hashrates. For reference, the hashrates were 174 exahashes per second for Bitcoin and 632 terahashes per second for Ethereum at their peak before the ban.

[32] We used block reward instead of miner revenue since miner revenue can be derived from the other covariates: the block reward, the average transaction fee, and the number of transactions.

[33] Only weekly metrics were available for Bitcoin, so the daily metrics were forward-filled from the weekly metrics.

inequality in block production among the remaining nodes, as measured by the Gini coefficient, remained the same. Additionally, the economic incentives (*i.e.*, token price, transaction fee, and block reward) have significant but negligible effects on the order of less than 10^{-12} per dollar. The effects from the number of transactions are also significant but negligible on the order of 10^{-8} per transaction for the Nodes and Gini metrics.

Dependent variable	*Entropy*		*Nodes*		*Gini*		*Nakamoto*		*HHI*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Treatment	-0.209***	-0.152**	5.050***	0.134	0.047***	0.005	-0.414***	-0.477***	0.012***	0.010*
	(0.049)	(0.048)	(1.229)	(1.112)	(0.010)	(0.010)	(0.094)	(0.117)	(0.003)	(0.004)
Bitcoin	0.020	0.296**	-33.980***	-33.920***	-0.327***	-0.370***	1.577***	1.776***	-0.037***	-0.055***
	(0.018)	(0.103)	(1.091)	(1.951)	(0.009)	(0.025)	(0.053)	(0.298)	(0.002)	(0.007)
Post	0.020	-0.045	-6.893***	-3.501**	-0.033***	-0.011	-0.060	0.089	0.001	0.004
	(0.020)	(0.028)	(1.061)	(1.200)	(0.006)	(0.008)	(0.061)	(0.059)	(0.003)	(0.003)
Intercept	3.722***	3.672***	53.740***	60.209***	0.801***	0.848***	2.990***	2.902***	0.131***	0.140***
	(0.015)	(0.069)	(0.948)	(1.627)	(0.005)	(0.013)	(0.008)	(0.249)	(0.002)	(0.007)
Block Size (MB)		0.147		1.167		0.002		0.467*		-0.010
		(0.075)		(1.148)		(0.015)		(0.212)		(0.005)
Hashrate (MH/s)		-0.004***		-0.034***		0.000		-0.006***		0.000***
		(0.000)		(0.007)		(0.000)		(0.001)		(0.000)
Number of Tx		0.000		-0.000**		-0.000**		0.000		-0.000
		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
Price ($)		0.000		0.000***		0.000***		0.000		-0.000
		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
Fee per Tx ($)		-0.000		-0.002		-0.000		-0.003		0.000*
		(0.001)		(0.020)		(0.000)		(0.003)		(0.000)
Reward ($)		0.000		-0.000**		-0.000**		-0.000		-0.000
		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)
Observations	1202	1202	1202	1202	1202	1202	1202	1202	1202	1202

Notes: *p<0.05; **p<0.01; ***p<0.001. Standard errors are clustered by blockchain and month.

Table A.1: Robustness check of difference-in-difference estimation of China's mining ban in Section 4.1 with consensus layer covariates.

B Multi-Period Difference-in-Difference

To account for the rolling enforcement of China's policy change banning crypto mining, we decomposed our difference-in-difference estimation in Section 4.1 into multiple periods. Instead of simply having a single date that divides the time series into $\text{Post}_t = 0$ before May 15, 2021, and $\text{Post}_t = 1$ after, we created both During_t and Post_t variables. The During_t variable is a binary indicator variable that equals one from May 15, 2021, to the end of July 2021 and represents the days that the ban was being enforced. The Post_t variable is a binary indicator variable as before, but it now equals one starting on August 2021. We also reformulated the Exposure_{it} as the cross-term $(\text{During}_t \times \text{Exposure}_i)$, instead of $(\text{Post}_t \times \text{Exposure}_i)$, to account for the exposure occurring during the rolling ban. We then specified a multi-period difference-in-difference estimation by

$$\begin{aligned} \text{Entropy}_{it} = \delta_1(\text{During}_t \times \text{Bitcoin}_i) + \delta_2(\text{Post}_t \times \text{Bitcoin}_i) \\ + \beta_1 \text{Exposure}_{it} + \beta_2 \text{Bitcoin}_i + \beta_3 \text{During}_t + \beta_4 \text{Post}_t + \epsilon_{it}, \end{aligned} \tag{B.1}$$

where the variables are the same as explained in Equation 1.

When we estimated Equation B.1, we found a similar result as in Section 4.1 of the main manuscript. In Column 1 of Table B.1, we observed that there were no significant effects during the rollout, as seen in the following terms: During, (During × Bitcoin), and (During × Exposure). However, we did see a negative effect of (Post × Bitcoin) after the rollout. This negative effect is similar to our results in Section 4.1 but also slightly larger because it now excludes the more minor effects during the rollout of the ban. The consistency of the negative effect observed post-rollout with our earlier findings in Section 4.1 not only validates our initial results but also suggests that the impact of the ban on Bitcoin's decentralization was more pronounced when isolating the post-rollout period, thereby reinforcing the robustness of our analysis.

