THE
QUARTERLY JOURNAL
OF ECONOMICS

Vol. 140 2025 Issue 1

TRUST AT SCALE: THE ECONOMIC LIMITS OF CRYPTOCURRENCIES AND BLOCKCHAINS *

ERIC BUDISH

Satoshi Nakamoto (2008) invented a new kind of economic system that does not need the support of government or rule of law. Trust and security instead arise from a combination of cryptography and economic incentives, all in a completely anonymous and decentralized system. This article shows that Nakamoto's novel form of trust, while undeniably ingenious, is deeply economically limited. The core argument is three equations. A zero-profit condition on the quantity of honest blockchain "trust support" (work, stake, etc.) and an incentive-compatibility condition on the system's security against majority attack (the Achilles heel of all forms of permissionless consensus) together imply an equilibrium constraint, which says that the "flow" cost of blockchain trust has to be large at all times relative to the benefits of attacking the system. This is extremely expensive

* This paper originally circulated in June 2018 in a shorter form as Budish (2018). I am grateful to the editor Andrei Shleifer, the coeditor Stefanie Stantcheva, and six anonymous referees for their valuable advice. Thanks are also due to Susan Athey, Vitalik Buterin, Glenn Ellison, Gene Fama, Alex Frankel, Joshua Gans, Matt Gentzkow, Edward Glaeser, Austan Goolsbee, Hanna Halaburda, Zhiguo He, Joi Ito, Steve Kaplan, Anil Kashyap, Judd Kessler, Scott Kominers, Randy Kroszner, Robin Lee, Jacob Leshno, Andrew Lewis-Pye, Shengwu Li, Jens Ludwig, Neale Mahoney, Gregor Matvos, Paul Milgrom, Sendhil Mullainathan, Vipin Narang, Neha Narula, Ariel Pakes, David Parkes, Al Roth, Tim Roughgarden, John Shim, Scott Stornetta, Adi Sunderam, Chad Syverson, Alex Tabarrok, Nusret Tas, David Tse, Rakesh Vohra, and numerous seminar audiences. Ethan Che, Natalia Drozdoff, Matthew O'Keefe, Anand Shah, Peyman Shahidi, Jia Wan, and Tianyi Zhang provided excellent research assistance.

The Quarterly Journal of Economics (2025), 1–62. https://doi.org/10.1093/qje/qjae033.
Advance Access publication on October 16, 2024.

relative to traditional forms of trust and scales linearly with the value of attack. In scenarios that represent Nakamoto trust becoming a more significant part of the global financial system, the cost of trust would exceed global GDP. Nakamoto trust would become more attractive if an attacker lost the stock value of their capital in addition to paying the flow cost of attack, but this requires either collapse of the system (hardly reassuring) or external support from rule of law. The key difference between Nakamoto trust and traditional trust grounded in rule of law and complementary sources, such as reputations, relationships, and collateral, is economies of scale: society or a firm pays a fixed cost to enjoy trust over a large quantity of economic activity at low or zero marginal cost. *JEL codes:* A1, A12, B00, B21, C7, C73, D47, E42, E5, F3, G1, G2, K24, K42, L14, O3, P16.

I. Introduction

Economists have long widely agreed that the market system requires some form of government and rule of law for support. This is uncontroversial among even the most free-market-oriented thinkers. Adam Smith (1776) mostly argues for reducing government interference in markets, but he does not go all the way to zero, writing that "commerce and manufactures can seldom flourish long in any state" without a legal system, property rights and contract enforcement, as well as certain public goods. Hayek (1960) grapples at length with the paradox that to maximize freedom—which he defines as the absence of coercion—it is necessary to have a government with the power to coerce. Friedman (1962) famously described the government's role establishing the "rules of the game" for the market system and acting as its "umpire." There is significant debate in modern economics about what government should do beyond these basic supports for the market system (e.g., social insurance, correcting externalities), but that there is some role for government and rule of law is essentially taken for granted.

Satoshi Nakamoto (2008) invented a new kind of economic system that does not need the support of government or laws.[1] Trust and security instead arise from a combination of cryptog-

1. The antigovernment views of Nakamoto and other early Bitcoin enthusiasts have been widely documented. See Popper (2015) for one early account and satoshi.nakamotoinstitute.org as a primary source. A few example quotes from these sources give a sense: "It's completely decentralized with no server or central authority" (Nakamoto personal communication, January 8, 2009); "a new territory

raphy and economic incentives, all in a completely anonymous and decentralized system. Computer scientists call the innovation "permissionless consensus": a large, anonymous, freely entering and exiting set of participants around the world is incentivized to collectively maintain a common data set, enabling trust in the data set without the need for rule of law or any specific trusted party. This invention enabled cryptocurrencies, including Nakamoto's own creation, Bitcoin. The data structure maintained by the large set of participants is called a blockchain.[2]

It is no understatement to say that Nakamoto's invention captured the world's attention. One oft-cited figure is the $3 trillion of market capitalization of Bitcoin and other crypto assets at their peak, but even this figure seems to understate the amount of cultural, political, and commercial attention that has been paid to blockchains and cryptocurrencies. At the same time, the economic usefulness of Nakamoto's invention remains an open question. To date, the majority of cryptocurrency volume appears to be speculative, with the other most widely documented use case being black-market transactions (Makarov and Schoar 2021; Gensler 2021; Buterin 2022).[3] Moreover, the majority of this speculative volume has been through cryptocurrency exchanges—which are

of freedom" (Nakamoto personal communication, November 6, 2008); "outside the reach of any government" (Popper 2015, 48).

2. Not widely appreciated is that the blockchain data structure, without the novel method of trust, significantly predates Nakamoto, at least in terms of the core scientific ingredients if not popular and commercial appreciation of its potential usefulness (Haber and Stornetta 1991; Bayer, Haber, and Stornetta 1993). This data structure is sometimes called a permissioned blockchain and is in essence a well-architected database that is append-only, has clear rules about what parties can add what data, and uses cryptography to prove that past data have not been deleted or tampered with. See Section II.E and Budish and Sunderam (2023) for further discussion.

3. Makarov and Schoar (2021) find that about 75% of Bitcoin transaction volume since 2015 involves cryptocurrency exchanges or exchange-like entitites, once the data are cleaned to account for spurious volume (such as a user moving their own funds from one address to another). They conclude that "the vast majority of Bitcoin transactions between real entities are for trading and speculative purposes." In a data set from an earlier time period and using a different data cleaning and classification methodology, Foley, Karlsen, and Putniņš (2019) find that 46% of Bitcoin transactions that do not involve cryptocurrency exchanges relate to illegal activity. Many prominent public officials, such as U.S. Treasury Secretary Janet Yellen and SEC Chair Gary Gensler, have also described cryptocurrency activity to date as mostly speculative or black market (Cox 2021; Gensler 2021). Ethereum founder Vitalik Buterin wrote in a December 2022 essay that he

(at least in principle) centralized, trusted financial intermediaries.

This article studies the economics of Nakamoto's novel form of trust. Is it economically viable as an alternative to the traditional market system supported by rule of law? What are the fundamental economics of the fundamental computer science innovation in Nakamoto (2008)?

I find that—at least in its pure form, without any implicit protection from rule of law—Nakamoto's novel form of trust faces serious economic limits. It is unusually expensive in absolute terms relative to the stakes involved, and its expense scales linearly with the stakes involved. These results have an if-then implication: if permissionless consensus in its pure form were to become a more important part of the global economic and financial system than it has been to date, then the costs of securing the trust would become preposterous—more than global GDP in some scenarios. The analysis may also sharpen our conceptual understanding of what is special about traditional forms of trust that are grounded in rule of law and other complementary sources, such as reputations, relationships, and collateral. The key distinction will prove to be economies of scale in the production of trust.

The core of the argument is three simple equations. The first equation is a zero-profit condition that describes the quantity of "trust support" devoted to maintaining permissionless consensus as a function of the compensation paid by the protocol for trust support, assuming that all participants behave honestly. Trust support can take the form of providing computational work as in Bitcoin ("proof of work"), providing other kinds of computational resources as in proof-of-storage or proof-of-memory protocols, or posting cryptocurrency coins in a proof-of-stake protocol, such as Ethereum.[4] For a sense of magnitudes, in recent years the com-

is most excited about applications still to come in the future, not the ones that already exist, which he describes as "hyperfinancialized" (Buterin 2022).

4. See Lewis-Pye and Roughgarden (2024) for an overview of the computer science literature on permissionless consensus, including a description of all of the key protocols and forms of trust support. Throughout the article, when I use the phrase "Nakamoto trust" or "blockchain-based trust" I am referring to Nakamoto's overall idea of blockchain-based permissionless consensus, in its pure form without implicit support from the rule of law. I try to be clear when I am referring more specifically to Nakamoto (2008)'s specific implementation of permissionless consensus using longest-chain proof of work. In particular, the quantification anal-

pensation to Bitcoin trust support (known as miners) has averaged about $250,000 per block of data, or about $36 million a day, and Bitcoin miners performed an average of about 200 million trillion calculations per second as an equilibrium response to this compensation. The computer science details behind this process are complicated (see Section II), and vary to some extent by blockchain protocol, but the economics is standard free-entry logic. Variations of this first equation have appeared in numerous other prior papers.

The second equation is an incentive-compatibility condition: how much security does a given level of trust support produce? The Achilles heel of permissionless consensus is that it is vulnerable to "majority attack": Nakamoto (2008) and subsequent methods for creating an anonymous, decentralized consensus about the state of a data set rely on a majority of the resources devoted to maintaining the data set to behave honestly. This is not an obscure point; it is in the abstract of the famous Nakamoto (2008) paper and has been understood to be an issue more generally with distributed consensus systems since famous computer science research in the 1980s. Intuitively, permissionless consensus, whether in the form invented by Nakamoto or other variations such as proof of stake, always relies on some implicit version of majority or supermajority voting to adjudicate what the state is in case of a dispute. The second equation captures that it must not be economically profitable for a potential attacker to provide a majority of the total trust support and manipulate the state.

It is worth briefly contrasting the approach to security taken in my equation (2) with the approach taken in the prior computer science literature. Famous early research on distributed consensus, such as Pease, Shostak, and Lamport (1980), Lamport, Shostak, and Marshall (1982), and Dwork, Lynch, and Stockmeyer (1988), stated results of the following form: as long as less than proportion ρ of the servers maintaining consensus are faulty (Byzantine), then the consensus protocol has good properties of safety (transactions, once confirmed, will not be reverted) and liveness (transactions are confirmed within a reasonable amount of time). Nakamoto (2008) gives results of the same

ysis in Section IV is specific to Nakamoto's longest-chain proof-of-work protocol, and the key economic security difference between proof-of-work and proof-of-stake protocols is analyzed in detail in Sections V.A and V.B and discussed later in the introduction.

form—indeed, with an improvement from $\rho = \frac{1}{3}$ to $\rho = \frac{1}{2}$. However, the early computer science literature was studying what are now called permissioned consensus systems, for example, a company keeping its database servers in sync with each other. Treating security as a 0–1 property that holds if and only if an attacker is less than some bound ρ seems reasonable for permissioned consensus—for example, it would tell a company to use enough redundant hardware and enough overlapping security measures that the risk of proportion ρ of its servers failing at the same time is sufficiently small. In contrast, Nakamoto (2008) and subsequent blockchain researchers are studying *permissionless* consensus protocols, in which anybody can freely enter and exit the system at any time, anonymously, without protection from rule of law. As soon as the system is permissionless, one has to ask what is an attacker's incentive to acquire proportion ρ of the total system resources. This rethink of the economic security of a permissionless consensus protocol that I introduce applies to proof of work, proof of stake, or any other approach to permissionless consensus. They are all vulnerable to some form of majority attack, so they all need to be studied with both a free-entry condition like my equation (1) and an incentive-compatibility condition like my equation (2). In effect, this article argues that Nakamoto (2008) made an error conceptualizing security in the same way as the 1980s consensus literature and not as an economic incentive-compatibility constraint.[5]

My third equation is an equilibrium constraint that connects equations (1) and (2), that is, connects the zero-profit condition to the incentive-compatibility condition. The reason these equations can be linked is that the amount of honest trust support appears in both. In equation (1), the amount of honest trust support reflects the recurring payments to this trust support. In equation (2), the amount of honest trust support determines the cost of majority attack. Equation (3) tells us that the recurring payments to the honest trust support in the zero-profit equilibrium must be large relative to the value of attacking the system.

5. Some permissionless consensus protocols are also vulnerable to forms of dishonest play by small participants, typically to obtain a disproportionate share of block rewards as opposed to stealing large sums of money. See Eyal and Sirer (2014), Carlsten et al. (2016), Biais et al. (2019), and Saleh (2021) for well-known studies of this issue and of the conditions under which honest play is a Nash equilibrium for small participants.

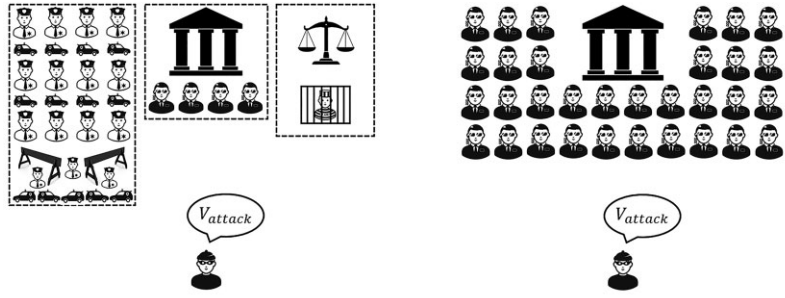

FIGURE I

Comparison of Trust Models: Nakamoto versus Traditional

This is a very expensive form of trust! The recurring payments to trust support are a "flow," so equation (3) tells us that the flow costs of maintaining the trust must exceed the value of attack at all times.[6] Moreover, this required flow cost scales linearly with the value of attack. An intuition is that Nakamoto trust is memoryless—it is as secure at a moment in time as the amount of trust support at that moment. Under idealized attack circumstances (assuming away several reasonable frictions), I obtain an even stronger result, which is that the net cost of attack is zero—because the attacker also earns the trust-support compensation that would have gone to the honest participants.

The essential difference between the Nakamoto trust model and the traditional model grounded in rule of law is shown in Figure I. A criminal is thinking of robbing a bank. In the traditional model, the criminal must consider how many security guards he will need to overcome. Then he will have to take into account that the bank will call in reinforcements from police, and

6. A 2016 blog post by Vitalik Buterin contains an informal statement of this insight: "The size of the mining network has to be so large that attacks are inconceivable. Attackers of size less than X are discouraged from appearing by having the network constantly spend X every single day. I reject this logic because (i) it kills trees, and (ii) it fails to realize the cypherpunk spirit." See Section III.C for the full quotation and how it can be translated into this paper's model.

that, if caught, he will go to jail. Similarly, consider a country thinking of whether to invade another country. They will have to consider the soldiers at the border (analog of security guards) but also that the invaded country will call in military reinforcements (analog of police) and that the invaded country may launch a counterattack (analog of Beckerian deterrence from courts).[7] In contrast, the Nakamoto model is just to have a very large number of security guards at the bank or soldiers at the border. This works, but it is very expensive and scales terribly with the stakes.

Notice two sources of scale economies in the traditional model. First, the police do not have to be present at the particular bank to provide security—they can provide security to many locations at once as long as they are not all attacked simultaneously. Second, the courts can deter crime with just the credible threat to prosecute and imprison—a fixed-cost investment in court capacity can deter a large quantity of potential criminal activity. This is the essence of Becker's (1968) model of optimal deterrence, and is central to Hayek's (1960) resolution of the paradox noted above that freedom requires a government with the power to coerce.[8]

Notice as well a subtle additional source of inefficiency in the Nakamoto trust model—the full scale of the trust support is present for all transactions, whether for large sums or small. This is like having the same number of security guards outside the local bank branch as outside Fort Knox or the Federal Reserve.

7. I thank Edward Glaeser for drawing this connection to military strategy and Vipin Narang for a helpful discussion about the topic.

8. Hayek's resolution is that a government's credible threat of coercion, in response to violations of clear, predictable, and symmetrically enforced laws, is both (i) not a violation of freedom, and (ii) sufficient to secure freedom. "The *threat* of coercion has a very different effect from that of *actual and unavoidable* coercion ... The great majority of the threats to coercion that a free society must employ are of this avoidable kind ... The sanctions of the law are designed only to prevent a person from doing certain things or to make him perform obligations that he has voluntarily incurred.... Provided that I know beforehand ... I need never be coerced." (Hayek 1960, 209–210, emphasis added). "It is the cases that never come before the courts, not those that do, that are the measure of the certainty of the rule of law" (Hayek 1960, 316). "There is little difference between the knowledge that if he builds a bonfire on the floor of his living room his house will burn down, and the knowledge that if he sets his neighbor's house on fire he will find himself in jail. Like the laws of nature the laws of the state provide fixed features in the environment in which he has to move" (Hayek 1960, 221).

