Systemic Risk by Design?
Causal Evidence on Endogenous Blockchain Security

Pablo D. Azar* Maryam Farboodi†

September 5, 2025

Abstract

Can blockchain incentives replace the role of courts and governments in guaranteeing the finality of large transactions? We test this question using Proof-of-Work blockchains, asking whether their "market for security" falters when economic stakes are highest. We provide the first causal evidence that it does. Using major crypto hacks and network crises as plausibly exogenous instruments for transaction fees—the prize for an attack—our IV estimates show that a 10% increase in USD-denominated fees causally increases the rate of likely-strategic transaction reversals by 12.5%. These findings provide the first empirical confirmation of the central economic critique of PoW, showing its security degrades under stress. This reveals a novel channel for systemic risk in decentralized finance, contrasting sharply with traditional financial infrastructures.

1 Introduction

Market systems have long been understood to require the support of a government and its rule of law to enforce contracts. This legal framework provides the ultimate guarantee for the finality of settlement, ensuring that once a high-value transaction is complete, it is irreversible. Proof-of-Work (PoW) blockchains represent a radical departure from this model, introducing a new economic system that generates trust from a combination of cryptography and economic incentives—a mechanism designed to align private incentives with the public goal of maintaining a secure ledger. This creates a novel "market for security," in which the rational, self-interested actions of decentralized miners are meant to produce an honest and canonical history of transactions. While technologically ingenious, the economic properties of this new form of trust—and its viability as a foundation for a modern financial system—remain largely untested.

This paper tests whether Proof-of-Work's "market for security" can fulfill its promise to replace traditional legal enforcement by asking a fundamental question about its design: is the enforcement of finality in Proof-of-Work robust when the economic stakes are highest? The mechanism's central tension is that high-value transactions create two opposing economic forces. They can fund

*Federal Reserve Bank of New York; pablo.azar@ny.frb.org. The views expressed here are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

†MIT Sloan School of Management; farboodi@mit.edu.

larger rewards for honest miners, attracting more computational power to secure the network, but they also create a more lucrative prize for a rational attacker who could profit by subverting the consensus and reversing those same transactions. We provide the first causal evidence that the mechanism fails this stress test; its guarantee of finality weakens precisely when the transactions it secures are most valuable, revealing a deep economic limit to this form of decentralized trust.

Using on-chain data, we provide the first causal evidence that the incentives to deviate from honesty in Proof-of-Work increase with the value of transactions. We exploit a unique feature of the Ethereum blockchain's protocol during its seven-year Proof-of-Work era that makes these settlement reversals directly observable. The ledger publicly recorded valid blocks that had won the mining race but were nevertheless not appended to the canonical chain—providing a high-frequency, on-chain measure of security degradation. We term these events stolen blocks and construct an hourly panel of their frequency. Our identification strategy leverages plausibly exogenous shocks to transaction demand—such as major crypto ecosystem hacks and network-wide crises—as instruments for transaction fees. This approach allows us to cleanly identify the causal channel, confirming that the mechanism's security degrades under economic stress.

Our instrumental variable estimates establish a direct causal link: higher economic stakes actively undermine settlement finality. We find that a 10% increase in USD-denominated transaction fees causally increases the hourly rate of stolen blocks by 1.07%. This relationship is particularly pronounced for settlement reversals that are more likely to be strategic rather than accidental; for these events, the elasticity is an order of magnitude larger, with a 10% increase in fees yielding a 12.5% increase in likely-strategic security failures. During a network-wide crisis, the resulting surge in transaction fees precipitates an estimated 2.9% degradation in settlement finality. This quantitative evidence supports our central argument: the economic incentives designed to secure the chain can, under stress, actively subvert it, weakening the guarantee of finality precisely when market participants value certainty the most.

Our paper makes a direct empirical contribution to the mostly theoretical literature on the economics of blockchain incentives. Specifically, we provide the first causal test of the central economic critique that their security is inherently expensive because the "flow" cost of maintaining the system must be large relative to the one-shot value of attacking it (Budish, 2025). While a robust literature has modeled the game-theoretic properties of protocols (Biais et al., 2019; Abadi and Brunnermeier, 2018), identified specific vulnerabilities (Eyal and Sirer, 2014), and analyzed specific on-chain markets (Huberman et al., 2021), the fundamental economic security model has not been subjected to a causal empirical test. Our finding that security degrades as transaction value increases provides direct causal evidence for the equilibrium constraints formalized in Budish (2025), confirming the theory's prediction that the system's security is weakest when the economic stakes are highest, and providing context for the ecosystem's search for alternatives like Proof-of-Stake (Saleh, 2021).