While the above multi-period analysis does not reveal any significant effects during the

Dependent variable	*Entropy*		*Nodes*		*Gini*		*Nakamoto*		*HHI*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
During × Bitcoin	0.007		5.944***		0.029**		-0.145		0.005	
	(0.043)		(1.735)		(0.010)		(0.074)		(0.003)	
During × Bitcoin × Day		0.001		0.186***		0.001		-0.005		0.000
		(0.001)		(0.041)		(0.000)		(0.002)		(0.000)
Post × Bitcoin	-0.285***		4.729**		0.053***		-0.505**		0.015*	
	(0.065)		(1.622)		(0.012)		(0.154)		(0.006)	
Post × Bitcoin × Day		-0.002***		0.011		0.000**		-0.003***		0.000***
		(0.000)		(0.007)		(0.000)		(0.001)		(0.000)
Exposure	-0.001		0.092		0.004*		-0.032		0.000	
	(0.008)		(0.126)		(0.002)		(0.017)		(0.000)	
Exposure × Day		-0.007*		-0.357***		-0.001		0.010		-0.000
		(0.003)		(0.095)		(0.001)		(0.007)		(0.000)
	(0.020)	(0.029)	(1.283)	(1.308)	(0.007)	(0.010)	(0.010)	(0.050)	(0.002)	(0.003)
Bitcoin	0.020	0.045*	-33.980***	-32.731***	-0.327***	-0.316***	1.577***	1.564***	-0.037***	-0.037***
	(0.026)	(0.022)	(1.313)	(1.156)	(0.010)	(0.009)	(0.062)	(0.060)	(0.003)	(0.002)
During	-0.012	0.023	-5.469***	1.230	-0.014*	0.015	0.018	-0.054	-0.004	-0.002
	(0.027)	(0.039)	(1.648)	(1.300)	(0.007)	(0.011)	(0.011)	(0.090)	(0.002)	(0.004)
Post	0.031	-0.014	-7.399***	-2.537	-0.040***	-0.018	-0.084	0.036	0.002	0.001
	(0.030)	(0.054)	(1.442)	(2.230)	(0.008)	(0.019)	(0.112)	(0.147)	(0.005)	(0.007)
Day		0.000		-0.011*		-0.000		-0.000		0.000
		(0.000)		(0.005)		(0.000)		(0.000)		(0.000)
Intercept	3.722***	3.742***	53.740***	51.530***	0.801***	0.791***	2.990***	2.941***	0.131***	0.131***
Observations	1202	1202	1202	1202	1202	1202	1202	1202	1202	1202

Notes: *p<0.05; **p<0.01; ***p<0.001. Standard errors are clustered by blockchain and month.

Table B.1: China bans crypto mining

rollout, this may be because the rollout is gradual. That is, this simple "difference-in-means" approach may not reveal the nuanced dynamics of the rollout, as we can visually see happening in Figure 3. Thus, to get an even more fine-grain understanding of the multi-period dynamics, we extended Equation B.1 by adding the time variable Day_t, by itself and as cross-terms with $(\text{During}_t \times \text{Chain}_i)$ and $(\text{Post}_t \times \text{Chain}_i)$. We also added the time variable as a cross-term with the control variable Exposure_{it} because it is indeed the exposure that is being rolled out gradually during the ban. As such, we are even more explicitly estimating the effect of virtualization, or lack thereof, *in response to* the exposure while, at the same time, even controlling for the time-varying effects of the exposure itself. We specified this new time-varying multi-period difference-in-difference estimation by

$$\begin{aligned}\text{Entropy}_{it} = {} & \delta_1(\text{During}_t \times \text{Bitcoin}_i \times \text{Day}_t) + \delta_2(\text{Post}_t \times \text{Bitcoin}_i \times \text{Day}_t) \\ & + \beta_1(\text{Exposure}_{it} \times \text{Day}) + \beta_2\text{Bitcoin}_i + \beta_3\text{During}_t + \beta_4\text{Post}_t \\ & + \beta_5\text{Day}_t + \epsilon_{it}.\end{aligned} \tag{B.2}$$

By estimating Equation B.2, we first found that the exposure itself has the greatest negative effect of -0.007 bits per day. This initial result confirms the simple intuition that the direct shutdown of nodes should have a negative effect on decentralization. In fact, in this initial period, Bitcoin received this negative effect of the exposure slightly less than Ethereum did by 0.001 bits per day ($p = 0.045$), and this is visible in Figure 3, where we can see that despite Bitcoin's greater exposure, both Bitcoin and Ethereum experienced a similar decrease in entropy throughout the rollout period. However, after the rollout period, Bitcoin experienced a significant negative effect of -0.002 bits per day, consistent with our other results. Thus, although both Bitcoin and Ethereum initially saw a decline in decentralization, Bitcoin's more pronounced negative impact in the post-rollout period highlights the differing recovery dynamics and underscores the role of virtualization in influencing a blockchain's resilience to external shocks. Our decomposition analysis in Section D of the Appendix reveals the behavior of individual nodes and mining pools that may further elucidate the effects of the ban even more precisely.