There are many other sources of trust that are familiar to economists besides rule of law, such as reputations, brands, relationships, collateral, and cultural or organizational norms.[9] The contrast versus Nakamoto trust is similar: in each case the trust is more secure than the flow level of investment (e.g., a brand is more trustworthy than its advertising expenditure in the past 24 hours). It bears emphasizing that these sources of trust often work in conjunction with rule of law, sometimes implicitly. For example, a customer trusts Starbucks to provide good coffee because of its brand but also because it is illegal for a different entity to impersonate Starbucks' name and imagery. In a Levin (2003) relational contract, the employee trusts that if they put in high effort they will get paid a performance bonus, but in the background, it is also the case that the employee knows the employer will pay at least the promised minimum because of rule of law, and the employer trusts the employee not to rob the company because of rule of law.[10]

A blockchain would be significantly more economically secure than described so far if both (i) the capital used to maintain the permissionless consensus is specialized, and (ii) an attack causes the attacker's specialized capital to lose its value or be confiscated. The first condition is satisfied for many major blockchains (e.g., Bitcoin ASICs or stake for proof-of-stake protocols). If the second condition obtains as well, the attacker's cost is not just the flow cost of conducting the attack but the stock value of their specialized capital—analogously to how a firm that cheats its customers loses the stock value of its reputation or brand. A modified version of the incentive-compatibility condition equation (2) and some simple calculations suggest that the potential eco-

9. Foundational work on trust from rule of law includes Schelling (1960), Becker (1968), Hart (1995), Shleifer and Vishny (1997), and La Porta et al. (1998). Important work on other sources of trust includes Nelson (1974), Kreps et al. (1982), Fudenberg, Levine, and Maskin (1994), Tadelis (1999) on brands and reputations; Baker, Gibbons, and Murphy (2002), Levin (2003) on relationships; Kandori (1992), Holmstrom and Milgrom (1994), La Porta et al. (1997), Guiso, Sapienza, and Zingales (2006) on norms.

10. Formally, in Levin (2003)'s model, the employer pays the employee at least the fixed salary w_t no matter what, and the employee's lowest action, denoted $e_t = 0$, harms the firm only through poor effort, not theft. If the firm could pay the worker nothing or the worker could rob the firm, the scope for cooperation in the relational contract would be far worse (formally, each party's "renege" option would be much more attractive, undermining the relational contract's ability to be self-enforcing).

nomic security improvement is three to four orders of magnitude. This difference is large enough to plausibly explain why major blockchains such as Bitcoin and Ethereum have not been majority attacked to date.

However, how exactly does the attacker lose their capital? One possibility is that the value of the cryptocurrency collapses because of the attack, and as a result the value of specialized capital collapses, too. But vulnerability to collapse is itself a serious problem—hardly a reassuring foundation for a novel economic system—and raises the possibility of an attack motivated by this collapse per se. A second possibility is legal punishment, which certainly would work but concedes that Nakamoto trust fails without support from rule of law. A third possibility is that the protocol could algorithmically confiscate or block the attacker's capital in a targeted punishment, while leaving honest participants' capital untouched. In a companion computer science paper (Budish, Lewis-Pye, and Roughgarden 2024) we show that this approach (i) is impossible for a broad class of permissionless consensus protocols that includes Bitcoin's; (ii) is possible for a class of permissionless consensus protocols that includes proof-of-stake Ethereum's, but only under a strong assumption about network reliability and only if the attacker is less than proportion $\rho = \frac{2}{3}$ of the total. Hence, the current article's framework is a lens through which we can understand the economic goals of Ethereum's recent adoption of proof of stake and also its economic limits—for attackers of size less than $\rho = \frac{2}{3}$, an attack costs a stock not a flow, but if an attacker is incentivized to exceed $\rho = \frac{2}{3}$ then some external source of trust is required for security, such as rule of law.

The remainder of this article is organized as follows. Section II provides a description of Nakamoto (2008) and related concepts. Section III presents the heart of the economic critique of Nakamoto's novel form of trust, equations (1)–(3). Section IV uses the model to quantify the cost of keeping Nakamoto (2008)'s specific blockchain design secure against double-spending attacks. Section V analyzes the case of specialized capital that can lose its value because of the attack. Section VI contrasts Nakamoto trust with traditional trust grounded in the rule of law. Section VII concludes. Online Appendix A discusses responses to this article's argument since it first circulated in 2018. Online Appendix B provides technical results in support of the double-spending attack analysis. Online Appendix C compiles lists of cryptocurrency ma-

jority attacks, thefts of cryptocurrency financial institutions, and collapses of cryptocurrency financial instutions.

II. Overview of Nakamoto (2008) and Related Concepts

Sections II.A–II.D provide an overview of Nakamoto (2008). Some of the detail is specific to Bitcoin and Nakamoto's specific form of permissionless consensus, longest-chain proof of work, while also trying to describe the key ideas at a more conceptual level. Section II.E discusses three related concepts: permissioned blockchains, smart contracts, and proof-of-stake consensus. The overall goal of this section is to provide an overview of the relevant computer science concepts that is self-contained and at a sufficient level of detail to justify the economics analysis in the rest of the article.[11] Readers already familiar with this material may skip this section without much loss.

II.A. Transactions

The first step in describing Nakamoto (2008) is to describe transactions and the limitations of other methods of keeping track of transactions.

1. *Elements of a Transaction.* The key elements of a cryptocurrency transaction are the sender of funds, the receiver of funds, the transaction amount, and a cryptographic signature. The sender and receiver are represented as alphanumeric strings called addresses; addresses are somewhat analogous to account numbers. The cryptographic signature uses standard ideas from public-key cryptography to prove that the transaction was initiated by the sender; that is, the signature could only be created by someone who knows the sender's private key for that address. The cryptographic signature also encodes the other transaction details, including the receiver and the transaction amount; it is like not only signing a check but also signing the seal of the envelope that contains the check, so the recipient and amount cannot be subsequently altered.

11. Readers interested in additional computer science detail should consult sources such as the original Nakamoto (2008) paper, Lewis-Pye and Roughgarden (2024)'s survey of the computer science literature on permissionless consensus, and Roughgarden's (2023) online course. There are several surveys aimed specifically at economists including Halaburda et al. (2022) and Böhme et al. (2015).

2. *Limitations of a Shared Public Spreadsheet of Transactions.* Imagine keeping track of such transactions on a shared public spreadsheet, such as a Google Doc. The cryptographic signature provides a certain level of trust in the data, in that only Alice, or someone in possession of Alice's private key, can add correctly signed transactions in which Alice is the sender of funds. However, there are three vulnerabilities:

(i) Alice could add a transaction in which she sends money she does not have.
(ii) Alice could add multiple transactions at the same or similar time, in which she sends money she does have but to multiple parties at the same time.
(iii) Alice could delete previous transactions from the shared public spreadsheet; either her own or others'.

Thus, while a shared public spreadsheet of transactions could be used among parties that trust each other—for example, a modern version of the babysitting co-op parable in Krugman (1998)—this system is not suitable for tracking transactions among parties that do not have such a level of trust.

3. *Limitations of a Trusted Party.* Imagine keeping track of transactions through a widely trusted party that keeps track of balances, such as a central bank (e.g., a Central Bank Digital Currency or CBDC). This approach addresses the three vulnerabilities described above with respect to the shared public spreadsheet: the trusted party can ensure that only valid transactions are added to the ledger and that previous transactions are not deleted. However, the limitation is that it requires such a trusted party. The central goal of Nakamoto (2008) is to have a trusted ledger of transactions that does not require any specific trusted party.

II.B. What is the Nakamoto Blockchain?

This section describes the Nakamoto (2008) blockchain (i.e., proof-of-work longest-chain consensus) in four steps.

1. *Pending Transactions List.* Users submit transactions to a pending transactions list, called the mempool. One can think of the mempool as the shared public spreadsheet discussed already. Transactions in the mempool are not considered official yet.

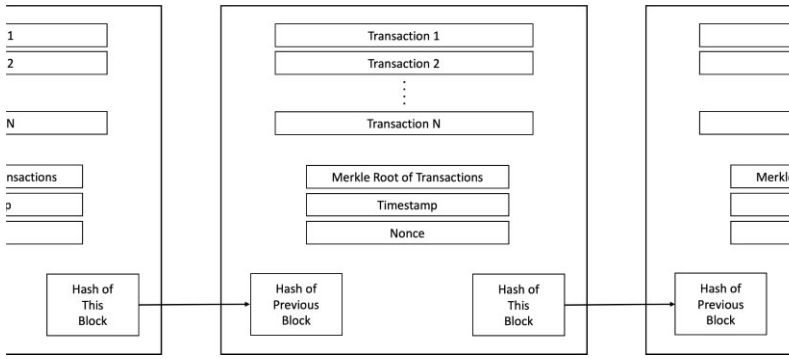

FIGURE II

Illustration of the Blockchain Data Structure

See Section II.A for a description of transactions and Section II.B for a description of the overall blockchain data structure and the other elements in the diagram.

2. *Valid Blocks.* Any computer around the world can compete for the right to add transactions from the mempool to a data structure called the blockchain. The computational competition is described in the next step.

The word "blockchain" references that transactions are added in blocks (for Bitcoin, consisting of about 1,000–2,000 transactions), and each block of transactions "chains" to the previous block by including a hash of the data in the previous block (see Figure II). This use of hashes to chain together a sequence of blocks of data was invented by Haber and Stornetta (1991) and Bayer, Haber, and Stornetta (1993). Since the hash of the current block depends on the data in the previous block, which in turn includes its hash of the block before that, and so on, any change to any element in the history of transactions affects the value of the hash of the current block.

For a block of transactions to be valid, the following three criteria must all be true:

(i) Each individual transaction must be properly signed: the cryptographic signature could only be generated by a user in possession of the sender's private key.
(ii) Each individual transaction must be properly funded: given all transactions in previous blocks in the chain, the

sender must be in possession of the cryptocurrency she or he is sending.

(iii) The transactions in a block must not contradict each other: there cannot be two or more transactions in a block in which a common sender sends the same cryptocurrency to multiple receivers.

3. *Bitcoin Mining Computational Tournament.* In Nakamoto (2008), the competition to add new blocks boils down to a massive, brute-force search for a lucky random alphanumeric string. More precisely, Bitcoin miners—where "miners" is just the terminology for computational power that attempts to add new blocks of transactions to the Bitcoin blockchain—choose a valid block of Bitcoin transactions from the mempool that they want to chain to the previous block of transactions and search for an alphanumeric string (called a nonce) such that when that alphanumeric string, in combination with all of the other data in the block they are trying to add, is all hashed together using the hash function SHA-256, the result has a very large number of leading zeros.

For readers unfamiliar with hash functions, it is highly recommended to go to a website like https://www.movable-type.co.uk/scripts/sha256.html to get a feel for how they work. For example, the hash of the title of this paper is 09b23bf1eb4b7cda..., which has one leading zero. A block added to the Bitcoin blockchain in January 2024, block 824601, has the hash

0000000000000000000237dcb2a4e6ebb21b499cd81a1ec94b49053c8636be34

which has 19 leading zeros. Because each digit in the hash can take on values 0–9 and a–f, and the SHA-256 hash function is pseudo-random, the likelihood of finding an alphanumeric string that produces a hash with 19 leading zeros is 1 out of 16^{19}, which is about 1 out of 75 billion trillion. The number of leading zeros required is calibrated by the Bitcoin system every roughly two weeks, based on the current amount of computational power devoted to Bitcoin mining, to ensure that blocks are successfully mined on average every 10 minutes. (This calibration can be finer than is possible using just zeros, e.g., 19 leading zeros and a 20th digit weakly less than 7.)

When a miner finds a lucky alphanumeric string, they publicly broadcast their block to all of the other Bitcoin miners. Other Bitcoin miners can quickly check whether the block is valid; that

is, whether the set of transactions in the block meet the criteria listed above in step 2, and whether the alphanumeric string indeed produces a valid hash with enough leading zeros. Critically, while finding a lucky alphanumeric string is extremely computationally intensive, checking the validity of a given block is computationally trivial. For this reason, a valid block is "proof of work"—proof that the miner who found the block did a large amount of computational work in expectation.

The lucky miner who broadcast the valid block gets compensated in two ways. First, the miner is compensated with new Bitcoins. This is called the "block reward," which was originally 50 Bitcoins per block, and halves every roughly four years, most recently in April 2024 to 3.125 Bitcoins per block. Second, the miner earns transactions fees associated with the transactions they included in their block. The economics of these transactions fees are considered in depth in Huberman, Leshno, and Moallemi (2021); users who place a high value on getting their transaction added to the blockchain quickly can ensure faster service by offering a larger transaction fee, so there is an auction-theoretic flavor to the fees, as well as queueing and congestion issues.

4. *Longest-Chain Rule.* Once a valid block is broadcast and the other miners have checked its validity, miners are supposed to move on to mining the next block. To induce this behavior, Nakamoto (2008) proposed the longest-chain rule—the rule that if there are multiple chains of blocks, the longest chain, as measured by the amount of computational work, is the official consensus record of transactions.

Intuitively, Nakamoto's longest-chain rule provides a decentralized way to coordinate miners' efforts. If miners focus their attention on the current longest chain, and they find a lucky alphanumeric string and mine a block, then their new block will be part of the new longest chain, and hence new official record, and the miner will earn the block reward. Nakamoto (2008) shows formally that as long as a majority of computational power follows the longest-chain rule, then the longest chain will outpace attackers with probability that converges to one exponentially in the honest-majority's share and the deficit the attacker must overcome.

A related intuition is that the longest-chain rule provides a decentralized way to adjudicate disputes—computational power "votes" on the true state, and the majority rules.

The game-theoretic validity of longest-chain consensus has received considerable scholarly attention. The most general treatment to date is Biais et al. (2019), who show that honest mining on the longest chain is indeed a Nash equilibrium, although there can be other equilibria as well. Carlsten et al. (2016) show that longest-chain consensus is an equilibrium only if the block reward component of miner compensation is large enough. Kroll, Davey, and Felten (2013) provide credible intuition for why longest-chain consensus is a Nash equilibrium, though without a formal game-theoretic model.

However, these analyses explicitly assume that all miners are "small"—that is, they assume away the possibility of majority attack. Majority attack, discussed next, is at the heart of this article's analysis.

II.C. Vulnerability to Majority Attack

Nakamoto's blockchain is vulnerable to attack by an adversary with 51% or more of the computational power. This is because the adversary, whenever they like, can create an alternative chain of blocks that will outpace the honest chain of blocks with probability one, and hence become the new consensus. This vulnerability is widely understood—it is even in the abstract of the Nakamoto (2008) paper (excerpted below). Moreover, that Nakamoto's blockchain is vulnerable to attack is not surprising to computer scientists in the sense that previous approaches to distributed consensus, based on the Byzantine fault tolerance (BFT) paradigm, were also vulnerable to attack by a too-large adversary except under very restrictive assumptions (Pease, Shostak, and Lamport 1980; Lamport, Shostak, and Pease 1982; Dolev and Strong 1983; Fischer, Lynch, and Paterson 1985; Dwork, Lynch, and Stockkmeyer 1988).[12]

12. The exception in which communication among honest parties is secure even in the presence of an unbounded adversary requires that the honest parties have access to a communication network that never experiences delays longer than a known bound (the synchronous communications model) and have access in advance of communication to all honest parties' cryptographic public keys (the public key infrastructure assumption). These assumptions are frequently satisfied in practical applications with preexisting trust (e.g., secure military communications) but are widely viewed to be incompatible with the kind of communication Nakamoto (2008) is trying to facilitate over the internet among parties without pre-existing trust. See Roughgarden (2023) for an accessible treatment of key results on BFT consensus.

The canonical attack Nakamoto (2008) worried about is called a double-spend: the attacker sends Bitcoins in transactions on the original honest chain, and then deletes those transactions from the consensus record with their alternative chain, allowing them to spend the same currency twice. Section IV will describe double-spending attacks in detail and analyze their economic implications.

Eyal and Sirer (2014) show that Bitcoin is also vulnerable to a form of minority attack, in which a large-enough miner can sometimes profit, in expectation, from holding back a solved block so that they can work on extending it in private, while other miners therefore focus their attention on what is probabilistically not the longest chain. However, the purpose of the Eyal and Sirer (2014) minority attack is more circumscribed in that its goal is to obtain a disproportionate share of mining rewards, not manipulate the blockchain to double spend. A somewhat analogous issue arises with longest-chain proof-of-stake consensus (Saleh 2021).

II.D. Nakamoto (2008): Summary

The abstract of Nakamoto (2008) succinctly summarizes the accomplishment and its vulnerability:

> A purely peer-to-peer version of electronic cash would allow online payments to be sent directly from one party to another without going through a financial institution. Digital signatures provide part of the solution, but the main benefits are lost if a trusted third party is still required to prevent double-spending. We propose a solution to the double-spending problem using a peer-to-peer network. The network timestamps transactions by hashing them into an ongoing chain of hash-based proof-of-work, forming a record that cannot be changed without redoing the proof-of-work. The longest chain serves not only as proof of the sequence of events witnessed, but proof that it came from the largest pool of CPU power. *As long as a majority of CPU power is controlled by nodes that are not cooperating to attack the network, they'll generate the longest chain and outpace attackers.* The network itself requires minimal structure. Messages are broadcast on a best effort basis and nodes can leave and rejoin the network at will, accepting the longest proof-of-work chain as proof of what happened while they were gone. (p.1, emphasis added)

The accomplishment is a "purely peer-to-peer version of electronic cash" without the use of a "trusted third party." Trust in the integrity of the data emerges from the hash-based proof-of-work,

conducted by an unstructured network with free entry and exit. The longest chain is the official record of "what happened"—that is, the (permissionless) consensus.