This empirical confirmation of PoW's design flaw has sharp implications for both the design of financial market infrastructure and our understanding of the economic role of law. The risk we identify—that security degrades as value increases—runs directly counter to the core principles of traditional payment systems, which are designed to enhance stability and prevent collapse during periods of stress (Rochet and Tirole, 1996; Mankiw, 1986; Duffie, 2019; Koeppl and Monnet, 2010), and introduces a novel channel for systemic risk (Adrian and Brunnermeier, 2016). Ultimately, our findings serve as an empirical testament to the long-standing view that legal institutions are crucial for enforcing contracts and enabling trust (La Porta et al., 1998; Dixit, 2004). While

decentralized systems can be remarkably effective at aggregating information (Hayek, 1945), our results question whether they can provide decentralized *enforcement*, especially for the kind of security from theft (Gennaioli et al., 2015) that underpins traditional finance, as opposed to relying on state-backed deterrence (Becker, 1968).

This debate over the institutional foundations of trust is not merely academic. The recent, rapid integration of blockchain-based assets into traditional finance—through products like Bitcoin and Ethereum ETFs, the growth of tokenized money market funds, and legislative initiatives like the GENIUS Act—makes our findings an urgent matter for financial stability. The core insight of institutional economics is that the "rules of the game" are a primary determinant of economic outcomes (North, 1990; Coase, 1960). By importing a new set of technological rules to perform a function traditionally reserved for the legal system, the financial system is unwittingly adopting the incentive flaw we identify. Understanding the true, incentive-based properties of these new digital institutions is therefore a first-order concern, and it is this understanding that our paper provides.

2 Institutional Details: The Economics of Selfish Mining

The incentive structure of Proof-of-Work (PoW) mining is the central institutional feature that enables our analysis; while mining rewards are typically modest, the distribution of these rewards has a thick right tail, creating rare, lottery-like prizes that can incentivize strategic deviations from the protocol. In a PoW system, miners engage in a continuous, winner-take-all contest to solve a computational puzzle, with the victor earning the right to append a new block of transactions to the chain and claim a reward. This economic expenditure on computation is the fundamental source of the ledger's security, as reversing a recently confirmed block requires a costly computational attack. For a more detailed description of the mechanics of PoW blockchains, see Appendix A.

This reward structure, however, creates an incentive for miners to engage in *selfish mining*. A rational miner who discovers a new block can strategically withhold it from the network to gain a head start in the race to find the *next* block. If successful, this strategy allows the miner to earn a disproportionate share of total rewards by creating a secret, longer chain that ultimately orphans the blocks found by honest competitors. The theoretical possibility of such a strategy and its implications for blockchain security have been established in the computer science and economics literature (Eyal and Sirer, 2014; Budish, 2025). Our paper provides the first causal test of how the economic incentive to engage in this behavior varies under stress.

While the incentive to selfishly mine is a general feature of PoW, our analysis focuses on Ethereum because its protocol provides a unique empirical laboratory for observing the outcomes of these strategic contests. Most PoW blockchains, including Bitcoin, are designed to discard blocks that lose the mining race, leaving no public record of a competing chain. Ethereum's PoW consensus mechanism, however, was explicitly designed to address this issue by retaining a record of certain recently orphaned blocks, which it termed "uncles." The stated motivation for this design was to improve network security and promote decentralization; by rewarding miners for this otherwise wasted work, the protocol reduced the inherent advantage of large, well-connected mining pools and allowed for shorter block times without the same loss of security (**?**). This feature provides an exceptional, on-chain window into the resolution of the temporary forks that arise from selfish mining attempts, giving us a direct measure of settlement reversals.

We leverage this unique observability to construct our primary dependent variable. In our

framework, the "theft" is the act of a later block supplanting an earlier one. We therefore define the event based on the block that makes it into the canonical chain.

Definition 1. *A **stolen block** is a canonical block whose timestamp is later than that of one of its included uncle blocks. We attribute the "theft" to this canonical block, as it has successfully displaced an earlier, valid block from the main chain.*

This timing indicates that a valid, earlier block was supplanted by a later one, forcing a reversal of the transactions it contained—the blockchain analogue of a settlement failure. To isolate events that are more likely to be strategic overrides rather than accidental byproducts of network latency, we also construct stricter versions of the variable that require the canonical block's timestamp to be at least 10, 30, or 60 seconds after that of its stolen uncle.