C Bitcoin Hashrate Mining Maps

Here, we validated that the hashrates observed in Figure 2, which we used to estimate the Exposure covariate, do indeed represent exposure to China's mining ban. While we do not have geographical hashrates for Ethereum, we used the data available for Bitcoin as validation for our exposure measures using total hashrates. To validate our exposure measures, we used the Bitcoin Mining Map collected by The Cambridge Centre for Alternative Finance (The Cambridge Centre for Alternative Finance, 2023). The Centre has partnered with several Bitcoin mining pools and uses the IP (Internet Protocol) addresses of mining facility operators to identify the geographical location of a miner.

Figure C.1 shows the monthly hashrates by country from April 2021 to September 2021, around the time of China's mining ban. First, we observed on the map that monthly hashrates decreased from 43.98% in April 2021 to 34.25% in May 2021. Then, China's hashrate disappears to 0% for the months of June and July until it recovers 22.29% in August. As an alternative visualization, Figure C.2 shows a chart from Statista that displays the same data as in Figure C.1 as a bar chart (Cambridge Centre for Alternative Finance, 2022). This trend accurately matches the decrease in total hashrates from May 2021 to July 2021 in Figure 2. Second, the maximum percentage of China's hashrate was 43.98%, which is close to the exposure measure of 51.14% for Bitcoin. While it is conceivable that the partnered mining pools colluded with Chinese miners to hide their IP, that seems unlikely given the similarity between the hashrate of Chinese miners in the map and the total hashrate. The close alignment between the observed hashrate percentages in China and the total hashrate data further validates our exposure measure and our analysis surrounding the impact of China's mining ban on blockchain decentralization.