The vulnerability is majority attack—the construction relies on the assumption that "a majority of CPU power is controlled by nodes that are not cooperating to attack the network."

II.E. Clarifications and Discussion

1. *Permissioned Blockchains.* As interest in Bitcoin and Nakamoto's blockchain surged, many started to use the phrase "blockchain" to describe similarly architected databases maintained by known, trusted parties—that is, without the central scientific innovation of Nakamoto (2008). This concept is sometimes known as a permissioned or private blockchain, or sometimes as distributed ledger technology (see Bakos and Halaburda 2023). An IBM marketing campaign called it "Blockchain for Business." Goldman Sachs called such blockchains "The New Technology of Trust" (Goldman Sachs 2018).

Many researchers and observers view this use of the phrase "blockchain" as hype for what is in essence just an append-only distributed database with well-defined permissions, and with cryptography to protect against data tampering as in Haber and Stornetta (1991). Financial columnist Matt Levine memorably wrote:

> If you announce that you are updating the database software used by a consortium of banks to track derivatives trades, the New York Times will not write an article about it. If you say that you are blockchaining the blockchain software used by a blockchain of blockchains to blockchain blockchain blockchains, the New York Times will blockchain a blockchain about it. (Levine 2017)

As should be clear, this article's critique is of blockchains and permissionless consensus in the sense of Nakamoto (2008), not of distributed databases with trust grounded in traditional sources. It should be uncontroversial that well-architected databases are economically useful, even if there is debate about what to call them. See Budish and Sunderam (2023) for discussion of the potential value of the blockchain data structure in the context of traditional finance.

2. *Smart Contracts.* Notice that Nakamoto's novel form of trust is not specific to currency transactions. One can replace "Al-

ice sends Bob 10 Bitcoins, signed by Alice" with any executable computer instruction signed by Alice. This idea is often called a "smart contract" (see Buterin 2014a).

The analysis framework of this article applies analogously to blockchains that allow smart contracts though the attack possibilities may differ.

3. *Proof of Stake.* The computational work Bitcoin miners must perform to add a new block serves the role of sybil resistance, that is, making it expensive to add new identities to the permissionless system. Without sybil resistance, an attacker could create infinitely many identities.

Since Nakamoto (2008) there have been several other approaches taken to sybil resistance for permissionless consensus, the most prominent of which is proof of stake. Roughly, instead of voting for the correct chain with computational work, participants vote for the correct chain with stake in the cryptocurrency. Ethereum, the second-most valuable cryptocurrency project after Bitcoin, switched from proof of work to proof of stake in fall 2022, and its founder has been discussing the potential benefits of proof of stake since as early as 2014 (Buterin 2014b, 2016).

One motivation for proof of stake over proof of work is to reduce environmental externalities. The computational work that powers Bitcoin consumes on the order of 0.3%–0.8% of all global electricity, which is a fairly astonishing figure.[13] As will become clear, the environmental issue is orthogonal to the economic security concerns raised herein about Nakamoto (2008)—whether the cost of trust support is the cost of computational power or the opportunity cost of holding stake the arguments in Section III go through unchanged.

The more important aspect of proof of stake for the purpose of my argument is that stakes are not as memoryless as computational work: stakes can be locked up on chain for a period of time, like collateral, and observably persist for the time they are locked, like reputation. This opens up the possibility of punishing attackers by confiscating their locked stakes (slashing), which makes attacks more expensive and hence the blockchain more secure.

13. De Vries (2018); Digiconomist (2022). The 0.8% figure is Digiconomist (2022)'s main estimate, whereas the 0.3% figure is based on its best-case analysis under the assumption that all Bitcoin mining equipment is maximally energy efficient.

Intuitively, this is an attempt to algorithmically mimic the traditional trust that is created by law in combination with financial collateral (Buterin 2014b, 2016; Buterin and Griffith 2019). The analysis in Section V shows that if slashing works, it can improve the cost of security by several orders of magnitude.

However, recent research suggests that this approach to security, while intuitively compelling, faces its own limitations. Tas et al. (2023) show as their Theorem 1 that it is impossible to guarantee that a large-enough attacker can be successfully slashed by any positive amount before the attacker is able to withdraw their stake. Budish, Lewis-Pye, and Roughgarden (2024) derive a possibility result for slashing the attacker's stake without punishing honest participants, but the result requires a strong assumption about the nature of the networking environment and the assumption that the attacker's majority is smaller than $\frac{2}{3}$ of the total locked-up stake. If the attacker is too large, then the attacker can circumvent the punishment; a rough intuition is that a large-enough attacker controls the protocol's legal system.[14] An interpretation of these results is that a proof-of-stake blockchain can successfully mimic traditional trust to deter sub-$\frac{2}{3}$ attackers, but that rule of law must step in in the case of a large-enough attacker. See further discussion in Section V.B.

Proof of stake also raises a wide variety of implementation complexities that in turn can create protocol-specific opportunities for dishonest play other than the majority attacks that are central to this article (and universal across all forms of permissionless consensus). See Roughgarden (2023), lectures 12.1–12.24, for detailed discussion of the potential issues and how they are addressed in various protocols. Of special note, the earliest versions of proof of stake used a version of Nakamoto longest-chain consensus and faced a problem called "nothing at stake," under which forks might persist indefinitely in Nash equilibrium

14. The network reliability assumption in Budish, Lewis-Pye and Roughgarden (2024), Theorem 5.1 is that it is possible to impose a delay period between when a user (including honest users) asks to unlock their stake and when they can use it in a transaction, such that the delay period is longer than any feasible attack. Otherwise the attacker could withdraw their stake and spend it before their attack is detected and punished, as in Tas et al. (2023) and the negative results Theorems 4.1 and 4.2 in Budish, Lewis-Pye, and Roughgarden (2024). An implication is that for a proof-of-stake cryptocurrency to have significant real-world economic utility, only a fraction of the total stake can be locked up for trust support with the rest unlocked for potential use.

even with only small participants; see Saleh (2021) for a careful analysis.

III. Nakamoto Trust: A Critique in Three Equations

Sections III.A–III.C present the three-equation critique of Nakamoto's novel form of trust. Section III.D presents a result that shows that the net cost of attack of Nakamoto trust may be zero under strong assumptions. Section III.E presents a one-shot game version of the analysis that captures the essence of the critique of Nakamoto trust while abstracting from many details. Section III.F compares Nakamoto trust to trust from repeated interaction.

III.A. Zero-Profit Condition (Honest Play)

Our conceptual question here is: how much "trust support" will maintain the permissionless consensus if we restrict all participants to behave honestly?

Let there be a large finite number I of honest participants who follow a given permissionless consensus protocol automatically. We may think of I as representing all people who could potentially provide part of the decentralized support for Nakamoto trust. For example, I is the number of people connected to the internet around the world.

Each player i chooses a quantity of "trust support" $x_i \in \mathbb{R}^+$, which we may think of as their quantity of computational work in a proof-of-work blockchain, or their quantity of some other costly action in another blockchain such as stake, storage, and memory. Let $N = \sum_{i=1}^{I} x_i$ denote the total quantity of trust support. A player can choose a quantity of zero if they like, which is how we can think about people not participating in the decentralized trust. Our equilibrium concept for N will be a zero-profit condition. This is meant to capture the permissionless, free-entry/free-exit nature of Nakamoto trust. Nash equilibrium is studied in the one-shot game analysis of Section III.E and is very similar.

Let c denote the cost per unit time to supply one unit of trust support. For example, for proof of work, this is the cost per unit time to run one unit of computational power, including variable costs such as electricity and a rental cost of capital for capital equipment. We sometimes use the notation $c = rC + \eta$, where rC is the rental cost of capital and η is the variable cost of electricity.

For proof of stake, c is the opportunity cost per unit time to supply one unit of stake.

Let p_{block} denote the economic reward paid to a participant who successfully adds a new block of transactions, that is, wins a computational tournament in the case of proof of work. We treat p_{block} as exogenous and derive constraints on it below. Participants' probability of winning the next reward p_{block} is equal to their share of trust support. Specifically, player i wins the next reward with probability $\frac{x_i}{N}$. For the present purposes, we consider the compensation to providers of trust support in aggregate, without distinguishing between whether this compensation is in the form of newly issued cryptocurrency (which are a form of seignorage tax on holders of the currency) or transaction fees.

Let D denote the block difficulty level, defined as the number of units of trust-support-time needed, in expectation, to add one new block. For proof of work, we may assume that new blocks arrive Poisson. That is, if there are N units of trust support, blocks are solved according to a Poisson point process with mean $\frac{D}{N}$. For some proof-of-stake protocols, blocks arrive deterministically at interval $\frac{D}{N}$.[15]

Note a potential source of confusion is that costs c are incurred per unit time, whereas rewards p_{block} are earned per block. The next two concepts help map between objects that are per unit time and objects that are per block. First, we can define profits per unit of trust support per unit time as

$$\frac{1}{N}\frac{D}{N}p_{block} - c,$$

because some unit of trust support solves a block every $\frac{D}{N}$ time in expectation and each of the N units is equally likely to be the winner.

Second, define honest equilibrium as follows:

DEFINITION 1. A zero-profit honest Nakamoto trust equilibrium consists of quantities $\{x_i^*\}_{i=1}^I$ and a difficulty level D^* such that participants (i) solve one block per unit time (as a normalization), and (ii) earn zero economic profits in expectation.

15. A difference between Bitcoin longest-chain proof of work and Ethereum BFT-style proof of stake is that in the latter the protocol directly observes the quantity of trust support N, because the trust support is provided by stake that is locked on chain. In the terminology of Lewis-Pye and Roughgarden (2024) this makes Ethereum a quasi-permissionless protocol.

PROPOSITION 1. Let $N^* = \sum_{i=1}^{I} x_i^*$. In any zero-profit honest Nakamoto trust equilibrium,

$$N^* c = p_{block} \tag{1}$$

and $D^* = N^*$.

Proof. For participants to solve one block per unit time (condition (i)) requires $D^* = N^*$. This in turn implies that profits per unit of trust support per unit time are $\frac{1}{N}\frac{D}{N} p_{block} - c = \frac{1}{N^*} p_{block} - c$. For these profits to be zero (condition (ii)) implies $N^* c = p_{block}$. □

Proposition 1 is widely known and is the standard characterization of a rent-seeking tournament: the prize in the tournament, p_{block}, is dissipated by expenditures aimed at winning the prize, $N^* c$.[16] Prat and Walter (2021) provide empirical support that equation (1) describes actual equilibrium behavior in the Bitcoin mining market, with some additional nuances related to capital adjustment costs.

III.B. Incentive Compatibilty Condition (Majority Attack)

Our conceptual question here is: how much security is generated by the amount of honest trust-support characterized in equation (1)? As discussed in Section II.C, it is widely understood that an agent who provides a majority of the trust support could successfully attack Nakamoto (2008)'s blockchain, for example, by double spending. This is an issue more generally with all forms of permissionless consensus: intuitively, the decentralized trust support "votes" on the true state of the blockchain, and the vote can be manipulated by a majority or supermajority. In this subsection, we focus on the direct costs of attacking a permissionless consensus protocol. In Section V we will consider the possibility that as a post-attack consequence of the attack, the attacker's trust-support capital loses its value or is confiscated.

Consider an additional player, the attacker, not restricted to honest play. This player can attack by choosing AN^* units of trust

16. See Kroll, Davey, and Felten (2013), 8; Huberman, Leshno, and Moallemi (2021), Theorem 1; Easley, O'Hara, and Basu (2019), equation (2); Chiu and Koeppl (2022), Lemma 1; Ma, Gans, and Tourky (2019), equation (7); and Halaburda et al. (2022), equation (4). It is also straightforward to allow for heterogeneous mining costs. Let $c(\cdot)$ denote a continuous weakly increasing function where $c(n)$ gives the per block cost of the nth unit of computational power. Then equation (1) becomes $N^* c(N^*) = p_{block}$. The marginal unit of computational power earns zero economic profits.

support, $A > 1$, for an $\frac{A}{A+1}$ majority at cost AN^*c per unit time. Denote the expected duration of the attack by $t(A)$; the timing details of attacks will vary by protocol, and we derive a closed form for $t(A)$ for longest-chain proof of work in Section IV. Call $AN^*c \cdot t(A)$ the gross cost of attack. The attacker can choose A to minimize $A \cdot t(A)$. Call this optimum $A^* \cdot t(A^*)$; Online Appendix B studies this optimum numerically for longest-chain proof of work.

Let V_{attack} denote the value of attack. For now, let us think about this value of attack in the abstract but have in mind that the value of attack grows as the blockchain's economic usefulness grows.

DEFINITION 2. Nakamoto trust is incentive compatible (IC) against an outsider attack, on a gross-cost basis, if the gross cost of attack exceeds the benefits of attack:

$$A^*N^*c \cdot t(A^*) > V_{attack}. \tag{2}$$

Two brief remarks are required. First, condition (2) is the IC constraint for an outside attacker, but an attack could also come from an insider, that is, part of the current honest trust support N^*. The outside attacker IC condition (2) seems more attactive conceptually, because it treats the honest participants as small and dispersed per the Nakamoto ideal and in equilibrium with the level of compensation p_{block}. That said, an inside attacker might be more realistic in practice; for example, for Bitcoin there is evidence that mining is concentrated (Makarov and Schoar 2021; Cong, He, and Li 2021). Second, the left side of condition (2) is the gross cost of attack. However, the attacker would earn rewards for the blocks in their new chain, which subsidizes their attack. We come back to the distinction between gross and net costs of attack in Section III.D.

III.C. Equilibrium Constraint

In the hoped-for equilibrium in which participants are honest, the amount of trust support devoted to maintaining the blockchain is characterized by the zero-profit equilibrium (1).[17] The IC condition (2) relates this amount of trust support to the level of security generated. Since N^*c appears in both the zero-

17. If trust-support providers earn positive economic profits in equilibrium, then the quantity N^* characterized by the zero-profit condition in equation (1) is an upper bound, and equation (3) obtains as is.

profit condition (1) and the IC condition (2), we can combine the two equations:

THEOREM 1. (Equilibrium Constraint) The zero-profit honest Nakamoto trust equilibrium condition (1) and the IC against outsider attack condition (2) together imply the equilibrium constraint:

$$p_{block} > \frac{V_{attack}}{A^* \cdot t(A^*)}. \tag{3}$$

In words: the equilibrium per block payment to participants for providing trust support on the blockchain must be large relative to the benefits of attacking the blockchain.

Proof. Constraint (3) follows directly from combining equations (1) and (2). □

Economically, this is a very expensive form of trust. Imagine if users of the Visa network had to pay fees to Visa, every ten minutes, that were large relative to the value of a successful attack on the Visa network. Imagine a brand was only as trustworthy as its flow investment in advertising. Imagine that a country were only as secure as its flow expenditure on soldiers at the border.

Another contrast is trust that is supported by rule of law. In such cases, the cost of cheating to the cheating party is related not to the direct costs of conducting the crime but to the costs of potentially getting caught and punished (Becker 1968). A government able to credibly impose large punishments (the parameter f in Becker's model) can deter large attacks or crimes at comparatively low cost. As emphasized, the ingenious aspect of the Nakamoto (2008) form of trust is that it is completely anonymous and decentralized, without any reliance on the rule of law, relationships, or other traditional sources. This aspect also makes the Nakamoto (2008) form of trust much more expensive.

From a computer security perspective, the key thing to note about constraint (3) is that the security of the blockchain is linear in the amount of expenditure on trust support. For instance, if attacking the system grows 1,000 times more attractive, then the cost of securing the system must grow 1,000 times as well. In contrast, in many other contexts investments in computer security yield convex returns, such as traditional uses of cryptography—analogously to how a lock on a door increases the security of a house by more than the cost of the lock. It is much more expen-

sive to break modern cryptography than it is to implement it![18] Imagine if the cost of attacking Visa was that you had to have as much computational power as Visa for a few hours.

A blog post by Ethereum founder Vitalik Buterin (2016) deserves credit for an early informal statement of this equilibrium constraint:

> Because proof of work security can only come from block rewards, and incentives to miners can only come from the risk of them losing their future block rewards, proof of work necessarily operates on a logic of massive power incentivized into existence by massive rewards. Recovery from attacks in PoW is very hard: the first time it happens, you can hard fork to change the PoW and thereby render the attacker's ASICs useless, but the second time you no longer have that option, and so the attacker can attack again and again. Hence, the size of the mining network has to be so large that attacks are inconceivable. *Attackers of size less than X are discouraged from appearing by having the network constantly spend X every single day.* I reject this logic because (i) it kills trees, and (ii) it fails to realize the cypherpunk spirit—cost of attack and cost of defense are at a 1:1 ratio, so there is no defender's advantage. (emphasis added).

Buterin's language can be translated into the formal analysis of this article by dollarizing both uses of X: "Attackers of size less than X" can be interpreted as attackers with an attack opportunity worth less than V_{attack} (not as the size of the attacker's computational power, or else the two X's are not comparable). "Having the network constantly spend X every single day" can be interpreted as assuming that an attack takes one day, that is, $A^* \cdot t(A^*)$ is one day's worth of block-compute costs, and then requiring for security that $A^* \cdot t(A^*) \cdot p_{block} > V_{attack}$, that is, equation (3).

III.D. Zero Net Attack Cost Theorem

What we may call the net cost of attack can differ from the gross cost of attack, modeled above, for three potential reasons: block rewards, attacker cost frictions, and effects of the attack on the value of the cryptocurrency itself.