The validity of this measure is strengthened by the protocol's own incentive structure. A miner who found a canonical block had a small financial incentive to include recent uncles to earn an inclusion reward. Crucially, because the reporting of timestamps did not affect these payoffs, there was no incentive for miners to misrepresent the timing of events. This combination of direct on-chain observability and incentive alignment provides the empirical leverage for our analysis. The same economic incentives that drive selfish mining on Ethereum are present, and are arguably even stronger, on the Bitcoin network. The critical institutional feature that creates this incentive—the thick-tailed nature of block rewards—is a shared property of both systems. As shown in Figure 2, the distribution of daily rewards for both blockchains is highly non-normal and right-skewed. Statistical tests confirm this visual evidence decisively. The kurtosis of daily rewards in USD, a measure of a distribution's tail thickness, is 40.2 for Ethereum and an even more extreme 300.9 for Bitcoin, far exceeding the value of 3 expected under a normal distribution. For all reward series in both native and USD terms, the Jarque-Bera test for normality is strongly rejected ($p \leq 0.001$). This shared feature suggests that while we cannot directly observe the outcomes of selfish mining on the Bitcoin blockchain, the underlying economic mechanism is the same, implying our findings for Ethereum likely apply to the entire PoW ecosystem.

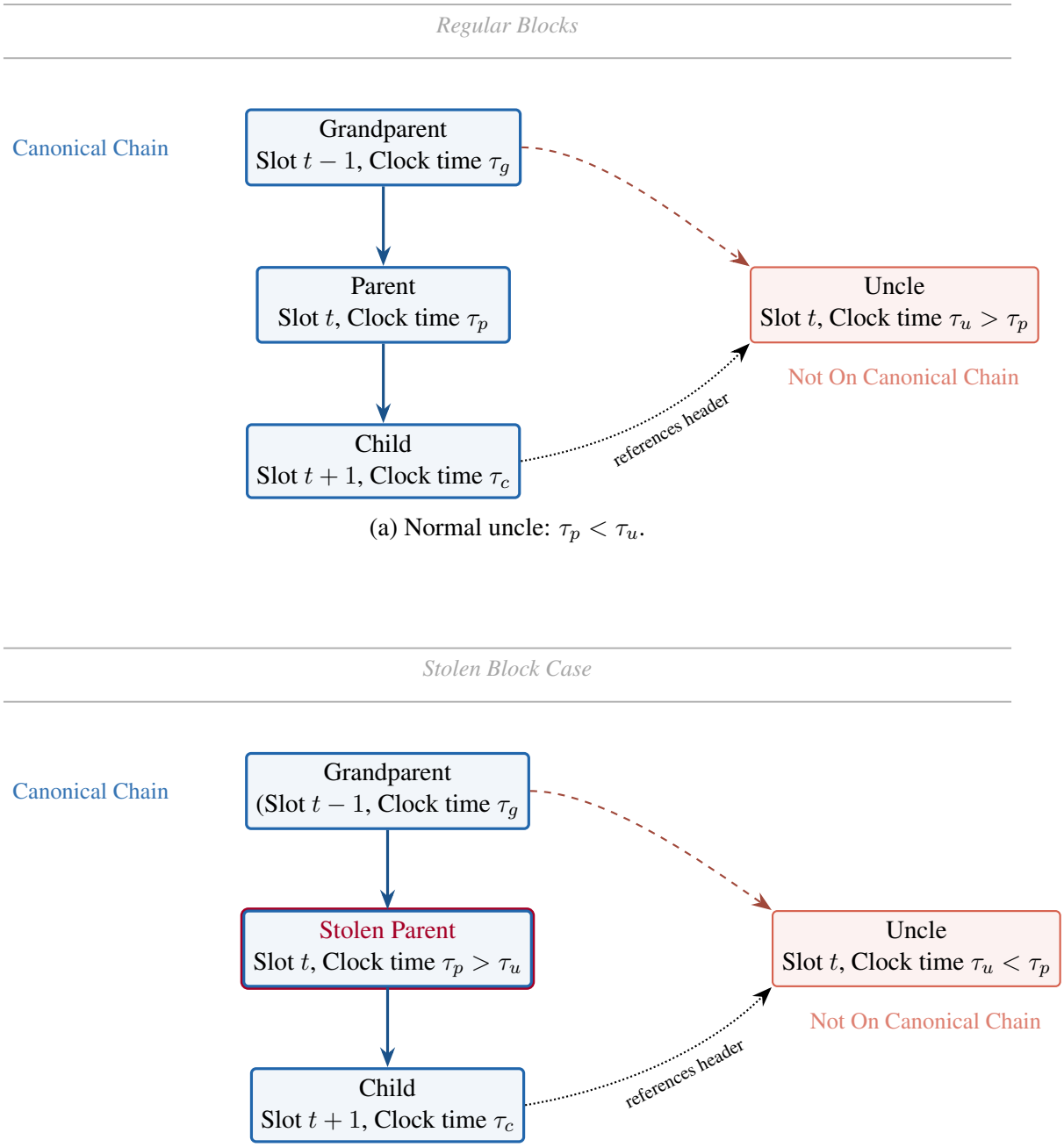
3 Empirical Design

To test the central hypothesis of our paper—that the security of the Proof-of-Work mechanism degrades as the economic value it secures increases—we require a credible source of variation in the incentives for miners to attack the chain. This section details our empirical strategy. We first describe the data and key variables before turning to our identification strategy, which uses plausibly exogenous shocks to transaction demand as an instrument for miner fees.