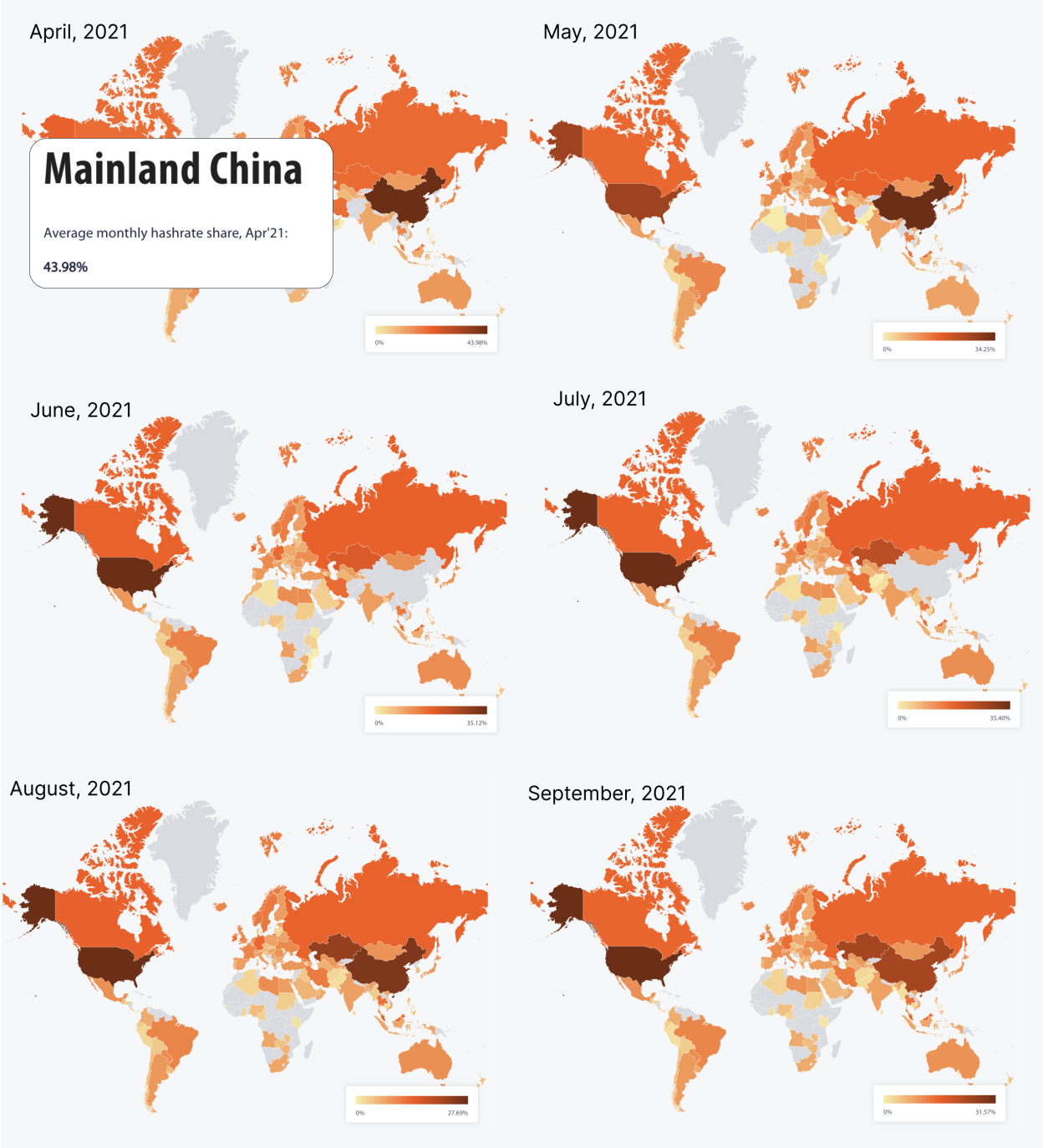

Figure C.1: The Bitcoin Mining Map by The Cambridge Centre for Alternative Finance from April to September 2021.

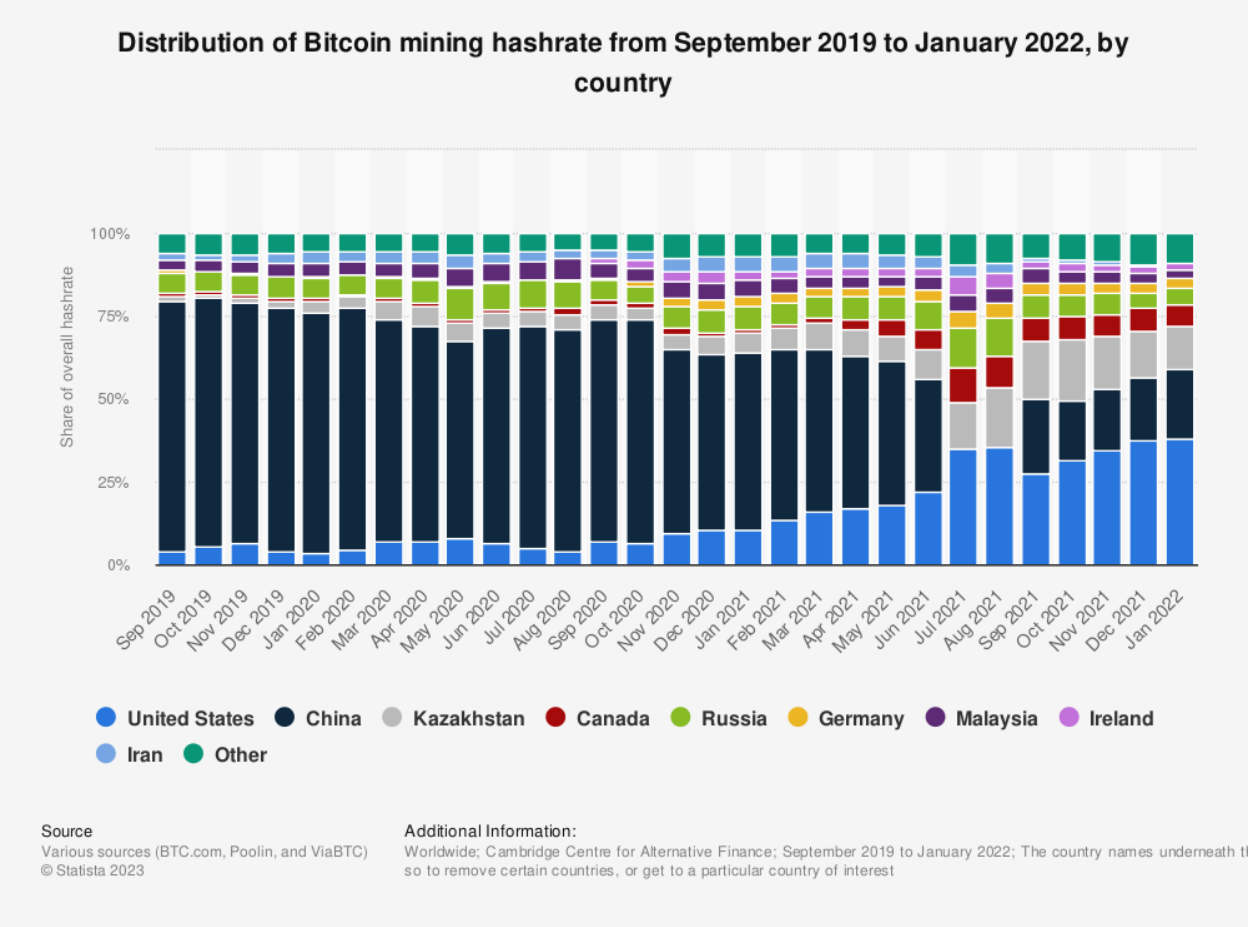

Figure C.2: Distribution of Bitcoin mining hashrate from September 2019 to January 2022, by country.