18. For example, to break an SHA-256 encrypted data set through brute force would require $2^{256} \approx 10^{77}$ calculations. I estimate that if you had a trillion Amazon Web Services' worth of computer power (about $65 billion trillion of capital), running for the age of the universe (14 billion years), that would get you to about 10^{45} hashes.

First, the attacker earns block rewards from their attack. That is, after the attacker's alternative chain replaces the honest chain, the attacker earns the block rewards associated with the blocks in the new longest chain. These block rewards in effect subsidize the attack. An A attacker who attacks for t time performs $At \cdot N^*$ units of trust support. If the difficulty stays constant at $D' = D^* = N^*$, then this corresponds to At block rewards in expectation. If the difficulty on the attacker chain adjusts upward, that is, $D' > D^*$, then the attacker will earn $At \cdot \frac{N^*}{D'} < At$ block rewards.

Second, the attacker may face frictions relative to the cost of honest trust support. For example, in proof of work, the attacker's compute power might be less energy efficient than the honest miners' compute power. In either proof of work or proof of stake, there might be frictions associated with starting and stopping the attack. Let $\kappa \geqslant 0$ parameterize the attacker's cost inefficiency relative to honest mining, such that their total cost of attack is $(1 + \kappa)At \cdot N^*c$.

Third, the attack may harm the value of the cryptocurrency. This reduces the value of the attacker's block rewards and reduces the value of the cryptocurrency the attacker is left with after double spending. If we let $\Delta_{attack} \geqslant 0$ parameterize this decline, this reduces the value of the attacker's block rewards by $\Delta_{attack}At \cdot N^*c$ and reduces the benefit of a double-spending attack originally worth V_{attack} by $\Delta_{attack}V_{attack}$. If the attacker's capital is specific to the attacked cryptocurrency (e.g., ASICs, stake), then the attack would reduce the value of the attacker's capital as well; we return to this issue in Section V.

In the ideal case for the attacker with respect to these three sources of cost difference, we have the following remarkable conclusion:

THEOREM 2. (Zero Net Attack Cost) If the attacker does not face any cost frictions relative to the costs of honest participants ($\kappa = 0$), the attack concludes without any difficulty adjustment ($D' = D^*$), and the attack does not cause the value of the cryptocurrency to fall ($\Delta_{attack} = 0$), then the net cost of attack is zero.

Proof. The attacker's trust-support cost of attack is $(1 + \kappa)At \cdot N^*c$. The net value of the attacker's block rewards is $At \cdot \frac{N^*}{D'} p_{block}(1 - \Delta_{attack})$. The reduction in the value of the cryptocurrency the attacker is left with after double spending is

$\Delta_{attack}V_{attack}$. If $\kappa = \Delta_{attack} = 0$ and $D' = D^*$, then substituting $N^*c = p_{block}$ and $D^* = N^*$ from Proposition 1 yields a net cost of $At \cdot N^*c - At \cdot p_{block} = 0$. □

The intuition behind this result is that the attacker is fully compensated for their trust-support costs for the same reason that honest participants are fully compensated for their trust-support costs under honest play. In effect, permissionless consensus treats the attacker as if they are an honest participant, because the majority determines the truth.

Tabarrok (2019), Auer (2019), Moroz et al. (2020), and Jacob Leshno (personal communication) derive similar results to Theorem 2 building off of Budish (2018).[19] Bonneau (2016)'s analysis of "bribery" attacks deserves early credit for the intuition that the net cost of attacking Bitcoin might be very small because of the block rewards subsidy, as does a blog post by Buterin (2017) who stated the idea in a footnote. Recent work of Gans and Halaburda (2023) generalizes the zero net attack cost result and, by incorporating transaction fees into their model, shows that it is possible for an inside attacker to have a slighty negative net attack cost.

To be clear, zero attack frictions seems unrealistic. But zero friction is often useful as a benchmark case, and the result does reinforce that Nakamoto trust is economically implausible when taken literally.

III.E. One-Shot Game Nash Equilibrium Analysis

The foregoing analysis uses a price-theoretic zero-profit equilibrium concept for honest play and contains several details that are helpful for mapping to the specifics of Nakamoto (2008) but are complex. As a complement to that approach, consider the following stylized one-shot game, which yields a Nash equilibrium solution and abstracts from some of the complexities.[20]

There are I players. Each player i chooses a quantity x_i of trust support (work, stake, etc.) and a "posture" $a_i \in$

19. The analysis in the June 2018 draft artificially constrained the attacker to earn at most t block rewards. The June 2018 draft also did not have explicit cost frictions. Rather, the assumption that an attacker earns at most t block rewards is like an implicit cost friction, related to starting and stopping the attack, of $(A-1)t \cdot N^*c$. As a result, the June 2018 draft had slightly different simulated net costs than here and did not have Theorem 2.

20. I am grateful to Rakesh Vohra for suggesting this one-shot game approach.

$\{Honest, Attack\}$. Costs are c per unit of trust support. Define $N = \sum_{i=1}^{I} x_i$. Payoffs are as follows. If there is a player i with $x_i > \frac{N}{2}$ and $a_i = Attack$, this player gets a payoff of V_{attack}, gross of their costs. All other players get zero. Else, each player gets a payoff of $\frac{x_i}{N} p$.

Our question is: under what conditions does there exist a Nash equilibrium, denoted $\{(a_i^*, x_i^*)\}_{i=I}^{I}$, in which all players choose $a_i^* = Honest$? Call such a profile, if one exists, an honest equilibrium.

LEMMA 1. If there is an honest equilibrium, then $N^*c \leqslant p$ (analog of equation (1)).

Proof. Toward a contradiction, assume there is an honest equilibrium with $N^*c > p$. Choose any player i with $x_i^* > 0$. Player i's net payoff is $\frac{x_i^*}{N^*} p - x_i^* c < x_i^* c - x_i^* c = 0$. So the player has a profitable deviation by choosing $x_i' = 0$ instead. Contradiction. □

In words, this lemma tells us that the amount spent on trust support N^*c will be no greater than the compensation paid for this trust support p, analogously to equation (1).

LEMMA 2. If there is an honest equilibrium with $x_i^* = 0$ for some player i, then $N^*c \geqslant V_{attack}$ (analog of equation (2)).

Proof. Assume there is an honest equilibrium with $x_i^* = 0$. Lemma 1 implies $N^*c \leqslant p$. Consider the possible deviation by player i to choose quantity $x_i' = N^* + \epsilon$ for $\epsilon > 0$ and posture $a_i = Attack$. Since $x_i' > N^*$ player $i's$ attack succeeds. Player i's net payoff from the deviation x_i' is thus $V_{attack} - (N^*c + \epsilon c)$. Player i's net payoff from the conjectured honest equilibrium is 0. Hence to avoid contradiction we need $V_{attack} - (N^*c + \epsilon c) < 0 \rightarrow N^*c > V_{attack} + \epsilon c$. Taking the limit as $\epsilon \rightarrow 0$ yields the desired result. □

PROPOSITION 2. A necessary condition for an honest equilibrium is $p \geqslant \frac{V_{attack}}{1+\frac{1}{I}}$. In the limit as $I \rightarrow \infty$ this is $p \geqslant V_{attack}$ (analog of equation (3)).

Proof. Conjecture an honest equilibrium. Player i's payoff in the honest equilibrium is $\frac{x_i^*}{N^*} p - x_i^* c$. Consider a deviation by i in which they attack by choosing $a_i' = Attack$ and $x_i' = N_{j \neq i}^* + \epsilon = \sum_{j \neq i} x_j^* + \epsilon$ for some $\epsilon > 0$. For this to be worse for player i requires $V_{attack} - N^*{}_{j \neq i} c - \epsilon \leqslant \frac{x_i^*}{N^*} p - x_i^* c$. Rearrang-

ing, using Lemma 1, and noting that $\min(x_i^*) \leqslant \frac{1}{I}N^*$ yields $V_{attack} \leqslant p(1+\frac{1}{I})$. □

An interpretation of the timing of this game is that p and c now represent, respectively, blockchain compensation and trust-support costs for an amount of time commensurate with the duration of an attack, that is, the analog of $A^* \cdot t(A^*)$ in equation (3). The result says that the cost of running the blockchain for an attack duration amount of time must exceed the value of attacking it.

Note that this simple model purposefully restricts attention to honest play or majority attack. Many prior game-theoretic analyses of blockchains assume that all players are small, eliminating the possibility of majority attack, but allow for a richer set of strategies for such small players than just honestly following the protocol (e.g., Eyal and Sirer 2014; Carlsten et al. 2016; Biais et al. 2019; Saleh 2021). Since Proposition 2 is a necessary condition for the existence of an honest equilibrium given the possibility of majority attack, the result still holds even with an enlarged strategy space including behaviors such as selfish mining in Eyal and Sirer (2014) or nothing at stake in Saleh (2021).

III.F. Comparison to Trust from Repeated Interaction

Since Schelling (1956) and Aumann (1959), economists have studied trust that is facilitated by repeated mutually beneficial interaction. Suppose two agents play a repeated prisoner's dilemma game where each participant's per period payoff from cooperation is τ (for trust), a participant who defects while the other cooperates earns $b > \tau$ (for betray), if both players defect they earn zero, and the discount factor is δ. In the well-known grim-trigger equilibrium, cooperation is possible if

$$b - \tau \leqslant \frac{\delta}{1-\delta}\tau, \tag{4}$$

that is, if the one-shot benefit of cheating is smaller than the net present value of the trusting relationship.

To compare this form of trust from repeated interaction, "Schelling trust," to Nakamoto's permissonless consensus, let us rewrite equation (4) using $V_{attack} \equiv b$ and $V_{trust} \equiv \frac{1}{1-\delta}\tau$, and consider repeated play of the one-shot game we just analyzed in Section III.E. We thus have the comparison of incentive-

compatibility conditions:

$$\begin{aligned} &\textit{Schelling Trust}: \; V_{trust} \geqslant V_{attack} \\ &\textit{Nakamoto Trust}: \; p \geqslant V_{attack}. \end{aligned} \tag{5}$$

The reader will observe from expression (5) that the first essential difference between Nakamoto (2008) trust and trust from mutually beneficial repeated interaction is on the cost-of-attack side. In Schelling (1956), the cost of attack is the loss of the net present value of the mutually beneficial relationship.[21] The more valuable the relationship, the more value is secured against attack. In Nakamoto trust, by contrast, the repeated interaction among providers of trust support (e.g., miners) is zero profit, as if $\tau = 0$. Instead, the cost of attack comes from a different source, which is that for a short period of time (a single play of the game in Section III.E) an attacker has to be larger than the honest trust-support providers who are themselves playing a zero-profit game (in the limit as $I \to \infty$). Put differently, in Schelling (1956) the cost of attack is a stock (the net present value of the relationship) whereas in Nakamoto (2008) the cost of attack is a flow (the recurring costs of honest trust support for a short period of time).

The second essential difference between these two forms of trust is the cost of honest play. In Nakamoto trust, honest play costs the same p per period as it costs to attack—that is the equilibrium cost to induce the zero-profit trust support on a continual basis. In Schelling (1956), conditional on an existing trust relationship, the recurring cost of honest play is zero. Of course, in many contexts, developing a trust relationship in the first place may take investment (e.g., investment in a brand), so one can think of the per period cost as the interest on this fixed cost investment.

Table I summarizes this comparison. It is also worth noting that Schelling trust on its own likely is not enough for high-value financial applications. In financial applications, one can think of τ as the gains from trade in a transaction, and think of b as like the nominal value of the transaction (Budish and Sunderam 2023).

21. From Schelling (1956), p.300: "What makes many agreements enforceable is only the recognition of future opportunities for agreement that will be eliminated if mutual trust is not created and maintained, and whose value outweighs the momentary gain from cheating in the present instance." See Wolitzky (forthcoming) for a wide-ranging survey of models of trust from repeated interaction.

TABLE I
COMPARISON OF NAKAMOTO TRUST AND SCHELLING TRUST

	Cost of attack	Cost of honest play
Nakamoto trust	p, a flow cost. Attacker pays cost of honest trust support for a short period of time.	p per period. Cost to induce the zero-profit honest trust support on a recurring basis.
Schelling trust	$V_{trust} = \frac{1}{1-\delta}\tau$, a stock cost. Attacker loses the net present value of the trust relationship.	0 per period, given preexisting trust relationship.

Notes. See discussion in the text of Section III.F. Entries for Nakamoto trust are based on repeated play of the one-shot game analyzed in Section III.E in the limit as the number of players grows large.

So it takes a lot of repeated interaction to sustain trust for high stakes—it is helpful to have some support from the rule of law and collateral. We will return to this topic in Section VI (see especially Table VI).

IV. ANALYSIS OF DOUBLE-SPENDING ATTACKS

Equation (3) tells us that the possibility of a double-spending attack places equilibrium constraints on Nakamoto's novel form of trust. This section is an effort to understand these constraints quantitatively. For this quantification exercise, I focus specifically on Nakamoto (2008)'s proof-of-work longest-chain protocol. Section IV.A briefly describes the mechanics of a double-spend attack. Section IV.B analyzes the expected duration and cost of double-spend attacks. Section IV.C analyzes the implied equilibrium cost of Nakamoto trust using two related thought experiments for the value of attack. First, given a potential dollar value of V_{attack}, what is the equilibrium cost of security in dollars and as a percentage of V_{attack}? Second, given a potential dollar volume of honest transactions using Nakamoto trust, denoted V_{honest}, and an assumption that an attacker can double spend this honest volume for a period of time, what is the equilibrium cost of security as a percentage of honest transaction volume?

A caveat for this section is that there is no perfect way to model V_{attack}. These thought experiments reflect my best attempt to analyze the equilibrium implications of the analysis in Section III as transparently as possible. Future research may improve on these efforts.

IV.A. Mechanics of a Double-Spend Attack

The canonical attack Nakamoto (2008) worries about is called a double spend, in which an attacker is able to spend the same currency multiple times by effectively deleting some of their transactions from the record. Such attacks are also called safety or consistency violations in the distributed-consensus literature. In a double-spending attack on a longest-chain consensus protocol, the attacker engages in the following actions in sequence:

(i) The attacker sends cryptocurrency in exchange for goods or assets, potentially in many transactions over many blocks.
(ii) The attacker allows those transactions to be added to the blockchain in the usual way, described in Section II.B.
(iii) The attacker works in secret to create an alternative longest chain that does not include the transactions in (i). Instead they send the cryptocurrency to other accounts they control.
(iv) The attacker waits for any escrow periods to elapse, so they receive the goods or assets they transacted for in (i).
(v) The attacker releases their alternative longest chain. The attacker now has both the goods or assets they transacted for and their cryptocurrency.

See Figure III for an illustration.

IV.B. Duration and Gross Cost of Attack

The denominator of the right side of equation (3), $A^* \cdot t(A^*)$, is the cost of a double-spend attack in units of equilibrium per block trust-support costs N^*c. It is possible to obtain a closed-form expression for the expected duration t of a double-spending attack as a function of the attacker majority A and other parameters that affect duration. Specifically, let e denote the escrow period, and let k denote the number of blocks in which the attacker places transactions that they will subsequently revert (i.e., the number of blocks in step (i) of the attack). Assume that honest participants produce new blocks as a Poisson process with arrival rate one and the attacker produces new blocks as a Poisson process with arrival rate A.

PROPOSITION 3. For Nakamoto (2008) proof-of-work longest-chain consensus, the expected duration t of the double-spending attack as a function of the attacker majority A, escrow period e, and number of blocks in which the attacker

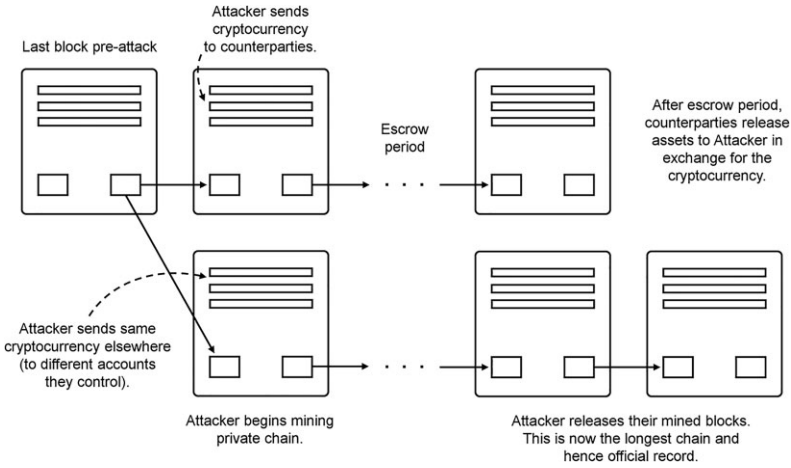

FIGURE III
Illustration of Double-Spending Attack

places transactions k, is given by:

$$t(A,e,k) = (k+e) + \left[\sum_{i=0}^{k+e}\left(\frac{i+1}{A-1}\right)\cdot\frac{(2(k+e)-1-i)!}{(k+e-i)!(k+e-1)!}\right.$$
$$\left.\times\left(\frac{A}{1+A}\right)^{k+e-i}\left(\frac{1}{1+A}\right)^{k+e}\right]. \quad (6)$$

As the attacker majority grows large ($A \to \infty$), $t(A,e,k)$ converges to $k+e$. In the limit as $A \to_+ 1$, we have $t(A,e,k) \to \infty$.