3.1 Data, Sample, and Key Variables

Our analysis is based on a novel hourly panel dataset constructed from on-chain data for the Ethereum blockchain, spanning its entire Proof-of-Work history from July 2015 to its transition to Proof-of-Stake in September 2022, for a total of $62,298$ hourly observations.

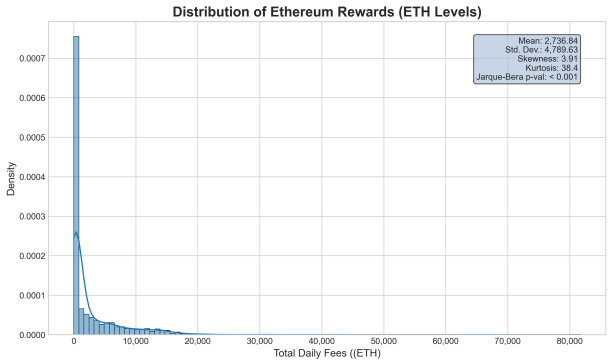
Dependent Variable: Finality Failures. Our primary dependent variable is the number of stolen blocks—a direct, on-chain measure of settlement reversals. As detailed in Section 2, Ethereum's



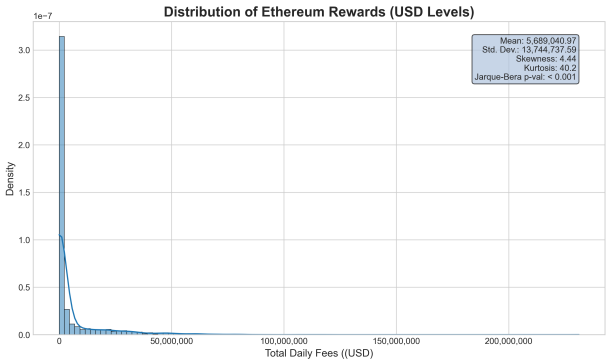
(a) Normal uncle: $\tau_p < \tau_u$.

(b) "Stolen" case: $\tau_p > \tau_u$; the canonical parent (amber outline) arrives later.

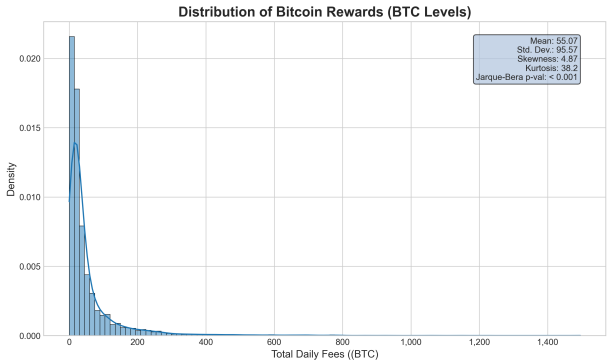
Figure 1: Ethereum uncles and "stolen" blocks. Blue nodes form the canonical chain; the rust node is an *uncle* created at the same height as the canonical parent. Solid arrows denote parent–child links; the dashed arc is the competing fork; the dotted arc indicates the nephew's inclusion of the uncle header. Panel (a) shows the usual ordering $\tau_p < \tau_u$. Panel (b) highlights a *stolen* case with an amber outline, where the canonical parent arrives later ($\tau_p > \tau_u$).



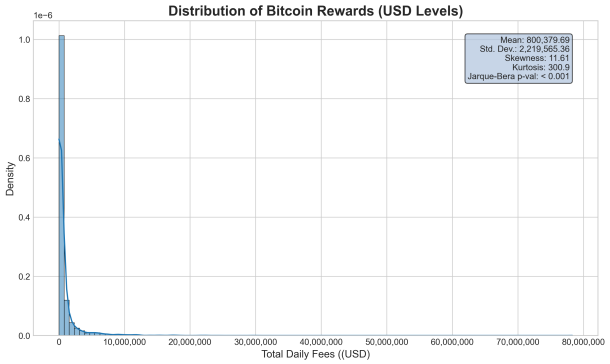
(a) Ethereum Rewards (ETH)



(b) Ethereum Rewards (USD)



(c) Bitcoin Rewards (BTC)



(d) Bitcoin Rewards (USD)

Figure 2: The top panel shows distributions for Ethereum daily rewards in native (ETH) and USD terms. The bottom panel shows the same for Bitcoin. All distributions exhibit significant right-skewness and kurtosis far exceeding that of a normal distribution, indicating thick tails.

protocol publicly recorded certain orphaned blocks ("uncles"), allowing us to observe when a block that was mined and propagated first was subsequently displaced from the canonical chain by a block mined later. We define the hourly count of stolen blocks, S_t, as the total number of canonical blocks in hour t whose timestamp is later than that of one of their included uncle blocks. Our main specifications use the outcome variable $\log(S_t)$. As shown in Table 1, finality failures are a frequent feature of the system, with a mean of 19.3 stolen blocks per hour, and the distribution is right-skewed, characterized by periods of calm punctuated by intense spikes.