D Decomposing China's Mining Ban

To better understand how China's mining ban may have affected Bitcoin, we decomposed the effects of the ban to the level of individual nodes. This analysis does not take the rolling nature of the ban into account and is certainly not causal. Indeed, because of Bitcoin's pseudonymity, it is difficult to determine the identity or geographic location of nodes. Moreover, mining pools can obfuscate the distribution of rewards to their participants and make it even more difficult to find definitive, causal observations at the node level. However, by exploring node-level data, we can check for the robustness of our interpretations of our results and also investigate more nuanced effects of the ban, such as those to mining pools versus individual miners.

First, we examined the nodes that may have individually been affected by the ban. To construct the number of nodes lost after the shock shown in Figure D.1, we determined the unique node addresses that received block rewards from 30 days prior to the shock on May 15, 2021, up to the day of the shock to see which nodes are no longer receiving any block rewards. If we conservatively assume that the initially high number of 18 nodes in Figure D.1A simply did not yet have a chance to receive rewards, we can see that there are around 7 nodes that no longer receive rewards at the steady state. Those nodes represent, however, only 0.0147 percent of the estimated hash power of the network: a far cry from the 51.1% decrease in total hashrate we observed in Section 3.3.

Given that individual miners do not seem to be much affected by our initial estimation, we estimated whether mining pools may have experienced changes in their mining throughout the ban. In Figure D.2A, we display the daily number of blocks mined per node address around the time of the ban with lowess trend lines. Though we cannot make any causal claims, we can clearly observe that some nodes are producing more and some fewer nodes around the time of the ban. We also show summary statistics for 30 days prior and 30 days after the ban in Figure D.2B. These varying trends in block production around the time of the ban suggest that mining pools may have experienced differential impacts, underscoring

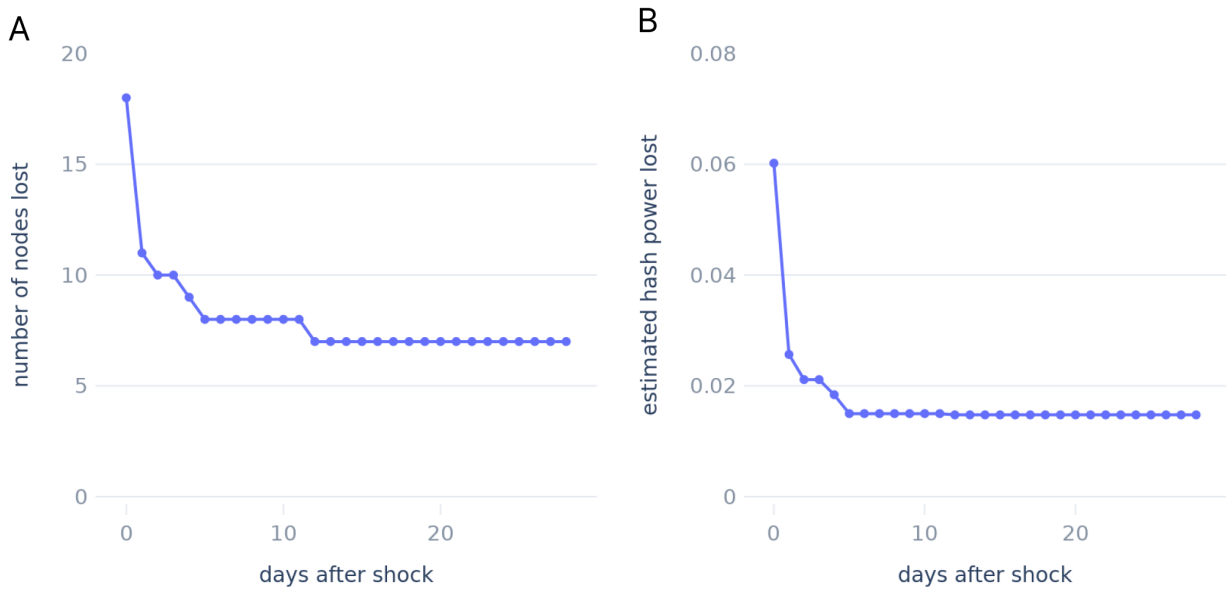

Figure D.1: (A) The number of nodes lost after the initial shock. (B) Estimated hash power of the nodes lost after the initial shock.

the impact of the ban on mining pools versus individual miners.