Proof. See Online Appendix B. □

Prior work by Grunspan and Pérez-Marco (2018, 2022) derives closely related mathematical results for the analysis of an attacker with less than 50% of the total hash power. Expression (6) can be understood as follows. In the attacker's best case, their attack takes $k+e$ time. This best case occurs if, as soon as the honest participants have produced $k+e$ blocks and all of the assets the attacker has transacted for have been released from escrow, the attacker is ready with their alternative longest chain of at least $k+e+1$ blocks. Suppose, on the other hand, that the attacker is behind the honest chain by $i \geqslant 0$ blocks at the time the honest participants produce their $k+e$ block. Given the Pois-

son arrival processes, it will take the attacker $\frac{i+1}{A-1}$ of time in expectation to strictly surpass the honest chain. The last part of the expression gives the probability that the attacker's deficit is i blocks.

Table II provides example calculations of duration t and the gross trust-support-cost term At. For example, if $k+e=12$ blocks (2 hours), then attacker majorities of $A=1.2-1.5$ generate average attack durations of 13.4–19.7 blocks, or about 2–3.5 hours, and average gross costs of 20.1–23.6 times N^*c. Notably, smaller majorities lead to significantly longer attack durations and higher costs. For example, if $A=1.05$, which corresponds to a 51.2% attacker majority, the average duration is 56.6 blocks (9.5 hours) and the average gross cost is 59.4 times N^*c.

As $k+e$ grows larger, the average attack duration t gets proportionally closer to the best-case duration $k+e$ for all values of A. The intuition for this is simple law of large numbers. However, even for very large values of $k+e$, such as $k+e=1{,}008$ which corresponds to a full week, the gross-cost-minimizing attacker majority is larger than 51%. It is true that a 51% majority is enough to ensure statistically that the attack will eventually succeed, but a cost-minimizing attacker will choose a somewhat larger majority.

Online Appendix B provides numerical analysis of the cost-minimizing attacker majority A^* and minimum gross costs $A^* \cdot t(A^*)$ as a function of $k+e$.

IV.C. Implied Equilibrium Cost of Nakamoto Trust

1. *Thought Experiment 1: What Is the Equilibrium Cost of Security for a Given Value of V_{attack}?* A majority attacker will not use their majority to double spend for a cappucino at Starbucks. They will use their majority to conduct transactions that are as large as possible given the current uses of cryptocurrencies, possibly using many different addresses over many blocks of transactions. V_{attack}, therefore, can be understood as a statistic on the economic usefulness of Nakamoto trust. The higher the economic throughput of the blockchain for honest participants, the more value an attacker can double spend.[22]

22. This point likely seems obvious, but it was missed in past academic literature on double-spending attacks. The computer science literature did not explicitly model the economic benefits of attack, and therefore missed that the value of attack might scale with Bitcoin's usefulness (Rosenfeld 2014; Eyal and Sirer 2014;

TABLE II

EXPECTED DURATION AND GROSS COST OF ATTACK

	# Blocks of double spending + escrow period ($k+e$)					
Attacker majority	1	6	12	36	144	1,008
Expected duration of attack in units of honest block time t						
$A = 1.05$ (51%)	25.51	42.43	56.59	97.91	229.61	1,074.55
$A = 1.10$ (52%)	13.02	22.88	31.73	60.14	167.31	1,010.90
$A = 1.20$ (55%)	6.79	13.24	19.69	43.10	146.79	1,008.00
$A = 1.33$ (57%)	4.34	9.60	15.35	38.04	144.18	1,008.00
$A = 1.50$ (60%)	3.08	7.84	13.43	36.47	144.00	1,008.00
$A = 2.00$ (67%)	1.89	6.45	12.20	36.01	144.00	1,008.00
$A = 5.00$ (83%)	1.12	6.00	12.00	36.00	144.00	1,008.00
Expected gross cost of attack in units of per block honest trust support N^*c						
$A = 1.05$ (51%)	26.78	44.56	59.42	102.81	241.09	1,128.28
$A = 1.10$ (52%)	14.32	25.16	34.90	66.16	184.04	1,111.99
$A = 1.20$ (55%)	8.14	15.89	23.63	51.72	176.15	1,209.60
$A = 1.33$ (57%)	5.78	12.77	20.41	50.59	191.76	1,340.64
$A = 1.50$ (60%)	4.62	11.76	20.14	54.71	216.01	1,512.00
$A = 2.00$ (67%)	3.78	12.89	24.39	72.02	288.00	2,016.00
$A = 5.00$ (83%)	5.59	30.02	60.00	180.00	720.00	5,040.00

Notes. Expected duration t as a function of the attacker majority A, escrow period e, and the number of blocks in which the attacker places transactions k is computed using formula (6) in the text and double-checked using a computational simulation. Each block corresponds to 10 minutes in expectation. In clock time, the columns correspond to 10 minutes, one hour, two hours, six hours, one day, one week.

TABLE III
COST TO SECURE AGAINST ATTACK: BASE CASE ANALYSIS

	Per block	Per day	Per year	Per transaction
Security costs as % of value secured	5.00%	720%	262,801%	0.0025%
To secure:				
$1 thousand	$50.0	$7.2 K	$2.6 M	$0.025
$1 million	$50.0 K	$7.2 M	$2.6 B	$25.0
$1 billion	$50.0 M	$7.2 B	$2.6 T	$25.0 K
$10 billion	$500.0 M	$72.0 B	$26.3 T	$250.0 K
$100 billion	$5.0 B	$720.0 B	$262.8 T	$2.5 M

Notes. See equation (3) and the text of Section IV.C for description. The base case scenario assumes $k + e = 12$ blocks (two hours). The abbreviations K = thousand; M = million; B = billion; T = trillion.

In this first thought experiment, I simply consider a wide range of dollar values for V_{attack}. I use $1,000 as the low end of this range, representing Bitcoin's early days when even buying a pizza was remarkable. I use $100 billion to represent the high end of this range. Although arbitrary, this seems a reasonable order of magnitude for a large-scale attack on the global financial system—for example, the U.S. financial system has daily transaction volume that conservatively exceeds $4 trillion (Budish and Sunderam 2023). This figure also represents about 7% of Bitcoin's peak market capitalization.

Table III presents results for a base case scenario in which $k + e = 12$. This corresponds to an escrow period of one hour ($e = 6$), as is standard practice for Bitcoin, and an assumption that the attacker spreads their double-spend transactions over one hour's worth of blocks ($k = 6$).[23] To keep the Nakamoto blockchain secure in this base case requires a per block cost that is 5.00% of the value secured against a double-spending attack. This fol-

Bonneau 2016). In economics, a model of Chiu and Koeppl (2022) assumes that an attack involves just a single transaction and holds this transaction size fixed. The authors conclude that the system becomes more secure as its economic value grows relative to this fixed transaction size. This is like noting that it is less attractive to engage in a double-spending attack for a cappucino in 2024 than it was in 2009.

23. A prior draft did not have the k parameter and instead assumed that the attacker placed all their double-spend transactions in a single block. This is equivalent to assuming $k = 1$, making $k + e = 7$ the base case in that version with Bitcoin's standard escrow period of $e = 6$. This accounts for the difference in the numbers in Table III.

TABLE IV
COST TO SECURE AGAINST ATTACK: SENSITIVITY ANALYSIS

	# blocks of double spending + escrow period ($k+e$)					
	1	6	12	36	144	1,008
Per block security costs as % of V_{attack}	26.74%	8.55%	5.00%	1.99%	0.57%	0.09%
Annual security costs if V_{attack} equals:						
\$1 thousand	\$14.1 M	\$4.5 M	\$2.6 M	\$1.0 M	\$301.3 K	\$47.8 K
\$1 million	\$14.1 B	\$4.5 B	\$2.6 B	\$1.0 B	\$301.3 M	\$47.8 M
\$1 billion	\$14.1 T	\$4.5 T	\$2.6 T	\$1.0 T	\$301.3 B	\$47.8 B
\$10 billion	\$140.5 T	\$44.9 T	\$26.3 T	\$10.5 T	\$3.0 T	\$478.2 B
\$100 billion	\$1.4 Q	\$449.1 T	\$262.8 T	\$104.7 T	\$30.1 T	\$4.8 T

Notes. Each block corresponds to 10 minutes in expectation. In clock time, the columns correspond to 10 minutes, 1 hour, 2 hours, 6 hours, 1 day, 1 week. The abbreviations K = thousand; M = million; B = billion; T = trillion; Q = quadrillion.

lows directly from equation (3), rewritten as $\frac{p_{block}}{V_{attack}} \geqslant \frac{1}{A^* \cdot t(A^*)}$, with $t(A^*) = 13.99$ at the attacker-optimal choice of $A^* = 1.43$ (or 59%). Per transaction, assuming 2,000 transactions per block, the cost is 0.0025% of the value secured against attack.

These costs likely sound economically plausible. But consider how they scale with time and the amount of value secured. A cost of 5% per block amounts to 720% of the value secured against attack per day, and about 263,000% of the value secured against attack per year. For example, to secure the system against a \$1 billion attack requires \$2.63 trillion of annual security expense. To secure the system against a \$40 billion attack requires \$105 trillion a year, or all of 2023 global GDP per the World Bank.

The per transaction fee of 0.0025% likely sounds very small, but this is a percentage of the value secured against attack, not the size of the transaction. For example, if an attack could be worth \$1 billion, then each transaction must pay 0.0025% of \$1 billion which is \$25,000 of security costs. The intuition is that every transaction has to implicitly pay for the costs of the large standing army—as if a transaction for a cappucino has the security required by Fort Knox.

Table IV presents a sensitivity analysis. As the escrow period e grows, and the number of blocks k it takes the attacker to execute their double-spend transactions grows, the cost of security declines. Whereas the cost of security is 5% of V_{attack} per block in the base case, the cost declines to 0.09% of V_{attack} if $e + k$ amounts to a week ($e + k = 1{,}008$). However, even in the scenario

where $e + k$ equals a week, the cost to secure against large attacks remains high. To secure against a $10 billion attack requires an annual security cost of $478 billion, which is more than the United States' annual spending on prisons, courts, and police (see Section VI). To secure against a $100 billion attack requires an annual security cost of 4.5% of global GDP.

2. *Thought Experiment 2: What Is the Equilibrium Cost of Security as a Percent of Honest Transaction Volume?* As a second thought experiment, assume that honest users of a cryptocurrency transact an average of V_{honest} of volume per block, and that a majority attacker can double spend exactly this average honest volume per block, for k blocks total. That is, $V_{attack} = kV_{honest}$. This is a simple structural model in which the value of a majority attack grows with economic usefulness, as measured by the amount of honest transaction volume. I caution that this thought experiment has likely biases in both directions. It is conservative in that a majority attacker would seek to transact the largest possible amounts, not just average amounts.[24] It is aggressive in that a majority attacker would likely share block space with honest users whose transactions happen to be in the mempool at the time the attacker starts double spending.

Table V presents results for this thought experiment. In the base case in which the escrow period is one hour ($e = 6$) and the attacker can double spend for one hour worth of blocks ($k = 6$), that is, $V_{attack} = kV_{honest}$ corresponds to one hour of average honest transaction volume, the equilibrium cost of security is 30% of V_{honest} per block.[25] This percentage cost increases with the num-

24. Let honest transaction volume be drawn from a distribution with support on $[\underline{V}, \overline{V}]$ and mean V_{honest}. For the system to facilitate honest transaction volume for any quantity drawn from the distribution, it must be the case that it is not worthwhile to double spend even for the maximum honest transaction value $\overline{V}$. We can use data on the real-world distribution of blockchain transaction volume to understand the empirical relationship between observed $\overline{V}$ and V_{honest}. If we look at Bitcoin volume in BTC in 2023, the ratio of maximum-to-mean volume $\frac{\overline{V}}{V_{honest}}$ is about 7.5 at the one-hour level, 5.2 at the two-hour level, and 1.85 at the one-day level. Makarov and Schoar (2021) find that about 90% of Bitcoin volume is spurious. Under the strong assumption that the attacker double spends the maximum observed transaction volume and it is all nonspurious, then these ratios of $\frac{\overline{V}}{V_{honest}}$ could be multiplied by a factor of 10.

25. Online Appendix Table C.1 compiles data on double-spend attacks on forks of Bitcoin and Ethereum. The Bitcoin Gold attacks had a length of largest reorga-

TABLE V
COSTS TO SECURE AGAINST ATTACK AS PERCENT OF HONEST VOLUME

V_{attack} # blocks of honest volume (k)	Escrow period (e)					
	1	6	12	36	144	1,008
1	18.07%	7.61%	4.69%	1.94%	0.57%	0.09%
6	45.68	30.00	21.55	10.45	3.31	0.54
12	56.26	43.10	33.85	18.58	6.39	1.08
36	70.01	62.67	55.73	38.82	16.78	3.17
144	82.02	79.50	76.66	67.14	43.26	11.52
1,008	91.63	91.20	90.68	88.67	80.63	46.87

Notes. Each block corresponds to 10 minutes in expectation. In clock time, the columns correspond to 10 minutes, one hour, two hours, six hours, one day, and one week.

ber of blocks the attacker can double spend for and decreases with the escrow period. For example, if the attacker can double spend one hour of average honest transaction volume ($k = 6$), but the escrow period is one day ($e = 144$), then the required security cost is just 3.31% of honest transaction volume—roughly in line with the costs of credit card transactions in the traditional financial system. In the other direction, if the escrow period is the standard one hour ($e = 6$) but the attacker can double spend one day's worth of average honest transaction volume ($k = 144$), then the required security cost is 79% of V_{honest} per block. For any fixed escrow period, the cost converges to 100% of V_{honest} as $k \to \infty$.

IV.D. Discussion

The double-spending analysis is consistent with the modest early use cases of Bitcoin, in which Bitcoin was primarily used by hobbyists and for small-scale black market activity (e.g., online gambling, Silk Road). In these early days, the amount that could be gained in a double-spending attack was not very high, because there were not high-value transaction opportunities. If a double-spending attack could gain at most $1,000, then the implicit cost per transaction in the base case necessary to secure the trust is just $0.025.

nization of 16–22 blocks (roughly 2.5–4 hours). The Ethereum Classic attacks had lengths of 140 blocks to 7,000 blocks (30 minutes to 1 day). The details of these attacks are thinly reported but may give a sense of the range of realistic values of k and e.

The double-spending analysis is also consistent with larger-scale black market uses of cryptocurrencies, especially as black market users may be most willing to pay the high implicit costs. For example, if a double-spending attack could gain at most $10 million, then the implicit cost per transaction in the base case needs to be about $250. This is modest relative to the costs of transporting large amounts of cash (Rogoff 2017).

Where the analysis suggests greater skepticism is the use of cryptocurrencies and Nakamoto trust as a major component of the mainstream global financial system (again, considering its pure form without support from the rule of law). If cryptocurrencies and Nakamoto trust were to become more integrated with the mainstream global financial system, then it would be possible to move amounts of value that are ordinary in the scheme of global finance, and hence it would be possible to double spend for amounts of value that are ordinary in the scheme of global finance. The analysis suggests that this scenario is unrealistic because of the way the trust model scales. To secure the system against attacks of $1 billion—which is less than 0.2% of daily trading volume in the U.S. Treasury market alone—requires a per transaction security cost of $25,000 and an annual security cost of $2.6 trillion. To secure against attacks of $40 billion requires an annual security cost of all of global GDP. While market power and fees in traditional finance are clearly an important economic issue (Greenwood and Scharfstein 2013; Philippon 2015), and Huberman, Leshno, and Moallemi (2021) are careful to remind us to compare the costs of the Nakamoto trust model against the costs of market power in traditional finance, it is clear from these calculations that Nakamoto trust is very expensive relative to traditional trust. I return to this comparison between Nakamoto trust and traditional trust in Section VI.

A conceptual insight that is reinforced by the double-spending analysis is that blockchain security should be thought of not as a 0-1 variable that breaks at a threshold ρ, as in the distributed consensus literature, but as more like a (high) percentage tax.

V. Specific Capital and Collapse

Nakamoto (2008) envisioned that ordinary computers would be used to maintain the Bitcoin blockchain, famously using the phrase "one-CPU-one-vote." Since 2013, however, Bitcoin min-

ing has been dominated by specialized computer chips called ASICs (application-specific integrated circuits). ASICs have the SHA-256 hash function etched directly onto hardware, which makes them extremely efficient at Bitcoin mining and useless for any other application that does not involve performing a large number of SHA-256 hashes.[26] Many of the other largest cryptocurrencies use a proof-of-stake consensus protocol, most prominently Ethereum since fall 2022. With proof of stake, the capital used to maintain the blockchain is intrinsically specialized to the blockchain, as the capital is literally units of the cryptocurrency.

If the capital used to maintain the blockchain is specialized (as opposed to repurposable), and a majority attack causes the attacker's capital to lose its value or be confiscated, then the attacker cost model needs to be modified. In addition to charging the attacker the flow cost of the attack, we also need to charge the attacker the stock value of their loss of specialized capital. This makes an attack significantly more expensive.

Section V.A redoes the theoretical analysis of Section III under an alternative incentive constraint that includes the stock value of the attacker's capital. Section V.B considers three logical possibilities for how an attacker might lose their capital value: collapse of the cryptocurrency, internal-to-protocol punishment without collapse, and external-to-protocol responses to the attack. Section V.C discusses the analysis.