To distinguish potentially strategic attacks from reversals caused by network latency, we also construct a stricter measure, counting only stolen blocks where the timestamp difference is at least 10 seconds. This stricter definition isolates events that are more likely to be the result of deliberate, strategic withholding of blocks by miners.

Independent Variable: The Economic Incentive to Attack. The core independent variable is the economic prize available to a potential attacker. We measure this as the total value of transaction fees paid to miners in a given hour, Fees_t. We analyze fees in both their native denomination (ETH) and in U.S. dollars (USD), using the $\log$ of the hourly total in our regressions. Table 1 shows that, like our dependent variable, fees are extremely heavy-tailed.

Table 1: Summary Statistics

	Count"	Mean	SD	Min	P10	P25	Median	P75	P90	Max	Skewness
Panel A: Dependent Variables											
Stolen Blocks (Count)	62,298	19.31	11.74	0.00	9.00	12.00	16.00	21.00	37.00	106.00	1.93
Log Stolen Blocks	62,296	2.82	0.51	0.00	2.20	2.48	2.77	3.04	3.61	4.66	0.44
Stolen ($\geq$10s, Count)	62,298	14.94	10.97	0.00	6.00	8.00	12.00	17.00	31.00	102.00	2.00
Log Stolen Blocks ($\geq$ 10s)	62,282	2.49	0.64	0.00	1.79	2.08	2.48	2.83	3.43	4.62	0.16
Panel B: Independent Variables											
Hourly Fees (ETH)	62,298	113.96	284.06	0.01	1.00	6.48	20.12	124.98	368.29	31345.19	39.31
Log Hourly Fees (ETH)	62,298	3.06	2.13	-4.35	0.00	1.87	3.00	4.83	5.91	10.35	-0.22
Hourly Fees (USD)	62,253	236990.23	785834.67	0.02	8.41	668.67	5150.20	116952.83	813056.60	8.57e+07	38.00
Log Hourly Fees (USD)	62,253	8.27	4.28	-4.01	2.13	6.51	8.55	11.67	13.61	18.27	-0.56
Panel C: Instruments											
Hack or Crisis Dummy	62,298	0.02	0.15	0.00	0.00	0.00	0.00	0.00	0.00	1.00	6.40
Hack Event (0/1)	62,298	0.02	0.14	0.00	0.00	0.00	0.00	0.00	0.00	1.00	6.78
Crisis Event (0/1)	62,298	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	1.00	20.73

The table reports summary statistics for the main variables used in the analysis. The data consist of an hourly panel for the Ethereum blockchain spanning its Proof-of-Work era from July 2015 to September 2022, for a total of 62,298 hourly observations. Panel A describes the dependent variables: the hourly count of stolen blocks, which measures settlement reversals, and its stricter variant counting only reversals with a time difference of at least 10 seconds, along with their logarithmic transformations. Panel B describes the main independent variables: total hourly transaction fees in both native currency (ETH) and U.S. dollars (USD), in levels and logs. Panel C describes the instrumental variables: indicator variables for days with major hacks or network-wide crises

Instrumental Variable: Exogenous Shocks to Transaction Demand. To identify a causal relationship, we require a source of variation in fees that is exogenous to the unobserved technological factors and strategic considerations that might also affect the rate of stolen blocks. We construct an instrument based on major, publicly reported crypto ecosystem hacks and network-wide crisis events. Our instrument, $Event_t$, is an indicator variable equal to one for all hours on the day of such an event. These events are rare, occurring in approximately 2% of our sample hours.

3.2 Identification Strategy: Isolating the Causal Effect of Fees

Estimating the causal effect of transaction fees on the rate of stolen blocks presents a significant endogeneity challenge. A simple OLS regression of stolen blocks on fees would be biased by unobserved confounding variables. For example, a surge in legitimate user activity could increase both network congestion—which may lead to more accidental forks and thus a higher measured rate of stolen blocks—and the fees users are willing to pay for transaction priority. This would induce a spurious positive correlation, leading an OLS estimate to overstate the true causal effect of fees on strategic security degradation.

To overcome this challenge, we employ a two-stage least squares (2SLS) identification strategy. Our strategy uses the timing of major hacks and network crises as an instrument for transaction fees. The validity of this approach rests on two standard assumptions: instrument relevance and the exclusion restriction.