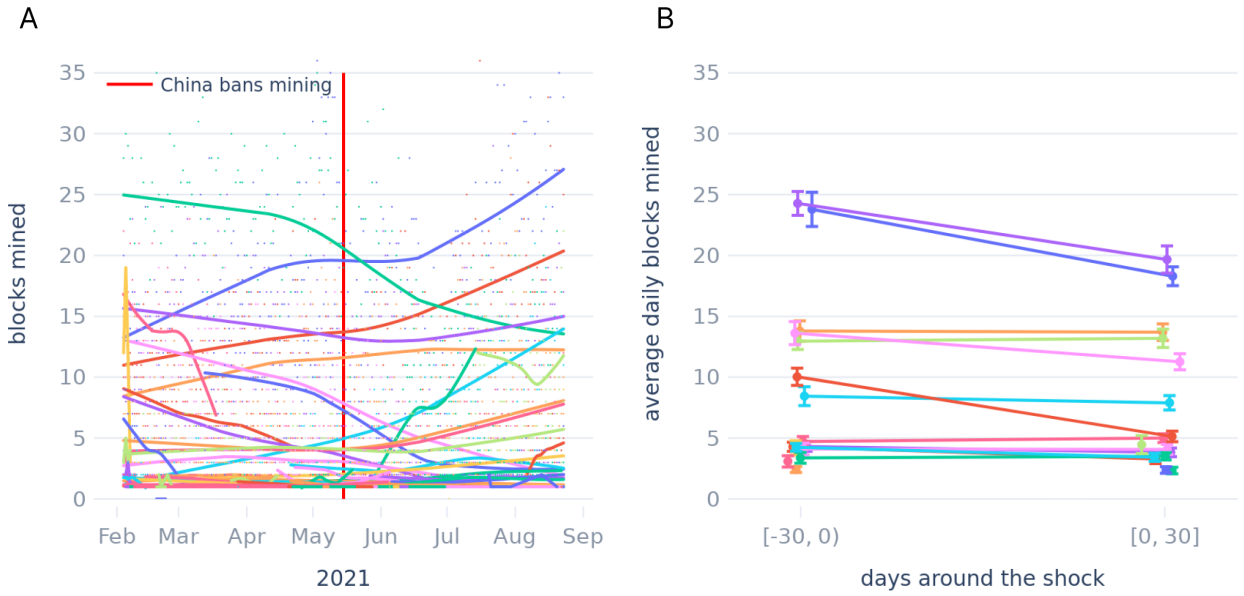

Figure D.2: Number of blocks mined by Bitcoin miners around the time of the shock. (A) Daily number of blocks mined with Lowess trend lines. (B) The average number of blocks mined 30 days before and after the shock with standard errors.

While many individual nodes cannot be identified due to Bitcoin's pseudonymity, some larger nodes, such as mining pools, are labeled, allowing us to more accurately assess the

impact of the ban on these specific entities within the blockchain network. We examined three examples of mining pools that have experienced a reduction in the daily nodes mined from our above analysis: the Binance pool, the Huobi pool, and an unknown pool.[34] While we cannot definitively identify the location of these pools, and even more so for their individual participants, we do know that the cryptocurrency exchanges Binance and Huobi were both initially based in China but moved their headquarters abroad later on.

Figure D.3 displays the balance of BTC in the pools, obtained from blockchain.com, as a proxy for mining activity, since mining results in an increased balance of BTC. We observed a clear, lasting reduction in the BTC balance in May 2021 for the Binance pool. For the Huobi Pool, we saw a slight reduction in BTC balance in May 2021 and a near-complete reduction from June to the end of July, which directly matches our analysis of total hashrates and shock exposure in Section 3.3. For the Unkown Pool, we see a similar but less pronounced decrease in the BTC balance in May 2021 and then a larger decrease in July 2021. In contrast to these mining pools, some mining pools, such as those indicated by the prominent, rising blue and red lines in Figure D.2A, see even greater mining activity after the ban. The observed variations in the trends of mining pools not only align with our analysis of the impact of virtualization on decentralization in response to a major policy shock, but they also reveal the complex dynamics of blockchain decentralization in response to policy shocks.

As an interesting side note, the mining pools that experienced greater activity after the ban are labeled as those of Antpool and ViaBTC. Both companies were, and still are, headquartered in China. In contrast, Binance and Huobi, both of which experienced marked decreases in mining (Figure D.3), are no longer based in China, which suggests a consolidation of Bitcoin mining among Chinese miners. Whether or not such consolidation of Bitcoin mining was deliberate or happenstance, it may further support our virtualization hypothesis in either case. A deliberate consolidation would require high frictions, such as the costs and logistical challenges of relocating specialized ASIC hardware and significant electricity consumption, to

[34]The charts for these exact nodes can be found for the Binance, Huobi, and the unlabeled pools. Accessed October 30, 2023.