V.A. Analysis if an Attacker Loses Their Capital

Let $c = rC + \eta$ denote the cost per unit time to supply one unit of trust support, where C denotes the fixed cost of specialized capital (e.g., ASICs, stake), r denotes the rental cost of capital per unit time (risk-adjusted interest expense plus depreciation), and η denotes variable costs per unit time (e.g., electricity). The honest Nakamoto trust equilibrium condition (1) can be rewritten:

$$N^*(rC + \eta) = p_{block}. \tag{7}$$

26. These specialized chips are so much more efficient than general purpose chips that execute SHA-256 in software that, I estimate, even if one controlled all of Amazon Web Services that would amount to about 0.05% of Bitcoin's hash rate. This calculation is based on Amazon Web Services owning $65 billion of technology capital per its 2021 10-K filing, the calculations below that the Bitcoin capital stock is about $12.5 billion, and a conservative assumption that specialized ASIC chips are at least 10,000 times more economically efficient at SHA-256 hashing than general-purpose computers.

An outside attacker would need at least N^*C worth of specialized capital to conduct the attack, whereas an inside attacker would need at least $\frac{N^*C}{2}$ of capital.

For this subsection, let us focus on an outside attacker who loses all of their capital as a result of the attack, for example, because of a total collapse of the cryptocurrency—this is the case for which the cost of attack is highest. Given how small the flow costs of attack are, as analyzed in Section IV, let us ignore these and focus only on the stock cost of the specialized capital. This yields an approximate attack cost of N^*C and an approximate incentive compatibility condition of

$$N^*C > V_{attack}. \tag{8}$$

We can compute N^*C as a function of p_{block}. Let $\mu = \frac{rC}{rC+\eta}$ denote the capital share of trust support. The honest equilibrium equation (7) can be rewritten:

$$N^*C = \frac{\mu p_{block}}{r}. \tag{9}$$

Hence we can derive a modified version of equation (3) :

$$p_{block} > \frac{r}{\mu} V_{attack}. \tag{10}$$

This is several orders of magnitude more secure than equation (3) because r is the interest rate per unit time. Here is an example calculation for Bitcoin. Assume the capital share of mining is $\mu = 0.4$ (De Vries 2018; Digiconomist 2022), and the annual discount rate for ASICs is 50% (ASICs depreciate quickly and mining is risky), which implies that the per unit time discount rate is $r \approx 0.0008\%$. This means $\frac{r}{\mu} = 0.002\%$. Now compare $\frac{r}{\mu}$ on the right side of equation (10) to the $\frac{1}{A^* \cdot t(A^*)}$ factor on the right side of equation (3). If we use the base case value of $\frac{1}{A^* \cdot t(A^*)} = 5.0\%$, we have a roughly 2,500-fold improvement in the cost of security.

If we use these same values for μ and r and use p_{block} of \$250,000, then equation (9) implies a capital stock of \$12.5 billion, which about matches what is implied by current prices for state-of-the-art ASIC machines.[27] This suggests these magnitudes are reasonable.

27. A Bitmain Antminer S19j XP has a current retail price of \$4,983 (https://www.bitmain.com/, accessed August 31, 2023) and it would take about 2.3 million of these machines to match Bitcoin's current hash rate, for a total capital cost of

V.B. Issue: How Does an Attacker Lose Their Capital?

1. *Collapse of the Cryptocurrency.* One way the attacker can lose their capital value is if the majority attack causes a significant decline in the value of the cryptocurrency. For example, a majority attack on a major cryptocurrency would be widely reported and might cause a market crash.

Mathematically, suppose the majority attack causes a proportional decline in the value of the cryptocurrency of Δ_{attack}. If the specialized capital is stake, and the attacker is unable to withdraw their stake before the decline occurs, then their stake declines in value by $\Delta_{attack}N^*C$. If the specialized capital is hardware, and the attacker is unable to liquidate their hardware before the decline occurs, then the decline can be faster than rate Δ_{attack}. Specifically, if we assume that the market is in equilibrium (7) before the attack and we assume that the postattack equilibrium value of specialized capital reflects a permanent decline in block rewards of Δ_{attack}, then the specialized capital declines in value by proportion $\max(1, \frac{\Delta_{attack}}{\mu})$.[28]

Thus, if Δ_{attack} is large, this significantly increases the attacker's costs. The Bitcoin Wiki classifies the majority attack into its "probably not a problem" category for this reason (Bitcoin Wiki 2020).[29] However, vulnerability to collapse is undesirable for two related reasons. First, collapse harms honest holders of the cryptocurrency and the specialized capital. Second, there is the possibility of an attack motivated by this harm per se, that is, a sabotage attack. The possibility of a sabotage attack was first raised by Rosenfeld (2014).[30]

---

about $11.5 billion at retail prices. I do not have any information on how retail prices relate to the prices paid by large-scale miners.

28. Starting with the equilibrium $N^*(rC + \eta) = p_{block}$, we want to recompute the postattack equilibrium value of capital C' given a decline in rewards to $(1 - \Delta_{attack})p_{block}$. This can be computed as $N^*rC' = (1 - \Delta_{attack})p_{block} - N^*\eta = (1 - \Delta_{attack})p_{block} - (1 - \mu)p_{block}$. So the ratio $\frac{C'}{C} = \frac{(1-\Delta_{attack})-(1-\mu)}{1-(1-\mu)} = \frac{\mu-\Delta_{attack}}{\mu} = 1 - \frac{\Delta_{attack}}{\mu}$. If Δ_{attack} exceeds μ, then the postattack capital value is zero and some capital will be mothballed given the decline in compensation. A richer model would reflect the stochastic nature of p_{block} and the associated option value of capital, along the lines of Prat and Walter (2021), both before and after the attack.

29. "A miner with more than 50% hash power is incentived [*sic*] to reduce their mining power and refrain from attacking in order for their mining equipment and bitcoin income to retain its value."

30. "In this section we will assume $q < p$ [i.e., that the attacker does not have a majority]. Otherwise, all bets are off with the current Bitcoin protocol... The hon-

What is the value of a sabotage attack on a significant cryptocurrency such as Bitcoin or Ethereum? It is hard to say, of course, but easy to imagine that the magnitudes are already large and would be larger still if cryptocurrencies become more significantly integrated into the global financial system. Open interest in CME Bitcoin futures as of April 2024 is about 28,000 contracts, each tracking five Bitcoins, worth about $9 billion at current prices. According to data from The Block, open interest in Bitcoin futures aggregated across the major crypto exchanges is about $24 billion as of April 2024 and about $9.5 billion for Ethereum futures.[31] These figures give a sense of magnitudes for what could be made from a short-selling attack.

The market capitalization of cryptocurrencies gives another sense of magnitudes for the amount of economic harm an attacker could cause. Bitcoin's market capitalization has been as high as about $1.4 trillion and Ethereum's as high as about $500 billion. Across all crypto assets tracked by CoinMarketCap, market capitalization peaked at about $3 trillion. Paypal co-founder Peter Thiel (2022) recently predicted that Bitcoin will be worth more than $100 trillion.

Last, Ethereum founder Vitalik Buterin described a future in which it is "just considered normal for there to be *trillion dollar assets* that are managed on Ethereum" (Klein 2022, emphasis added). If indeed assets of that magnitude are managed on Ethereum or other blockchains, without implicit or explicit protections from the rule of law, then the value and risk of sabotage would be large.

---

est miners, who no longer receive any rewards, would quit due to lack of incentive; this will make it even easier for the attacker to maintain his dominance. This will cause either the collapse of Bitcoin or a move to a modified protocol. As such, *this attack is best seen as an attempt to destroy Bitcoin*, motivated not by the desire to obtain Bitcoin value, but rather wishing to maintain entrenched economical systems or obtain speculative profits from holding a short position" (p. 10, emphasis added).

31. CME open interest data is available via its website. I found open interest data from crypto exchanges at https://www.theblock.co/data/crypto-markets/futures/. I believe this to be a credible source but am less confident in it than I am in the CME figures. For what it's worth, when I wrote the June 2018 draft of this article, CME + CBOE open interest was about $160 million, and crypto exchange futures did not, to my knowledge, exist at the time. That is, futures market open interest has grown by two orders of magnitude in the past few years.

2. *Internal-to-Protocol Punishment without Collapse.* It would be very attractive to get the security benefits of an attack costing a stock not a flow, that is, an attack that costs $O(N^*C)$ yielding equilibrium constraint (10) rather than costing $O(N^*c)$ yielding equilibrium constraint (3), without needing the cryptocurrency to collapse. This is what Ethereum proof-of-stake consensus with "slashing" is trying to accomplish.

Slashing relies on two distinct departures from Nakamoto (2008)'s version of permissionless consensus. First, the trust-support capital is stake that exists on-chain, as opposed to computational equipment that exists in the physical world off-chain. Second, Ethereum's consensus protocol requires, roughly speaking, at least $\frac{2}{3}$ of all of the trust-support capital to sign a block before it is added to the record, in what is known as the BFT paradigm. These features together mean that in the event of a double-spending attack, the attacker will leave cryptographic proof that they have violated the protocol. If blocks A and A' conflict each other (e.g., send the same currency to two different places), then for both to be confirmed, at least $\frac{2}{3}$ of all of the stake has to have signed A and $\frac{2}{3}$ has to have signed A', so at least $\frac{1}{3}$ of all of the stake has to have signed both conflicting transactions. The idea of slashing is simply to confiscate the stake that has signed conflicting transactions—to "slash" it from the record. See Figure IV.

Intuitively, slashing is trying to mimic the traditional trust combination of collateral plus rule of law. A counterparty posts collateral, and if they cheat the legal system can seize the collateral. The difficulty is that a large-enough attacker can effectively control or stall the protocol's legal system. An impossibility result of Tas et al. (2023) (Theorem 1) shows that is is impossible for a proof-of-stake protocol to achieve any positive amount of what they call "slashable safety," meaning confiscating any positive amount of the attacker's stake, if the attacker can be large enough. The proof shows that a large-enough attacker (specifically, with at least $\rho = \frac{2}{3}$) can always withdraw all of the capital that signed conflicting transactions before the confirmation of any transactions that slash their stake.

Budish, Lewis-Pye, and Roughgarden (2024) provide a possibility result for proof of stake with slashing to successfully confiscate the attacker's capital that signed conflicting transactions if both (i) the attacker is strictly less than $\rho = \frac{2}{3}$ of the total stake,

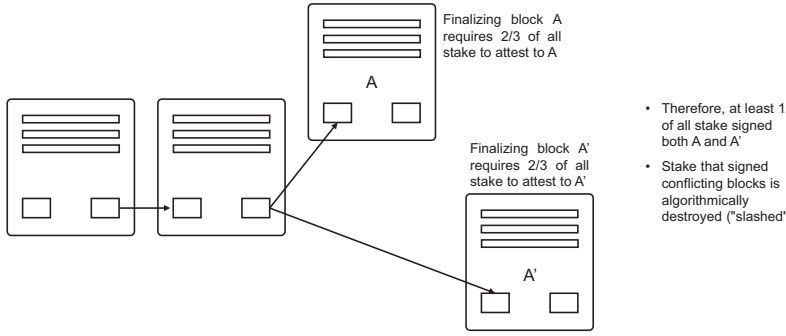

FIGURE IV

Proof of Stake with Slashing

Double-spending attacks (or more generally safety violations) leave on-chain proof of malfeasance in BFT-style proof-of-stake consensus. This proof can be the basis for the protocol's algorithmically confiscating the attacker's capital, unless the attacker is so large that they can prevent the punishment from happening before they remove their guilty stake from the system.

and (ii) there is a known finite bound on the maximum possible network delay.[32] The protocol used to prove the result has a lock-up period for stake that is long relative to the maximum possible honest network delay and constructs an in-protocol recovery from attack procedure that a sub-$\frac{2}{3}$ attacker cannot thwart. A rough intuition for how recovery is possible is that an attack requires at least $\frac{1}{3}$ of the total stake to sign conflicting transactions (per Figure IV), so a sub-$\frac{2}{3}$ attacker is thus left with $< \frac{2}{3} - \frac{1}{3} = \frac{1}{3}$ of stake that is not accused of attack, which is less than 50% of the total nonaccused stake.

A separate issue with BFT-style consensus is what are known as liveness attacks. Because $\frac{2}{3}$ of all stake has to sign each block,

32. Assumption (ii) is known as the synchronous communications assumption and is considered a strong assumption in the distributed-consensus literature (see Lewis-Pye and Roughgarden 2024). Budish, Lewis-Pye, and Roughgarden (2024) prove an impossibility result for slashing for the partially synchronous communications model of Dwork, Lynch, and Stockmeyer (1988). Budish, Lewis-Pye, and Roughgarden (2024) also prove additional impossibility results for economic security from targeted punishment for any permissionless consensus protocol in the class of fully permissionless or dynamically available protocols. The fully permissionless class includes Bitcoin proof of work and the dynamically available class includes some other forms of proof of stake based on longest-chain consensus rather than the BFT paradigm.

an attacker with greater than $\rho = \frac{1}{3}$ of the total stake can effectively halt transactions for a significant period of time.[33] If Ethereum were to become more integrated into the traditional global financial system, this would seem to be a significant source of attack risk. Unlike for a safety violation as depicted in the figure, a liveness attack does not leave any cryptographic proof of the attacker's malfeasance. The silent stake could be having a legitimate network outage or computer failure. For these reasons, Ethereum does slash stake that is silent for a long period of time, but very slowly. By my calculations, one could halt Ethereum for a day at a cost of 0.09% of the total honest stake, two days at a cost of 0.37% of the honest stake, a week at a cost of 4.68% of the honest stake, and two weeks at a cost of 20.53% of the honest stake.[34]

I conclude that Ethereum proof of stake with slashing achieves a security improvement over Bitcoin that can be understood through the lens of this article's analysis: a sub-$\frac{2}{3}$ double-spend attacker incurs costs that are $O(N^*C)$ even if $\Delta_{attack} = 0$. However, to deter $> \frac{2}{3}$ attackers or to deter liveness attacks in the event Ethereum becomes more integrated with global financial markets likely requires a source of trust support external to the protocol, such as the rule of law.[35]

33. It is possible to improve the safety of BFT-style consensus against double-spending attacks by requiring $1 - \rho_l > \frac{2}{3}$ of stake to sign each block, but this lowers the threshold at which the consensus is vulnerable to liveness attacks to $\rho_l < \frac{1}{3}$. See further discussion in Budish, Lewis-Pye, and Roughgarden (2024). This intrinsic trade-off between safety and liveness is familiar in the distributed-consensus literature more broadly (Lewis-Pye and Roughgarden 2024).

34. These calculations are based on the description of the inactivity penalty mechanism provided by Edgington (2023), which is linked to from Ethereum.org. The latest version of this mechanism was implemented in Ethereum's major update in September 2022 (Bellatrix) and has remained unchanged as of its most recent major update in April 2024 (Deneb). Note that if Ethereum were to increase its security against double-spend attacks as discussed in the previous note, it would lower the costs of liveness attacks. For example, I calculate that if Ethereum required 90% of stake to sign each block (instead of $\frac{2}{3}$), then a liveness attack could halt Ethereum for a day at a cost of 0.02% of honest stake, two days at a cost of 0.08% of honest stake, a week at a cost of 1.04% of honest stake, and two weeks at a cost of 4.56% of honest stake.

35. Ethereum's official developer documentation discusses the risk of $\frac{2}{3}$ attackers in some detail, writing, "As the supermajority stakeholder, the attacker would always control the contents of the finalized blocks, with the power to spend, rewind and spend again, censor certain transactions and reorg the chain at will." The discussion concludes, "The only real defenses here are the enormous cost of

3. *External-to-Protocol Punishment.* The third logical possibility for how an attacker could lose their capital value is by a punishment external to the blockchain protocol. For example, an attacker who engaged in the short-selling attacks discussed above might face legal consequences, which could include prison and significant financial penalties. This would certainly work in the sense of getting the security benefits of equilibrium constraint (10) rather than equation (3), but the whole question of this article is whether Nakamoto's novel form of trust works economically without government and the rule of law.

V.C. Discussion

The theoretical analysis of Section V.A shows that *if* the attacker loses their capital as a result of the attack, permissionless consensus becomes significantly more economically attractive. A 2,500-times improvement in the cost of trust is enough to transform Nakamoto's novel form of trust from being extremely expensive relative to traditional forms of trust to competitive with traditional trust. For example, referring to Table III, the cost to secure against an attack of $100 billion goes from about 2.5 times global GDP to just over $100 billion a year, which is the same order of magnitude as spending on police, courts, and prisons in the United States (see Section VI). Or, referring to Table V, if we assume an attacker can double-spend for $k = 6$ blocks of honest volume and the escrow period is $e = 6$ blocks, then the cost to secure against attack goes from a 30% tax on honest transaction volume to just 0.012% of honest transaction volume, which is cheaper than most payment fees.

The difficulty is *how* the attacker loses their capital. Collapse risk does not seem like a plausible foundation for a novel economic system, especially when one considers the possibility of sabotage. In-protocol punishment without collapse would be very attractive but faces impossibility theorems that suggest it is only a partial solution. External-to-protocol punishment certainly could work but begs the question of what is accomplished if per-

66% of the total staked ether, and the option to fall back to the social layer to coordinate adoption of an alternative fork" (Ethereum.org 2023). My understanding of the term "social layer" is that it is meant to encompass any form of trust support external to the protocol itself. My own speculation is that this would include the rule of law if large sums of money are under dispute.

missionless consensus is ultimately reliant on traditional rule of law.