Instrument Relevance. The first stage requires that our instrument has a strong and statistically significant effect on the endogenous variable, $\log(\text{Fees}_t)$. This condition is readily met. Hacks and crises create a sudden, urgent demand for block space as users rush to move assets, mitigate losses, or respond to market volatility. This surge in demand predictably drives up transaction fees. As reported in Table 3, our first-stage regressions confirm this intuition. The occurrence of an event increases log USD-denominated fees by 0.272 points. The first-stage F-statistic of 118.2 for our preferred USD specification is well above conventional thresholds for weak instruments, confirming the instrument's relevance and mitigating concerns of weak instrument bias.

The Exclusion Restriction. The more crucial assumption for causal identification is the exclusion restriction, which requires that our instrument, $Event_t$, affects the rate of stolen blocks *only* through its effect on transaction fees, conditional on our fixed effects. We argue this assumption is plausible because our instruments are fundamentally demand-side shocks for block space. A major hack, for instance, does not directly alter the supply-side conditions of mining; it does not change the consensus algorithm, the physical infrastructure of the network, or the speed of block propagation. Its primary channel of influence on a miner's strategic decision to attempt a block theft is by dramatically increasing the economic prize for successfully doing so—that is, by increasing the value of the fees contained in the block. Any slow-moving changes in the technological environment, such as the gradual adoption of new mining hardware or software, are absorbed by the inclusion of quarterly fixed effects in our model. Our identification is therefore based on the sharp, high-frequency variation in fees generated by these plausibly exogenous events.

3.3 Econometric Specification

We estimate our main specification using a 2SLS model. The first-stage regression is:

$$\log(\text{Fees}_t) = \pi_1 Event_t + \gamma_{q(t)} + u_t$$

where t indexes the hour, $Event_t$ is our instrument, and $\gamma_{q(t)}$ is a full set of quarter-by-year fixed effects. The second-stage equation, which captures the causal relationship of interest, is:

$$\log(\text{StolenBlocks})_t = \beta \widehat{\log(\text{Fees}_t)} + \gamma_{q(t)} + \varepsilon_t$$

where $\widehat{\log(\text{Fees}_t)}$ is the predicted value from the first stage. The coefficient of interest, β, is the elasticity of stolen blocks with respect to transaction fees.

4 Empirical Results

We now present our main empirical findings. We first establish that higher transaction fees causally increase the rate of finality failures. We then show that this effect is robust to a battery of alternative specifications and is economically and statistically significant. Finally, we provide evidence that this mechanism is most pronounced for settlement reversals that are more likely to be strategic in nature.

Table 2 presents our main results. Columns (1) and (2) show the OLS estimates, which reveal a strong positive correlation between fees and stolen blocks. As anticipated, however, these estimates are likely biased upwards by confounders such as network congestion.

Our preferred IV estimates are reported in columns (3) and (4). Focusing on the USD specification in column (4) of Panel A, we find a positive and statistically significant coefficient of 0.107. This result indicates that the relationship is causal: higher economic stakes actively undermine settlement finality. The coefficient can be interpreted as an elasticity: a 10% increase in USD-denominated transaction fees causes a statistically significant 1.07% increase in the hourly rate of stolen blocks.

The economic magnitude of this effect is substantial, particularly during periods of market stress. We can use our first-stage estimates to quantify the impact of a network-wide crisis. The onset of a crisis event causes an immediate increase in log USD fees of 0.272 points. Our second-stage estimate implies this surge in fees leads to a total increase in log stolen blocks of $0.107 \times 0.272 \approx 0.029$, which corresponds to a 2.9% increase in the rate of finality failures. This represents a significant degradation of the ledger's security precisely when market participants value certainty the most, providing direct evidence that Proof-of-Work's security model fails its most important stress test.

4.1 Robustness and the Nature of Stolen Blocks

We now investigate the robustness of our central finding. The remaining panels of Table 2 demonstrate that our result is not an artifact of specific modeling choices or driven by outlier events.

In Panel B, we re-estimate our model while excluding "The DAO Hack," one of the most significant events in Ethereum's history. The resulting elasticity of 0.070 remains positive and significant, demonstrating that our findings are not driven by a single influential event. Panel C shows that the result also holds when we restrict our instrument to include only hack events, excluding broader crises.

A potential concern is that our measure of stolen blocks conflates strategic, economically-motivated attacks with accidental forks arising from network latency. To address this, we turn to our stricter definition of the dependent variable, which only counts reversals where the canonical block's timestamp is at least 10 seconds later than the orphaned uncle's. These events are far more likely to represent deliberate, selfish mining behavior than random network noise.