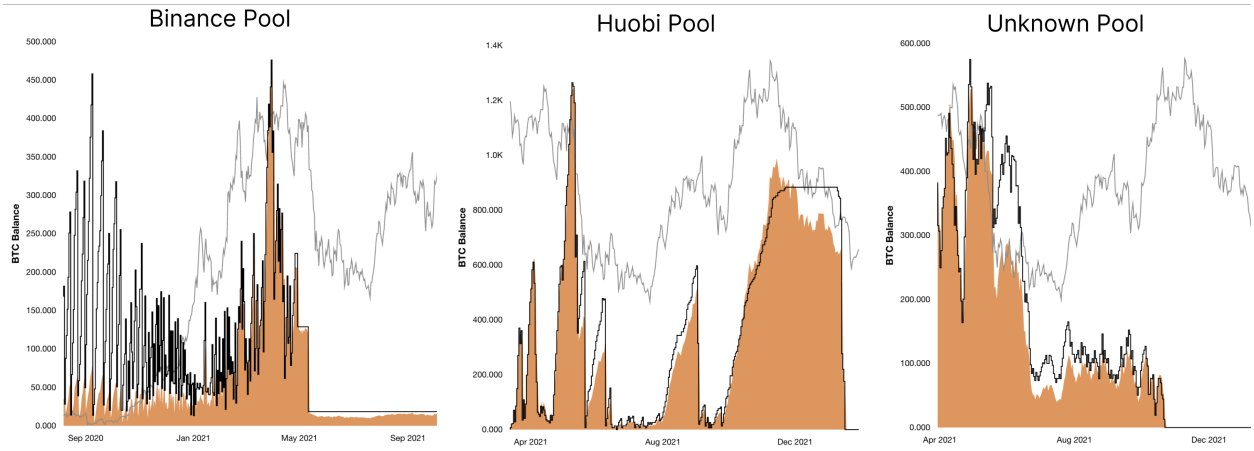

Figure D.3: The BTC balance of major Bitcoin mining pools during China's mining ban.

prevent miners from migrating out of the jurisdiction. On the other hand, if the consolidation was happenstance, it still underscores the rigidity and lack of adaptability in less virtualized systems like Bitcoin. Either way, the observed consolidation among Chinese miners lends further credence to our virtualization hypothesis, highlighting the role of virtualization in shaping the decentralization dynamics of blockchain networks.

E Lagged Effects of China's mining ban

Table E.1 shows the coefficients of the lagged difference-in-difference estimation for China's mining ban according to Equation 3.

Dependent variable	*Entropy*	*Nodes*	*Gini*	*Nakamoto*	*HHI*
Treated$_{\tau=-180}$	0.004	0.685*	0.019	0.060	0.001
	(0.037)	(0.275)	(0.011)	(0.168)	(0.003)
Treated$_{\tau=-120}$	0.072*	1.050***	0.005	0.033	-0.005*
	(0.029)	(0.319)	(0.004)	(0.121)	(0.002)
Treated$_{\tau=-60}$	-0.050***	0.389	0.029*	-0.314**	0.006***
	(0.013)	(0.500)	(0.011)	(0.119)	(0.002)
Treated$_{\tau=0}$	-0.022	6.317***	0.015	0.142	-0.002
	(0.040)	(1.555)	(0.012)	(0.115)	(0.004)
Treated$_{\tau=60}$	-0.065	-1.314	-0.008	-0.035	0.003
	(0.046)	(0.683)	(0.007)	(0.093)	(0.003)
Treated$_{\tau=120}$	-0.114*	-0.817	0.019***	-0.500***	0.010***
	(0.048)	(0.671)	(0.003)	(0.092)	(0.003)
Treated$_{\tau=180}$	-0.211***	-2.869***	-0.021*	-0.079	0.010**
	(0.057)	(0.765)	(0.009)	(0.087)	(0.003)
Bitcoin	-0.001	-34.825***	-0.344***	1.584***	-0.037***
	(0.032)	(1.292)	(0.012)	(0.106)	(0.003)
Post	0.020	-6.893***	-0.033***	-0.060	0.001
	(0.027)	(1.443)	(0.008)	(0.084)	(0.004)
Intercept	3.722***	53.740***	0.801***	2.990***	0.131***
	(0.020)	(1.285)	(0.007)	(0.010)	(0.002)
Observations	2002				

Notes: *p<0.05; **p<0.01; ***p<0.001. Standard errors are clustered by blockchain and month.

Table E.1: Lagged effects of China's ban on crypto mining.

F Lagged Effects of Hetzner's Shutdown

Table F.1 shows the coefficients of the lagged difference-in-difference estimation for Hetzner's shutdown of Solana nodes according to Equation 3.

Dependent variable	*Entropy*	*Nodes*	*Gini*	*Nakamoto*	*HHI*
Treated$_{\tau=-30}$	0.009***	53.400***	0.002***	0.300***	0.000***
	(0.001)	(3.507)	(0.000)	(0.015)	(0.000)
Treated$_{\tau=-20}$	0.015***	31.300***	-0.002***	0.400***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Treated$_{\tau=-10}$	0.011***	25.300***	0.000***	0.400***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Treated$_{\tau=0}$	-0.359	-337.217**	0.013	-6.593***	-0.005
	(0.217)	(104.606)	(0.218)	(1.945)	(0.019)
Treated$_{\tau=10}$	0.194***	75.900***	-0.037***	-0.300***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Treated$_{\tau=20}$	0.148***	5.600***	-0.023***	6.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Treated$_{\tau=30}$	0.037***	11.527***	0.001***	2.791***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	162	162	162	162	162

Notes: *p<0.05; **p<0.01; ***p<0.001. Standard errors are clustered by blockchain and day.

Table F.1: Lagged effects of Hetzner's shutdown of Solana nodes.