VI. Comparison of Nakamoto Trust and Traditional Trust

In this section, I return to the contrast discussed in the introduction between Nakamoto trust and traditional trust supported by the rule of law and complementary sources, such as reputations and relationships. The essential difference is economies of scale.

VI.A. Beckerian Deterrence as an Economy of Scale

For concreteness, consider a financial transaction between two parties of size V, but where one of the parties has an opportunity to cheat and steal the other party's assets. Specifically, Party 1 chooses an action from the set {*Engage*, *Don't Engage*}; Party 2 chooses an action from the set {*Honest*, *Cheat*}. If the players choose *Engage* and *Honest*, then both parties get a payoff of $\tau > 0$, representing the net benefit of an honest transaction; but if Party 1 chooses *Engage* and Party 2 chooses *Cheat*, then Party 2 gets a payoff of V and Party 1 gets a payoff of $-V$, representing that Party 2 has stolen Party 1's assets. If Party 1 chooses *Don't Engage*, then both parties get a payoff of zero. Clearly, in the game as described so far, the only equilibrium is for Party 1 to choose *Don't Engage*, so the parties will forgo the benefit of transacting. (See Dixit 2004 for analysis of similar game forms.)

Now add a legal system with the power to enforce contracts. Specifically, if Party 2 plays *Cheat*, Party 1 can pay a cost c_l to adjudicate the transaction in court, and the court can perfectly observe whether Party 2 played *Cheat* or *Honest*. If the court observes that Party 2 played *Cheat* it can compel Party 2 to return Party 1's assets and punish Party 2 with a large fine f. In this scenario, Party 1's payoff is $-c_l$, the cost of bringing the matter to court, and Party 2's payoff is $-f$, the cost of the fine.

Clearly, the legal system makes it an equilibrium for the parties to transact honestly; the credible threat of a large fine deters Player 2 from cheating. Since the players will transact honestly on the equilibrium path, the court need not even involve itself with most transactions in the first place. This is the point emphasized in the introduction about the economies of scale in traditional trust that is implicit in Hayek (1960) and Becker (1968).

The key insight is: a society that pays a fixed cost of operating a court system can facilitate honest transactions that have zero marginal cost of security because of the deterrence effect. This is a scale economy for traditional trust.

We can translate this conceptual point about scale economies for traditional trust into the language of our earlier analysis. Consider the stripped-down model of Section III.E augmented in two ways. First, as in Section IV.C, add a parameter V_{honest} that represents the average volume transacted per period by honest users of the system if trust is secured. Second, model traditional trust as costing a fixed cost F plus a variable cost per unit transacted of c, such that society's cost of traditional trust is $F + cV_{honest}$ per period. In Figure I, the c can be interpreted as the cost of the security guards outside the bank and the F can be interpreted as society's cost of police and courts.

The two trust models' cost per unit volume are thus:

$$Traditional\ Trust: \frac{F}{V_{honest}} + c$$

$$Nakamoto\ Trust: \frac{V_{attack}}{V_{honest}}.$$

Traditional trust has scale economies to the extent that fixed costs F that support trust can scale over a large quantity of transaction volume V_{honest}. Beckerian deterrence of crime is a leading example. Similar scale economies of trust arise in the private sector from fixed-cost investments in brands, reputation, relationships, or collateral. Often such investments work in conjunction with societal fixed-cost investments in rule of law. For example, a firm's brand, reputation, or relationships can serve as a credible commitment to provide high quality on those dimensions of quality that are not contractible (Nelson 1974; Fudenberg, Levine, and Maskin 1994; Tadelis 1999; Baker, Gibbons, and Murphy 2002; Levin 2003), while laws or contracts can cover the dimensions that are contractible.

Collateral is a particularly important example to discuss in our context. Imagine that a bank intermediates the transaction above between Party 1 and Party 2 and has general-purpose collateral on its balance sheet that exceeds the value of the transaction. The bank can be trusted not to abscond with the parties' assets, without any appeal to reputation or brand, if a court can compel it to compensate the parties out of its general-purpose collateral if it cheats. Moreover, the cost of general-purpose collat-

eral as a source of trust support is low because collateral earns a market rate of return. Under the assumptions of the Modigliani and Miller (1958) theorem, the cost of collateral as a source of trust support is *zero*. Estimates from the empirical literature on the magnitudes of violations of the Modigliani-Miller theorem find that the cost of collateral is not literally zero but is less than 1% per year of the collateral amount, which likely translates to less than 0.01% of transaction volume (see Budish and Sunderam 2023). That is, $\frac{F}{V_{honest}} < 0.0001$ for collateral.

Nakamoto trust, by contrast, only enjoys economies of scale with transaction volume if the scope for attacking the system V_{attack} does not grow with the system's usefulness for honest participants V_{honest}, which seems unlikely without support from the rule of law. In the stylized financial transaction, the size of the attack opportunity equals the size of the honest transaction, that is, $\frac{V_{attack}}{V_{honest}} = 1$. Indeed, the ratio $\frac{V_{attack}}{V_{honest}}$ could easily exceed 1, as a majority attacker will engage in large transactions whereas V_{honest} measures the size of average transactions.

Table VI presents a summary comparison of these different forms of trust.

VI.B. Sense of Magnitudes for Finance

Total annual spending on police, prisons, and courts, at the state, local, and federal level, is about $300 billion a year in the United States (Urban Institute 2023). Real value added in the U.S. financial industry is about $800 billion a year.[36] So we could use $1 trillion a year as a conservative upper bound for the cost of trust in the U.S. financial sector, since the former figure includes spending that is unrelated to finance and the latter figure includes spending that is unrelated to trust.[37] We can use $1 quadrillion a year as a conservative lower bound for transaction

36. This figure is taken from the U.S. Bureau of Economic Analysis "Real Value Added by Industry" data, lines 56–57 ("Federal Reserve banks, credit intermediation, and related activities" and "Securities, commodity contracts, and investments"). Real value added measures payments to both capital and labor. See Philippon (2015) on why it is a useful measure for the cost of the financial sector.

37. Gennaioli, Shleifer, and Vishny (2015) distinguish between financial sector trust in the sense of security from expropriation or theft and trust in the sense of confidence to take risks. They argue that high fees and market power in the financial sector, especially as relates to investment management (see Greenwood and Scharfstein 2013), can often be intrepreted as demand for the latter kind of

TABLE VI

COMPARISON OF TRUST MODELS

	Cost to break trust	Variable cost of honest play	Fixed cost investments that support honest play
Nakamoto trust in its original form (Nakamoto 2008)	Flow. Attacker pays cost of honest trust support $p_{block} = N^* c$ for a short period of time.	p_{block} per period	—
Repeated interaction (Schelling 1956; Aumann 1959)	Stock. Attacker loses the net present value of the trust relationship $V_{trust} = \frac{1}{1-\delta}\tau$.	Zero, given preexisting trust relationship	Investment in trust relationship, brand, reputation
Credible deterrence (Hayek 1960; Becker 1968)	Arbitrarily high up to value of life and assets. Attacker can be punished by the state.	"Security guards" in Figure I. Variable costs of monitoring and enforcement associated with specific locations or transactions.	Investment in government and legal system
Collateral	Stock + punishment. Attacker loses the value of their collateral plus any additional punishment.	Zero, under conditions of Modigliani-Miller theorem. Requires legal enforcement mechanism, which entails fixed costs.	Collateral value
Nakamoto with collapse	Stock. Attacker loses specialized capital worth $N^* C$.	p_{block} per period	ASICs or other specialized computational equipment
Proof of stake in its original form	Flow. Attacker pays opportunity cost of honest capital $p_{block} = N^* c$ for a short period of time.	p_{block} per period	Cryptocurrency stake
Proof of stake with slashing (idealized)	Stock. Attacker loses specialized capital worth $N^* C$.	p_{block} per period	Cryptocurrency stake

Notes. Nakamoto trust in its original form: see analysis in Sections III.A and III.B. Repeated interaction: see analysis in Section III.F. Credible deterrence: see discussion in Section VI.A. Collateral: see discussion in Section VI.A. Nakamoto with collapse: see analysis in Section V.A and discussion in Section V.B. Proof of stake in its original form: see analysis in Sections III.A and III.B and discussion in Section II.E. Proof of stake with slashing (idealized): see analysis in Section V.A and discussion in Section V.B; "idealized" means ignoring the issues that arise if the attacker is large enough to prevent the protocol from punishing them before they withdraw their guilty stake and ignoring the risk of liveness attacks.

volume in the U.S. financial sector (Budish and Sunderam 2023). We can thus upper bound the cost of trust support $\frac{F}{V_{honest}} + c$ in traditional finance by 0.1% of transaction volume. Clearly this is just a rough ballpark. Many fees in traditional finance, especially for large transactions, are on the order of 0.01% or less of transaction volume.[38]

VII. Conclusion

Nakamoto (2008)'s novel form of trust—a completely anonymous and decentralized permissionless consensus, without any support from government or trusted intermediaries—is *ingenious but expensive*. Equation (3) says that for the trust to be meaningful requires that the flow cost of maintaining the trust must be large relative to the one-shot value of attacking it. This is like a very large implicit tax. Moreover, the cost of Nakamoto trust scales linearly with the value of the attack—for example, securing against a $1 billion attack is 1,000 times more expensive than securing against a $1 million attack. This equilibrium constraint suggests that if cryptocurrencies were to become a more significant part of the global financial system than they have been to date, then their costs would have to grow to absurd levels (absent implicit support from rule of law). In the base case analysis considered in Section IV, it would take all of global GDP to secure the

trust, confidence to take risks. The former kind of trust, security from theft, is the aspect I have in mind here as relevant for the comparison to Nakamoto trust.

38. For example, Hu, Pan, and Wang (2021) report that the median fee charged in the treasury repo market is 2 basis points (0.02%) on an annual basis, which translates to about 0.00005% per day. Interactive brokers' fees for large foreign exchange transactions are about 0.001%. Budish, Lee, and Shim (2024) find that the average exchange trading fee in the U.S. stock market is $0.0001 per share per side, or about 0.0001% on a $100 share of stock. Fees are higher for retail transactions but still often relatively small. Visa's annual operating expenses are less than 0.1% of their annual transaction volume and their revenue is about 0.25% of volume (Visa 2022). Asset management fees for hundreds of Vanguard index funds are less than 0.10% (https://investor.vanguard.com/investment-products/list/all). There are numerous other fees in finance that are much higher, but these tend not to be fees for trust (in the sense of security from theft, see footnote 37) but fees that reflect market power over consumers (Campbell 2006; Greenwood and Scharfstein 2013) or consumers' willingness to pay for financial advice (Gennaioli, Shleifer, and Vishny 2015). A famous paper of Philippon (2015) estimates the cost of the financial sector as a percentage of the value of real intermediation (as opposed to transaction volume) and finds that this is about 1.5%–2%.

system against a $40 billion attack. Traditional trust, whether from rule of law, reputations, relationships, collateral, and so on, is by no means perfect but is a bargain relative to Nakamoto (see Section VI).

A conceptual insight of this article for computer science is that the economic security of a permissionless consensus protocol should be thought of not as a 0-1 variable that simply breaks at a threshold ρ, as in the classic distributed-consensus literature (e.g., Lamport, Shostak, and Pease 1982; Dwork, Lynch, and Stockmeyer 1988), but as an incentive-compatibility constraint. This is what allows for if-then equilibrium reasoning about the plausibility of permissionless consensus at scale.

Nakamoto trust would be a lot more economically attractive if an attacker lost the stock value of their capital in addition to paying the flow cost of attack, as considered in Section V. However, this requires either that security rests on the possibility of outright collapse, or that the blockchain is ultimately reliant on an external source of trust support if there is a large-enough attacker, such as rule of law.

It bears emphasis that the article's analysis is consistent with the continued use of cryptocurrencies and blockchains for black market purposes, and more generally in use cases where users are willing to pay the high implicit costs of anonymous, decentralized trust.

This study's analysis is also consistent with the usefulness of the blockchain data structure *without* Nakamoto (2008)'s novel form of trust. This is often called distributed ledger technology or a permissioned blockchain (see footnote 2 and Section II.E). Indeed, what this article highlights is that it is exactly the aspect of Nakamoto (2008) that is so innovative relative to these kinds of distributed databases—the anonymous, decentralized trust—that is the source of its economic limits. As one specific example, it is completely consistent with the present analysis that central bank digital currencies (CBDCs) could be of high economic value. CBDCs take some technical inspiration from cryptocurrencies but are anchored in traditional trust from rule of law and the reputation of central banks, and thus do not face the scaling problem of Nakamoto trust highlighted here.

At a broader level, this article builds on the view tracing all the way back to Adam Smith that government and laws are essential ingredients for the market system. A central point this article has tried to emphasize is a fairly simple one implicit in Hayek

(1960) and Becker (1968), though I have not seen it stated explicitly in this language, which is that traditional trust supported by the rule of law enjoys economies of scale. Society pays the fixed cost of the apparatus of the rule of law, or firms pay the fixed cost of building a brand or reputation or holding collateral (each of which works in conjunction with laws), and these fixed-cost assets can provide trust over a large number of economic activities at low or zero marginal cost.

VII.A. Directions for Future Research

I highlight two directions for future research. First and most directly, is there a "solution" to this paper's critique of Nakamoto trust? Informally, is there a way to generate trust in a public data set (or specifically a cryptocurrency) that has some of the anonymity and decentralization aspects of Nakamoto (2008) while being signficantly less economically constrained by the arguments in this article?[39] Online Appendix A describes several of the responses this study has received since it first circulated in 2018. The most promising responses combine blockchain-based trust with traditional trust in some way. For example, Ethereum's new protocol algorithmically mimics the combination of collateral plus rule of law to punish sub-$\frac{2}{3}$ attackers but must rely on some external source of trust support to punish large-enough attackers. If large sums of money were in dispute, it seems likely that this external source of trust support would involve the rule of law. Another interesting response is to concede that blockchain trust is intrinsically very expensive, per my argument, but only use it for occasional large transactions with long escrow periods (Layer 1), while most transactions are conducted off-chain (Layer 2), supported by external sources of trust. The idea of Bitcoin as a "digital gold" held by traditional institutional investors also fits this paradigm. Bitcoin exchange traded funds have over $1 billion of volume per trading day, traded on traditional regulated stock exchanges and managed by institutions like Blackrock and Fidelity. An open conceptual question about these responses is what the permissionless consensus part adds given the ultimate reliance on external sources of trust support. Are these combinations of permissionless consensus and traditional sources of trust

39. Recent efforts to formalize this question include Leshno, Pass, and Shi (2023) and Budish, Lewis-Pye, and Roughgarden (2024).

superior to existing alternatives? Does permissionless consensus expand the production possibilities frontier for trust?

The second direction for research is a broader conceptual question. How should economists and computer scientists model trust that comes from a combination of technology and rule of law? More generally, how should researchers understand trust when it comes from multiple sources in the same transaction that work in complement?[40] This is often the case in practice, with trust arising from some combination of the rule of law, reputations, relationships, brands, collateral, norms, technology, and so on, often implicitly and without drawing notice. For example, a large financial institution like JP Morgan is trusted in a particular transaction because of its reputation and relationships, its balance sheet collateral, its technology, and ultimately the protection that comes from the legal system. Even the most ordinary of transactions, like buying a cup of coffee at the local coffee shop, enjoys trust from multiple sources. The consumer trusts the coffee shop to provide quality coffee because of reputational incentives and perhaps implicitly food safety laws. The coffee shop trusts the consumer's payment if cash because counterfeiting is technologically complex and illegal, and if electronic because of traditional cryptography and because the financial intermediary has reputational, relational, and legal reason to follow through. The employee trusts their employer to follow through with promised compensation because of laws and the implicit relational contract. The customer and the employee trust the other not to rob them because of laws and social norms. All this trust for a cup of coffee! These multiple layers of trust that work together for even the most ordinary of economic transactions are likely part of what makes the traditional market system so robust and, dare I say, beautiful.

UNIVERSITY OF CHICAGO BOOTH SCHOOL OF BUSINESS, UNITED STATES

40. One simple preliminary exploration of this question is in Budish and Sunderam (2023), section 4.1 who conceptualize trust as getting to cooperate-cooperate in a prisoner's dilemma (as discussed in La Porta et al. 1997), and model (i) technology as eliminating some actions from the possibility set, (ii) law as changing the payoffs to some actions via punishment, and (iii) reputation as the differential incentive to cooperate if play is repeated versus one shot (as in the traditional folk theorem arguments of Aumann 1959; Fudenberg, Levine, and Maskin 1994; and others).

Supplementary Material

An Online Appendix for this article can be found at *The Quarterly Journal of Economics* online.

REFERENCES

Auer, Raphael, "Beyond the Doomsday Economics of 'Proof-of-Work' in Cryptocurrencies," Globalization Institute Working Paper No. 355, 2019. https://doi.org/10.24149/gwp355.

Aumann, R. J., "Acceptable Points in General Cooperative n-Person Games," in *Contributions to the Theory of Games IV, Annals of Mathematics Study 40*, (Princeton, NJ: Princeton University Press, 1959), 287–324.