The results, presented in Panel D, are striking. The estimated elasticity for these "strategic" reversals is 1.253—an order of magnitude larger than our baseline estimate. This implies that a 10% increase in transaction fees leads to a 12.5% increase in the rate of likely-strategic settlement reversals. This finding provides powerful evidence in favor of our proposed economic mechanism. While higher fees may have a modest effect on the overall rate of network forks, their effect on the incentive to strategically attack the chain is dramatic. The economic incentives designed to secure the chain are, under stress, actively encouraging strategic deviations that undermine it.

In sum, our empirical investigation provides strong causal evidence for a fundamental flaw in the Proof-of-Work consensus mechanism. The security of the ledger is not constant but is instead endogenous to the value of the transactions it is meant to secure. When the economic stakes are highest, the guarantee of finality is at its weakest.

Table 2: The Causal Effect of Fees on Stolen Blocks

	(1) OLS (Native)	(2) OLS (USD)	(3) IV (Native)	(4) IV (USD)
Panel A: Main Results				
log(Total Fees)	0.114***	0.097***	0.151***	0.107***
	(0.002)	(0.002)	(0.044)	(0.031)
KP rk Wald F			93.912	149.463
Panel B: Robustness: Excluding The DAO Hack				
log(Total Fees)	0.113***	0.096***	0.102**	0.070**
	(0.002)	(0.002)	(0.051)	(0.034)
KP rk Wald F			71.919	123.316
Panel C: Robustness: Instrument Excludes Crisis Events				
log(Total Fees)	0.114***	0.097***	0.144*	0.080*
	(0.002)	(0.002)	(0.080)	(0.044)
KP rk Wald F			35.813	84.270
Panel D: Robustness: Stolen Block Threshold $\geq$ 10s				
log(Total Fees)	2.751***	2.295***	1.761***	1.253***
	(0.052)	(0.040)	(0.600)	(0.427)
KP rk Wald F			93.917	149.473

The table reports OLS and 2SLS (IV) estimates of the effect of hourly transaction fees on the hourly count of "stolen blocks". The sample is an hourly panel of the Ethereum blockchain from July 2015 to September 2022. The dependent variable in Panels A-C is the log of the total hourly stolen block count. Panel B excludes The DAO Hack event from the sample , and Panel C uses an alternative instrument that includes only hack events. All specifications include quarter-by-year fixed effects. In Panel D, the dependent variable uses a stricter definition, counting only stolen blocks with a timestamp difference of at least 10 seconds. The endogenous variable is the log of total hourly transaction fees, measured in native currency (ETH) in columns (1) and (3) and in U.S. dollars (USD) in columns (2) and (4). The instrumental variable for the 2SLS estimates is an indicator for days with major crypto ecosystem hacks or network-wide crises. Robust standard errors are reported in parentheses. For IV specifications, the Kleibergen-Paap (KP) rk Wald F-statistic for weak instruments is reported.

Table 3: First-Stage Regressions

Dependent Variable:	(1) log(Fees, Native)	(2) log(Fees, USD)
Panel A: Main Instrument		
Event Dummy	0.193***	0.272***
	(0.020)	(0.022)
Panel B: Robustness: Excluding The DAO Hack		
Event Dummy	0.162***	0.236***
	(0.019)	(0.021)
Panel C: Robustness: Instrument Excludes Crisis Events (Hacks Only)		
Event Dummy (Hacks)	0.113***	0.204***
	(0.019)	(0.022)

The table reports the first-stage OLS estimates corresponding to the 2SLS specifications in Table 2. The dependent variable is the log of total hourly transaction fees, in native currency (ETH) in column (1) and U.S. dollars (USD) in column (2). The main independent variable is the instrument. Panel A shows the effect of our main instrument, an indicator for a hack or crisis day, on the full sample. Panel B reports the same relationship on a sample that excludes The DAO Hack event. Panel C reports the effect of an alternative instrument that includes only hack events. All specifications are estimated on an hourly panel from July 2015 to September 2022 and include quarter-by-year fixed effects. Robust standard errors are reported in parentheses.

5 Discussion and Conclusion

This paper provides the first causal evidence that the economic security of Proof-of-Work (PoW) blockchains is endogenous to the value of the transactions they are meant to secure. Using a novel, high-frequency measure of settlement reversals derived from on-chain data, we show that plausibly exogenous spikes in transaction fees—driven by major hacks and network crises—causally degrade the finality of the ledger. The effect is not only statistically significant but economically large, particularly for reversals that are most likely to be strategic: our estimates imply that a 10% increase in USD-denominated fees increases the rate of these likely-strategic attacks by a striking 12.5%. Security weakens precisely when it is most needed.

Our findings make several contributions. First, we provide a direct empirical test of the central economic critique of PoW systems, confirming the theoretical arguments of Budish (2025) that security is inherently expensive because the cost of attack must be countered by the flow of rewards. By demonstrating that the incentive to attack scales with the prize, our results move this foundational debate from theory to causal evidence.