G Price Correlation between BTC and ETH

Here, we measured the price correlation between BTC, ETH, and other financial assets to understand the level of price correlation between BTC and ETH. Using price data from Yahoo Finance, we found that BTC and ETH had a relatively high Pearson's correlation coefficient of 0.93 (Figure G.1).[35] This is in comparison with correlations of 0.89 and 0.86 with the S&P 500 for BTC and ETH, respectively. Correlations are shown for other prominent financial assets in Figure G.1 for comparison. TLT is the iShares 20+ Year Treasury Bond ETF, and VNQ is the Vanguard Real Estate Index Fund ETF.

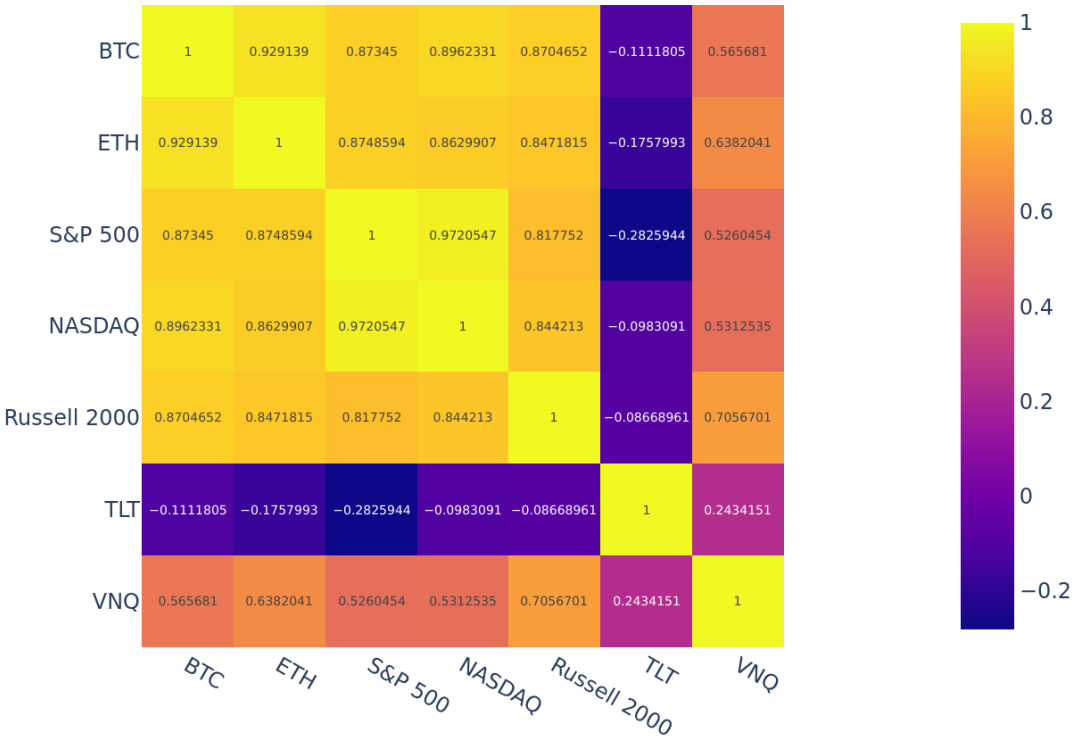

Figure G.1: Pearson's correlation coefficients between BTC, ETH, the S&P 500 Index, NASDAQ Composite, Russell 2000, TLT (the iShares 20+ Year Treasury Bond ETF), and VNQ (the Vanguard Real Estate Index Fund ETF).

[35] Data obtained from Yahoo Finance for BTC, ETH, S&P Index, NASDAQ Composite, Russell 2000, TLT, and VNQ. Accessed April 5, 2024.

H Price Analysis during Shocks

Here, we tested the hypothesis that market sentiment influences a blockchain's recovery from shocks. Below are estimations of Equation 5 on the daily price of tokens native to the blockchains studied here. Using price data around key events for BTC, SOL, and ETH, we found that these shocks had no positive impact on prices; some even dropped significantly.

Dependent variable		*Entropy*	
	BTC	SOL	ETH
Intercept	49353.00***	30.415***	1686.777***
	(1046.509)	(0.921)	(34.231)
Post	-1288.26	-7.246***	-350.015***
	(1456.120)	(1.282)	(47.629)
Days	126.316***	-0.054*	1.209
	(29.837)	(0.026)	(0.976)
(Post × Days)	191.234***	-0.201***	0.238
	(41.683)	(0.037)	(1.363)
Observations	121	121	121

Notes: *p<0.05; **p<0.01; ***p<0.001. Standard errors are clustered by blockchain and month.

Table H.1: Event studies on the token price of BTC during China's ban, SOL during Hetzner's shutdown, and ETH during the Merge.

Appendix References

Cambridge Centre for Alternative Finance. 2022. Distribution of bitcoin mining hashrate from september 2019 to january 2022, by country [graph]. In Statista. URL https://www.statista.com/statistics/1200477/bitcoin-mining-by-country/. Retrieved October 30, 2023.

The Cambridge Centre for Alternative Finance. 2023. Bitcoin mining map. URL https://ccaf.io/cbnsi/cbeci/mining_map. Accessed: October 30, 2023.