Baker, George, Robert Gibbons, and Kevin J. Murphy, "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics*, 117 (2002), 39–84. https://doi.org/10.1162/003355302753399445.

Bakos, Yannis, and Hanna Halaburda, "Permissioned vs Permissionless Blockchain Platforms: Tradeoffs in Trust and Performance," NYU Stern School of Business Working Paper, 2023. https://ssrn.com/abstract=3789425.

Bayer, Dave, Stuart Haber, and W. Scott Stornetta, "Improving the Efficiency and Reliability of Digital Time-Stamping," in *Sequences II: Methods in Communication, Security and Computer Science*, (Berlin: Springer, 1993), 329–334.

Becker, Gary S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 76 (1968), 169–217. https://doi.org/10.1086/259394.

Biais, Bruno, Christophe Bisière, Matthieu Bouvard, and Catherine Casamatta, "The Blockchain Folk Theorem," *Review of Financial Studies*, 32 (2019), 1662–1715. https://doi.org/10.1093/rfs/hhy095.

Bitcoin Magazine, "Peter Thiel—Bitcoin Keynote—Bitcoin 2022 Conference," YouTube, April 7, 2022. https://www.youtube.com/watch?v=ko6K82pXcPA.

Bitcoin Wiki, "Weaknesses → Probably Not a Problem → Attacker Has a Lot of Computing Power," Bitcoin Wiki, 2020. https://en.bitcoin.it/wiki/Weaknesses#Attacker_has_a_lot_of_computing_power.

Böhme, Rainer, Nicolas Christin, Benjamin Edelman, and Tyler Moore, "Bitcoin: Economics, Technology, and Governance," *Journal of Economic Perspectives*, 29 (2015), 213–238. https://doi.org/10.1257/jep.29.2.213.

Bonneau, Joseph, "Why Buy When You Can Rent? Bribery Attacks on Bitcoin Consensus," in *Proceedings of the 20th International Conference on Financial Cryptography and Data Security (FC 2016)*, (Berlin: Springer, 2016), 19–26.

Budish, Eric, "The Economic Limits of Bitcoin and the Blockchain," NBER Working Paper no. 24717, 2018. https://doi.org/10.3386/w24717.

Budish, Eric, Robin S. Lee, and John J. Shim, A Theory of Stock Exchange Competition and Innovation: Will the Market Fix the Market?," *Journal of Political Economy*, 132 (2024), 1209–1246. https://doi.org/10.1086/727284.

Budish, Eric, Andrew Lewis-Pye, and Tim Roughgarden, "The Economic Limits of Permissionless Consensus," in *Proceedings of the 25th ACM Conference on Economics and Computation (EC'24)*, (2024). https://arxiv.org/abs/2405.09173.

Budish, Eric, and Adi Sunderam, "Blockchain Technology and Stablecoins in Traditional Finance," in *Sveriges Riksbank 7th Annual Macroprudential Conference*(2023). https://arxiv.org/abs/2405.09173.

Buterin, Vitalik, Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform," Ethereum Foundation White Paper, 2014a. https://ethereum.org/en/whitepaper/.

———, "Slasher: A Punitive Proof-of-Stake Algorithm," Ethereum, 2014b. https://blog.ethereum.org/2014/01/15/slasher-a-punitive-proof-of-stake-algorithm/.

———, "A Proof of Stake Design Philosophy," *Medium*, December 30, 2016. https://medium.com/@VitalikButerin/a-proof-of-stake-design-philosophy-506585978d51.

———, "Minimal Slashing Conditions," *Medium*, March 2, 2017. https://medium.com/@VitalikButerin/minimal-slashing-conditions-20f0b500fc6c.

———, "What in the Ethereum Application Ecosystem Excites Me," Vitalik (blog), December 5, 2022. https://vitalik.eth.limo/general/2022/12/05/excited.html.

Buterin, Vitalik, and Virgil Griffith, "Casper the Friendly Finality Gadget," arXiv working paper 1710.09437, 2019. https://doi.org/10.48550/arXiv.1710.09437.

Campbell, John Y., "Household Finance," *Journal of Finance*, 61 (2006), 1553–1604. https://doi.org/10.1111/j.1540-6261.2006.00883.x.

Carlsten, Miles, Harry Kalodner, S. Matthew Weinberg, and Arvind Narayanan, "On the Instability of Bitcoin Without the Block Reward"," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, (2016), 154–167. https://doi.org/10.1145/2976749.2978408.

Chiu, Jonathan, and Thorsten V. Koeppl, "The Economics of Cryptocurrency: Bitcoin and Beyond," *Canadian Journal of Economics*, 55 (2022), 1762–1798. https://doi.org/10.1111/caje.12625.

Cong, Lin William, Zhiguo He, and Jiasun Li, "Decentralized Mining in Centralized Pools," *Review of Financial Studies*, 34 (2021), 1191–1235. https://doi.org/10.1093/rfs/hhaa040.

Cox, Jeff, "Yellen Sounds Warning about 'Extremely Inefficient' Bitcoin," *CNBC*, February 22, 2021. https://www.cnbc.com/2021/02/22/yellen-sounds-warning-about-extremely-inefficient-bitcoin.html.

De Vries, Alex, "Bitcoin's Growing Energy Problem," *Joule*, 2 (2018), 801–805. https://doi.org/10.1016/j.joule.2018.04.016.

Digiconomist, "Bitcoin Energy Consumption Index," *Digiconomist*, 2022. https://digiconomist.net/bitcoin-energy-consumption.

Dixit, Avinash K., *Lawlessness and Economics: Alternative Modes of Governance*, (Princeton, NJ: Princeton University Press, 2004).

Dolev, D., and H. R. Strong, "Authenticated Algorithms for Byzantine Agreement," *SIAM Journal on Computing*, 12 (1983), 656–666. https://doi.org/10.1137/0212045.

Dwork, Cynthia, Nancy Lynch, and Larry Stockmeyer, "Consensus in the Presence of Partial Synchrony," *Journal of the ACM*, 35 (1988), 288–323. https://doi.org/10.1145/42282.42283.

Easley, David, Maureen O'Hara, and Soumya Basu, "From Mining to Markets: The Evolution of Bitcoin Transaction Fees," *Journal of Financial Economics*, 134 (2019), 91–109. https://doi.org/10.1016/j.jfineco.2019.03.004.

Edgington, Ben, "Inactivity Leak," 2023. https://eth2book.info/capella/part2/incentives/inactivity/.

Ethereum.org, "Ethereum Proof-of-Stake Attack and Defense," August 15, 2023. https://ethereum.org/en/developers/docs/consensus-mechanisms/pos/attack-and-defense/.

Eyal, Ittay, and Emin Gun Sirer, "Majority Is not Enough: Bitcoin Mining Is Vulnerable," in *Proceedings of the 18th International Conference on Financial Cryptography and Data Security (FC 2014)*, (Berlin: Springer, 2014), 436–454.

Fischer, Michael J., Nancy A. Lynch, and Michael S. Paterson, "Impossibility of Distributed Consensus with One Faulty Process," *Journal of the ACM*, 32 (1985), 374–382. https://doi.org/10.1145/3149.214121.

Foley, Sean, Jonathan R. Karlsen, and Tālis J. Putniņš, "Sex, Drugs, and Bitcoin: How Much Illegal Activity Is Financed through Cryptocurrencies?," *Review of Financial Studies*, 32 (2019), 1798–1853. https://doi.org/10.1093/rfs/hhz015.

Friedman, Milton, *Capitalism and Freedom*, (Chicago: University of Chicago Press, 1962).

Fudenberg, Drew, David Levine, and Eric Maskin, "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62 (1994), 997–1039. https://doi.org/10.2307/2951505.

Gans, Joshua S., and Hanna Halaburda, "'Zero Cost' Majority Attacks on Permissionless Blockchains," NBER Working Paper no. 31473, 2023. https://doi.org/10.3386/w31473.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny, "Money Doctors," *Journal of Finance*, 70 (2015), 91–114. https://doi.org/10.1111/jofi.12188.
Gensler, Gary, "Remarks Before the Aspen Security Forum," SEC, 2021. https://www.sec.gov/news/public-statement/gensler-aspen-security-forum-2021-08-03.
Goldman Sachs, "Blockchain—The New Technology of Trust," 2018.
Greenwood, Robin, and David Scharfstein, "The Growth of Finance," *Journal of Economic Perspectives*, 27 (2013), 3–28. https://doi.org/10.1257/jep.27.2.3.
Grunspan, Cyril, and Ricardo Pérez-Marco, "Double Spend Races," *International Journal of Theoretical and Applied Finance*, 21 (2018), 1850053. https://doi.org/10.1142/S021902491850053X.
———, "On Profitability of Nakamoto Double Spend," *Probability in the Engineering and Informational Sciences*, 36 (2022), 732–746. https://doi.org/10.1017/S026996482100005X.
Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "Does Culture Affect Economic Outcomes?," *Journal of Economic Perspectives*, 20 (2006), 23–48. https://doi.org/10.1257/jep.20.2.23.
Haber, Stuart, and W. Scott Stornetta, "How to Time-Stamp a Digital Document," *Journal of Cryptography*, 3 (1991), 99–111. https://doi.org/10.1007/3-540-38424-3_32.
Halaburda, Hanna, Guillaume Haeringer, Joshua S. Gans, and Neil Gandal, "The Microeconomics of Cryptocurrencies," *Journal of Economic Literature*, 60 (2022), 971–1013. https://doi.org/10.1257/jel.20201593.
Hart, Oliver, *Firms, Contracts, and Financial Structure*, (Oxford: Clarendon Press, 1995). https://doi.org/10.1093/0198288816.001.0001.
Hayek, Friedrich A., *The Constitution of Liberty*, (Chicago: University of Chicago Press, 1960).
Holmstrom, Bengt, and Paul Milgrom, "The Firm as an Incentive System," *American Economic Review*, 84 (1994), 972–991.
Hu, Grace Xing, Jun Pan, and Jiang Wang, "Tri-Party Repo Pricing," *Journal of Financial and Quantitative Analysis*, 56 (2021), 337–371. https://doi.org/10.1017/S0022109019000863.
Huberman, Gur, Jacob D. Leshno, and Ciamac Moallemi, "Monopoly without a Monopolist: An Economic Analysis of the Bitcoin Payment System," *Review of Economic Studies*, 88 (2021), 3011–3040. https://doi.org/10.1093/restud/rdab014.
Kandori, Michihiro, "Social Norms and Community Enforcement," *Review of Economic Studies*, 59 (1992), 63–80. https://doi.org/10.2307/2297925.
Klein, Ezra, "Ethereum's Founder on What Crypto Can—and Can't—Do," *New York Times* podcast, 2022. https://www.nytimes.com/2022/09/30/podcasts/transcript-ezra-klein-interviews-vitalik-buterin.html.
Kreps, David M., Paul Milgrom, John Roberts, and Robert Wilson, "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory*, 27 (1982), 245–252. https://doi.org/10.1016/0022-0531(82)90029-1.
Kroll, Joshua A., Ian C. Davey, and Edward W. Felten, "The Economics of Bitcoin Mining or, Bitcoin in the Presence of Adversaries," in *12th Workshop on the Economics of Information Security*, (2013). https://econinfosec.org/archive/weis2013/papers/KrollDaveyFeltenWEIS2013.pdf.
Krugman, Paul, "Baby-Sitting the Economy," *Slate*, August 14, 1998. https://slate.com/business/1998/08/baby-sitting-the-economy.html.
La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert W. Vishny, "Trust in Large Organizations," *American Economic Review Papers and Proceedings*, 87 (1997), 333–338.
———, " Law and Finance," *Journal of Political Economy*, 106 (1998), 1113–1155. https://doi.org/10.1086/250042.
Lamport, Leslie, Robert Shostak, and Marshall Pease, "The Byzantine Generals Problem," *ACM Transactions on Programming Languages and Systems*, 4 (1982), 382–401. https://doi.org/10.1145/357172.357176.

Leshno, Jacob D., Rafael Pass, and Elaine Shi, "On the Viability of Open-Source Financial Rails: Economic Security of Permissionless Consensus," Cryptology ePrint Archive Working Paper 2023/1516, 2023. https://eprint.iacr.org/2023/1516.
Levin, Jonathan, "Relational Incentive Contracts," *American Economic Review*, 93 (2003), 835–857. https://doi.org/10.1257/000282803322157115.
Levine, Matt, "Bank Blockchains and an Alibaba Box," Bloomberg View, January 10, 2017. https://www.bloomberg.com/view/articles/2017-01-10/bank-blockchains-and-an-alibaba-box.
Lewis-Pye, Andrew, and Tim Roughgarden, Permissionless Consensus," arXiv working paper 2304.14701, 2024. https://doi.org/10.48550/arXiv.2304.14701.
Ma, June, Joshua S. Gans, and Rabee Tourky, "Market Structure in Bitcoin Mining," Rotman School of Management Working Paper no. 3103104, 2019. https://dx.doi.org/10.2139/ssrn.3103104.
Makarov, Igor, and Antoinette Schoar, "Blockchain Analysis of the Bitcoin Market," NBER Working Paper no. 29396, 2021. https://doi.org/10.3386/w29396.
Modigliani, Franco, and Merton H. Miller, "The Cost of Capital, Corporation Finance and the Theory of Investment," *American Economic Review*, 48 (1958), 261–297.
Moroz, Daniel J., Daniel J. Aronoff, Neha Narula, and David C. Parkes, "Double-Spend Counterattacks: Threat of Retaliation in Proof-of-Work Systems," arXiv working paper 2002.10736, 2020. https://doi.org/10.48550/arXiv.2002.10736.
Nakamoto, Satoshi, "Bitcoin: A Peer-to-Peer Electronic Cash System," Bitcoin.org White Paper, 2008.
Nelson, Phillip, "Advertising as Information," *Journal of Political Economy*, 82 (1974), 729–754. https://doi.org/10.1086/260231.
Pease, M., R. Shostak, and L. Lamport, "Reaching Agreement in the Presence of Faults," *Journal of the ACM*, 27 (1980), 228–234. https://doi.org/10.1145/322186.322188.
Philippon, Thomas, "Has the US Finance Industry Become Less Efficient? On the Theory and Measurement of Financial Intermediation," *American Economic Review*, 105 (2015), 1408–1438. https://doi.org/10.1257/aer.20120578.
Popper, Nathaniel, *Digital Gold: Bitcoin and the Inside Story of the Misfits and Millionaires Trying to Reinvent Money*, (New York: Harper, 2015).
Prat, Julien, and Benjamin Walter, "An Equilibrium Model of the Market for Bitcoin Mining," *Journal of Political Economy*, 129 (2021), 2415–2452. https://doi.org/10.1086/714445.
Rogoff, Kenneth, *The Curse of Cash*, (Princeton, NJ: Princeton University Press, 2017).
Rosenfeld, Meni, "Analysis of Hashrate-Based Double-Spending," arXiv working paper 1402.2009, 2014. https://doi.org/10.48550/arXiv.1402.2009.
Roughgarden, Tim, "Online Course: Foundations of Blockchains," online course, 2023. https://www.youtube.com/playlist?list=PLEGCF-WLh2RLOHv_xUGLqRts_9JxrckiA.
Saleh, Fahad, "Blockchain without Waste: Proof-of-Stake," *Review of Financial Studies*, 34 (2021), 1156–1190. https://doi.org/10.1093/rfs/hhaa075.
Schelling, Thomas C., "An Essay on Bargaining," *American Economic Review*, 46 (1956), 281–306.
———, *The Strategy of Conflict*, (Cambridge, MA: Harvard University Press, 1960).
Shleifer, Andrei, and Robert W. Vishny, "A Survey of Corporate Governance," *Journal of Finance*, 52 (1997), 737–783. https://doi.org/10.1111/j.1540-6261.1997.tb04820.x.
Smith, Adam, *The Wealth of Nations*, (New York: Penguin Classics, 1982).
Tabarrok, Alex, "Bitcoin Is Less Secure than Most People Think," Marginal Revolution, 2019. https://marginalrevolution.com/marginalrevolution/2019/01/bitcoin-much-less-secure-people-think.html.

Tadelis, Steven, "What's in a Name? Reputation as a Tradeable Asset," *American Economic Review*, 89 (1999), 548–563. https://doi.org/10.1257/aer.89.3.548.
Tas, Ertem Nusret, David Tse, Fangyu Gai, Sreeram Kannan, Mohammad Ali Maddah-Ali, and Fisher Yu, "Bitcoin-Enhanced Proof-of-Stake Security: Possibilities and Impossibilities," arXiv working paper 2207.08392, 2023. https://doi.org/10.48550/arXiv.2207.08392.
Urban Institute, "Criminal Justice Expenditures: Police, Corrections, and Courts," 2023. https://www.urban.org/policy-centers/cross-center-initiatives/state-and-local-finance-initiative/state-and-local-backgrounders/criminal-justice-police-corrections-courts-expenditures.
Visa, "Visa Inc. Fiscal 2022 Annual Report," 2022. https://s29.q4cdn.com/385744025/files/doc_downloads/2022/Visa-Inc-Fiscal-2022-Annual-Report.pdf.
Wolitzky, Alexander, "Cooperation in Large Societies," in *Advances in Economics and Econometrics: Twelfth World Congress* (Cambridge U.K.: Cambridge University Press, forthcoming).