Second, we contribute to the literature on financial market infrastructure and systemic risk. Modern payment and settlement systems are designed to be bastions of stability, engineered to prevent—not encourage—failures during periods of market stress. Our findings show that PoW operates on the opposite principle, introducing a novel channel for systemic risk where the integrity of the infrastructure is inversely related to its usage under duress. As blockchain-based assets become increasingly integrated into the traditional financial system through ETFs and tokenization, this inherent incentive flaw represents a latent, and poorly understood, financial stability concern.

Finally, our results speak to a long-standing question in institutional economics: what are the ultimate sources of trust and enforcement in market economies? Proponents have framed blockchains as a new institutional technology that can provide "trustless" enforcement, replacing the role of traditional legal systems. Our evidence serves as an empirical testament to the classic view that state-backed legal institutions are crucial for enforcing high-value contracts and property rights. While decentralized systems may be effective at aggregating information, our findings question their ability to provide decentralized enforcement, revealing deep economic limits to this new institutional form. The "market for security" in PoW, we find, can perversely incentivize theft precisely when the rewards are highest, suggesting that the problem of securing property rights has not been solved by technology, but merely reformulated.

This analysis opens several avenues for future research. While our paper focuses on PoW, the general principle of endogenous security may apply to other consensus mechanisms, such as Proof-of-Stake, warranting further empirical investigation. Furthermore, the ability to observe security degradation in real-time using on-chain data provides a rich new laboratory for studying the economics of crime, deterrence, and institutional design. Ultimately, the integration of these new technological "rules of the game" into the financial system is not a neutral act. It is an institutional substitution that, as we show, carries with it a fundamental incentive flaw. Understanding these properties is a first-order concern for ensuring a stable financial future.

References

ABADI, J., AND M. BRUNNERMEIER (2018): "Blockchain, Coordination, and Capital," NBER Working Paper w24440, National Bureau of Economic Research.

ADRIAN, T., AND M. K. BRUNNERMEIER (2016): "CoVaR," *American Economic Review*, 106(7), 1705–1741.

BECKER, G. S. (1968): "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 76(2), 169–217.

BIAIS, B., C. BISIÈRE, M. BOUVARD, AND C. CASAMATTA (2019): "The Blockchain Folk Theorem," *The Review of Financial Studies*, 32(5), 1662–1715.

BUDISH, E. (2025): "Trust at Scale: The Economic Limits of Cryptocurrencies and Blockchains," *The Quarterly Journal of Economics*, 140(1), 1–62.

COASE, R. H. (1960): "The Problem of Social Cost," *Journal of Law and Economics*, 3, 1–44.

DIXIT, A. K. (2004): *Lawlessness and Economics: Alternative Modes of Governance*. Princeton University Press.

DUFFIE, D. (2019): "Digital Currencies and the Future of Payments," *Journal of Economic Perspectives*, 33(4), 41–62.

EYAL, I., AND E. G. SIRER (2014): "Majority is Not Enough: Bitcoin Mining Is Vulnerable," in *Financial Cryptography and Data Security*, pp. 436–454. Springer.

GENNAIOLI, N., A. SHLEIFER, AND R. W. VISHNY (2015): "Money Doctors," *The Journal of Finance*, 70(1), 91–114.

HAYEK, F. A. (1945): "The Use of Knowledge in Society," *American Economic Review*, 35(4), 519–530.

HUBERMAN, G., J. D. LESHNO, AND C. MOALLEMI (2021): "Monopoly without a Monopolist: An Economic Analysis of the Bitcoin Payment System," *The Review of Economic Studies*, 88(6), 3011–3040.

KOEPPL, T. V., AND C. MONNET (2010): "The Emergence and Future of Central Counterparties," *Bank of Canada Review*, 2010(Autumn), 27–36.

LA PORTA, R., F. LOPEZ-DE-SILANES, A. SHLEIFER, AND R. W. VISHNY (1998): "Law and Finance," *Journal of Political Economy*, 106(6), 1113–1155.

MANKIW, N. G. (1986): "The Allocation of Credit and Financial Collapse," *The Quarterly Journal of Economics*, 101(3), 455–470.

NORTH, D. C. (1990): *Institutions, Institutional Change and Economic Performance*. Cambridge University Press.

ROCHET, J.-C., AND J. TIROLE (1996): "Interbank Lending and Systemic Risk," *Journal of Money, Credit and Banking*, 28(4), 733–762.

SALEH, F. (2021): "Blockchain without Waste: Proof-of-Stake," *The Review of Financial Studies*, 34(3), 1156–1190.

SMITH, A. (1776): *An Inquiry into the Nature and Causes of the Wealth of Nations*. W. Strahan and T. Cadell.